A White Noise Approach to Evolutionary Ecology

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

3/1/2020

Abstract

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

1 Introduction

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

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

Current theoretical approaches to synthesize evolution and ecology have capitalized on the fact that biological fitness plays a key role in determining both sets of dynamics. While correlation of fitness and genotype is the 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 1970; Roughgarden 1979), evolutionary game theory (Hofbauer and Sigmund 1998; Nowak 2006; Lion 2018), quantitative genetics (Lande 1982; Doebeli 1996; Lion 2018) and a unifying framework for these three distinct approaches to evolutionary theory (Champagnat, Ferrière, and Méléard 2006) which is intimately related to the approach we take here.

Reviewing these accomplishments reveals a beautiful synthesis of evolution and population ecology. However, 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 population's response to random genetic drift is derived in discrete time under the assumption of constant effective population size using arguments based on properties of random samples (Lande 1976). Though 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 has been 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 vexing mathematical challenge. Here we close this gap by combining the calculus of white noise with results on rescaled limits of branching Brownian motion processes (BBM) and stochastic partial differential equations (SPDE). Our goal has two components: 1) Establish a novel synthetic approach to theoretical evolutionary ecology that provides a formal connection between demographic stochasticity and random genetic drift in the context of quantitative traits. To show that our approach can be used to develop useful biological insights we derive a model of coevolution in an ecological network and use it to investigate the relationship between competition coefficients and selection gradients. 2) Communicate some useful properties of space-time white noise, BBM and SPDE to as wide of audience as possible. With this goal in mind we will not provide a rigorous treatment of any of these deep subjects. Instead, we introduce a set of heuristics that only require the basic concepts of Reimann integration, partial differentiation and some exposure to Brownian motion and stochastic ordinary differential equations (SDE). For a concise introduction to SDE and Brownian motion, we recommend the primer by Evans (2014). Rigorous treatments of SPDE and rescaled limits of BBM can be found in Walsh (1986) and Dawson (1993) respectively.

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

the derived equations of population dynamics with classical niche theory to formulate a model of coevolution across a guild of S species participating in exploitative competition along a common resource continuum. In SM §5.8 we apply a classical theorem on rescaled limits of BBM that allow for ecological interactions to provide a rigorous derivation. To gain biological insight, in §3.2 we numerically integrate our model of coevolution for $S = 100$ species, tracking the dynamics of traits and abundances, under scenarios of weak and strong competition. We include an account of the natural history of the simulated system and discuss the significance of demographic stochasticity for structuring ecological communities. In §3.3 we provide expressions for selection gradients and competition coefficients implied by our model and use these expressions to investigate the relationship between the strengths of competition and coevolution. Finally, §4 concludes with a summary of accomplishments, a few remarks on the limits of this approach and future directions to incorporate more explicitly the genetic architecture of populations, feedbacks with ecosystem processes and the macroevolutionary history of interacting lineages.

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 stuctures, such as species abundance distributions and interaction networks, can be computed.

2.1 Deterministic dynamics

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

$$\frac{d\nu_i}{dt} = m_i \nu_i. \tag{1}$$

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

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

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

Two important observations of these equations include 1) mean fitness $\bar{m}$ determines the abundance dynamics of the entire population and 2) allele i will increase (decrease) in frequency if $m_i > \bar{m}$ ($< \bar{m}$). Equation (3) is known in the field of evolutionary game theory as the replicator equation (Hofbauer and Sigmund 1998; Nowak 2006; Lion 2018; Taylor and Jonker 1978; Schuster and Sigmund 1983). 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. Using a matrix of transition rates between differing types, it is straight-forward to extend the replicator equation to include mutation, which is known as the replicator-mutator equation (Nowak 2006).

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

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

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

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

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

where $\bar{x}(t)$ and $\sigma^2(t)$ are respectively the mean trait and phenotypic variance at time $t \geq 0$. Second, we assume selection is determined by the growth rate $m(\nu(x,t),x)$ that is differentiable with respect to both arguments and satisfies $m(y,x) \leq r$ for some $r \in \mathbb{R}$ and for all $x \in \mathbb{R}$, $y \geq 0$. From here on we abbreviate $m(\nu(x,t),x)$ to $m(\nu,x)$. Third, we assume mutation is captured by diffusion with coefficient $\frac{\mu}{2}$. With these technicalities aside, the demographic dynamics can be modelled by the PDE

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

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

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

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

Using our assumptions on mutation and rate of growth, we show in SM §5.1 $\nu(x,t)$ is twice differentiable with respect to x and $\int_{-\infty}^{+\infty}\nu(x,t)dx < \infty$ for all $t \geq 0$. This implies that we are justified in swapping the order of differentiation and integration and the result $\int_{-\infty}^{+\infty}\frac{\partial^2}{\partial x^2}\nu(x,t)dx = 0$ can be derived from the fundamental theorem of calculus. Biological reasoning agrees with this latter result since mutation neither creates nor destroys individuals, but merely changes their type from their parental type. Taking the same approach, we derive the dynamics of the phenotypic distribution $p(x,t)$ in response to selection and muation as

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

This result closely resembles Kimura's continuum-of-alleles model (Kimura 1965; Bürger 1986). 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), equation (9) can also be seen as a special case of the continuum-of-alleles model.

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

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

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

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

Equation (11) is a continuous time analog of the well known Robertson-Price equation without transmission bias (Robertson 1966; Price 1970; Frank 2012; Queller 2017; Lion 2018). The covariance of fitness and phenotype creates change in $\bar{x}$ to maximize mean fitness $\bar{m}$. Since this change is driven by a covariance

with respect to phenotypic diversity, the response in mean trait to selection is mediated by the phenotypic variance. In particular, when $\sigma^2 = 0$, $\bar{x}$ will not respond to selection. The result $\int_{-\infty}^{+\infty} x \frac{\partial^2}{\partial x^2} p(x,t) dx = 0$ can be found by applying integration by parts. Following the approach taken to calculate the evolution of $\bar{x}$, we find the response of phenotypic variation to this model of selection and mutation is

$$\frac{d}{dt}\sigma^2(t) = \mathrm{Cov}_t\Big(m(\nu, x), (x - \bar{x})^2\Big) + \mu. \tag{12}$$

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

In SM §5.6, we extend these results to include the effects of demographic stochasticity. Similar to the approach taken by Champagnat, Ferriére and Méléard (2006), we begin with a BBM that models populations as discrete sets of reproducing individuals whose vital rates depend on their trait value as well as the state of the entire population. Taking a large population size limit and keeping our assumption of single dimensional traits, we employ a pair of classical results that show, under the appropriate rescaling in time, phenotypic space and population density, a sequence of rescaled BBM converges to a limiting process that can be characterized by a SPDE (Méléard and Roelly 1993; Li 1998). The limiting processes of rescaled BBM have been referred to as measure-valued Markov processes (Dawson 1993) or superprocesses (Etheridge 2000). Under the simplifying assumptions inherited from our treatment of deterministic dynamics and an additional assumption on demographic stochasticity, we obtain as a special case a relatively simple expression for an SPDE that generalizes equation (7). The simplicity of our special case allows us to use properties of space-time white noise processes to derive a set of SDE that generalize equations (8), (11) and (12) to include the effects of demographic stochasticity. Classical expressions for the effects of random genetic drift on the evolution of mean traits are obtained as a special case.

In the following section we provide the necessary mathematical tools needed to derive SDE from SPDE. 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. In addition to the notions of Reimann integration and partial differentiation already employed, the reader will only need some elementary probability and an intuitive understanding of SDE, including Brownian motion. Because space-time white noise, denoted by $\dot{W}(x,t)$, appears in the SPDE characterizing diffusion limits of BBM, we begin by defining $\dot{W}(x,t)$ and illustrating its relevant properties including a set of heuristics for performing calculations. Treating only the simplist of cases, we then provide a brief review of BBM, their diffusion limits and the SPDE that characterize them. For those not interested in the white noise calculus or superprocesses and would rather jump straight into more biologically relevant results, we recommend skipping to §2.3.1.

2.2 White noise calculus and superprocesses

2.2.1 Definition and basic properties of white noise

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

However, it turns out that realizations of white noise do not exist as functions in the classical sense. Indeed, since Brownian motion is nowhere differentiable with respect to time, white noise cannot be formally under-

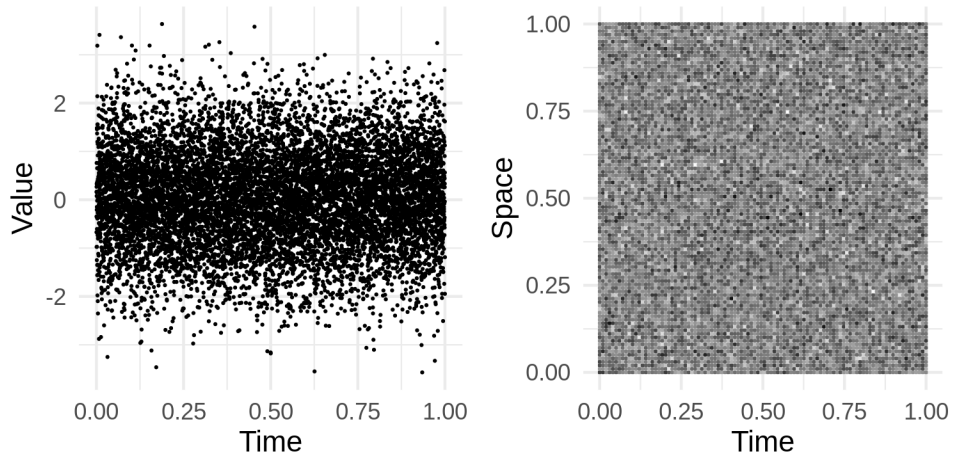

Figure 1: Approximations of sample paths of temporal white noise (left) and space-time white noise (right) with brightness scaled to value.

stood as its time derivative. Thus our notation $\dot{W}$ is only meant to aid intuition and not be taken as formal. A formal understanding is possible by considering white noise as a *measure*-valued process (Dawson 1975; Walsh 1986) or as a *generalized* process that acts on classically defined functions or stochastic processes to return either random variables or stochastic processes (Krylov and Rozovskii 1981; Da Prato and Zabczyk 2014). Since a measure-valued process can be defined from a generalized process and a generalized process can be defined from a measure-valued process, the distinction between the two is more or less a matter of perspective. However, we find the perspective of white-noise as a generalized process to be a more efficient route for developing heuristics to help with some routine calculations involved with deriving SDE from SPDE. Hence, the notion of a generalized process provides the general idea implemented here. Although the treatments of SPDE provided by Krylov and Rozovskii (1981) and Da Prato and Zabczyk (2014) extend the theory of SDE to formally treat SPDE in a general and elegant fashion, they require the navigation of an enourmous amount of technical definitions and detailed proofs. To extract some particularly useful results from this theory relevant to our goal of synthesizing the stochastic dynamics of biological populations, we provide a streamlined approach by capitolizing on the solid ground these authors have established. For instance, instead of rigorously proving properties of white-noise, we simply define them to be so, taking solice in the fact that the technical details have been worked out elsewhere.

Before diving in, we shed a bit of light on the idea of a generalized process. A generalized process is the stochastic analog of a generalized function, such as the Dirac delta function δ. Just as a generalized function operates on classical functions to return a value (e.g., $\delta(f) = f(0)$), a generalized process acts on a set of functions to return a classically defined stochastic process. For a breif primer on the theory of generalized functions, see the addendum to chapter 3 of Kolmogorov and Fomin (1999).

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

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

and make use of the convention $f = g$ if $\|f - g\|_t = 0$ for all $t \geq 0$.

We define a generalized stochastic process $\mathbf{W}$ that maps processes $f \in \mathscr{N}_2$ to real-valued stochastic processes indexed by time $t \geq 0$, but not by space. To evaluate $\mathbf{W}$ for a process $f \in \mathscr{N}_2$ and some time $t \geq 0$ we write

$\mathbf{W}_t(f)$. Specifically, for any $f, g \in \mathscr{N}_2$, we define $\mathbf{W}(f)$ and $\mathbf{W}(g)$ to be Gaussian processes satisfying, for any $t, t_1, t_2 \geq 0$,

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

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

where $t_1 \wedge t_2 = \min(t_1, t_2)$ and $\mathbb{C}$ denotes covariance with respect to the underlying probability space. In particular, denoting $\mathbb{V}$ the variance operator with respect to the underlying probability space, we have $\mathbb{V}\big(\mathbf{W}_t(f)\big) = \|f\|_t^2$ for all $t \geq 0$ and $f \in \mathscr{N}_2$. The operators $\mathbb{E}$ and $\mathbb{C}$ are to be distinguished from $\bar{f}(t)$ and $\mathrm{Cov}_t(f, g)$ 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 and since we assume the $\mathscr{N}_2$ processes are continuous in time, the processes $\mathbf{W}(f)$ and $\mathbf{W}(g)$ are well defined. As an example, if $f \in \mathscr{N}_2$ is independent of time, then $\mathbf{W}(f)$ is a Brownian motion with variance at time $t \geq 0$ equal to $\|f\|_t^2 = t\ \mathbb{E}(\int_{\mathbb{R}} f^2(x,0)dx)$. With the generalized process $\mathbf{W}$ defined, we define the space-time white noise $\dot{W}(x,t)$ implicitly via the stochastic integral

$$\text{“}\int_0^t \int_{\mathbb{R}} f(x,s)\dot{W}(x,s)dxds\text{”} = \mathbf{W}_t(f),\ \forall\ f \in \mathscr{N}_2,\ t \geq 0. \tag{16}$$

We place quotations in the above expression to emphasize its informal nature and that it should not be confused with classical Riemann integration. Following this definition of white noise, we compute its value by sampling it using $\mathscr{N}_2$ processes. 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 $\mathbf{W}_t(I_{D\times[0,+\infty)})$ for the deterministic process

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

Since

$$\|I_{D\times[0,+\infty))}\|_t^2 = \mathbb{E}\left(\int_0^t \int_{\mathbb{R}} I^2_{D\times[0,+\infty)}(x,s)dxds\right) = t\int_D dx = t|D| < +\infty, \tag{18}$$

where $|D|$ denotes the length of D, we have $I_{D\times[0,+\infty)} \in \mathscr{N}_2$. Thus, using equations (14) and (15) and adopting the informal notation introduced in equation (16), we can write the following

$$\mathbb{E}\left(\int_0^t \int_D \dot{W}(x,s)dxds\right) = 0, \tag{19}$$

$$\mathbb{V}\left(\int_0^t \int_D \dot{W}(x,s)dxds\right) = t|D|. \tag{20}$$

Using this informal notation, equations (14) and (15) can be rewritten as

$$\mathbb{E}\left(\int_0^t \int_{\mathbb{R}} f(x,s)\dot{W}(x,s)dxds\right) = 0, \tag{21}$$

$$\mathbb{C}\left(\int_0^{t_1} \int_{\mathbb{R}} f(x,s)\dot{W}(x,s)dxds, \int_0^{t_2} \int_{\mathbb{R}} g(x,s)\dot{W}(x,s)dxds\right) = \int_0^{t_1\wedge t_2} \int_{\mathbb{R}} f(x,s)g(x,s)dxds. \tag{22}$$

To relate these formulae to the common notation used for SDE, we write

$$d\hat{\mathbf{W}}_t(f) = \frac{1}{\|f\|_t}\left(\int_{\mathbb{R}} f(x,t)\dot{W}(x,t)dx\right)dt, \tag{23}$$

so that

$$\int_0^t d\hat{\mathbf{W}}_s(f) = \int_0^t \int_{\mathbb{R}} \frac{f(x,s)}{\sqrt{\int_{\mathbb{R}} f^2(s,y)dy}}\dot{W}(x,s)dxds. \tag{24}$$

This implies $\mathbb{E}\big(\int_0^t d\hat{\mathbf{W}}_s(f)\big) = 0$, $\mathbb{C}\big(\int_0^{t_1} d\hat{\mathbf{W}}_s(f), \int_0^{t_2} d\hat{\mathbf{W}}_s(f)\big) = t_1 \wedge t_2$ and, as a function of t, $\int_0^t d\hat{\mathbf{W}}_s(f)$ is a standard Brownian motion for any $f \in \mathscr{N}_2$. Hence, $d\hat{\mathbf{W}}_t(f)$ is analogous to the traditional shorthand used to denote stochastic differentials. Thus, equation (22) effectively extends Itô's multiplication table to:

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

	$d\hat{\mathbf{W}}_t(f)$	$d\hat{\mathbf{W}}_t(g)$	dt
$d\hat{\mathbf{W}}_t(f)$	dt	$\frac{\left(\int_{\mathbb{R}} f\ g\ dx\right)dt}{\|f\|_t \|g\|_t}$	0
$d\hat{\mathbf{W}}_t(g)$	$\frac{\left(\int_{\mathbb{R}} f\ g\ dx\right)dt}{\|f\|_t \|g\|_t}$	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 §5.6 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 particular SPDE naturally arises as the diffusion limit of a BBM.

2.2.2 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 (i.e., 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 populations of individuals occuring in some continuous space (Watanabe 1968; Dawson 1975; Perkins 1992, 1995; Méléard and Roelly 1993; Li 1998; Bertoin and Le Gall 2003; Etheridge 2008; Barton and Etheridge 2019). This space can represent geographic space or, relevant to our context, phenotypic space. In the following subsection, we describe the BBM process, which is a particularly important branching process with spatial stucture. This process has been very useful in the study of SPDE due to its simplifying assumption that individuals do not interact. However, this assumption imposes an unfortunate restriction by precluding the modelling of ecological interactions. We therefore follow our discussion of BBM with a review of a few important results on spatially structured branching processes that account for interactions.

Branching Brownian motion

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

1) **Branching rate:** In our formulation of BBM we assume Lifetimes of individuals are exponentially distributed with death rate $\lambda > 0$ and reproduction occurs simultaneously with death. Biologically, this

implies individuals are semelparous An alternative formulation treats birth and death events separately to model iteroparity. However, under the appropriate rescaling, both approaches converge to the same diffusion limit. We therefore choose the former approach for the sake of simplicity.

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

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

To obtain a SPDE from a BBM we take a diffusion limit. There are several ways to do this, but a simple approach is to rescale the mass of individuals and time by $1/n$, diffusion by $\mu \to \mu/n$, branching rate by $\lambda \to n\lambda$, fitness by $\mathscr{W} \to \mathscr{W}^{1/n}$ and consider the limit as $n \to \infty$. Denoting the rescaled process by $X_t^{(n)}(D)$, the limiting process $\mathscr{X}_t = \lim_{n\to\infty} X_t^{(n)}$ is called a super-Brownian motion and is also a non-negative finite measure-valued process (Watanabe 1968). However, instead of returning the number of individuals alive in a region of space, super-Brownian motion returns the *mass* of the population concentrated in a region of space. Since we have rescaled individual mass by $1/n$ and took the limit $n \to \infty$, individuals are no longer discrete units. Instead, the particle view of the population gets replaced by a blob spread across d-dimensional space. In particular, the value of $\mathscr{X}_t(D)$ is a continuously varying non-negative random variable for any $t \geq 0$ and $D \subset \mathbb{R}^d$.

Unfortunately, just as with cream cheese spread across too much toast, the blob perspective of the population may exhibit some complicated spatial discontinuities. However, for spatial dimension $d = 1$, it turns out that $\mathscr{X}_t$ is absolutely continuous with respect to the Lebesgue measure for each $t \geq 0$ (Konno and Shiga 1988; Reimers 1989). This means that we can write $\mathscr{X}_t(D) = \int_D \nu(x,t)dx$ for some density process $\nu(x,t)$. Setting $\lambda = 1$ and $m = \ln \mathscr{W}$ this density process satisfies the SPDE

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

Since $\nu(x,t)$ is not generally differentiable in x or t, the derivatives in expression (25) are taken in the *weak* sense (sensu Evans 2010). That is, to rigorously interpret SPDE (25), we integrate the solution $\nu(x,t)$ against functions $f \in C_b^2(\mathbb{R})$, where $C_b^2(\mathbb{R})$ is the set of bounded and twice continuously differentiable functions on $\mathbb{R}$. Hence, equation (25) is just an abbreviation for

$$\int_{\mathbb{R}} \nu(x,t)f(x)dx - \int_{\mathbb{R}} \nu(x,0)f(x)dx = \int_0^t \int_{\mathbb{R}} \nu(x,s)\left(mf(x) + \frac{\mu}{2}\frac{\partial^2}{\partial x^2}f(x)\right)dsdx + \int_0^t \int_{\mathbb{R}} f(x)\sqrt{V\nu(x,s)}\dot{W}(x,s)dxds, \ \forall\ f \in C_b^2(\mathbb{R}). \tag{26}$$

This expression is referred to as the *mild* solution of (25). For more on the general theory of SPDE see Walsh (1986) and Da Prato and Zabczyk (2014). Note that since $\nu(x,t)$ is the density of a finite measure, it is integrable for each $t \geq 0$. Thus, since for some $M > 0$, $|f(x)| \leq M$ for every $x \in \mathbb{R}$, setting $\varphi(x,t) = f(x)\sqrt{V\nu(x,t)}$ implies $\varphi \in \mathscr{N}_2$. Hence, the white noise integral on the right-hand side of equation (26)

can be understood using the heuristics introduced above. Evaluating equation (26) 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 generalization of BBM 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. Konna and Shiga (1988) and Reimers (1989) independently proved Dawson's suggestion was indeed correct. The diffusion limit of this more general branching process (in arbitrary spatial dimension) is referred to as a Dawson-Watanabe superprocess (Etheridge 2000). Conditioning a Dawson-Watanabe superprocess to have constant mass returns a Fleming-Viot process (Etheridge and March 1991; Perkins 1991) which has been popular in studies of spatial population genetics. In particular, an extension of Fleming-Viot processes, known as Λ-Fleming-Viot processes, were introduced by Bertoin and Le Gall (2003) and coined by Etheridge (2008) where it was used to resolve some technical challenges in modelling isolation by distance (Felsenstein 1975; see also Barton, Etheridge, and Véber 2013; and Barton and Etheridge 2019). Although this provides an impressive list of accomplishments, the Dawson-Watanabe superprocess falls short of our needs. In particular this process assumes individuals do not interact and thus precludes its ability to model ecological interactions. However, this concern has been addressed, leading to constuctions of superprocesses that account for interactions among individuals. In the next subsection we summarize the main results in this area and introduce the SPDE that provides the basis for our approach to theoretical evolutionary ecology.

Interacting superprocesses

The existence of diffusion limits for a class of measure-valued branching processes involving interactions among individuals has been treated by Méléard and Roelly (1992, 1993). The interactions can manifest as dependencies of the spatial movement or reproductive law of individuals on their position x and the state of the whole population described by X_t. An important result of Méléard and Roelly (1992, 1993) is a theorem that provides sufficient conditions to construct a sequence of rescaled measure-valued branching processes that converge to a generalization of the Dawson-Watanabe superprocess that includes interactions. The rescaling is analogous to that described above for non-interacting Dawson-Watanabe superprocesses, but now the reproductive law described by $\mathscr{W}(X_t, x)$ and $V(X_t, x)$, branching rate $\lambda(X_t, x)$ and diffusion parameter $\sqrt{\mu(X_t, x)}$ are allowed to depend on the whole population X_t and individual location x. In Figure 2 we demonstrate this rescaling in discrete time for a population experiencing stabilizing selection and logistic growth. Since time is discretized, the process we simulate is formally a branching random walk. For further details on our simulation see SM §5.5.

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

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

$$\frac{\partial}{\partial t}\nu(x,t) = m(\nu, x)\nu(x,t) + \frac{\mu}{2}\frac{\partial^2}{\partial x^2}\nu(x,t) + \sqrt{V\nu(x,t)}\dot{W}(x,t). \tag{27}$$

Comparing equation (27) to equation (3.5) of Li (1998), our m and V correspond to Li's b and c respectively. It is important to note that, under the assumptions made in Méléard and Roelly (1992, 1993) and Li (1998), equation (27) is only formal when $m(y, x)$ is bounded across all combinations of $y \geq 0$ and $x \in \mathbb{R}$. However, recalling our condition $m(y, x) \leq r \in \mathbb{R}$, the growth rates we consider are only bounded above. Yet, in the proof of the construction of the interacting superprocess as the limit of rescaled branching diffusions, Méléard

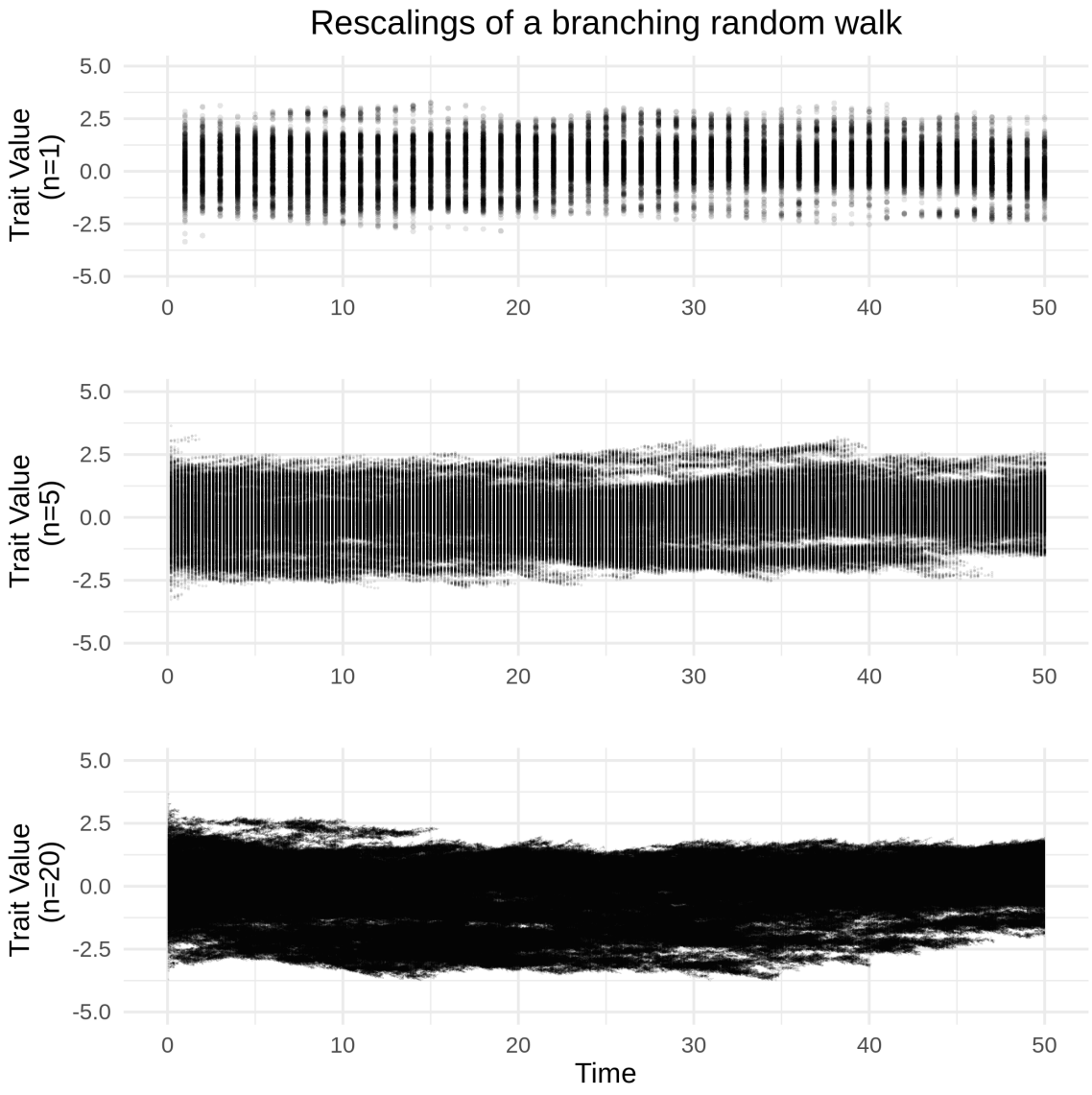

Figure 2: Rescaled sample paths of a branching random walk under stabilizing selection and logistic growth. The top plot displays a sample path without scaling ($n = 1$), the middle plot shows a sample path rescaled by $n = 5$ and the bottom plot shows a sample path rescaled by $n = 20$.

and Roelly (1992, 1993) assumed $m(y,x)$ to be bounded to guarantee the total mass process will have finite mean and variance, for finite $t \geq 0$. This allowed the authors to employ a tightness criterion for sequences of measures and show the rescaled processes converge to a superprocess with finite total mass. Li's (1998) result builds directly off of Méléard and Roelly's construction, inheriting the assumption of boundedness for $m(y,x)$. However, in Li (1998), the sufficiency of $m(y,x)$ being bounded above is even more clear since Li works explicitly with a common upperbound for both $m(y,x)$ and V. Hence, one can repeat the necessary proofs replacing the assumption that $m(y,x)$ is bounded with the assumption that $m(y,x)$ is merely bounded above to derive the same results.

With solutions to SPDE (27) well defined for any growth rate bounded above, we can calculate the total mass process $N(t)$ using the mild solution of (27) with $f(x) \equiv 1 \in C_b^2(\mathbb{R})$ (the symbol "$\equiv$" means equal to for every x). That is,

$$\begin{aligned} N(t) = N(0) + \int_0^t \int_{\mathbb{R}} \nu(x,s) \left(m(\nu,x) \cdot 1 + \frac{\mu}{2} \frac{\partial^2}{\partial x^2} 1 \right) + 1\sqrt{V\nu(x,s)}\dot{W}(x,s)dsdx \\ = N(0) + \int_0^t \bar{m}(s)N(s)dt + \int_0^t \sqrt{VN(s)}d\hat{\mathbf{W}}_s(\sqrt{\nu(x,s)}), \end{aligned} \tag{28}$$

where

$$\bar{m}(t) = \frac{1}{N(t)} \int_{\mathbb{R}} m(\nu,x)\nu(x,t)dx, \tag{29}$$

and

$$\int_0^t d\hat{\mathbf{W}}_s(\sqrt{\nu(x,s)}) = \int_0^t \int_{\mathbb{R}} \frac{\sqrt{\nu(x,s)}}{\sqrt{\int_{\mathbb{R}} \nu(x,s)dx}} \dot{W}(x,s)dxds. \tag{30}$$

Setting $W_1(t) = \hat{\mathbf{W}}_t(\sqrt{\nu(x,t)})$, we can use traditional stochastic differential notation to write

$$dN = \bar{m}Ndt + \sqrt{VN}dW_1. \tag{31}$$

To find the associated SDE for $\bar{x}(t)$ and $\sigma^2(t)$, we want to repeat the same approach for $f(x) = x, x^2$ and apply Itô's lemma. However, for these cases $f \notin C_b^2(\mathbb{R})$ since f will not be bounded. However, if we can show that $\int_{\mathbb{R}}(|x| + x^2 + x^4)\nu(x,t)dx < +\infty$ for all $t > 0$ given this condition is satisfied by $\nu(x,0)$, then we can apply the mild solution of (27) to derive SDE for $\bar{x}(t)$ and $\sigma^2(t)$. To illustrate, let us suppose this is the case. Setting $\tilde{x}(t) = \int_{\mathbb{R}} x\nu(x,t)dx$, we have

$$\tilde{x}(t) = \tilde{x}(0) + \int_0^t \int_{\mathbb{R}} \nu(x,s)m(\nu,x)x + x\sqrt{V\nu(x,s)}\dot{W}(x,s)dxds. \tag{32}$$

Similarly, setting $\tilde{\tilde{\sigma}}^2(t) = \int_{\mathbb{R}} x^2\nu(x,t)dx$, we have

$$\tilde{\tilde{\sigma}}^2(t) = \tilde{\tilde{\sigma}}^2(0) + \int_0^t \int_{\mathbb{R}} \nu(x,s)\left(m(\nu,x)x^2 + \mu\right) + x^2\sqrt{V\nu(x,s)}\dot{W}(x,s)dxds. \tag{33}$$

Since $\bar{x}(t) = \tilde{x}(t)/N(t)$ and $\sigma^2(t) = \tilde{\tilde{\sigma}}^2(t)/N(t) - \bar{x}^2(t)$, we can use Itô's lemma to derive SDE for $\bar{x}(t)$ and $\sigma^2(t)$, which we perform in SM §5.6. We do not make any general assertions about the existence or uniqueness of $\bar{x}(t)$ or $\sigma^2(t)$.

2.3 Equations of evolutionary and demographic dynamics

In SM §5.6 we show SDE for $N(t)$, $\bar{x}(t)$ and $\sigma^2(t)$ can be expressed as

$$dN(t) = \bar{m}(t)N(t)dt + \sqrt{VN(t)}dW_1(t), \tag{34a}$$

$$d\bar{x}(t) = \mathrm{Cov}_t\Big(x, m(\nu, x)\Big)dt + \sqrt{V\frac{\sigma^2(t)}{N(t)}}dW_2(t), \tag{34b}$$

$$d\sigma^2(t) = \left(\mathrm{Cov}_t\Big((x-\bar{x}(t))^2, m(\nu,x)\Big) + \mu - V\frac{\sigma^2(t)}{N(t)}\right)dt + \sqrt{V\frac{\overline{(x-\bar{x}(t))^4} - \sigma^4(t)}{N(t)}}dW_3(t), \tag{34c}$$

where W_1, W_2 and W_3 are standard Brownian motions. We note that conditions on the growth rate m to guarantee existence and uniqueness of solutions to (34b) and (34c) have yet to be investigated. However, our results on the deterministic PDE suggest that $m(y, x)$ bounded above and differentiable in both arguments is sufficient. Dividing by dt one can interpret equations (34) as if they are ordinary differential equations, but this not technically rigorous since Brownian motion is nowhere differentiable with respect to time. In SM §5.6 we show that in general W_1 is independent of both W_2 and W_3, but W_2 and W_3 covary.

There is quite a bit we can learn from expressions (34). Firstly, setting $V = 0$ recovers the deterministic dynamics derived in §2.1. Alternatively, one can take $N(t) \to \infty$ to recover the deterministic dynamics for $\bar{x}(t)$ and $\sigma^2(t)$. Characteristically, we note the effect of demographic stochasticity on abundance grows with $\sqrt{N(t)}$. Hence, dividing by N, we find the effects of demographic stochasticity on the per-capita growth rate diminish with increased abundance. Relating the response to demographic stochasticity derived here to the effect of random genetic drift derived in classic quantitative genetic theory, if we set $\sigma^2(t) = \sigma^2$ and $N(t) = N$ constant with respect to time, then integrating the stochastic term in equation (34b) 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 reproductive variance is unity ($V = 1$) 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 (34c), particularly $-V\sigma^2(t)/N(t)$. The component of SDE (34c) describing random fluctuations in $\sigma^2(t)$ is more complicated and is proportional to the root of the difference between the centralized fourth moment of $p(x, t)$ and $\sigma^4(t)$.

These expressions can be used to investigate the dynamics of the mean and variance for general $\nu(x, t)$. However, in the next subsection we simplify these expressions by approximating $\nu(x, t)$ with a Gaussian curve. By assuming $\nu(x, t)$ is Gaussian for $t \geq 0$, we guarantee the existence of $\bar{x}(t)$ and $\sigma^2(t)$ for all $t \geq 0$. Furthermore, in SM §5.6 we show that under the Gaussian case W_1, W_2 and W_3 are independent.

2.3.1 Particular results assuming a Gaussian phenotypic distribution

By assuming $\nu(x, t)$ can be approximated by a Gaussian curve for each $t \geq 0$, expressions (34a), (34b) and (34b) transform into efficient tools for deriving the dynamics of populations given a fitness function $m(\nu, x)$. 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 (27) will only be a Gaussian curve in expectation, assuming a Gaussian initial condition. Yet this difficulty is not as challenging as it may first appear. Indeed, since SPDE (27) 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 (i.e., when the variance in reproductive output is much smaller than the population size). In SM §5.5 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

$$\nu(x,t) = \frac{N(t)}{\sqrt{2\pi\sigma^2(t)}} \exp\left(-\frac{\left(x-\bar{x}(t)\right)^2}{2\sigma^2(t)}\right). \tag{35}$$

Under this assumption we find in SM §5.4 the results (suppressing the dependency on t)

$$\mathrm{Cov}(x,m) = \sigma^2\left(\frac{\partial\bar{m}}{\partial\bar{x}} - \overline{\frac{\partial m}{\partial\bar{x}}}\right), \tag{36}$$

$$\mathrm{Cov}\Big((x-\bar{x})^2, m\Big) = 2\sigma^4\left(\frac{\partial\bar{m}}{\partial\sigma^2} - \overline{\frac{\partial m}{\partial\sigma^2}}\right) \tag{37}$$

and $\overline{(x-\bar{x})^4} = 3\sigma^4$. Equation (36) is the continuous time equivalent to equation (9) in Lande (1976). In particular, these results imply

$$d\bar{x} = \sigma^2\left(\frac{\partial\bar{m}}{\partial\bar{x}} - \overline{\frac{\partial m}{\partial\bar{x}}}\right)dt + \sqrt{V\frac{\sigma^2}{N}}dW_2, \tag{38a}$$

$$d\sigma^2 = 2\sigma^4\left(\frac{\partial\bar{m}}{\partial\sigma^2} - \overline{\frac{\partial m}{\partial\sigma^2}}\right)dt + \left(\mu - V\frac{\sigma^2}{N}\right)dt + \sigma^2\sqrt{\frac{2V}{N}}dW_3. \tag{38b}$$

These equations allow us to derive the response in trait mean and variance by taking derivatives of fitness, a much more straightforward operation than calculating a covariance for general phenotypic distributions. 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 the mean trait, but here we see an analogous 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. This genetic variance represents the component of the variance in expressed traits σ^2 explained by additive effects of different alleles among genetic loci encoding for the focal phenotype (Roughgarden 1979; Bulmer 1980; Lynch and Walsh 1998). It is therefore fitting that G is referred to as the additive genetic variance.

2.3.2 The evolution of additive genetic variance

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

$$\psi(x,g) = \frac{1}{\sqrt{2\pi\eta}} \exp\left(-\frac{(x-g)^2}{2\eta}\right). \tag{39}$$

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

evolution of $\nu(x,t)$ to modelling the evolution of $\rho(g,t)$. Once $\rho(g,t)$ is determined, we can compute $\nu(x,t)$ via

$$\nu(x,t) = \int_{-\infty}^{+\infty} \rho(g,t)\psi(x,g)dg. \tag{40}$$

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

$$m^*(\rho,g) = \int_{-\infty}^{+\infty} m(\nu,x)\psi(x,g)dx. \tag{41}$$

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 the SPDE

$$\dot{\rho}(g,t) = m^*(\rho,g)\rho(g,t) + \frac{\mu}{2}\frac{\partial^2}{\partial^2 g}\rho(g,t) + \sqrt{V\rho(g,t)}\dot{W}(g,t). \tag{42}$$

We assume $\rho(g,t)$ is Gaussian which implies its mode coincides with $\bar{x}$. Furthermore, since $\sigma^2 = G + \eta$, we can use equation (41) and the chain rule from calculus to find

$$\frac{\partial \bar{m}}{\partial G} = \frac{\partial \bar{m}}{\partial \sigma^2}\frac{\partial \sigma^2}{\partial G} = \frac{\partial \bar{m}}{\partial \sigma^2}, \tag{43a}$$

$$\overline{\frac{\partial m}{\partial G}} = \overline{\frac{\partial m}{\partial \sigma^2}\frac{\partial \sigma^2}{\partial G}} = \overline{\frac{\partial m}{\partial \sigma^2}}. \tag{43b}$$

Thus, equations (38) become

$$d\bar{x} = G\left(\frac{\partial \bar{m}}{\partial \bar{x}} - \overline{\frac{\partial m}{\partial \bar{x}}}\right)dt + \sqrt{V\frac{G}{N}}dW_2, \tag{44a}$$

$$dG = 2G^2\left(\frac{\partial \bar{m}}{\partial G} - \overline{\frac{\partial m}{\partial G}}\right)dt + \left(\mu - V\frac{G}{N}\right)dt + G\sqrt{\frac{2V}{N}}dW_3. \tag{44b}$$

From expressions (44) we see that, under our model of inheritance, focusing on additive genetic variance G instead the variance in expressed traits σ^2 makes no structural changes to the basic equations describing the dynamics of populations.

3 A model of diffuse coevolution

3.1 Formulation

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

The dynamics of phenotypic distributions and abundances have been derived above and so the only task remaining is the formulation of a fitness function. 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 §5.8.

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 resouce 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

$$u_i(\zeta, x_i) = \frac{U_i}{\sqrt{2\pi w_i}} \exp\left(-\frac{(x_i - \zeta)^2}{2w_i}\right). \tag{45}$$

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

$$e_i(\zeta) = Q_i \exp\left(-\frac{A_i}{2}(\theta_i - \zeta)^2\right). \tag{46}$$

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

$$\int_{-\infty}^{+\infty} e_i(\zeta) u_i(\zeta, x_i) d\zeta = \frac{Q_i U_i}{\sqrt{A_i w_i + 1}} \exp\left(-\frac{A_i}{2(A_i w_i + 1)}(\theta_i - x_i)^2\right). \tag{47}$$

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

$$\mathcal{O}_{ij}(x_i, x_j) = \int_{-\infty}^{+\infty} u_i(\zeta, x_i) u_j(\zeta, x_j) d\zeta = \frac{U_i U_j}{\sqrt{2\pi(w_i + w_j)}} \exp\left(-\frac{(x_i - x_j)^2}{2(w_i + w_j)}\right). \tag{48}$$

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

for the individual in species i by $c_i\mathcal{O}_{ij}(x_i, x_j)$ for some $c_i \geq 0$. In SM §5.8 we combine this niche model with equations (34a), (44a) and (44b) to find

$$dN_i = \left\{ R_i - \frac{a_i}{2}\Big((\bar{x}_i - \theta_i)^2 + G_i + \eta_i\Big) - c_i \sum_{j=1}^{S} N_j U_i U_j \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2} \right\} N_i dt + \sqrt{V_i N_i} dW_1, \quad (49a)$$

$$d\bar{x}_i = \left\{ a_i G_i(\theta_i - \bar{x}_i) - c_i G_i \Big(\sum_{j=1}^{S} N_j U_i U_j b_{ij} (\bar{x}_j - \bar{x}_i) \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2} \Big) \right\} dt + \sqrt{V_i \frac{G_i}{N_i}} dW_2, \quad (49b)$$

$$dG_i = \left\{ c_i {G_i}^2 \left(N_i U_i^2 b_{ii} \sqrt{\frac{b_{ii}}{2\pi}} + \sum_{j=1}^{S} N_j U_i U_j b_{ij} \left(1 - b_{ij}(\bar{x}_i - \bar{x}_j)^2\right) \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2} \right) \right.$$
$$\left. + \mu_i - a_i {G_i}^2 - V_i \frac{G_i}{N_i} \right\} dt + G_i \sqrt{\frac{2V_i}{N_i}} dW_3, \quad (49c)$$

where

$$R_i = \ln\left(\frac{Q_i U_i}{\sqrt{1 + A_i w_i}} \right), \quad (50a)$$

$$a_i = \frac{A_i}{1 + A_i w_i}, \quad (50b)$$

$$b_{ij}(t) = b_{ji}(t) = \big(w_i + w_j + \eta_i + \eta_j + G_i(t) + G_j(t)\big)^{-1}, \quad (50c)$$

$$c_i \geq 0. \quad (50d)$$

Despite the convoluted appearance of system (49), there are some nice features that reflect biological reasoning. For example, the dynamics of abundance are just a generalization of Lotka-Volterra dynamics. In particular, the effect of competition with species j on the fitness of species i grows linearly with N_j. However, as biotic selection pushes $\bar{x}_i$ away from $\bar{x}_j$, the effect of competition with species j on the fitness of species i rapidly diminishes, reflecting a reduction in niche overlap. The divergence of $\bar{x}_i$ and $\bar{x}_j$ due to competition is referred to in the community ecology literature as character displacement. We also see that the fitness of species i drops quadratically with the difference between $\bar{x}_i$ and the abiotic optimum θ_i. Hence, abiotic selection acts to pull $\bar{x}_i$ towards θ_i. The response in mean trait $\bar{x}_i$ to natural selection is proportional to the amount of heritable variation in the population, represented by the additive genetic variance G_i. However, we have that G_i is itself a dynamic quantity. Under our model, abiotic stabilizing selection erodes away heritable variation at a rate that is independent of both N_i and $\bar{x}_i$. The effect of competition on G_i is a bit more complicated. When $b_{ij}(\bar{x}_i - \bar{x}_j)^2 < 1$, competition with species j acts as diversifying selection which tends to increase the amount of heritable variation. However, when $b_{ij}(\bar{x}_i - \bar{x}_j)^2 > 1$, competition with species j acts as directional selection and reduces G_i. In the following subsections we demonstrate the behavior of system (49) by plotting numerical solutions and investigate implications for the relationship between the strength of ecological interactions and selection.

3.2 Community dynamics

For the sake of illustration we numerically integrated system (49) for a richness of $S = 100$ species. We assumed homogeneous model parameters across species in the community as summarized by Table 2. We

repeated numerical integration under the two scenarios of weak and strong competition. For the first scenario of weak competition we set $c = 1.0 \times 10^{-7}$ and for the second scenario of strong competition we set $c = 5.0 \times 10^{-6}$. With these two sets of model parameters, we simulated our model for 1000.0 units of time. For both scenarios, we initialized the trait means to $\bar{x}_i = 0.0$, additive genetic variances to $G_i = 10.0$ and abundances to $N_i = 1000.0$ for each $i = 1, \ldots, S$.

Temporal dynamics for each scenario are provided in Figure 3. This figure suggests weaker competition leads to smoother dynamics and a higher degree of organization within the community. Considering expression (49a) we note that, all else equal, relaxed competition allows for larger growth rates which promote greater abundances. From (49a) we also note that the per-capita effects on demographic stochasticity diminish with abundance. To see this, divide both sides by N_i. Inspecting expressions (49b) and (49c), we see that larger abundances also erode the effects of demographic stochasticity on the evolution of mean trait and additive genetic variance. These effects were already noted in §2.3, and thus are not a consequence of our model of coevolution per-se, but we revisit them here since Figure 3 demonstrates the importance of demographic stochasticity in structuring ecological communities even when populations are very large. Hence, contrary to the common assumption that stochastic effects can be ignored for large populations, we find that minute asymmetries by demographic stochasticity remain significant drivers of community structure. In particular, we initialized the species with identical state variables and model parameters, but found an enourmous amount of asymmetry and even some potential phase changes. In the following two paragraphs we describe the natural history of the community as illustrated in Figure 3.

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

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

Finding ways to interpret simulated dynamics provides a useful arena to exercise biological reasoning. How-

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

3.3 The relation between the strength of ecological interactions and coevolution

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

$$\alpha_{ij}(t) = \frac{c_i}{r_i(t)} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p_i(x,t) p_j(y,t) \mathcal{O}_{ij}(x,y) dx dy = \frac{c_i U_i U_j}{r_i(t)} \sqrt{\frac{b_{ij}(t)}{2\pi}} \exp\left(-\frac{b_{ij}(t)}{2} \big(\bar{x}_i(t) - \bar{x}_j(t)\big)^2 \right), \tag{51}$$

where

$$r_i(t) = R_i - \frac{a_i}{2}\Big((\bar{x}_i(t) - \theta_i)^2 + G_i(t) + \eta_i \Big). \tag{52}$$

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

$$dN_i(t) = r_i(t) \left(1 - \sum_{j=1}^{S} \alpha_{ij}(t) N_j(t) \right) N_i(t) dt + \sqrt{V_i N_i(t)} dW_1(t). \tag{53}$$

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

$$\beta_{ij}(t) = c_i U_i U_j N_j(t) b_{ij}(t) \big(\bar{x}_i(t) - \bar{x}_j(t)\big) \sqrt{\frac{b_{ij}(t)}{2\pi}} \exp\left(-\frac{b_{ij}(t)}{2} \big(\bar{x}_i(t) - \bar{x}_j(t)\big)^2 \right). \tag{54}$$

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

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

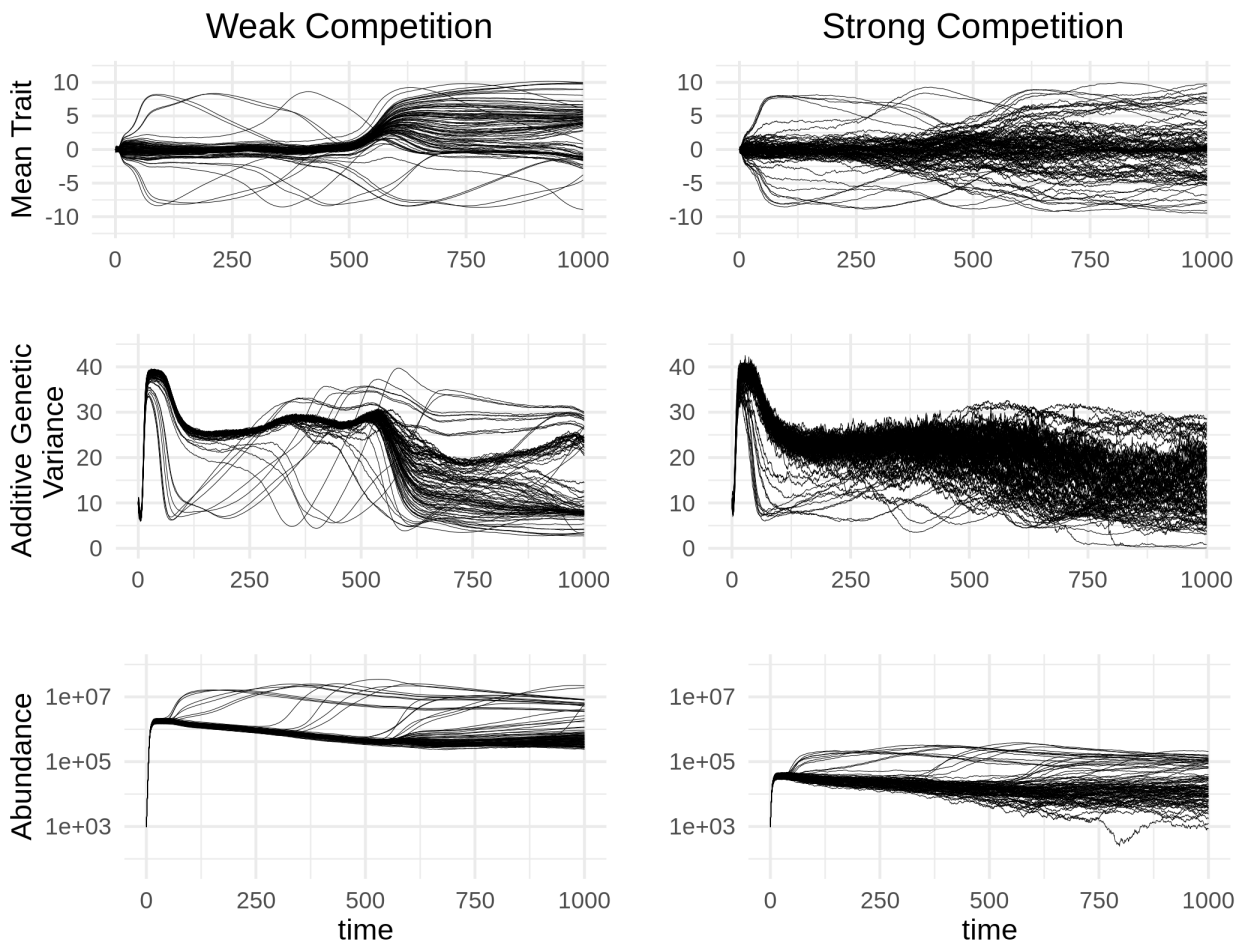

Figure 3: Temporal dynamics of mean trait (top), additive genetic variance (middle) and abundance (bottom) for the scenario of weak competition (left) and strong competition (right). Red lines indicate average trend across species.

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.

Metric of pairwise coevolution

Below we 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 §5.8 we show that the standardized stabilizing selection gradient induced by species j on species i can be computed as

$$\gamma_{ij}(t) = c_i U_i U_j N_j(t) b_{ij}(t) \left(1 - b_{ij}(t)\big(\bar{x}_i(t) - \bar{x}_j(t)\big)^2\right) \sqrt{\frac{b_{ij}(t)}{2\pi}} \exp\left(-\frac{b_{ij}(t)}{2}\big(\bar{x}_i(t) - \bar{x}_j(t)\big)^2\right). \tag{55}$$

To measure the total evolutionary change in species i induced by species j, we form the metric $\Psi_{ij} = |\beta_{ij}| + |\gamma_{ij}|$. Figure 4 displays the joint distribution of $\mathfrak{C}_{ij} = \Psi_{ij}\Psi_{ji}$, our metric of pairwise coevolution, and the product of competition coefficients $\alpha_{ij}\alpha_{ji}$ for two simulated communities, both with richness $S = 1000$. The solid contour represents the case of strong selection ($c = 5 \times 10^{-7}$) and the dashed contour represents the case of weak sensitivity to competition ($c = 1 \times 10^{-8}$). In both cases we see that $\mathfrak{C}_{ij}$ and $\alpha_{ij}\alpha_{ji}$ are, essentially, unrelated. However, this depiction is only representative of a specific set of parameters. Next, we provide analytical approximations of the covariance between selection gradients and competition coefficients and a numerical estimate for the relationship between pairwise coevolution and competition coefficients for a range of parameters.

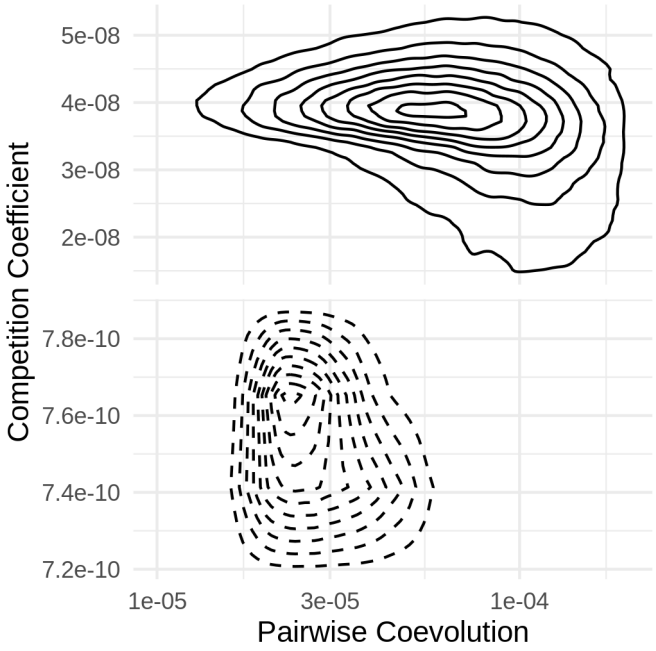

Figure 4: Bivariate distributions of competition coffecients (y-axis) and pairwise coevolution (x-axis) under the scenarios of weak competition (dashed line) and strong competition (solid line) after simulating for $t = 1000.0$ units of time. Simulations were ran for $S = 1000$ species. Parameters are the same as in Table 2, except we used $c = 1 \times 10^{-8}$ for weak competition and $c = 5 \times 10^{-7}$ for strong competition in order to account for increased species richness.

Covariance of selection and competition as a function of diversity

We now make use of the expressions derived for competition coeffecients 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 $V_{\bar{X}}$ and density $f_{\bar{X}}$. Such assumptions are typical when deriving analytical results in the field of theoretical coevolutionary community ecology (Guimarães, Jordano, and Thompson 2011; Guimarães et al. 2017; Nuismer, Jordano, and Bascompte 2012; Nuismer, Week, and Aizen 2018). If $\bar{\bar{x}}$ is near θ and $V_{\bar{X}}$ is much smaller than $|2R/a - G - \eta|$, then we may approximate r_i with

$$\bar{r} = \int_{-\infty}^{+\infty} \Big(R - \frac{a}{2}\Big((\bar{x}-\theta)^2 + G + \eta\Big)\Big) f_{\bar{X}}(\bar{x}) d\bar{x} = R - \frac{a}{2}\Big((\bar{\bar{x}}-\theta)^2 + V_{\bar{X}} + G + \eta\Big). \tag{56}$$

In SM §5.9 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. We find that the covariance between linear selection gradients and competition coeffecients are zero due to the symmetry implied by our assumptions. However, setting $\alpha(\bar{x}_i, \bar{x}_j) = \alpha_{ij}$, $\beta(\bar{x}_i, \bar{x}_j) = \beta_{ij}$ and $\gamma(\bar{x}_i, \bar{x}_j) = \gamma_{ij}$, the covariances between the magnitude of linear selection gradients and competition coefficients and between stabilizing selection gradients and competition coefficients can be written as

$$\mathrm{Cov}_{f_{\bar{X}}}(\alpha, |\beta|) = \frac{2c^2b^2U^4N}{\pi\bar{r}}\sqrt{\frac{V_{\bar{X}}}{2\pi}}\left(\frac{1}{(1+8bV_{\bar{X}})^{3/4}} - \frac{1}{(1+4bV_{\bar{X}})^{3/4}(1+2bV_{\bar{X}})^{1/2}}\right), \tag{57a}$$

$$\mathrm{Cov}_{f_{\bar{X}}}(\alpha, \gamma) = \frac{c^2b^2U^4N}{2\pi\bar{r}}(1-2bV_{\bar{X}})\left(\frac{1}{\sqrt{1+4bV_{\bar{X}}}} - \frac{1}{1+2bV_{\bar{X}}}\right). \tag{57b}$$

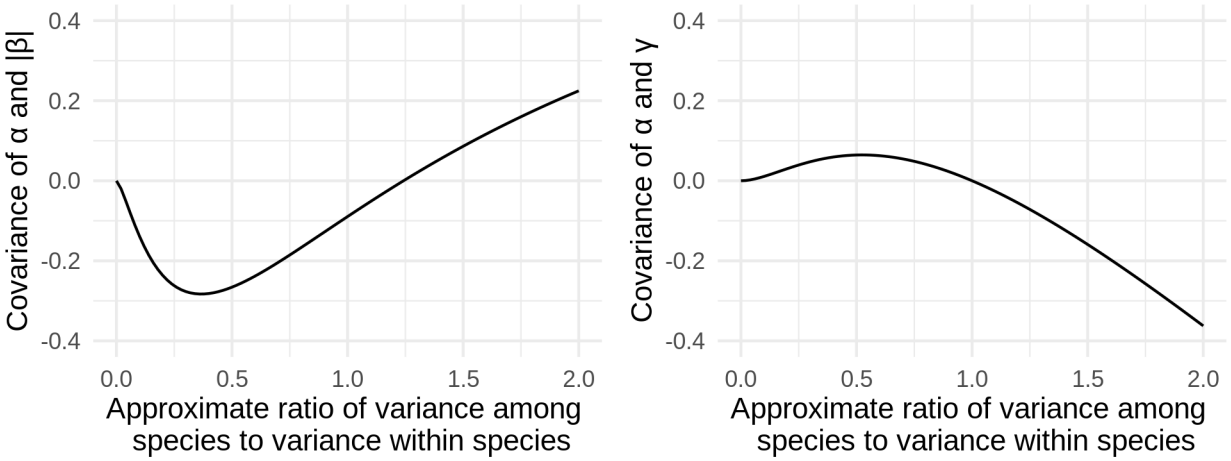

Figure 5: Curves representing the covariance between the magnitude of linear selection gradients and competition coefficients (left) and between stabilizing selection gradients and competition coefficients (right) as a function of $2bV_{\bar{X}}$ which is approximately equal to the ratio of variance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

For fixed $c, b, \bar{r}$ and N, we visualize these relationships in Figure 5. To gain insight into the relationship between selection gradients and competition coefficients, note that our assumptions in this section imply $b^{-1} = 2(\sigma^2 + w)$. If we further assume $\sigma^2 + w \approx \sigma^2$, then $2bV_{\bar{X}} \approx V_{\bar{X}}/\sigma^2$. That is, when populations are generalists and are comprised of specialist individuals, the value $2bV_{\bar{X}}$ is approximately equal to the ratio of interspecific mean trait variation to intraspecific individual trait variation. Hence, for both covariances

we see that there is no relationship between selection gradients and competition coefficients when this ratio is zero. From equation (57a) we can use numerical optimization to find that when $V_{\bar{X}}/\sigma^2 \approx 1.25$ the relationship between the magnitudes of linear selection gradients and competition coefficients disappears, but when (approximately) $V_{\bar{X}}/\sigma^2 < 1.25$ (> 1.25), this covariance becomes negative (positive). Equation (57b) states that when $V_{\bar{X}}/\sigma^2$ is approximately equal to one (or slightly larger), there is no expected relationship between competition coefficients and quadratic selection gradients. However, when $V_{\bar{X}}/\sigma^2 < 1.0$ (> 1.0), then we expect a positive (negative) relationship between α and γ. These results are true regardless of the chosen parameter values. In SM §5.9 we use simulations of system (49) to show that these results do not qualitatively differ when allowing for heterogeneous population sizes and additive genetic variances across species.

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

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

In summary, we see the relation between competition coefficients and selection is highly non-trivial and depends on the relative magnitudes of different ecological processes shaping the community. However, this does not address the relation between competition coefficients and coevolution per se. In SM §5.9 we show that calculating a formula for the covariance between competition coefficients and the metric of coevolution $\mathfrak{C}$ introduced above provides a difficult analytical challenge. Instead of confronting this challenge we build on our numerical approach used to justify analytical approximations of $\mathrm{Cov}_{f_{\bar{X}}}(\alpha, |\beta|)$ and $\mathrm{Cov}_{f_{\bar{X}}}(\alpha, \gamma)$ to approximate the correlation of α and $\mathfrak{C}$. This numerical approach inherits the assumptions of homogeneous background parameters such as the mutation rate μ and abiotic optima θ, but allows us to relax the assumption that N and G are constant across species and time.

In particular, we numerically integrated system (49) for $T_1 = 1000.0$ units of time and then continued to integrate for $T_2 = 1000.0$ units of time. We then calculate the covariance between the quantities α and $\mathfrak{C}$ across $S = 100$ species for each of the last T_2 time steps. We assume the temporal average of these covariances across the last T_2 units of time approximates the expectation at equilibrium. We repeated this approach for randomly drawn a and c until our sample size reached 1000. In Figure 6 we plot the temporally averaged values of $\mathrm{Cov}_{f_{\bar{X}}}(\alpha, \mathfrak{C})$ against the strength of competition c. Using a cubic regression, we see the correlation of coevolutionary selection gradients and competition coefficeints is negative at variance ratios below 0.5, zero at variance ratios between 0.5 and 1.0, and drops below zero again at variance ratios above 1.0.

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 selection gradients and competition coefficients, finding that these quantities exhibit interesting relationships which shed light on the feedback between the structure and dynamics of ecological communities.

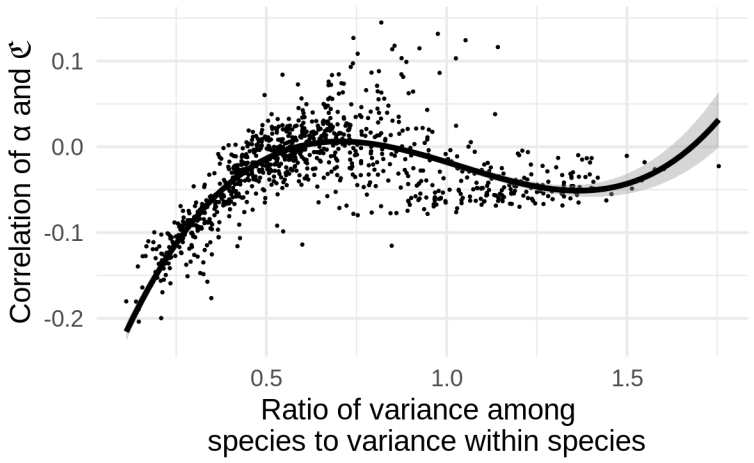

Figure 6: Numerical estimates for the correlation of the strength of coevolution measured by $\mathfrak{E}$ and competition coefficeints α plotted against the variance of mean traits among species divided by the mean variance of traits within species. Each dot represents the result from a single simulation.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

Our approach has the merit of rigorously synthesizing the dynamics of abundance and phenotypic evolution for populations experiencing demographic stochasticity. Yet, there remains biological details and their associated technical challenges that need to be confronted for gaining a more thorough and rigorous understanding of ecological communities. We touch on just four of them here and provide suggestions for future research to approach these challenges.

Limitations of diffusion limits

As noted by Feller (1951), diffusion limits provide reasonable approximations for large populations, but relatively small populations require discrete models. Hence, as a diffusion limit, SPDE (27) restricts parameter values to regions that maintain large population sizes. Although this suggests an important restriction on any model developed under this framework by implying populations are not at risk of extinction, the SDE describing abundance dynamics technically permits extinction. However, for small abundances, pathology emerges in the evolution of trait means and variances. In particular, stochastic components of the SDE describing the evolution of $\bar{x}$ and σ^2 diverge towards infinite values as $N \to 0$. Hence, studies of evolutionary rescue and colonization-extinction dynamics that incorporate phenotypic evolution should be pursued via a different approach. A natural alternative can be developed utilizing the underlying BBM that SPDE (27) is a diffusion limit of (see section 2.2.2). This process explicitly tracks individuals throughout their life-history and captures reproduction as branching events. Hence, BBM processes model population size as a non-negative integer instead of a continuously varying number. In particular, the pathological behavior described for the diffusion limit does not occur for BBM as population size goes to zero.

The genetic architecture and distributions of traits

Our treatment of inheritance and our approach to model coevolution rest on the assumptions of normally distributed breeding values and expressed phenotypes along with asexual reproduction. However, real traits are not encoded by an infinite number of loci each contributing an additive infinitesimal effect (as required by the infinitesimal model), mutations are not inherited as normally distributed deviations from parental breeding values (as required by the Gaussian descendents model) and many populations of interest exhibit non-random sexual reproduction. 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 in sexually reproducing populations. However, Gaussian approximations are convenient since they are defined by their mean and variance. Future work investigating the effects of non-normally dis-

tributed traits on the structure and dynamics of ecological communities will therefore 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).

Another 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. 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 Gavrilets 2006; Nuismer, Ridenhour, and Oswald 2007). However, work to investigate the role of complex genetic architectures in mediating feedbacks between the dynamics of population abundances and the distributions of traits mediating ecological interactions has apparently only just begun (Patel and Bürger 2019).

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 direction 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. Using sub-alpine flower communities as an example, one can observe a very strict ordering of phenology across a broad geographic range. 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? Although it would be exciting to find that these species repeatedly coevolved this pattern in each location, a more parsimoniuous 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 occuring over large spans of time. This suggests a road forward to connect the theory we have introduced to models of macroevolutionary trait evolution.

Some approaches to modelling macroevolutionary trait change simply repurpose microevolutionary models by blindly rescaling time from the units of generations to millions of years [Nuismer and Harmon (2014); Luke, can you think of others?]. Such an approach makes the implicit assumption that trait evolution is statistically self-similar (sensu Falconer 2014) so that the stochastic evolution of traits on macroevolutionary time scales has the same properties of trait evolution on microevolutionary time scales. 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

$$dX_t = a(\theta_t - X_t)dt + bdW_t \tag{58}$$

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 (if we zoom in close enough, they look like Brownian motion), 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 an application of 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 (Sterner 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)

Throughout this supplement, we set use dot notation for time derivatives so that $\dot{f}(x,t) = \frac{\partial}{\partial t} f(x,t)$ and set $\Delta = \frac{\partial^2}{\partial x^2}$, except in §5.9.1.3 where Δ represents a random variable.

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

Recall that $m(\nu, x)$ is shorthand for $m(\nu(x,t), x)$. That is, $m : [0, +\infty) \times \mathbb{R} \to \mathbb{R}$. Following our assumptions of the main text, we have that $m(y, x)$ is differentiable with respect to both x and y and there exists $r \in \mathbb{R}$ such that $m(y, x) \leq r$ for each $y \geq 0$ and $x \in \mathbb{R}$. Hence, we can apply Proposition 1.21 of Cantrell and Cosner (2004) to show that $\nu(x,t)$ remains differentiable with respect to t and twice differentiable with respect to x for all $t \geq 0$.

As in the main text, we also assume the initial condition $\nu(x, 0)$ satisfies

$$N(0) = \int_{\mathbb{R}} \nu(x,0)dx < +\infty, \tag{59}$$

$$-\infty < \bar{x}(0) = \int_{\mathbb{R}} xp(x,0)dx < +\infty, \tag{60}$$

$$\sigma^2(0) = \int_{\mathbb{R}} (x - \bar{x}(0))^2 p(x,0)dx < +\infty, \tag{61}$$

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

$$\dot{\nu}(x,t) = m(\nu, x)\nu(x,t) + \frac{\mu}{2}\Delta\nu(x,t). \tag{62}$$

Replacing m with it's upper bound $r \in \mathbb{R}$, equation (62) reduces to a simple parabolic equation that can be solved using standard techniques (Farlow 1993). In particular, when $m(\nu, x) \equiv 0$ denote the solution to (62) by $\nu_0(x,t)$. Then, denoting

$$\Phi(x,t) = \frac{\exp\left(-x^2/2\mu t\right)}{\sqrt{2\pi\mu t}}, \tag{63}$$

we have

$$\nu_0(x,t) = \int_{\mathbb{R}} \Phi(x-y,t)\nu(y,0)dy. \tag{64}$$

In the more general case, when $m(\nu, x) \equiv r \in \mathbb{R}$, equation (62) has the solution $\nu_r(x,t) = e^{rt}\nu_0(x,t)$. Hence, $\nu_r(x,t) \geq 0$ for all $x \in \mathbb{R}$ and $\int_{\mathbb{R}} \nu_r(x,t)dx = e^{rt}N(0) < +\infty$ for all $t \geq 0$. Furthermore, denoting $N_r(t) = \int_{\mathbb{R}} \nu_r(x,t)dx$, $p_r(x,t) = \nu_r(x,t)/N_r(t)$, $\bar{x}_r(t) = \int_{\mathbb{R}} xp_r(x,t)dx$ and $\sigma_r^2(t) = \int_{\mathbb{R}} (x - \bar{x}_r(t))^2 p_r(x,t)dx$, we have

$$\bar{x}_r(t) = \int_{\mathbb{R}} x \int_{\mathbb{R}} \Phi(x-y,t)p(y,0)dydx = \int_{\mathbb{R}} yp(y,0)dy = \bar{x}(0), \tag{65}$$

$$\sigma_r^2(t) = \int_{\mathbb{R}} (x - \bar{x}_r(t))^2 \int_{\mathbb{R}} \Phi(x-y,t)p(y,0)dydx = \int_{\mathbb{R}} \Big((y - \bar{x}(0))^2 + \mu t\Big)p(y,0)dy = \sigma^2(0) + \mu t. \tag{66}$$

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

$$\nu(x,t) - \nu(x,0) = \int_0^t m(\nu, x)\nu(x,s) + \frac{\mu}{2}\Delta\nu(x,s)ds > \int_0^t r\nu_r(x,s) + \frac{\mu}{2}\Delta\nu_r(x,s)ds = \nu_r(x,t) - \nu(x,0) \tag{67}$$

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

$$0 \leq \int_{\mathbb{R}} x^2 \nu(x,t)dx \leq \int_{\mathbb{R}} x^2 \nu_r(x,t)dx < +\infty \tag{68}$$

and in particular

$$0 \leq \sigma^2(t) + \bar{x}^2(t) = \frac{1}{N(t)} \int_{\mathbb{R}} x^2 \nu(x,t)dx < +\infty \tag{69}$$

for each $t \geq 0$.

5.2 The relation between diffusion and convolution with a Gaussian kernel

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

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

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

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

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

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

Hence, by the fundamental theorem of calculus,

$$f(x,t) + \int_t^{t+1} \dot{f}(x,s)ds = f(x,t+1) \tag{SM1.4}$$

$$= \int_{\mathbb{R}^d} \Phi(x-y,t+1) g(y) dy = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \Phi(x-y,1) \Phi(y-z,t) g(z) dz dy$$

$$= \int_{\mathbb{R}^d} \Phi(x-y,1) f(t,y) dy.$$

In particular,

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

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

Picking up from the main text §2.1,

$$\begin{aligned}\dot{\sigma}^2(t) = \frac{d}{dt}\int_{\mathbb{R}}(x-\bar{x}(t))^2 p(x,t)dx &= \int_{\mathbb{R}} 2(x-\bar{x}(t))\dot{\bar{x}}(t) + (x-\bar{x}(t))^2\dot{p}(x,t)dx \\ &= \int_{\mathbb{R}}(x-\bar{x}(t))^2\left((m(\nu,x)-\bar{m}(t))p(x,t) + \frac{\mu}{2}\frac{\partial^2}{\partial x^2}p(x,t)\right)dx \\ = \int_{\mathbb{R}}\left((x-\bar{x}(t))^2 - \sigma^2(t) + \sigma^2(t)\right)(m(\nu,x)-\bar{m}(t))p(x,t) &+ (x-\bar{x}(t))^2\frac{\mu}{2}\frac{\partial^2}{\partial x^2}p(x,t)dx \\ &= \mathrm{Cov}_t\Big((x-\bar{x}(t))^2, m(\nu,x)\Big) + \frac{\mu}{2}\int_{\mathbb{R}}(x-\bar{x}(t))^2\frac{\partial^2}{\partial x^2}p(x,t)dx. \quad (70)\end{aligned}$$

Applying integration by parts twice yields

$$\int_{-\infty}^{+\infty}(x-\bar{x}(t))^2\frac{\partial^2}{\partial x^2}p(x,t)dx = 2. \quad (71)$$

5.4 Simplifying covariances with fitness under the assumption of a Gaussian density

By assuming

$$p(x,t) = \frac{\exp\left(-\frac{(x-\bar{x}(t))^2}{2\sigma^2(t)}\right)}{\sqrt{2\pi\sigma^2(t)}} \quad (72)$$

we have

$$\begin{aligned}\sigma^2\left(\frac{\partial\bar{m}}{\partial\bar{x}} - \overline{\frac{\partial m}{\partial\bar{x}}}\right) &= \sigma^2\left(\frac{\partial}{\partial\bar{x}}\int_{\mathbb{R}} m(\nu,x)p(x,t)dx - \int_{\mathbb{R}} p(x,t)\frac{\partial}{\bar{x}}m(\nu,x)dx\right) \\ &= \sigma^2\int_{\mathbb{R}} m(\nu,x)\frac{\partial}{\partial\bar{x}}p(x,t)dx = \sigma^2\int_{\mathbb{R}}\frac{x-\bar{x}(t)}{\sigma^2}m(\nu,x)p(x,t)dx \\ &= \int_{\mathbb{R}}(x-\bar{x})(m(\nu,x)-\bar{m})p(x,t)dx = \mathrm{Cov}_t(m,x), \quad (73)\end{aligned}$$

and

$$\begin{aligned}2\sigma^4\left(\frac{\partial\bar{m}}{\partial\sigma^2} - \overline{\frac{\partial m}{\partial\sigma^2}}\right) &= 2\sigma^4\left(\frac{\partial}{\partial\sigma^2}\int_{\mathbb{R}} m(\nu,x)p(x,t)dx - \int_{\mathbb{R}} p(x,t)\frac{\partial}{\partial\sigma^2}m(\nu,x)dx\right) \\ = 2\sigma^4\int_{\mathbb{R}}\frac{(x-\bar{x})^2-\sigma^2}{2\sigma^4}m(\nu,x)p(x,t)dx &= \int_{\mathbb{R}}\left((x-\bar{x})^2-\sigma^2\right)\big(m(\nu,x)-\bar{m}\big)p(x,t)dx \\ &= \mathrm{Cov}_t\Big((x-\bar{x})^2, m\Big). \quad (74)\end{aligned}$$

5.5 Simulating the rescaled process and numerical evidence of approximate normality

Here we use a numerical argument to suggest, for

$$r - \frac{a}{2}(\theta - x)^2 - c\int_{\mathbb{R}} \nu(x,t)dx \leq m(\nu, x) \leq r - \frac{a}{2}(\theta - x)^2,\ \forall(\nu, x) \in C^+_{1,c}(\mathbb{R}) \times \mathbb{R}, \tag{75}$$

the density process $\nu(x,t)$ defined by SPDE (27) of the main text satisfies $\int_{\mathbb{R}}(|x| + x^2)\nu(x,t)dx < \infty$. From SM §5.1, under the assumption $m(\nu, x) = r - \frac{a}{2}(\theta - x)^2 - c\int_{\mathbb{R}} \nu(x,t)dx$, we can derive the differential equations

$$\dot{\bar{x}} = aG(\theta - \bar{x}) \tag{76a}$$

$$\dot{G} = \mu - aG^2 \tag{76b}$$

$$\dot{N} = \left(r - \frac{a}{2}\left((\theta - \bar{x})^2 + G + \eta\right) - cN\right)N. \tag{76c}$$

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

$$\hat{\bar{x}} = \theta, \tag{77a}$$

$$\hat{G} = \sqrt{\frac{\mu}{a}}, \tag{77b}$$

$$\hat{N} = \frac{1}{c}\left(r - \frac{1}{2}(\eta a + \sqrt{\mu a})\right), \tag{77c}$$

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

5.5.1 Description of simulation

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

$$\nu(x,t) = \sum_{i=1}^{N(t)} \delta(x - \xi_i(t)), \tag{78}$$

where $\delta(x - \xi_i(t))$ denotes the point mass located at $\xi_i(t)$. To allow for imperfect heritability, we also track the set of breeding values which, at time t, is denoted by $\{\gamma_1(t), \ldots, \gamma_{N(t)}(t)\}$ and should not be confused with the quadratic selection gradients discussed in §5.8 of the main text. Then the distribution of breeding values can be written as

$$\rho(g,t) = \sum_{i=1}^{N(t)} \delta(g - \gamma_i(t)). \tag{79}$$

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

$$\frac{s(1-q)}{q}, \ \frac{s(1-q)}{q^2}. \tag{80}$$

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

$$q(\nu, \xi_i) = \frac{\mathscr{W}(\nu, \xi_i)}{V}, \ s(\nu, \xi_i) = \frac{\mathscr{W}^2(\nu, \xi_i)}{V - \mathscr{W}(\nu, \xi_i)}. \tag{81}$$

The imposes the restriction $V > \mathscr{W}(\nu, \xi_i)$. For each offpsring produced by the individual with breeding value $\gamma_i(t)$, we assign indepently drawn breeding values normally distributed around $\gamma_i(t)$ with variance μ. After breeding values have been assigned, we randomly draw trait values for each offpsring as described above. For an overview of our model of inheritance, see §2.3.2 of the main text. This summarizes the basic structure of our simulation. To impose selection and density dependent growth rates, we set

$$\mathscr{W}(\nu, \xi_i) = \exp\left(r - \frac{a}{2}(\theta - \xi_i)^2 - c\int_{\mathbb{R}} \nu(x,t)dx\right), \tag{82}$$

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

Rescaling

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

$$\mathscr{W}(\nu, \xi_i) \to \mathscr{W}^{(n)}(\nu, \xi_i) = \exp\left(\frac{r}{n} - \frac{a}{2n}(\theta - \xi_i)^2 - \frac{c}{n^2}N(t)\right) = \exp\left(\frac{r}{n} - \frac{a}{2n}(\theta - \xi_i)^2 - \frac{c}{n}N^{(n)}(t)\right). \tag{83}$$

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

$$\hat{N} = \frac{n^2}{c}\left(\frac{r}{n} - \frac{1}{2n}(\eta a + \sqrt{\mu a})\right) = \frac{n}{c}\left(r - \frac{1}{2}(\eta a + \sqrt{\mu a})\right). \tag{84}$$

The deterministic equilibrium of the rescaled abundance is then

$$\hat{N}^{(n)} = \frac{1}{c}\left(r - \frac{1}{2}(\eta a + \sqrt{\mu a})\right). \tag{85}$$

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

$$\lim_{n\to\infty} n\left(\mathscr{W}^{(n)}(\nu, \xi_i) - 1\right) = r - \frac{a}{2}(\theta - \xi_i)^2 - cN^{(\infty)}(t). \tag{86}$$

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

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

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

5.5.2 Evidence of normality

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

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

Picking up from §2.2.2 of the main text, we have

$$\tilde{x}(t) = \int_{\mathbb{R}} x\nu(x,t)dx, \quad \tilde{\tilde{\sigma}}^2(t) = \int_{\mathbb{R}} x^2\nu(x,t)dx \tag{87}$$

and

$$\tilde{x}(t) = \tilde{x}(0) + \int_0^t \int_{\mathbb{R}} \nu(x,s)m(\nu,x)x + x\sqrt{V\nu(x,s)}\dot{W}(x,s)dxds, \tag{88}$$

$$\tilde{\tilde{\sigma}}^2(t) = \tilde{\tilde{\sigma}}^2(0) + \int_0^t \int_{\mathbb{R}} \nu(x,s)\left(m(\nu,x)x^2 + \mu\right) + x^2\sqrt{V\nu(x,s)}\dot{W}(x,s)dxds. \tag{89}$$

5.6.1 Derivation for trait mean

We make use of the notation

$$\begin{cases} \|N\|_2 = \sqrt{V\int_{\mathbb{R}} \nu(x,t)dx} = \sqrt{VN} \\ \|\tilde{x}\|_2 = \sqrt{V\int_{\mathbb{R}} x^2\nu(x,t)dx} \\ \langle \tilde{x}, N\rangle = V\int_{\mathbb{R}} x\nu(x,t)dx = \bar{x}VN. \end{cases} \tag{90}$$

Rewriting formula (88) as an SDE provides

$$d\tilde{x} = \left(\overline{xm}N + \frac{\mu}{2}\int_{\mathbb{R}} x\Delta\nu(x,t)dx\right)dt + \|\tilde{x}\|_2 d\tilde{W}_2, \tag{91}$$

where,

$$d\tilde{W}_2 = d\hat{\mathbf{W}}_{\sqrt{Vx^2\nu}} = \frac{1}{\|\tilde{x}\|_2}\int_{\mathbb{R}} x\sqrt{V\nu(x,t)}\dot{W}(x,t)dxdt. \tag{92}$$

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

$$d\bar{x} = d\left(\frac{\tilde{x}}{N}\right) = \frac{\tilde{x}}{N}\left(\frac{d\tilde{x}}{\tilde{x}} - \frac{dN}{N} - \frac{d\tilde{x}}{\tilde{x}}\frac{dN}{N} + \Big(\frac{dN}{N}\Big)^2\right) = \frac{d\tilde{x}}{N} - \bar{x}\frac{dN}{N} - \frac{d\tilde{x}}{N}\frac{dN}{N} + \bar{x}\Big(\frac{dN}{N}\Big)^2. \tag{93}$$

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

$$\begin{aligned} d\bar{x} &= \overline{xm}dt + \frac{\|\tilde{x}\|_2}{N}d\tilde{W}_2 - \bar{x}\left(\bar{m}dt + \sqrt{\frac{V}{N}}dW_1\right) - \frac{\langle \tilde{x}, N\rangle}{N^2}dt + \bar{x}\frac{\|N\|_2^2}{N^2}dt \\ &= (\overline{xm} - \bar{x}\bar{m})dt + \frac{\|\tilde{x}\|_2}{N}d\tilde{W}_2 - \bar{x}\sqrt{\frac{V}{N}}dW_1 - V\frac{\bar{x}}{N}dt + V\frac{\bar{x}}{N}dt \\ &= \mathrm{Cov}_t(x,m) + \frac{\|\tilde{x}\|_2}{N}d\tilde{W}_2 - \bar{x}\sqrt{\frac{V}{N}}dW_1. \end{aligned} \tag{94}$$

999 Note that

$$\begin{aligned} \frac{\|\tilde{x}\|_2}{N}d\tilde{W}_2 - \bar{x}\sqrt{\frac{V}{N}}dW_1 &= \frac{1}{N}\int_{\mathbb{R}} x\sqrt{V\nu(x,t)}\dot{W}(x,t)dx - \frac{\bar{x}}{N}\int_{\mathbb{R}}\sqrt{V\nu(x,t)}\dot{W}(x,t)dx \\ &= \int_{\mathbb{R}}\frac{x-\bar{x}}{N}\sqrt{V\nu(x,t)}\dot{W}(x,t)dx \end{aligned} \tag{95}$$

1000 and

$$\mathbb{V}\left(\int_{\mathbb{R}}\frac{x-\bar{x}}{N}\sqrt{V\nu(x,t)}\dot{W}(x,t)dx\right) = \frac{V}{N}\int_{\mathbb{R}}(x-\bar{x})^2 p(x,t)dx = V\frac{\sigma^2}{N}. \tag{96}$$

1001 Hence, by setting

$$dW_2 = \frac{\int_{\mathbb{R}}\frac{(x-\bar{x})}{N}\sqrt{V\nu(x,t)}\dot{W}(x,t)dx}{\sqrt{V\sigma^2/N}} \tag{97}$$

1002 we can write

$$d\bar{x} = \mathrm{Cov}_t(x,m)dt + \sqrt{V\frac{\sigma^2}{N}}dW_2. \tag{98}$$

1003 **5.6.2 Derivation for trait variance**

1004 We make use of the notation

$$\begin{cases} \|\tilde{\tilde{\sigma}}^2\|_2 = \sqrt{V\int_{\mathbb{R}} x^4\nu(x,t)dx} \\ \langle\tilde{\tilde{\sigma}}^2, N\rangle = V\int_{\mathbb{R}} x^2\nu(x,t)dx = \overline{x^2}VN. \end{cases} \tag{99}$$

1005 Applying formula (89) provides

$$d\tilde{\bar{\sigma}}^2 = \left(\overline{x^2 m}N + \mu N\right)dt + \|\tilde{\bar{\sigma}}^2\|_2 d\tilde{W}_3 \tag{100}$$

where

$$d\tilde{W}_3 = d\hat{\mathbf{W}}_{\sqrt{Vx^4\nu}} = \frac{1}{\|\tilde{\bar{\sigma}}^2\|_2}\int_{\mathbb{R}} x^2\sqrt{V\nu(x,t)}\dot{W}(x,t)dx. \tag{101}$$

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

$$d\overline{x^2} = d\left(\frac{\tilde{\bar{\sigma}}^2}{N}\right) = \frac{\tilde{\bar{\sigma}}^2}{N}\left(\frac{d\tilde{\bar{\sigma}}^2}{\tilde{\bar{\sigma}}^2} - \frac{dN}{N} - \frac{d\tilde{\bar{\sigma}}^2}{\tilde{\bar{\sigma}}^2}\frac{dN}{N} + \left(\frac{dN}{N}\right)^2\right) = \frac{d\tilde{\bar{\sigma}}^2}{N} - \overline{x^2}\frac{dN}{N} - \frac{d\tilde{\bar{\sigma}}^2}{N}\frac{dN}{N} + \overline{x^2}\left(\frac{dN}{N}\right)^2. \tag{102}$$

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

$$\begin{aligned} d\overline{x^2} &= \left(\overline{x^2 m} + \mu\right)dt + \frac{\|\tilde{\bar{\sigma}}^2\|_2}{N}d\tilde{W}_3 - \overline{x^2}\left(\bar{m}dt + \sqrt{\frac{V}{N}}dW_1\right) - \frac{\langle\tilde{\bar{\sigma}}^2, N\rangle}{N^2}dt + \overline{x^2}\frac{\|N\|_2^2}{N^2}dt \\ &= \left(\overline{x^2 m} - \overline{x^2}\bar{m}dt + \mu\right)dt + \frac{\|\tilde{\bar{\sigma}}^2\|_2}{N}d\tilde{W}_3 - \overline{x^2}\sqrt{\frac{V}{N}}dW_1 - \overline{x^2}\frac{V}{N}dt + \overline{x^2}\frac{V}{N}dt \\ &= \left(\mathrm{Cov}_t\left(x^2, m\right) + \mu\right)dt + \frac{\|\tilde{\bar{\sigma}}^2\|_2}{N}d\tilde{W}_3 - \overline{x^2}\sqrt{\frac{V}{N}}dW_1. \end{aligned} \tag{103}$$

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

$$\begin{aligned} d\sigma^2 &= d\overline{x^2} - 2\bar{x}d\bar{x} - (d\bar{x})^2 = \left(\mathrm{Cov}_t\left(x^2, m\right) + \mu\right)dt + \frac{\|\tilde{\bar{\sigma}}^2\|_2}{N}d\tilde{W}_3 - \overline{x^2}\sqrt{\frac{V}{N}}dW_1 \\ &\quad - 2\bar{x}\left(\mathrm{Cov}_t(x,m) + \mu dt + \sqrt{\frac{V\sigma^2}{N}}dW_2\right) - \left(\mathrm{Cov}_t(x,m)dt + \mu dt + \sqrt{\frac{V\sigma^2}{N}}dW_2\right)^2 \\ &= \left(\mathrm{Cov}_t\left(x^2 - 2\bar{x}x, m\right) + \mu\right)dt + \frac{\|\tilde{\bar{\sigma}}^2\|_2}{N}d\tilde{W}_3 - \overline{x^2}\sqrt{\frac{V}{N}}dW_1 - 2\bar{x}\sqrt{\frac{V\sigma^2}{N}}dW_2 - \left(\frac{V\sigma^2}{N}\right)dt \\ &= \left(\mathrm{Cov}_t\left(x - \bar{x})^2, m\right) + \mu - \frac{V\sigma^2}{N}\right)dt + \frac{\|\tilde{\bar{\sigma}}^2\|_2}{N}d\tilde{W}_3 - \overline{x^2}\sqrt{\frac{V}{N}}dW_1 - 2\bar{x}\sqrt{\frac{V\sigma^2}{N}}dW_2. \end{aligned} \tag{104}$$

In light of

$$\begin{aligned} \frac{\|\tilde{\bar{\sigma}}^2\|_2}{N}d\tilde{W}_3 - \overline{x^2}\sqrt{\frac{V}{N}}dW_1 - 2\bar{x}\sqrt{\frac{V\sigma^2}{N}}dW_2 &= \frac{1}{N}\int_{\mathbb{R}}\left(x^2 - \tilde{\sigma}^2 - 2\bar{x}(x-\bar{x})\right)\sqrt{V\nu(x,t)}\dot{W}(x,t)dx \\ &= \frac{1}{N}\int_{\mathbb{R}}\left((x-\bar{x})^2 - \sigma^2\right)\sqrt{V\nu(x,t)}\dot{W}(x,t)dx \end{aligned} \tag{105}$$

and

$$\frac{1}{N}\int_{\mathbb{R}}\Big(\left((x-\bar{x})^2-\sigma^2\right)\sqrt{V\nu(x,s)}\Big)^2dx=\frac{V}{N}\left(\int_{\mathbb{R}}((x-\bar{x})^4-2(x-\bar{x})^2\sigma^2+\sigma^4)p(x,t)dx\right)$$
$$=\frac{V}{N}\left(\overline{(x-\bar{x})^4}-\sigma^4\right) \quad (106)$$

we set

$$dW_3=\frac{\int_{\mathbb{R}}\left((x-\bar{x})^2-\sigma^2\right)\sqrt{V\nu(x,t)}\dot{W}(x,t)dx}{V\left(\overline{(x-\bar{x})^4}-\sigma^4\right)} \quad (107)$$

so that

$$d\sigma^2=\left(\mathrm{Cov}_t\Big((x-\bar{x})^2,m\Big)+\mu-V\frac{\sigma^2}{N}\right)dt+\sqrt{V\frac{\overline{(x-\bar{x})^4}-\sigma^4}{N}}dW_3. \quad (108)$$

Table 1 of the main text implies

$$dW_1dW_2=\frac{\int_{\mathbb{R}}(x-\bar{x})\nu(x,t)dx}{\sqrt{N\sigma^2}}dt=0, \quad (109)$$

$$dW_1dW_3=\frac{\int_{\mathbb{R}}\big((x-\bar{x})^2-\sigma^2\big)\nu(x,t)dx}{\sqrt{\overline{(x-\bar{x})^4}-\sigma^4}}dt=0, \quad (110)$$

$$dW_2dW_3=\frac{\int_{\mathbb{R}}(x-\bar{x})\big((x-\bar{x})^2-\sigma^2\big)p(x,t)dx}{\sqrt{\sigma^2\big(\overline{(x-\bar{x})^4}-\sigma^4\big)}}dt=\frac{N\overline{(x-\bar{x})^3}}{\sqrt{\sigma^2\big(\overline{(x-\bar{x})^4}-\sigma^4\big)}}dt. \quad (111)$$

In particular, when p is a Gaussian curve $dW_2dW_3=0$.

5.7 Relating fitness of expressed traits to fitness of breeding values

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

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

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

5.8 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 §5.6, 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 $i = 1, \ldots, S$ 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 $\mathscr{B}_{ij}$ a function that maps X 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 can be summarized by

$$\mathscr{B}_{ij}(X) = C_i^{\sum_{l\neq j} \mathcal{O}_{ii}(x_{ij},x_{il}) + \sum_{k\neq i}\sum_{l=1}^{n_k} \mathcal{O}_{ik}(x_{ij},x_{kl})}, \tag{112}$$

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 $\mathscr{W}_{ij}(X)$ as the average number of offspring left by the j-th individual of species i, we have

$$\mathscr{W}_{ij}(X) = \mathscr{A}_i(x_{ij})\mathscr{B}_{ij}(X), \tag{113}$$

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

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

$$\mathscr{W}_{1j}^{(\boldsymbol{n})}(X) = \mathscr{A}_1(x_{1j})^{1/n_1} \exp\left(\frac{\ln C_1}{n_1^2}\sum_{l\neq j}\mathcal{O}_{11}(x_{1j},x_{1l}) + \frac{\ln C_1}{n_1}\sum_{k\neq i}\frac{1}{n_k}\sum_{l=1}^{n_k}\mathcal{O}_{1k}(x_{1j},x_{kl})\right). \tag{114}$$

For large n_1, we have the approximation

$$\mathscr{W}_{1j}^{(\boldsymbol{n})}(X) \approx \mathscr{A}_1(x_{1j})^{1/n_1}\left(1 + \frac{\ln C_1}{n_1^2}\sum_{l\neq j}\mathcal{O}_{11}(x_{1j},x_{1l}) + \frac{\ln C_1}{n_1}\sum_{k\neq 1}\frac{1}{n_k}\sum_{l=1}^{n_k}\mathcal{O}_{1k}(x_{1j},x_{kl})\right). \tag{115}$$

Hence

$$\lim_{n_1\to\infty} n_1\left(\mathscr{W}_{1j}^{(\boldsymbol{n})}(X) - 1\right) = \ln\mathscr{A}_1(x_{1j}) + \left(\int_{\mathbb{R}}\mathcal{O}_{11}(x_{1j},y)\nu_1(y,t)dy + \sum_{k\neq 1}\frac{1}{n_k}\sum_{l=1}^{n_k}\mathcal{O}_{1k}(x_{1j},x_{kl})\right)\ln C_1. \tag{116}$$

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

$$m_1(\boldsymbol{\nu}, x_{1j}) := \lim_{\boldsymbol{n}\to\boldsymbol{\infty}} n_1 \left(\mathscr{W}_{1j}^{(\boldsymbol{n})}(X) - 1 \right) = \ln \mathscr{A}_1(x) + \left(\sum_{k=1}^{S} \int_{\mathbb{R}} \mathcal{O}_{1k}(x_{1j}, y) \nu_k(y,t) dy \right) \ln C_1. \tag{117}$$

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

$$\bar{\mathcal{O}}_{ij}(x,t) = \frac{\int_{\mathbb{R}} \mathcal{O}_{ij}(x,y)\nu_j(y,t)dy}{\int_{\mathbb{R}} \nu_j(y,t)dy}. \tag{118}$$

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

$$\bar{\mathcal{O}}_{ij}(x,t) = \frac{\int_{\mathbb{R}} \mathcal{O}_{ij}(x,y)\nu_j(y,t)dy}{\int_{\mathbb{R}} \nu_j(y,t)dy} = \frac{U_i U_j}{\sqrt{2\pi(w_i + w_j + {\sigma_j}^2(t))}} \exp\left(-\frac{(x - \bar{x}_j(t))^2}{2(w_i + w_j + {\sigma_j}^2(t))} \right). \tag{119}$$

Setting

$$\sigma_i^2(t) = G_i(t) + \eta_i, \tag{120a}$$
$$R_i = \ln\left(\frac{Q_i U_i}{\sqrt{1 + A_i w_i}} \right), \tag{120b}$$
$$a_i = \frac{A_i}{1 + A_i w_i}, \tag{120c}$$
$$\tilde{b}_{ij}(t) = \frac{1}{w_i + w_j + \sigma_j^2(t)}, \tag{120d}$$
$$c_i = -\ln C_i, \tag{120e}$$

we get

$$m_i(\boldsymbol{\nu}, x) = R_i - \frac{a_i}{2}(x - \theta_i)^2 - c_i \sum_{j=1}^{S} N_j(t) U_i U_j \sqrt{\frac{\tilde{b}_{ij}(t)}{2\pi}} e^{-\frac{\tilde{b}_{ij}(t)}{2}(x - \bar{x}_j(t))^2}. \tag{121}$$

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

For the remainder of the derivation we suppress notation indicating dependency on $\boldsymbol{\nu}$, x and t. From (121) we calculate

$$\frac{\partial m_i}{\partial \bar{x}_i} = c_i N_i U_i^2 \tilde{b}_{ii}(x - \bar{x}_i) \sqrt{\frac{\tilde{b}_{ii}}{2\pi}} e^{-\frac{\tilde{b}_{ii}}{2}(x - \bar{x}_i)^2} \tag{122}$$

$$\begin{aligned} \frac{\partial m_i}{\partial G_i} &= \frac{c_i N_i U_i^2}{2} \left(\frac{(x - \bar{x}_i)^2 - G_i - \eta_i - 2w_i}{(G_i + \eta_i + 2w_i)^2} \right) \sqrt{\frac{\tilde{b}_{ii}}{2\pi}} e^{-\frac{\tilde{b}_{ii}}{2}(x - \bar{x}_i)^2} \\ &= \frac{c_i N_i U_i^2 \tilde{b}_{ii}^2}{2} \left((x - \bar{x}_i)^2 - {\sigma_i}^2 - 2w_i \right) \sqrt{\frac{\tilde{b}_{ii}}{2\pi}} e^{-\frac{\tilde{b}_{ii}}{2}(x - \bar{x}_i)^2}. \end{aligned} \tag{123}$$

Note that

$$\sqrt{\frac{\tilde{b}_{ii}}{2\pi}}\exp\left(-\frac{\tilde{b}_{ii}}{2}(x-\bar{x}_i)^2\right)\sqrt{\frac{1}{2\pi{\sigma_i}^2}}\exp\left(-\frac{(x-\bar{x}_i)^2}{2{\sigma_i}^2}\right)$$
$$=\sqrt{\frac{1}{2\pi({\sigma_i}^2+1/\tilde{b}_{ii})}}\sqrt{\frac{{\sigma_i}^2+1/\tilde{b}_{ii}}{2\pi{\sigma_i}^2/\tilde{b}_{ii}}}\exp\left(-\frac{{\sigma_i}^2+1/\tilde{b}_{ii}}{2{\sigma_i}^2/\tilde{b}_{ii}}(x-\bar{x}_i)^2\right)$$
$$=\sqrt{\frac{1}{4\pi({\sigma_i}^2+w_i)}}\sqrt{\frac{2({\sigma_i}^2+w_i)}{2\pi{\sigma_i}^2({\sigma_i}^2+2w_i)}}\exp\left(-\frac{{\sigma_i}^2({\sigma_i}^2+2w_i)}{4({\sigma_i}^2+w_i)}(x-\bar{x}_i)^2\right). \quad (124)$$

Hence,

$$\overline{\frac{\partial m_i}{\partial \bar{x}_i}}=0, \quad (125)$$

$$\overline{\frac{\partial m_i}{\partial G_i}}=\frac{c_iN_iU_i^2}{2({\sigma_i}^2+2w_i)^2}\left(\frac{({\sigma_i}^2+2w_i){\sigma_i}^2}{2(w_i+{\sigma_i}^2)}-{\sigma_i}^2-2w_i\right)\sqrt{\frac{b_{ii}}{2\pi}}$$
$$=\frac{c_iN_iU_i^2}{2({\sigma_i}^2+2w_i)}\left(\frac{{\sigma_i}^2}{2({\sigma_i}^2+w_i)}-1\right)\sqrt{\frac{b_{ii}}{2\pi}}=-\frac{c_iN_iU_i^2b_{ii}}{2}\sqrt{\frac{b_{ii}}{2\pi}}, \quad (126)$$

where

$$b_{ij}=\frac{1}{w_i+w_j+{\sigma_i}^2+{\sigma_j}^2}. \quad (127)$$

The average fitness for species i is

$$\bar{m}_i=R_i-\frac{a_i}{2}\Big((\bar{x}_i-\theta_i)^2+G_i+\eta_i\Big)-c_i\sum_{j=1}^{S}N_jU_iU_j\sqrt{\frac{b_{ij}}{2\pi}}e^{-\frac{b_{ij}}{2}(\bar{x}_i-\bar{x}_j)^2}. \quad (128)$$

Thus,

$$\frac{\partial\bar{m}_i}{\partial\bar{x}_i}=a_i(\theta_i-\bar{x}_i)-c_i\sum_j N_jU_iU_jb_{ij}(\bar{x}_j-\bar{x}_i)\sqrt{\frac{b_{ij}}{2\pi}}e^{-\frac{b_{ij}}{2}(\bar{x}_i-\bar{x}_j)^2}, \quad (129)$$

$$\frac{\partial\bar{m}_i}{\partial G_i}=-\frac{a_i}{2}+\frac{c_i}{2}\sum_{j=1}^{S}N_jU_iU_jb_{ij}\left(1-b_{ij}(\bar{x}_i-\bar{x}_j)^2\right)\sqrt{\frac{b_{ij}}{2\pi}}e^{-\frac{b_{ij}}{2}(\bar{x}_i-\bar{x}_j)^2}. \quad (130)$$

In particular

$$\frac{\partial\bar{m}_i}{\partial G_i}-\overline{\frac{\partial m_i}{\partial G_i}}=-\frac{a_i}{2}+\frac{c_i}{2}\left(N_iU_i^2b_{ii}\sqrt{\frac{b_{ii}}{2\pi}}+\sum_{j=1}^{S}N_jU_iU_jb_{ij}\left(1-b_{ij}(\bar{x}_i-\bar{x}_j)^2\right)\sqrt{\frac{b_{ij}}{2\pi}}e^{-\frac{b_{ij}}{2}(\bar{x}_i-\bar{x}_j)^2}\right). \quad (131)$$

Applying equations (34a), (44a) and (44b) of the main text recovers system (49) of the main text.

5.9 The relation between competition coefficients and selection

5.9.1 Derivation of analytical approximations

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

$$f_1(x) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(\mu_1 - x)^2}{2\sigma_1^2}\right),\ f_2(x) = \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{(\mu_2 - x)^2}{2\sigma_2^2}\right), \tag{132}$$

then

$$f_1(x)f_2(x) = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} \exp\left(-\frac{(\mu_1 - \mu_2)^2}{2(\sigma_1^2 + \sigma_2^2)}\right) \frac{1}{\sqrt{2\pi\tilde{\sigma}^2}} \exp\left(-\frac{(\tilde{\mu} - x)^2}{2\tilde{\sigma}^2}\right), \tag{133}$$

where

$$\tilde{\mu} = \frac{\sigma_2^2\mu_1 + \sigma_1^2\mu_2}{\sigma_1^2 + \sigma_2^2},\ \tilde{\sigma}^2 = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2}. \tag{134}$$

5.9.1.1 Caclulating $\mathrm{Cov}_{f_{\bar{X}}}(\alpha, \gamma)$

Recalling

$$\alpha(\bar{x}_i, \bar{x}_j) = \frac{c}{\bar{r}}\sqrt{\frac{b}{2\pi}} \exp\left(-\frac{b}{2}\big(\bar{x}_i - \bar{x}_j\big)^2\right), \tag{135}$$

$$\gamma(\bar{x}_i, \bar{x}_j) = cNb\left(1 - b\big(\bar{x}_i - \bar{x}_j\big)^2\right)\sqrt{\frac{b}{2\pi}} \exp\left(-\frac{b}{2}\big(\bar{x}_i - \bar{x}_j\big)^2\right), \tag{136}$$

we have

$$\begin{aligned}\bar{\alpha} &= \int_{\mathbb{R}}\int_{\mathbb{R}} \alpha(\bar{x}_i, \bar{x}_j) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\ &= \frac{c}{\bar{r}} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi(b^{-1} + V_{\bar{X}})}} \exp\left(-\frac{(\bar{\bar{x}} - \bar{x}_j)^2}{2(b^{-1} + V_{\bar{X}})}\right) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j = \frac{c/\bar{r}}{\sqrt{2\pi(b^{-1} + 2V_{\bar{X}})}},\end{aligned} \tag{137}$$

$$
\begin{aligned}
\bar{\gamma} &= \int_{\mathbb{R}}\int_{\mathbb{R}} \gamma(\bar{x}_i,\bar{x}_j) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= cNb \int_{\mathbb{R}} \left\{ 1 - \left[\left(\frac{\bar{\bar{x}} + bV_{\bar{X}}\bar{x}_j}{1+bV_{\bar{X}}} - \bar{x}_j \right)^2 + \frac{V_{\bar{X}}}{1+bV_{\bar{X}}} \right] \right\} \frac{1}{\sqrt{2\pi(b^{-1}+V_{\bar{X}})}} \exp\left(-\frac{(\bar{\bar{x}}-\bar{x}_j)^2}{2(b^{-1}+V_{\bar{X}})} \right) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j \\
&= cNb \int_{\mathbb{R}} \left\{ 1 - \left[\left(\frac{\bar{\bar{x}} - \bar{x}_j}{1+bV_{\bar{X}}} \right)^2 + \frac{V_{\bar{X}}}{1+bV_{\bar{X}}} \right] \right\} \frac{1}{\sqrt{2\pi(b^{-1}+V_{\bar{X}})}} \exp\left(-\frac{(\bar{\bar{x}}-\bar{x}_j)^2}{2(b^{-1}+V_{\bar{X}})} \right) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j \\
&= cNb \left(1 - \frac{(1+bV_{\bar{X}})V_{\bar{X}}}{1+2bV_{\bar{X}}} \frac{1}{(1+bV_{\bar{X}})^2} - \frac{V_{\bar{X}}}{1+bV_{\bar{X}}} \right) \frac{1}{\sqrt{2\pi(b^{-1}+2V_{\bar{X}})}} \\
&= cNb \left[1 - \left(\frac{1}{1+2bV_{\bar{X}}} + 1 \right) \frac{V_{\bar{X}}}{1+bV_{\bar{X}}} \right] \frac{1}{\sqrt{2\pi(b^{-1}+2V_{\bar{X}})}} \\
&= cNb \left(1 - \frac{2V_{\bar{X}}}{1+2bV_{\bar{X}}} \right) \sqrt{\frac{b}{2\pi(1+2bV_{\bar{X}})}}, \qquad (138)
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}_{f_{\bar{X}}}(\alpha) &= \int_{\mathbb{R}}\int_{\mathbb{R}} \left(\bar{\alpha} - \alpha(\bar{x}_i,\bar{x}_j)\right)^2 f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \frac{c^2}{\bar{r}^2} \left(\sqrt{\frac{b}{4\pi}} \int_{\mathbb{R}}\int_{\mathbb{R}} \sqrt{\frac{b}{\pi}} \exp\left(-b(\bar{x}_i-\bar{x}_j)^2\right) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j - \frac{1}{2\pi(b^{-1}+2V_{\bar{X}})} \right) \\
&= \frac{c^2}{\bar{r}^2} \left(\sqrt{\frac{b}{4\pi}} \int_{\mathbb{R}} \sqrt{\frac{1}{2\pi(\frac{1}{2b}+V_{\bar{X}})}} \exp\left(-b(\bar{\bar{x}}-\bar{x}_j)^2\right) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j - \frac{1}{2\pi(b^{-1}+2V_{\bar{X}})} \right) \\
&= \frac{c^2}{\bar{r}^2} \left(\sqrt{\frac{b}{4\pi}} \sqrt{\frac{1}{2\pi(\frac{1}{2b}+2V_{\bar{X}})}} - \frac{1}{2\pi(b^{-1}+2V_{\bar{X}})} \right) = \frac{c^2 b}{2\pi\bar{r}^2} \left(\frac{1}{\sqrt{1+4bV_{\bar{X}}}} - \frac{1}{1+2bV_{\bar{X}}} \right), \qquad (139)
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Cov}_{f_{\bar{X}}}(\alpha,\gamma) &= \int_{\mathbb{R}}\int_{\mathbb{R}} \left(\bar{\alpha} - \alpha(\bar{x}_i,\bar{x}_j)\right)\left(\bar{\gamma} - \gamma(\bar{x}_i,\bar{x}_j)\right) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \frac{c^2 Nb}{2\bar{r}} \sqrt{\frac{b}{\pi}} \int_{\mathbb{R}}\int_{\mathbb{R}} \left(1 - b(\bar{x}_i-\bar{x}_j)^2\right) \sqrt{\frac{b}{\pi}} \exp\left(-b(\bar{x}_i-\bar{x}_j)^2\right) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j - \bar{\alpha}\bar{\gamma} \\
&= \frac{c^2 Nb}{2\bar{r}} \sqrt{\frac{b}{\pi}} \frac{1-2bV_{\bar{X}}}{\sqrt{2\pi((2b)^{-1}+2V_{\bar{X}})}} - \frac{c^2 Nb}{\bar{r}} \frac{1-2bV_{\bar{X}}}{2\pi(b^{-1}+2V_{\bar{X}})} \\
&= \frac{c^2 b^2 N}{2\pi\bar{r}} (1-2bV_{\bar{X}}) \left(\frac{1}{\sqrt{1+4bV_{\bar{X}}}} - \frac{1}{1+2bV_{\bar{X}}} \right). \qquad (140)
\end{aligned}
$$

5.9.1.2 Caclulating $\mathrm{Cov}_{f_{\bar{X}}}(\alpha, |\beta|)$

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

$$
\overline{|\beta|} = \sqrt{\frac{2\mathrm{Var}_{f_{\bar{X}}}(\beta)}{\pi}} \exp\left(-\frac{\bar{\beta}^2}{2\mathrm{Var}_{f_{\bar{X}}}(\beta)} \right) - \bar{\beta}\left[1 - 2\Phi\left(\frac{\bar{\beta}}{\mathrm{Var}_{f_{\bar{X}}}(\beta)} \right) \right] \qquad (141)
$$

$$
\mathrm{Var}_{f_{\bar{X}}}(|\beta|) = \bar{\beta}^2 + \mathrm{Var}_{f_{\bar{X}}}(\beta) - \overline{|\beta|}^2. \qquad (142)
$$

Recall that

$$\beta(\bar{x}_i,\bar{x}_j) = cNb\big(\bar{x}_i-\bar{x}_j\big)\sqrt{\frac{b}{2\pi}}\exp\left(-\frac{b}{2}\big(\bar{x}_i-\bar{x}_j\big)^2\right) \tag{143}$$

and hence

$$\begin{aligned}\bar{\beta} &= \int_{\mathbb{R}}\int_{\mathbb{R}} \beta(\bar{x}_i,\bar{x}_j) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\ &= cNb\int_{\mathbb{R}} (\bar{\bar{x}}-\bar{x}_j)\frac{1}{\sqrt{2\pi(b^{-1}+V_{\bar{X}})}}\exp\left(-\frac{(\bar{\bar{x}}-\bar{x}_j)^2}{2(b^{-1}+V_{\bar{X}})}\right) f_{\bar{X}}(\bar{x}_j)d\bar{x}_j = 0, \end{aligned} \tag{144}$$

$$\begin{aligned}\mathrm{Var}_{f_{\bar{X}}}(\beta) &= \int_{\mathbb{R}}\int_{\mathbb{R}} \big(\bar{\beta}-\beta(\bar{x}_i,\bar{x}_j)\big)^2 f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} c^2N^2b^2(\bar{x}_i-\bar{x}_j)^2\frac{b}{2\pi}\exp\left(-b(\bar{x}_i-\bar{x}_j)^2\right) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\ &= \sqrt{\frac{b}{4\pi}}c^2N^2b^2\int_{\mathbb{R}}\left[\left(\frac{\bar{\bar{x}}+2bV_{\bar{X}}\bar{x}_j}{1+2bV_{\bar{X}}}-\bar{x}_j\right)^2+\frac{V_{\bar{X}}}{1+2bV_{\bar{X}}}\right]\frac{\exp\left(-\frac{(\bar{\bar{x}}-\bar{x}_j)^2}{2(\frac{1}{2b}+V_{\bar{X}})}\right)}{\sqrt{2\pi(\frac{1}{2b}+V_{\bar{X}})}} f_{\bar{X}}(\bar{x}_j)d\bar{x}_j \\ &= \sqrt{\frac{b}{4\pi}}c^2N^2b^2\int_{\mathbb{R}}\left[\frac{(\bar{\bar{x}}-\bar{x}_j)^2}{(1+2bV_{\bar{X}})^2}+\frac{V_{\bar{X}}}{1+2bV_{\bar{X}}}\right]\frac{\exp\left(-\frac{(\bar{\bar{x}}-\bar{x}_j)^2}{2(\frac{1}{2b}+V_{\bar{X}})}\right)}{\sqrt{2\pi(\frac{1}{2b}+V_{\bar{X}})}} f_{\bar{X}}(\bar{x}_j)d\bar{x}_j \\ &= \sqrt{\frac{b}{4\pi}}c^2N^2b^2\left[\frac{(1+2bV_{\bar{X}})V_{\bar{X}}}{1+4bV_{\bar{X}}}\frac{1}{(1+2bV_{\bar{X}})^2}+\frac{V_{\bar{X}}}{1+2bV_{\bar{X}}}\right]\frac{1}{\sqrt{2\pi(\frac{1}{2b}+2V_{\bar{X}})}} \\ &= \frac{b}{\pi}\frac{c^2N^2b^2}{\sqrt{1+4bV_{\bar{X}}}}\frac{V_{\bar{X}}}{1+2bV_{\bar{X}}}\left(\frac{1}{1+4bV_{\bar{X}}}+1\right) = \frac{2c^2N^2b^3V_{\bar{X}}}{\pi(1+4bV_{\bar{X}})^{3/2}}. \end{aligned} \tag{145}$$

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

$$\overline{|\beta|} = \sqrt{\frac{2}{\pi}}\frac{cNb^{3/2}}{(1+4bV_{\bar{X}})^{3/4}}\sqrt{\frac{2V_{\bar{X}}}{\pi}} = \frac{2}{\pi}\frac{cNb^{3/2}}{(1+4bV_{\bar{X}})^{3/4}}\sqrt{V_{\bar{X}}}, \tag{146}$$

$$\mathrm{Var}_{f_{\bar{X}}}(|\beta|) = \frac{c^2N^2b^3}{(1+4bV_{\bar{X}})^{3/2}}\frac{2V_{\bar{X}}}{\pi}\left(1-\frac{2}{\pi}\right). \tag{147}$$

We also calculate

$$\begin{aligned}\mathrm{Cov}_{f_{\bar{X}}}(\alpha,\beta) &= \int_{\mathbb{R}}\int_{\mathbb{R}} \big(\bar{\alpha}-\alpha(\bar{x}_i,\bar{x}_j)\big)\big(\bar{\beta}-\beta(\bar{x}_i,\bar{x}_j)\big) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\ &= \frac{c^2Nb}{2\bar{r}}\sqrt{\frac{b}{\pi}}\int_{\mathbb{R}}\int_{\mathbb{R}} (\bar{x}_i-\bar{x}_j)\sqrt{\frac{b}{\pi}}\exp\left(-b(\bar{x}_i-\bar{x}_j)^2\right) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j = 0. \end{aligned} \tag{148}$$

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

$$\begin{aligned}
\mathrm{Cov} f_{\bar{X}}(\alpha, |\beta|) &= \int_{\mathbb{R}}\int_{\mathbb{R}} \alpha(\bar{x}_i, \bar{x}_j)|\beta(\bar{x}_i, \bar{x}_j)| f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j - \bar{\alpha}\overline{|\beta|} \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{c}{\bar{r}}\sqrt{\frac{b}{2\pi}} \exp\left(-\frac{b}{2}(\bar{x}_i - \bar{x}_j)^2\right) cNb|\bar{x}_i - \bar{x}_j|\sqrt{\frac{b}{2\pi}} \exp\left(-\frac{b}{2}(\bar{x}_i - \bar{x}_j)^2\right) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j - \bar{\alpha}\overline{|\beta|} \\
&= \frac{c^2 Nb}{\bar{r}}\sqrt{\frac{b}{4\pi}} \int_{\mathbb{R}}\int_{\mathbb{R}} |\bar{x}_i - \bar{x}_j|\sqrt{\frac{b}{\pi}} \exp\left(-b(\bar{x}_i - \bar{x}_j)^2\right) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j - \bar{\alpha}\overline{|\beta|}. \quad (149)
\end{aligned}$$

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

$$\int_{\mathbb{R}}\int_{\mathbb{R}} (\bar{x}_i - \bar{x}_j)\sqrt{\frac{b}{\pi}} \exp\left(-b(\bar{x}_i - \bar{x}_j)^2\right) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j = 0 \quad (150)$$

and

$$\begin{aligned}
&\int_{\mathbb{R}}\int_{\mathbb{R}} (\bar{x}_i - \bar{x}_j)^2 \frac{b}{\pi} \exp\left(-2b(\bar{x}_i - \bar{x}_j)^2\right) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&\qquad = \sqrt{\frac{2b}{\pi}} \int_{\mathbb{R}}\int_{\mathbb{R}} (\bar{x}_i - \bar{x}_j)^2 \frac{1}{\sqrt{2\pi\frac{1}{4b}}} \exp\left(-\frac{(\bar{x}_i - \bar{x}_j)^2}{2\frac{1}{4b}}\right) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \sqrt{\frac{2b}{\pi}} \int_{\mathbb{R}} \left[\left(\frac{\bar{\bar{x}} + 4bV_{\bar{X}}\bar{x}_j}{1 + 4bV_{\bar{X}}} - \bar{x}_j\right)^2 + \frac{V_{\bar{X}}}{1 + 4bV_{\bar{X}}}\right] \frac{1}{\sqrt{2\pi(\frac{1}{4b} + V_{\bar{X}})}} \exp\left(-\frac{(\bar{\bar{x}} - \bar{x}_j)^2}{2(\frac{1}{4b} + V_{\bar{X}})}\right) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j \\
&\qquad = \sqrt{\frac{2b}{\pi}} \int_{\mathbb{R}} \left[\left(\frac{\bar{\bar{x}} - \bar{x}_j}{1 + 4bV_{\bar{X}}}\right)^2 + \frac{V_{\bar{X}}}{1 + 4bV_{\bar{X}}}\right] \frac{1}{\sqrt{2\pi(\frac{1}{4b} + V_{\bar{X}})}} \exp\left(-\frac{(\bar{\bar{x}} - \bar{x}_j)^2}{2(\frac{1}{4b} + V_{\bar{X}})}\right) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j \\
&\qquad\qquad = \sqrt{\frac{2b}{\pi}} \left[\frac{(1 + 4bV_{\bar{X}})V_{\bar{X}}}{1 + 8bV_{\bar{X}}} \frac{1}{(1 + 4bV_{\bar{X}})^2} + \frac{V_{\bar{X}}}{1 + 4bV_{\bar{X}}}\right] \frac{1}{\sqrt{2\pi(\frac{1}{4b} + 2V_{\bar{X}})}} \\
&\qquad\qquad\qquad = \sqrt{\frac{2b}{\pi}} \frac{2V_{\bar{X}}}{1 + 8bV_{\bar{X}}} \sqrt{\frac{4b}{2\pi(1 + 8bV_{\bar{X}})}} = \frac{b}{\pi} \frac{4V_{\bar{X}}}{(1 + 8bV_{\bar{X}})^{3/2}}. \quad (151)
\end{aligned}$$

Hence

$$\int_{\mathbb{R}}\int_{\mathbb{R}} |\bar{x}_i - \bar{x}_j|\sqrt{\frac{b}{\pi}} \exp\left(-b(\bar{x}_i - \bar{x}_j)^2\right) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j = \sqrt{\frac{2}{\pi}}\sqrt{\frac{b}{\pi}\frac{4V_{\bar{X}}}{(1 + 8bV_{\bar{X}})^{3/2}}} = \frac{2}{\pi}\frac{\sqrt{2bV_{\bar{X}}}}{(1 + 8bV_{\bar{X}})^{3/4}} \quad (152)$$

and

$$\begin{aligned}\mathrm{Cov}_{f_{\bar{X}}}(\alpha,|\beta|) &= \frac{c^2Nb}{\bar{r}}\sqrt{\frac{b}{4\pi}}\frac{2}{\pi}\frac{\sqrt{2bV_{\bar{X}}}}{(1+8bV_{\bar{X}})^{3/4}} - \bar{\alpha}\overline{|\beta|} \\ &= \frac{2c^2Nb^2}{\pi\bar{r}(1+8bV_{\bar{X}})^{3/4}}\sqrt{\frac{V_{\bar{X}}}{2\pi}} - \frac{c}{\bar{r}}\sqrt{\frac{b}{2\pi(1+2bV_{\bar{X}})}}\frac{2}{\pi}\frac{cNb^{3/2}}{(1+4bV_{\bar{X}})^{3/4}}\sqrt{V_{\bar{X}}} \\ &= \frac{2c^2Nb^2}{\pi\bar{r}(1+8bV_{\bar{X}})^{3/4}}\sqrt{\frac{V_{\bar{X}}}{2\pi}} - \frac{2c^2Nb^2}{\pi\bar{r}(1+4bV_{\bar{X}})^{3/4}}\sqrt{\frac{V_{\bar{X}}}{2\pi(1+2bV_{\bar{X}})}} \\ &= \frac{2c^2Nb^2}{\pi\bar{r}}\sqrt{\frac{V_{\bar{X}}}{2\pi}}\left(\frac{1}{(1+8bV_{\bar{X}})^{3/4}} - \frac{1}{(1+4bV_{\bar{X}})^{3/4}(1+2bV_{\bar{X}})^{1/2}}\right). \end{aligned} \tag{153}$$

5.9.1.3 Starting the caclulation of $\mathrm{Cov}_{f_{\bar{X}}}(\alpha,\mathfrak{C})$

We have

$$\mathfrak{C}(\bar{x}_i,\bar{x}_j) = c^2N^2b^2\Big(|\bar{x}_i-\bar{x}_j| + \big|1-b(\bar{x}_i-\bar{x}_j)^2\big|\Big)^2\exp\left(-\frac{b}{2}(\bar{x}_i-\bar{x}_j)^2\right). \tag{154}$$

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

$$\begin{aligned}\mathfrak{C}(\delta,0) &= c^2N^2b^2\Big(|\delta| + \big|1-b\delta^2\big|\Big)^2\exp\left(-\frac{b}{2}\delta^2\right) \\ &= c^2N^2b^2\Big(\delta^2 + 2\big||\delta| - b|\delta|^3\big| + \big(1-b\delta^2\big)^2\Big)\exp\left(-\frac{b}{2}\delta^2\right). \end{aligned} \tag{155}$$

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

$$\int_{\mathbb{R}}\big||\delta| - b|\delta|^3\big|\exp\left(-\frac{b}{2}\delta^2\right)f_\Delta(\delta)d\delta. \tag{156}$$

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

5.9.2 Numerical estimates for heterogeneous N and G

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

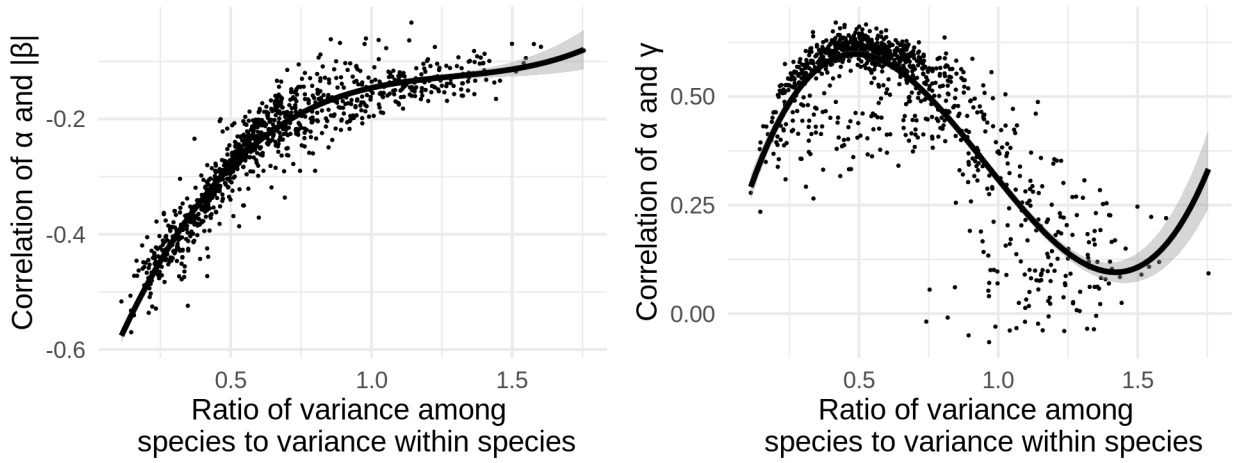

Figure 7: Numerical estimate for the correlations of selection gradients and competition coefficients.

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.

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. 1986. "On the Maintenance of Genetic Variation: Global Analysis of Kimuras Continuum-of-Alleles Model." *Journal of Mathematical Biology* 24 (3). Springer Nature: 341–51.

———. 2000. *The Mathematical Theory of Selection, Recombination, and Mutation*. Wiley.

Cantrell, Robert Stephen, and Chris Cosner. 2004. *Spatial Ecology via Reaction-Diffusion Equations*. Wiley.

Champagnat, Nicolas, Régis Ferrière, and Sylvie Méléard. 2006. "Unifying Evolutionary Dynamics: From Individual Stochastic Processes to Macroscopic Models." *Theoretical Population Biology* 69 (3). Elsevier BV: 297–321.

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. 1970. *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.

Da Prato, Giuseppe, and Jerzy Zabczyk. 2014. *Stochastic Equations in Infinite Dimensions*. Cambridge University Press.

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.

———. 1993. "Measure-Valued Markov Processes." In *École d'été de Probabilités de Saint-Flour Xxi-1991*, 1–260. Springer.

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.

Doebeli, Michael. 1996. "Quantitative Genetics and Population Dynamics." *Evolution* 50 (2). Wiley: 532–46.

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.

Falconer, Kenneth. 2014. *Fractal Geometry: Mathematical Foundations and Applications*. Wiley.

Farlow, Stanley J. 1993. *Partial Differential Equations for Scientists and Engineers*. Dover.

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.

Fisher, R. A. 1923. "XXI.-on the Dominance Ratio." *Proceedings of the Royal Society of Edinburgh* 42. Cambridge University Press: 321–41.

Fitzpatrick, Connor R., Anurag A. Agrawal, Nathan Basiliko, Amy P. Hastings, Marney E. Isaac, Michael Preston, and Marc T. J. Johnson. 2015. "The Importance of Plant Genotype and Contemporary Evolution for Terrestrial Ecosystem Processes." *Ecology* 96 (10). Wiley: 2632–42.

Fitzpatrick, Connor R., Anna V. Mikhailitchenko, Daniel N. Anstett, and Marc T. J. Johnson. 2017. "The Influence of Range-Wide Plant Genetic Variation on Soil Invertebrate Communities." *Ecography* 41 (7). Wiley: 1135–46.

Frank, S. A. 2012. "Natural Selection. IV. The Price Equation." *Journal of Evolutionary Biology* 25 (6). Wiley: 1002–19.

Fridley, Jason D. 2017. "Plant Energetics and the Synthesis of Population and Ecosystem Ecology." Edited by David Gibson. *Journal of Ecology* 105 (1). Wiley: 95–110.

Fussman, G. F., M. Loreau, and P. A. Abrams. 2007. "Eco-Evolutionary Dynamics of Communities and Ecosystems." *Functional Ecology* 21 (3). Wiley: 465–77.

Guimarães, Paulo R., Pedro Jordano, and John N. Thompson. 2011. "Evolution and Coevolution in Mutualistic Networks." *Ecology Letters* 14 (9). Wiley: 877–85.

Guimarães, Paulo R., Mathias M. Pires, Pedro Jordano, Jordi Bascompte, and John N. Thompson. 2017. "Indirect Effects Drive Coevolution in Mutualistic Networks." *Nature* 550 (7677). Springer Science; Business Media LLC: 511–14.

Harmon, Luke J., Cecilia S. Andreazzi, Florence Débarre, Jonathan Drury, Emma E. Goldberg, Ayana B. Martins, Carlos J. Melián, et al. 2019. "Detecting the Macroevolutionary Signal of Species Interactions." *Journal of Evolutionary Biology* 32 (8). Wiley: 769–82.

Harte, John. 2011. *Maximum Entropy and Ecology*. Oxford University Press.

Harte, John, and Erica A. Newman. 2014. "Maximum Information Entropy: A Foundation for Ecological Theory." *Trends in Ecology & Evolution* 29 (7). Elsevier BV: 384–89.

Hickerson, M.J., B.C. Carstens, J. Cavender-Bares, K.A. Crandall, C.H. Graham, J.B. Johnson, L. Rissler, P.F. Victoriano, and A.D. Yoder. 2010. "Phylogeography's Past, Present, and Future: 10 Years After Avise, 2000." *Molecular Phylogenetics and Evolution* 54 (1). Elsevier BV: 291–301.

Hofbauer, Josef, and Karl Sigmund. 1998. *Evolutionary Games and Population Dynamics*. Cambridge University Press.

Holt, Robert D. 1987. "On the Relation Between Niche Overlap and Competition: The Effect of Incommensurable Niche Dimensions." *Oikos* 48 (1). JSTOR: 110.

Kendall, David G. 1966. "Branching Processes Since 1873." *Journal of the London Mathematical Society* s1-41 (1). Wiley: 385–406.

Kimmel, Marek, and David E. Axelrod. 2015. *Branching Processes in Biology*. Springer New York.

Kimura, M. 1965. "A Stochastic Model Concerning the Maintenance of Genetic Variability in Quantitative Characters." *Proceedings of the National Academy of Sciences* 54 (3). Proceedings of the National Academy of Sciences: 731–36.

Kimura, M., and J. F. Crow. 1978. "Effect of Overall Phenotypic Selection on Genetic Change at Individual Loci." *Proceedings of the National Academy of Sciences* 75 (12). Proceedings of the National Academy of Sciences: 6168–71.

Kirkpatrick, Mark, Toby Johnson, and Nick Barton. 2002. "General Models of Multilocus Evolution." *Genetics* 161 (4). Genetics: 1727–50.

Kolmogorov, A.N., and S.V. Fomin. 1999. *Elements of the Theory of Functions and Functional Analysis*. v. 1. Dover.

Konno, N., and T. Shiga. 1988. "Stochastic Partial Differential Equations for Some Measure-Valued Diffusions." *Probability Theory and Related Fields* 79 (2). Springer Nature: 201–25.

Kopp, Michael, and Sergey Gavrilets. 2006. "Multilocus Gentics and the Coevolution of Quantitative Traits." *Evolution* 60 (7). Wiley: 1321–36.

Kölzsch, Andrea, Adriana Alzate, Frederic Bartumeus, Monique de Jager, Ellen J. Weerman, Geerten M. Hengeveld, Marc Naguib, Bart A. Nolet, and Johan van de Koppel. 2015. "Experimental Evidence for Inherent Lévy Search Behaviour in Foraging Animals." *Proceedings of the Royal Society B: Biological Sciences* 282 (1807). The Royal Society: 20150424.

Kraft, Nathan J. B., William K. Cornwell, Campbell O. Webb, and David D. Ackerly. 2007. "Trait Evolution, Community Assembly, and the Phylogenetic Structure of Ecological Communities." *The American Naturalist* 170 (2). University of Chicago Press: 271–83.

Krylov, N. V., and B. L. Rozovskii. 1981. "Stochastic Evolution Equations." *Journal of Soviet Mathematics* 16 (4). Springer Science; Business Media LLC: 1233–77.

Lande, Russel. 1976. "Natural Selection and Random Genetic Drift in Phenotypic Evolution." *Evolution* 30 (2). Wiley: 314–34.

———. 1980. "The Genetic Covariance between Characters Maintained by Pleiotropic Mutations." *Genetics* 94 (1): 203–15.

Lande, Russell. 1982. "A Quantitative Genetic Theory of Life History Evolution." *Ecology* 63 (3). Wiley: 607–15.

Lande, Russell, and Stevan J. Arnold. 1983. "The Measurement of Selection on Correlated Characters." *Evolution* 37 (6). JSTOR: 1210.

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.

Landis, Michael J., and Joshua G. Schraiber. 2017. "Pulsed Evolution Shaped Modern Vertebrate Body Sizes." *Proceedings of the National Academy of Sciences* 114 (50). Proceedings of the National Academy of Sciences: 13224–9.

Levins, Richard. 1968. *Evolution in Changing Environments: Some Theoretical Explorations. (MPB-2) (Monographs in Population Biology)*. Princeton University Press.

Li, Zeng-Hu. 1998. "Absolute Continuity of Measure Branching Processes with Interaction." *Chinese Journal of Applied Probability and Statistics* 14. Citeseer: 231–42.

Lion, Sébastien. 2018. "Theoretical Approaches in Evolutionary Ecology: Environmental Feedback as a Unifying Perspective." *The American Naturalist* 191 (1). University of Chicago Press: 21–44.

Loreau, Michel. 2010. *From Populations to Ecosystems: Theoretical Foundations for a New Ecological Synthesis*. Princeton University Press.

Lynch, Michael, and Bruce Walsh. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates is an imprint of Oxford University Press.

MacArthur, Robert H. 1969. "Species Packing, and what Competition Minimizes." *Proceedings of the National Academy of Sciences* 64 (4). Proceedings of the National Academy of Sciences: 1369–71.

———. 1970. "Species Packing and Competitive Equilibrium for Many Species." *Theoretical Population Biology* 1 (1). Elsevier BV: 1–11.

———. 1972. *Geographical Ecology*. Princeton University Press.

MacArthur, Robert H., and Richard Levins. 1967. "The Limiting Similarity, Convergence, and Divergence of Coexisting Species." *The American Naturalist* 101 (921). University of Chicago Press: 377–85.

Manceau, Marc, Amaury Lambert, and Hélène Morlon. 2016. "A Unifying Comparative Phylogenetic Framework Including Traits Coevolving Across Interacting Lineages." *Systematic Biology*, December. Oxford University Press (OUP), syw115.

Marx, Hannah E., Cédric Dentant, Julien Renaud, Romain Delunel, David C. Tank, and Sébastien Lavergne. 2017. "Riders in the Sky (Islands): Using a Mega-Phylogenetic Approach to Understand Plant Species Distribution and Coexistence at the Altitudinal Limits of Angiosperm Plant Life." *Journal of Biogeography* 44 (11). Wiley: 2618–30.

McPeek, Mark A. 2017. *Evolutionary Community Ecology*. Princeton University Press.

Meester, Luc De, Kristien I. Brans, Lynn Govaert, Caroline Souffreau, Shinjini Mukherjee, Héléne Vanvelk, Konrad Korzeniowski, et al. 2018. "Analyzing Eco-Evolutionary Dynamics - the Challenging Complexity of the Real World." *Functional Ecology*. Wiley.

Méléard, M, and S Roelly. 1992. "Interacting Branching Measure Processes." *Stochastic Partial Differential Equations and Applications (G. Da Prato and L. Tubaro, Eds.)*, 246–56.

———. 1993. "Interacting Measure Branching Processes. Some Bounds for the Support." *Stochastics and Stochastic Reports* 44 (1-2). Informa UK Limited: 103–21.

Mubayi, Anuj, Christopher Kribs, Viswanathan Arunachalam, and Carlos Castillo-Chavez. 2019. "Studying Complexity and Risk Through Stochastic Population Dynamics: Persistence, Resonance, and Extinction in Ecosystems." In *Handbook of Statistics*, 157–93. Elsevier.

Nowak, Martin A. 2006. *Evolutionary Dynamics: Exploring the Equations of Life*. Belknap Press.

Nuismer, Scott L., Michael Doebeli, and Danny Browning. 2005. "The Coevolutionary Dynamics of Antagonistic Interactions Mediated by Quantitative Traits with Evolving Variances." *Evolution* 59 (10). The Society for the Study of Evolution: 2073.

Nuismer, Scott L., and Luke J. Harmon. 2014. "Predicting Rates of Interspecific Interaction from Phylogenetic Trees." Edited by Jerome Chave. *Ecology Letters* 18 (1). Wiley: 17–27.

Nuismer, Scott L., Pedro Jordano, and Jordi Bascompte. 2012. "Coevolution and the Architecture of Mutualistic Networks." *Evolution* 67 (2). Wiley: 338–54.

Nuismer, Scott L., Benjamin J. Ridenhour, and Benjamin P. Oswald. 2007. "Antagonistic Coevolution Mediated by Phenotypic Differenes Between Quantitative Traits." *Evolution* 61 (8). Wiley: 1823–34.

Nuismer, Scott L., Bob Week, and Marcelo A. Aizen. 2018. "Coevolution Slows the Disassembly of Mutualistic Networks." *The American Naturalist* 192 (4). University of Chicago Press: 490–502.

Nuland, Michael E. Van, Ian M. Ware, Joseph K. Bailey, and Jennifer A. Schweitzer. 2019. "Ecosystem Feedbacks Contribute to Geographic Variation in Plant-Soil Eco-Evolutionary Dynamics Across a Fertility Gradient." Edited by Franziska Brunner. *Functional Ecology* 33 (1). Wiley: 95–106.

Page, Karen M., and Martin A. Nowak. 2002. "Unifying Evolutionary Dynamics." *Journal of Theoretical Biology* 219 (1). Elsevier BV: 93–98.

Parachnowitsch, Amy L, Jessamyn S Manson, and Nina Sletvold. 2018. "Evolutionary Ecology of Nectar." *Annals of Botany* 123 (2). Oxford University Press (OUP): 247–61.

Patel, Swati, and Reinhard Bürger. 2019. "Eco-Evolutionary Feedbacks Between Prey Densities and Linkage Disequilibrium in the Predator Maintain Diversity." *Evolution* 73 (8). Wiley: 1533–48.

Perkins, Edwin A. 1991. "Conditional Dawson-Watanabe Processes and Fleming-Viot Processes." In *Seminar on Stochastic Processes, 1991*, 143–56. Birkhäuser Boston.

———. 1992. "Measure-Valued Branching Diffusions with Spatial Interactions." *Probability Theory and Related Fields* 94 (2). Springer Science; Business Media LLC: 189–245.

———. 1995. *On the Martingale Problem for Interactive Measure-Valued Branching Diffusions*. Amer Mathematical Society.

Price, George R. 1970. "Selection and Covariance." *Nature* 227 (5257). Springer Nature: 520–21.

Queller, David C. 2017. "Fundamental Theorems of Evolution." *The American Naturalist* 189 (4). University of Chicago Press: 345–53.

Reimers, Mark. 1989. "One Dimensional Stochastic Partial Differential Equations and the Branching Measure Diffusion." *Probability Theory and Related Fields* 81 (3). Springer Nature: 319–40.

Robertson, Alan. 1966. "A Mathematical Model of the Culling Process in Dairy Cattle." *Animal Science* 8 (1). Cambridge University Press: 95–108.

Roughgarden, Joan. 1979. *Theory of Population Genetics and Evolutionary Ecology: An Introduction*. Macmillan.

Rudman, Seth M., Matthew A. Barbour, Katalin Csilléry, Phillip Gienapp, Frederic Guillaume, Nelson G. Hairston Jr, Andrew P. Hendry, et al. 2017. "What Genomic Data Can Reveal About Eco-Evolutionary Dynamics." *Nature Ecology & Evolution* 2 (1). Springer Nature: 9–15.

Schuster, Peter, and Karl Sigmund. 1983. "Replicator Dynamics." *Journal of Theoretical Biology* 100 (3). Elsevier BV: 533–38.

Skovmand, Lotte H., Charles C.Y. Xu, Maria R. Servedio, Patrik Nosil, Rowan D.H. Barrett, and Andrew P. Hendry. 2018. "Keystone Genes." *Trends in Ecology & Evolution* 33 (9). Elsevier BV: 689–700.

Sterner, R.W., and J.J. Elser. 2008. "Ecological Stoichiometry: Overview." In *Encyclopedia of Ecology*, 1101–16. Elsevier.

Taylor, Peter D., and Leo B. Jonker. 1978. "Evolutionary Stable Strategies and Game Dynamics." *Mathematical Biosciences* 40 (1-2). Elsevier BV: 145–56.

Tilman, David. 1982. *Resource Competition and Community Structure*. Princeton University Press.

Turelli, Michael. 1984. "Heritable Genetic Variation via Mutation-Selection Balance: Lerchs Zeta Meets the Abdominal Bristle." *Theoretical Population Biology* 25 (2). Elsevier: 138–93.

———. 1986. "Gaussian Versus Non-Gaussian Genetic Analyses of Polygenic Mutation-Selection Balance." In *Evolutionary Processes and Theory*, 607–28. Academic Press.

———. 2017. "Commentary: Fisher's Infinitesimal Model: A Story for the Ages." *Theoretical Population Biology* 118 (December). Elsevier BV: 46–49.

Turelli, Michael, and Nick Barton. 1994. "Genetic and statistical analyses of strong selection on polygenic traits: what, me normal?" *Genetics* 138 (3): 913–41.

Walsh, John B. 1986. "An Introduction to Stochastic Partial Differential Equations." In *Lecture Notes in Mathematics*, 265–439. Springer Berlin Heidelberg.

Watanabe, Shinzo. 1968. "A Limit Theorem of Branching Processes and Continuous State Branching Processes." *Journal of Mathematics of Kyoto University* 8 (1). Duke University Press: 141–67.

Wright, Sewall. 1931. "Evolution in Mendelian Populations." *Genetics* 16 (2). Genetics: 97–159.

Xiao, Xiao, Daniel J. McGlinn, and Ethan P. White. 2015. "A Strong Test of the Maximum Entropy Theory of Ecology." *The American Naturalist* 185 (3). University of Chicago Press: E70–E80.