Supplementary Material for
A White Noise Approach to Evolutionary Ecology

Bob Week Scott L. Nuismer Luke J. Harmon Stephen M. Krone

Contents

1 Solutions to DAGA

1.1 Sufficient Conditions for Finite Moments Under DAGA

In this section we investigate the conditions under which the trait mean $\bar{x}(t)$, trait variance $\sigma^2(t)$ and abundance $N(t)$ remain finite for finite time $t \geq 0$ when they evolve according to DAGA.

The growth rate expression $m(\nu, x)$ is actually shorthand for the more accurate expression $m\big((K\nu)(x,t), x\big)$ where K is an operator that accounts for nonlocal effects, such as resource competition, on growth rates (Champagnat et al., 2006; Volpert, 2014). In particular, we consider operators of the form $(K\nu)(x,t) = \int_{\mathbb{R}} \kappa(x-y)\nu(y,t)dy$ for some non-negative and bounded function κ. Hence, $K\nu$ is a non-negative function whenever ν is a non-negative function. In particular, this implies m is actually a bivariate function of two real numbers $h \geq 0$ and $x \in \mathbb{R}$. To ensure existence and uniqueness of solutions to DAGA, we assume the existence of $R \in \mathbb{R}$ such that $m(h,x) \leq R$ across all $h \geq 0$ and $x \in \mathbb{R}$ along with a twice continuously differentiable and integrable initial condition $u(x)$ that satisfies

$$0 < \int_{\mathbb{R}} (|x| + x^2) u(x) dx < +\infty. \tag{SM.1}$$

In particular, this implies finite initial moments $N(0), |\bar{x}(0)|, \sigma^2(0) < +\infty$ and positive initial abundance and trait variance $0 < N(0), \sigma^2(0)$. Following DAGA, we consider the Cauchy problem

$$\begin{cases} \dot{\nu}(x,t) = m(\nu,x)\nu(x,t) + \frac{\mu}{2}\Delta\nu(x,t) & t > 0 \\ \nu(x,0) = u(x) & t = 0. \end{cases} \tag{SM.2}$$

We assume the operator F defined by $\nu(x,t) \mapsto m(\nu,x)\nu(x,t)$ is locally Lipschitz continuous. This implies that, given two abundance densities $\nu_1(x), \nu_2(x)$ with total abundances $N_1 = \int \nu_1(x)dx$, $N_2 = \int \nu_2(x)dx$ and a positive number $M > 0$, there exists a constant $L_M > 0$ depending on M such that when $N_1, N_2 \leq M$, then

$$\int_{\mathbb{R}} |m(\nu_1,x)\nu_1(x) - m(\nu_2,x)\nu_2(x)|\, dx \leq L_M \int_{\mathbb{R}} |\nu_1(x) - \nu_2(x)| dx. \tag{SM.3}$$

To be specific, we define the domain of the Laplacian as $D(\Delta) = C^2(\mathbb{R}) \cap L^1(\mathbb{R})$ with the norm $\|v\| = \int_{\mathbb{R}} |v(x)|dx$ and define F as an operator on $D(\Delta)$. That is, F acts on functions that are integrable and twice continuously differentiable. Then Theorem 2.5.6 of Zheng (2004) implies for some maximal $T > 0$, the Cauchy problem (SM.2) admits a unique classical solution $\nu(x,t)$ for $t \in [0,T)$. This implies the solution $\nu(x,t)$ is continuously differentiable with respect to t and twice continuously differentiable with respect to x for all $t \in [0,T)$. Furthermore, Theorem 2.5.6 of Zheng (2004) implies either $T = +\infty$ and $N(t) < +\infty$ for all $t > 0$ or $T < +\infty$ and $\lim_{t \uparrow T} N(t) = +\infty$. The latter case corresponds to the notion of "blow-up".

In this section we show that our assumptions on mutation, initial conditions and growth rate, $m(h,x) \leq R$ for all $h \geq 0$ and $x \in \mathbb{R}$ in particular, implies $T = +\infty$ and $N(t) < +\infty$ for all $t > 0$. Replacing m with the upper bound $R \in \mathbb{R}$, PDE (SM.2) reduces to a simple parabolic equation that can be solved using elementary techniques (Farlow, 1993). In particular, when $m(h,x) \equiv R = 0$ we denote the solution to (SM.2) by $\nu_0(x,t)$. Then, denoting

$$\Phi(x,t) = \frac{\exp\left(-x^2/2\mu t\right)}{\sqrt{2\pi\mu t}}, \tag{SM.4}$$

we have

$$\nu_0(x,t) = \int_{\mathbb{R}} \Phi(x-y,t)u(y)dy. \tag{SM.5}$$

In the more general case, when $m(\nu, x) \equiv R \in \mathbb{R}$, equation (SM.2) 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

$$N_R(t) = \int_{\mathbb{R}} \nu_R(x,t)dx, \tag{SM.6a}$$

$$p_R(x,t) = \nu_R(x,t)/N_R(t), \tag{SM.6b}$$

$$\bar{x}_R(t) = \int_{\mathbb{R}} x p_R(x,t)dx, \tag{SM.6c}$$

$$\sigma_R^2(t) = \int_{\mathbb{R}} (x - \bar{x}_R(t))^2 p_R(x,t)dx, \tag{SM.6d}$$

we have

$$\bar{x}_R(t) = \int_{\mathbb{R}} x \int_{\mathbb{R}} \Phi(x-y,t)p_R(y,0)dydx = \int_{\mathbb{R}} y p_R(y,0)dy = \bar{x}(0), \tag{SM.7}$$

$$\sigma_R^2(t) = \int_{\mathbb{R}} (x - \bar{x}_R(t))^2 \int_{\mathbb{R}} \Phi(x-y,t)p_R(y,0)dydx = \int_{\mathbb{R}} \Big((y-\bar{x}(0))^2 + \mu t\Big) p_R(y,0)dy = \sigma^2(0) + \mu t. \tag{SM.8}$$

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 $t > 0$ such that $\nu(x,t) > \nu_R(x,t)$. Then

$$\nu(x,t) - u(x) = \int_0^t m(\nu,x)\nu(x,s) + \frac{\mu}{2}\Delta\nu(x,s)ds > \int_0^t R\nu_R(x,s) + \frac{\mu}{2}\Delta\nu_R(x,s)ds = \nu_R(x,t) - u(x) \tag{SM.9}$$

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

$$0 \leq \int_{\mathbb{R}} x^2\nu(x,t)dx \leq \int_{\mathbb{R}} x^2\nu_R(x,t)dx < +\infty. \tag{SM.10}$$

Furthermore, since $\nu(x,t)$ is a classical solution of Cauchy problem (SM.2) and since we assumed $N(0) > 0$, we conclude $N(t) > 0$ for all finite $t > 0$. Hence, for each $t > 0$,

$$0 \leq \sigma^2(t) + \bar{x}^2(t) = \frac{1}{N(t)} \int_{\mathbb{R}} x^2\nu(x,t)dx < +\infty. \tag{SM.11}$$

1.2 Equilibrium moments for a population experiencing logistic growth and stabilizing selection under DAGA

Here we show, under DAGA, the population moments $N, \bar{x}$ and σ^2 evolve to the following asymptotically stable equilibrium

$$\hat{N} = \tfrac{1}{c}(R - \tfrac{1}{2}\sqrt{a\mu}), \tag{SM.12a}$$

$$\hat{\bar{x}} = \theta, \tag{SM.12b}$$

$$\hat{\sigma}^2 = \sqrt{\tfrac{\mu}{a}}, \tag{SM.12c}$$

given the initial condition $N(0) > 0$ and growth rate

$$m(\nu, x) = R - \frac{a}{2}(\theta - x)^2 - c\int_{\mathbb{R}} \nu(y,t)dy = R - \frac{a}{2}(\theta - x)^2 - cN(t) \tag{SM.13}$$

that satisfies $\theta \in \mathbb{R}$, $a, c, \mu > 0$ and $R > \frac{1}{2}\sqrt{\mu a}$. Following equation (SM.13), mean fitness becomes

$$\bar{m}(t) = R - \frac{a}{2}\Big[\big(\theta - \bar{x}(t)\big)^2 + \sigma^2(t)\Big] - cN(t), \tag{SM.14}$$

and the ODE for $N(t)$ becomes

$$\frac{d}{dt}N(t) = \Big\{R - \frac{a}{2}\Big[\big(\theta - \bar{x}(t)\big)^2 + \sigma^2(t)\Big] - cN(t)\Big\}N(t). \tag{SM.15}$$

Solving for equilibrium total abundance $\hat{N}$ amounts to setting $\frac{d}{dt}N(t) = 0$ and solving for $N(t)$. Ignoring the equilibrium $N(t) = 0$, this reduces to solving $\bar{m}(t) = 0$ for $N(t)$, which, assuming finite equilibrial $\hat{\bar{x}}$ and $\hat{\sigma}^2$, returns

$$\hat{N} = \frac{1}{c}\left\{R - \frac{a}{2}\left[(\theta - \hat{\bar{x}})^2 - \hat{\sigma}^2\right]\right\}. \tag{SM.16}$$

Unfortunately, deriving ODE for $\bar{x}(t)$ and $\sigma^2(t)$ leads to expressions involving higher moments and finding ODE for these higher moments will lead to expressions involving yet even higher moments. To avoid this infinite regression, we find the equilibrium abundance density $\hat{\nu}(x)$ by solving $\frac{\partial}{\partial t}\nu(x,t) = 0$ for $\nu(x,t)$. This implies the following ordinary differential equation

$$\frac{d^2}{dx^2}\hat{\nu}(x) = \left(\frac{2c}{\mu}\hat{N} + \frac{a}{\mu}(\theta - x)^2 - \frac{2R}{\mu}\right)\hat{\nu}(x) \tag{SM.17}$$

which has the solution

$$\hat{\nu}(x) = \frac{\hat{N}}{\sqrt{2\pi}}\left(\frac{a}{\mu}\right)^{\frac{1}{4}} \exp\left(-\sqrt{\frac{a}{\mu}}\frac{(\theta - x)^2}{2}\right). \tag{SM.18}$$

From this expression we infer $\hat{\bar{x}} = \theta$ and $\hat{\sigma}^2 = \sqrt{\frac{\mu}{a}}$. Hence $\hat{N} = \frac{1}{c}\left(R - \frac{1}{2}\sqrt{a\mu}\right)$.

To show this equilibrium is locally stable, we use linear stability analysis. Since $\hat{\nu}(x)$ is Gaussian, we do not run into the same issue with higher moments as above. Furthermore, following equations (23) of the main text, ODE for $\bar{x}(t)$ and $\sigma^2(t)$ can now be expressed as

$$\frac{d}{dt}\bar{x}(t) = \sigma^2(t)\left(\frac{\partial\bar{m}(t)}{\partial\bar{x}(t)} - \overline{\frac{\partial m(t)}{\partial\bar{x}(t)}}\right) = a\sigma^2(t)(\theta - \bar{x}(t)), \tag{SM.19a}$$

$$\frac{d}{dt}\sigma^2(t) = 2\sigma^4(t)\left(\frac{\partial\bar{m}(t)}{\partial\sigma^2(t)} - \overline{\frac{\partial m(t)}{\partial\sigma^2(t)}}\right) + \mu = \mu - a\sigma^4(t). \tag{SM.19b}$$

These expressions confirm our findings that $\hat{\bar{x}} = \theta$ and $\hat{\sigma}^2 = \sqrt{\frac{\mu}{a}}$. Furthermore, calculating

$$\frac{\partial}{\partial\sigma^2(t)}\frac{d}{dt}\sigma^2(t) = -2a\sigma^2(t) \tag{SM.20}$$

and evaluating at $\sigma^2(t) = \hat{\sigma}^2$ demonstrates the equilibrium phenotypic variance is stable when $a, \mu > 0$. Hence, calculating

$$\frac{\partial}{\partial \bar{x}(t)} \frac{d}{dt} \bar{x}(t) = -a\sigma^2(t) \tag{SM.21}$$

and evaluating at $\sigma^2(t) = \hat{\sigma}^2$ and $\bar{x}(t) = \hat{\bar{x}}$ demonstrates the equilibrium phenotypic mean is stable when $a, \mu > 0$. Finally, calculating

$$\frac{\partial}{\partial N(t)} \frac{d}{dt} N(t) = R - \frac{a}{2}\Big[\big(\theta - \bar{x}(t)\big)^2 + \sigma^2(t)\Big] - 2cN(t) \tag{SM.22}$$

and evaluating at $\sigma^2(t) = \hat{\sigma}^2$, $\bar{x}(t) = \hat{\bar{x}}$ and $N(t) = \hat{N}$ demonstrates the equilibrium total abundance is stable when $a, c, \mu > 0$, and $R > \frac{1}{2}\sqrt{a\mu}$.

2 Space-Time White Noise

Here we introduce some heuristics for performing calculations with respect to space-time white noise processes. We then compare these heuristics to results rigorously presented by Da Prato and Zabczyk (2014).

2.1 Heuristics for the White Noise Calculus

We define $\mathscr{N}_2$ as the set of stochastic processes $f(x,t)$ that are continuous in t and satisfy $\mathbb{E}\left(\int_0^t \int_{\mathbb{R}} |f(x,s)|^2 dxds\right) < +\infty$ 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

$$\|f\|_t = \sqrt{\mathbb{E}\left(\int_0^t \int_{\mathbb{R}} |f(x,s)|^2 dxds\right)}, \tag{SM.23}$$

and make use of the convention $f = g$ if $\|f - g\|_t = 0$ for all $t \geq 0$. Since the abundance density process $\nu(x,t)$ satisfying SAGA is continuous in t and integrable with respect to x for each $t \geq 0$, it also satisfies $\sqrt{\nu} \in \mathscr{N}_2$. This enables us to utilize the heuristics developed in this section for the calculation of SDE describing the stochastic dynamics of $N(t)$, $\bar{x}(t)$ and $\sigma^2(t)$. To begin developing these heuristics, we introduce a generalized process that captures the essence of space-time white noise in a mathematically tractable format.

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 $t, t_1, t_2 \geq 0$,

$$\mathbb{E}\big(\mathbf{W}_t(f)\big) = \mathbb{E}\big(\mathbf{W}_t(g)\big) = 0, \tag{SM.24a}$$

$$\mathbb{C}\big(\mathbf{W}_{t_1}(f), \mathbf{W}_{t_2}(g)\big) = \mathbb{E}\left(\int_0^{t_1 \wedge t_2} \int_{\mathbb{R}} f(x,s)g(x,s)dxds\right), \tag{SM.24b}$$

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 expectations and covariances with respect to phenotypic diversity such as $\bar{x}$ and $\mathrm{Cov}(m, x)$. In particular, since we model phenotypic diversity as a random process, the phenotypic moments $\bar{x}$ and $\mathrm{Cov}(m, x)$ are random variables and $\mathbb{E}(\bar{x})$, $\mathbb{E}(\mathrm{Cov}(m, x))$ denote the expectations of these random variables with respect to the underlying probability space.

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

$$\text{``}\int_0^t \int_{\mathbb{R}} f(x, s)\dot{W}(x, s)dxds\text{''} = \mathbf{W}_t(f),\ \forall\ f \in \mathscr{N}_2,\ t \geq 0. \tag{SM.25}$$

We place quotations in the above expression to emphasize its informal nature and that it should not be confused with classical Riemann integration. Using this informal notation, equations (SM.24a) and (SM.24b) can be rewritten as

$$\mathbb{E}\left(\int_0^t \int_{\mathbb{R}} f(x, s)\dot{W}(x, s)dxds\right) = 0, \tag{SM.26a}$$

$$\mathbb{C}\left(\int_0^{t_1} \int_{\mathbb{R}} f(x, s)\dot{W}(x, s)dxds, \int_0^{t_2} \int_{\mathbb{R}} g(x, s)\dot{W}(x, s)dxds\right) = \int_0^{t_1 \wedge t_2} \int_{\mathbb{R}} f(x, s)g(x, s)dxds. \tag{SM.26b}$$

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

$$\hat{f}(x, t) = \frac{f(x, t)}{\sqrt{\int_{\mathbb{R}} f^2(y, t)dy}} \quad \text{and} \quad d\hat{\mathbf{W}}_t(f) = \left(\int_{\mathbb{R}} \hat{f}(x, t)\dot{W}(x, t)dx\right)dt \tag{SM.27}$$

so that

$$\int_0^t d\hat{\mathbf{W}}_s(f) = \int_0^t \int_{\mathbb{R}} \frac{f(x, s)}{\sqrt{\int_{\mathbb{R}} f^2(y, s)dy}}\dot{W}(x, s)dxds. \tag{SM.28}$$

This implies

$$\mathbb{E}\left(\int_0^t d\hat{\mathbf{W}}_s(f)\right) = 0,\ \mathbb{C}\left(\int_0^{t_1} d\hat{\mathbf{W}}_s(f), \int_0^{t_2} d\hat{\mathbf{W}}_s(f)\right) = t_1 \wedge t_2 \tag{SM.29}$$

and in particular, 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 (SM.26b) effectively extends Itô's multiplication table to Table S1.

Table S1: An extension of Itô's multiplication table.

$\times$	$d\hat{\mathbf{W}}_t(f)$	$d\hat{\mathbf{W}}_t(g)$	dt
$d\hat{\mathbf{W}}_t(f)$	dt	$\left(\int_{\mathbb{R}} \hat{f}(x, t)\hat{g}(x, t)dx\right)dt$	0
$d\hat{\mathbf{W}}_t(g)$	$\left(\int_{\mathbb{R}} \hat{f}(x, t)\hat{g}(x, t)dx\right)dt$	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 §3.4 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 show how our heuristics of space-time white noise relate to results rigorously presented by Da Prato and Zabczyk (2014).

2.2 Comparing the White Noise Heuristics to the Infinite-Dimensional Stochastic Calculus of Da Prato and Zabczyk (2014)

Our above approach is inspired by the treatment provided in §4.2 of Da Prato and Zabczyk (2014). Here the authors develop a stochastic integral of operator-valued processes. In particular, they consider processes indexed by time $t \geq 0$ valued as Hilbert-Schmidt operators $\Phi(t)$ and, denoting $\Phi^*(t)$ the adjoint of $\Phi(t)$, define the norm

$$\|\Phi\|_t = \sqrt{\mathbb{E}\left(\int_0^t \mathrm{Tr}[\Phi(s)\Phi^*(s)]ds\right)}, \ t \geq 0. \tag{SM.30}$$

In our case we only consider the so-called multiplication operators. That is, processes that consist of operators $\Phi(t)$ having the form $\Phi(t)g(x) = \varphi(x,t)g(x)$ such that $\varphi(\cdot,t) \in L^2(\mathbb{R})$ a.s. for each $t \geq 0$. In this case $\Phi(t) = \Phi^*(t)$ and

$$\|\Phi\|_t = \|\varphi\|_t = \sqrt{\mathbb{E}\left(\int_0^t \int_{\mathbb{R}} \varphi^2(x,s)dxds\right)}, \ t \geq 0. \tag{SM.31}$$

Da Prato and Zabczyk (2014) form the space $\mathscr{N}_W^2(0,T)$ of Hilbert-Schmidt operator-valued predictable processes $\Phi(t)$ that satisfy $\|\Phi\|_T < +\infty$ for some $T > 0$. This corresponds to our more specialized space $\mathscr{N}_2$ that consists of $L^2(\mathbb{R})$-valued processes $\varphi(x,t)$ such that $\|\varphi\|_t < +\infty$ for all $t \geq 0$. In their treatment, $W(t)$ plays a similar role to our generalized process $\mathbf{W}_t$. For $\Phi \in \mathscr{N}_W^2(0,T)$, they denote the stochastic integral for $t \in [0,T]$ by $\Phi \cdot W(t)$. Hence, for $\Phi(t)g(x) = \varphi(x,t)g(x)$ as above, $\mathbf{W}_t(\varphi) = \Phi \cdot W(t)$. The authors then prove the following:

Proposition 4.28 Assume that $\Phi_1, \Phi_2 \in \mathscr{N}_W^2(0,T)$. Then

$$\mathbb{E}(\Phi_i \cdot W(t)) = 0, \ \ \mathbb{E}(\|\Phi_i \cdot W(t)\|^2) < +\infty, \ \ \forall \ t \in [0,T].$$

Corollory 4.29 Under the same assumptions as Proposition 4.28,

$$\mathbb{C}(\Phi_1 \cdot W(t), \Phi_2 \cdot W(s)) = \mathbb{E}\left(\int_0^{t\wedge s} \mathrm{Tr}[\Phi_2(r)\Phi_1^*(r)]dr\right), \ \ \forall \ t,s \in [0,T].$$

Simplifying these expressions for the multiplication operators described above returns equations (SM.26a) and (SM.26b) above.

3 From Branching Processes to SDE

3.1 Calculating the Deterministic and Variance Components of Offspring Number

Here we provide calculations used to obtain results on the mean and variance of offspring number and birth/death rates mentioned in §2.3 of the main text. In particular, since we set the expected reproductive output at the k-th stage of rescaling to $w^{(k)}(X,x) = w^{1/k}(X,x)$, the associated variance in reproductive output will be

$$v_{(k)}^2(X,x) = \sum_{j=0}^{\infty}(j - w^{1/k}(X,x))^2 f_j^{(k)}(X,x), \tag{SM.32}$$

where $f_j^{(k)}(X,x)$ is the probability that an individual having trait value x in a population with state X gives birth to j offspring in the k-th stage of rescaling. In particular, under the Poisson model of demographic stochasticity we have $v_{(k)}^2(X,x) = w^{1/k}(X,x)$. Under the birth-death model we can use the relationship $w^{1/k}(X,x) = 2f_2^{(k)}(X,x) = 2(1 - f_0^{(k)}(X,x))$ to find

$$v_{(k)}^2(X,x) = 2w^{1/k}(X,x) - w^{2/k}(X,x). \tag{SM.33}$$

Following our assumptions in §2.3 of the main text, a result of Méléard and Roelly (1992, 1993) implies we can compute the reproductive variance in the diffusion-limit as

$$V(\mathscr{X},x) = \rho(\mathscr{X},x)\lim_{k\to\infty}\left[v_{(k)}^2(X^{(k)},x) + (w^{1/k}(X^{(k)},x) - 1)^2\right]. \tag{SM.34}$$

In particular, both the Poisson and birth-death models of demographic stochasticity imply $V(\mathscr{X},x) = \rho(\mathscr{X},x)$.

This is not always the case. In particular, assuming binomially distributed number of offspring with probability parameter $u(X,x)$, similar calculations show $V(\mathscr{X},x) = \rho(\mathscr{X},x)(1 - u(\mathscr{X},x))$. Hence, alternative models of demographic stochasticity may lead to novel forms of noise-induced selection. However, for the sake of simplicity we ignore this in the main text.

Aside: To compute the results of this section, it is crucial to note that in Méléard and Roelly (1992, 1993) $\mathbb{N} = \{0,1,\dots\}$ as opposed to the more common convention $\mathbb{N} = \{1,2,\dots\}$.

3.2 Simulating the Rescaled Process

In this section we describe an approach to simulating rescaled individual-based models. We follow the basic model structure and notation outlined in §2.3 of the main text. We assume offspring number at branching events are Poisson distributed and a constant branching rate across all trait values so that selection enters only through the expected number of offspring $w(X_t,x)$. We further assume selection is stabilizing around zero with strength a and individuals compete independently of trait value so that growth is logistic. Fitness for the unscaled process can then be written as

$$w(X_t,x) = \exp\left(R - \frac{a}{2}x^2 - cn(t)\right), \tag{SM.35}$$

where R is the innate growth rate, c is the strength of competition and $n(t) = X_t(\mathbb{R})$ is the number of individuals at time t.

For the rescaled process, we choose an initial population mass of N_0 (which can be any non-negative real number) and a rescaling parameter k (which can be any positive integer). The rescaled population process is denoted $X_t^{(k)}$ and is defined so that $X_0^{(k)} = \frac{N_0}{kn(0)}\sum_{i=1}^{kn(0)}\delta_{x_i}$. The additional trait values $x_{n(0)+1},\dots,x_{kn(0)}$ can be appended by any preferred method so long as the initial mass $X_0^{(k)}$ admits a density as $k \to \infty$. Here we assume all initial trait values are drawn iid from a normal distribution with mean zero and variance $\sqrt{\mu/a}$. We set $n^{(k)}(t) = \frac{kn(0)}{N_0}X_t^{(k)}(\mathbb{R})$ so that $n^{(k)}(t)$ tracks the number of discrete individuals at time t of the rescaled process $X_t^{(k)}$. In particular, $n^{(k)}(0) = kn(0)$.

Following our approach to rescaling fitness outlined in the main text, we have

$$w^{(k)}(X_t^{(k)}, x) = \exp\left[\frac{1}{k}\left(R - \frac{a}{2}x^2 - cN_0\frac{n^{(k)}(t)}{kn(0)}\right)\right]. \tag{SM.36}$$

Since each individual is assigned an exponentially distributed lifetime, to iterate the rescaled process we search for the individual with the shortest lifetime. We then draw offspring to replace the chosen individual and assign each of them trait values, fitnesses and exponentially distributed lifetimes. We repeat these steps until some stopping criterion is satisfied. Here we chose to stop the process after 5 units of time. Results of these simulations for $k = 1, 5, 10$ are shown in Figure S1. Code for performing these simulations can be found at:

`https://github.com/bobweek/white.noise.community.ecology`

3.3 Calculating SDE From SPDE

Assuming a growth rate $m(\nu, x)$ such that solutions to SAGA are well defined, we can calculate the total mass process $N(t)$ using the weak solution of SAGA with $f(x) \equiv 1$ (Walsh, 1986; Etheridge, 2000; Evans, 2010). This implies

$$\begin{aligned} N(t) - N(0) &= \int_0^t \int_{\mathbb{R}} \nu(x,s)\left(m(\nu,x)\cdot 1 + \frac{\mu}{2}\frac{\partial^2}{\partial x^2}1\right) + 1\sqrt{V(\nu,x)\nu(x,s)}\dot{W}(x,s)dsdx \\ &= \int_0^t \bar{m}(s)N(s)dt + \int_0^t \sqrt{N(s)}d\hat{\mathbf{W}}_s\left(\sqrt{V(\nu,x)\nu(x,s)}\right), \end{aligned} \tag{SM.37}$$

where the population growth rate is calculated as

$$\bar{m}(t) = \frac{1}{N(t)}\int_{\mathbb{R}} m(\nu,x)\nu(x,t)dx, \tag{SM.38}$$

and $\hat{\mathbf{W}}_s\left(\sqrt{V(\nu,x)\nu(x,s)}\right)$ is a standard Brownian motion (see SM §2.1) given by

$$\int_0^t d\hat{\mathbf{W}}_s\left(\sqrt{V(\nu,x)\nu(x,s)}\right) = \int_0^t \int_{\mathbb{R}} \frac{\sqrt{V(\nu,x)\nu(x,s)}}{\sqrt{\int_{\mathbb{R}} V(\nu,y)\nu(y,s)dy}}\dot{W}(x,s)dxds. \tag{SM.39}$$

Setting $W_N(t) = \hat{\mathbf{W}}_t(\sqrt{V(\nu,x)\nu(x,t)})$ and $\bar{V}(t) = \frac{1}{N(t)}\int_{\mathbb{R}} V(\nu,x)\nu(x,t)dx$, we can use traditional stochastic differential notation to write

$$dN = \bar{m}Ndt + \sqrt{\bar{V}N}dW_N. \tag{SM.40}$$

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. But, if we can show $\int_{\mathbb{R}}(|x| + x^2 + x^4)\nu(x,t)dx < +\infty$ for all $t > 0$ given this condition is satisfied by $\nu(x,0)$, then we can apply the weak solution of SAGA 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

$$\tilde{x}(t) = \tilde{x}(0) + \int_0^t \int_{\mathbb{R}} \left[\nu(x,s)m(\nu,x)x + x\sqrt{V(\nu,x)\nu(x,s)}\dot{W}(x,s)\right]dxds. \tag{SM.41}$$

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

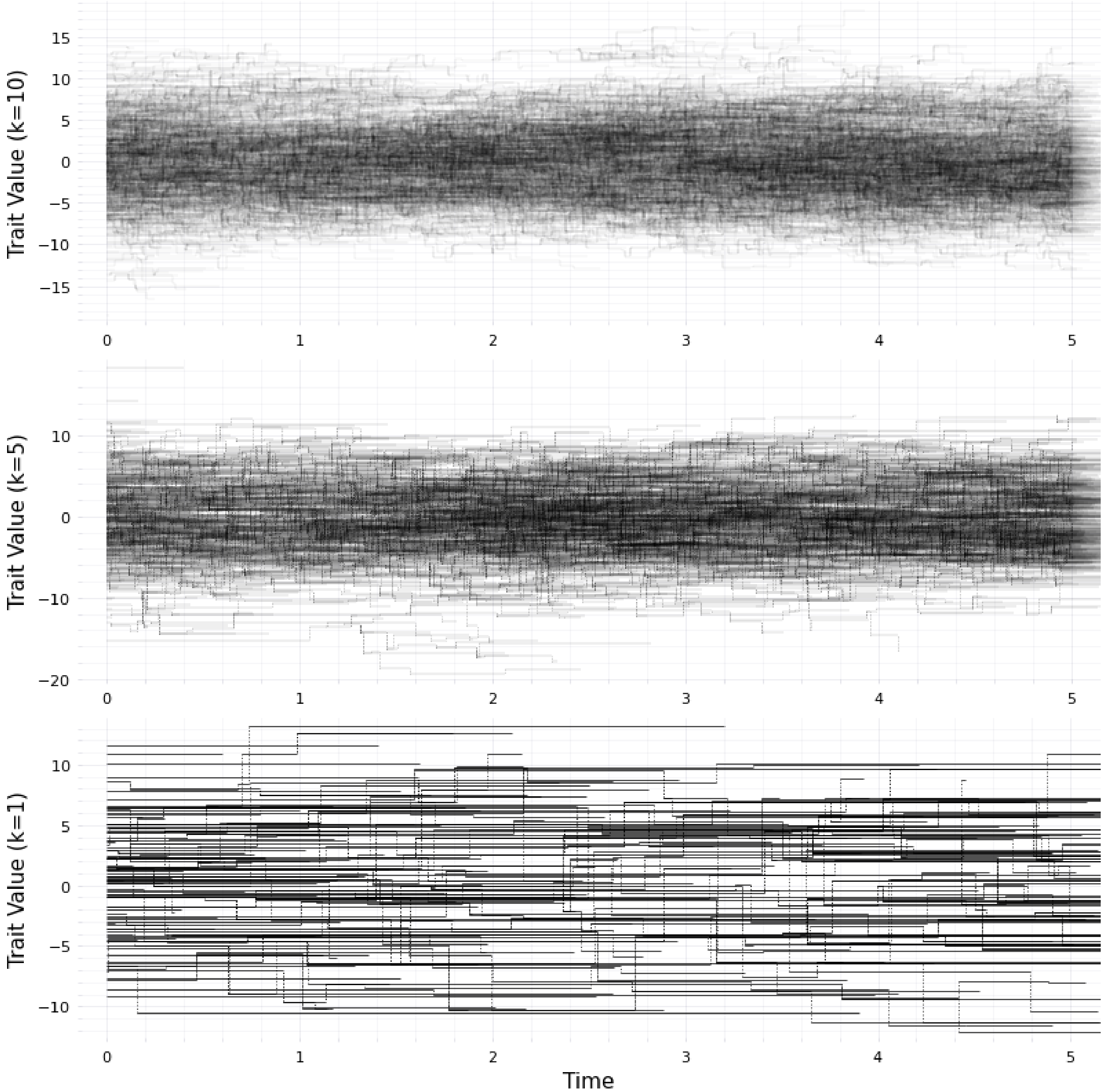


Figure S1: Rescaled sample paths of an individual-based model that includes stabilizing selection and logistic growth. Vertical locations horizontal lines indicate individual trait values and lengths of these lines indicate lifetimes. Dotted vertical lines connect parents to offspring. The bottom plot displays a sample path without scaling ($k = 1$), the middle plot shows a sample path rescaled by $k = 5$ and the top plot shows a sample path rescaled by $k = 10$.

$$\tilde{\sigma}^2(t) = \tilde{\sigma}^2(0) + \int_0^t \int_{\mathbb{R}} \left[\nu(x,s) \left(m(\nu,x)x^2 + \mu \right) + x^2 \sqrt{V(\nu,x)\nu(x,s)} \dot{W}(x,s) \right] dxds. \tag{SM.42}$$

Since $\bar{x}(t) = \tilde{x}(t)/N(t)$ and $\sigma^2(t) = \tilde{\sigma}^2(t)/N(t) - \bar{x}^2(t)$, we can use Itô's lemma to derive SDE for $\bar{x}(t)$ and $\sigma^2(t)$, which we perform in SM §3.4. We make no attempt in finding sufficient conditions to ensure $\int_{\mathbb{R}} (|x| + x^2 + x^4)\nu(x,t)dx < +\infty$ and hence make no general assertions about the existence or uniqueness of $\bar{x}(t)$ or $\sigma^2(t)$. Regardless, when we assume $\nu(x,t)$ can be approximated by a Gaussian curve in x for all $t \geq 0$, this implies $\int_{\mathbb{R}} |x|^n \nu(x,t)dx < +\infty$ for all $n \in \{1,2,\dots\}$ and for all $t \geq 0$. Hence, although not necessary, a Gaussian phenotypic distribution is sufficient to guarantee the existence of $\bar{x}(t)$ and $\sigma^2(t)$ for all $t > 0$ and thus will make for an important initial approximation. However, we do not assume Gaussian phenotypic distributions in the following section.

3.4 Calculation of SDE for $\bar{x}$ and σ^2

Here we calculate the stochastic dynamics of $\bar{x}$ and σ^2 under SAGA using the white noise heuristics developed above. In particular, we combine weak solutions of SPDE, an extension of Itô's multiplication table summarized in Table S1 and Itô's quotient rule. These calculations require the abundance density $\nu(x,t)$ to have finite first, second and fourth phenotypic moments. Hence, we assume

$$\int_{\mathbb{R}} \nu(x,t)(|x| + x^2 + x^4)dx < +\infty. \tag{SM.43}$$

In the following two subsections we use Itô's quotient rule to derive expressions for the evolution of $\bar{x} = \tilde{x}/N$ and $\sigma^2 = \tilde{\sigma}^2 - \bar{x}^2$. Following these two subsections we investigate stochastic dependencies between the processes N, $\bar{x}$ and σ^2.

3.4.1 Calculation for Trait Mean

We make use of the notation

$$\begin{cases} \|N\|_2 = \sqrt{\int_{\mathbb{R}} V(\nu,x)\nu(x,t)dx} = \sqrt{\bar{V}N}, \\ \|\tilde{x}\|_2 = \sqrt{\int_{\mathbb{R}} x^2 V(\nu,x)\nu(x,t)dx} = \sqrt{\overline{x^2V}N}, \\ \langle \tilde{x}, N \rangle = \int_{\mathbb{R}} xV(\nu,x)\nu(x,t)dx = \overline{xV}N. \end{cases} \tag{SM.44}$$

Rewriting formula (SM.41) as an SDE provides

$$d\tilde{x} = \overline{xm}Ndt + \|\tilde{x}\|_2 d\tilde{W}_{\tilde{x}}, \tag{SM.45}$$

where $\overline{xm}(t) = \frac{1}{N(t)} \int_{\mathbb{R}} xm(\nu,x)\nu(x,t)dx$ and

$$d\tilde{W}_{\tilde{x}}(t) = d\hat{\mathbf{W}}_t(\sqrt{x^2V(\nu,x)\nu(x,t)}) = \frac{1}{\|\tilde{x}\|_2} \int_{\mathbb{R}} x\sqrt{V(\nu,x)\nu(x,t)}\dot{W}(x,t)dxdt. \tag{SM.46}$$

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

$$d\bar{x} = d\left(\frac{\tilde{x}}{N}\right) = \frac{\tilde{x}}{N}\left(\frac{d\tilde{x}}{\tilde{x}} - \frac{dN}{N} - \frac{d\tilde{x}}{\tilde{x}}\frac{dN}{N} + \left(\frac{dN}{N}\right)^2\right) = \frac{d\tilde{x}}{N} - \bar{x}\frac{dN}{N} - \frac{d\tilde{x}}{N}\frac{dN}{N} + \bar{x}\left(\frac{dN}{N}\right)^2. \tag{SM.47}$$

From Table S1 we have $d\tilde{x}dN = \langle \tilde{x}, N \rangle$ and $dN^2 = \|N\|_2^2$. Hence,

$$\begin{aligned} d\bar{x} &= \overline{xm}dt + \frac{\|\tilde{x}\|_2}{N}d\tilde{W}_{\tilde{x}} - \bar{x}\left(\bar{m}dt + \sqrt{\frac{\bar{V}}{N}}dW_N\right) - \frac{\langle \tilde{x}, N\rangle}{N^2}dt + \bar{x}\frac{\|N\|_2^2}{N^2}dt \\ &= \mathrm{Cov}_t(m,x)dt - \frac{1}{N}\left(\overline{xV} - \bar{x}\bar{V}\right)dt + \left(\frac{\|\tilde{x}\|_2}{N}dW_{\tilde{x}} - \bar{x}\sqrt{\frac{\bar{V}}{N}}dW_N\right). \end{aligned} \tag{SM.48}$$

Setting $\mathrm{Cov}_t(V,x) = \frac{1}{N(t)}\int_{\mathbb{R}}(x-\bar{x}(t))(V(\nu,x)-\bar{V})\nu(x,t)dx = \overline{xV} - \bar{x}\bar{V}$, we can then write

$$d\bar{x} = \Big(\mathrm{Cov}_t(m,x) - \frac{1}{N}\mathrm{Cov}_t(V,x)\Big)dt + \left(\frac{\|\tilde{x}\|_2}{N}dW_{\tilde{x}} - \bar{x}\sqrt{\frac{\bar{V}}{N}}dW_N\right). \tag{SM.49}$$

Note that

$$\begin{aligned} &\frac{\|\tilde{x}\|_2}{N}d\tilde{W}_{\tilde{x}} - \bar{x}\sqrt{\frac{\bar{V}}{N}}dW_N \\ &\qquad = \frac{1}{N}\int_{\mathbb{R}} x\sqrt{V(\nu,x)\nu(x,t)}\dot{W}(x,t)dx - \frac{\bar{x}}{N}\int_{\mathbb{R}}\sqrt{V(\nu,x)\nu(x,t)}\dot{W}(x,t)dx \\ &\qquad = \int_{\mathbb{R}}\frac{(x-\bar{x})}{N}\sqrt{V(\nu,x)\nu(x,t)}\dot{W}(x,t)dx. \end{aligned} \tag{SM.50}$$

Furthermore, by setting $\mathrm{Cov}_t\big(V,(x-\bar{x})^2\big) = \int_{\mathbb{R}}\left((x-\bar{x})^2 - \sigma^2\right)\left(V(\nu,x)-\bar{V}\right)p(x,t)dx$ we can write

$$\begin{aligned} &\mathbb{V}\left(\int_{\mathbb{R}}\frac{(x-\bar{x})}{N}\sqrt{V(\nu,x)\nu(x,t)}\dot{W}(x,t)dx\right) \\ &\qquad = \frac{1}{N}\int_{\mathbb{R}}(x-\bar{x})^2V(\nu,x)p(x,t)dx \\ &\qquad = \frac{1}{N}\Big(\mathrm{Cov}_t\big(V,(x-\bar{x})^2\big) + \bar{V}\sigma^2\Big) \end{aligned} \tag{SM.51}$$

Hence, by setting

$$dW_{\bar{x}} = \frac{\int_{\mathbb{R}}\frac{(x-\bar{x})}{N}\sqrt{V(\nu,x)\nu(x,t)}\dot{W}(x,t)dx}{\sqrt{\frac{1}{N}(\mathrm{Cov}_t\big(V,(x-\bar{x})^2\big) + \bar{V}\sigma^2)}} \tag{SM.52}$$

we can finally write

$$d\bar{x} = \Big(\mathrm{Cov}_t(m,x) - \frac{1}{N}\mathrm{Cov}_t(V,x)\Big)dt + \sqrt{\frac{1}{N}\left(\mathrm{Cov}_t\big(V,(x-\bar{x})^2\big) + \bar{V}\sigma^2\right)}dW_{\bar{x}}. \tag{SM.53}$$

3.4.2 Calculation for Trait Variance

We make use of the notation

$$\begin{cases} \|\tilde{\sigma}^2\|_2 = \sqrt{\int_{\mathbb{R}} x^4V(\nu,x)\nu(x,t)dx} \\ \langle\tilde{\sigma}^2, N\rangle = \int_{\mathbb{R}} x^2V(\nu,x)\nu(x,t)dx = \overline{x^2V}N. \end{cases} \tag{SM.54}$$

Applying formula (SM.42) provides

$$d\tilde{\tilde{\sigma}}^2 = \left(\overline{x^2 m}N + \mu N\right) dt + \|\tilde{\tilde{\sigma}}^2\|_2 d\tilde{\tilde{W}}_{\tilde{\tilde{\sigma}}^2} \tag{SM.55}$$

where

$$d\tilde{\tilde{W}}_{\tilde{\tilde{\sigma}}^2}(t) = d\hat{\mathbf{W}}_t\big(\sqrt{x^4 V(\nu, x)\nu(x,t)}\big) = \frac{1}{\|\tilde{\tilde{\sigma}}^2\|_2}\int_{\mathbb{R}} \sqrt{x^4 V(\nu, x)\nu(x,t)}\,\dot{W}(x,t)dx. \tag{SM.56}$$

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

$$d\overline{x^2} = d\left(\frac{\tilde{\tilde{\sigma}}^2}{N}\right) = \frac{\tilde{\tilde{\sigma}}^2}{N}\left(\frac{d\tilde{\tilde{\sigma}}^2}{\tilde{\tilde{\sigma}}^2} - \frac{dN}{N} - \frac{d\tilde{\tilde{\sigma}}^2}{\tilde{\tilde{\sigma}}^2}\frac{dN}{N} + \left(\frac{dN}{N}\right)^2\right) = \frac{d\tilde{\tilde{\sigma}}^2}{N} - \overline{x^2}\frac{dN}{N} - \frac{d\tilde{\tilde{\sigma}}^2}{N}\frac{dN}{N} + \overline{x^2}\left(\frac{dN}{N}\right)^2. \tag{SM.57}$$

Table S1 implies $d\tilde{\tilde{W}}_{\tilde{\tilde{\sigma}}^2} dW_N = \langle \tilde{\tilde{\sigma}}^2, N\rangle$ and hence

$$\begin{aligned}
d\overline{x^2} &= \left(\mu + \overline{x^2 m}\right) dt + \frac{\|\tilde{\tilde{\sigma}}^2\|_2}{N} d\tilde{\tilde{W}}_{\tilde{\tilde{\sigma}}^2} - \overline{x^2}\left(\bar{m}dt + \sqrt{\frac{\bar{V}}{N}}dW_N\right) - \frac{\langle \tilde{\tilde{\sigma}}^2, N\rangle}{N^2}dt + \overline{x^2}\frac{\|N\|_2^2}{N^2}dt \\
&= \left(\mu + \overline{x^2 m} - \overline{x^2}\bar{m}\right) dt + \frac{\|\tilde{\tilde{\sigma}}^2\|_2}{N} d\tilde{\tilde{W}}_{\tilde{\tilde{\sigma}}^2} - \overline{x^2}\sqrt{\frac{\bar{V}}{N}}dW_N - \frac{\overline{x^2 V}}{N}dt + \frac{\overline{x^2}\bar{V}}{N}dt \\
&= \left(\mu + \mathrm{Cov}_t(m, x^2) - \frac{1}{N}\mathrm{Cov}_t(V, x^2)\right) dt + \frac{\|\tilde{\tilde{\sigma}}^2\|_2}{N} d\tilde{\tilde{W}}_{\tilde{\tilde{\sigma}}^2} - \overline{x^2}\sqrt{\frac{\bar{V}}{N}}dW_N,
\end{aligned} \tag{SM.58}$$

where $\mathrm{Cov}_t(V, x^2) = \int_{\mathbb{R}}\big(x^2 - \overline{x^2}\big)\big(V(\nu, x) - \bar{V}\big)\nu(x,t)dx = \overline{x^2 V} - \overline{x^2}\bar{V}$. Setting $F(y,z) = y - z^2$, we use Itô's formula on $\sigma^2 = F(\overline{x^2}, \bar{x}) = \overline{x^2} - \bar{x}^2$ to obtain

$$
\begin{aligned}
d\sigma^2 &= d\overline{x^2} - 2\bar{x}d\bar{x} - (d\bar{x})^2 \\
&= \left(\mu + \mathrm{Cov}_t\big(m, x^2\big) - \frac{1}{N}\mathrm{Cov}_t\big(V, x^2\big)\right) dt + \frac{\|\tilde{\tilde{\sigma}}^2\|_2}{N} d\tilde{\tilde{W}}_{\tilde{\tilde{\sigma}}^2} - \overline{x^2}\sqrt{\frac{\bar{V}}{N}} dW_N \\
&\quad -2\bar{x}\left[\left(\mathrm{Cov}_t(m, x) - \frac{1}{N}\mathrm{Cov}_t(V, x)\right) dt + \sqrt{\frac{1}{N}\big(\mathrm{Cov}_t(V, (x-\bar{x})^2) + \bar{V}\sigma^2\big)} dW_{\bar{x}}\right] \\
&\quad - \left[\left(\mathrm{Cov}_t(m, x) - \frac{1}{N}\mathrm{Cov}_t(V, x)\right) dt + \sqrt{\frac{1}{N}\big(\mathrm{Cov}_t(V, (x-\bar{x})^2) + \bar{V}\sigma^2\big)} dW_{\bar{x}}\right]^2 \\
&= \left(\mu + \mathrm{Cov}_t\big(m, x^2 - 2\bar{x}x\big) - \frac{1}{N}\mathrm{Cov}_t\big(V, x^2 - 2\bar{x}x\big) - \frac{1}{N}\mathrm{Cov}_t\big(V, (x-\bar{x})^2\big) - \frac{1}{N}\bar{V}\sigma^2\right) dt \\
&\quad + \frac{\|\tilde{\tilde{\sigma}}^2\|_2}{N} d\tilde{\tilde{W}}_{\tilde{\tilde{\sigma}}^2} - \overline{x^2}\sqrt{\frac{\bar{V}}{N}} dW_N - 2\bar{x}\sqrt{\frac{1}{N}\big(\mathrm{Cov}_t(V, (x-\bar{x})^2) + \bar{V}\sigma^2\big)} dW_{\bar{x}} \\
&= \left(\mu + \mathrm{Cov}_t\big(m, (x-\bar{x})^2\big) - \frac{1}{N}\left(2\mathrm{Cov}_t\big(V, (x-\bar{x})^2\big) + \bar{V}\sigma^2\right)\right) dt \\
&\quad + \frac{\|\tilde{\tilde{\sigma}}^2\|_2}{N} d\tilde{\tilde{W}}_{\tilde{\tilde{\sigma}}^2} - \overline{x^2}\sqrt{\frac{\bar{V}}{N}} dW_N - 2\bar{x}\sqrt{\frac{1}{N}\big(\mathrm{Cov}_t(V, (x-\bar{x})^2) + \bar{V}\sigma^2\big)} dW_{\bar{x}}. \qquad \text{(SM.59)}
\end{aligned}
$$

Since

$$
\begin{aligned}
&\frac{\|\tilde{\tilde{\sigma}}^2\|_2}{N} d\tilde{\tilde{W}}_{\tilde{\tilde{\sigma}}^2} - \overline{x^2}\sqrt{\frac{\bar{V}}{N}} dW_N - 2\bar{x}\sqrt{\frac{1}{N}\big(\mathrm{Cov}_t(V, (x-\bar{x})^2) + \bar{V}\sigma^2\big)} dW_{\bar{x}} \\
&\qquad = \frac{1}{N}\int_{\mathbb{R}} \left(x^2 - \overline{x^2} - 2\bar{x}(x-\bar{x})\right) \sqrt{V(\nu, x)\nu(x,t)}\dot{W}(x,t)dx \\
&\qquad = \frac{1}{N}\int_{\mathbb{R}} \left((x-\bar{x})^2 - \sigma^2\right) \sqrt{V(\nu, x)\nu(x,t)}\dot{W}(x,t)dx \qquad \text{(SM.60)}
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{V}\left(\frac{1}{N}\int_{\mathbb{R}} \left((x-\bar{x})^2 - \sigma^2\right) \sqrt{V(\nu, x)\nu(x,t)}\dot{W}(x,t)dx\right) \\
&\qquad = \frac{1}{N}\int_{\mathbb{R}} \left((x-\bar{x})^2 - \sigma^2\right)^2 V(\nu, x)p(x,t)dx \\
&\qquad = \frac{1}{N}\left(\mathrm{Cov}_t\Big(V, \big[(x-\bar{x})^2 - \sigma^2\big]^2\Big) + \bar{V}\left[\overline{(x-\bar{x})^4} - \sigma^4\right]\right) \qquad \text{(SM.61)}
\end{aligned}
$$

we set

$$dW_{\sigma^2} = \frac{\frac{1}{N}\int_{\mathbb{R}} \left((x-\bar{x})^2-\sigma^2\right)\sqrt{V(\nu,x)\nu(x,t)}\,\dot{W}(x,t)dx}{\sqrt{\frac{1}{N}\left(\mathrm{Cov}_t\Big(V,\left[(x-\bar{x})^2-\sigma^2\right]^2\Big)+\bar{V}\left[\overline{(x-\bar{x})^4}-\sigma^4\right]\right)}}. \tag{SM.62}$$

Hence, we can finally write

$$d\sigma^2 = \left(\mu + \mathrm{Cov}_t\big(m,(x-\bar{x})^2\big) - \frac{1}{N}\left(2\mathrm{Cov}_t\big(V,(x-\bar{x})^2\big)+\bar{V}\sigma^2\right)\right)dt$$

$$+\frac{1}{N}\left(\mathrm{Cov}_t\Big(V,\left[(x-\bar{x})^2-\sigma^2\right]^2\Big)+\bar{V}\left[\overline{(x-\bar{x})^4}-\sigma^4\right]\right)dW_{\sigma^2}. \tag{SM.63}$$

3.5 Stochastic Dependencies Between N, $\bar{x}$ and σ^2

Table S1 implies

$$dW_N dW_{\bar{x}} = \frac{\frac{1}{N}\int_{\mathbb{R}}(x-\bar{x})V(\nu,x)\nu(x,t)dx}{\sqrt{\bar{V}\mathrm{Cov}_t\big(V,(x-\bar{x})^2\big)+\bar{V}^2\sigma^2}}dt = \left(\bar{V}\overline{V(x-\bar{x})^2}\right)^{-1/2}\mathrm{Cov}_t(V,x)dt, \tag{SM.64a}$$

$$dW_N dW_{\sigma^2} = \frac{\frac{1}{N}\int_{\mathbb{R}}\left((x-\bar{x})^2-\sigma^2\right)V(\nu,x)\nu(x,t)dx}{\sqrt{\bar{V}\mathrm{Cov}_t\big(V,[(x-\bar{x})^2-\sigma^2]^2\big)+\bar{V}^2\left[\overline{(x-\bar{x})^4}-\sigma^4\right]}}dt$$

$$= \left(\bar{V}\overline{V\left[(x-\bar{x})^2-\sigma^4\right]^2}\right)^{-1/2}\mathrm{Cov}_t\big(V,(x-\bar{x})^2\big)dt, \tag{SM.64b}$$

$$dW_{\bar{x}} dW_{\sigma^2} = \frac{\int_{\mathbb{R}}(x-\bar{x})\left[(x-\bar{x})^2-\sigma^2\right]V(\nu,x)p(x,t)dx}{\sqrt{\left\{\mathrm{Cov}_t\big(V,(x-\bar{x})^2\big)+\bar{V}\sigma^2\right\}\left\{\mathrm{Cov}_t\big(V,[(x-\bar{x})^2-\sigma^2]^2\big)+\bar{V}\left[\overline{(x-\bar{x})^4}-\sigma^4\right]\right\}}}dt$$

$$= \left\{\left(\overline{V\left[(x-\bar{x})^2-\sigma^4\right]^2}\right)\left(\overline{V(x-\bar{x})^2}\right)\right\}^{-1/2}\left[\bar{V}\overline{(x-\bar{x})^3}+\mathrm{Cov}_t\Big(V,(x-\bar{x})^2-\sigma^2 x\Big)\right]dt. \tag{SM.64c}$$

where

$$\overline{V(x-\bar{x})^2} = \int_{\mathbb{R}}(x-\bar{x})^2 V(\nu,x)p(x,t)dx, \tag{SM.65a}$$

$$\overline{V[(x-\bar{x})^4-\sigma^4]^2} = \int_{\mathbb{R}}[(x-\bar{x})^2-\sigma^4]^2 V(\nu,x)p(x,t)dx. \tag{SM.65b}$$

Hence, stochastic fluctuations in the evolution of abundance N are in general correlated with those in mean trait $\bar{x}$ and trait variance σ^2. However, for a constant reproductive variance $V(\nu,x)\equiv V>0$, the stochastic fluctuations in N become uncorrelated from both $\bar{x}$ and σ^2, but the stochastic fluctuations in the evolutions of $\bar{x}$ and σ^2 may be correlated. This is not the case when traits are normally distributed as equation (SM.64c) would then imply $dW_{\bar{x}}dW_{\sigma^2} = 0$.

3.6 Replacing Covariances with Fitness Gradients

Here we show how to rewrite SDE for $\bar{x}$ and σ^2 in terms of fitness gradients instead of fitness covariances by assuming Gaussian populations In particular, we can use the Guassian population assumption to rewrite the expressions

$$\mathrm{Cov}_t(m,x),\ \mathrm{Cov}_t(m,(x-\bar{x})^2),\ \mathrm{Cov}_t(V,x),\ \mathrm{Cov}_t(V,(x-\bar{x})^2), \tag{SM.66}$$

and

$$\begin{aligned}&\int_{\mathbb{R}} \left((x-\bar{x})^2-\sigma^2\right)^2 V(\nu,x)p(x,t)dx \\ &\qquad = \overline{\left((x-\bar{x})^2-\sigma^2\right)^2 V} = \mathrm{Cov}_t\left(V,\left((x-\bar{x})^2-\sigma^2\right)^2\right) + \bar{V}\left[\overline{(x-\bar{x})^4}-\sigma^4\right]\end{aligned} \tag{SM.67}$$

in terms of parital derivatives of either $m(\nu,x)$ or $V(\nu,x)$ with respect to either $\bar{x}$ or σ^2. For brevity we write $\partial_{\bar{x}} = \frac{\partial}{\partial \bar{x}}$ and $\partial_{\sigma^2} = \frac{\partial}{\partial \sigma^2}$. Then, for any function $f(x,\bar{x})$ that allows $\partial_{\bar{x}} \int_{\mathbb{R}} f(x,\bar{x})p(x,t)dx = \int_{\mathbb{R}} \partial_{\bar{x}}[f(x,\bar{x})p(x,t)]dx$, we have

$$\begin{aligned}\partial_{\bar{x}}\bar{f} &= \partial_{\bar{x}} \int_{\mathbb{R}} f(x,\bar{x})p(x,t)dx \\ &= \int_{\mathbb{R}} p(x,t)\partial_{\bar{x}}f(x,\bar{x}) + \frac{(x-\bar{x})}{\sigma^2}f(x,\bar{x})p(x,t)dx = \overline{\partial_{\bar{x}}f} + \frac{1}{\sigma^2}\mathrm{Cov}_t(f,x).\end{aligned} \tag{SM.68}$$

Hence, $\mathrm{Cov}_t(f,x) = \sigma^2\left(\partial_{\bar{x}}\bar{f} - \overline{\partial_{\bar{x}}f}\right)$. Given $m(\nu,x)$ and $V(\nu,x)$ allow swapping the order of differentiation with respect to $\bar{x}$ and integration with respect to x, we can apply this result to rewrite $\mathrm{Cov}_t(m,x)$ and $\mathrm{Cov}_t(V,x)$.

Suppose that $h(x,\sigma^2)$ satisfies $\partial_{\sigma^2}\bar{h} = \int_{\mathbb{R}} \partial_{\sigma^2}[h(x,\sigma^2)p(x,t)]dx$. Then

$$\begin{aligned}\partial_{\sigma^2}\bar{h} &= \int_{\mathbb{R}} p(x,t)\partial_{\sigma^2}h(x,\sigma^2) + h(x,\sigma^2)\partial_{\sigma^2}p(x,t)dx \\ &= \overline{\partial_{\sigma^2}h} + \int_{\mathbb{R}} \frac{(x-\bar{x})^2-\sigma^2}{2\sigma^4}h(x,\sigma^2)p(x,t)dx = \overline{\partial_{\sigma^2}h} + \frac{1}{2\sigma^4}\mathrm{Cov}_t\left(h,(x-\bar{x})^2\right).\end{aligned} \tag{SM.69}$$

Hence, $\mathrm{Cov}_t\left(h,(x-\bar{x})^2\right) = 2\sigma^4\left(\partial_{\sigma^2}\bar{h} - \overline{\partial_{\sigma^2}h}\right)$. When $m(\nu,x)$ and $V(\nu,x)$ allow the swapping of differentiation with respect to σ^2 and integration with respect to x, we can use this result to rewrite $\mathrm{Cov}_t(m,(x-\bar{x})^2)$ and $\mathrm{Cov}_t(V,(x-\bar{x})^2)$.

Finally, we are left with expression (SM.67). As a first guess, we expand $\partial^2_{\sigma^2}\bar{V}$ in hopes of writing the covariance appearing in (SM.67) in terms of gradients of reproductive variance. We can apply the above results to get

$$
\begin{aligned}
\partial^2_{\sigma^2}\bar{V} &= \partial_{\sigma^2}\int_{\mathbb{R}} p(x,t)\partial_{\sigma^2}V(\nu,x) + V(\nu,x)\partial_{\sigma^2}p(x,t)dx \\
&= \int_{\mathbb{R}} p(x)\partial^2_{\sigma^2}V(\nu,x) + [\partial_{\sigma^2}V(\nu,x)][\partial_{\sigma^2}p(x,t)] + \partial_{\sigma^2}\left[\frac{((x-\bar{x})^2-\sigma^2)}{2\sigma^4}V(\nu,x)p(x,t)\right]dx \\
&= \overline{\partial^2_{\sigma^2}V} + \int_{\mathbb{R}} \frac{((x-\bar{x})^2-\sigma^2)}{2\sigma^4}p(x,t)\partial_{\sigma^2}V(\nu,x) + \left(\frac{1}{2\sigma^4}-\frac{(x-\bar{x})^2}{\sigma^6}\right)V(\nu,x)p(x,t) \\
&\quad + \frac{((x-\bar{x})^2-\sigma^2)}{2\sigma^4}p(x,t)\partial_{\sigma^2}V(\nu,x) + \frac{((x-\bar{x})^2-\sigma^2)}{2\sigma^4}V(\nu,x)\partial_{\sigma^2}p(x,t)dx \\
&= \overline{\partial^2_{\sigma^2}V} + \frac{1}{\sigma^4}\mathrm{Cov}_t\left(\partial_{\sigma^2}V,(x-\bar{x})^2\right) + \frac{\bar{V}}{2\sigma^4} \\
&\quad - \frac{1}{\sigma^6}\left[\mathrm{Cov}_t\left(V,(x-\bar{x})^2\right) + \bar{V}\sigma^2\right] + \left(\frac{(x-\bar{x})^2-\sigma^2}{2\sigma^4}\right)^2 V(\nu,x)p(x,t)dx \\
&= \overline{\partial^2_{\sigma^2}V} + 2\left(\partial_{\sigma^2}\overline{\partial_{\sigma^2}V} - \overline{\partial^2_{\sigma^2}V}\right) - \frac{\bar{V}}{2\sigma^4} - \frac{2}{\sigma^2}\left(\partial_{\sigma^2}\bar{V} - \overline{\partial_{\sigma^2}V}\right) \\
&\quad + \frac{1}{4\sigma^8}\left[\mathrm{Cov}_t\left(V,((x-\bar{x})^2-\sigma^2)^2\right) + \bar{V}\left[\overline{(x-\bar{x})^4}-\sigma^4\right]\right].
\end{aligned}
\tag{SM.70}
$$

Hence, when trait values follow a normal distribution, we can write

$$
\begin{aligned}
&\mathrm{Cov}_t\left(V,((x-\bar{x})^2-\sigma^2)^2\right) + \bar{V}\left[\overline{(x-\bar{x})^4}-\sigma^4\right] \\
&\qquad = 2\sigma^4\left[2\sigma^4\left(\partial^2_{\sigma^2}\bar{V} - 2\partial_{\sigma^2}\overline{\partial_{\sigma^2}V} + \overline{\partial^2_{\sigma^2}V}\right) + 4\sigma^2\left(\partial_{\sigma^2}\bar{V} - \overline{\partial_{\sigma^2}V}\right) + \bar{V}\right].
\end{aligned}
\tag{SM.71}
$$

4 Imperfect Inheritance

In this section we describe our approach to modelling imperfect inheritance, which follows the simplest assumptions of classical quantitative genetics. For further reading on the theory of inheritance and genetic architecture of quantitative traits, we recommend Bürger (2000) for a review of mathematical developments and Walsh and Lynch (2018) as a less technical but comprehensive overview.

4.1 Inheritance

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 (called genotypic value in Bulmer, 1971; Walsh and Lynch, 2018) of an individual is the sum of additive effects of the alleles carried by the individual on its expressed trait. Hence, if the trait is encoded by L loci and the additive effect at locus l is a_l, then $g = \sum_{l=1}^{L} a_l$. The additive genetic variance G is just the variance of breeding values in a population (Bulmer, 1971; Walsh and Lynch, 2018). Following Lande (1975), we assume a mutation at locus l occurs with probability M and replaces the additive effect a_l with $a_l + \kappa_l$ where κ_l is normally distributed with a mean of zero and variance μ/M. Hence, we adopt the Gaussian allelic model of mutation. Next, we implement an infinitesimal approximation by assuming breeding values are determined by an infinite number of loci. Although very general infinitesimal approximations have been provided by Barton et al. (2017), for the sake of simplicity we employ a less technical approach. In particular, we rescale the mutational effects κ_l by $1/\sqrt{L}$ and take the limit $L \to \infty$. Then, denoting g' the breeding value of an offspring produced by a parent with breeding value g and I_l the indicator variable determining whether or not a mutation occurs at locus l, we have

$$g' = g + \lim_{L\to\infty} \frac{1}{\sqrt{L}} \sum_{l=1}^{L} I_l \kappa_l. \tag{SM.72}$$

This limit implies that g' has expected value g and variance μ. Thus, our assumptions yield the Gaussian descendants approximation coined by Turelli (2017). For a detailed treatment of breeding values, additive genetic variances and more general genetic architectures see Walsh and Lynch (2018).

4.2 Development

Our treatment of the relationship between breeding values and expressed traits follows classical quantitative genetic assumptions such as those used by Bulmer (1971) to investigate the effect of selection on genetic variation. In particular, we ignore epistatic interactions so that effects at different loci combine additively. Since our treatment assumes haploid asexuals, there are no contributions of dominance or inbreeding depression to phenotypic variance. We assume expressed traits for given individuals are normally distributed around their breeding values with a fixed variance E. Hence, phenotypic variance decomposes as $\sigma^2 = G + E$. The variance E is referred to as the variance of environmental deviation (Walsh and Lynch, 2018). For a fixed breeding value g, we denote the probability density of a randomly drawn expressed trait x by $\psi(x, g)$ so that

$$\psi(x, g) = \frac{1}{\sqrt{2\pi E}} \exp\left(-\frac{(x-g)^2}{2E}\right). \tag{SM.73}$$

4.3 Selection On Breeding Values

To include the relationship between breeding values and expressed traits in our framework, we write $\Xi_t = \sum_{i=1}^{n(t)} \delta_{g_i}$ as the population process of breeding values and $\Xi_t^{(k)}$ as the k-th rescaling of Ξ_t. We assume Ξ_t and $\Xi_t^{(k)}$ follow the same basic structure outlined in §2.3 of the main text. When the diffusion-limit $\lim_{k\to\infty} \Xi_t^{(k)}$ exists and admits a density, we write $\xi(g, t)$ as the abundance density of breeding values. Then, following the above treatment of development, we now have

$$X_t = \sum_{i=1}^{n(t)} \delta_{x_i} = \sum_{i=1}^{n(t)} \delta_{g_i + z_i}, \tag{SM.74}$$

where the z_i are iid random variables normally distributed around zero with variance E. Hence, $\nu(x, t) = \int_{\mathbb{R}} \xi(g, t)\psi(x, g)dg$ when $\xi(g, t)$ exists. Assuming finite abundance, this implies

$$\int_{\mathbb{R}} \nu(x, t)dx = \int_{\mathbb{R}} \xi(g, t)dg = N(t). \tag{SM.75}$$

We therefore switch our focus from directly modelling the evolution of $\nu(x, t)$ to modelling the evolution of $\xi(g, t)$.

Since selection acts on expressed phenotypes, we use the assumed relationship between breeding values and expressed traits to calculate the fitness of breeding values. Writing $w^*(\Xi_t, g_i)$ and $\rho^*(\Xi_t, g_i)$ respectively as the expected offspring number and branching rate of the i-th individual in the population process Ξ_t at time t, the above model of development implies $w^*(\Xi_t, g_i) = w(X_t, x_i)$ and $\rho^*(\Xi_t, g_i) = \rho(X_t, x_i)$. Thus, given $\xi(g,t)$ exists in the diffusion-limit $\lim_{k\to\infty} \Xi_t^{(k)}$, the growth rate $m^*(\xi, g)$ of breeding value g given abundance density $\xi(g,t)$ can be written as

$$m^*(\xi, g) = \int_{\mathbb{R}} \psi(x,g) \lim_{k\to\infty} k\rho(X_t^{(k)}, x)(w^{1/k}(X_t^{(k)}, x) - 1)dx = \int_{\mathbb{R}} m(\nu, x)\psi(x,g)dx. \tag{SM.76}$$

This is similar to equation (1.4) in chapter 5 of Bürger (2000), except we focus on growth rates instead of individual fitness. Similarly, for the reproductive variance $V^*(\xi, g)$ of breeding value g given abundance density $\xi(g,t)$, we find

$$V^*(\xi, g) = \int_{\mathbb{R}} V(\nu, x)\psi(x,g)dx. \tag{SM.77}$$

With the relationships between $m(\nu, x)$ and $m^*(\xi, g)$ and $V(\nu, x)$ and $V^*(\xi, g)$ established, the evolution of the density of breeding values $\xi(g,t)$ is given by the SPDE

$$\dot{\xi}(g,t) = m^*(\xi, g)\xi(g,t) + \frac{\mu}{2}\frac{\partial^2}{\partial g^2}\xi(g,t) + \sqrt{V^*(\xi,g)\xi(g,t)}\dot{W}(g,t). \tag{SM.78}$$

Recall from above that $\nu(x,t)$ is now defined indirectly by (SM.78) via the relation $\nu(x,t) = \int \xi(g,t)\psi(x,g)dg$. Equation (SM.78) is a stochastic generalization of DAGA, the deterministic PDE (4) from §2.1. However, equation (SM.78) describes the evolution of the distribution of breeding values instead of expressed characters. Regardless, whether modelling expressed characters or breeding values, we refer to SPDE of the form (SM.78) as Stochastic Asexual Gaussian allelic models with Abundance dynamics (abbreviated SAGA). In SM §2.1 we developed some heuristics to perform calculations with respect to the space-time white noise term $\dot{W}$.

4.4 Evolution

Setting $\bar{g}(t) = \frac{1}{N(t)} \int_{\mathbb{R}} g\xi(g,t)dg$ and $G(t) = \frac{1}{N(t)} \int_{\mathbb{R}} (g - \bar{g}(t))^2 \xi(g,t)dg$, the above model of development implies $\bar{x}(t) = \bar{g}(t)$ and $\sigma^2(t) = G(t) + E$. Hence, we can use equations (SM.76) and (SM.77) along with the chain rule from calculus to justify equations (24) of the main text.

5 Motivation for Diffuse Coevolution Model

In this section we provide motivation for our model of diffuse coevolution driven by resource competition. Since most of the calculations needed have been completed in SM §3.4, we focus here on calulcating growth rates of each species as a function of trait values and abundance densities of across all species in the community. We then use this fitness function to calculate selection gradients.

5.1 Individual Fitness

We begin by considering populations of discrete individuals, as is the case with branching processes. In particular, we begin by assuming population size n_i is an integer for each species $i = 1, \dots, S$ before passing to the large population size limit. We assume the competitive effects on fitness for each individual accumulate multiplicatively. The magnitude of these negative effects on the growth rate of species i due to competition with species j increases with the degree of niche-overlap between species i and j. The rate of this increase is mediated by the sensitivity parameter $c_{ij} > 0$, which we refer to as the strength of competition on species i due to species j.

We model niche space using the real line $\mathbb{R}$ and represent locations along this axis with the symbol ζ. We assume individuals of species i sample the niche axis following a probability distribution with density $u_i(\zeta, x)$, x being the average niche location sampled which we call the niche-center. In particular, we assume individuals sample their environment following a normal distribution so that the niche-use curve corresponding to individuals of species i can be written

$$u_i(\zeta, x) = \frac{U_i}{\sqrt{2\pi\lambda_i}} \exp\left(-\frac{(\zeta - x)^2}{2\lambda_i}\right), \tag{SM.79}$$

where U_i represents total niche use (since $U_i = \int u_i(\zeta, x)d\zeta$) and λ_i represents niche breadth (the width of the bell curve u_i). We define the niche-overlap between individuals of species i and j with niche centers x_i and x_j respectively as

$$\mathcal{O}_{ij}(x_i, x_j) = \int_{\mathbb{R}} u_i(\zeta, x_i)u_j(\zeta, x_j)d\zeta = \frac{U_iU_j}{\sqrt{2\pi(\lambda_i + \lambda_j)}} \exp\left(-\frac{(x_i - x_j)^2}{2(\lambda_i + \lambda_j)}\right). \tag{SM.80}$$

Denoting x_{ij} the niche-center of the j-th individual belonging to species i, the process tracking the set of trait values of species i can be expressed as

$$X_i = \sum_{j=1}^{n_i} \delta_{x_{ij}}, \tag{SM.81}$$

where individuals branch at exponentially distributed time intervals as described in §2.3 of the main text. Hence, each X_i and n_i depend on time. However, we suppress that dependency here to simplify notation.

We write $\boldsymbol{X} = (X_1, \dots, X_S)$ to denote the S-tuple of branching processes representing the community of interacting species. We denote by $\mathscr{B}_i(\boldsymbol{X}, x)$ a function that maps $\boldsymbol{X}$ and x to the cumulative effect of all competitive interactions on the fitness of an individual belonging to species i with trait value x. 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

$$\begin{aligned}\mathscr{B}_i(\boldsymbol{X}, x_{ij}) &= \exp\left(c_{ii}\mathcal{O}_{ii}(x_{ij}, x_{ij}) - \sum_{k=1}^{S} \int_{\mathbb{R}} c_{ik}\mathcal{O}_{ik}(x_{ij}, y)X_k(dy)\right) \\ &= \exp\left(-\sum_{l \neq j} c_{ii}\mathcal{O}_{ii}(x_{ij}, x_{il}) - \sum_{k \neq i}\sum_{l=1}^{n_k} c_{ik}\mathcal{O}_{ik}(x_{ij}, x_{kl})\right). \end{aligned} \tag{SM.82}$$

To capture abiotic stabilizing selection we assume resources follow a Gaussian density along the niche axis. We also assume the concentration of resources is proportional to expected reproductive output. Combining these assumptions, we denote by $e_i(\zeta)$ the fitness benefits for individuals sampling at niche location ζ so that

$$e_i(\zeta) = w_{\max,i} \exp\left(-\frac{A_i}{2}(\theta_i - \zeta)^2\right), \tag{SM.83}$$

where $w_{\max,i}$ is the maximum expected reproductive output in the absence of competitive interactions, θ_i is the phenotypic optimum (location along niche axis of most abundant resources) and $A_i > 0$ determines the strength of abiotic stabilizing selection (the sharpness of the resource distribution) for species i. We calculate the effect of mismatch between resource use and resource distribution on the fitness of individuals in species i with niche center x as

$$\mathscr{A}_i(x) = \int_{\mathbb{R}} e_i(\zeta)u_i(\zeta, x)d\zeta = \frac{U_i w_{\max,i}}{\sqrt{1 + A_i\lambda_i}} \exp\left(-\frac{A_i}{1 + A_i\lambda_i}(\theta_i - x)^2\right). \tag{SM.84}$$

Writing $w_i(\boldsymbol{X}, x)$ as the average number of offspring left by an individual of species i with trait value x, the absolute fitness of individual j of species i is

$$\begin{aligned} w_i(\boldsymbol{X}, x_{ij}) &= \mathscr{A}_i(x_{ij})\mathscr{B}_i(\boldsymbol{X}, x_{ij}) \\ &= \frac{U_i w_{\max,i}}{\sqrt{1 + A_i\lambda_i}} \exp\left(-\frac{A_i}{1 + A_i\lambda_i}(\theta_i - x)^2 - \sum_{l\neq j} c_{ii}\mathcal{O}_{ii}(x_{ij}, x_{il}) - \sum_{k\neq i}\sum_{l=1}^{n_k} c_{ik}\mathcal{O}_{ik}(x_{ij}, x_{kl})\right). \end{aligned} \tag{SM.85}$$

5.2 Continuous-Time Growth Rate

Our framework for developing population-level models depends quantifying continuous-time growth rates associated with both particular trait values ($m(\nu, x)$) and the entire population ($\bar{m}$). Hence, transitioning from the individual-based formulation to the population-level model requires the calculation of continuous-time growth rates from the expectation of individual reproductive output, denoted above by $w_i(\boldsymbol{X}, x)$, and variance around this expectation, $v_i^2(\boldsymbol{X}, x)$, and branching rate $\rho_i(\boldsymbol{X}, x)$. For the sake of model simplicity, we make branching rates constant for each species so that noise-induced selection does not appear in the diffusion-limit. We also assume Poisson distributed offspring numbers so that $v_i^2(\boldsymbol{X}, x) = w_i(\boldsymbol{X}, x)$. Hence, calculations involved with transitioning from the individual-based model to the the population-level model will only depend on $w_i(\boldsymbol{X}, x)$. The key step in this transition makes use of equation (17) of the main text. To simplify notation, we set the initial number of individuals in each species of the individual-based model to a constant so that $n_i(0) = n$. We then conduct the rescaling using n instead of k. To make use of equation (17) in the SPDE context we require the diffusion limit of the n-th rescaling of X_i, denoted $X_i^{(n)}$, to converge to a superprocess with density ν_i. Following our description of rescaling in the main text, we consider the initial conditions of X_i and $X_i^{(n)}$, which we denote respectively by Y_i and $Y_i^{(n)}$, and set

$$Y_i^{(n)} = \frac{N_i(0)}{n}\sum_{j=1}^{n} \delta_{x_{ij}(0)}, \tag{SM.86}$$

for some positive value of initial total abundance $N_i(0) > 0$. As n increases we suppose each additional initial trait value $x_{ij}(0)$ is drawn independently from a common distribution with density $\nu_i(x, 0)$. For the remainder of this section we continue to work with the initial condition and suppress dependency on time to simplify notation. We denote by $\boldsymbol{Y}^{(n)} = (Y_1^{(n)}, \dots, Y_S^{(n)})$ the rescaled initial condition of the individual-based community $\boldsymbol{Y} = (Y_1, \dots, Y_S)$, by $\boldsymbol{\nu} = (\nu_1, \dots, \nu_S)$ the S-tuple of abundance densities describing the community of interacting species in the diffusion limit and by $m_i(\boldsymbol{\nu}, x)$ the growth rate associated with trait value x for species i in the diffusion limit. Then, following equation (17) of the main text, we have

$$m_i(\boldsymbol{\nu}, x) = \lim_{n\to\infty} n\left(w_i^{1/n}\left(\boldsymbol{Y}^{(n)}, x\right) - 1\right). \tag{SM.87}$$

To perform this calculation with respect to our formulation of individual fitness above, we note that large n gives the approximation

$$w_i^{1/n}(\boldsymbol{Y}^{(n)}, x) \approx \mathscr{A}_i(x)^{1/n}\left(1 + \frac{c_{ii}}{n^2}\mathcal{O}_{ii}(x, x) - \frac{1}{n}\sum_{k=1}^{S}\int_{\mathbb{R}} c_{ik}\mathcal{O}_{ik}(x, y)Y_k^{(n)}(dy)\right). \tag{SM.88}$$

Hence, the growth rate associated with trait value x in species i is

$$
\begin{aligned}
m_i(\boldsymbol{\nu}, x) &= \lim_{n\to\infty} n\left(w_i^{1/n}(\boldsymbol{Y}^{(n)}, x) - 1\right) \\
&= \lim_{n\to\infty} n\left(\mathscr{A}_i(x)^{1/n} - 1\right) + \mathscr{A}_i(x)^{1/n}\left(\frac{c_{ii}}{n}\mathcal{O}_{ii}(x,x) - \sum_{k=1}^{S}\int_{\mathbb{R}} c_{ik}\mathcal{O}_{ik}(x,y)Y_k^{(n)}(dy)\right) \\
&= \ln \mathscr{A}_i(x) - \left(\sum_{k=1}^{S}\int_{\mathbb{R}} c_{ik}\mathcal{O}_{ik}(x,y)\nu_k(y)dy\right). \qquad \text{(SM.89)}
\end{aligned}
$$

This calculation did not require the assumption of normally distributed trait values. However, we make this assumption in the next section on our route towards developing explicit expressions for the evolution of total abundances $N_1, \ldots, N_S$, mean traits $\bar{x}_1, \ldots, \bar{x}_S$ and additive genetic variances $G_1, \ldots, G_S$.

5.3 Computing Fitness Gradients

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

$$\overline{\mathcal{O}}_{ij}(x,t) = \frac{\int_{\mathbb{R}} \mathcal{O}_{ij}(x,y)\nu_j(y,t)dy}{\int_{\mathbb{R}} \nu_j(y,t)dy} = \frac{1}{N_j(t)}\int_{\mathbb{R}} \mathcal{O}_{ij}(x,y)\nu_j(y,t)dy. \qquad \text{(SM.90)}$$

In addition to our assumption that individuals of species i sample their environment via a normal distribution with density $u_i(\zeta, x)$, we further assume normally distributed trait values for each of the S species. In this case $\overline{\mathcal{O}}_{ij}(x,t)$ simplifies to

$$\overline{\mathcal{O}}_{ij}(x,t) = \frac{1}{N_j(t)}\int_{\mathbb{R}} \mathcal{O}_{ij}(x,y)\nu_j(y,t)dy = \frac{U_iU_j}{\sqrt{2\pi(\lambda_i + \lambda_j + \sigma_j{}^2(t))}}\exp\left(-\frac{(x-\bar{x}_j(t))^2}{2(\lambda_i+\lambda_j+\sigma_j{}^2(t))}\right), \qquad \text{(SM.91)}$$

where $\sigma_i^2(t)$ is the variance of niche-centers in species i at time t. Adopting the model of imperfect inheritance formulated in SM §4, we recall the expressed trait of an individual x_i is normally distributed around its breeding value g_i with variance E_i. We call E_i the environmental variance and G_i, which is the variance of breeding values, the additive genetic variance for species i. Under this model of inheritance, the variance of expressed traits decomposes as $\sigma_i^2(t) = G_i(t) + E_i$.

To simplify notation we set

$$R_i = \ln\left(\frac{U_i w_{\max,i}}{\sqrt{1+A_i\lambda_i}}\right), \qquad \text{(SM.92a)}$$

$$a_i = \frac{A_i}{1+A_i\lambda_i}, \qquad \text{(SM.92b)}$$

$$\tilde{b}_{ij}(t) = \frac{1}{\lambda_i+\lambda_j+\sigma_j^2(t)}, \qquad \text{(SM.92c)}$$

where R_i is the innate growth rate, a_i is the strength of abiotic stabilizing selection and $\tilde{b}_{ij}$ is an intermediate variable mediating the sensitivity of individual fitness for members of species i to niche-overlap with members of species j. With this notation, the growth rate $m_i(\boldsymbol{\nu}, x)$ can be expressed as

$$m_i(\boldsymbol{\nu}, x) = R_i - \frac{a_i}{2}(x-\theta_i)^2 - \sum_{j=1}^{S} c_{ij}N_jU_iU_j\sqrt{\frac{\tilde{b}_{ij}}{2\pi}}\exp\left(-\frac{\tilde{b}_{ij}}{2}(x-\bar{x}_j)^2\right). \qquad \text{(SM.93)}$$

For the remainder of the calculation we suppress notation indicating dependency on $\boldsymbol{\nu}$ and x. From (SM.93) we calculate

$$\frac{\partial m_i}{\partial \bar{x}_i} = c_{ii} N_i U_i^2 \tilde{b}_{ii} (x - \bar{x}_i) \sqrt{\frac{\tilde{b}_{ii}}{2\pi}} \exp\left(-\frac{\tilde{b}_{ii}}{2} (x - \bar{x}_i)^2 \right), \tag{SM.94}$$

$$\begin{aligned} \frac{\partial m_i}{\partial G_i} &= -\frac{c_{ii} N_i U_i^2}{2} \left(\frac{(x - \bar{x}_i)^2 - G_i - E_i - 2\lambda_i}{(G_i + E_i + 2\lambda_i)^2} \right) \sqrt{\frac{\tilde{b}_{ii}}{2\pi}} \exp\left(-\frac{\tilde{b}_{ii}}{2} (x - \bar{x}_i)^2 \right) \\ &= -\frac{c_{ii} N_i U_i^2 \tilde{b}_{ii}^2}{2} \left((x - \bar{x}_i)^2 - {\sigma_i}^2 - 2\lambda_i \right) \sqrt{\frac{\tilde{b}_{ii}}{2\pi}} \exp\left(-\frac{\tilde{b}_{ii}}{2} (x - \bar{x}_i)^2 \right). \end{aligned} \tag{SM.95}$$

Note that

$$\begin{aligned} &\sqrt{\frac{\tilde{b}_{ii}}{2\pi}} \exp\left(-\frac{\tilde{b}_{ii}}{2} (x - \bar{x}_i)^2 \right) \sqrt{\frac{1}{2\pi {\sigma_i}^2}} \exp\left(-\frac{(x - \bar{x}_i)^2}{2{\sigma_i}^2} \right) \\ &\qquad = \sqrt{\frac{1}{2\pi({\sigma_i}^2 + 1/\tilde{b}_{ii})}} \sqrt{\frac{{\sigma_i}^2 + 1/\tilde{b}_{ii}}{2\pi {\sigma_i}^2/\tilde{b}_{ii}}} \exp\left(-\frac{{\sigma_i}^2 + 1/\tilde{b}_{ii}}{2{\sigma_i}^2/\tilde{b}_{ii}} (x - \bar{x}_i)^2 \right) \\ &\qquad = \sqrt{\frac{1}{4\pi({\sigma_i}^2 + \lambda_i)}} \sqrt{\frac{2({\sigma_i}^2 + \lambda_i)}{2\pi {\sigma_i}^2 ({\sigma_i}^2 + 2\lambda_i)}} \exp\left(-\frac{{\sigma_i}^2 + \lambda_i}{{\sigma_i}^2({\sigma_i}^2 + 2\lambda_i)} (x - \bar{x}_i)^2 \right). \end{aligned} \tag{SM.96}$$

Hence, gradients of the growth rate m_i averaged across the phenotypic distribution p_i become

$$\overline{\frac{\partial m_i}{\partial \bar{x}_i}} = 0, \tag{SM.97}$$

$$\begin{aligned} \overline{\frac{\partial m_i}{\partial G_i}} &= -\frac{c_{ii} N_i U_i^2}{2({\sigma_i}^2 + 2\lambda_i)^2} \left(\frac{({\sigma_i}^2 + 2\lambda_i){\sigma_i}^2}{2(\lambda_i + {\sigma_i}^2)} - {\sigma_i}^2 - 2\lambda_i \right) \sqrt{\frac{b_{ii}}{2\pi}} \\ &= -\frac{c_{ii} N_i U_i^2}{2({\sigma_i}^2 + 2\lambda_i)} \left(\frac{{\sigma_i}^2}{2({\sigma_i}^2 + \lambda_i)} - 1 \right) \sqrt{\frac{b_{ii}}{2\pi}} = \frac{c_{ii} N_i U_i^2 b_{ii}}{4} \sqrt{\frac{b_{ii}}{2\pi}}, \end{aligned} \tag{SM.98}$$

where

$$b_{ij} = \frac{1}{\lambda_i + \lambda_j + {\sigma_i}^2 + {\sigma_j}^2}. \tag{SM.99}$$

The growth rate for species i is

$$\bar{m}_i = R_i - \frac{a_i}{2} \left((\bar{x}_i - \theta_i)^2 + G_i + E_i \right) - \sum_{j=1}^{S} c_{ij} N_j U_i U_j \sqrt{\frac{b_{ij}}{2\pi}} \exp\left(-\frac{b_{ij}}{2} (\bar{x}_i - \bar{x}_j)^2 \right). \tag{SM.100}$$

Thus, we find the following growth rate gradients

$$\frac{\partial \bar{m}_i}{\partial \bar{x}_i} = a_i(\theta_i - \bar{x}_i) - \sum_{j=1}^{S} c_{ij} N_j U_i U_j b_{ij} (\bar{x}_j - \bar{x}_i) \sqrt{\frac{b_{ij}}{2\pi}} \exp\left(-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2\right), \tag{SM.101}$$

$$\frac{\partial \bar{m}_i}{\partial G_i} = -\frac{a_i}{2} + \frac{1}{2} \sum_{j=1}^{S} c_{ij} N_j U_i U_j b_{ij} \left(1 - b_{ij}(\bar{x}_i - \bar{x}_j)^2\right) \sqrt{\frac{b_{ij}}{2\pi}} \exp\left(-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2\right). \tag{SM.102}$$

In particular, we find

$$\left(\frac{\partial \bar{m}_i}{\partial G_i} - \frac{\overline{\partial m_i}}{\partial G_i}\right) = -\frac{a_i}{2} + \frac{1}{2}\left(\frac{c_{ii} N_i U_i^2 b_{ii}}{2}\sqrt{\frac{b_{ii}}{2\pi}} + \sum_{j \neq i} c_{ij} N_j U_i U_j b_{ij} \left(1 - b_{ij}(\bar{x}_i - \bar{x}_j)^2\right) \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2}\right). \tag{SM.103}$$

Applying equations (21a), (24a) and (24b) of the main text recovers system (30) of the main text.

5.4 Initial Dynamics of Figure 1

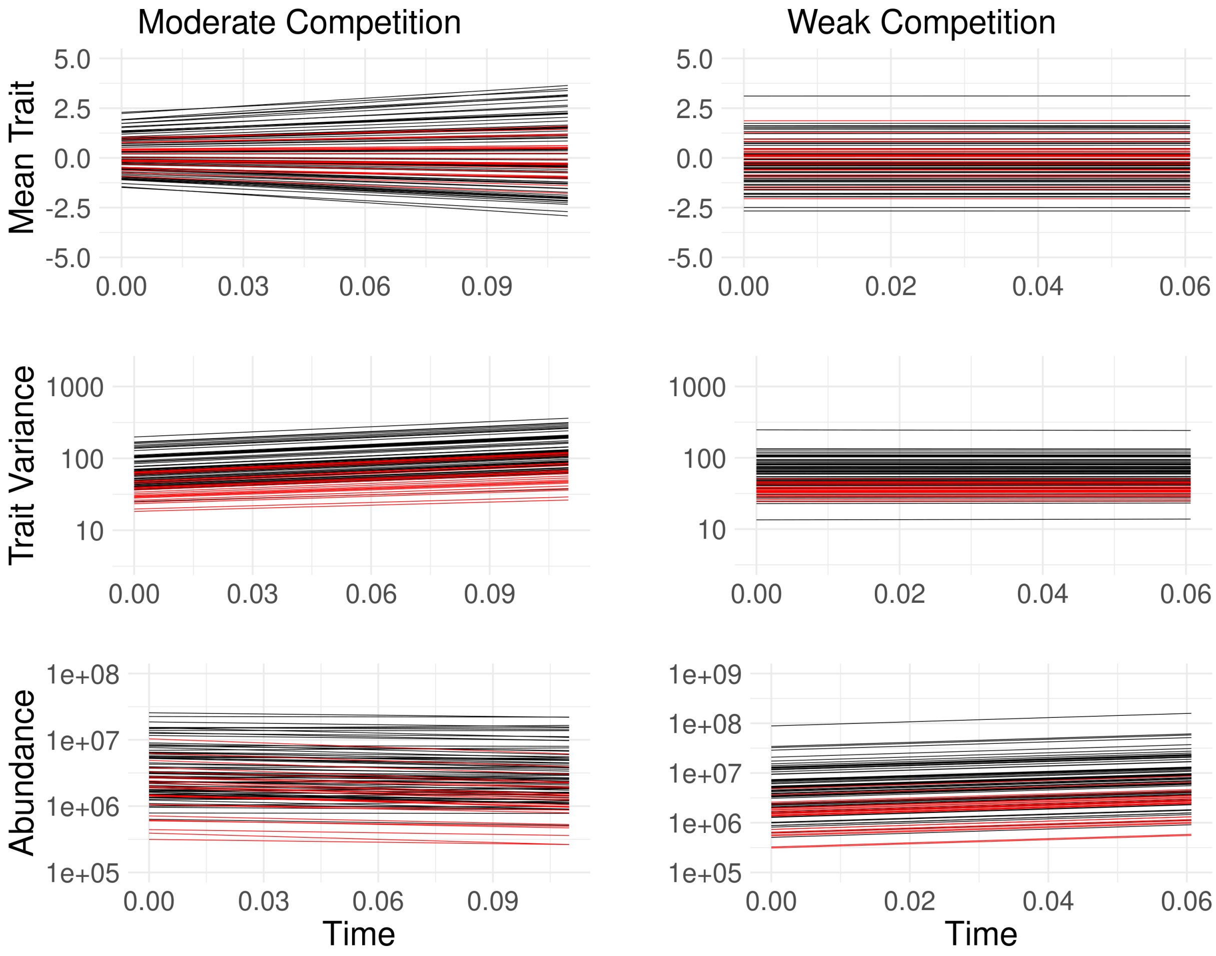


Figure S2: Early dynamics of numerical solution presented in Figure 1 of the main text. Black lines represent species that persisted to the end of numerical integration. Red lines represent species that became extinct at some time before the numerical integration ended. From this we see extinction is associated with lower initial trait variances and lower initial abundances.

5.5 Relaxing the Gaussian Population Approximation

In this subsection, we briefly explore the consequences of relaxing the assumption of normally distributed trait values within species. To do so, we apply SAGA to the classical niche framework developed above. In particular, we continue to quantify niche-overlap between individuals of species i and j by $\mathcal{O}_{ij}(x_i, x_j)$ following equation (SM.80). Thus, without making the Gaussian population assumption, we can still write the growth rate associated with niche location x for species i as

$$m_i(\boldsymbol{\nu}, x) = R_i - \frac{a_i}{2}(\theta_i - x)^2 - \sum_{j=1}^{S} c_{ij} \int_{\mathbb{R}} \mathcal{O}_{ij}(x,y)\nu_j(y)dy. \tag{SM.104}$$

Our goal here is to determine conditions under which the Gaussian population assumption remains approximately valid and also to explore the extent to which this assumption may be violated. Taking a numerical approach, we investigated five scenarios: 1) Symmetric competition (so that competition with conspecifics is just as strong as competition with heterospecifics), strong mutation, large population sizes and three species, 2) Asymmetric competition (so that competition with conspecifics is weaker than competition with heterospecifics), strong mutation, large population sizes and three species, 3) Symmetric competition, weak mutation, large population sizes and three species, 4) Symmetric competition, strong mutation, small population sizes and three species and 5) Symmetric competition, strong mutation, large population sizes and ten species. For each scenario we set $a_i = 2e-4$, $\theta_i = 0$ and $U_i = \lambda_i = V_i = 1$ for each species i and initialized abundance densities as smooth, compactly supported functions. In particular, we set

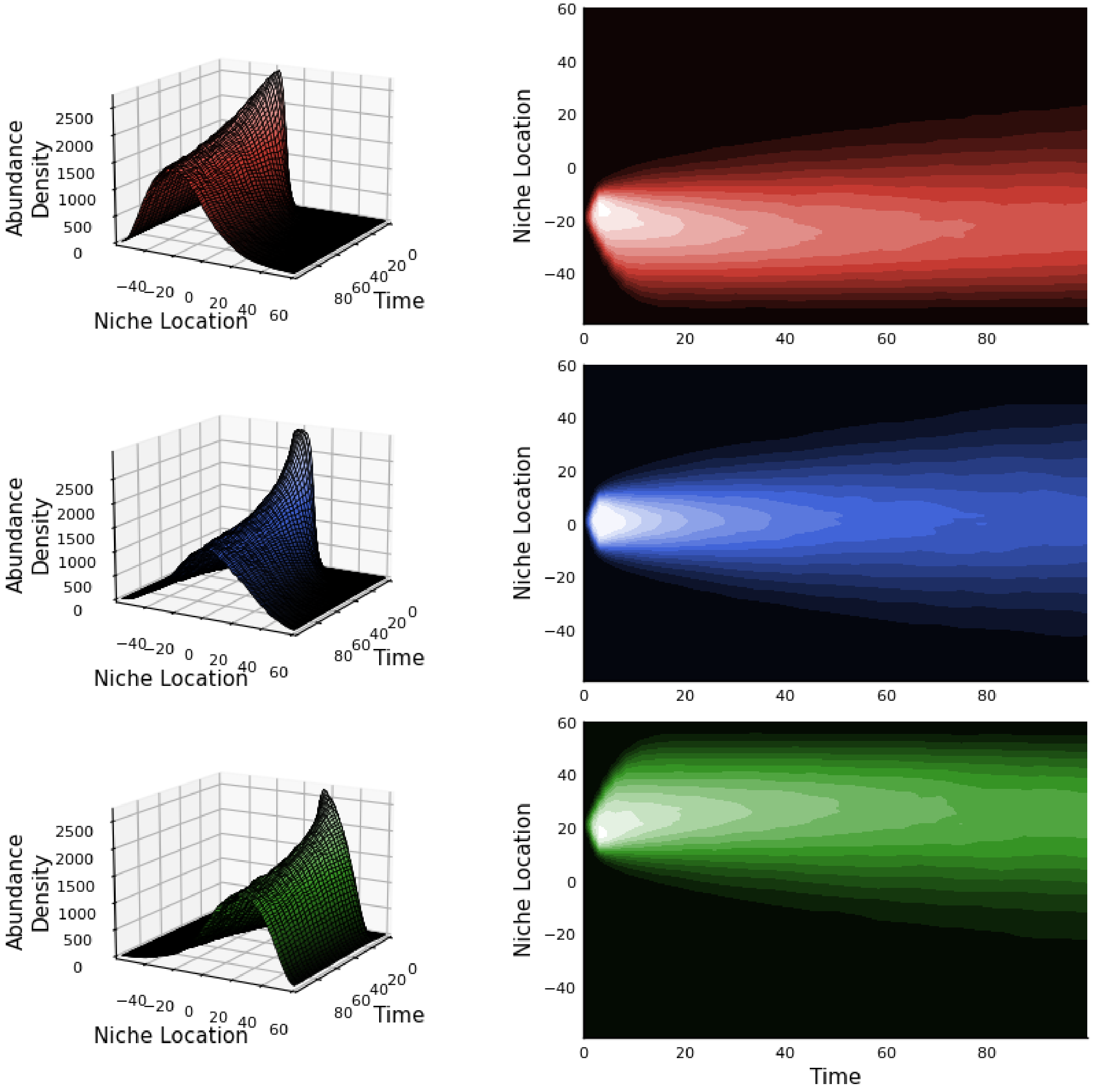
$$\nu_i(x,0) = \begin{cases} M_i \exp\left(1 - \left(1 - \frac{(x-\bar{x}_i(0))^2}{\varsigma_i}\right)^{-1}\right), & x \in (\bar{x}_i(0) - \sqrt{\varsigma_i}, \bar{x}_i(0) + \sqrt{\varsigma_i}) \\ 0, & x \notin (\bar{x}_i(0) - \sqrt{\varsigma_i}, \bar{x}_i(0) + \sqrt{\varsigma_i}). \end{cases} \tag{SM.105}$$

For each scenario involving three species we set $\bar{x}_1(0) = -20$, $\bar{x}_2(0) = 0$, $\bar{x}_3(0) = 20$ and $\varsigma_i = 4$ for $i = 1,2,3$. For the scenario of ten competing species, we evenly spaced $\bar{x}_1(0), \ldots, \bar{x}_{10}(0)$ between -45 and 45 and set $\varsigma_i = 1$ for $i = 1, \ldots, 10$. The remaining parameter values and initial conditions chosen for each scenario are provided in Table S2. We numerically integrated each scenario for 100 units of time. Results for each scenario are shown in Figures S3-S17.

Our results demonstrate the Gaussian population approximation may hold when mutation is strong and competition is symmetric even when population sizes are small and when many species are competing. When comparing the SPDE solutions that relax the Gaussian population assumption to solutions of corresponding ODE that require the Guassian population assumption, we find the solutions agree over a brief initial time interval. Afterwards the solutions then quickly diverge, leading to increased abundance and trait variance when the Gaussian population assumption is relaxed. However, many qualitative features of the solutions remain intact, such as the relationship between abundance and location along the niche axis.

Table S2: Values of model parameters used for numerical integration of SPDE model. Abbreviations correspond to: SC = Symmetric Competition, SM = Strong Mutation, LP = Large Population sizes, AC = Asymmetric Competition, WM = Weak Competition, SP = Small Population sizes, 3S = Three Species and 10S = Ten Species.

Scenario	S	R_i	μ_i	c_{ii}	$c_{ij},\ i \neq j$	M_i
1) SC, SM, LP, 3S	3	3	5	1e-3	1e-3	250
2) AC, SM, LP, 3S	3	3	5	5e-4	1e-3	250
3) SC, WM, LP, 3S	3	3	0.5	1e-3	1e-3	250
4) SC, SM, SP, 3S	3	1	5	5e-2	5e-2	10
5) SC, SM, LP, 10S	10	3	5	1e-4	1e-4	250



SPDE solutions for scenario 1

Figure S3: The evolution of abundance densities for three competing species (colored red, blue and green). Left column shows surface plots of each species separately and right column shows contour plots of the same numerical solution.

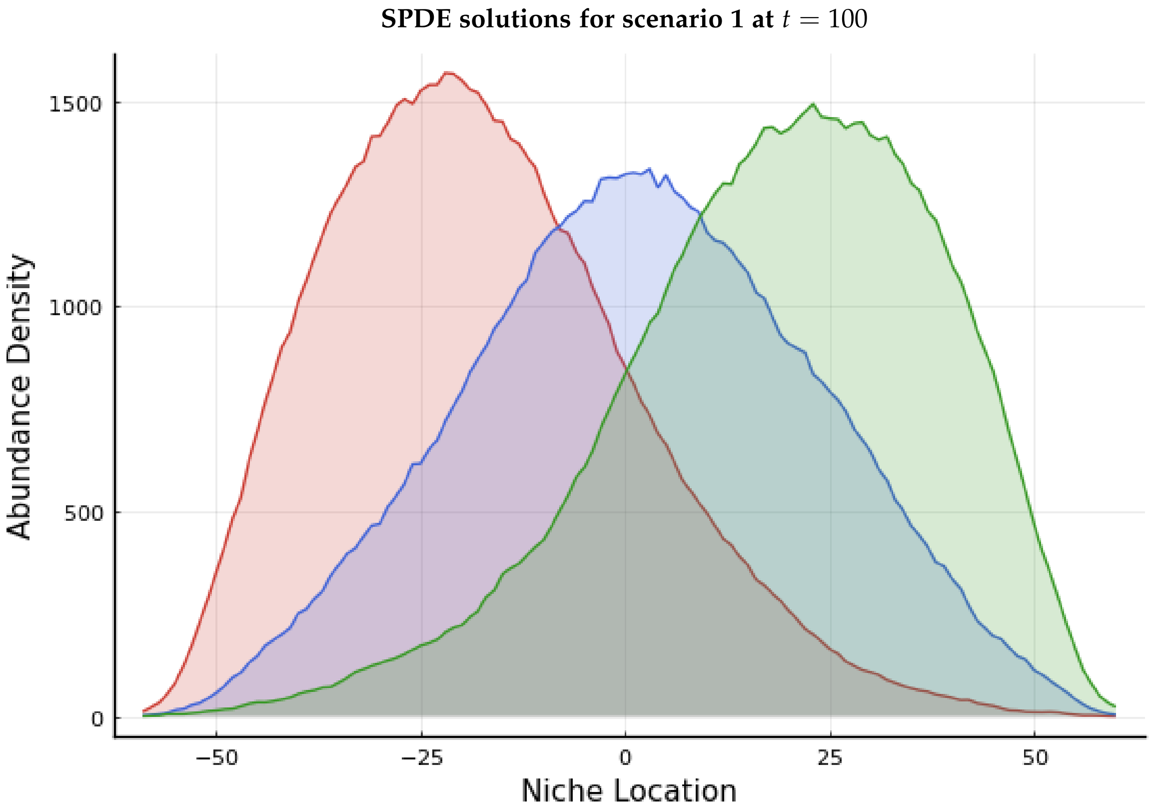
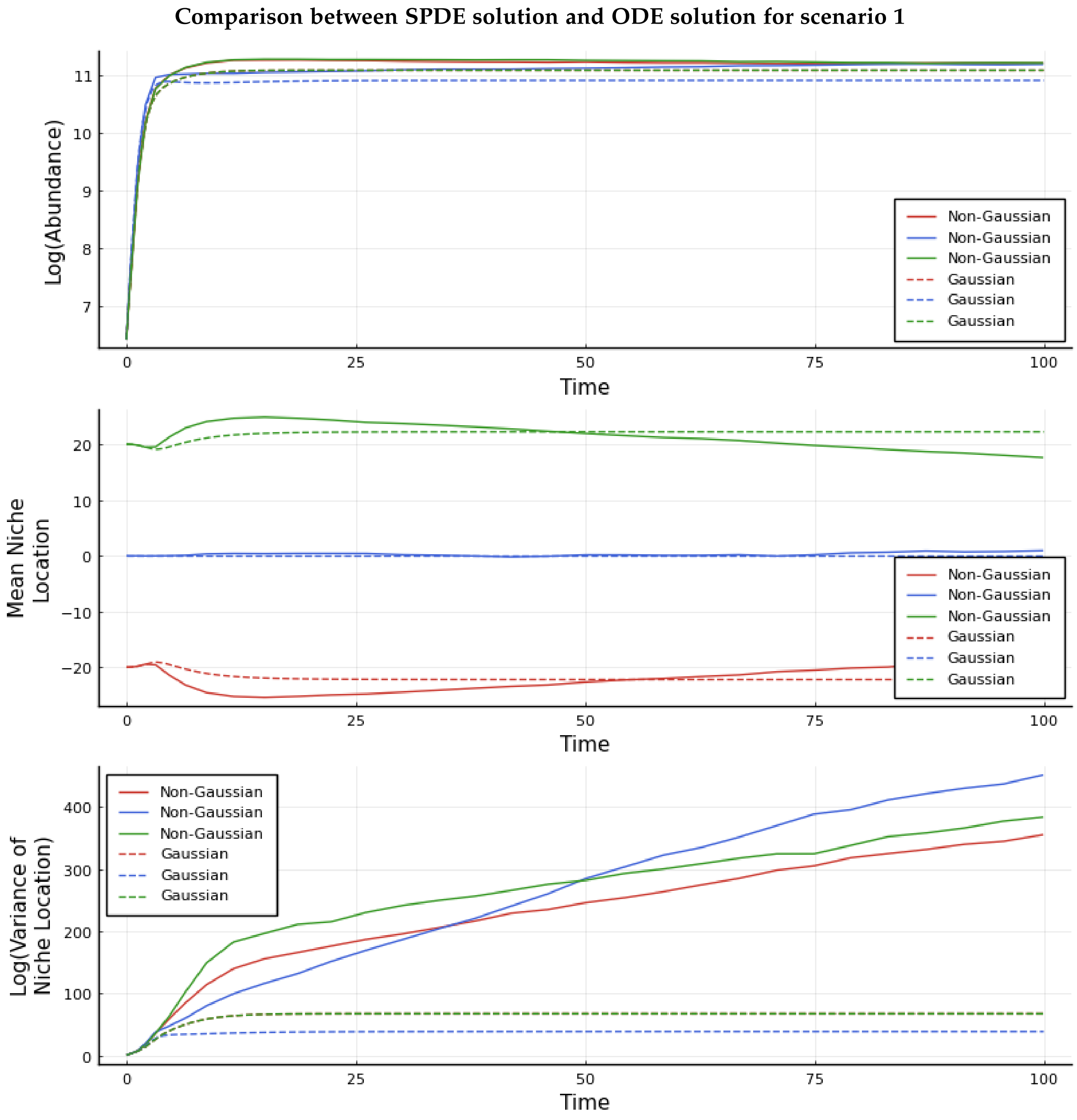


Figure S4: Abundance densities of the same three competing species as in Figure S3 at time $t = 100$. The red and green trait distributions appear slightly skewed towards the center which is likely due to the effects of abiotic stabilizing selection. The extreme overlap in niche space is due to symmetric competition.



Comparison between SPDE solution and ODE solution for scenario 1

Figure S5: Time series of $\ln N_i$ (top), $\bar{x}_i$ (middle) and $\ln \sigma_i^2$ (bottom). Solid lines correspond to the SPDE solutions for each color shown in Figure S3. Dashed lines are the corresponding ODE solutions which make use of the Gaussian population assumption. Some solutions may overlap and thus are hidden by the most recently plotted solution.

SPDE solutions for scenario 2

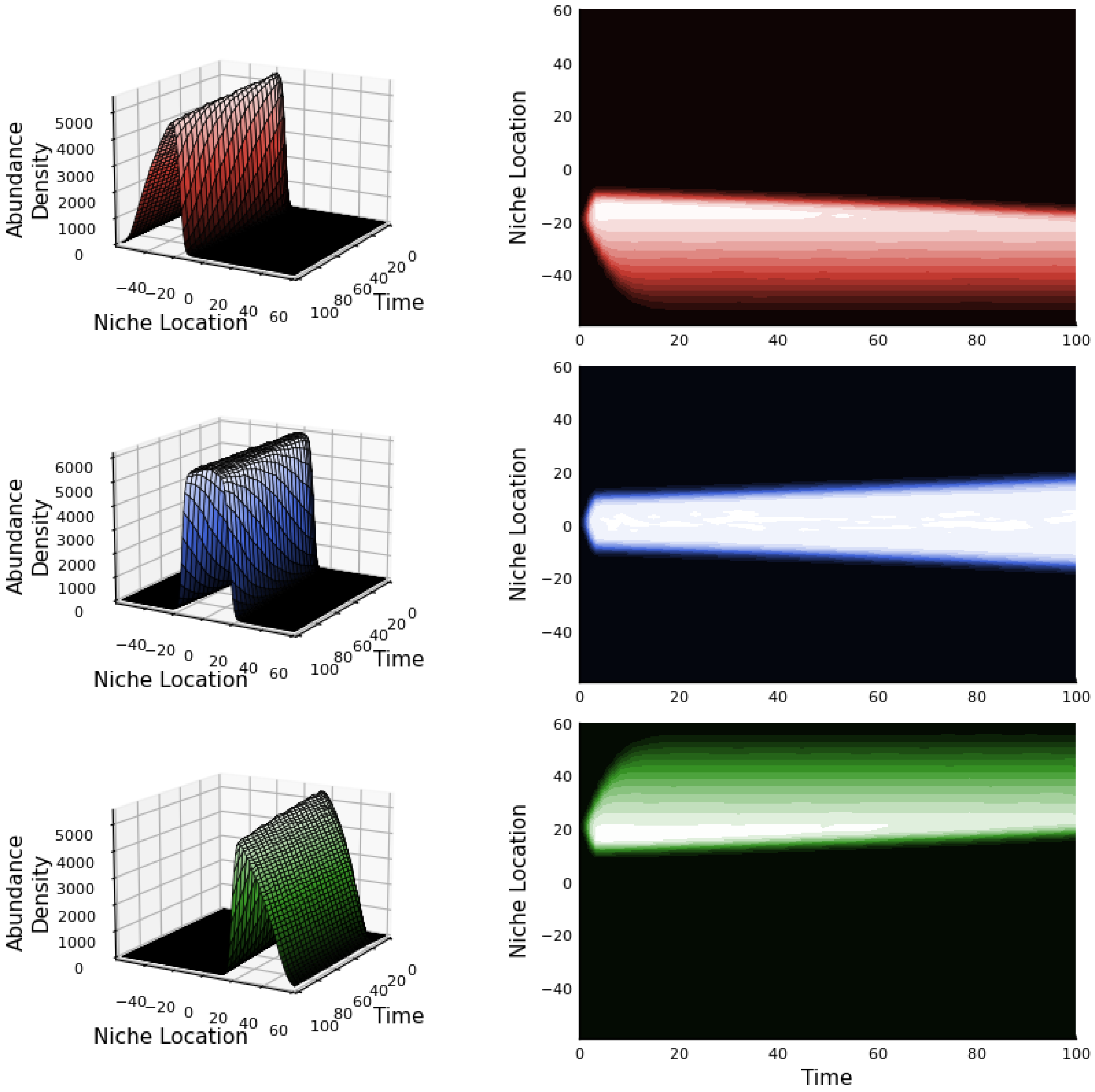


Figure S6: The evolution of abundance densities for three competing species (colored red, blue and green). Left column shows surface plots of each species separately and right column shows contour plots of the same numerical solution.

SPDE solutions for scenario 2 at $t = 100$

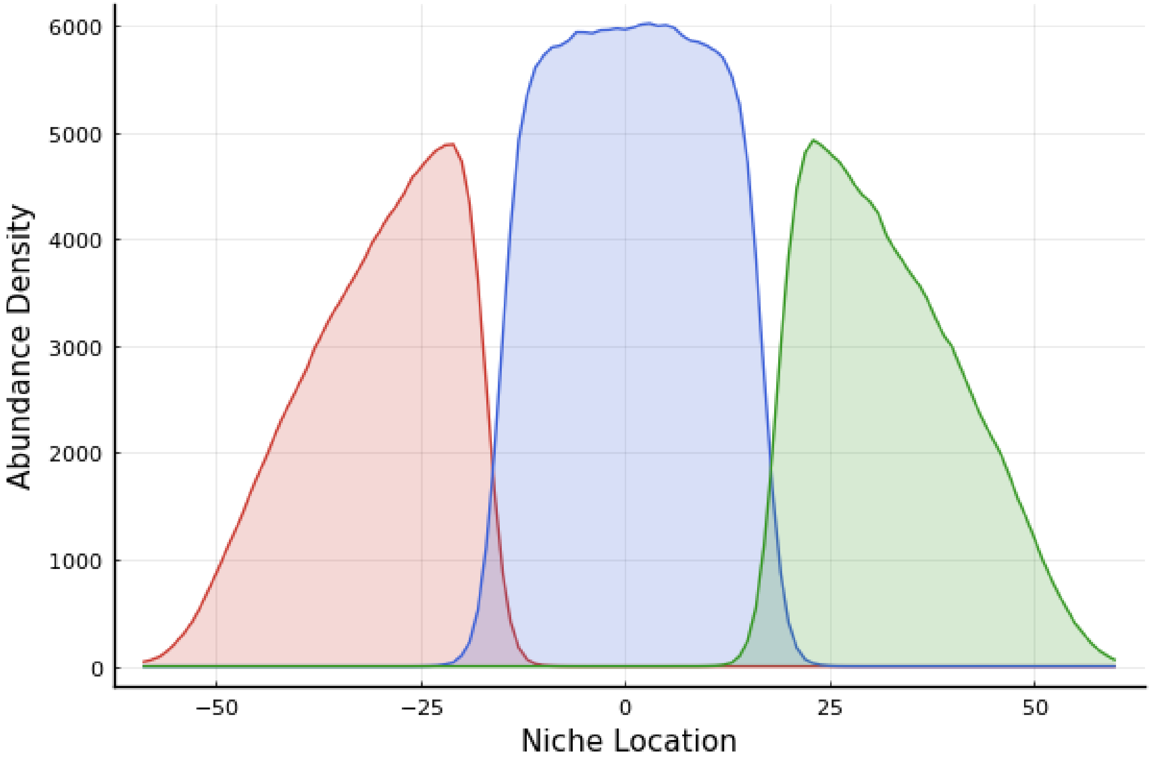
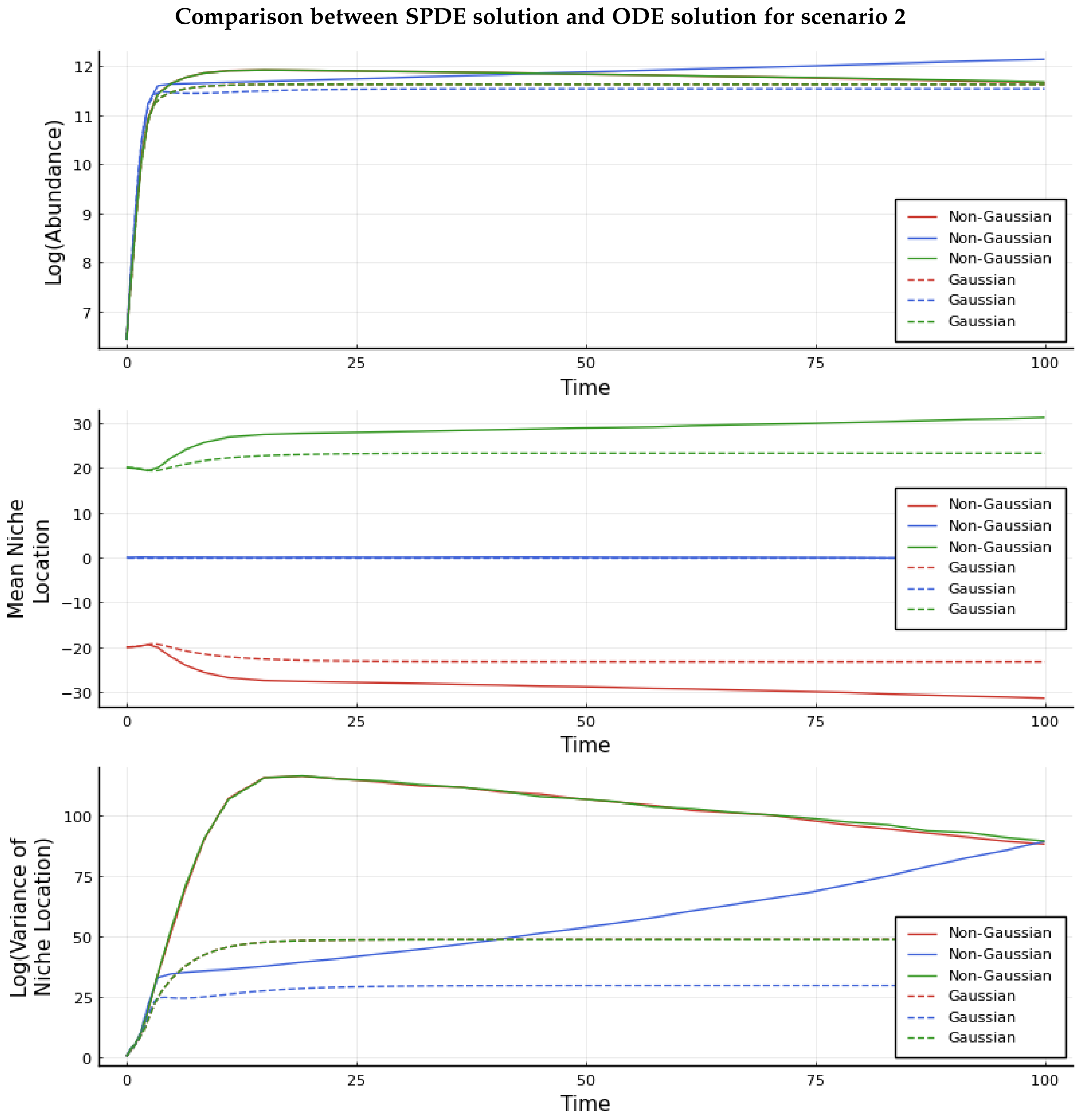


Figure S7: Abundance densities of the same three competing species as in Figure S6 at time $t = 100$. Under this scenario niche space becomes acutely partitioned due to asymmetric competition and the concave-downward curve of abiotic stabilizing selection becomes clearly visible. The red and green trait distributions become heavily skewed away from the center and the blue distribution becomes extremely platykurtic. The Gaussian population assumption may not provide a reasonable approximation for this scenario.



Comparison between SPDE solution and ODE solution for scenario 2

Figure S8: Time series of $\ln N_i$ (top), $\bar{x}_i$ (middle) and $\ln \sigma_i^2$ (bottom). Solid lines correspond to the SPDE solutions for each color shown in Figure S6. Dashed lines are the corresponding ODE solutions which make use of the Gaussian population assumption. Some solutions may overlap and thus are hidden by the most recently plotted solution.

SPDE solutions for scenario 3

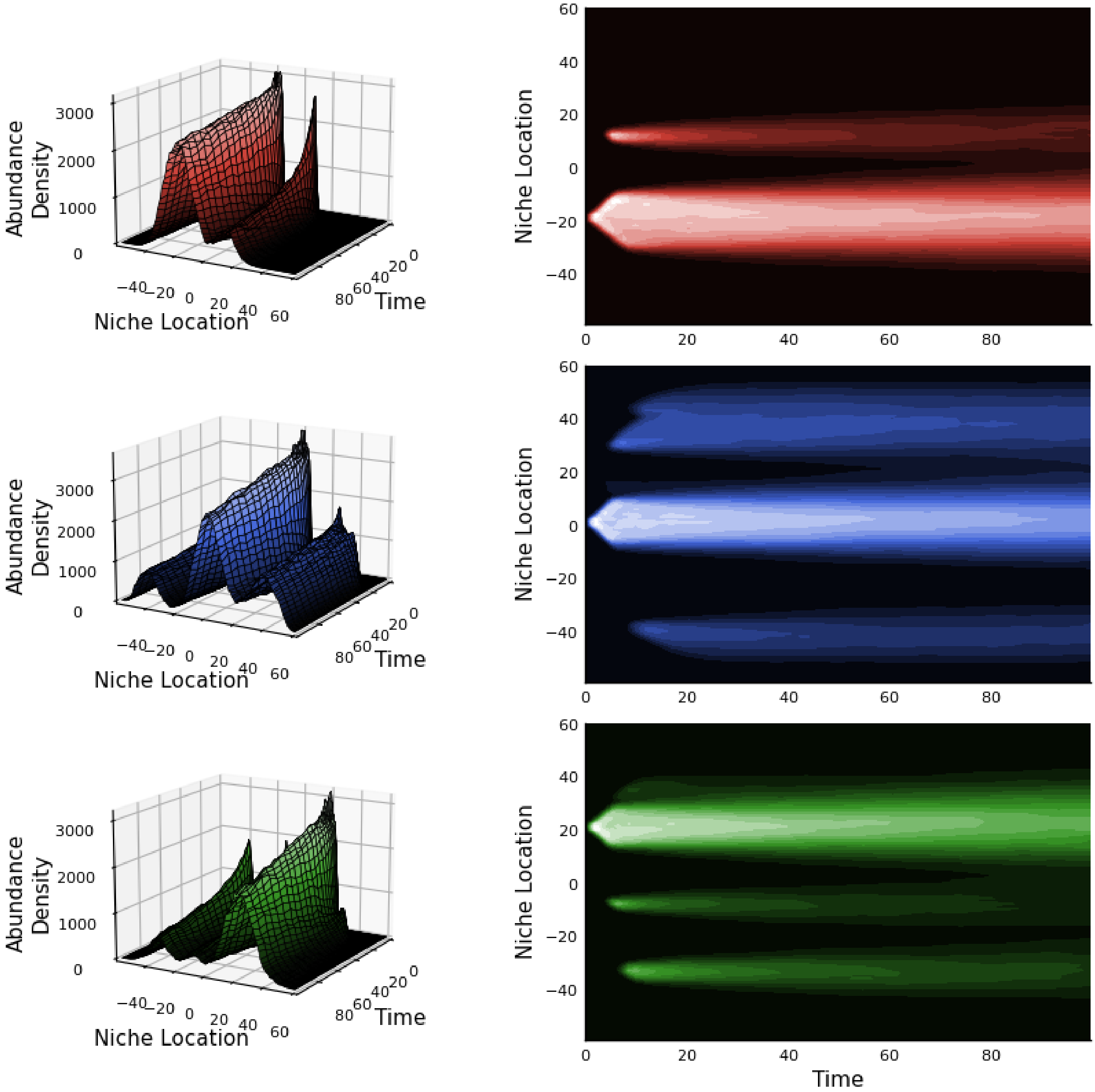


Figure S9: The evolution of abundance densities for three competing species (colored red, blue and green). Left column shows surface plots of each species separately and right column shows contour plots of the same numerical solution.

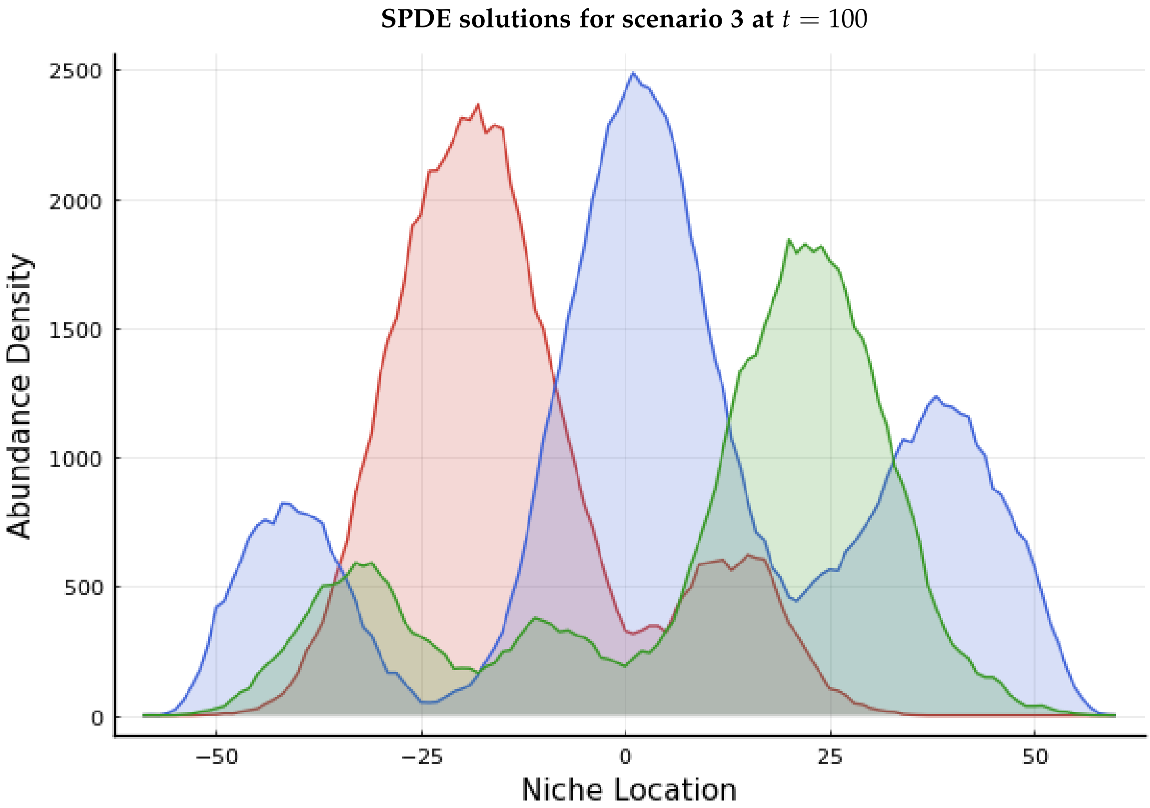


Figure S10: Abundance densities of the same three competing species as in Figure S9 at time $t = 100$. In this scenario with weak mutation, we see that phenotypic distributions evolve to become highly multimodal, clearly breaking the Gaussian population assumption.

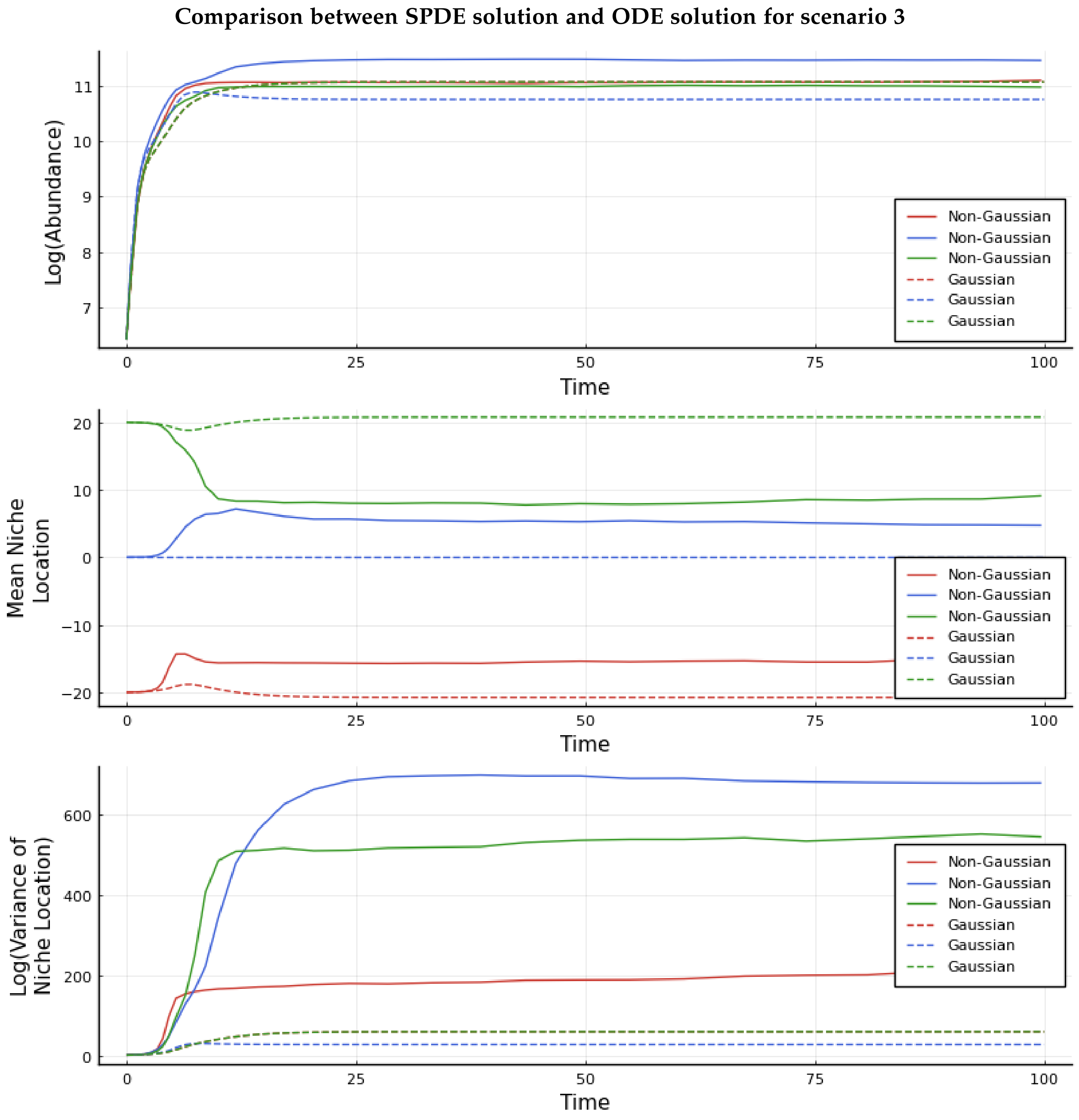
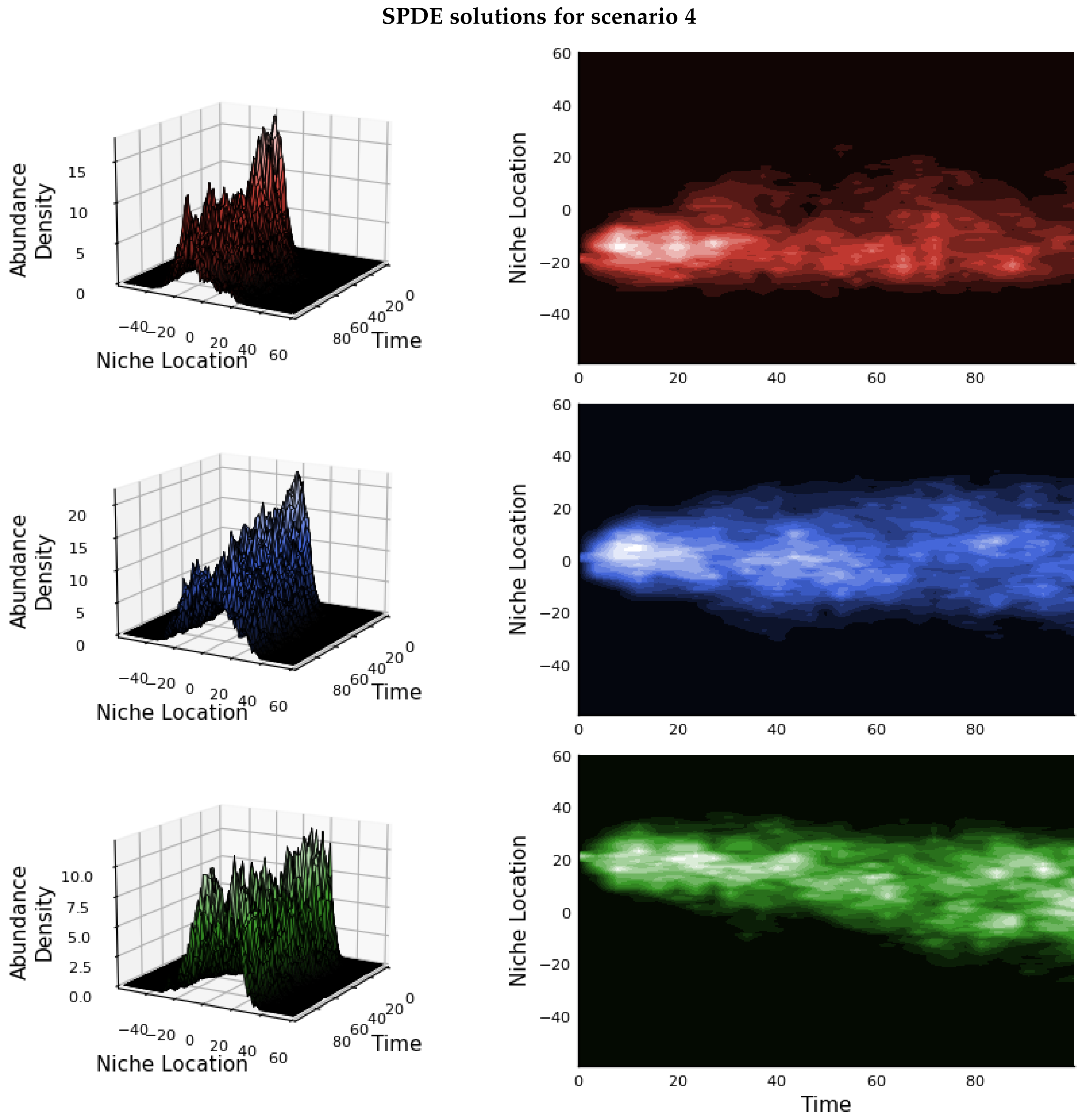

Comparison between SPDE solution and ODE solution for scenario 3

Figure S11: Time series of $\ln N_i$ (top), $\bar{x}_i$ (middle) and $\ln \sigma_i^2$ (bottom). Solid lines correspond to the SPDE solutions for each color shown in Figure S9. Dashed lines are the corresponding ODE solutions which make use of the Gaussian population assumption. Some solutions may overlap and thus are hidden by the most recently plotted solution.



SPDE solutions for scenario 4

Figure S12: The evolution of abundance densities for three competing species (colored red, blue and green) under scenario 4). Left column shows surface plots of each species separately and right column shows contour plots of the same numerical solution.

SPDE solutions for scenario 4 at $t = 100$

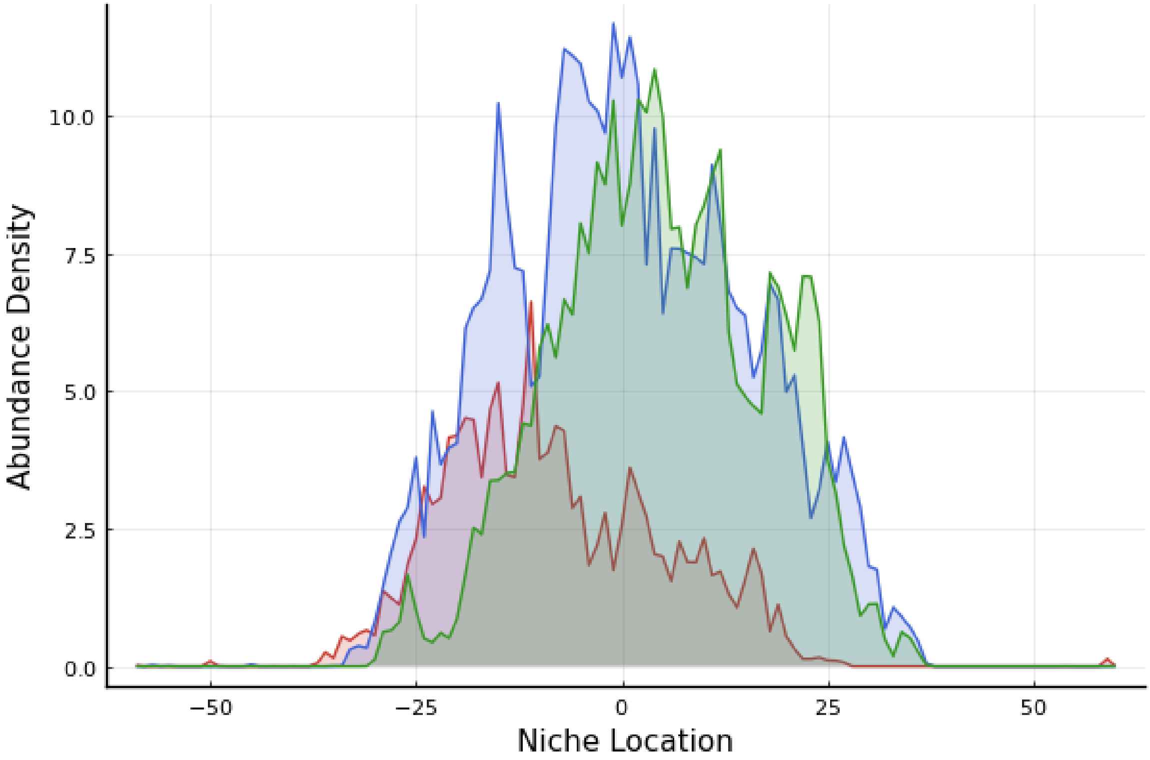
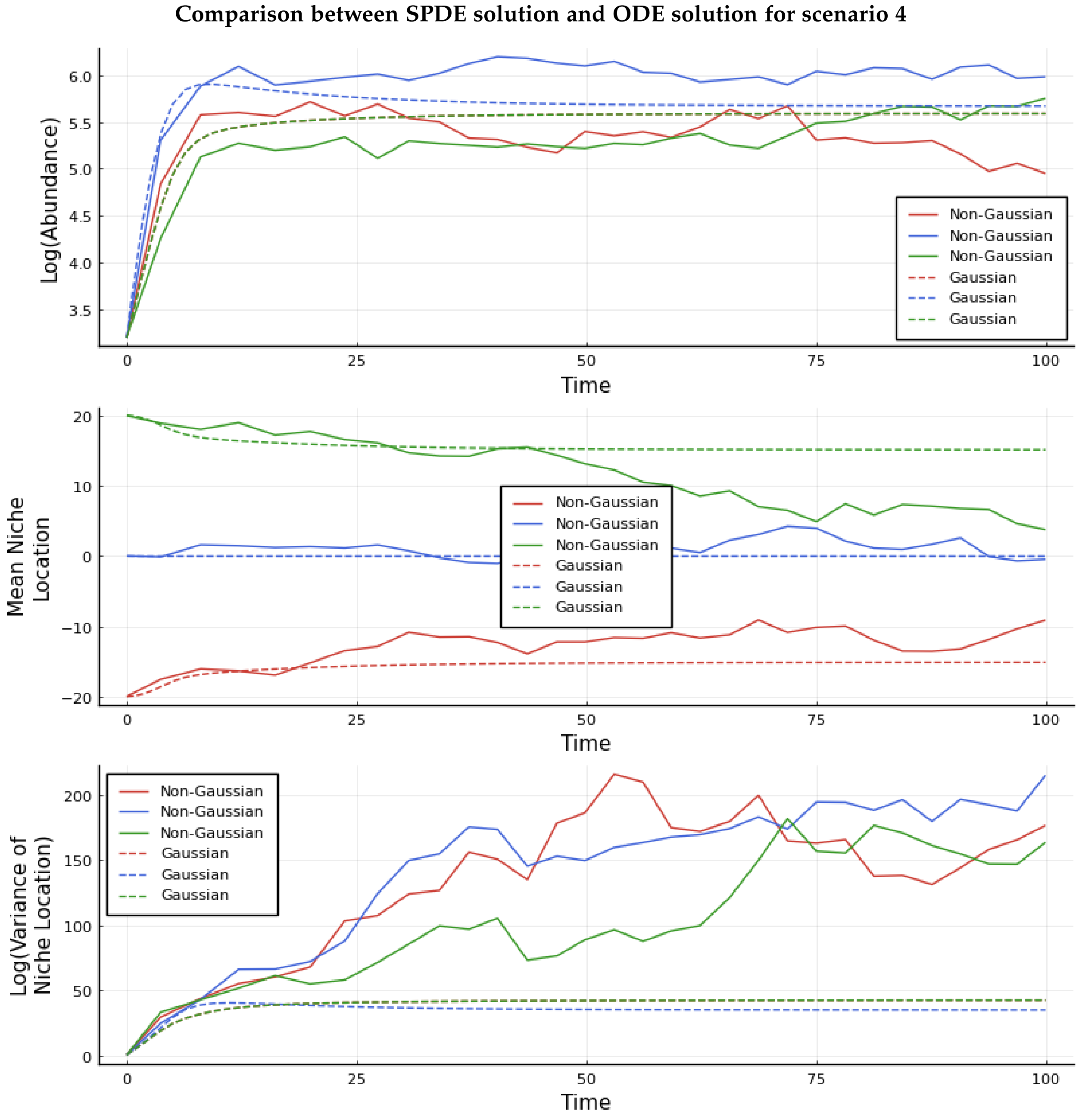


Figure S13: Abundance densities of the same three competing species as in Figure S12 at time $t = 100$.



Comparison between SPDE solution and ODE solution for scenario 4

Figure S14: Time series of $\ln N_i$ (top), $\bar{x}_i$ (middle) and $\ln \sigma_i^2$ (bottom). Solid lines correspond to the SPDE solutions for each color shown in Figure S12. Dashed lines are the corresponding ODE solutions which make use of the Gaussian population assumption. Some solutions may overlap and thus are hidden by the most recently plotted solution.

SPDE solutions for scenario 5

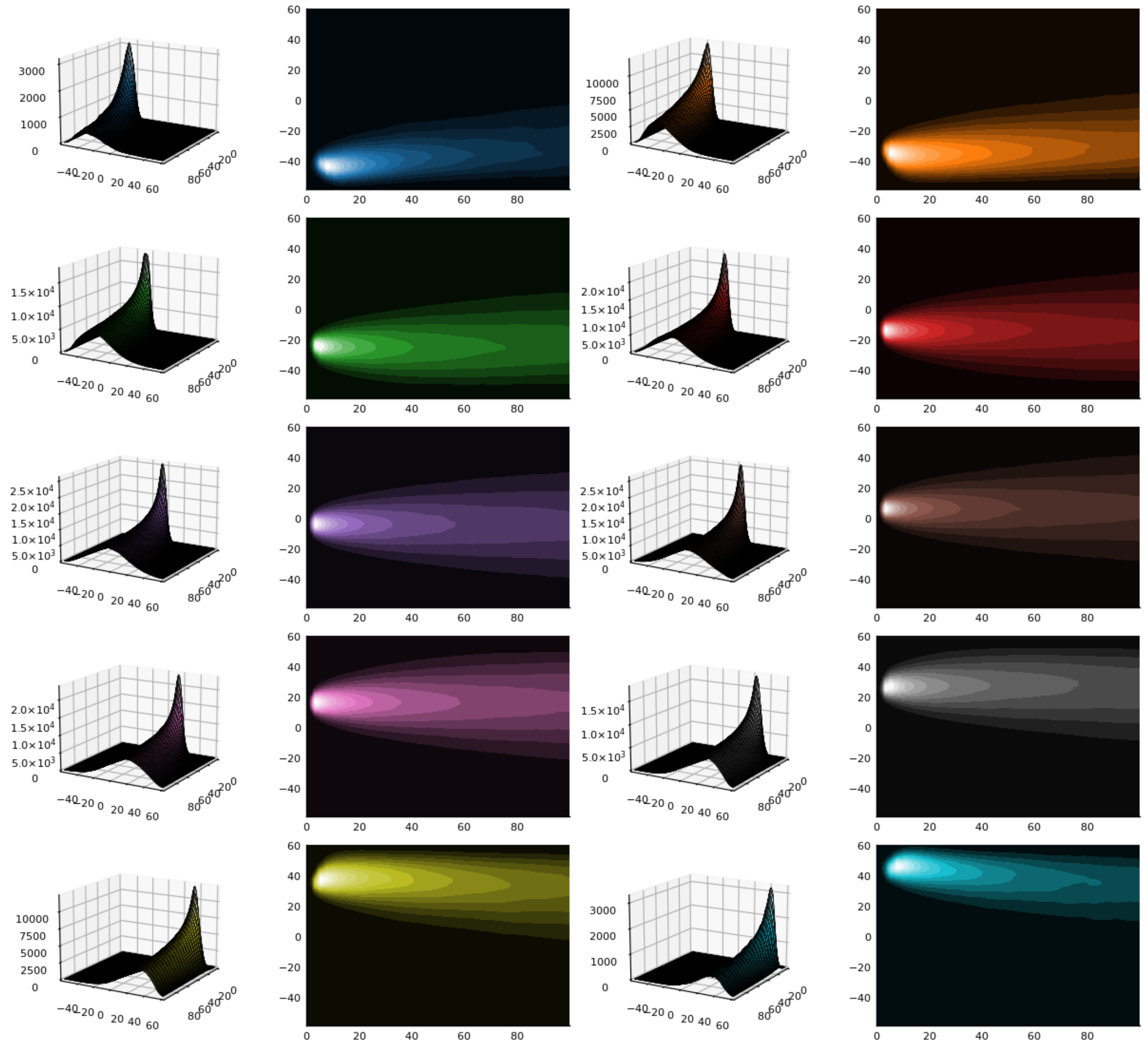


Figure S15: The evolution of abundance densities for ten competing species. Left and middle-right columns show surface plots of five species separately. Middle-left and right columns show contour plots of the same numerical solution.

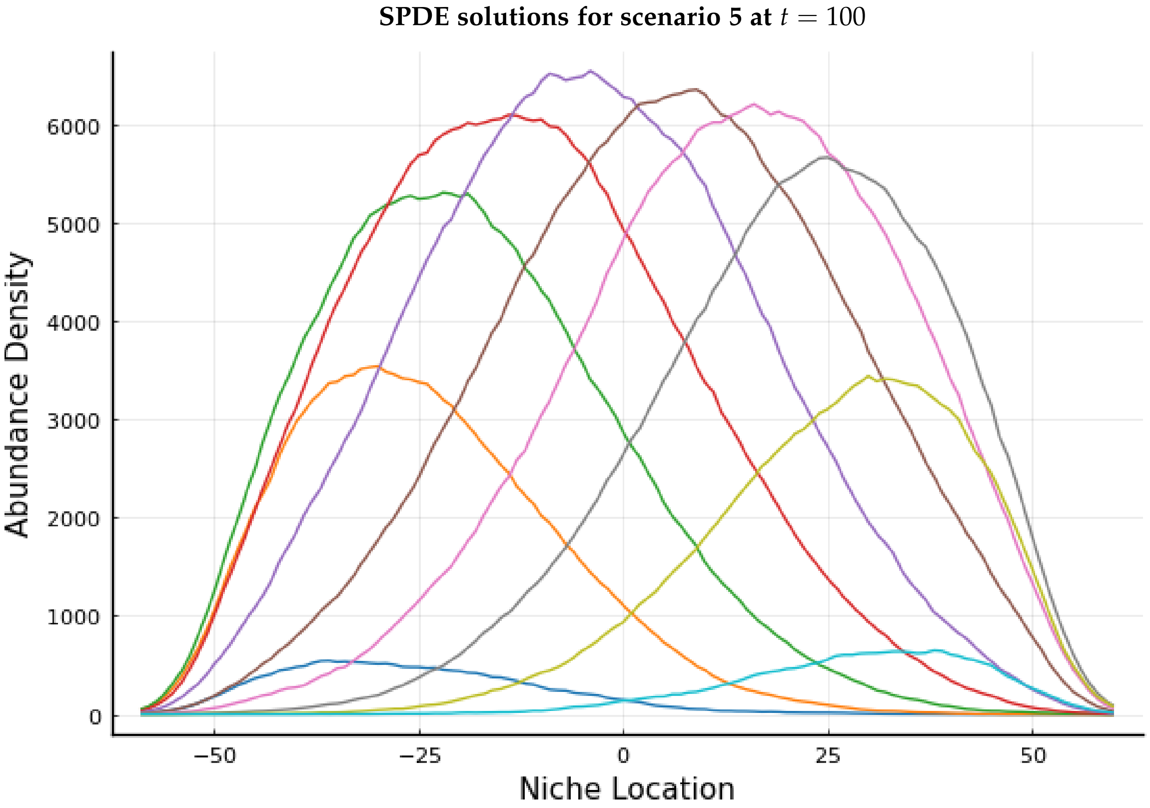


Figure S16: Abundance densities of the same ten competing species as in Figure S15 at time $t = 100$.

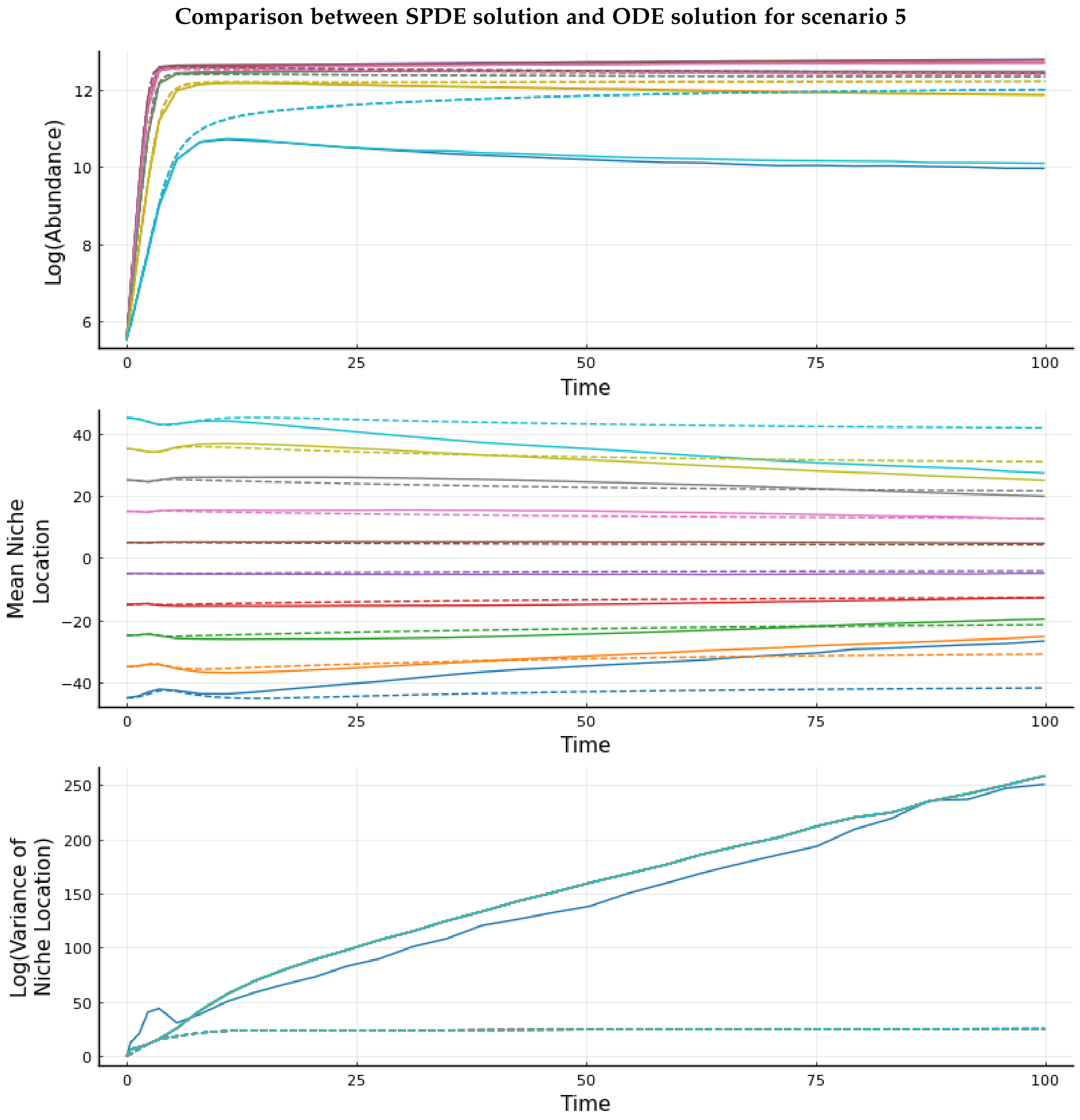

Comparison between SPDE solution and ODE solution for scenario 5

Figure S17: Time series of $\ln N_i$ (top), $\bar{x}_i$ (middle) and $\ln \sigma_i^2$ (bottom). Solid lines correspond to the SPDE solutions for each color shown in Figure S15. Dashed lines are the corresponding ODE solutions which make use of the Gaussian population assumption. Some solutions may overlap and thus are hidden by the most recently plotted solution.

6 Competition Coefficients and Selection Gradients

6.1 Definition of Selection Gradients

Our definition of selection gradients differs slightly from traditional definitions. In particular, Lande and Arnold (1983) express the linear selection gradient β in general as $\beta = \frac{1}{\sigma^2}\text{Cov}_t(w, x)$. This is convenient for discrete time models of mean trait evolution where the change in mean trait between generations is captured by

$$\Delta \bar{x} = \frac{G}{\sigma^2}\text{Cov}_t(w, x) = G\beta. \tag{SM.106}$$

However, in our case, we model mean trait evolution in continuous time via

$$\frac{d\bar{x}}{dt} = \frac{G}{\sigma^2}\text{Cov}_t(m, x). \tag{SM.107}$$

Hence, we define the linear selection gradient $\beta := \frac{1}{\sigma^2}\text{Cov}_t(m, x)$. Similarly, the quadratic selection gradient γ is expressed in Lande and Arnold (1983) as $\gamma = \frac{1}{\sigma^4}\text{Cov}_t(w, (x - \bar{x})^2)$. Then, in analogy to our definition of β, we define $\gamma := \frac{1}{\sigma^4}\text{Cov}_t(m, (x - \bar{x})^2)$.

6.2 Selection Gradients Under Abiotic Stabilizing Selection and Resource Competition

Here we provide expressions for selection gradients under our model of diffuse coevolution driven by resource competition. Combining our definitions of selection gradients with the results found in SM §5 (which require the Gaussian population assumption), our model of diffuse coevolution yields, for species i,

$$\beta_i = a_i(\theta_i - \bar{x}_i) - \sum_{j=1}^{S} c_{ij} N_j U_i U_j b_{ij} (\bar{x}_j - \bar{x}_i) \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_j - \bar{x}_i)^2}, \tag{SM.108a}$$

$$\gamma_i = -a_i + \frac{c_{ii} N_i U_i^2 b_{ii}}{2} \sqrt{\frac{b_{ii}}{2\pi}} + \sum_{j \neq i} c_{ij} N_j U_i U_j b_{ij} \left(1 - b_{ij}(\bar{x}_i - \bar{x}_j)^2\right) \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2}. \tag{SM.108b}$$

Note these selection gradients can be additively partitioned as $\beta_i = \beta_i^{(a)} + \sum_{j=1}^{S} \beta_{ij}$ and $\gamma_i = \gamma_i^{(a)} + \sum_{j=1}^{S} \gamma_{ij}$ where $\beta_i^{(a)}, \gamma_i^{(a)}$ denote the components due to abiotic stabilizing selection and β_{ij}, γ_{ij} denote the components due to interactions with species j. In particular, we find $\beta_i^{(a)} = a_i(\theta_i - \bar{x}_i)$, $\gamma_i^{(a)} = -a_i$ and

$$\beta_{ij} = c_{ij} N_j U_i U_j b_{ij} (\bar{x}_i - \bar{x}_j) \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2}, \tag{SM.109a}$$

$$\gamma_{ij} = c_{ij} N_j U_i U_j b_{ij} \left(1 - b_{ij}(\bar{x}_i - \bar{x}_j)^2\right) \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2},\ i \neq j \tag{SM.109b}$$

$$\gamma_{ii} = \frac{c_{ii} N_i U_i^2 b_{ii}}{2} \sqrt{\frac{b_{ii}}{2\pi}},\ i = j. \tag{SM.109c}$$

6.3 High Richness Approximations for Moments of Competition Coefficients and Selection Gradients

Here we derive covariances between competition coefficients and selection gradients following the model of diffuse coevolution derived in SM §5. We assume the community is very rich (i.e., the number of species S is very large) and that the distribution of mean traits is approximately normal and independent of the distribution of abundance. We denote by $\bar{\bar{x}}$, $V_{\bar{x}}$ the community-wide mean and variance of species mean traits and by $\bar{N}, V_N$ the community-wide mean and variance of species abundances. For simplicity we assume constant species trait variances and niche breadths so that $\sigma_i^2 = \sigma^2$ and $\lambda_i = \lambda$ for some $\sigma^2, \lambda > 0$. Thus $b_{ij} = b = 1/(2\sigma^2 + 2\lambda)$ for each $i, j = 1, \ldots, S$. Under these conditions, we can express competition coefficients, linear selection gradients and quadratic selection gradients respectively as

$$\alpha_{ij} = c_{ij} U_i U_j \sqrt{\frac{b}{2\pi}} e^{-\frac{b}{2}(\bar{x}_i - \bar{x}_j)^2}, \tag{SM.110a}$$

$$\beta_{ij} = c_{ij} U_i U_j N_j b (\bar{x}_i - \bar{x}_j) \sqrt{\frac{b}{2\pi}} e^{-\frac{b}{2}(\bar{x}_i - \bar{x}_j)^2}, \tag{SM.110b}$$

$$\gamma_{ij} = c_{ij} U_i U_j N_j b \big(1 - b(\bar{x}_i - \bar{x}_j)^2\big) \sqrt{\frac{b}{2\pi}} e^{-\frac{b}{2}(\bar{x}_i - \bar{x}_j)^2}, \; i \neq j. \tag{SM.110c}$$

To compute statistical distributions of these quantities we draw i and j independently from the set $\{1, \ldots, S\}$ each with probability $1/S$. Then the event $i = j$ occurs with probability $1/S^2$. We suppose S is large enough that we can safely ignore the event $i = j$.

Under our model of diffuse coevolution, the competition coefficients and selection gradients can be written in terms of the difference $D_{ij} = \bar{x}_i - \bar{x}_j$. By our assumption that i and j are drawn independently and that $\bar{x}_i$, $\bar{x}_j$ approximately follow a normal distribution with mean $\bar{\bar{x}}$ and variance $V_{\bar{x}}$, we see the distribution of D_{ij} is approximated by a normal distribution with mean zero and variance $2V_{\bar{x}}$.

We suppose the strengths of competition c_{ij} and niche-use parameters U_i are distributed independently of mean traits, abundances and each other. We write $\bar{c}$, $\bar{U}$ and V_c, V_U as the mean and variance of these parameters respectively.

6.3.1 Means and Variances of Competition Coefficients and Selection Gradients

Combining the above assumptions and notation, we can approximate the expectations of competition coefficients and selection gradients via

$$\bar{\alpha} = \frac{1}{S^2} \sum_{i,j=1}^{S} \alpha_{ij} \approx \bar{c}\bar{U}^2 \int_{\mathbb{R}} \sqrt{\frac{b}{2\pi}} e^{-\frac{b}{2}D^2} \frac{1}{\sqrt{4\pi V_{\bar{x}}}} e^{-\frac{D^2}{4V_{\bar{x}}}} dD = \bar{c}\bar{U}^2 \sqrt{\frac{b}{2\pi(2V_{\bar{x}}b + 1)}}, \tag{SM.111a}$$

$$\bar{\beta} = \frac{1}{S^2} \sum_{i,j=1}^{S} \beta_{ij} \approx \bar{c}\bar{U}^2 \bar{N} b \int_{\mathbb{R}} D \sqrt{\frac{b}{2\pi}} e^{-\frac{b}{2}D^2} \frac{1}{\sqrt{4\pi V_{\bar{x}}}} e^{-\frac{D^2}{4V_{\bar{x}}}} dD = 0, \tag{SM.111b}$$

$$\begin{aligned} \bar{\gamma} = \frac{1}{S^2} \sum_{i,j=1}^{S} \gamma_{ij} &\approx \bar{c}\bar{U}^2 \bar{N} b \int_{\mathbb{R}} \left(1 - bD^2\right) \sqrt{\frac{b}{2\pi}} e^{-\frac{b}{2}D^2} \frac{1}{\sqrt{4\pi V_{\bar{x}}}} e^{-\frac{D^2}{4V_{\bar{x}}}} dD \\ &= \bar{c}\bar{U}^2 \bar{N} b \sqrt{\frac{b}{2\pi(2V_{\bar{x}}b + 1)}} \left(\frac{1}{2V_{\bar{x}}b + 1}\right) = \frac{\bar{\alpha}\bar{N}b}{2V_{\bar{x}}b + 1}. \end{aligned} \tag{SM.111c}$$

Similarly, their variances can be approximated as

$$\begin{aligned}\mathrm{Var}(\alpha) = \overline{\alpha^2} - \bar{\alpha}^2 &\approx (V_c + \bar{c}^2)(V_U + \bar{U}^2)^2 \int_{\mathbb{R}} \frac{b}{2\pi} e^{-bD^2} \frac{1}{\sqrt{4\pi V_{\bar{x}}}} e^{-\frac{D^2}{4V_{\bar{x}}}} dD - \bar{\alpha}^2 \\ &= \frac{(V_c + \bar{c}^2)(V_U + \bar{U}^2)^2 b}{2\pi\sqrt{4V_{\bar{x}}b + 1}} - \bar{\alpha}^2, \qquad \text{(SM.112a)}\end{aligned}$$

$$\begin{aligned}\mathrm{Var}(\beta) = \overline{\beta^2} - \bar{\beta}^2 &\approx (V_c + \bar{c}^2)(V_U + \bar{U}^2)^2 (V_N + \bar{N}^2) b \int_{\mathbb{R}} D^2 \frac{b}{2\pi} e^{-bD^2} \frac{1}{\sqrt{4\pi V_{\bar{x}}}} e^{-\frac{D^2}{4V_{\bar{x}}}} dD \\ &= \frac{(V_c + \bar{c}^2)(V_U + \bar{U}^2)^2 (V_N + \bar{N}^2) b^2 V_{\bar{x}}}{\pi (4V_{\bar{x}}b + 1)^{3/2}}, \qquad \text{(SM.112b)}\end{aligned}$$

$$\begin{aligned}\mathrm{Var}(\gamma) = \overline{\gamma^2} - \bar{\gamma}^2 &\approx (V_c + \bar{c}^2)(V_U + \bar{U}^2)^2 (V_N + \bar{N}^2) b \int_{\mathbb{R}} \left(1 - bD^2\right)^2 \frac{b}{2\pi} e^{-bD^2} \frac{1}{\sqrt{4\pi V_{\bar{x}}}} e^{-\frac{D^2}{4V_{\bar{x}}}} dD - \bar{\gamma}^2 \\ &= \frac{(V_c + \bar{c}^2)(V_U + \bar{U}^2)^2 (V_N + \bar{N}^2) b^2}{\pi\sqrt{4V_{\bar{x}}b + 1}} \left(1 - 2b\left(\frac{V_{\bar{x}}}{4V_{\bar{x}}b + 1}\right) + 3b^2 \left(\frac{V_{\bar{x}}}{4V_{\bar{x}}b + 1}\right)^2\right) - \bar{\gamma}^2. \qquad \text{(SM.112c)}\end{aligned}$$

6.3.2 Mean and Variance of Absolute Values of Linear Selection Gradients

Since our above assumptions imply a certain degree of symmetry across the community, we find the average linear selection gradient is zero. Then, to extract information about the total quantity of linear selection occurring in the community, we consider the absolute values of linear selection gradients. Following the above assumptions we can express $|\beta_{ij}|$ as

$$|\beta_{ij}| = c_{ij} U_i U_j N_j b |D_{ij}| \sqrt{\frac{b}{2\pi}} e^{-\frac{b}{2} D_{ij}^2}. \qquad \text{(SM.113)}$$

The mean of $|\beta_{ij}|$ can then be approximated as

$$\begin{aligned}\overline{|\beta|} &\approx \bar{c}\bar{U}^2\bar{N}b \int_{\mathbb{R}} |D| \sqrt{\frac{b}{2\pi}} e^{-\frac{b}{2}D^2} \frac{1}{\sqrt{4\pi V_{\bar{x}}}} e^{-\frac{D^2}{4V_{\bar{x}}}} dD \\ &= \bar{c}\bar{U}^2\bar{N}b \sqrt{\frac{b}{2\pi(2V_{\bar{x}}b + 1)}} \int_{\mathbb{R}} |D| \sqrt{\frac{b}{2\pi(2V_{\bar{x}}b + 1)}} e^{-\frac{2V_{\bar{x}}b+1}{2b} D^2} dD. \qquad \text{(SM.114)}\end{aligned}$$

Computing the integral on the RHS is equivalent to computing the mean of the absolute value of a normally distributed random variable with mean zero and variance $\frac{b}{2V_{\bar{x}}b+1}$. It is well known that the absolute value $|Z|$ of a normally distributed random variable Z, itself taking mean zero and variance V_Z, has mean $\overline{|Z|} = \sqrt{2V_Z/\pi}$. Hence, we can use this information to compute

$$\overline{|\beta|} \approx \bar{c}\bar{U}^2\bar{N}b \sqrt{\frac{b}{2\pi(2V_{\bar{x}}b + 1)}} \sqrt{\frac{2b}{\pi(V_{\bar{x}}b + 1)}} = \frac{\bar{c}\bar{U}^2\bar{N}b^2}{\pi(2V_{\bar{x}}b + 1)}. \qquad \text{(SM.115)}$$

The variance $\mathrm{Var}(|\beta|)$ is a bit easier to calculate. In particular, we can approximate the variance of absolute values of β via

$$\mathrm{Var}(|\beta|) = \overline{|\beta|^2} - \overline{|\beta|}^2 = \overline{\beta^2} - \overline{|\beta|}^2 \approx \mathrm{Var}(\beta) - \overline{|\beta|}^2, \qquad \text{(SM.116)}$$

where we have capitalized on the result $\bar{\beta} \approx 0$.

6.3.3 Correlations Between Competition Coefficients and Selection Gradients

Following the above assumptions and notation, the covariance of competition coefficients α_{ij} and linear selection gradients β_{ij} can be approximated as

$$\mathrm{Cov}(\alpha,\beta) = \overline{\alpha\beta} - \bar{\alpha}\bar{\beta} \approx (V_c + \bar{c}^2)(V_U + \bar{U}^2)^2 \bar{N} b \int_{\mathbb{R}} D \frac{b}{2\pi} e^{-bD^2} \frac{1}{\sqrt{4\pi V_{\bar{x}}}} e^{-\frac{D^2}{4V_{\bar{x}}}} dD = 0. \tag{SM.117}$$

Again, this result follows from our assumptions on the distribution of model parameters across the community. To extract information about the covariance between competition coefficients and the magnitude of linear selection, we compute $\mathrm{Cov}(\alpha, |\beta|)$. This quantity can be approximated by

$$\begin{aligned}\mathrm{Cov}(\alpha,|\beta|) = \overline{\alpha|\beta|} - \bar{\alpha}\overline{|\beta|} &\approx (V_c + \bar{c}^2)(V_U + \bar{U}^2)^2 \bar{N} b \int_{\mathbb{R}} |D| \frac{b}{2\pi} e^{-bD^2} \frac{1}{\sqrt{4\pi V_{\bar{x}}}} e^{-\frac{D^2}{4V_{\bar{x}}}} dD - \bar{\alpha}\overline{|\beta|} \\ &= \frac{(V_c + \bar{c}^2)(V_U + \bar{U}^2)^2 \bar{N} b^2}{\pi(4V_{\bar{x}} b + 1)} \sqrt{\frac{V_{\bar{x}}}{\pi}} - \bar{\alpha}\overline{|\beta|},\end{aligned} \tag{SM.118}$$

where we have again made use of the properties of absolute values of normally distributed random variables.

The covariance between competition coefficients and quadratic selection gradients can be approximated by

$$\begin{aligned}\mathrm{Cov}(\alpha,\gamma) = \overline{\alpha\gamma} - \bar{\alpha}\bar{\gamma} &\approx (V_c + \bar{c}^2)(V_U + \bar{U}^2)^2 \bar{N} b \int_{\mathbb{R}} (1 - bD^2) \frac{b}{2\pi} e^{-bD^2} \frac{1}{\sqrt{4\pi V_{\bar{x}}}} e^{-\frac{D^2}{4V_{\bar{x}}}} dD - \bar{\alpha}\bar{\gamma} \\ &= \frac{(V_c + \bar{c}^2)(V_U + \bar{U}^2)^2 \bar{N} b^2 (2V_{\bar{x}} b + 1)}{2\pi(4V_{\bar{x}} b + 1)^{3/2}} - \bar{\alpha}\bar{\gamma}.\end{aligned} \tag{SM.119}$$

Finally, these approximations can be used to approximate the correlations between competition coefficients and selection gradients via

$$\mathrm{Corr}(\alpha,|\beta|) = \frac{\mathrm{Cov}(\alpha,|\beta|)}{\sqrt{\mathrm{Var}(\alpha)\mathrm{Var}(|\beta|)}}, \tag{SM.120a}$$

$$\mathrm{Corr}(\alpha,\gamma) = \frac{\mathrm{Cov}(\alpha,\gamma)}{\sqrt{\mathrm{Var}(\alpha)\mathrm{Var}(\gamma)}}. \tag{SM.120b}$$

References

N. Barton, A. Etheridge, and A. Véber. The infinitesimal model: Definition, derivation, and implications. Theoretical Population Biology, 118:50–73, Dec. 2017.

M. G. Bulmer. The effect of selection on genetic variability. The American Naturalist, 105(943):201–211, May 1971.

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

N. Champagnat, R. Ferrière, and S. Méléard. Unifying evolutionary dynamics: From individual stochastic processes to macroscopic models. Theoretical Population Biology, 69(3):297–321, May 2006.

G. Da Prato and J. Zabczyk. Stochastic Equations in Infinite Dimensions. Cambridge University Press, 2014.

A. M. Etheridge. An Introduction to Superprocesses. American Mathematical Society, aug 2000.

L. C. Evans. Partial Differential Equations: Second Edition. American Mathematical Society, 2010.

S. J. Farlow. Partial Differential Equations for Scientists and Engineers. Dover, sep 1993.

M. Kimura and J. F. Crow. Effect of overall phenotypic selection on genetic change at individual loci. Proceedings of the National Academy of Sciences, 75(12):6168–6171, 1978.

R. Lande. The maintenance of genetic variability by mutation in a polygenic character with linked loci. Genetical Research, 26(3):221–235, Dec. 1975.

R. Lande and S. J. Arnold. The measurement of selection on correlated characters. Evolution, 37(6):1210, Nov. 1983.

M. Méléard and S. Roelly. Interacting branching measure processes. Stochastic Partial Differential Equations and Applications (G. Da Prato and L. Tubaro, eds.), pages 246–256, 1992.

M. Méléard and S. Roelly. Interacting measure branching processes. some bounds for the support. Stochastics and Stochastic Reports, 44(1-2):103–121, Sept. 1993.

M. Turelli. Commentary: Fisher's infinitesimal model: A story for the ages. Theoretical Population Biology, 118:46–49, Dec. 2017.

V. Volpert. Elliptic Partial Differential Equations. Springer Basel, 2014.

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

J. B. Walsh. An introduction to stochastic partial differential equations. In Lecture Notes in Mathematics, pages 265–439. Springer Berlin Heidelberg, 1986.

S. Zheng. Nonlinear evolution equations. Chapman & Hall/CRC Press, Boca Raton, Fla, 2004.