



**HAL**  
open science

# A Theoretical Framework for Trait-Based Eco-Evolutionary Dynamics: Population Structure, Intraspecific Variation, and Community Assembly

Jonas Wickman, Thomas Koffel, Christopher A Klausmeier

► **To cite this version:**

Jonas Wickman, Thomas Koffel, Christopher A Klausmeier. A Theoretical Framework for Trait-Based Eco-Evolutionary Dynamics: Population Structure, Intraspecific Variation, and Community Assembly. *The American Naturalist*, 2023, 201 (4), pp.501-522. 10.1086/723406 . hal-04607714

**HAL Id: hal-04607714**

**<https://hal.science/hal-04607714v1>**

Submitted on 18 Jun 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A theoretical framework for trait-based eco-evolutionary dynamics: population structure, intraspecific variation, and community assembly

*Manuscript elements:* Figure 1, figure 2, figure 3, table 1, table 2, online appendices A, B, C, D, and Supplementary S1 (separate file).

*Keywords:* Moment methods, evolutionarily stable communities, adaptive dynamics, quantitative genetics, intraspecific trait variation (ITV)

*Manuscript type:* Article.

Prepared using the suggested L<sup>A</sup>T<sub>E</sub>X template for *Am. Nat.*

## Abstract

How is trait diversity in a community apportioned between and within co-evolving species? Disruptive selection may result in either a few species with large intraspecific trait variation (ITV) or many species with different mean traits but little ITV. Similar questions arise in spatially structured communities: heterogeneous environments could result in either a few species that exhibit local adaptation or many species with different mean traits but little local adaptation. To date, theory has been well-equipped to either include ITV or to dynamically determine the number of coexisting species, but not both. Here, we devise a theoretical framework that combines these facets, and apply it to the above questions of how trait variation is apportioned within and between species in unstructured and structured populations, using two simple models of Lotka-Volterra competition. For unstructured communities, we find that as the breadth of resources increases, ITV goes from being unimportant to crucial for characterizing the community. For spatially structured communities on two patches, we find that either no local adaptation, symmetric local adaptation, or asymmetric local adaptation prevails depending on how much the patches differ. Our framework provides a general approach to incorporate ITV in models of eco-evolutionary community assembly.

## Introduction

18 In recent years, the need to understand intraspecific variation of functional traits has been in-  
creasingly recognized as important for understanding the functioning of ecological communities  
(Albert et al., 2010; Violle et al., 2012). For example, intraspecific trait variation (ITV) has been  
21 shown to be important for detecting niche differentiation and environmental filtering (Paine  
et al., 2011) and for quantifying how pairwise species interactions affect the total biomass in ex-  
perimentally assembled communities (Kraft et al., 2014). ITV in functional traits has also been  
24 shown to often be substantial compared to variation between species (Siefert et al., 2015).

If species have heritable variation in traits that affect fitness, natural selection will act on that  
variation, changing the distribution of traits in the community over time. In the long run, a  
27 community with a fixed number of species will approach an attractor for the eco-evolutionary  
dynamics, in the simplest case an equilibrium. However, such an eco-evolutionary equilibrium  
could still be invisable by other species from outside the community with other trait distributions,  
30 and one of the resident species might be under selective forces that results in an ‘evolutionary  
branching’, where a species ends up at fitness minimum as a result of directional selection and  
consequently splits into two (Geritz et al., 1998). Thus, in the longer term, the species richness of a  
33 community at an eco-evolutionary equilibrium might not be stable. Eventually, a community that  
is in equilibrium and is stable to further addition of species may be reached. Such a community  
has been called an evolutionarily stable community (ESC); since such a community can persist  
36 over long time scales and serves as an attractor for eco-evolutionary dynamics, ESCs can serve  
as important model communities (Edwards et al., 2018).

If both the number of species and the trait variation within each species are driven by eco-  
39 evolutionary dynamics, this then raises questions regarding how trait variation is apportioned  
between and within species in ESCs. For example, what happens when selection goes from  
being stabilizing to being disruptive? In general, this can depend on many factors—including  
42 genetic architecture, phenotypic plasticity, and mating patterns (Rueffler et al., 2006)—but even

in simplified, purely phenotypic trait-based modeling, theory gives different answers depending on the basic assumptions the theory uses. When the theory includes ITV but fixes the number of species, weaker stabilizing selection will lead to more ITV (Kimura, 1965; Bürger, 1986), whereas  
45 in theory where the number of species is dynamic, but with no ITV, weaker stabilizing selection will lead to more species coexisting (Levins, 1962; MacArthur and Levins, 1964; Dieckmann and Doebeli, 1999). If one could incorporate both facets, the actual ESC would likely be some combination of the two, but whether the ESC will comprise a few species with great ITV or many species with little ITV can not currently be determined as theoretical approaches for assembling  
48 ESCs that take ITV into account are not yet developed.

This tension between ITV and species coexistence can also play out in spatially structured communities. Given sufficient heritable trait variation, spatially varying selection can lead species' traits to vary across space, resulting in different degrees of local adaptation (Kirkpatrick and Barton, 1997; Bruggeman, 2009; Norberg et al., 2012; Le Gland et al., 2020). However, when all individuals of a species are assumed to be identical, local adaptation is not possible, but variable spatial conditions can promote species coexistence (Troost et al., 2005; Débarre and Gandon, 2010; Fortelius et al., 2015; Wickman et al., 2017, 2019). Thus, to cover trait space on the regional scale, an ESC might comprise a few species with substantial local adaptation or many species  
54 with little local adaptation.

The current state of eco-evolutionary theory is not well equipped to deal with these questions. On one hand, quantitative genetics (Lande, 1979; Lande and Arnold, 1983) and other moment-equation-based frameworks (Wirtz and Eckhardt, 1996; Norberg et al., 2001; Savage et al., 2007; Sasaki and Dieckmann, 2011; Merico et al., 2014) readily incorporate ITV into the models. However, even the versions of these models that can treat multiple species (Sasaki and Dieckmann, 2011; Débarre et al., 2014) assume that the number of species is fixed rather than a dynamic outcome of eco-evolutionary community assembly. On the other hand, using a different modeling approach called adaptive dynamics (Metz et al., 1992; Dieckmann and Law, 1996; Geritz et al., 1998; Dercole and Rinaldi, 2008; Brännström et al., 2013) one can easily determine whether  
63  
66  
69

a community is closed to evolutionary branching and/or invasion by other species. Thus, using adaptive dynamics, one can assemble a community that is not only at an eco-evolutionary equilibrium, but which is also stable against further diversification into more species by invasion or evolutionary branching. However, by construction, adaptive dynamics assumes that all individuals within a species are identical, so it cannot take ITV into account. While model-specific studies in structured communities that combine species assembly with ITV have been carried out (Débarre et al., 2013), and the moment equations for a fixed number of species for a general category of class-structured communities has recently been developed (Lion et al., 2022), a framework for systematically assembling ESCs that take ITV into account does not exist.

In this paper we synthesize many recent advances in describing eco-evolutionary dynamics through moment equations taking the total density, mean traits, and trait variances into account, and adapt the community-assembly capabilities of adaptive dynamics to work with these moment equations. This yields a general framework for eco-evolutionary community assembly for class-structured communities that allows us to assemble ESCs that take ITV into account. The framework can handle complex models that include for example multiple traits and external resources, but in the main text we will focus on resolving the twin tensions described above regarding whether trait diversity will be apportioned within or between species, and within a local community or across a spatial gradient. We illustrate the framework using two simple models of Lotka-Volterra competition. First, we describe the framework in the simplest possible setting without population (spatial) structure, and use the first Lotka-Volterra model to investigate the simultaneous increase in ITV and number of species that result from weaker stabilizing selection. We then extend the framework to class-structured populations, which we illustrate with a two-patch spatially structured Lotka-Volterra model to investigate the tension between local adaptation and species coexistence as the two patches become increasingly different. These analyses show that rich patterns of intra- and interspecific trait variation emerge even in these simple models, and thus demonstrate the utility we believe this framework has for furthering our understanding of the interaction between eco-evolutionary dynamics and individual variation in

traits.

## Eco-evolutionary dynamics and community assembly in unstructured communities

99

In this section we will introduce the trait-space equations that underlie our approach, and describe how these equations can be approximated by moment equations that track the total densities, mean traits, and trait variances of several species. We will then explain how these equations can be used in conjunction with a community-assembly approach to determine the evolutionarily stable community (ESC). To explore the tension between intraspecific trait variation and species coexistence, we employ a simple Lotka-Volterra competition model.

102

105

### *Trait-space equations for unstructured communities*

We consider a single trait value  $x$  that encodes some property of an organism such as body mass or resource-uptake ability. The entire community can then be described by the trait-density distribution  $v(x)$ , which tells us how many individuals there are with trait  $x$ ; to be more precise,  $\int_a^b v(x)dx$  describes the total density of individuals that has a trait value in the interval  $a \leq x \leq b$ . We assume that the density of individuals grows due to births with per capita rate  $b(x, v)$  and declines due to mortality with per capita rate  $m(x, v)$ , where both birth and death rates can depend both on the trait of an individual,  $x$ , as well as the entire trait-density distribution in which the individual exists,  $v$ . We will assume that reproduction is clonal but that parents do not give birth to exact copies of themselves, but instead produce offspring with normally distributed traits. Thus, a parent with trait  $y$  will produce offspring with trait  $x$  according to

108

111

114

117

$$\mathcal{N}(x, y, M) := \frac{1}{\sqrt{2\pi M}} \exp\left(-\frac{(x-y)^2}{2M}\right), \quad (1)$$

where  $M$  is the mutation variance. We will use the notation  $\mathcal{N}(x, \mu, V)$  throughout for the probability-density function of a normal distribution with mean  $\mu$  and variance  $V$  evaluated at

120  $x$ . As in adaptive dynamics and trait-diffusion approaches (Merico et al., 2014; Le Gland et al.,  
 2020; Nordbotten et al., 2020), we do not specify the genetic makeup that would result in these  
 dynamics for phenotypic traits, but simply assume that new heritable variation is generated  
 123 through mutations that are associated with some ecological process, such as births.

We can now write down how the trait-density distribution will change over time, which is  
 given by the equation

$$126 \quad \frac{dv(x)}{dt} = \int_{-\infty}^{\infty} b(y, v)v(y)\mathcal{N}(x, y, M)dy - m(x, v)v(x), \quad (2)$$

where  $\mathcal{N}(x, y, M)$  is the normal mutation kernel with variance  $M$ , describing how different off-  
 spring are from their parents on average. We will refer to these types of equations as *trait-space*  
 129 *equations*. These models are similar to classic work in quantitative genetics (Kimura, 1965; Bürger,  
 1986), and models like this have been shown to be the limit of certain individual-based models  
 in the limit of large populations (Champagnat et al., 2006).

### 132 *Trait-space equations for the unstructured Lotka-Volterra model*

To take a concrete example, we will consider a simple model of Lotka-Volterra competition  
 (Lotka, 1925; Volterra, 1928). The model is similar to Lotka-Volterra models in many other recent  
 135 studies using for example ‘oligomorphic dynamics’ (Sasaki and Dieckmann, 2011), quantitative  
 genetics (Barabás et al., 2022), and adaptive dynamics (Ranjan and Klausmeier, 2022). For this  
 model, we assume that individuals experience a net density-independent per capita growth  $r(x)$   
 138 in the absence of other individuals due to the environment, which we split into a birth rate  
 $b(x)$  and background death rate  $\mu(x)$ . Additionally, all individuals experience extra density-  
 dependent mortality due to competition with all other individuals, at a per capita rate  $a(x, v)$ , in  
 141 a way such that competition is most intense between individuals with similar traits. Comparing  
 with Eq. 2, we thus have that  $m(x, v) = \mu(x) + a(x, v)$ , and the trait-space equations are given by

$$144 \quad \frac{dv(x)}{dt} = \int_{-\infty}^{\infty} b(y)v(y)\mathcal{N}(y, x, M)dy - \mu(x)v(x) - a(x, v)v(x), \quad (3a)$$



$$b(x) = r_0, \quad \mu(x) = \frac{x^2}{V_r}, \quad r(x) = b(x) - \mu(x) = r_0 - \frac{x^2}{V_r}, \quad (3b)$$

$$a(x, v) = \int_{-\infty}^{\infty} \alpha(y, x) v(y) dy, \quad \alpha(y, x) = \exp\left(-\frac{(y-x)^2}{2V_c}\right). \quad (3c)$$

147 Here,  $V_r$  is the environmental variance and  $V_c$  is the competition variance, which is a measure of  
 how broad competition is in trait space, with larger  $V_c$  increasing competition between individ-  
 uals of different traits. For  $V_c \rightarrow 0$ , individuals will compete only with other individuals with  
 150 exactly the same trait, and for  $V_c \rightarrow \infty$  individuals will compete equally with individuals with  
 any trait. Trait-matching competition of this type can be motivated by, for example, that birds  
 with similarly-sized beaks will compete more strongly since they compete for similarly-sized  
 153 seeds (MacArthur, 1972).

In general, the generic trait-space equations (Eq. 2) will not be analytically tractable; even  
 the equilibrium of Eq. 2 will be hard to ascertain, even for very simple examples (e.g., Kimura,  
 156 1965; Bürger, 1986). Numerically too, discretizing trait space and solving Eq. 2 often requires  
 significant computational power. For the more complicated models we will study later in this  
 manuscript especially, the feasibility of numerical explorations of the equations is highly limited.  
 159 Instead, we will study approximations of the trait-space equations where the moments of the  
 trait-density distribution are tracked. Such moment equations have the additional benefit that  
 they are often more ecologically interpretable than the trait-space equations.

### 162 *Moment equations for unstructured communities*

Here, we will follow the general approach of previous moment-based frameworks (Slatkin, 1980;  
 Wirtz and Eckhardt, 1996; Norberg et al., 2001; Bruggeman, 2009; Sasaki and Dieckmann, 2011;  
 165 Tirok et al., 2011; Merico et al., 2014; Débarre et al., 2014; Nordbotten et al., 2020) to derive  
 approximate equations that track the total density, mean trait, and trait variance of populations.  
 These approaches, and ours, all result in similar equations, but as the details differ we will here  
 168 detail the precise assumptions we make to derive these equations. We will assume that the  
 trait-density distribution  $v(x)$  can be decomposed into a set of  $S$  ‘species’, where each species

represents a peak in the trait-density distribution so that

$$v(x) = \sum_{i=1}^S v_i(x). \quad (4)$$

We will then assume that each such species can be approximated with a normal distribution so that

$$v_i(x) \approx u_i \mathcal{N}(x, \bar{x}_i, V_i), \quad (5)$$

where  $u_i = \int_{-\infty}^{\infty} v_i(x) dx$  is the total density of species  $i$ ,  $\bar{x}_i = (1/u_i) \int_{-\infty}^{\infty} x v_i(x) dx$  is the mean trait of species  $i$ , and  $V_i = (1/u_i) \int_{-\infty}^{\infty} (x - \bar{x}_i)^2 v_i(x) dx$  is the trait variance of species  $i$ . We will

use the notation

$$\tilde{v}(x) := \sum_{i=1}^S u_i \mathcal{N}(x, \bar{x}_i, V_i) \approx v(x) \quad (6)$$

for the approximate trait-density distribution. In words, the approximate community trait-

density distribution  $\tilde{v}(x)$  is the sum of all approximate trait-density distributions for each species,  $u_i \mathcal{N}(x, \bar{x}_i, V_i)$ . We will assume that each species is reproductively isolated so that individuals

born in species  $i$  end up as species  $i$  individuals. Finally, for the birth rate  $b(x, v)$  and mortality

rate  $m(x, v)$  we will introduce notation that describes the population-level per capita rates for a

normally-distributed population with mean trait  $\bar{x}$  and trait variance  $V$  in the environment set

by the resident community with trait-density  $\tilde{v}$ . These population-level rates are given by the

Gaussian integrals

$$\hat{b}(\bar{x}, V, \tilde{v}) := \int_{-\infty}^{\infty} b(x, \tilde{v}) \mathcal{N}(x, \bar{x}, V) dx, \quad (7a)$$

$$\hat{m}(\bar{x}, V, \tilde{v}) := \int_{-\infty}^{\infty} m(x, \tilde{v}) \mathcal{N}(x, \bar{x}, V) dx. \quad (7b)$$

Given these assumptions, without further approximation, we can derive the following equa-

tions for the total densities, mean traits, and trait variances for each species  $i = 1, \dots, S$ , which are

given by (see Appendix A for the derivation):

$$\frac{du_i}{dt} = \underbrace{(\hat{b}(\bar{x}_i, V_i, \tilde{v}) - \hat{m}(\bar{x}_i, V_i, \tilde{v}))}_{\text{population-level per capita net growth}} u_i, \quad (8a)$$

$$\frac{d\bar{x}_i}{dt} = V_i \left( \underbrace{\frac{\partial \hat{b}}{\partial \bar{x}}(\bar{x}_i, V_i, \bar{v}) - \frac{\partial \hat{m}}{\partial \bar{x}}(\bar{x}_i, V_i, \bar{v})}_{\text{directional selection}} \right), \quad (8b)$$

$$\frac{dV_i}{dt} = V_i^2 \left( \underbrace{\frac{\partial^2 \hat{b}}{\partial \bar{x}^2}(\bar{x}_i, V_i, \bar{v}) - \frac{\partial^2 \hat{m}}{\partial \bar{x}^2}(\bar{x}_i, V_i, \bar{v})}_{\text{stabilizing/disruptive selection}} \right) + \underbrace{\hat{b}(\bar{x}_i, V_i, \bar{v})M}_{\text{mutation}}. \quad (8c)$$

195 We will refer to these equations as *moment equations*. Equation 8a describes how the total density  $u_i$  of species  $i$  increases in response to its population-level birth rate  $\hat{b}(\bar{x}, V, \bar{v})$  evaluated at its mean trait  $\bar{x}_i$  and trait variance  $V_i$ , and how it decreases in response to the population-level

198 mortality rate  $\hat{m}(\bar{x}, V, \bar{v})$ . Note that the birth and mortality rates may depend on the approximate trait-density distribution  $\bar{v}$  of the entire community, not just conspecifics. Equation 8b describes how the mean trait  $\bar{x}_i$  of species  $i$  responds to directional selection. The mean trait will move

201 along the selection gradient induced by births and deaths towards mean traits that yield higher per capita net growth, similar to the gradient dynamics of the canonical equation of adaptive dynamics (Dieckmann and Law, 1996; Champagnat, 2003) or in quantitative genetics (Lande

204 and Arnold, 1983). The speed at which this directional selection acts depends on the selection gradient multiplied by the level of trait variance  $V_i$  in species  $i$ , but note that here the selection gradients themselves,  $\partial \hat{b} / \partial \bar{x}$  and  $\partial \hat{m} / \partial \bar{x}$ , can also depend on the trait variance (due to Eqs. 7).

207 Equation 8c describes how the trait variance  $V_i$  of species  $i$  responds to stabilizing/disruptive selection and mutations. Selection is stabilizing—resulting in a reduction in trait variance over time—if the sign of the sum of the second derivatives is negative and selection is disruptive—

210 resulting in an increase in trait variance over time—if the sign is positive. The second term corresponds to an increase in trait variance over time due to mutations, which depends on the population-level per capita birth rate  $\hat{b}$  as well as the mutation kernel variance  $M$ .

### 213 *Moment equations for the unstructured Lotka-Volterra model*

To get the moment equations for our Lotka-Volterra model, we first need to compute the population-level rates which can be calculated to be (see Appendix C)

$$216 \quad \hat{b}(\bar{x}, V) = r_0, \quad (9a)$$

$$\hat{r}(\bar{x}, V) = \hat{b}(\bar{x}, V) - \hat{\mu}(\bar{x}, V) = r_0 - \frac{\bar{x}^2 + V}{V_r}, \quad (9b)$$

$$\hat{a}(\bar{x}, V, \bar{v}) = \sqrt{2\pi V_c} \sum_{j=1}^S u_j \mathcal{N}(\bar{x}, \bar{x}_j, V + V_j + V_c). \quad (9c)$$

219 Having done this, we can plug these expressions into the general moment equations (Eqs. 8),  
which yields the equations

$$\frac{du_i}{dt} = \left[ \underbrace{\left( r_0 - \frac{\bar{x}_i^2 + V_i}{V_r} \right)}_{\hat{r}(\bar{x}_i, V_i) = \hat{b}(\bar{x}_i, V_i) - \hat{\mu}(\bar{x}_i, V_i)} - \underbrace{\sqrt{2\pi V_c} \sum_{j=1}^S u_j \mathcal{N}(\bar{x}_i, \bar{x}_j, V_i + V_j + V_c)}_{\hat{a}(\bar{x}_i, V_i, \bar{v})} \right] u_i, \quad (10a)$$

$$222 \quad \frac{d\bar{x}_i}{dt} = V_i \left[ \underbrace{-\frac{2\bar{x}_i}{V_r}}_{\frac{\partial \hat{r}}{\partial \bar{x}}(\bar{x}_i, V_i)} + \underbrace{\sqrt{2\pi V_c} \sum_{j=1}^S u_j \frac{\bar{x}_i - \bar{x}_j}{V_i + V_j + V_c} \mathcal{N}(\bar{x}_i, \bar{x}_j, V_i + V_j + V_c)}_{-\frac{\partial \hat{a}}{\partial \bar{x}}(\bar{x}_i, V_i, \bar{v})} \right], \quad (10b)$$

$$\begin{aligned} \frac{dV_i}{dt} = & V_i^2 \left[ \underbrace{-\frac{2}{V_r}}_{\frac{\partial^2 \hat{r}}{\partial \bar{x}^2}(\bar{x}_i, V_i)} + \underbrace{\sqrt{2\pi V_c} \sum_{j=1}^S u_j \frac{V_i + V_j + V_c - (\bar{x}_i - \bar{x}_j)^2}{(V_i + V_j + V_c)^2} \mathcal{N}(\bar{x}_i, \bar{x}_j, V_i + V_j + V_c)}_{-\frac{\partial^2 \hat{a}}{\partial \bar{x}^2}(\bar{x}_i, V_i, \bar{v})} \right] \\ & + \underbrace{r_0 M}_{\hat{b}(\bar{x}_i, V_i)}. \end{aligned} \quad (10c)$$

225 The growth rate of the total density  $u_i$  of species  $i$  (Eq. 10a) depends on the environmental growth  
rate, which is determined by the species' distance in trait mean  $\bar{x}_i$  from the optimal trait  $\bar{x}_i = 0$   
and decreased due to the species' trait variance  $V_i$ . The species then suffers additional mortality  
228 from intra- and inter-specific competition, which depends on the mean traits and trait variances  
of both the focal species  $i$  and its competitor  $j$ . Interspecific competition is most intense between  
species whose mean traits  $\bar{x}_i$  and  $\bar{x}_j$  are close, and is mediated by the trait variance of both species  
231  $i$  and  $j$  as well as the competition variance  $V_c$ .

The rate of change of the mean trait  $\bar{x}_i$  of species  $i$  (Eq. 10b) depends on the directional  
selection induced by both the environmental growth rate and competition. The first term, which  
234 is the selection induced by the environment, always exerts selective pressure towards the optimal  
trait  $\bar{x}_i = 0$ . The second term, which is the selection induced by competition, always exerts  
selective pressure for the mean traits of any species to separate from one another.

237 The rate of change of the trait variance  $V_i$  of species  $i$  (Eq. 10c) depends on stabilizing/disruptive  
selection induced by the environment and competition. The first term, due to the environment, is  
always stabilizing and acts to reduce trait variance over time. This is because the quadratic envi-  
240 ronmental growth rate universally disfavors trait variance. The second term, due to competition  
is stabilizing if  $(\bar{x}_i - \bar{x}_j)^2 > V_i + V_j + V_c$  and disruptive if the opposite is true. For intraspecific  
competition ( $i = j$ ) the difference between the means is zero, so intraspecific competition is al-  
243 ways disruptive. This is because a species with more trait variance will have its individuals more  
spread out in trait space and consequently suffer less intraspecific competition. If two different  
species are far enough apart, selection will act stabilizingly on both species to reduce their over-  
246 lap in trait space by reducing the trait variance of both species. The final term is the mutation  
term, which always acts to increase trait variance over time.

Together, equations 10 describe the dynamics of the first three moments (zeroth: total density,  
249 first: mean trait, second: trait variance) of each species in a community of  $S$  species. As we  
have made a normal approximation for the trait-density distribution for each species, no higher  
moments have to be tracked. We further discuss our choice of normal approximation in the  
252 discussion section.

### *Eco-evolutionary community assembly for unstructured communities*

So far we have derived how to treat a fixed number of species,  $S$ , and their eco-evolutionary  
255 dynamics, including the effects on the intraspecific trait variance of the species, but this then  
raises the question: how many species can coexist? For our Lotka-Volterra model (Eqs. 10) we  
can see an example of how there exists an intrinsic tension between large intraspecific variance  
and the number of species: Intraspecific competition will increase trait variance and interspecific  
258 competition can reduce it, thus, under weak stabilizing selection (e.g., small  $V_c$  or large  $V_r$ ) will  
we get many species with small trait variance or a few—or even just one—species with large trait  
261 variance? Likely the answer is some combination of the two, but without some method for being  
able to ascertain whether a community where intraspecific variation is taken into account is open

to the addition of more species, it is not possible to determine the relative importance of these  
264 two effects.

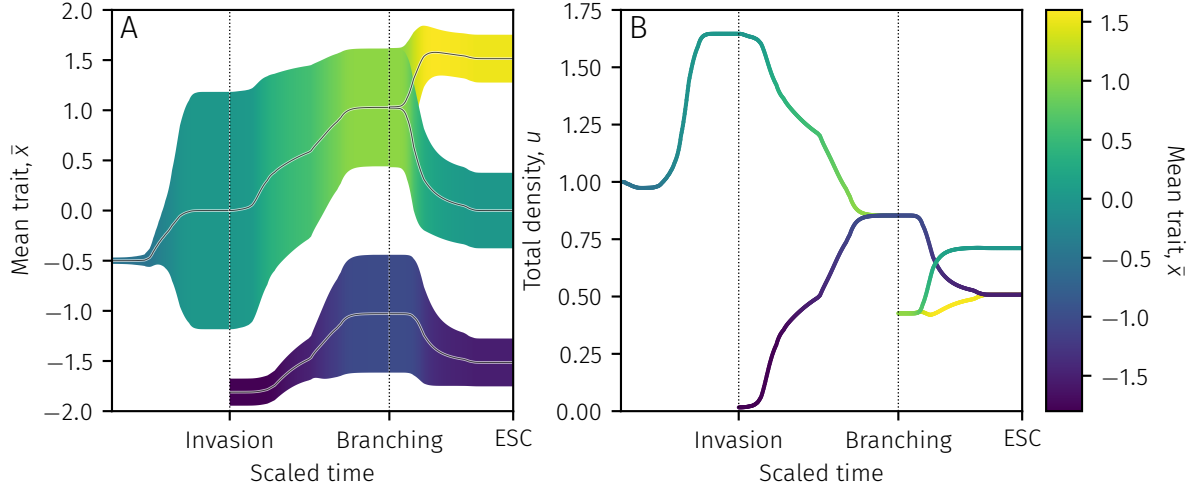
To be able to answer these questions we need to adapt the eco-evolutionary community-  
assembly methods available in adaptive dynamics. Specifically, we devise a method for determin-  
267 ing whether a species in a community can undergo evolutionary branching when in equilibrium  
and a method for determining whether a community is invasible by any other rare species.

### *Evolutionary branchings*

270 In adaptive dynamics, evolutionary branching occurs when a species has evolved a trait in accor-  
dance with directional selection until directional selection ceases and the species finds itself at a  
fitness minimum causing the species to split in two (Geritz et al., 1998). This procedure cannot  
273 be directly translated to our moment-equation framework due to the fact that we include trait  
variance. Instead, after Eqs. 8 have reached an equilibrium for  $S$  species we virtually split each  
species into two identical copies and examine the linear stability of the resulting  $S + 1$  species  
276 equilibrium. If such a split-species equilibrium is unstable, we say that the system has under-  
gone an evolutionary branching and keep the split-species pair. The details of the procedure are  
available in Appendix B. Figure 1 depicts an example of a branching in the Lotka-Volterra model.

### *Invasion analysis*

Even if no evolutionary branchings are possible, the community might not be closed to invasion  
by types that are farther away in trait space. We will here employ a scheme for carrying out  
282 a global invasion analysis inspired by Kremer and Klausmeier (2013). To determine whether a  
community is closed to invasion, we introduce a rare invading population with mean trait  $\bar{x}^{\text{inv}}$   
and trait variance  $V^{\text{inv}}$  into a resident community whose species are all in equilibrium. Let  $\tilde{v}^{\text{res}}$   
285 denote the sum-of-normals trait-density distribution of the residents. As we assume that the  
invader will initially remain rare, we do not need to track its total density, and the equations for



**Figure 1: Community assembly through invasion and branching in the unstructured Lotka-Volterra model.** The panels depict how mean traits  $\bar{x}_i$  and trait variances  $V_i$  (Panel A, one standard deviation is depicted as a filled area) and the total densities  $u_i$  (Panel B) evolve over time during a community-assembly procedure. To start with, a single species with initial moments  $u = 1$ ,  $\bar{x} = -0.5$ , and  $V = 0.001$  is introduced. Selection from the environment and competition leads to changes in the density, mean trait ('ecological character displacement'; Slatkin, 1980), and variance according to Eqs. 10, until it reaches an eco-evolutionary equilibrium (first dotted line). Then, the system is checked for whether any branchings or invasions are possible. In this case, no branching is possible, but an invader with positive invasion fitness is found and introduced. After that, Eqs. 10 are once again run for the two-species community until it reaches an eco-evolutionary equilibrium (second dotted line). Here, a branching is detected and the top species is split into two new species with slightly different mean traits. We then once again solve Eqs. 10 until equilibrium, during which the branching species diverge. At the final time point the system is once again at an eco-evolutionary equilibrium, and no branchings or invasions are possible meaning that we have reached a three-species eco-evolutionarily stable community (ESC). To show all pertinent events, time has been rescaled nonuniformly. **Parameter values:**  $V_r = 9.0$ ,  $M = 10^{-4}$ , and other parameters are as in Table 1.

its mean trait and trait variance are given by

$$\frac{d\bar{x}^{\text{inv}}}{dt} = V^{\text{inv}} \left( \frac{\partial \hat{b}}{\partial \bar{x}}(\bar{x}^{\text{inv}}, V^{\text{inv}}, \bar{v}^{\text{res}}) - \frac{\partial \hat{m}}{\partial \bar{x}}(\bar{x}^{\text{inv}}, V^{\text{inv}}, \bar{v}^{\text{res}}) \right), \quad (11a)$$

$$\frac{dV^{\text{inv}}}{dt} = (V^{\text{inv}})^2 \left( \frac{\partial^2 \hat{b}}{\partial \bar{x}^2}(\bar{x}^{\text{inv}}, V^{\text{inv}}, \bar{v}^{\text{res}}) - \frac{\partial^2 \hat{m}}{\partial \bar{x}^2}(\bar{x}^{\text{inv}}, V^{\text{inv}}, \bar{v}^{\text{res}}) \right) + \hat{b}(\bar{x}^{\text{inv}}, V^{\text{inv}}, \bar{v}^{\text{res}})M, \quad (11b)$$

where we note that the invader rates depend only on the resident densities, because we assume it to be so rare that its effect on itself is negligible.

As long as the invader starts out sufficiently rare, we can assume that its mean trait and trait variance will settle onto an attractor, in the simplest case an equilibrium, where its exponential

294 growth rate can be determined (c.f., Lin et al., 2020). Since the invader equations are nonlinear,  
several attractors may possibly exist. We will here assume equilibrium dynamics, and we can  
thus assume that a rare invader will eventually settle onto one of several equilibria. Assume  
297 that there are  $N$  such equilibria and denote invaders' means and variances by  $\bar{x}^{\text{inv},n}$  and  $V^{\text{inv},n}$   
respectively for equilibrium  $n = 1, \dots, N$ . Once an invader's trait distribution has reached an  
equilibrium mean and variance, we can calculate its per capita growth  $\lambda_n$  as

$$300 \quad \lambda_n = \hat{b}(\bar{x}^{\text{inv},n}, V^{\text{inv},n}, \bar{\sigma}^{\text{res}}) - \hat{m}(\bar{x}^{\text{inv},n}, V^{\text{inv},n}, \bar{\sigma}^{\text{res}}). \quad (12)$$

If  $\lambda_n > 0$ , the per capita growth rate of the invader will be positive in the environment set by the  
residents and the invader can successfully invade, and if  $\lambda_n \leq 0$ , the invader will not be able to  
303 invade. Note that all resident species will show up as neutrally stable ( $\lambda_n = 0$ ) equilibria for the  
invader, though not necessarily attractors.

In Fig. 1 we depict an invasion event in the Lotka-Volterra model. For the single resident  
306 species depicted at equilibrium at the dotted line marked "invasion" we can find two invader  
equilibria given by

$$\bar{x}^{\text{inv},1} \approx -1.81, \quad V^{\text{inv},1} \approx 0.0182, \quad \lambda_1 \approx 0.0962, \quad (13a)$$

$$309 \quad \bar{x}^{\text{inv},2} \approx 1.81, \quad V^{\text{inv},2} \approx 0.0182, \quad \lambda_2 \approx 0.0962. \quad (13b)$$

The Lotka-Volterra model is symmetric in trait space, and thus, with a single resident with  $\bar{x} = 0$ ,  
we find two invaders equidistant from the middle, both with positive invasion fitness. As in  
312 adaptive dynamics, we assume that only one invasion event is allowed to happen at once, and so  
for this example, invader one was chosen at random to invade the community.

In principle, this procedure gives us an invasion criterion for a community. If there exists  
315 any invader equilibrium with  $\lambda_n > 0$  we can add a species to the community at a low density  
with mean trait  $\bar{x}^{\text{inv},n}$  and trait variance  $V^{\text{inv},n}$ . However, in practice, exhaustively proving that  
all such equilibria have been found is in general not feasible, so we will use a heuristic. For  
318 a given resident community we compute the invasion fitness for all invaders across a range of



mean traits and with zero trait variance. This is equivalent to computing the invasion-fitness landscape in adaptive dynamics. We then find all the local maxima of this adaptive-dynamics fitness landscape and use those as the initial conditions for the invader equations (Eqs. 11), and solve the equations until they reach equilibrium. We can then compute the invasion fitness for all these equilibria to determine whether the community is closed to invasion. While theoretically not exhaustive, we have found this heuristic to work very well in practice, as the zero-variance invaders serve as good first approximations to where positive invasion fitness might be available.

### *Assembly protocol*

We can now use the moment equations, the branching condition, and the invasion process to assemble an eco-evolutionarily stable community. To assemble a globally eco-evolutionarily stable community we start with an arbitrarily specified community of  $S$  species with total densities  $u_i^0$ , mean traits  $\bar{x}_i^0$  and trait variances  $V_i^0$ . We then proceed along the following steps.

1. We let the community evolve according to Eqs. 8 until it reaches equilibrium.
2. We check each species for evolutionary branching. In case of a branching we split the species undergoing branching into two new species and then return to step 1, letting the new community of  $S + 1$  species evolve according to Eqs. 8.
3. In case the moment equations for the resident community reach equilibrium and there are no branchings, we use our invasion scheme to see if any invader with positive invasion fitness exists. If an invader with positive invasion fitness is found, it is added to the community with a small density, and we return to step 1, and let the new community evolve once again according to Eqs. 8.

We continue going through these steps until we have reached a community where Eqs. 8 are in equilibrium and no more invaders with positive invasion fitness can be found. This community is thus eco-evolutionarily stable. In Fig. 1 we depict an example of this process for the Lotka-Volterra model. Note that at the first dotted line in Fig. 1, we first calculate the branching criterion, but a branching is not possible for that configuration, and we thus move on to step 3.

345 to check for invasions.

The purpose of our algorithm is to find the final evolutionarily stable community (ESC). Our assembly process cannot accurately capture the temporal dynamics of the trait-space equations or other assembly processes, as we integrate the moment equations to equilibrium between each branching or invasion event. Moreover, we have necessarily made some choices regarding the order in which we carry out our assembly steps. We have chosen to check for evolutionary branchings before invasions, as the existence of a branching implies a successful invasion but not vice versa. When a single final ESC exists, the order will not matter much as the assembly process will converge on this ESC. Our Lotka-Volterra example models in this paper are of this type. However, for more complicated eco-evolutionary dynamics, multiple alternate ESCs may exist, and under such conditions choices made regarding assembly order may yield different final communities. This situation is, however, no different compared to community assembly in adaptive dynamics, and similarly to the situation there, care must be taken when multiple ESCs are present, using tools from, for example, bifurcation theory to capture all the possible ESCs.

### *Eco-evolutionarily stable communities in the unstructured Lotka-Volterra model*

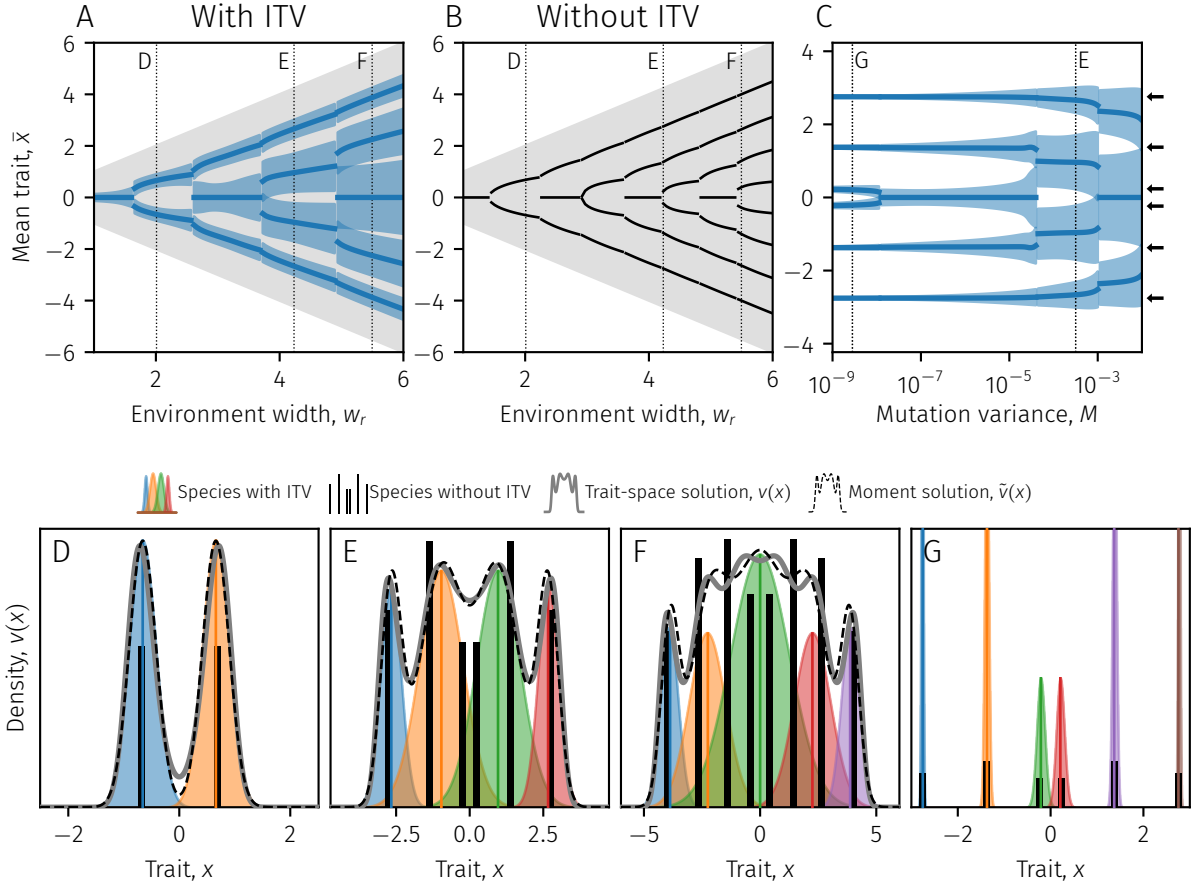
360 We now have all the tools necessary for assembling ESCs, so we will now turn to applying them to our unstructured Lotka-Volterra model (Eqs. 10). As we stated, the Lotka-Volterra model exhibits an inherent tension between the tendency towards multiple species and larger intraspecific trait variation (ITV) in each species under disruptive or weak stabilizing selection. Using adaptive dynamics, where no species has any ITV, Ranjan and Klausmeier (2022) studied a similar Lotka-Volterra model and found that as the environmental width became wider, an increasingly large number of species could co-exist in the assembled eco-evolutionarily stable community. Conversely, in another similar Lotka-Volterra model Barabás et al. (2022) found that ITV decreased when more species were included in a quantitative genetics models for a given environmental width, but had no way of systematically determine whether any of their communities were stable to invasions or evolutionary branchings. Here, we will use our moment equations together with

our assembly procedure to determine how community trait variation is partitioned into inter-  
and intraspecific terms in our Lotka-Volterra model.

To see how this tension manifests, we assemble eco-evolutionarily stable communities for a  
range of different environments. Specifically, we vary the environment variance  $V_r$  from 1 to  
36, although for ease of interpretation we will here present results using  $w_r = \sqrt{r_0 V_r}$  as the  
independent variable, which describes the half-width of where the net growth function  $r(x)$   
is positive (the *fundamental community* sensu Klausmeier et al. 2020). We depict the results in  
Fig. 2A. To further examine the role of ITV in eco-evolutionary community assembly, we also  
compare the results from our model with ITV to a model that does not incorporate ITV (adaptive  
dynamics, see Appendix C), depicted in Fig. 2B.

When  $w_r = 1$ , selection is strongly stabilizing, and only mutations prevent the trait variance  
of the single species from collapsing to zero. As  $w_r$  becomes wider, stabilizing selection weakens,  
and the disruptive selection from intraspecific competition initially increases trait variance, but  
eventually the one-species community becomes inviable and the community transitions into a  
two-species community. As we increase  $w_r$  the ESCs with ITV initially follows the ESCs without  
ITV closely (Fig. 2D). However, as  $w_r$  becomes increasingly large, the role of ITV becomes in-  
creasingly important, and the bifurcations into more species desynchronize between our model  
and the one without ITV (Fig. 2A–B). In particular, the species more centrally located in trait  
space show large differences for when ITV is included compared to when it is not.

Figure 2E–F show two communities where our model with ITV exhibits large discrepancies  
compared to when ITV is not taken into account, both in terms of the number of species present  
in the community as well as the values of the mean traits. One primary reason for this discrep-  
ancy is the fact that species close to the center of the environment develop large trait variances.  
This is in contrast to the species closer to the edges of the environment, which have less trait  
variance. These differences come about through intra- and interspecific competition. As can be  
seen in Eq. 10c, intraspecific competition always generates disruptive selection, and inter-specific  
competition can engender either stabilizing or disruptive selection, depending on how far apart



**Figure 2: Eco-evolutionarily stable communities for different environmental widths and mutation variances in the unstructured Lotka-Volterra model.** (A) For each value of environmental half-widths  $w_r = \sqrt{r_0 V_r}$  we depict the mean traits (solid blue lines) and one standard deviation (filled blue areas) for each species in the eco-evolutionarily stable community (ESC) computed assuming species with intraspecific variation (ITV) for that environmental width. The gray area depicts where the environment yields positive growth rate, i.e.,  $r(x) \geq 0$ . The mutation variance is  $M = 10^{-3.5}$ . (B) Same as panel A, but black lines here depict the mean traits in ESCs computed assuming species without ITV. (C) For each value of mutation variance  $M$  we depict the mean traits (solid blue lines) and one standard deviation (filled blue areas) for each species in the eco-evolutionarily stable community (ESC) for that mutation variance. Arrows on the right indicate the mean traits of the ESCs without ITV. The environmental width is  $w_r \approx 4.2$ . (D–G) ESCs computed assuming species with and without ITV for the values of  $w_r$  and  $M$  indicated by dotted lines in panels A, B, and C. The gray line depicts the numerical solution of the trait-space equations (Eqs. 3),  $v(x)$ , and the black broken line depicts the numerical solution of the moment equations (Eqs. 10),  $\tilde{v}(x) = \sum_{i=1}^S u_i \mathcal{N}(x, \bar{x}_i, V_i)$ , where the number of species  $S$  was determined through our assembly procedure. The colored areas depict the trait-densities of each species  $i = 1, \dots, S$  when ITV is included, and the black bars depict the densities and traits of species when ITV is not included. Parameter values other than  $w_r$  and  $M$  are as in Table 1.

the species' mean traits are. Additionally, all species are not equally prevalent since the environ-  
399 mental conditions are better towards the center of trait space, and the species there have larger  
total densities  $u$  compared to species close to the edge. This means that species close to the cen-  
ter will experience more intraspecific competition and less interspecific competition resulting in  
402 stronger disruptive selection, which in turn translates to more standing variation in the ESC for  
these species. This model is thus an example of when being able to keep track of both ITV and  
dynamically assembling a community with a variable number of species is required for under-  
405 standing how trait variation is partitioned over the long term. We also note that the substantial  
differences in trait variances between species in the ESCs for larger environmental widths means  
that models that incorporate ITV, but fixes the trait variances, would also have been insufficient  
408 for a good characterization of the distribution of traits.

Figure 2D–E show good agreement between the trait-space solution (in gray) and the moment-  
equation approximation (in black). Although qualitative agreement is usually good between the  
411 moment solution and trait-space solution for the ESC (e.g., Fig. 2F), the moment approximation  
becomes less accurate near transitions in species richness for the moment equation ESC. This  
is due to the fact that while the moment equations by construction always has a well-defined  
414 number of species, the trait-space equations have no such constraints and around the transition  
points in species richness for the moment equations, the trait-space ESCs tend to exhibit non-  
normal shapes that are ambiguous with regards to the number of species. We provide examples  
417 of this in Supplementary S1 where assembled communities can be seen for all different levels of  
environmental widths  $w_r$ .

The amount of standing variation in an ESC also depends on how much variation is gener-  
420 ated through mutation (Fig. 2C). For exceedingly low mutation variances  $M$ , the ESCs with ITV  
closely resemble those without ITV (Fig. 2G). However, for the species located in the middle of  
trait space, even very low levels of mutation variance can result in qualitative differences, with  
423 the two central species merging around  $M \approx 10^{-8}$ . To better understand why this merging hap-  
pens, we note as trait variances get larger due to increased mutation widths, both the directional

selection repulsing the two species (Eq. 10b) and the stabilizing selection generated by interspecific competition with other species (Eq. 10c) will get weaker. This will ultimately lead to the two species collapsing into one as the mutation variance gets bigger. As the mutation variance is further increased, other species will merge for the same reasons, but as stated above, more centrally located species are more sensitive to this pattern as they experience more intraspecific competition.

**Table 1:** Parameters for the two Lotka-Volterra models.

Symbol	Description	Value/Range	
		Unstructured	Two-patch
$r_0$	Maximal growth rate	1	1
$V_r$	Environment variance	[1,36]	4
$V_c$	Competition variance	1	1
$M$	Mutation variance	$10^{-3.5}$	$10^{-4}$
$x_1^{\text{opt}}$	Optimal trait, patch 1	-	[-2,0]
$x_2^{\text{opt}}$	Optimal trait, patch 2	-	[0,2]
$d$	Dispersal rate	-	$10^{-3}$

## Eco-evolutionary dynamics and community assembly in structured populations

For communities of unstructured populations we saw how weakened stabilizing selection could increase both the number of species and intraspecific trait variation (ITV), and that there was an inherent tension between these two forces. In spatially structured communities a related phenomenon spread over space can take place. In heterogeneous environments, spatially varying selection can lead to local adaptation within a species (Kirkpatrick and Barton, 1997; Bruggeman, 2009; Norberg et al., 2012; Le Gland et al., 2020), resulting in greater ITV across the landscape (essentially, increased beta trait diversity within a species). On the other hand, under the no-variance conditions of adaptive dynamics, variable local conditions can lead to coexistence of multiple species (Troost et al., 2005; Débarre and Gandon, 2010; Fortelius et al., 2015; Wickman

et al., 2017). Additionally, using models with fixed trait variance, large fixed variances have been shown to lead to communities with fewer surviving species but with more local adaptation, and smaller fixed variances to more surviving species but with less local adaptation (Norberg  
444 et al., 2012; Edwards et al., 2018). Thus, we here have a similar but distinct tension between ITV and diversification into multiple types, but now playing out over space in the form of local  
447 adaptation. To be able to model such scenarios, we need to generalize Eq. 2 to take population structure (including spatial structure) into account.

### *Trait-space equations for class-structured communities*

450 Since populations can be structured in ways other than spatial, for generality we will here assume that the community is class-structured, meaning that the community can be sorted into  $K$  discrete bins such as spatial patches, age classes, or size classes. The trait-density distribution in each class  
453  $k$  is given by  $v_k(x)$ ,  $k = 1, 2, \dots, K$ , which describes how abundant individuals with a given trait  $x$  in class  $k$  are. We also write  $\mathbf{v} := (v_1, \dots, v_K)$  for the vector of all trait-density distributions. For unstructured communities it was sufficient to consider a birth rate and a mortality rate, but in  
456 general structured communities other demographic processes can cause the trait-density  $v_k(x)$  in each class to change over time. We can consider separately the contribution of each process with per-capita rate  $f(x, \mathbf{v})$  from a source class  $s$  to a destination class  $d$ . Within-class processes  
459 ( $s = d$ ) may include local birth and mortality rates on a spatial patch, whereas between-class processes ( $s \neq d$ ) may include immigration from spatial patch  $s$  or from births from adult class  $s$  to juvenile class  $d$  in a stage-structured model. Each such process can have a mutation kernel  
462  $\mathcal{N}(x, y, M)$  associated with it (most processes will have  $M = 0$ ). The contribution of each process to the trait-space equations is given by

$$\frac{dv_d(x)}{dt} \stackrel{\pm}{=} \int_{-\infty}^{\infty} f(y, \mathbf{v}) v_s(y) \mathcal{N}(y, x, M) dy, \quad (14)$$

465 where the  $\stackrel{\pm}{=}$  operator means addition to the left-hand side, so that we sum up the rates of all processes to get the rate of change of  $v_d(x)$ . For (most) processes where  $M = 0$ , we will interpret

integration of the process rate against the mutation kernel in the delta-Dirac sense so that simply

$$\int_{-\infty}^{\infty} f(y, \mathbf{v}) v_s(y) \mathcal{N}(y, x, 0) dy = f(x, \mathbf{v}) v_s(x). \quad (15)$$

### *Trait-space equations for the two-patch Lotka-Volterra model*

To take a concrete example, we generalize our Lotka-Volterra competition model to take place on two patches, so that  $K = 2$ , with local births  $b_1$  and  $b_2$ , deaths  $\mu_1$  and  $\mu_2$ , competition  $a$  on each patch, and a constant symmetric rate of dispersal  $d$  between the patches. The trait-space equations are given by

$$\frac{dv_1(x)}{dt} = \int_{-\infty}^{\infty} b_1(y) v_1(y) \mathcal{N}(y, x, M) dy - \mu_1(x) v_1(x) - a(x, v_1) v_1(x) - d v_1(x) + d v_2(x), \quad (16a)$$

$$\frac{dv_2(x)}{dt} = \int_{-\infty}^{\infty} b_2(y) v_2(y) \mathcal{N}(y, x, M) dy - \mu_2(x) v_2(x) - a(x, v_2) v_2(x) + d v_1(x) - d v_2(x), \quad (16b)$$

$$b_k(x) = r_0, \quad \mu_k(x) = \frac{(x - x_k^{\text{opt}})^2}{V_r}, \quad r_k(x) = b_k(x) - \mu_k(x) = r_0 - \frac{(x - x_k^{\text{opt}})^2}{V_r}. \quad (16c)$$

This model is locally the same as the unstructured Lotka-Volterra model (Eqs. 3) on each patch apart from the optimal traits  $x_1^{\text{opt}}$  and  $x_2^{\text{opt}}$  now potentially being different. Comparing with the generic trait-space equations (Eq. 14), when patch one is the destination patch (Eq. 16a), we thus have five demographic processes, four for which patch one is the source patch, namely  $b_1$ ,  $-\mu_1$ ,  $-a$ , and  $-d$ , and one for which patch two is the source patch,  $d$ . Of these, only  $b_1$  has a non-zero mutation variance  $M$  associated with it.

### *Moment equations for class-structured communities*

As for the unstructured community, we will assume that we can approximate each trait-density distribution  $v_k(x)$  in each class with a sum of  $S$  normal distributions with total density  $u_{ik}$ , mean trait  $\bar{x}_{ik}$  and trait variance  $V_{ik}$  for species  $i = 1, \dots, S$  in class  $k = 1, \dots, K$ , so that

$$v_k(x) \approx \tilde{v}_k(x) = \sum_{i=1}^S u_{ik} \mathcal{N}(x, \bar{x}_{ik}, V_{ik}). \quad (17)$$

Note that a species will exist across all classes (even if its density could be very close to zero in some) so that the species richness in each class, as well as globally, is equal to  $S$ . We define  $\tilde{\mathbf{v}} =$



$(\tilde{v}_1, \dots, \tilde{v}_K)$  to be the vector of approximate densities. Also as for the unstructured communities we define the population-level per capita rate for a population with mean  $\bar{x}$  and trait variance  $V$

492 to be

$$\hat{f}(\bar{x}, V, \tilde{\mathbf{v}}) = \int_{-\infty}^{\infty} f(x, \tilde{\mathbf{v}}) \mathcal{N}(x, \bar{x}, V) dx. \quad (18)$$

We can now derive the moment equations for our class-structured system (see Appendix A),

495 which, when summing over all demographic processes, are given by

$$\frac{du_{id}}{dt} \stackrel{(i) \text{ population-level per capita rate}}{=} \hat{f}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}}) u_{is}, \quad (19a)$$

$$\frac{d\bar{x}_{id}}{dt} \stackrel{(ii) \text{ relative-density weight}}{=} \frac{u_{is}}{u_{id}} \left[ \underbrace{V_{is} \frac{\partial \hat{f}}{\partial \bar{x}}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}})}_{(iii) \text{ directional selection}} + \underbrace{\hat{f}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}})(\bar{x}_{is} - \bar{x}_{id})}_{(iv) \text{ mean-trait flow}} \right], \quad (19b)$$

$$\begin{aligned} \frac{dV_{id}}{dt} \stackrel{(ii)}{=} & \frac{u_{is}}{u_{id}} \left[ \underbrace{V_{is}^2 \frac{\partial^2 \hat{f}}{\partial \bar{x}^2}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}})}_{(v) \text{ stabilizing/disruptive selection}} + \underbrace{\hat{f}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}})(V_{is} - V_{id})}_{(vi) \text{ trait-variance flow}} \right. \\ & + \underbrace{\hat{f}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}})(\bar{x}_{is} - \bar{x}_{id})^2}_{(vii a) \text{ between-to-within class variation}} + \underbrace{2V_{is} \frac{\partial \hat{f}}{\partial \bar{x}}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}})(\bar{x}_{is} - \bar{x}_{id})}_{(vii b) \text{ class-local adaptation and directional selection interaction}} \\ & \left. + \underbrace{\hat{f}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}})M}_{(viii) \text{ mutation}} \right]. \quad (19c) \end{aligned}$$

### 501 *Interpreting the moment equations for class-structured communities*

As noted for other moment-based frameworks (e.g., Norberg et al., 2001), in addition to being more tractable, moment equations can also be more interpretable than the trait-space equations from which they are derived. We now go through the various terms and factors in the moment equations (Eqs. 19) and their interpretations. Table 2 contains a list of symbols and descriptions of the various quantities related to the class-structured trait-space and moment equations.

507 Equation 19a describes the rate of change of the total density  $u_{id}$  of species  $i$  in destination class  $d$ . Term (i), *population per capita growth*, describes the per capita rate of process  $\hat{f}$  between

**Table 2:** Quantities involved in the trait-space and moment equations for class-structured communities.

Symbol	Description	Definition	Label in Eqs. 19
$x$	Trait value		
$K$	Number of classes		
$v_k(x)$	Trait-density distribution in class $k = 1, \dots, K$		
$\mathbf{v}$	Vector of trait-density distributions	$\mathbf{v} = (v_1, \dots, v_K)$	
$f(x, \mathbf{v})$	Per capita rate of a demographic process		
$M$	Mutation variance associated with process $f$		
$\hat{f}(\bar{x}, V, \mathbf{v})$	Population-level rate for mean trait $\bar{x}$ and trait variance $V$ for rate $f$	$\hat{f}(\bar{x}, V, \mathbf{v}) = \int_{-\infty}^{\infty} f(x, \mathbf{v}) \mathcal{N}(x, \bar{x}, V) dx$	
$S$	Number of species		
$v_{ik}(x)$	Trait-density distribution of species $i$ in class $k$		
$u_{ik}$	Total density of species $i$ in class $k$	$u_{ik} = \int_{-\infty}^{\infty} v_{ik}(x) dx$	
$\bar{x}_{ik}$	Mean trait of species $i$ in class $k$	$\bar{x}_{ik} = (1/u_{ik}) \int_{-\infty}^{\infty} x v_{ik}(x) dx$	
$V_{ik}$	Trait variance of species $i$ in class $k$	$V_{ik} = (1/u_{ik}) \int_{-\infty}^{\infty} (x - \bar{x}_{ik})^2 v_{ik}(x) dx$	
$\tilde{v}_k(x)$	Approximate trait-density distribution in class $k$	$\tilde{v}_k(x) = \sum_{i=1}^S u_i \mathcal{N}(x, \bar{x}_{ik}, V_{ik})$	
$\tilde{\mathbf{v}}$	Vector of approximate trait-density distributions	$\tilde{\mathbf{v}} = (\tilde{v}_1, \dots, \tilde{v}_K)$	
$d$	Index of the destination class		
$s$	Index of the source class		
$\hat{f}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}})$	Population-level per capita growth		(i)
$u_{is}/u_{id}$	Relative-density weight		(ii)
$V_{is} \frac{\partial \hat{f}}{\partial \bar{x}}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}})$	Directional selection		(iii)
$\hat{f}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}})(\bar{x}_{is} - \bar{x}_{id})$	Mean-trait flow		(iv)
$V_{is}^2 \frac{\partial^2 \hat{f}}{\partial \bar{x}^2}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}})$	Stabilizing/disruptive selection		(v)
$\hat{f}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}})(V_{is} - V_{id})$	Trait-variance flow		(vi)
$\hat{f}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}})(\bar{x}_{is} - \bar{x}_{id})^2$	Between-to-within class variation		(viiia)
$2V_{is} \frac{\partial \hat{f}}{\partial \bar{x}}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}})(\bar{x}_{is} - \bar{x}_{id})$	Class-local adaptation and directional selection interaction		(viiib)
$(viiia) + (viiib)$	Effects of class-local adaptation on variance	$(vii) = (viiia) + (viiib)$	(vii)
$\hat{f}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}})M$	Mutation		(viii)

source class  $s$  and destination class  $d$  evaluated at the mean trait  $\bar{x}_{is}$  and trait variance  $V_{is}$  of  
510 species  $i$  in the source class  $s$ . The equations capture how the total population densities of  
species change both due to within-class processes ( $s = d$ ) such as local birth and death, and  
between-class processes ( $s \neq d$ ) such as dispersal between patches.

513 Equation 19b describes the rate of change of the mean trait  $\bar{x}_{id}$  of species  $i$  in destination class  
 $d$ . Term (ii), *relative-density weight*, weighs contributions by the relative total densities of classes  
 $d$  and  $s$ , so that if the density in the destination class  $d$  is much greater than that in the source  
516 class  $s$ , the ecological processes going from  $s$  to  $d$  will only have a marginal impact on the mean  
trait in class  $d$ . Conversely, if the total density in the source class  $s$  is much greater than in the  
destination class  $d$ , the ecological process will have a large impact on the mean trait in class  $d$ .  
519 The change of the mean trait is then governed by two terms, (iii) and (iv). The first term, (iii),  
*directional selection*, describes the effect of directional selection in the process  $\hat{f}$  in class  $s$ , pushing  
 $\bar{x}_{id}$  in the direction of maximum increase of  $\hat{f}$ , at a rate proportional to the trait variance of species  
522  $i$  in class  $s$ ,  $V_{is}$ . This effect comes about since the process in the source class will produce more  
trait-density on the side of the mean in which the slope is pointing, and less on the other side.  
The second term, (iv), *mean-trait flow*, describes how mean traits are homogenized by between-  
525 class transitions, where the rate of homogenization is governed by the per capita rate function  $\hat{f}$ ,  
so that the mean trait of the destination class,  $\bar{x}_{id}$ , will change in the direction of the mean trait  
in the source class,  $\bar{x}_{is}$ . For within-class processes ( $s = d$ ) we note that the relative-density weight  
528 (ii) = 1 and the mean-trait flow (iv) = 0, meaning that for within-class processes only directional  
selection (iii) is relevant. If the process under consideration is trait-independent, such as for a  
constant dispersal rate between patches, directional selection (iii) would be equal to zero, but  
531 mean-trait flow (iv) could still contribute towards changing the mean trait in class  $d$  if the mean  
traits in  $s$  and  $d$  differ.

Equation 19c describes the rate of change of the trait variance  $V_{id}$  of species  $i$  in destination  
534 class  $d$ . As for the mean traits, the changes are weighted by the relative densities between class  
 $d$  and  $s$ , term (ii). The dynamics of the trait variance then depends on five terms. The first

term,  $(v)$ , describes *stabilizing/disruptive selection* resulting from the process with rate  $\hat{f}$ . Roughly speaking, if individuals close to the mean trait  $\bar{x}_{is}$  of species  $i$  in class  $s$  contribute more than individuals away from this optimum to process  $\hat{f}$  then the curvature as measured by  $\partial^2 \hat{f} / \partial \bar{x}^2$  will be negative, which will contribute to a decrease in trait variances. Conversely, if individuals close to the mean contribute less, the curvature will be positive and this will contribute towards increases in trait variances. The second term,  $(vi)$ , *trait-variance flow*, homogenizes trait variance between classes, so that having trait-density flow from class  $s$  to  $d$  drive the variance  $V_{id}$  of species  $i$  in class  $d$  closer to the trait variance  $V_{is}$  in class  $s$ . Terms  $(vii)$ , *effects of class-local adaptation on variance*, describe the effects of variability in mean traits between classes on the trait variance within each class. In the case where the classes are spatial patches, a species having different mean traits for different patches would simply be referred to as ‘local adaptation’, and we adapt this moniker here to the broader context of any class-structured community. Term  $(viii)$ , *between-to-within class variation*, describes how trait variances are increased by the differences in mean traits between classes, converting between-class variance into within-class variance. Term  $(viib)$ , *class-local adaptation and directional selection interaction*, describes the interaction between directional selection and between-class differences in mean traits. Roughly speaking, trait variances will decrease when the mean trait difference and the selection gradient point in opposite directions, and increase if they are aligned. Finally, term  $(viii)$ , *mutation*, is the contribution to trait variance from mutations in process  $\hat{f}$ , which will contribute towards increasing trait variances. Note that typically for most processes under consideration  $M$  would be zero, as in our Lotka-Volterra example (Eqs. 16), where only birth processes are assumed to have mutations associated with them. How much mutations contribute towards trait variance increase also depends on the rate  $\hat{f}$ . Thus, for example, in a system with high birth and death rates with mutations associated with births the mutations would have a stronger impact on trait variance than in a system with low birth and death rates even if net per capita growth were the same in both systems. For within-class processes ( $s = d$ ), the relative density weight  $(ii)$  will be equal to one, and only stabilizing/disruptive selection  $(v)$  and mutations  $(viii)$  will be nonzero, making

these the only contributing factors. If the process  $\hat{f}$  under consideration is trait independent, sta-  
 564 bilizing/disruptive selection ( $v$ ) and class-local adaptation and directional selection interaction  
 ( $viib$ ) will be zero, but mutations ( $viii$ ), variance flow ( $vi$ ), and between-to-within class variation  
 ( $viii$ ) may still contribute to changes in the trait variances.

567 *Moment equations for the two-patch Lotka-Volterra model*

For a specific example, we take our two-patch Lotka-Volterra model with trait-space equations  
 given by Eqs. 16. We can, after identifying the various rates now plug these into the generic  
 570 class-structured moment equations (Eqs. 19) to yield the moment equations for the two-patch  
 Lotka-Volterra system. Below we display the moment equations for the dynamics on patch one;  
 the dynamics on patch two are nearly identical with patch index one swapped for patch index  
 573 two as necessary:

$$\frac{du_{i1}}{dt} = \left[ \underbrace{\left( r_0 - \frac{(\bar{x}_{i1} - x_1^{\text{opt}})^2 + V_{i1}}{V_r} \right)}_{(i), \hat{r}_1(\bar{x}_{i1}, V_{i1}) = \hat{b}_1(\bar{x}_{i1}, V_{i1}) - \hat{\mu}_1(\bar{x}_{i1}, V_{i1})} - \underbrace{\sqrt{2\pi V_c} \sum_{j=1}^S u_{j1} \mathcal{N}(\bar{x}_{i1}, \bar{x}_{j1}, V_{i1} + V_{j1} + V_c)}_{(i), \hat{a}(\bar{x}_{i1}, V_{i1}, \bar{v}_1)} \right] u_{i1} - \underbrace{d}_{(i)} u_{i1} + \underbrace{d}_{(i)} u_{i2}, \quad (20a)$$

$$\frac{d\bar{x}_{i1}}{dt} = V_{i1} \left[ \underbrace{-\frac{2(\bar{x}_{i1} - x_1^{\text{opt}})}{V_r}}_{(iii), \frac{\partial \hat{r}_1}{\partial \bar{x}}(\bar{x}_{i1}, V_{i1})} + \underbrace{\sqrt{2\pi V_c} \sum_{j=1}^S u_{j1} \frac{\bar{x}_{i1} - \bar{x}_{j1}}{V_{i1} + V_{j1} + V_c} \mathcal{N}(\bar{x}_{i1}, \bar{x}_{j1}, V_{i1} + V_{j1} + V_c)}_{(iii), -\frac{\partial \hat{a}}{\partial \bar{x}}(\bar{x}_{i1}, V_{i1}, \bar{v}_1)} \right] + \underbrace{\frac{u_{i2}}{u_{i1}} d(\bar{x}_{i2} - \bar{x}_{i1})}_{(ii) \quad (iv)}, \quad (20b)$$

$$\frac{dV_{i1}}{dt} = V_{i1}^2 \left[ \underbrace{-\frac{2}{V_r}}_{(v), \frac{\partial^2 \hat{r}_1}{\partial \bar{x}^2}(\bar{x}_{i1}, V_{i1})} + \underbrace{\sqrt{2\pi V_c} \sum_{j=1}^S u_{j1} \frac{V_{i1} + V_{j1} + V_c - (\bar{x}_{i1} - \bar{x}_{j1})^2}{(V_{i1} + V_{j1} + V_c)^2} \mathcal{N}(\bar{x}_{i1}, \bar{x}_{j1}, V_{i1} + V_{j1} + V_c)}_{(v), -\frac{\partial^2 \hat{a}}{\partial \bar{x}^2}(\bar{x}_{i1}, V_{i1}, \bar{v}_1)} \right] + \underbrace{\frac{u_{i2}}{u_{i1}} d(V_{i2} - V_{i1})}_{(ii) \quad (vi)} + \underbrace{\frac{u_{i2}}{u_{i1}} d(\bar{x}_{i2} - \bar{x}_{i1})^2}_{(ii) \quad (viii)} + \underbrace{r_0 M}_{(viii)}. \quad (20c)$$

Here, we have marked the various terms with their corresponding terms in the generic equations Eqs. 19. The total-density dynamics (Eq. 20a) describe the local net growth on patch one due to the environmental birth and death rates, the mortality from local competition, and dispersal to and from patch two. The mean-trait dynamics (Eq. 20b) describe local selection on patch one, where the selection due to the environmental growth rate drives the mean trait towards the patch one optimum  $x_1^{\text{opt}}$ , and the selection due to local competition drives each species mean trait apart from the mean traits of other species. The second term describes mean-trait flow from patch two driving the mean trait on patch one to become more similar on patch two, eroding local adaptation over time. The first term of the trait-variance dynamics (Eq. 20c) describes local stabilizing/disruptive selection, where selection due the environmental growth rate is universally stabilizing, and selection from local competition can be both stabilizing or disruptive in the same way as for the unstructured model. The next term describes trait-variance flow which drives the trait variance on patch one to become closer to the variance on patch two. The final term describes how between-patch variation, i.e., the difference between mean traits on the patches, is converted to within-patch variation by driving an increase in variance. Note that since the only between-patch process, with rate  $d$ , is trait independent, term  $(viiib)$  in the generic equations (Eq. 19c) does not arise in this two-patch model.

### *Eco-evolutionary community assembly for class-structured communities*

Our general approach for community assembly in class-structured communities closely resembles that for unstructured communities, albeit with more involved mathematical machinery. For evolutionary branchings, we perform the same kind of splitting and stability analysis as for the unstructured communities. The invasion analysis too proceeds along similar lines, but since the community is now structured we need to keep track of the frequency distribution of invaders across classes. As for the unstructured model, we can use these branching and invasion criteria to build up a community one species at a time until no more invasions or branchings are possible and we have an eco-evolutionarily stable community. The details of how the branching and

606 invasion analysis are carried out for structured communities are available in Appendix B.

### *Eco-evolutionarily stable communities in the two-patch Lotka-Volterra model*

When the optimal traits on the patches  $x_1^{\text{opt}}$  and  $x_2^{\text{opt}}$  differ between the patches there will be local  
609 selection towards different traits on the two patches (Eq. 20b). Under sufficiently low dispersal  
a single species can thus be expected to exhibit local adaptation, with its mean trait differing  
between patches. However, based on insights derived from adaptive dynamics, we expect to  
612 instead see multiple species as a response to heterogeneous local conditions. This then raises the  
question when these two possibilities are combined: with heterogeneous local conditions and  
weak dispersal, will we end up with fewer locally adapted species or more species with less local  
615 adaptation but differing in their mean traits?

In the two-patch Lotka-Volterra model (Eqs. 20) as well as the general class-structured mo-  
ment equations (Eqs. 19), we can see that directional selection (term *(iii)*) is multiplied by the  
618 within-patch trait variance of the species to determine the effect of directional selection on  
changes in the mean trait. We can also see that that there is a term that converts between-patch  
variation into within-patch variation (term *(viii)*). This means that if the trait optima on the two  
621 different patches are different, there will be selective pressure for the mean traits in a species  
to separate, and as they separate, term *(viii)* will increase within-class variation strengthening  
local directional selection, further increasing the selective pressure on trait separation. This, then,  
624 creates a positive feedback between local adaptation (between-class variation) and within-class  
variation. However, the potential of local adaptation as a strategy for covering more of trait space  
on the regional scale has two limitations potentially opening up a locally adapted species to in-  
627 vasion. First, the mean-trait-flow term *(iv)* acts as a barrier to local adaptation by exerting a force  
towards mean-trait homogenization and may prevent sufficient local adaptation from develop-  
ing, leaving unused trait space available for invasion. Conversely, if local adaptation, and hence  
630 within-patch trait variance, becomes too large this might make the species relatively maladapted  
by covering too much unfavorable trait space, again opening up the species to invasion. None of

these scenarios are universally favored above another, and local adaptation, multiple species, or  
633 neither will be model- and parameter-dependent.

To explore a specific instance of this tension between local adaptation and divergence into  
multiple species we use our Lotka-Volterra model (Eqs. 20) and assemble eco-evolutionarily sta-  
636 ble communities under conditions of weak dispersal and for a range of differences in the trait  
optima between the two patches. Specifically, we let  $V_r = 4$  on each patch, which is enough  
environmental breadth for two species to coexist in the unstructured model, and then make the  
639 patches more dissimilar by varying  $x_2^{\text{opt}} = -x_1^{\text{opt}}$  from 0 to 2. The results are depicted in Fig. 3,  
and parameter values are listed in Table 1.

Generally, the outcomes when we vary the patch dissimilarity fall into three categories with  
642 regards to local adaptation. The first is no local adaptation (for example, Fig. 3E–F). The second  
is symmetric local adaptation, meaning that a species is split roughly equally between the two  
patches, and the mean traits between the patches differ (for example, Fig. 3C). The third is  
645 asymmetric local adaptation, where species are significantly more prevalent on one patch, and  
their mean trait on the sink patch differs from that of the main patch (for example, Fig. 3D). To  
better illustrate the effects of the inclusion of intraspecific trait variation (ITV), we also depict  
648 ESCs for a model which does not incorporate ITV (adaptive dynamics, see Appendix C). Note  
that without ITV, all individuals in a species are identical, meaning that a species necessarily has  
the same mean trait on both patches. Below we describe in detail how the ESCs change as the  
651 patch optima become further separated.

When the optima are both equal to zero, the patches are identical and the two-patch system is  
functionally equivalent to an unstructured model. Conversely, when the optima are equal to  $\pm 2$   
654 there is no overlap in trait space of positive growth rates between the two patches, and for any  
species having a positive environmental net growth rate on one patch, the other patch will be a  
pure sink. To the very left in Fig. 3A–B the communities with and without ITV agree in terms of  
657 mean traits, and since the patches are identical, two species exist across the region. However, after  
the patch optima separate only slightly, the community without ITV diverges into a four-species



system, with two species roughly corresponding to the community that would have evolved on  
660 patch one in isolation and one community corresponding to the one that would have evolved on  
patch two in isolation. In contrast, the communities with ITV remain as a two-species system all  
the way up to roughly  $x_k^{\text{opt}} = \pm 0.25$ , but with both species exhibiting symmetric local adaptation  
663 between the patches (Fig. 3C).

The community with ITV does split into four species around  $x_k^{\text{opt}} = \pm 0.25$ , but the resulting  
four-species community is still characterized by significant asymmetric local adaptation in at least  
666 two species for the interval between approximately  $0.25 \leq \pm x_k^{\text{opt}} \leq 0.5$  (Fig. 3D). The two species  
depicted in blue and orange are here primarily patch-one specialists and the species depicted in  
green and red are primarily patch-two specialists. The asymmetry in local adaptation can come  
669 about through the weighted effect of densities on mean-trait flow between the patches, terms (ii)  
and (iii) in Eqs. 20, where the species density on the off-patch exerts relatively little pressure on  
the mean trait on the main patch to diverge from its local optimum.

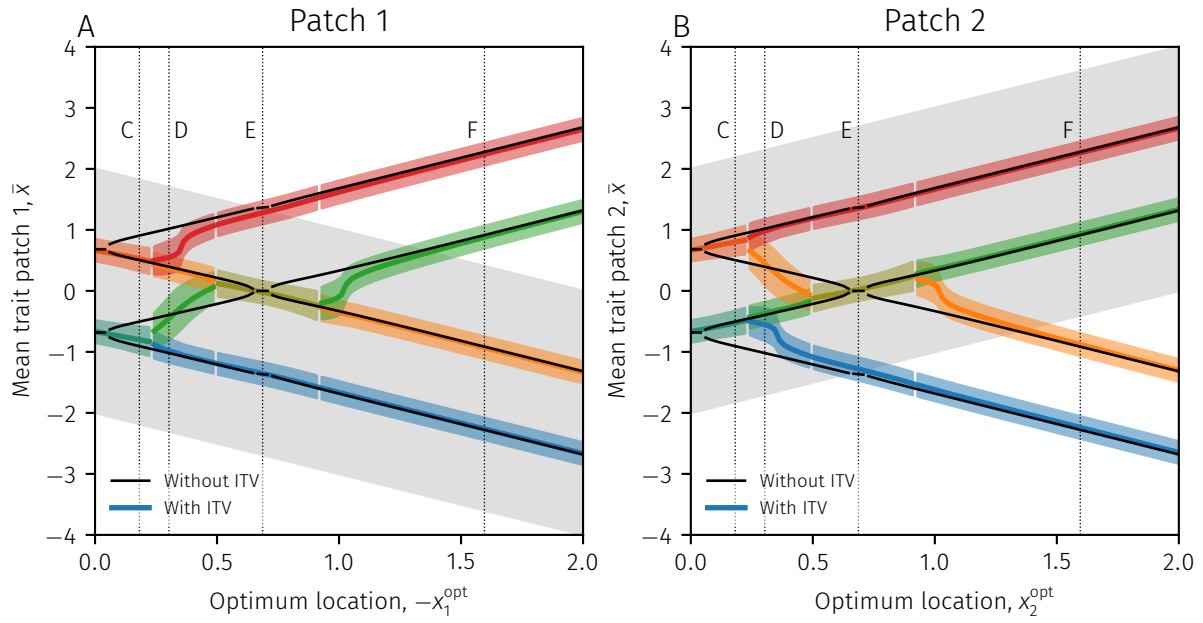
672 Around  $x_k^{\text{opt}} = \pm 0.7$  the two sets of lines describing the community without ITV on each patch  
cross, and there is a small interval where the community without ITV will have only three species  
before the lines separate after the cross into four species again. For the community with ITV, the  
675 interval with three species is significantly larger, spanning roughly  $0.5 \leq \pm x_k^{\text{opt}} \leq 0.9$ . For  
this interval, the community is characterized by two outer species (blue and red) with little local  
adaptation and a central species (olive) that exhibits significant and symmetrical local adaptation,  
678 except in the middle of the interval, where there is no significant local adaptation in any of the  
three species (Fig. 3E).

After  $x^{\text{opt}} = \pm 0.9$  the community with ITV once again splits into four species, where the cen-  
681 tral two species (orange and green) are initially characterized by asymmetrical local adaptation.  
As the environmental optima separate further the means converge onto those for the community  
without ITV, and for large separations  $|x_k^{\text{opt}}| > 1.1$  the system for both the communities with and  
684 without ITV are characterized by two species (blue and orange) being present almost entirely on  
patch one, and two species (green and red) being present almost entirely on patch two (Fig. 3F).

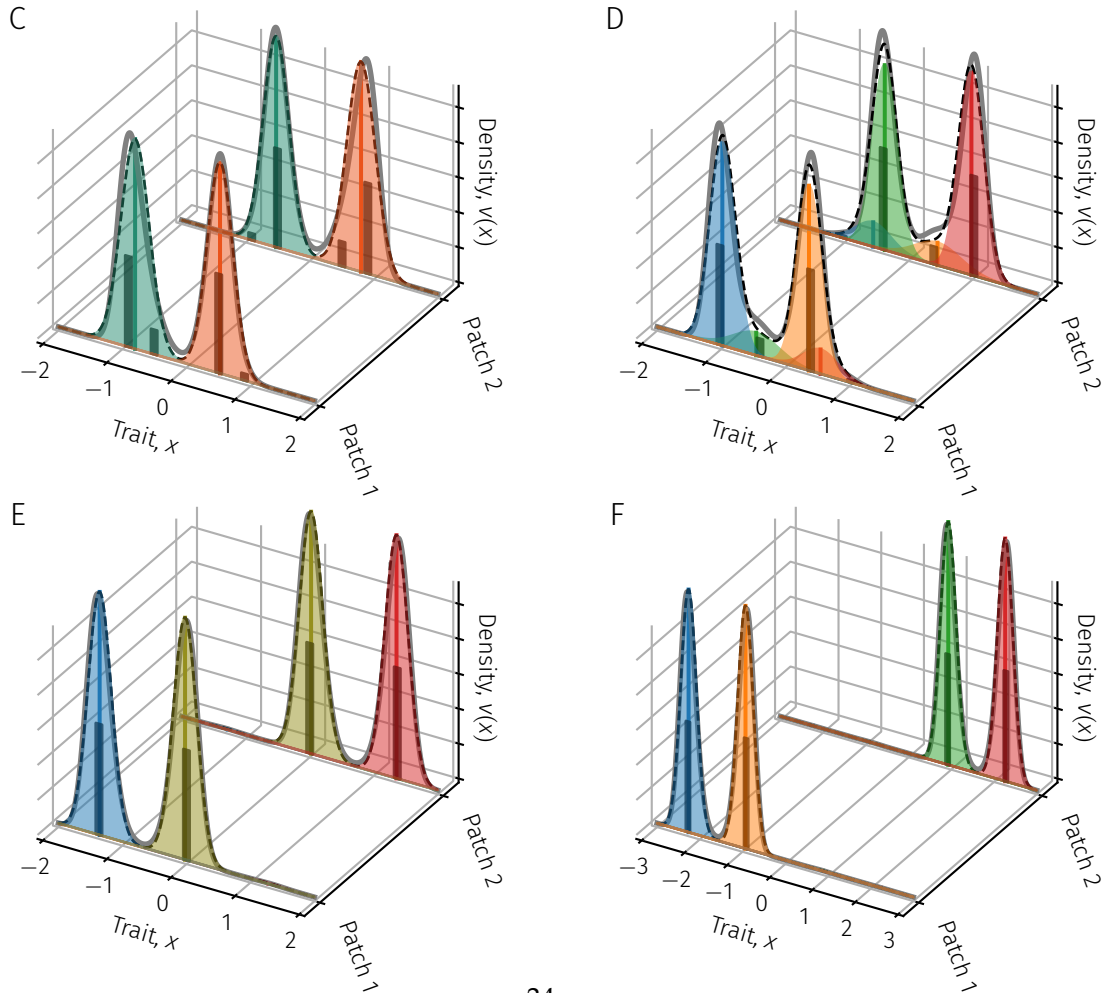
Due to the huge asymmetry in densities across the two patches, the mean-trait flow from the  
687 main patch to the sink patch becomes so strong that local adaptation can effectively not occur on  
the sink patch.

Across the range of patch dissimilarities, this model then showcased a range of behaviors  
690 with regards to local adaptation and diversification into multiple species. The fact that we could  
roughly sort these outcomes into three qualitatively different regimes is particularly interesting,  
and we could not have come across this behavior without the use of our framework.

693 As for the unstructured model, the two-patch moment equation ESCs mostly have good  
agreement with the trait-space equations, with some issues around the transition points in the  
number of species (see Supplementary S1). One issue that did not arise in the unstructured  
696 model is that very close to the transition from two to four species ( $x_k^{\text{opt}} \approx \pm 0.25$ ), the two-species  
community depicted is invasible, but after the invasion, one of the original species is depressed  
to extinction, after which the invader assumes the total density, mean, and variance of the extinct  
699 resident, yielding a never-ending cycle of invasions and extinctions. Similarly, the species are also  
branchable, but after an initial divergence in trait space, one of the branching species goes extinct  
and the remaining species assumes the distribution of the original branched resident, yielding a  
702 never-ending branching–extinction cycle. For these kind of cycles, we define the ESC to be the  
community to which the cycle returns between branchings or invasions that is in equilibrium  
with respect to Eqs. 19, which is the community we depict in Fig. 3A–B for  $x_k^{\text{opt}} \approx \pm 0.25$ . This  
705 cycling phenomenon only occurs in a very small sliver of parameter space.



Species with ITV    Species without ITV    Trait-space solution,  $v_k(x)$     Moment solution,  $\bar{v}_k(x)$



**Figure 3: Eco-evolutionarily stable communities for different environmental optima in the two-patch Lotka-Volterra model.** (A–B) For each value of environmental optima  $x_1^{\text{opt}}$  and  $x_2^{\text{opt}}$  we depict the mean traits (solid colored lines) and one standard deviation (filled colored areas) for each species in the eco-evolutionarily stable community (ESC) computed assuming species with intraspecific trait variation (ITV) for that pair of optima. Depicted as black lines are the ESCs for communities computed assuming species without ITV. Note that without ITV, local adaptation is precluded, and these black lines are the same in panels A and B. The gray area depicts where the environment yields positive growth rate, i.e.,  $r_k(x) \geq 0$ , on patch one (panel A), and patch two (panel B). The species in the communities with ITV are color coded so that the same species is depicted in the same color for both patches for any given value of  $x_k^{\text{opt}}$ . (C–F) ESCs for the values of  $x_k^{\text{opt}}$  indicated by the dotted lines in panels A and B. The gray line depicts the numerical solution of the trait-space equations (Eqs. 16),  $v_k(x)$ , and the black broken line depicts the numerical solution of the moment equations (Eqs. 20),  $\tilde{v}_k(x) = \sum_{i=1}^S u_{ik} \mathcal{N}(x, \bar{x}_{ik}, V_{ik})$ , where the number of species  $S$  was determined through our assembly procedure. The colored areas depict the trait-density of each species  $i = 1, \dots, S$  for the moment equations, and the black bars depict the traits and densities of the ESC without ITV. Parameter values are as in Table 1.

708

## Generalizing the framework

For the purposes of exposition and to focus on our main questions regarding the tension between intraspecific trait variation, local adaptation, and diversification into multiple species we have kept our example models simple. However, several more complications can be handled by generalizing the framework. First, more than one trait may affect performance, so tracking a single scalar trait may not always be sufficient, and in such cases we would need some way of deriving the dynamics for a mean-trait vector and a trait variance-covariance matrix. In Appendix A and Appendix B we show how the framework can be expanded to include multiple traits in this way. Second, in our Lotka-Volterra models the rate functions were simple enough that we could calculate the population-level rates necessary for the moment equations using Gaussian integration analytically. This will in general not be possible, and in Appendix A, section *Series solutions and Taylor approximations*, we show how to derive Taylor-approximations of arbitrary order to accommodate such situations. Third, our Lotka-Volterra models do not include any dependence on any external variables such as abiotic resources, and in Appendix A, section *Environmental variables*, we show how external variables can be included in the framework. To provide an example that includes all these complications we have in Appendix D briefly explored a stage-structured model with one juvenile and one adult stage in two traits that compete for two

723

abiotic resources.

726 We have here focused on fully heritable trait variation, but our framework can be extended  
to include non-heritable environmental variation. While we have not worked this case out to the  
same level of generality as we have for fully heritable traits, in Appendix A, section *Environmental,*  
729 *non-heritable trait variation*, we provide a sketch for the simplest case of one trait in unstructured  
populations. While it is relatively straight-forward to incorporate the generation of non-heritable  
variation into the model, it requires additional moment equations, and the derivation of the  
732 moment equations from the trait-space equations becomes more involved.

## Discussion

In this paper we have presented a general framework for eco-evolutionary community assembly  
735 for class-structured communities that incorporates intraspecific trait variation. We have done so  
by deriving moment equations for the total density, mean trait, and trait variance for each species  
in a community, combined with a procedure for determining whether additional species need to  
738 be added to a community in order for it to be closed to further invasion. Through examples, we  
demonstrated the application of the framework in an unstructured and a two-patch-structured  
Lotka-Volterra competition model, where we saw how less stabilizing conditions could result in  
741 different combinations of more intraspecific variation, local adaptation, and the addition of more  
species to the assembled eco-evolutionarily stable community (ESC).

### *Intra- and interspecific trait variation*

744 In recent years, the role of intraspecific vs. interspecific variation in traits has received increasing  
attention in functional and community ecology, with two broad questions at the center. First,  
how much of trait variation is intraspecific and how much is interspecific (Albert et al., 2011;  
747 Siefert et al., 2015; Griffiths et al., 2016; Gaudard et al., 2019; Xavier Jordani et al., 2019)? And  
second, how important is intraspecific trait variation (ITV) for higher-level outcomes such as

species coexistence and ecosystem functioning (Bolnick et al., 2011; Violle et al., 2012; Turcotte  
750 and Levine, 2016; Raffard et al., 2019)?

Regarding the first question, ITV has in general been found to account for substantial por-  
tions of trait variation, especially in plants (Siefert et al., 2015; Westerband et al., 2021). However,  
753 in some study systems of ants (Gaudard et al., 2019) and beetles (Griffiths et al., 2016) ITV was  
found to be negligible, and even among plant studies the preponderance of ITV is highly variable  
(Westerband et al., 2021), and patterns of intra- and interspecific variation can be highly idiosyn-  
756 cratic (Costa-Pereira et al., 2018; Umaña and Swenson, 2019). For the insect–plant discrepancy,  
Gaudard et al. (2019) suggested that one explanation could be related to the higher plasticity in  
plants that arises as a consequence of plants being sessile. In the context of individual niche spe-  
759 cialization, several factors have been proposed as important for accounting for the level of ITV  
including intra- and interspecific competition, ecological opportunity (the diversity of resources),  
and predation (Araújo et al., 2011). While these factors can no doubt play a role in shaping ITV,  
762 they can also, in turn be shaped by ITV. In our unstructured Lotka-Volterra model (Eqs. 10, Fig. 2)  
ecological opportunity can roughly be said to correspond to the environmental width  $w_r$  which  
we take to be fixed, but intra- and interspecific competition both shape and are shaped by intra-  
765 and interspecific trait variation through eco-evolutionary feedbacks, so that neither can be said  
to be the strict cause of the other. In field studies, idiosyncratic patterns of trait variation are  
to some extent to be expected due to the myriad factors that can be present in natural systems.  
768 Interestingly, however, even for our simple model with just one parameter varied (Fig. 2A), ITV  
is sometimes, but not always, important for characterizing the trait variation in the community.  
When the width of the environment is small, yielding one or two species ( $w_r \leq 2.3$ ), ITV is small  
771 so the trait means of the species of our model incorporating ITV agree very well with a model  
not incorporating ITV, implying that ITV is not important for characterizing the trait distribution  
in the community. For larger environmental widths, however, ITV becomes much more substan-  
774 tial, and the mean trait values, and even the number of species in the community, no longer  
agree with the model without ITV, implying that the role of ITV is crucial for characterizing the

trait distribution in the community. Given the simplicity of our model, our results thus suggest that whether ITV is important for characterizing communal trait variation could be highly system-specific and may resist broad-scale explanation. However, a more complete theoretical exploration of whether broad patterns in different models can be found that serve as good indicators for the importance of ITV is outside the scope of this paper, and more research will be required to determine the extent to which such patterns exist. Our framework would serve as a good tool for carrying out such explorations.

Scaling up to structured communities, Albert et al. (2011) proposed that in a nested sampling design, ITV should saturate as the spatial or ecological scale is increased since at a sufficiently large scale, species will cover their entire range meaning that interspecific trait variation should become the more important factor. Although our two-patch Lotka-Volterra model (Eqs. 20, Fig. 3) is not set up for one-to-one comparisons with this hypothesis, we observe some related phenomena as the environments on the two patches become increasingly different (Fig. 3). Initially, ITV, mostly in the form of local adaptation, increases as the patches separate. More interspecific variation is then added as more species join the community, but beyond a certain point the patches are too different to permit local adaptation and interspecific variation dominates. To wit, not much theoretical attention has been paid to what conditions generically promote intra- vs. interspecific variation on the regional scale (but see Norberg et al., 2012; Edwards et al., 2018), and a more systematic exploration of this question would be an interesting application of our framework.

Regarding the second question concerned with how ITV affects higher-level outcomes, species coexistence has been proposed both to be negatively and positively related to ITV (Violle et al., 2012). In two theoretical models of competition with fixed intraspecific variances, Hart et al. (2016) and Barabás and D'Andrea (2016) found that coexistence is (mostly) hampered by ITV. In our framework, both the number of species and ITV are dynamic outcomes, and it is thus not strictly possible to speak causally about how species coexistence affects ITV or vice versa within the framework. As we have done, however, it is possible to compare a model that includes ITV (our framework) to one that does not (adaptive dynamics), and in this sense it seems unlikely

that the inclusion of ITV would yield a higher number of coexisting species. The available trait  
804 space must be apportioned, and if it can be effectively covered by fewer species by incorporating  
ITV, the number of coexisting species will be lower.

In a two-patch version of the Hart et al. (2016) model, Uriarte and Menge (2018) found that  
807 ITV could promote regional species coexistence. This contradicts our findings, and in our two-  
patch model, ITV in the form of local adaptation can preclude the coexistence of species by a  
single species covering more trait space. A crucial difference between our models is that Uriarte  
810 and Menge (2018) assumed that the trait means and variances were fixed on each patch, whereas  
in our models these are outcomes of the dynamics. As for the unstructured case, it is hard  
envisioning a scenario using our model where the inclusion of intraspecific variance would lead  
813 to more coexisting species than the reference model without ITV. A scenario like that of Uriarte  
and Menge (2018) is more likely if trait variation is plastic and driven by the environment as  
opposed to heritable. The nature of ITV could thus be an important factor too in determining  
816 whether ITV can promote coexistence, as has also been observed in an apparent-competition  
model (Schreiber et al., 2011).

Ecosystem functioning has also been shown to depend on the trait distributions of commu-  
819 nities (Mouillot et al., 2011; Gross et al., 2017). Our modeling framework opens up the door  
for theoretical explorations of the relationship between trait variation both within and between  
species and various ecosystem functions. Similarly to coexistence, ecosystem function and the  
822 trait variation are both outcomes of the system's dynamics, so that neither is the strict cause of  
the other. However, this enables explorations of what mechanisms and environments create and  
sustain either, neither, or both of trait diversity and ecosystem functioning.

### 825 *Relationship to other theoretical approaches*

Our framework builds on, and connects to, several other strands of eco-evolutionary theory.  
Considering a single class and setting all mutations to zero, our framework reduces to the  
828 community-ecology framework of Wirtz and Eckhardt (1996) and Norberg et al. (2001), where



the moments of the trait distribution of an ensemble of species is tracked. This approach often includes an immigration term from a fixed species pool to maintain trait variance in the community. Such an immigration term could easily be incorporated into our framework by designating one of the classes a “species-pool class” and letting the internal rates of this class all be zero to keep the species pool fixed. Note that our formulation also includes cases where the species pool too consists of structured populations, so that, for example, juveniles and adults in a stage-structured model could immigrate at different rates from the species pool.

Although the assumptions going into the model at the outset are different, the moment equations derived for our model are also very similar to the equations derived under the assumptions of quantitative genetics. Assuming a single class, no mutations, and that phenotypic variation equals genetic variation (no environmental variation), our equations mirror those of Barabás et al. (2022), who derived equations for mean traits and trait variance-covariance under the assumptions of quantitative genetics for multiple traits. These similarities between trait-space approaches and quantitative genetics have been noted before (e.g., Débarre et al., 2013) and the two approaches have complementary strengths and weaknesses. Most notably, the assumptions of normality are less ad hoc in quantitative genetics, and each species is, as described by a normal distribution, well defined (Turelli and Barton, 1994; Barton et al., 2017). Similarly, the reproductive isolation of species, which in our moment equations is an approximation, is in quantitative genetics based on the biological species concept. This, however, comes at the cost that diversification into multiple peaks cannot be easily incorporated as opposed to our approach here.

Our moment equations also closely resemble those derived in trait-diffusion approaches (Merico et al., 2014; Le Gland et al., 2020), where mutations are generated by a diffusion process in trait space. Merico et al. (2014) derived the moment equations for well-mixed, single-species, single-trait populations and Le Gland et al. (2020) extended this to multiple traits and spatial structure by way of reaction-diffusion equations in continuous space. For mutation kernels with small variance-covariance matrices without covariances, our mutation convolution integral is well approximated by such trait-diffusion processes (Kimura, 1965; Débarre et al., 2013), and our

assembly framework is easily adapted to this setting, equipping the trait-diffusion approaches with a way of assembling eco-evolutionarily stable communities of several species.

858 Finally, our framework also produces similar equations to those in ‘oligomorphic dynam-  
ics’ (Sasaki and Dieckmann, 2011; Débarre et al., 2014; Lion et al., 2022), where a trait-density  
distribution is also decomposed into a number of ‘species’ to track their moments. Compared  
861 to oligomorphic dynamics, the major innovation we present here is our invasion-analysis and  
community-assembly framework, but there are also differences in how the moment equations  
are derived. First, we assume that mutations are associated with some specific ecological pro-  
864 cess (typically births) whereas in oligomorphic dynamics, mutations are assumed to be an inde-  
pendent process. Second, rather than assuming that each species is reproductively isolated, in  
oligomorphic dynamics it is instead assumed that new individuals are allocated to each species  
867 in proportion to each species’ density for given trait. For structured populations, additional as-  
sumptions are required (Lion et al., 2022). These assumptions do however ultimately produce the  
same general shapes for the moment equations as we derived here. Finally, rather than assuming  
870 that each species is normally distributed, oligomorphic dynamics assumes a small-variance ap-  
proximation and can derive equations for an arbitrary set of moments under these assumptions.  
To get a closed system of moment equations a so-called moment-closure approximation is then  
873 used, where using normal distributions is one such possible closure.

Our choice of approximating each species’ trait density as a normal distribution at the outset  
rather than expanding around the mean of each species for an arbitrary number of moments has  
876 both advantages and disadvantages. The advantages are that our moment equations are more  
stable when variances become large compared to a small-variance expansion, and that we can  
more easily integrate mutations into birth processes. The disadvantage is that the flexibility in  
879 shape is more limited. In principle, functions other than normal distributions could be used to  
derive the moment equations. For example, Klauschies et al. (2018) and Cropp and Norbury  
(2021) used beta distributions as the approximating distribution to close their moment equa-  
882 tions, and there are also approaches that use more involved methods for making non-normal

approximations for the trait-space distributions at equilibrium (Mirrahimi and Gandon, 2020). Assumptions other than normality are likely possible while still retaining the core features of our approach, but each such differing assumption would require the re-derivation of nearly all moment equations. While a shortcoming, our comparisons between the trait-space equations and moment equations indicate that as long as normality in the distributions of birthed phenotypes from a parent phenotype is assumed, then our additional assumption that species' trait distributions are normal seem not to affect the accuracy of the moment approximations by any large degree, as our eco-evolutionarily stable communities assembled by moment equations agreed very well with the corresponding trait-space equations (Figs. 2D–F, 3C–F). We note that although we assume normality in each species, the community trait-density distribution is the sum of these normals, which makes the community distribution much more flexible, see Fig. 3D for an example. The primary exception to good accuracy in the moment equations is when the number of species is ambiguous. In these instances the trait-space equations would yield solutions that could not easily be approximated by a sum of normal distributions that were assumed to be reproductively isolated (see Supplementary S1 for examples).

While these drawbacks should be kept in mind, we nevertheless believe that our framework makes substantial progress in eco-evolutionary modeling with intraspecific trait variation. In a phytoplankton model Peeters and Straile (2018) compared trait-space equations without mutations to single-species moment equations, and concluded that single-species moment equations failed to provide any useful information when considering parts of parameter space where the trait-space equations diverged into multiple species. In a similar model using the trait-diffusion approach Le Gland et al. (2020) noted that their trait-space equations sometimes exhibited multimodality and speculated on the utility of modeling multiple modes, making the selection of how many modes to include based on functional groups. While multi-species moment models are not new (Sasaki and Dieckmann, 2011; Norberg et al., 2012; Barabás and D'Andrea, 2016), our assembly approach obviates the need for a-priori decisions on how many modes or species to include by using our assembly process. It also gives an alternate approach for deriving the equations for

moment dynamics for general class-structured populations, compared to Lion et al. (2022).

912 Taken together we believe that our framework of unifying the community-assembly techniques of adaptive dynamics with the moment-equation approach to including intraspecific trait variation could be of great use to theoreticians and modelers seeking to take advantage of facets of both eco-evolutionary modeling frameworks. We also believe that being able to assemble 915 eco-evolutionarily stable communities that accounts for intraspecific trait variation could help address many ecological questions regarding the extent and importance of intra- and interspecific trait variation.

## Online Appendix A: Derivation of moment equations

In this appendix we will derive the generic moment equations (Eqs. 19 in the main text) from the generic trait-space equations (Eqs. 14 in the main text). However, we will here treat a more general case than we do in the main text for which there might be multiple traits under consideration. We will thus consider a *trait vector*  $x$  where each component of the trait vector  $x_1, x_2, \dots, x_n$  describes some property of the organism such as body mass, the propensity to consume a particular resource, or anything that can be encoded by a real number. This means that we will have a *mean-trait vector*  $\bar{x}$ , and a *trait variance-covariance matrix*  $V$ , where, for example,  $V_{22}$  is the variance in trait  $x_2$ , and  $V_{12} = V_{21}$  is the covariance between  $x_1$  and  $x_2$ . This more general formulation reduces to the case we treat in the main text when the trait vector  $x$  is one dimensional.

### *Calculating the time derivatives and general notation*

#### *Time derivatives of the moments*

Let the trait-density distribution of a community be given by  $v(x)$  where  $x \in \mathbb{R}^n$  is a trait vector. Let  $\psi$  and  $\omega$  index the trait components of  $x$ , and let dots over symbols denote time derivatives. We now wish to consider the following moments of the trait-density distribution  $v$ :

$$u = \int v(x) dx \quad \text{Total density} \quad (\text{A.1a})$$

$$\bar{x}_\psi = \frac{1}{u} \int x_\psi v(x) dx \quad \text{Mean-trait vector component } \psi \quad (\text{A.1b})$$

$$V_{\psi\omega} = \frac{1}{u} \int (x - \bar{x})_\psi (x - \psi)_\omega v(x) dx \quad \text{Variance-covariance component } \psi\omega \quad (\text{A.1c})$$

All integrals are over the entirety of  $\mathbb{R}^n$ . Calculating the moment time derivatives then yields:

For the total density  $u$ :

$$\dot{u} = \frac{d}{dt} \left[ \int v(x) dx \right] = \int \dot{v}(x) dx \quad (\text{A.2})$$

For the mean trait component  $\bar{x}_\psi$ :

$$\dot{\bar{x}}_\psi = \frac{d}{dt} \left[ \frac{1}{u} \int x_\psi v(x) dx \right] = -\frac{1}{u^2} \dot{u} \int x_\psi v(x) dx + \frac{1}{u} \int x_\psi \dot{v}(x) dx = \quad (\text{A.3a})$$

$$= -\frac{1}{u} \int \dot{v} dx \bar{x}_\psi + \frac{1}{u} \int x_\psi \dot{v}(x) dx = \frac{1}{u} \int (x - \bar{x})_\psi \dot{v}(x) dx \quad (\text{A.3b})$$

942

$$\iff$$

$$u \dot{\bar{x}}_\psi = \int (x - \bar{x})_\psi \dot{v}(x) dx \quad (\text{A.3c})$$

For the variance-covariance component  $V_{\psi\omega}$ :

945

$$V_{\psi\omega} = \frac{1}{u} \int (x - \bar{x})_\psi (x - \bar{x})_\omega v(x) dx \quad (\text{A.4a})$$

$$\iff$$

$$u V_{\psi\omega} = \int (x - \bar{x})_\psi (x - \bar{x})_\omega v(x) dx \quad (\text{A.4b})$$

948

$$\implies$$

$$\begin{aligned} u \dot{V}_{\psi\omega} + V_{\psi\omega} \dot{u} &= -\dot{\bar{x}}_\psi \int (x - \bar{x})_\omega v(x) dx - \dot{\bar{x}}_\omega \int (x - \bar{x})_\psi v(x) dx \\ &\quad + \int (x - \bar{x})_\psi (x - \bar{x})_\omega \dot{v}(x) dx \end{aligned} \quad (\text{A.4c})$$

951

$$= \int (x - \bar{x})_\psi (x - \bar{x})_\omega \dot{v}(x) dx \quad (\text{A.4d})$$

Together we thus have

$$\dot{u} = \int \dot{v}(x) dx \quad (\text{A.5a})$$

954

$$u \dot{\bar{x}}_\psi = \int (x - \bar{x})_\psi \dot{v}(x) dx \quad (\text{A.5b})$$

$$u \dot{V}_{\psi\omega} + V_{\psi\omega} \dot{u} = \int (x - \bar{x})_\psi (x - \bar{x})_\omega \dot{v}(x) dx \quad (\text{A.5c})$$

This way of expressing the relationships between the time derivatives of the moments and the time derivative of the trait distribution makes it easier to derive the moment equations for the various models, as we can plug in the right-hand side of any equation in place of  $\dot{v}$  in Eqs. A.5.

### *A note on notation*

The calculations for deriving the moment equations are not difficult in the sense that they require any complicated mathematical concepts or techniques. They are however rather complex in terms of notation and book keeping. To alleviate this problem, we introduce some useful notation.

963 First, we frequently employ so-called Einstein summation notation. This means that, unless  
 otherwise indicated, repeated indices are summed over, so that for example

$$A_{\alpha\beta}B_{\beta\gamma} := \sum_{\beta} A_{\alpha\beta}B_{\beta\gamma}. \quad (\text{A.6})$$

966 Second, we let indices on scalar quantities denote components of gradients, i.e., if  $f$  is a function  
 from  $\mathbb{R}^n$  to  $\mathbb{R}$ , then

$$f_{\alpha} = \frac{\partial f(x)}{\partial x_{\alpha}}, \quad \alpha \in \{1, 2, \dots, n\} \quad (\text{A.7})$$

969 Thus, for example, in index notation we would write  $\nabla_x g$  as  $g_{\alpha}$ . Combining these two notational  
 conventions, we can for example write the matrix–vector product  $V\nabla_x g$  as  $V_{\psi\alpha}g_{\alpha}$  and the matrix–  
 matrix–matrix product  $V\nabla_x^2 g V$  as  $V_{\psi\alpha}g_{\alpha\beta}V_{\beta\omega}$ .

972 We will use Greek letters to indicate trait components. For these indices the Einstein summa-  
 tion convention will always apply. We will use Latin indices  $i$  and  $j$  for species, and  $k$  and  $l$  for  
 classes, and finally  $m$  for processes. The Einstein summation convention will not apply to these  
 975 indices.

### *Deriving the moment equations*

To make it easier to derive the moment equations from the trait-space equations, we will first set  
 978 out to derive some intermediate results.

First, let the probability-density function of an  $n$ -dimensional multivariate normal distribu-  
 tion be denoted by

$$981 \quad \mathcal{N}(x, \bar{x}, V) = \frac{1}{\sqrt{(2\pi)^n \det(V)}} \exp\left(-\frac{1}{2}(x - \bar{x})^T V^{-1}(x - \bar{x})\right), \quad (\text{A.8})$$

where  $x$  is the argument vector,  $\bar{x}$  is the mean vector, and  $V$  is the variance-covariance matrix  
 of the multivariate normal distribution, and  $T$  denotes matrix transposition. We can then cal-  
 984 culate the first and second derivatives of this normal distribution with respect to its argument  
 components:

$$\mathcal{N}_{\psi} = \frac{\partial \mathcal{N}}{\partial x_{\psi}} = -V_{\psi\alpha}^{-1}(x - \bar{x})_{\alpha} \mathcal{N}(x, \bar{x}, V) \quad (\text{A.9a})$$

$$\mathcal{N}_{\psi\omega} = \frac{\partial^2 \mathcal{N}}{\partial x_\psi \partial x_\omega} = \left[ V_{\psi\alpha}^{-1} (x - \bar{x})_\alpha V_{\omega\beta}^{-1} (x - \bar{x})_\beta - V_{\psi\omega}^{-1} \right] \mathcal{N}(x, \bar{x}, V) \quad (\text{A.9b})$$

Furthermore, if  $f(y)$  is the rate of an ecological process for an individual with trait  $y$ , let the population level per capita rate for a population with mean-trait vector  $x$  and variance-covariance matrix  $W$  be given by

$$\hat{f}(x, W) = \int f(y) \mathcal{N}(y, x, W) dy. \quad (\text{A.10})$$

We can now differentiate the population-level rate  $\hat{f}$  with respect to the mean-trait components:

$$\hat{f}_\psi = \frac{\partial \hat{f}}{\partial x_\psi} = W_{\psi\alpha}^{-1} \int (y - x)_\alpha f(y) \mathcal{N}(y, x, W) dy \quad (\text{A.11a})$$

$$\hat{f}_{\psi\omega} = \frac{\partial^2 \hat{f}}{\partial x_\psi \partial x_\omega} = \int \left[ W_{\psi\alpha}^{-1} (y - x)_\alpha W_{\omega\beta}^{-1} (x - y)_\beta - W_{\psi\omega}^{-1} \right] f(y) \mathcal{N}(y, x, W) dy \quad (\text{A.11b})$$

Using these results we can calculate two integrals that will appear later in our derivation:

$$\int (x - \bar{x})_\psi f(x) \mathcal{N}(x, \bar{x}, V) dx = V_{\psi\alpha} V_{\alpha\beta}^{-1} \int (x - \bar{x})_\beta f(x) \mathcal{N}(x, \bar{x}, V) dx \quad (\text{A.12a})$$

$$= V_{\psi\alpha} \hat{f}_\alpha(\bar{x}, V) \quad (\text{A.12b})$$

and

$$\int (x - \bar{x})_\psi (x - \bar{x})_\omega f(x) \mathcal{N}(x, \bar{x}, V) dx \quad (\text{A.13a})$$

$$= V_{\psi\alpha} V_{\omega\beta} \int V_{\alpha\gamma}^{-1} (x - \bar{x})_\gamma V_{\beta\epsilon}^{-1} (x - \omega)_\epsilon f(x) \mathcal{N}(x, \bar{x}, V) dx \quad (\text{A.13b})$$

$$= V_{\psi\alpha} V_{\omega\beta} \int [V_{\alpha\gamma}^{-1} (x - \bar{x})_\gamma V_{\beta\epsilon}^{-1} (x - \omega)_\epsilon - V_{\alpha\beta}^{-1} + V_{\alpha\beta}^{-1}] f(x) \mathcal{N}(x, \bar{x}, V) dx \quad (\text{A.13c})$$

$$= V_{\psi\alpha} \hat{f}_{\alpha\beta}(\bar{x}, V) V_{\beta\omega} + \hat{f}(\bar{x}, V) V_{\psi\omega} \quad (\text{A.13d})$$

where we have used that  $V$  is symmetric so that  $V_{\psi\omega} = V_{\omega\psi}$ .

Assuming that we have a (multi-)normally distributed trait-density distribution  $v(x) := u \mathcal{N}(x, \bar{x}, V)$  with total density  $u$ , mean  $\bar{x}$ , and variance-covariance  $V$  we can use the above results in turn to calculate three integrals that will appear in the derivation of the moment equations:

$$\int f(x) v(x) dx = \int f(x) u \mathcal{N}(x, \bar{x}, V) dx = \hat{f}(\bar{x}, V) u \quad (\text{A.14})$$



$$\int (x - \bar{x})_\psi f(x) v(x) dx = \int (x - \bar{x})_\psi f(x) u \mathcal{N}(x, \bar{x}, V) dx \quad (\text{A.15a})$$

$$= u V_{\psi\alpha} \hat{f}_\alpha(\bar{x}, V) \quad (\text{A.15b})$$

$$\int (x - \bar{x})_\psi (x - \bar{x})_\omega f(x) v(x) dx = \int (x - \bar{x})_\psi (x - \bar{x})_\omega f(x) u \mathcal{N}(x, \bar{x}, V) dx \quad (\text{A.16a})$$

$$= u \left( V_{\psi\alpha} \hat{f}_{\alpha\beta}(\bar{x}, V) V_{\beta\omega} + \hat{f}(\bar{x}, V) V_{\psi\omega} \right) \quad (\text{A.16b})$$

In the trait-space equations, we also need to evaluate the integrals over the normally distributed mutation kernels. In total, three such integrals will need to be evaluated, the first of which is:

$$\int \int f(y) v(y) \mathcal{N}(y, x, M) dy dx = \quad (\text{A.17a})$$

$$= \int f(y) v(y) \int \mathcal{N}(y, x, M) dx dy \quad (\text{A.17b})$$

$$= \int f(y) v(y) dy = \{\text{Eq. A.14}\} = \hat{f}(\bar{x}, V) u \quad (\text{A.17c})$$

The second integral is

$$\int (x - \bar{x})_\psi \int f(y) v(y) \mathcal{N}(y, x, M) dy dx = \int f(y) v(y) \int (x - \bar{x})_\psi \mathcal{N}(y, x, M) dx dy \quad (\text{A.18a})$$

$$= \int f(y) v(y) \int [(x - y)_\psi + (y - \bar{x})_\psi] \mathcal{N}(x, y, M) dx dy \quad (\text{A.18b})$$

$$= \int (y - \bar{x})_\psi f(y) v(y) dy = \{\text{Eqs. A.15}\} = u V_{\psi\alpha} \hat{f}_\alpha(\bar{x}, V) \quad (\text{A.18c})$$

The third integral is

$$\int (x - \bar{x})_\psi (x - \bar{x})_\omega \int f(y) v(y) \mathcal{N}(y, x, M) dy dx \quad (\text{A.19a})$$

$$= \int f(y) v(y) \int (x - \bar{x})_\psi (x - \bar{x})_\omega \mathcal{N}(x, y, M) dx dy \quad (\text{A.19b})$$

$$= \int f(y) v(y) \int [(x - y)_\psi (x - y)_\omega + (y - \bar{x})_\psi (x - y)_\omega \quad (\text{A.19c})$$

$$+ (x - y)_\psi (y - \bar{x})_\omega + (y - \bar{x})_\psi (y - \bar{x})_\omega] \mathcal{N}(x, y, M) dx dy \quad (\text{A.19d})$$

1029 
$$= M_{\psi\omega} \int f(\mathbf{y})v(\mathbf{y})d\mathbf{y} + \int (\mathbf{y} - \bar{x})_{\psi}(\mathbf{y} - \bar{x})_{\omega} f(\mathbf{y})v(\mathbf{y})d\mathbf{y} \quad (\text{A.19e})$$

$$= \{\text{Eq. A.14 and Eqs. A.16}\} \quad (\text{A.19f})$$

$$= u \left( M_{\psi\omega} \hat{f}(\bar{x}, V) + V_{\psi\alpha} \hat{f}_{\alpha\beta}(\bar{x}, V) V_{\beta\omega} + \hat{f}(\bar{x}, V) V_{\psi\omega} \right) \quad (\text{A.19g})$$

1032 We will here formulate a more explicit version of the trait-space and moment equations  
 (Eqs. 14 and 19 respectively in the main text). We assume that there are  $K$  classes, that  $v_k(x)$   
 is the trait-density distribution in class  $k$ , and that for each class pair there are  $N_{kl}$  processes  
 1035 where  $l$  is the source class and  $k$  is the destination class. We index each such process by  $f_{klm}(x, \mathbf{v})$ ,  
 $m \in \{1, \dots, N_{kl}\}$ , where  $\mathbf{v} := (v_1, v_2, \dots, v_K)$ . For each process we let  $M_{klm}$  be the mutation variance-  
 covariance matrix associated with that process, which for most processes will be  $M_{klm} = 0$ . By  
 1038 summing over all processes the trait-space equations are then given by

$$\frac{dv_k(x)}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \int f_{klm}(\mathbf{y})v_l(\mathbf{y})\mathcal{N}(x, \mathbf{y}, M_{klm})d\mathbf{y}. \quad (\text{A.20})$$

Comparing with the trait-space equations in the main text (Eq. 14), we here make explicit the  
 1041 summation over all the processes that we left implicit in the main text to obviate the need for  
 excessive notational clutter. For the purposes of deriving the moment equations however, it is  
 clearer to leave these summations in explicitly.

1044 We now make the assumptions that the trait-density distribution in each class  $v_k(x)$  can be  
 decomposed into a sum of component distributions (species)

$$v_k(x) = \sum_{i=1}^S v_{ik}(x), \quad (\text{A.21})$$

1047 that each such component distribution can be approximated with a normal distribution

$$v_{ik}(x) \approx \tilde{v}_{ik}(x) := u_{ik}\mathcal{N}(x, \bar{x}_{ik}, V_{ik}) \quad (\text{A.22})$$

and that each species is reproductively isolated so that each species can be treated separately.  
 1050 Using the results derived above we can now calculate the moment equations for these trait-space  
 equations.

For  $u$ :

$$1053 \quad \frac{du_{ik}}{dt} = \int \dot{v}_{ik}(x) dx \approx \int \dot{\check{v}}_{ik}(x) dx = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \int \int f_{klm}(y) \tilde{v}_{il}(y) \mathcal{N}(x, y, M_{klm}) dy dx \quad (\text{A.23a})$$

$$= \{\text{Eqs. A.17}\} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \hat{f}_{klm}(\bar{x}_{il}, V_{il}) u_{il} \quad (\text{A.23b})$$

For  $\bar{x}$ :

$$1056 \quad u_{ik} \frac{d\bar{x}_{ik\psi}}{dt} = \int (x - \bar{x}_{ik})_{\psi} \dot{v}_{ik}(x) dx \approx \int (x - \bar{x}_{ik})_{\psi} \dot{\check{v}}_{ik}(x) dx \quad (\text{A.24a})$$

$$= \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \int (x - \bar{x}_{ik})_{\psi} \int f_{klm}(y) \tilde{v}_{il}(y) \mathcal{N}(x, y, M_{klm}) dy dx \quad (\text{A.24b})$$

$$= \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \int [(x - \bar{x}_{il})_{\psi} + (\bar{x}_{il} - \bar{x}_{ik})_{\psi}] \int f_{klm}(y) \tilde{v}_{il}(y) \mathcal{N}(x, y, M_{klm}) dy dx \quad (\text{A.24c})$$

$$1059 \quad = \{\text{Eqs. A.17 and Eqs. A.18}\}$$

$$= \sum_{l=1}^K \sum_{m=1}^{N_{kl}} u_{il} \left( V_{\psi\alpha} \hat{f}_{klm\alpha}(\bar{x}_{il}, V_{il}) + \hat{f}_{klm}(\bar{x}_{il}, V_{il}) (\bar{x}_{il} - \bar{x}_{ik})_{\psi} \right) \quad (\text{A.24d})$$

$\iff$

$$1062 \quad \frac{d\bar{x}_{ik\psi}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left( V_{\psi\alpha} \hat{f}_{klm\alpha}(\bar{x}_{il}, V_{il}) + \hat{f}_{klm}(\bar{x}_{il}, V_{il}) (\bar{x}_{il} - \bar{x}_{ik})_{\psi} \right) \quad (\text{A.24e})$$

For  $V$ :

$$u_{ik} \frac{dV_{ik\psi\omega}}{dt} + V_{ik\psi\omega} \frac{du_{ik}}{dt} = \int (x - \bar{x}_{ik})_{\psi} (x - \bar{x}_{ik})_{\omega} \dot{v}_{ik}(x) dx \quad (\text{A.25a})$$

$$1065 \quad \approx \int (x - \bar{x}_{ik})_{\psi} (x - \bar{x}_{ik})_{\omega} \dot{\check{v}}_{ik}(x) dx \quad (\text{A.25b})$$

$$= \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \int (x - \bar{x}_{ik})_{\psi} (x - \bar{x}_{ik})_{\omega} \int f_{klm}(y) \tilde{v}_{il}(y) \mathcal{N}(x, y, M_{klm}) dy dx \quad (\text{A.25c})$$

$$= \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \int [(x - \bar{x}_{il})_{\psi} (x - \bar{x}_{il})_{\omega} + (\bar{x}_{il} - \bar{x}_{ik})_{\psi} (x - \bar{x}_{il})_{\omega}$$

$$1068 \quad + (x - \bar{x}_{il})_{\psi} (\bar{x}_{il} - \bar{x}_{ik})_{\omega} + (\bar{x}_{il} - \bar{x}_{ik})_{\psi} (\bar{x}_{il} - \bar{x}_{ik})_{\omega}] \int f_{klm}(y) \tilde{v}_{il}(y) \mathcal{N}(x, y, M_{klm}) dy dx \quad (\text{A.25d})$$

$$\begin{aligned}
&= \{\text{Eqs. A.17, Eqs. A.18, and Eqs. A.19}\} \\
&= \sum_{l=1}^K \sum_{m=1}^{N_{kl}} u_{il} \left[ M_{klm\psi\omega} \hat{f}_{klm}(\bar{x}_{il}, V_{il}) + V_{il\psi\alpha} \hat{f}_{klm\alpha\beta}(\bar{x}_{il}, V_{il}) V_{il\beta\omega} + \hat{f}_{klm}(\bar{x}_{il}, V_{il}) V_{il\psi\omega} \right. \\
1071 \quad &+ V_{il\psi\alpha} \hat{f}_{klm\alpha}(\bar{x}_{il}, V_{il})(\bar{x}_{il} - \bar{x}_{ik})_{\omega} + V_{il\omega\alpha} \hat{f}_{klm\alpha}(\bar{x}_{il}, V_{il})(\bar{x}_{il} - \bar{x}_{ik})_{\psi} \\
&\left. + \hat{f}_{klm}(\bar{x}_{il}, V_{il})(\bar{x}_{il} - \bar{x}_{ik})_{\psi}(\bar{x}_{il} - \bar{x}_{ik})_{\omega} \right] \tag{A.25e}
\end{aligned}$$

$\Leftrightarrow$

$$\begin{aligned}
1074 \quad \frac{dV_{ik\psi\omega}}{dt} &= \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left[ M_{klm\psi\omega} \hat{f}_{klm}(\bar{x}_{il}, V_{il}) + V_{il\psi\alpha} \hat{f}_{klm\alpha\beta}(\bar{x}_{il}, V_{il}) V_{il\beta\omega} \right. \\
&+ \hat{f}_{klm}(\bar{x}_{il}, V_{il})(V_{il\psi\omega} - V_{ik\psi\omega}) \\
&+ V_{il\psi\alpha} \hat{f}_{klm\alpha}(\bar{x}_{il}, V_{il})(\bar{x}_{il} - \bar{x}_{ik})_{\omega} + V_{il\omega\alpha} \hat{f}_{klm\alpha}(\bar{x}_{il}, V_{il})(\bar{x}_{il} - \bar{x}_{ik})_{\psi} \\
1077 \quad &\left. + \hat{f}_{klm}(\bar{x}_{il}, V_{il})(\bar{x}_{il} - \bar{x}_{ik})_{\psi}(\bar{x}_{il} - \bar{x}_{ik})_{\omega} \right] \tag{A.25f}
\end{aligned}$$

The complete moment equations for  $S$  species in  $K$  classes thus read

$$\frac{du_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \hat{f}_{iklm} u_{il} \tag{A.26a}$$

$$1080 \quad \frac{d\bar{x}_{ik\psi}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left( V_{il\psi\alpha} \hat{f}_{iklm\alpha} + \hat{f}_{iklm}(\bar{x}_{il} - \bar{x}_{ik})_{\psi} \right) \tag{A.26b}$$

$$\frac{dV_{ik\psi\omega}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left[ V_{il\psi\alpha} \hat{f}_{iklm\alpha\beta} V_{il\beta\omega} + \hat{f}_{iklm}(V_{il\psi\omega} - V_{ik\psi\omega}) + M_{klm\psi\omega} \hat{f}_{iklm} \right. \tag{A.26c}$$

$$+ V_{il\psi\alpha} \hat{f}_{iklm\alpha}(\bar{x}_{il} - \bar{x}_{ik})_{\omega} + V_{il\omega\alpha} \hat{f}_{iklm\alpha}(\bar{x}_{il} - \bar{x}_{ik})_{\psi} \tag{A.26d}$$

$$1083 \quad \left. + \hat{f}_{iklm}(\bar{x}_{il} - \bar{x}_{ik})_{\psi}(\bar{x}_{il} - \bar{x}_{ik})_{\omega} \right] \tag{A.26e}$$

$$\hat{f}_{iklm} := \hat{f}_{klm}(\bar{x}_{il}, V_{il}, \tilde{\mathbf{v}}) \tag{A.26f}$$

$$\hat{f}_{iklm\alpha} := \frac{\partial \hat{f}_{klm}}{\partial x_{\alpha}}(\bar{x}_{il}, V_{il}, \tilde{\mathbf{v}}) \tag{A.26g}$$

$$1086 \quad \hat{f}_{iklm\alpha\beta} := \frac{\partial^2 \hat{f}_{klm}}{\partial x_{\alpha} \partial x_{\beta}}(\bar{x}_{il}, V_{il}, \tilde{\mathbf{v}}) \tag{A.26h}$$

Here, we have omitted the explicit dependence of the process rates on its arguments for notational clarity. We have also added the dependence of the process rates  $\hat{f}_{klm}$  on the approximate trait-density distributions  $\tilde{\mathbf{v}} = (\tilde{v}_1, \dots, \tilde{v}_K)$  where  $\tilde{v}_k(x) = \sum_{i=1}^S u_{ik} \mathcal{N}(x, \bar{x}_{ik}, V_{ik})$ . For convenience we

did not include this dependence during the derivation, as their inclusion does not affect the derivation. We do however make one note with regards to this. When we define the population-

1092 level rate

$$\hat{f}_{klm}(x, W, \tilde{\mathbf{v}}) = \int f_{klm}(y, \tilde{\mathbf{v}}) \mathcal{N}(y, x, W) dy \quad (\text{A.27})$$

it technically describes the population-level per capita rate of a normally distributed population  
 1095 with mean-trait vector  $x$  and variance-covariance matrix  $W$  that is negligibly rare in the resident  
 community whose trait-density distribution is given by  $\tilde{\mathbf{v}}$  across all classes. In other words,  
 the rare population with mean  $x$  and variance-covariance  $W$  does not add to the trait-density  
 1098 distribution  $\tilde{\mathbf{v}}$ . This is analogous to how invasion fitness for a rare invader is defined in adaptive  
 dynamics (Metz et al., 1992; Geritz et al., 1998). This notation helps us get frequency-dependent  
 interactions (such as the competition in our Lotka-Volterra model examples) correct when we  
 1101 differentiate the population-level rates. Note also that while we in the appendices denote the rare  
 invader's mean trait and variance by  $x$  and  $W$  for reasons of separating these abstract quantities  
 of the invader from the realized quantities of an extant resident population, in the main text we  
 1104 instead use  $\bar{x}$  and  $V$  for the rare invader population to emphasize that they are the means and  
 variance respectively.

It can be easier to interpret these equations when not written in index notation, and so we  
 1107 let  $\nabla_x f$  denote the gradient vector of  $f$  with respect to  $x$ , and let  $\nabla_x^2 f$  denote the Hessian matrix  
 of  $f$  with respect to  $x$ , and rewrite the equations in the language of vectors and matrices which  
 yields

$$\frac{du_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \hat{f}_{iklm} u_{il} \quad (\text{A.28a})$$

$$\frac{d\bar{x}_{ik\psi}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left( V_{il} \nabla_x \hat{f}_{iklm} + \hat{f}_{iklm} (\bar{x}_{il} - \bar{x}_{ik}) \right) \quad (\text{A.28b})$$

$$\frac{dV_{ik\psi\omega}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left[ V_{il} \nabla_x^2 \hat{f}_{iklm} V_{il} + \hat{f}_{iklm} (V_{il} - V_{ik}) + M_{klm} \hat{f}_{iklm} \right. \quad (\text{A.28c})$$

$$\left. + V_{il} \nabla_x \hat{f}_{iklm} (\bar{x}_{il} - \bar{x}_{ik})^T + (\bar{x}_{il} - \bar{x}_{ik}) \left[ V_{il} \nabla_x \hat{f}_{iklm} \right]^T \right. \quad (\text{A.28d})$$

$$\left. + \hat{f}_{iklm} (\bar{x}_{il} - \bar{x}_{ik}) (\bar{x}_{il} - \bar{x}_{ik})^T \right] \quad (\text{A.28e})$$

$$\hat{f}_{iklm} := \hat{f}_{iklm}(\bar{x}_{il}, V_{il}, \tilde{\mathbf{v}}) \quad (\text{A.28f})$$

1116

$$\nabla_x \hat{f}_{iklm} := \nabla_x \hat{f}_{iklm}(\bar{x}_{il}, V_{il}, \tilde{\mathbf{v}}) \quad (\text{A.28g})$$

$$\nabla_x^2 \hat{f}_{iklm} := \nabla_x^2 \hat{f}_{iklm}(\bar{x}_{il}, V_{il}, \tilde{\mathbf{v}}) \quad (\text{A.28h})$$

where  $T$  denotes vector transpose, and all multiplication is taken to be matrix multiplication.

1119

When the trait is scalar, i.e,  $x \in \mathbb{R}$ , these equations simplify to

$$\frac{du_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \hat{f}_{iklm} u_{il} \quad (\text{A.29a})$$

$$\frac{d\bar{x}_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left( V_{il} \partial_x \hat{f}_{iklm} + \hat{f}_{iklm} (\bar{x}_{il} - \bar{x}_{ik}) \right) \quad (\text{A.29b})$$

1122

$$\frac{dV_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left[ V_{il}^2 \partial_x^2 \hat{f}_{iklm} + \hat{f}_{iklm} (V_{il} - V_{ik}) + M_{klm} \hat{f}_{iklm} \right. \quad (\text{A.29c})$$

$$\left. + 2V_{il} \partial_x \hat{f}_{iklm} (\bar{x}_{il} - \bar{x}_{ik}) + \hat{f}_{iklm} (\bar{x}_{il} - \bar{x}_{ik})^2 \right] \quad (\text{A.29d})$$

$$\hat{f}_{iklm}(x, W, \tilde{\mathbf{v}}) = \int f(y, \tilde{\mathbf{v}}) \mathcal{N}(y, x, W) dy \quad (\text{A.29e})$$

1125

$$\hat{f}_{iklm} := \hat{f}_{iklm}(\bar{x}_{il}, V_{il}, \tilde{\mathbf{v}}) \quad (\text{A.29f})$$

$$\partial_x \hat{f}_{iklm} := \frac{\partial \hat{f}_{iklm}}{\partial x}(\bar{x}_{il}, V_{il}, \tilde{\mathbf{v}}) \quad (\text{A.29g})$$

$$\partial_x^2 \hat{f}_{iklm} := \frac{\partial^2 \hat{f}_{iklm}}{\partial x^2}(\bar{x}_{il}, V_{il}, \tilde{\mathbf{v}}) \quad (\text{A.29h})$$

1128

Finally, we can describe the double summation over source classes  $l$  and processes  $m$  implicitly by letting  $d$  be the index for destination classes,  $s$  be the index for source classes,  $f$  be any process to be summed over, and  $M$  be its associated mutation variance, and  $\pm$  be the operator that describes

1131

addition to. This then yields

$$\frac{du_{id}}{dt} \pm \hat{f}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}}) u_{is}, \quad (\text{A.30a})$$

$$\frac{d\bar{x}_{id}}{dt} \pm \frac{u_{is}}{u_{id}} \left[ V_{is} \frac{\partial \hat{f}}{\partial \bar{x}}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}}) + \hat{f}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}}) (\bar{x}_{is} - \bar{x}_{id}) \right], \quad (\text{A.30b})$$

1134

$$\begin{aligned} \frac{dV_{id}}{dt} \pm \frac{u_{is}}{u_{id}} \left[ V_{is}^2 \frac{\partial^2 \hat{f}}{\partial \bar{x}^2}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}}) + \hat{f}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}}) (V_{is} - V_{id}) \right. \\ \left. + \hat{f}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}}) (\bar{x}_{is} - \bar{x}_{id})^2 + 2V_{is} \frac{\partial \hat{f}}{\partial \bar{x}}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}}) (\bar{x}_{is} - \bar{x}_{id}) \right] \end{aligned}$$

$$+ \hat{f}(\bar{x}_{is}, V_{is}, \tilde{\mathbf{v}})M \Big], \quad (\text{A.30c})$$

$$\hat{f}(\bar{x}, V, \tilde{\mathbf{v}}) = \int f(x, \tilde{\mathbf{v}}) \mathcal{N}(x, \bar{x}, V) dx. \quad (\text{A.30d})$$

These correspond to Eq. 19 in the main text.

### *Series solutions and Taylor approximations*

The population per capita rates

$$\hat{f}(x, W) = \int f(y) \mathcal{N}(y, x, W) dy \quad (\text{A.31})$$

are in most cases not going to be analytically computable. Owen (1980) contains a large list of Gaussian integrals. Here, we briefly note that for example exponential functions, Gaussian functions, and polynomials can be analytically integrated against a Gaussian, whereas, for example, a Type-II functional response of an organism with trait  $x \in \mathbb{R}$  consuming a resource  $R$  where the affinities depend on the trait  $x$

$$f(x) = \frac{a(x)R}{1 + a(x)hR} \quad (\text{A.32})$$

has, to wit, no analytically closed expression for simple forms for  $a(x)$  such as linear or exponential when the functional response is integrated against the Gaussian.

For the cases where the population-rate integrals cannot be evaluated directly, we may Taylor expand the process rates. To this end let  $\alpha^n$  denote an ordered set of indices such that

$$\alpha^n := \alpha_1 \alpha_2 \cdots \alpha_n \quad (\text{A.33})$$

where each  $\alpha_i$  indexes the trait-components. Thus, for example

$$f_{\alpha^4} = f_{\alpha_1 \alpha_2 \alpha_3 \alpha_4} \quad (\text{A.34})$$

denotes the 4-tensor of mixed partial derivatives of order 4 of  $f$ . For a vector  $x$  in trait space we let

$$x^{\alpha^n} := x_{\alpha_1} x_{\alpha_2} \cdots x_{\alpha_n} \quad (\text{A.35})$$

1158 denote the outer product of  $n$  copies of the vector. For a matrix  $W$  acting on trait space we also  
 1161 let

$$W^{\alpha^{2n}} := W_{\alpha_1\alpha_2} W_{\alpha_3\alpha_4} \cdots W_{\alpha_{2n-1}\alpha_{2n}} \quad (\text{A.36})$$

1161 denote the outer product of  $n$  copies of the matrices.

We can now Taylor-expand

$$\hat{f}(x, W) = \int f(y) \mathcal{N}(y, x, W) dy = \int \sum_{n=0}^{\infty} \frac{f_{\alpha^n}(x)}{n!} (y-x)^{\alpha^n} \mathcal{N}(y, x, W) dy \quad (\text{A.37a})$$

1164 The central moments of a multivariate normal distribution are  $n$ -tensors and are given in e.g.,  
 Triantafyllopoulos (2002). However, when contracted against the fully symmetric tensor  $f_{\alpha^n}$  these  
 expressions simplify considerably so that

$$1167 \int f_{\alpha^n}(x) (y-x)^{\alpha^n} \mathcal{N}(y, x, W) dy = \begin{cases} 0 & \text{if } n \text{ is odd} \\ (n-1)!! f_{\alpha^n} W^{\alpha^n} & \text{if } n \text{ is even} \end{cases} \quad (\text{A.38})$$

where “!!” is the double factorial, which is defined by

$$n!! = n(n-2)(n-4) \cdots 1, \quad 0!! = 1, \quad (-1)!! = 1 \quad (\text{A.39})$$

1170 For the Taylor expansion we thus have

$$\int \sum_{n=0}^{\infty} \frac{f_{\alpha^n}(x)}{n!} (y-x)^{\alpha^n} \mathcal{N}(y, x, W) dy = \sum_{\substack{n=0 \\ n \text{ even}}}^{\infty} \frac{f_{\alpha^n}(x)}{n!} (n-1)!! W^{\alpha^n} \quad (\text{A.40a})$$

$$= \sum_{n=0}^{\infty} \frac{f_{\alpha^{2n}}(x)}{(2n)!} (2n-1)!! W^{\alpha^{2n}} = \sum_{n=0}^{\infty} \frac{f_{\alpha^{2n}}(x)}{2^n n!} W^{\alpha^{2n}} \quad (\text{A.40b})$$

1173 and hence

$$\hat{f}(x, W) = \sum_{n=0}^{\infty} \frac{f_{\alpha^{2n}}(x)}{2^n n!} W^{\alpha^{2n}} = f(x) + \frac{1}{2} f_{\alpha\beta}(x) W_{\alpha\beta} + \frac{1}{8} f_{\alpha\beta\gamma\epsilon}(x) W_{\alpha\beta} W_{\gamma\epsilon} + \dots \quad (\text{A.41a})$$

$$\hat{f}_{\psi}(x, W) = \sum_{n=0}^{\infty} \frac{f_{\alpha^{2n}\psi}(x)}{2^n n!} W^{\alpha^{2n}} = f_{\psi}(x) + \frac{1}{2} f_{\psi\alpha\beta}(x) W_{\alpha\beta} + \frac{1}{8} f_{\psi\alpha\beta\gamma\epsilon}(x) W_{\alpha\beta} W_{\gamma\epsilon} + \dots \quad (\text{A.41b})$$

$$1176 \hat{f}_{\psi\omega}(x, W) = \sum_{n=0}^{\infty} \frac{f_{\alpha^{2n}\psi\omega}(x)}{2^n n!} W^{\alpha^{2n}} = f_{\psi\omega}(x) + \frac{1}{2} f_{\psi\omega\alpha\beta}(x) W_{\alpha\beta} + \frac{1}{8} f_{\psi\omega\alpha\beta\gamma\epsilon}(x) W_{\alpha\beta} W_{\gamma\epsilon} + \dots \quad (\text{A.41c})$$



A Taylor expansion of order  $n$  then corresponds to truncating to derivatives of  $f$  with no higher order than  $n$ , so that e.g., a Taylor-approximation of order three would yield

$$1179 \quad \hat{f}(x, W) = \sum_{n=0}^{\infty} \frac{f_{\alpha^{2n}}(x)}{2^n n!} W^{\alpha^{2n}} \approx f(x) + \frac{1}{2} f_{\alpha\beta}(x) W_{\alpha\beta} \quad (\text{A.42a})$$

$$\hat{f}_{\psi}(x, W) = \sum_{n=0}^{\infty} \frac{f_{\alpha^{2n}\psi}(x)}{2^n n!} W^{\alpha^{2n}} \approx f_{\psi}(x) + \frac{1}{2} f_{\psi\alpha\beta}(x) W_{\alpha\beta} \quad (\text{A.42b})$$

$$\hat{f}_{\psi\omega}(x, W) = \sum_{n=0}^{\infty} \frac{f_{\alpha^{2n}\psi\omega}(x)}{2^n n!} W^{\alpha^{2n}} \approx f_{\psi\omega}(x) \quad (\text{A.42c})$$

1182 For a scalar trait, let  $f^{(n)}$  denote the  $n$ th derivative of  $f$  with respect to  $x$ . The expressions then simplify to

$$\hat{f}(x, W) = \sum_{n=0}^{\infty} \frac{f^{(2n)}(x)}{2^n n!} W^n = f(x) + \frac{1}{2} f^{(2)}(x) W + \frac{1}{8} f^{(4)}(x) W^2 + \dots \quad (\text{A.43a})$$

$$1185 \quad \hat{f}^{(1)}(x, W) = \sum_{n=0}^{\infty} \frac{f^{(2n+1)}(x)}{2^n n!} W^n = f^{(1)}(x) + \frac{1}{2} f^{(3)}(x) W + \frac{1}{8} f^{(5)}(x) W^2 + \dots \quad (\text{A.43b})$$

$$\hat{f}^{(2)}(x, W) = \sum_{n=0}^{\infty} \frac{f^{(2n+2)}(x)}{2^n n!} W^n = f^{(2)}(x) + \frac{1}{2} f^{(4)}(x) W + \frac{1}{8} f^{(6)}(x) W^2 + \dots \quad (\text{A.43c})$$

### *Variable mutation variance-covariance matrices and mutation bias*

1188 For completeness we here presents result of more complicated mutation dynamics where the mutation variance-covariance matrix is allowed to depend on the trait of the parent so that some phenotypes generate a wider array of offspring than others in trait space, and where mutations  
1191 are not symmetric so that there is a bias of mutations in some direction in trait space.

First, we let the mutation variance-covariance matrix vary with trait so that  $M(y)$  is a positive-definite matrix for all  $y$ . The mutation kernel is then given by  $\mathcal{N}(x, y, M(y))$ . Note that for each  
1194 fixed  $y$ , this is just the probability density function of a normal distribution over  $x$ .

Working through the problem in the same way as we did for constant mutation matrices (Eqs. A.17–A.18) we can then evaluate the three integrals used to account for mutation, the first  
1197 of which is:

$$\int \int f(y) \mathcal{N}(y, \bar{x}, V)(y, x, M(y)) dy dx = \int f(y) \mathcal{N}(y, \bar{x}, V) \int \mathcal{N}(y, x, M(y)) dx dy \quad (\text{A.44a})$$

$$= \int f(y)\mathcal{N}(y, \bar{x}, V)dy = \int f(x)\mathcal{N}(x, \bar{x}, V)dx = \hat{f}(\bar{x}, V) \quad (\text{A.44b})$$

1200 The second integral is

$$\int (x - \bar{x})_\psi \int f(y)\mathcal{N}(y, \bar{x}, V)(y, x, M(y))dydx = \int f(y)\mathcal{N}(y, \bar{x}, V) \int (x - \bar{x})_\psi \mathcal{N}(y, x, M(y))dxdy \quad (\text{A.45a})$$

$$= \int f(y)\mathcal{N}(y, \bar{x}, V) \int [(x - y)_\psi + (y - \bar{x})_\psi] \mathcal{N}(x, y, M(y))dxdy \quad (\text{A.45b})$$

$$1203 = \int (y - \bar{x})_\psi f(y)\mathcal{N}(y, \bar{x}, V)dy = \int (x - \bar{x})_\psi f(x)\mathcal{N}(x, \bar{x}, V)dx = V_{\psi\alpha} \hat{f}_\alpha(\bar{x}, V) \quad (\text{A.45c})$$

The third integral is

$$\int (x - \bar{x})_\psi (x - \bar{x})_\omega \int f(y)\mathcal{N}(y, \bar{x}, V)(y, x, M(y))dydx \quad (\text{A.46a})$$

$$1206 = \int f(y)\mathcal{N}(y, \bar{x}, V) \int (x - \bar{x})_\psi (x - \bar{x})_\omega \mathcal{N}(x, y, M(y))dxdy \quad (\text{A.46b})$$

$$= \int f(y)\mathcal{N}(y, \bar{x}, V) \int [(x - y)_\psi (x - y)_\omega + (y - \bar{x})_\psi (x - y)_\omega] \quad (\text{A.46c})$$

$$+ (x - y)_\psi (y - \bar{x})_\omega + (y - \bar{x})_\psi (y - \bar{x})_\omega] \mathcal{N}(x, y, M(y))dxdy \quad (\text{A.46d})$$

$$1209 = \int f(y)M_{\psi\omega}(y)\mathcal{N}(y, \bar{x}, V)dy + \int (y - \bar{x})_\psi (y - \bar{x})_\omega f(y)\mathcal{N}(y, \bar{x}, M)dy \quad (\text{A.46e})$$

$$= \widehat{fM}_{\psi\omega}(\bar{x}, V) + V_{\psi\alpha} \hat{f}_{\alpha\beta}(\bar{x}, V) V_{\beta\omega} + \hat{f}(\bar{x}, V) V_{\psi\omega} \quad (\text{A.46f})$$

The mutation term in the moment equations when the mutation variance-covariance matrix is

1212 trait-dependent will thus read

$$(\widehat{fM})_{iklm\psi\omega} = \int f_{klm}(x)M_{klm\psi\omega}(x)\mathcal{N}(x, \bar{x}_{il}, V_{il})dx \quad (\text{A.47})$$

To second order we have

$$1215 \int f(y)M_{\psi\omega}(y)\mathcal{N}(y, \bar{x}, V)dy \quad (\text{A.48a})$$

$$\approx f(\bar{x})M_{\psi\omega}(\bar{x}) + \frac{1}{2}V_{\alpha\beta} [fM_{\psi\omega}]_{\alpha\beta}(\bar{x}) \quad (\text{A.48b})$$

$$= f(\bar{x})M_{\psi\omega}(\bar{x}) + \frac{1}{2}V_{\alpha\beta} \left[ f_{\alpha\beta}(\bar{x})M_{\psi\omega}(\bar{x}) + f_\alpha(\bar{x})M_{\psi\omega\beta}(\bar{x}) + f_\beta(\bar{x})M_{\psi\omega\alpha}(\bar{x}) + f(\bar{x})M_{\psi\omega\alpha\beta}(\bar{x}) \right] \quad (\text{A.48c})$$

1218 Apart from the term in the moment equations that add to variance-covariance through mutation,  
 which is now given by Eq. A.47, the rest of the moment equations remain unchanged under the  
 assumption that the variance-covariance matrix is trait dependent.

1221 We now turn to the case where there is also mutation bias so that a parent with trait  $y$  on  
 average produces offspring with trait  $y + \mu(y)$  with variance-covariance  $M(y)$ . Mathematically,  
 this implies that a rate  $f$  produces new individuals with trait  $x$  at a rate

$$1224 \quad \int f(y)v(y)\mathcal{N}(y, x - \mu(y), M(y))dy \quad (\text{A.49})$$

Once again, we can evaluate the mutation integrals, the first of which yields

$$\int \int f(y)\mathcal{N}(y, \bar{x}, V)(y, x - \mu(y), M(y))dydx = \int f(y)\mathcal{N}(y, \bar{x}, V) \int \mathcal{N}(x, y + \mu(y), M(y))dxdy \quad (\text{A.50a})$$

$$1227 \quad = \int f(y)\mathcal{N}(y, \bar{x}, V)dy = \int f(x)\mathcal{N}(x, \bar{x}, V)dx = \hat{f}(\bar{x}, V) \quad (\text{A.50b})$$

The second mutation integral is

$$\int (x - \bar{x})_\psi \int f(y)\mathcal{N}(y, \bar{x}, V)(y, x - \mu(y), M(y))dydx \quad (\text{A.51a})$$

$$1230 \quad = \int f(y)\mathcal{N}(y, \bar{x}, V) \int (x - \bar{x})_\psi \mathcal{N}(x, y + \mu(y), M(y))dxdy \quad (\text{A.51b})$$

$$= \int f(y)\mathcal{N}(y, \bar{x}, V) \int [(x - (y + \mu(y)))_\psi + (y + \mu(y) - \bar{x})_\psi] \mathcal{N}(x, y + \mu(y), M(y))dxdy \quad (\text{A.51c})$$

$$= \int [(y - \bar{x})_\psi + \mu(y)_\psi] f(y)\mathcal{N}(y, \bar{x}, V)dy = V_{\psi\alpha} \hat{f}_\alpha(\bar{x}, V) + \widehat{f\mu}_\psi(\bar{x}, V) \quad (\text{A.51d})$$

1233 where

$$\widehat{f\mu}_\psi(\bar{x}, V) = \int f(y)\mu_\psi(y)\mathcal{N}(y, \bar{x}, V)dy \quad (\text{A.52})$$

For the third mutation integral, let  $\tilde{y} := y + \mu(y)$ . We then get

$$1236 \quad \int (x - \bar{x})_\psi (x - \bar{x})_\omega \int f(y)\mathcal{N}(y, \bar{x}, V)(y, x - \mu(y), M(y))dydx \quad (\text{A.53a})$$

$$= \int f(y)\mathcal{N}(y, \bar{x}, V) \int (x - \bar{x})_\psi (x - \bar{x})_\omega \mathcal{N}(x, \tilde{y}, M(y))dxdy \quad (\text{A.53b})$$

$$= \int f(y)\mathcal{N}(y, \bar{x}, V) \int [(x - \tilde{y})_\psi (x - \tilde{y})_\omega + (\tilde{y} - \bar{x})_\psi (x - \tilde{y})_\omega] \quad (\text{A.53c})$$

$$1239 \quad + (x - \tilde{y})_{\psi}(\tilde{y} - \bar{x})_{\omega} + (\tilde{y} - \bar{x})_{\psi}(\tilde{y} - \bar{x})_{\omega}] \mathcal{N}(x, \tilde{y}, M(y)) dx dy \quad (\text{A.53d})$$

$$= \int f(y) M_{\psi\omega}(y) \mathcal{N}(y, \bar{x}, V) dy + \int (\tilde{y} - \bar{x})_{\psi}(\tilde{y} - \bar{x})_{\omega} f(y) \mathcal{N}(y, \bar{x}, V) dy \quad (\text{A.53e})$$

$$= f \widehat{M}_{\psi\omega}(\bar{x}, V) + \int (y - \bar{x})_{\psi}(y - \bar{x})_{\omega} f(y) \mathcal{N}(y, \bar{x}, V) dy \quad (\text{A.53f})$$

$$1242 \quad + \int \mu(y)_{\psi}(y - \bar{x})_{\omega} f(y) \mathcal{N}(y, \bar{x}, V) dy + \int (y - \bar{x})_{\psi} \mu(y)_{\omega} f(y) \mathcal{N}(y, \bar{x}, V) dy \quad (\text{A.53g})$$

$$+ \int \mu(y)_{\psi} \mu(y)_{\omega} f(y) \mathcal{N}(y, \bar{x}, V) dy \quad (\text{A.53h})$$

$$= f \widehat{M}_{\psi\omega}(\bar{x}, V) + V_{\psi\alpha} \hat{f}_{\alpha\beta}(\bar{x}, V) V_{\beta\omega} + \hat{f}(\bar{x}, V) V_{\psi\omega} \quad (\text{A.53i})$$

$$1245 \quad + V_{\psi\alpha} (\widehat{\mu_{\omega} f})_{\alpha}(\bar{x}, V) + V_{\omega\alpha} (\widehat{\mu_{\psi} f})_{\alpha}(\bar{x}, V) + \widehat{\mu_{\psi} \mu_{\omega} f}(\bar{x}, V) \quad (\text{A.53j})$$

When inserted into the full moment equations where each process  $f_{klm}$  has an attached mutation bias vector  $\mu_{klm}(y)$  for the class-structured community this yields the moment equations

$$1248 \quad \frac{du_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \hat{f}_{iklm} u_{il} \quad (\text{A.54a})$$

$$\frac{d\bar{x}_{ik\psi}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left( V_{il\psi\alpha} \hat{f}_{iklm\alpha} + \hat{f}_{iklm} (\bar{x}_{il} - \bar{x}_{ik})_{\psi} + \widehat{f\mu}_{iklm\psi} \right) \quad (\text{A.54b})$$

$$\frac{dV_{ik\psi\omega}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left( f \widehat{M}_{iklm\psi\omega} + V_{il\psi\alpha} \hat{f}_{iklm\alpha\beta} V_{il\beta\omega} + \hat{f}_{iklm} (V_{il\psi\omega} - V_{ik\psi\omega}) \right) \quad (\text{A.54c})$$

$$1251 \quad + V_{il\psi\alpha} (\widehat{f\mu}_{iklm\omega})_{\alpha} + V_{il\omega\alpha} (\widehat{f\mu}_{iklm\psi})_{\alpha} + \widehat{\mu\mu} f_{iklm\psi\omega} \quad (\text{A.54d})$$

$$+ (\bar{x}_{il} - \bar{x}_{ik})_{\psi} (V_{il\omega\alpha} \hat{f}_{iklm\alpha} + \widehat{f\mu}_{iklm\omega}) + (\bar{x}_{il} - \bar{x}_{ik})_{\omega} (V_{il\psi\alpha} \hat{f}_{iklm\alpha} + \widehat{f\mu}_{iklm\psi}) \quad (\text{A.54e})$$

$$+ \hat{f}_{iklm} (\bar{x}_{il} - \bar{x}_{ik})_{\psi} (\bar{x}_{il} - \bar{x}_{ik})_{\omega} \quad (\text{A.54f})$$

1254 Here, all functions with subscript  $iklm$  are evaluated at  $\bar{x}_{il}$  and  $V_{il}$ .

In vector-matrix form these equations are written as

$$\frac{du_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \hat{f}_{iklm} u_{il} \quad (\text{A.55a})$$

$$1257 \quad \frac{d\bar{x}_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left( V_{il} \nabla_x \hat{f}_{iklm} + \hat{f}_{iklm} (\bar{x}_{il} - \bar{x}_{ik}) + \widehat{f\mu}_{iklm} \right) \quad (\text{A.55b})$$

$$\frac{dV_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left( f \widehat{M}_{iklm} + V_{il} \nabla_x^2 \hat{f}_{iklm} V_{il} + \hat{f}_{iklm} (V_{il} - V_{ik}) \right) \quad (\text{A.55c})$$

$$+ V_{il} \nabla_x \widehat{f\mu}_{iklm} + [V_{il} \nabla_x \widehat{f\mu}_{iklm}]^T + \widehat{\mu\mu}^T f_{iklm} \quad (\text{A.55d})$$

$$1260 \quad + (\bar{x}_{il} - \bar{x}_{ik})(V_{il}\nabla_x \hat{f}_{iklm} + \widehat{f}\widehat{\mu}_{iklm})^T + (V_{il}\nabla_x \hat{f}_{iklm} + \widehat{f}\widehat{\mu}_{iklm})(\bar{x}_{il} - \bar{x}_{ik})^T \quad (\text{A.55e})$$

$$+ \hat{f}_{iklm}(\bar{x}_{il} - \bar{x}_{ik})(\bar{x}_{il} - \bar{x}_{ik})^T \quad (\text{A.55f})$$

For scalar traits, these simplify somewhat to

$$1263 \quad \frac{du_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \hat{f}_{iklm} u_{il} \quad (\text{A.56a})$$

$$\frac{d\bar{x}_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left( V_{il} \partial_x \hat{f}_{iklm} + \hat{f}_{iklm}(\bar{x}_{il} - \bar{x}_{ik}) + \widehat{f}\widehat{\mu}_{iklm} \right) \quad (\text{A.56b})$$

$$\frac{dV_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left( \widehat{f}\widehat{M}_{iklm} + V_{il}^2 \partial_x^2 \hat{f}_{iklm} + \hat{f}_{iklm}(V_{il} - V_{ik}) \right) \quad (\text{A.56c})$$

$$1266 \quad + 2V_{il} \partial_x \widehat{f}\widehat{\mu}_{iklm} + \widehat{\mu}^2 \hat{f}_{iklm} \quad (\text{A.56d})$$

$$+ 2(\bar{x}_{il} - \bar{x}_{ik})(V_{il} \partial_x \hat{f}_{iklm} + \widehat{f}\widehat{\mu}_{iklm}) \quad (\text{A.56e})$$

$$+ \hat{f}_{iklm}(\bar{x}_{il} - \bar{x}_{ik})^2 \quad (\text{A.56f})$$

1269

### *Environmental variables*

In our examples in the main text we have not included non-evolving environmental variables (but see Appendix D). Such environmental variables include e.g., abiotic resources, toxins, or natural enemies whose traits can be considered fixed over the time-scale of interest. Here we present a brief account of how these are included in the trait-space and moment equations when the rate functions of the environmental variables depend linearly on the species' trait-density distributions, as would be the case in classical resource-competition theory (Tilman, 1982) and contemporary niche theory (Chase and Leibold, 2003; Koffel et al., 2021).

Environmental variables may or may not follow the same class structure as the ecological communities. For example, in a two-patch model, resources may follow the same class structure as the species and are distributed among the patches, but in our stage-structured model (Appendix D) the resources are unstructured. Furthermore, each environmental variable might have its own structure. Thus, let  $E_{jn_j}$  be the density of environmental variable  $j$  in environmental class  $n_j$ , with  $N_E$  total environmental variables and  $K_{Ej}$  environmental variable classes for environmental

variable  $j$ . We let  $\mathbf{E}$  denote the vector of all environmental variable densities in all environmental  
 1284 classes. We note here that it would be equally possible to simply enumerate all environmental  
 variables  $E_1, E_2, \dots$  without loss of generality. However, for the times when the structure for the  
 species mirrors that of the environmental variables (as in e.g., a patch model where each patch  
 1287 has, say, a resource and a natural enemy) structuring the environmental variables is useful.

Together with the trait-space equations for the ecological communities this yields the equation  
 system

$$1290 \quad \frac{dv_k(x)}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \int f_{klm}(y, \mathbf{v}, \mathbf{E}) v_l(y) \mathcal{N}(x, y, M_{klm}) dy \quad (\text{A.57a})$$

$$\frac{dE_{jn_j}}{dt} = \rho_{jn_j}(\mathbf{E}) + \sum_{l=1}^K \int I_{jn,l}(\mathbf{E}, x) v_l(x) dx \quad (\text{A.57b})$$

Here,  $\rho_{jn_j}(\mathbf{E})$  represent the dynamics of environmental variable  $j$  in environmental class  $n_j$  in the  
 1293 absence of any species, and  $I_{jn,l}(\mathbf{E}, y)$  is the impact of a an individual in class  $l$  with trait  $y$  on  
 environmental variable  $j$  in environmental class  $n_j$ .

As before, we now make the assumption that the trait-density distribution  $v_l(x)$  can be ap-  
 1296 proximated by a sum of normal distributions:

$$v_l(x) \approx \tilde{v}_l(x) = \sum_{i=1}^S u_{il} \mathcal{N}(x, \bar{x}_{il}, V_{il}) \quad (\text{A.58})$$

and, similarly to how we derived the moment equations, introduce the population-level per  
 1299 capita impact from a population with mean trait  $\bar{x}_{il}$  and variance-covariance  $V_{il}$ :

$$\hat{I}_{ijn,l} := \int I_{jn,l}(\mathbf{E}, x) \mathcal{N}(x, \bar{x}_{il}, V_{il}) dx \quad (\text{A.59})$$

If necessary, this can then be Taylor-expanded to arbitrary order. In particular, to second order  
 1302 this yields:

$$\hat{I}_{ijn,l} \approx I_{jn,l}(\mathbf{E}, \bar{x}_{il}) + \frac{1}{2} \nabla_x^2 I_{jn,l}(\mathbf{E}, \bar{x}_{il}) : V_{il} \quad (\text{A.60})$$

Thus, the moment equations including environmental variables read

$$1305 \quad \frac{du_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \hat{f}_{iklm} u_{il} \quad (\text{A.61a})$$

$$\frac{d\bar{x}_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left( V_{il} \nabla_x \hat{f}_{iklm} + \hat{f}_{iklm} (\bar{x}_{il} - \bar{x}_{ik}) \right) \quad (\text{A.61b})$$

$$\begin{aligned} \frac{dV_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left( V_{il} \nabla_x^2 \hat{f}_{iklm} V_{il} + \hat{f}_{iklm} (V_{il} - V_{ik}) + \hat{f}_{iklm} M_{klm} \right. \\ \left. + \hat{f}_{iklm} (\bar{x}_{il} - \bar{x}_{ik}) (\bar{x}_{il} - \bar{x}_{ik})^T \right. \\ \left. + V_{il} \nabla_x \hat{f}_{iklm} (\bar{x}_{il} - \bar{x}_{ik})^T + (\bar{x}_{il} - \bar{x}_{ik}) \left[ V_{il} \nabla_x \hat{f}_{iklm} \right]^T \right) \end{aligned} \quad (\text{A.61c})$$

$$\frac{dE_{jn_j}}{dt} = \rho_{jn_j} + \sum_{i=1}^S \sum_{l=1}^K \hat{I}_{jn_jil} u_{il} \quad (\text{A.61d})$$

1308

1311 Here, the arguments of all rate functions have been suppressed for notational clarity, and functions are evaluated at  $\bar{x}_{il}$ ,  $V_{il}$ , and  $\mathbf{E}$ .

### *Environmental, non-heritable trait variation*

1314

We have throughout this work assumed that all trait variation is heritable, but our framework can be extended to include non-heritable environmental variation. While we have not worked out how to incorporate environmental variation to the same detail and generality as for the case without it, below we present a sketch of one way of incorporating environmental variation for the simplest case of a single trait in an unstructured community, and derive moment equations for this case.

1317

1320

To derive equations that include environmental variation, we must consider the joint genotype-phenotype distribution  $v(x, \xi)$ , where  $x$  is a genotype value and  $\xi$  is a phenotype value. The distribution  $v(x, \xi)$  thus describes how many individuals there are with genotype value  $x$  and phenotype value  $\xi$ , or, perhaps more intuitively, for any fixed  $x$ ,  $v(x, \xi)$  gives the distribution of phenotype values for the genotype value  $x$ . Furthermore, we will define

1323

$$v_G(x) := \int_{-\infty}^{\infty} v(x, \xi) d\xi, \quad (\text{A.62a})$$

1326

$$v_P(\xi) := \int_{-\infty}^{\infty} v(x, \xi) dx, \quad (\text{A.62b})$$

to be the distributions of genotypes and phenotypes in the community respectively.

Next, we must make assumptions regarding how environmental variation is generated in the  
 1329 community. We will do this in a similar manner to how we incorporated mutations into our  
 trait-space equations. We will assume that the per capita birth rate is given by  $b(\xi, v_P)$  and thus  
 depends on the phenotype value of the individual and the distribution of phenotype values in  
 1332 the community,  $v_P(\xi)$ . Thus, the total rate at which parents with genotype  $y$  birth offspring is

$$\int_{-\infty}^{\infty} b(\eta, v_P)v(y, \eta)d\eta, \quad (\text{A.63})$$

where we have integrated over all phenotypes  $\eta$  for the given genotype  $y$ . Assuming once again  
 1335 that each parent with genotype  $y$  gives birth to a normally-distributed clutch of offspring due to  
 mutations, the distribution of birthed genotypes in the community can then be calculated to be

$$B(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b(\eta, v_P)v(y, \eta)d\eta\mathcal{N}(x, y, M)dy. \quad (\text{A.64})$$

1338 Finally, we will assume that there is some fixed amount of normal environmental variation  $E_b$   
 assigned at birth around each genotype value. This means that new offspring will be produced  
 at a rate for the joint distribution given by

$$1341 \quad B(x)\mathcal{N}(\xi, x, E_b). \quad (\text{A.65})$$

Assuming that mortality rates as well as birth rates depend only on individuals' phenotype  
 values and the phenotype distribution in the community we can thus write down how the joint  
 1344 genotype–phenotype distribution will change over time:

$$\frac{dv(x, \xi)}{dt} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b(\eta, v_P)v(y, \eta)d\eta\mathcal{N}(x, y, M)dy\mathcal{N}(\xi, x, E_b) - m(\xi, v_P)v(x, \xi). \quad (\text{A.66})$$

This is the trait-space equation for the case where environmental variation is included.

### 1347 *Moment equations with environmental variation*

Just as we did without environmental variation, we will assume that the joint genotype–phenotype  
 distribution  $v(x, \xi)$  can be broken up into  $S$  ‘species’, or peaks, so that

$$1350 \quad v(x, \xi) = \sum_{i=1}^S v_i(x, \xi) \quad (\text{A.67})$$



and we will then approximate each peak by

$$\tilde{v}_i(x, \xi) := u_i \mathcal{N}(x, \bar{x}_i, G_i) \mathcal{N}(\xi, x, E_i) \approx v_i(x, \xi), \quad (\text{A.68})$$

1353 where  $u_i$  is the total density of individuals in species  $i$ ,  $\bar{x}_i$  is the mean genotype of species  $i$ ,  $G_i$   
is the genetic variance in species  $i$ , and  $E_i$  is the environmental variance in species  $i$ . In words,  
each species is assumed to have a normal distribution of genotype values, and to have a normal  
1356 distribution of phenotype values around each genotype value. We write

$$\tilde{v}(x, \xi) := \sum_{i=1}^S \tilde{v}_i(x, \xi) \approx v(x, \xi) \quad (\text{A.69})$$

for the approximated trait-density distribution. The distribution of phenotype values in each  
1359 species is given by

$$\tilde{v}_{P_i}(\xi) = \int_{-\infty}^{\infty} \tilde{v}_i(x, \xi) dx = \int_{-\infty}^{\infty} u_i \mathcal{N}(x, \bar{x}_i, G_i) \mathcal{N}(\xi, x, E_i) dx = u_i \mathcal{N}(\xi, \bar{x}, G_i + E_i) \quad (\text{A.70})$$

and we thus let  $V_i := G_i + E_i$  be the phenotypic trait variance in species  $i$ . Note that under these  
1362 assumptions, the mean phenotype in species  $i$ ,  $\bar{\xi}_i$ , is equal to the mean genotype in species  $i$ ,  $\bar{x}_i$ ,  
and so we shall only need to track the mean genotype  $\bar{x}_i$  for each species, and can refer to  $\bar{x}_i$  as  
simply the “mean trait” in species  $i$ , as we did for the case with no environmental variation.

1365 Like for the case without environmental variation we need to define population-level rates  
for a normally-distributed population with mean trait  $\bar{x}$  and phenotypic variance  $V$ , which are  
given by

$$1368 \quad \hat{b}(\bar{x}, V, \tilde{v}_P) = \int_{-\infty}^{\infty} b(\xi, \tilde{v}_P) \mathcal{N}(\xi, \bar{x}, V) d\xi, \quad (\text{A.71a})$$

$$\hat{m}(\bar{x}, V, \tilde{v}_P) = \int_{-\infty}^{\infty} m(\xi, \tilde{v}_P) \mathcal{N}(\xi, \bar{x}, V) d\xi. \quad (\text{A.71b})$$

Given this, we can derive (see subsection below) the moment equations for the total densities  $u_i$ ,  
1371 mean traits  $\bar{x}_i$ , genetic variance  $G_i$ , and environmental variance  $E_i$  to be

$$\frac{du_i}{dt} = \left[ \hat{b}(\bar{x}_i, V_i, \tilde{v}_P) - \hat{m}(\bar{x}_i, V_i, \tilde{v}_P) \right] u_i, \quad (\text{A.72a})$$

$$\frac{d\bar{x}_i}{dt} = G_i \left[ \frac{\partial \hat{b}}{\partial \bar{x}}(\bar{x}_i, V_i, \tilde{v}_P) - \frac{\partial \hat{m}}{\partial \bar{x}}(\bar{x}_i, V_i, \tilde{v}_P) \right], \quad (\text{A.72b})$$

$$1374 \quad \frac{dG_i}{dt} = G_i^2 \left[ \frac{\partial^2 \hat{b}}{\partial \bar{x}^2}(\bar{x}_i, V_i, \tilde{v}_P) - \frac{\partial^2 \hat{m}}{\partial \bar{x}^2}(\bar{x}_i, V_i, \tilde{v}_P) \right] + \hat{b}(\bar{x}_i, V_i, \tilde{v}_P)M, \quad (\text{A.72c})$$

$$\frac{dE_i}{dt} = \hat{b}(\bar{x}_i, V_i, \tilde{v}_P) (E_b - E_i) - (V_i^2 - G_i^2) \frac{\partial^2 \hat{m}}{\partial \bar{x}^2}(\bar{x}_i, V_i, \tilde{v}_P), \quad (\text{A.72d})$$

$$V_i = G_i + E_i. \quad (\text{A.72e})$$

1377 The first three equations for  $u_i$ ,  $\bar{x}_i$ , and  $G_i$  closely mirror those we derived without environ-  
 mental variation, but the population-level birth and mortality rates and their derivatives must  
 be evaluated using the total phenotypic variance  $V_i$ . When  $E_i = 0$ ,  $V_i = G_i$  and we have com-  
 1380 plete agreement with our equations that do not incorporate environmental variation as expected.  
 More notably, under our assumption of constant environmental variation generated at birth  $E_b$ ,  
 we must take the standing environmental variance  $E_i$  for each species to be a dynamic variable  
 1383 that changes over time in response to births and deaths. Roughly speaking, the first term of  
 Eq. A.72d describes how births change the standing environmental variation  $E_i$  to become closer  
 to the amount generated at birth,  $E_b$ . The second term of Eq. A.72d describes the effects of trait-  
 1386 dependent mortality on the standing environmental variation. If mortality is trait independent  
 so that  $\partial^2 \hat{m} / \partial \bar{x}^2 = 0$ , then the standing environmental variation will only be driven by the birth  
 process, and  $E_i$  will settle onto becoming equal to  $E_b$  for all species, and once  $E_i = E_b$ , it will  
 1389 remain so regardless of how the other state variables may change over time. If, however, there is  
 stabilizing selection due to trait dependence of the mortality term so that  $-\partial^2 \hat{m} / \partial \bar{x}^2 < 0$ , then  
 this will have the effect of reducing the standing environmental variation over time, and in equi-  
 1392 librium there will be a balance between the environmental variation generated through births  
 and the reduction of environmental variation due to stabilizing mortality selection.

While outside the scope of this paper, using the same basic idea one could work out the  
 1395 moment equations for class-structured populations with multiple traits including environmental  
 variation, and invasion and branching criteria could be worked out in analogous ways to how  
 we did so for the case without environmental variation.

Before we derive the expressions for the rates of change for the moments, we must compute some intermediary results that will be useful in computing the various integrals involved when

1401 calculating the moment equations.

First, let  $f \in \{b, m\}$  be either the function describing the birth rate or the mortality rate. We then define

$$1404 \quad \hat{f}(\bar{x}, G, E, v_p) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\zeta, v_p) \mathcal{N}(x, \bar{x}, G) \mathcal{N}(\zeta, x, E) d\zeta dx, \quad (\text{A.73a})$$

$$\hat{f}(\bar{x}, V, v_p) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\zeta, v_p) \mathcal{N}(\zeta, \bar{x}, V) d\zeta, \quad (\text{A.73b})$$

and note that when  $V = G + E$  these expressions are equal. For the first derivative of  $\hat{f}$  we

1407 calculate

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x}) f(\zeta, v_p) \mathcal{N}(x, \bar{x}, G) \mathcal{N}(\zeta, x, E) d\zeta dx \quad (\text{A.74a})$$

$$= G \int_{-\infty}^{\infty} f(\zeta, v_p) \frac{\zeta - \bar{x}}{V} \mathcal{N}(\zeta, \bar{x}, V) d\zeta = G \frac{\partial \hat{f}}{\partial \bar{x}}(\bar{x}, V, v_p). \quad (\text{A.74b})$$

1410 For the second derivatives of  $\hat{f}$  we may similarly calculate

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^2 f(\zeta, v_p) \mathcal{N}(x, \bar{x}, G) \mathcal{N}(\zeta, x, E) d\zeta dx = G^2 \frac{\partial^2 \hat{f}}{\partial \bar{x}^2}(\bar{x}, V, v_p) + G \hat{f}(\bar{x}, V, v_p), \quad (\text{A.75a})$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\zeta - \bar{x})^2 f(\zeta, v_p) \mathcal{N}(x, \bar{x}, G) \mathcal{N}(\zeta, x, E) d\zeta dx = V^2 \frac{\partial^2 \hat{f}}{\partial \bar{x}^2}(\bar{x}, V, v_p) + V \hat{f}(\bar{x}, V, v_p). \quad (\text{A.75b})$$

Under the assumption that the mean genotype  $\bar{x}_i$  equals the mean phenotype  $\bar{\zeta}_i$  (or simply, the mean trait) the moments are given by

$$u_i = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} v_i(x, \zeta) dx d\zeta, \quad (\text{A.76a})$$

$$\bar{x}_i = \frac{1}{u_i} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x v_i(x, \zeta) dx d\zeta, \quad (\text{A.76b})$$

$$G_i = \frac{1}{u_i} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x}_i)^2 v_i(x, \zeta) dx d\zeta, \quad (\text{A.76c})$$

$$V_i = \frac{1}{u_i} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\zeta - \bar{x}_i)^2 v_i(x, \zeta) dx d\zeta, \quad (\text{A.76d})$$

$$E_i = V_i - G_i = \frac{1}{u_i} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [(\zeta - \bar{x}_i)^2 - (x - \bar{x}_i)^2] v_i(x, \zeta) dx d\zeta \quad (\text{A.76e})$$

$$= \frac{1}{u_i} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [2(\zeta - x)(x - \bar{x}_i) + (\zeta - x)^2] v_i(x, \zeta) dx d\zeta. \quad (\text{A.76f})$$

1413 Now, assuming our normal approximations so that

$$\bar{v}_i(x, \zeta) := u_i \mathcal{N}(x, \bar{x}_i, G_i) \mathcal{N}(\zeta, x, E_i) \approx v_i(x, \zeta), \quad (\text{A.77a})$$

$$\bar{v}(x, \zeta) := \sum_{i=1}^S \bar{v}_i(x, \zeta) \approx v(x, \zeta), \quad (\text{A.77b})$$

$$1416 \quad \bar{v}_P(\zeta) := \int_{-\infty}^{\infty} \bar{v}(x, \zeta) dx, \quad (\text{A.77c})$$

we can derive the moment equations. Below, we calculate in turn the moment equations for the total density  $u_i$ , the mean trait  $\bar{x}_i$ , the genetic variance  $G_i$ , and the environmental variance  $E_i$  for each species  $i$ .

Total density,  $u_i$ :

$$\frac{du_i}{dt} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{dv_i(x, \zeta)}{dt} d\zeta dx \approx \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{d\bar{v}_i(x, \zeta)}{dt} d\zeta dx \quad (\text{A.78a})$$

$$1422 \quad = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b(\eta, \bar{v}_P) \bar{v}_i(y, \eta) d\eta \mathcal{N}(x, y, M) dy \mathcal{N}(\zeta, x, E_b) \right. \\ \left. - m(\zeta, \bar{v}_P) \bar{v}_i(x, \zeta) \right] d\zeta dx = \quad (\text{A.78b})$$

{integrate in order  $\zeta \rightarrow x \rightarrow y$  in birth term}

$$1425 \quad = \left[ \hat{b}(\bar{x}_i, V_i, \bar{v}_P) - \hat{m}(\bar{x}_i, V_i, \bar{v}_P) \right] u_i \quad (\text{A.78c})$$

Mean trait,  $\bar{x}_i$ :

$$\frac{d}{dt}(u_i \bar{x}_i) = u_i \frac{d\bar{x}_i}{dt} + \bar{x}_i \frac{du_i}{dt} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \frac{dv_i(x, \zeta)}{dt} d\zeta dx \quad (\text{A.79a})$$

$$1428 \quad = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x}_i) \frac{dv_i(x, \zeta)}{dt} d\zeta dx + \bar{x}_i \frac{du_i}{dt} \quad (\text{A.79b})$$

and we thus get

$$u \frac{d\bar{x}_i}{dt} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x}_i) \frac{dv_i(x, \zeta)}{dt} d\zeta dx \approx \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x}_i) \frac{d\bar{v}_i(x, \zeta)}{dt} d\zeta dx \quad (\text{A.80a})$$

$$1431 \quad = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x}_i) \left[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b(\eta, \bar{v}_P) \bar{v}_i(y, \eta) d\eta \mathcal{N}(x, y, M) dy \mathcal{N}(\zeta, x, E_b) \right. \\ \left. - m(\zeta, \bar{v}_P) \bar{v}_i(x, \zeta) \right] d\zeta dx = \quad (\text{A.80b})$$

{integrate in order  $\zeta \rightarrow x$  in birth term}

$$1434 \quad = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y - \bar{x}_i) b(\eta, \bar{v}_P) \bar{v}_i(y, \eta) d\eta dy - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x}_i) m(\zeta, \bar{v}_P) \bar{v}_i(x, \zeta) d\zeta dx \quad (\text{A.80c})$$

{Eqs. A.74}

$$= u_i G_i \frac{\partial \hat{b}}{\partial \bar{x}}(\bar{x}_i, V_i, \bar{v}_P) - u_i G_i \frac{\partial \hat{m}}{\partial \bar{x}}(\bar{x}_i, V_i, \bar{v}_P) \quad (\text{A.80d})$$

1437  $\iff$

$$\frac{d\bar{x}_i}{dt} = G_i \left[ \frac{\partial \hat{b}}{\partial \bar{x}}(\bar{x}_i, V_i, \bar{v}_P) - \frac{\partial \hat{m}}{\partial \bar{x}}(\bar{x}_i, V_i, \bar{v}_P) \right] \quad (\text{A.80e})$$

Genetic variance,  $G_i$ :

$$1440 \quad \frac{d}{dt}(u_i G_i) = u_i \frac{dG_i}{dt} + G_i \frac{du_i}{dt} = \frac{d}{dt} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x}_i)^2 v_i(x, \zeta) d\zeta dx \quad (\text{A.81a})$$

$$\approx \frac{d}{dt} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x}_i)^2 \bar{v}_i(x, \zeta) d\zeta dx \quad (\text{A.81b})$$

$$= \frac{d\bar{x}_i}{dt} \underbrace{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} 2(x - \bar{x}_i) \bar{v}_i(x, \zeta) d\zeta dx}_{=0} + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x}_i)^2 \frac{d\bar{v}_i}{dt} d\zeta dx \quad (\text{A.81c})$$

$$1443 \quad = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x}_i)^2 \left[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b(\eta, \bar{v}_P) \bar{v}_i(y, \eta) d\eta \mathcal{N}(x, y, M) dy \mathcal{N}(\zeta, x, E_b) \right. \\ \left. - m(\zeta, \bar{v}_P) \bar{v}_i(x, \zeta) \right] d\zeta dx = \quad (\text{A.81d})$$

{integrate in order  $\zeta \rightarrow x$  in birth term}

$$1446 \quad = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b(\eta, \bar{v}_P) \bar{v}_i(y, \eta) [M + (y - \bar{x}_i)^2] d\eta dy \\ - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} m(\zeta, \bar{v}_P) \bar{v}_i(x, \zeta) (x - \bar{x}_i)^2 d\zeta dx \quad (\text{A.81e})$$

{Eq. A.75}

$$1449 \quad = u_i \left[ \hat{b}(\bar{x}_i, V_i, \bar{v}_P) M + G_i^2 \frac{\partial^2 \hat{b}}{\partial \bar{x}^2}(\bar{x}_i, V_i, \bar{v}_P) + G_i \hat{b}(\bar{x}_i, V_i, \bar{v}_P) \right. \\ \left. - G_i^2 \frac{\partial^2 \hat{m}}{\partial \bar{x}^2}(\bar{x}_i, V_i, \bar{v}_P) - G_i \hat{m}(\bar{x}_i, V_i, \bar{v}_P) \right] \quad (\text{A.81f})$$

$\iff$

$$1452 \quad \frac{dG_i}{dt} = G_i^2 \left[ \frac{\partial^2 \hat{b}}{\partial \bar{x}^2}(\bar{x}_i, V_i, \bar{v}_P) - \frac{\partial^2 \hat{m}}{\partial \bar{x}^2}(\bar{x}_i, V_i, \bar{v}_P) \right] + \hat{b}(\bar{x}_i, V_i, \bar{v}_P) M \quad (\text{A.81g})$$

Environmental variance,  $E_i$ :

$$\frac{d}{dt}(u_i E_i) = u_i \frac{dE_i}{dt} + E_i \frac{du_i}{dt} \quad (\text{A.82a})$$

$$1455 \quad = \frac{d}{dt} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [2(\zeta - x)(x - \bar{x}_i) + (\zeta - x)^2] v_i(x, \zeta) dx d\zeta \quad (\text{A.82b})$$

$$\approx \frac{d}{dt} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [2(\zeta - x)(x - \bar{x}_i) + (\zeta - x)^2] \bar{v}_i(x, \zeta) dx d\zeta \quad (\text{A.82c})$$

$$= - \frac{d\bar{x}_i}{dt} \underbrace{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} 2(\zeta - x) \bar{v}_i(x, \zeta) dx d\zeta}_{=0}$$

$$1458 \quad + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [2(\zeta - x)(x - \bar{x}_i) + (\zeta - x)^2] \frac{d\bar{v}_i(x, \zeta)}{dt} dx d\zeta \quad (\text{A.82d})$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [2(\zeta - x)(x - \bar{x}_i) + (\zeta - x)^2] \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b(\eta, \bar{v}_p) \bar{v}_i(y, \eta) d\eta \mathcal{N}(x, y, M) dy \mathcal{N}(\zeta, x, E_b) d\zeta dx$$

$$- \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [(\zeta - \bar{x}_i)^2 - (x - \bar{x}_i)^2] m(\zeta, \bar{v}_p) \bar{v}_i(x, \zeta) d\zeta dx \quad (\text{A.82e})$$

1461 {integrate in order  $\zeta \rightarrow x$  in birth term, use Eqs. A.75 in mortality term}

$$= E_b \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b(\eta, \bar{v}_p) \bar{v}_i(y, \eta) d\eta dy$$

$$- u_i \left[ V_i^2 \frac{\partial^2 \hat{m}}{\partial \bar{x}^2}(\bar{x}_i, V_i, \bar{v}_p) + V_i \hat{m}(\bar{x}_i, V_i, \bar{v}_p) - G_i^2 \frac{\partial^2 \hat{m}}{\partial \bar{x}^2}(\bar{x}_i, V_i, \bar{v}_p) - G_i \hat{m}(\bar{x}_i, V_i, \bar{v}_p) \right] \quad (\text{A.82f})$$

$$1464 \quad = u_i E_b \hat{b}(\bar{x}_i, V_i, \bar{v}_p) - u_i (V_i^2 - G_i^2) \frac{\partial^2 \hat{m}}{\partial \bar{x}^2}(\bar{x}_i, V_i, \bar{v}_p) - u_i E_i \hat{m}(\bar{x}_i, V_i, \bar{v}_p) \quad (\text{A.82g})$$

$\iff$

$$\frac{dE_i}{dt} = \hat{b}(\bar{x}_i, V_i, \bar{v}_p)(E_b - E_i) - (V_i^2 - G_i^2) \frac{\partial^2 \hat{m}}{\partial \bar{x}^2}(\bar{x}_i, V_i, \bar{v}_p) \quad (\text{A.82h})$$

1467 Together, these are the moment equations (Eqs. A.72) when environmental variance is included.

## Online Appendix B: Evolutionary branching and invasion analysis

1470 In this appendix we will derive ways of examining whether a community of species at eco-  
evolutionary equilibrium is closed to invasion by further species. We refer to these methods  
as heuristics since following these procedures cannot guarantee a complete closure to further  
1473 invasions, and because the methodology is somewhat *ad hoc*. Nevertheless, in the examples we  
have tested (see the main text) these methods have so far proven to be highly accurate in the  
sense that they almost always produce communities with the same number of species for the  
1476 moment equations as the trait-space equations produce peaks in eco-evolutionary equilibrium.  
The exceptions, as shown for the two-patch example in the main text, primarily happen when  
the number of species is ambiguous (a distribution with several peaks, but where the peaks are  
1479 not well separated), for narrow ranges of parameters near bifurcation points.

We shall follow the general ideas of adaptive dynamics (Metz et al., 1992; Dieckmann and  
Law, 1996; Geritz et al., 1998; Dercole and Rinaldi, 2008; Brännström et al., 2013) and work out  
1482 two different ways of investigating when bifurcation into more species can take place. The first  
condition is local in trait space and known as an *evolutionary branching* where a species splits  
into two new species that subsequently diverge in trait space. It is easier to calculate than the  
1485 second condition, which treats global invasion analysis where the invader does not necessarily  
have a trait that is close to any resident species. The first condition implies the second, but not  
necessarily vice versa.

### 1488 *Evolutionary branchings*

We shall here work out a criterion for when a species characterized by the moment equations can  
be split into two new species locally in trait space in such a way that they subsequently diverge  
1491 in their traits.

Consider a community of  $S$  species in  $K$  classes being in a convergence-stable eco-evolutionary

equilibrium so that

$$1494 \quad 0 = \frac{d u_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \hat{f}_{iklm} u_{il} \quad (\text{B.1a})$$

$$0 = \frac{d \bar{x}_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left( V_{il} \nabla_x \hat{f}_{iklm} + \hat{f}_{iklm} (\bar{x}_{il} - \bar{x}_{ik}) \right) \quad (\text{B.1b})$$

$$0 = \frac{d V_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left( V_{il} \nabla_x^2 \hat{f}_{iklm} V_{il} + \hat{f}_{iklm} M_{klm} + \right. \quad (\text{B.1c})$$

$$1497 \quad \left. + V_{il} \nabla_x \hat{f}_{iklm} (\bar{x}_{il} - \bar{x}_{ik})^T + (\bar{x}_{il} - \bar{x}_{ik}) \left[ V_{il} \nabla_x \hat{f}_{iklm} \right]^T + \right. \quad (\text{B.1d})$$

$$\left. + \hat{f}_{iklm} \left[ V_{il} - V_{ik} + (\bar{x}_{il} - \bar{x}_{ik})(\bar{x}_{il} - \bar{x}_{ik})^T \right] \right) \quad (\text{B.1e})$$

for all  $i \in \{1, 2, \dots, S\}$  and  $k \in \{1, 2, \dots, K\}$ .

1500 To test whether a given species  $i$  within this community can undergo an evolutionary branching we first split this species into two new species  $i_1$  and  $i_2$  with densities, mean-trait vectors, and variance-covariances given by

$$1503 \quad u_{i_1 k} = u_{i_2 k} = \frac{u_{ik}}{2} \quad (\text{B.2a})$$

$$\bar{x}_{i_1 k} = \bar{x}_{i_2 k} = \bar{x}_{ik} \quad (\text{B.2b})$$

$$V_{i_1 k} = V_{i_2 k} = V_{ik} \quad (\text{B.2c})$$

1506 We then organize all the state variables,  $u$ ,  $\bar{x}$ , and  $V$  into a vector  $w$ , so that the entries of  $w$  are the total densities, the mean trait components and the variance-covariance components for all species and classes in the split system. We let  $F(w)$  be the function given by the right-hand side of Eq. B.1 and linearize around the equilibrium  $w^*$ ,

$$1509 \quad 0 = \frac{dw}{dt} \approx F(w^*) + J_F(w^*)(w - w^*) \quad (\text{B.3})$$

where  $J_F(w^*)$  is the Jacobian matrix of  $F$  evaluated at  $w^*$ . Normally, in dynamical systems theory, 1512 we would say that the community with the split species is stable if the dominant eigenvalue (eigenvalue with largest real part) of the Jacobian matrix has a negative real part. However, here, due to the splitting, there will always be a neutral direction corresponding to shifting mass 1515 between species  $i_1$  and  $i_2$ , so there will always exist a zero eigenvalue of the Jacobian. Thus, if the



dominant eigenvalue is zero there is no branching. However, if there exists an eigenvalue with positive real part this indicates that the split-species equilibrium is unstable and that a branching in species  $i$  will take place.

Figure B.1 shows an evolutionary branching for the two-patch Lotka-Volterra model (Eqs. 20, Fig. 3, main text). The initial species evolves to an eco-evolutionary equilibrium with substantial variance and significant local adaptation between the two patches. There, we split the species and determine the dominant eigenvalue  $\lambda_d$  as described above. For this example, the dominant eigenvalue is positive showing that the one-species equilibrium is unstable to the addition of a species. Let  $\phi$  be the eigenvector associated with the dominant eigenvalue  $\lambda_d$ . The components of the eigenvector are

$$\begin{array}{ccc} \text{Patch 1} & \text{Patch 2} & \\ u_{11}^\phi = 0.4468 & u_{12}^\phi = -0.4468 & \text{(B.4a)} \end{array}$$

$$u_{21}^\phi = -0.4468 \quad u_{22}^\phi = 0.4468 \quad \text{(B.4b)}$$

$$\bar{x}_{11}^\phi = -0.2021 \quad \bar{x}_{12}^\phi = -0.2021 \quad \text{(B.4c)}$$

$$\bar{x}_{21}^\phi = 0.2021 \quad \bar{x}_{22}^\phi = 0.2021 \quad \text{(B.4d)}$$

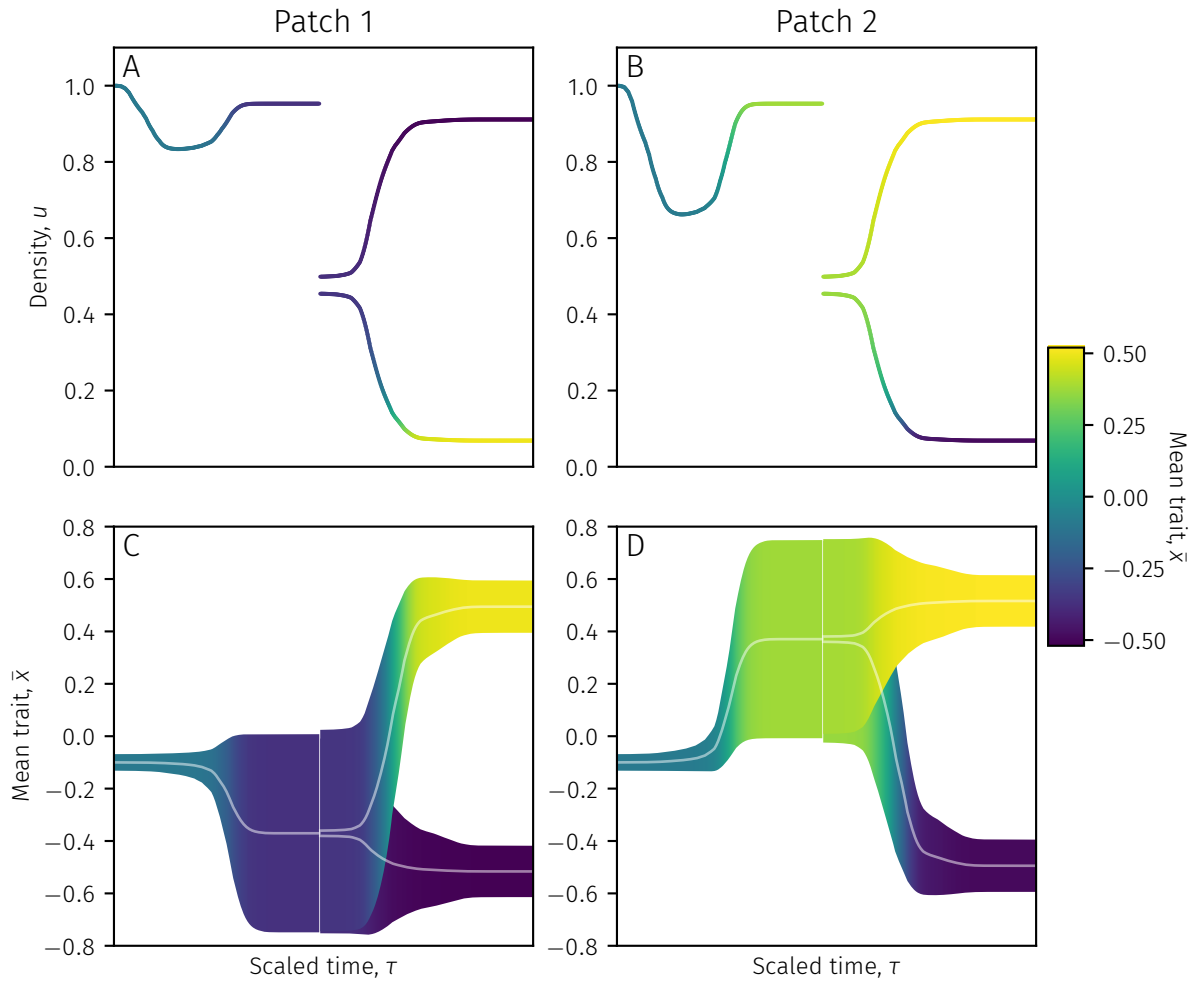
$$V_{11}^\phi = -0.09774 \quad V_{12}^\phi = 0.09774 \quad \text{(B.4e)}$$

$$V_{21}^\phi = 0.09774 \quad V_{22}^\phi = -0.09774 \quad \text{(B.4f)}$$

In other words, the unstable direction corresponds to one species becoming relatively more abundant on patch 1 and with a lower trait, and one species becoming relatively more abundant on patch 2 with a higher trait. We construct the new two-species system after the branching by letting

$$w_{\text{new}} = w^* + \alpha\phi \quad \text{(B.5)}$$

where  $\alpha$  is some small number, for this example  $\alpha = 0.05$ , then solve for the  $S + 1$ -species eco-evolutionary equilibrium.



**Figure B.1: Illustration of branching procedure for the two-patch Lotka-Volterra model (Eqs. 20, Fig. 3, main text).** The panels depicts the densities (A–B) and mean traits (C–D) in the two-patch model on both patches where the left two panels depict patch 1 and the right two panels depict patch 2. Panels (C–D) also depict one standard deviation around the mean as filled areas. The initial species evolves to an eco-evolutionary equilibrium after which a branching takes place, the species is split, and the two new species once again evolve to an eco-evolutionary equilibrium. To illustrate the branching we have rescaled time non-uniformly to ensure the interesting aspects of the branching are clearly visible. **Parameter values:**  $x_1^{\text{opt}} = -0.5$ ,  $x_2^{\text{opt}} = 0.5$ ,  $V_r = 1.0$ ,  $r_0 = 1.0$ ,  $V_c = 1.0$ ,  $d = 0.05$ , and  $M = 10^{-4}$ .

## Global invasion dynamics

Local branching in trait space will not always be sufficient to ensure that community is globally  
 1542 closed to further invasions. Here, we will work out a criterion for when an additional species with  
 an arbitrary trait can be added to the community in such a way that this added species grows  
 from being initially rare. This corresponds to local-but-not-global ESSs in adaptive dynamics  
 1545 (Geritz et al. 1999; Klausmeier et al. 2020) .

Let a discrete structured community of  $S$  species in  $K$  classes be in eco-evolutionary equilib-  
 rium so that

$$1548 \quad 0 = \frac{d u_{ik}}{d t} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \hat{f}_{iklm} u_{il} \quad (\text{B.6a})$$

$$0 = \frac{d \bar{x}_{ik}}{d t} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left( V_{il} \nabla_x \hat{f}_{iklm} + \hat{f}_{iklm} (\bar{x}_{il} - \bar{x}_{ik}) \right) \quad (\text{B.6b})$$

$$0 = \frac{d V_{ik}}{d t} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left( V_{il} \nabla_x^2 \hat{f}_{iklm} V_{il} + \hat{f}_{iklm} M_{klm} + \right. \quad (\text{B.6c})$$

$$1551 \quad \left. + V_{il} \nabla_x \hat{f}_{iklm} (\bar{x}_{il} - \bar{x}_{ik})^T + (\bar{x}_{il} - \bar{x}_{ik}) \left[ V_{il} \nabla_x \hat{f}_{iklm} \right]^T + \right. \quad (\text{B.6d})$$

$$\left. + \hat{f}_{iklm} \left[ V_{il} - V_{ik} + (\bar{x}_{il} - \bar{x}_{ik})(\bar{x}_{il} - \bar{x}_{ik})^T \right] \right) \quad (\text{B.6e})$$

for all  $i \in \{1, 2, \dots, S\}$  and  $k \in \{1, 2, \dots, K\}$ . Denote the trait-density distributions of these  
 1554 resident populations in equilibrium by

$$\tilde{v}_k^{\text{res}}(x) = \sum_{i=1}^S u_{ik} \mathcal{N}(x, \bar{x}_{ik}, V_{ik}) \quad (\text{B.7})$$

and let  $\tilde{\mathbf{v}}^{\text{res}} = (\tilde{v}_1^{\text{res}}, \dots, \tilde{v}_K^{\text{res}})$  be the vector of all resident trait-density distributions.

1557 We would now like to devise a procedure to determine whether the resident community  
 is closed to invasions. In analogy with adaptive dynamics (Metz et al., 1992; Dieckmann and  
 Law, 1996; Geritz et al., 1998; Dercole and Rinaldi, 2008; Brännström et al., 2013) we will let a  
 1560 very rare invader enter this community, and try to determine whether this invader can grow  
 in the environment set by the residents while still rare. However, we also need to account for  
 the internal dynamics of the invader, i.e., its mean traits and variance-covariances across classes.

1563 Although the invader will be assumed to be so rare that it does not effect its own dynamics or any  
of the residents, in the case of structured populations, we must still keep track of its distribution  
across classes to calculate the relative-density weight term in the mean trait and trait variance  
1566 equations. To this end let the class-frequency of the invader in class  $k$  be given by

$$p_k^{\text{inv}} := \frac{u_k^{\text{inv}}}{\sum_{l=1}^K u_l^{\text{inv}}}, \quad (\text{B.8})$$

where  $\sum_{k=1}^K p_k^{\text{inv}} \equiv 1$ . We can then calculate

$$1569 \quad \frac{dp_k^{\text{inv}}}{dt} = \frac{d}{dt} \left( \frac{u_k^{\text{inv}}}{\sum_{l=1}^K u_l^{\text{inv}}} \right) = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \hat{f}_{klm}^{\text{inv}} p_l^{\text{inv}} - p_k^{\text{inv}} \sum_{j=1}^K \sum_{l=1}^K \sum_{m=1}^{N_{jl}} \hat{f}_{jlm}^{\text{inv}} p_l^{\text{inv}}. \quad (\text{B.9})$$

We note that the mean-trait vector and variance-covariance matrix equations in the moment  
equations depend on the densities only through their ratios, and by construction we have that

$$1572 \quad \frac{p_l^{\text{inv}}}{p_k^{\text{inv}}} \equiv \frac{u_l^{\text{inv}}}{u_k^{\text{inv}}}. \quad (\text{B.10})$$

Let  $\bar{x}_k^{\text{inv}}$  and  $V_k^{\text{inv}}$  denote the mean-trait vector and variance-covariance matrix respectively of the  
rare invader in class  $k$ , and let  $\hat{f}_{klm}^{\text{inv}} := \hat{f}_{klm}(\bar{x}_l^{\text{inv}}, V_l^{\text{inv}}, \bar{\mathbf{v}}^{\text{res}})$ . Note that because we assume that  
1575 the invader is very rare, the per capita rates of the invader do not depend on the trait-density  
distribution of the invader itself. Similarly, we assume that the invader's rarity is such that it will  
not affect the residents, and  $\bar{\mathbf{v}}^{\text{res}}$  will be constant.

1578 We can then formulate the invader equations as

$$\frac{dp_k^{\text{inv}}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \hat{f}_{klm}^{\text{inv}} p_l^{\text{inv}} - p_k^{\text{inv}} \sum_{j=1}^K \sum_{l=1}^K \sum_{m=1}^{N_{jl}} \hat{f}_{jlm}^{\text{inv}} p_l^{\text{inv}}, \quad (\text{B.11a})$$

$$\frac{d\bar{x}_k^{\text{inv}}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{p_l^{\text{inv}}}{p_k^{\text{inv}}} \left( V_l^{\text{inv}} \nabla_x \hat{f}_{klm}^{\text{inv}} + \hat{f}_{klm}^{\text{inv}} (\bar{x}_l^{\text{inv}} - \bar{x}_k^{\text{inv}}) \right) \quad (\text{B.11b})$$

$$1581 \quad \frac{dV_k^{\text{inv}}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{p_l^{\text{inv}}}{p_k^{\text{inv}}} \left( V_l^{\text{inv}} \nabla_x^2 \hat{f}_{klm}^{\text{inv}} V_l^{\text{inv}} + \hat{f}_{klm}^{\text{inv}} M_{klm} + \right. \\
+ V_l^{\text{inv}} \nabla_x \hat{f}_{klm}^{\text{inv}} (\bar{x}_l^{\text{inv}} - \bar{x}_k^{\text{inv}})^T + (\bar{x}_l^{\text{inv}} - \bar{x}_k^{\text{inv}}) \left[ V_l^{\text{inv}} \nabla_x \hat{f}_{klm}^{\text{inv}} \right]^T + \\
\left. + \hat{f}_{klm}^{\text{inv}} \left[ V_l^{\text{inv}} - V_k^{\text{inv}} + (\bar{x}_l^{\text{inv}} - \bar{x}_k^{\text{inv}})(\bar{x}_l^{\text{inv}} - \bar{x}_k^{\text{inv}})^T \right] \right) \quad (\text{B.11c})$$

1584 Under equilibrium dynamics, an invader will eventually assume an equilibrium distribution, but  
the nonlinearity of the equations means that there could potentially exist several such equilibria.  
Assume that the invader equations have  $N$  total equilibria, and denote each such equilibrium by  
1587  $p_k^{\text{inv}^*n}$ ,  $\bar{x}_k^{\text{inv}^*n}$ ,  $V_k^{\text{inv}^*n}$  for the class frequencies, mean trait vectors, and variance-covariance matrices  
of the invaders for equilibrium  $n = 1, \dots, N$  and class  $k = 1, \dots, K$ .

Now, let the  $K \times K$  matrix  $\hat{F}^{\text{inv}^*n}$  with entries  $\hat{F}_{kl}^{\text{inv}^*n}$  be defined by

$$1590 \quad \hat{F}_{kl}^{\text{inv}^*n} = \sum_{m=1}^{N_{kl}} \hat{f}_{klm}^{\text{inv}^*n} = \sum_{m=1}^{N_{kl}} \hat{f}_{klm}(\bar{x}_l^{\text{inv}^*n}, V_l^{\text{inv}^*n}, \bar{\mathbf{v}}^{\text{res}}) \quad (\text{B.12})$$

and let the vector  $\mathbf{p}^{\text{inv}^*n} := (p_1^{\text{inv}^*n}, p_2^{\text{inv}^*n}, \dots, p_K^{\text{inv}^*n})$ . Then, from Eq. B.11a, we have in equilibrium  
that

$$1593 \quad \hat{F}^{\text{inv}^*n} \mathbf{p}^{\text{inv}^*n} = \left( \sum_{j=1}^K \sum_{l=1}^K \sum_{m=1}^{N_{jl}} \hat{f}_{jlm}^{\text{inv}^*n} p_l^{\text{inv}^*n} \right) \mathbf{p}^{\text{inv}^*n}. \quad (\text{B.13})$$

From this expression we see that  $\mathbf{p}^{\text{inv}^*n}$  is an eigenvector of  $\hat{F}^{\text{inv}^*n}$ , and we can see that the  
eigenvalue is

$$1596 \quad \lambda_n = \sum_{j=1}^K \sum_{l=1}^K \sum_{m=1}^{N_{jl}} \hat{f}_{jlm}^{\text{inv}^*n} p_l^{\text{inv}^*n}. \quad (\text{B.14})$$

If we now turn to the density dynamics of the invader we have that if the initial densities of the  
invader,  $\mathbf{u}_0^{\text{inv}^*n}$ , are proportional to an equilibrium class-frequency distribution,  $\mathbf{u}_0^{\text{inv}^*n} = c_0 \mathbf{p}^{\text{inv}^*n}$ ,

1599 then

$$\frac{d u_k^{\text{inv}^*n}}{d t} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \hat{f}_{klm}^{\text{inv}^*n} u_l^{\text{inv}^*n} \iff \frac{d \mathbf{u}^{\text{inv}^*n}}{d t} = \hat{F}^{\text{inv}^*n} \mathbf{u}^{\text{inv}^*n} \implies \mathbf{u}^{\text{inv}^*n}(t) = c_0 e^{\lambda_n t} \mathbf{p}^{\text{inv}^*n}. \quad (\text{B.15})$$

In other words,  $\lambda_n$  is the long-term exponential growth rate of the invader while still rare for  
1602 equilibrium  $n$ . This means that  $\lambda_n$  is the invasion fitness of invader equilibrium  $n$ . Thus, if  
 $\lambda_n > 0$ , the invader can successfully invade the community of residents, and if  $\lambda_n \leq 0$  then the  
invasion will not succeed.

1605 In principle, this then gives us a way of determining whether a resident community is in-  
vasible. For a given resident community, we find all the equilibria of the invader equations  
(Eqs. B.11), and then use Eq. B.14 to calculate the invasion fitness of each such equilibrium. If  
1608 any positive invasion fitness is found, the community is invadable, and if no positive invasion

fitness equilibria are found, the community is not invasible. In practice, however, proving that one has found all the equilibria of the invader equations will in general not be feasible. Thus, we need some heuristic that is easy to carry out and can still find the relevant equilibria most of the time. To this end, we start by computing the adaptive-dynamics fitness landscape, where we consider the invaders to have no variance and fixed mean-trait vectors with no local adaptation so that  $V_k^{\text{inv}} = 0$  for all  $k$  and  $\bar{x}_k = \bar{x}^{\text{AD}}$  for all  $k$ . Then, we can immediately calculate the growth rate of this zero-variance invader as the dominant eigenvalue of the matrix  $F^{\text{AD}}$ , with entries given by

$$F_{kl}^{\text{AD}} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} f_{klm}(\bar{x}^{\text{AD}}, \tilde{\mathbf{v}}^{\text{res}}). \quad (\text{B.16})$$

See Caswell (2001) for details on invasion analysis in structured populations in the adaptive-dynamics framework. We can then calculate this adaptive-dynamics invasion fitness for a range of trait vectors  $\bar{x}^{\text{AD}}$  in some region of trait space that is pertinent to the problem at hand. We then find all mean-trait vectors that yield local maxima in the AD-fitness landscape. Let these be denoted by  $\bar{x}^{\text{AD,max},q}$  for the  $q$ :th such local maximum. Finally, we use these mean-trait vectors as initial conditions for the invader equations (Eqs. B.11) by setting

$$\mathbf{p}_{t=0}^{\text{inv},q} = \mathbf{p}^{\text{AD,max},q} \quad (\text{B.17a})$$

$$\bar{x}_{k,t=0}^{\text{inv},q} = \bar{x}^{\text{AD,max},q} \quad \text{for all } k \quad (\text{B.17b})$$

$$V_{k,t=0}^{\text{inv},q} = 0 \quad \text{for all } k \quad (\text{B.17c})$$

where  $\mathbf{p}^{\text{AD,max},q}$  is the normalized eigenvector associated with dominant eigenvalue of the matrix with entries at  $k, l$  given by

$$\sum_{l=1}^K \sum_{m=1}^{N_{kl}} f_{klm}(\bar{x}^{\text{AD,max},q}, \tilde{\mathbf{v}}^{\text{res}}). \quad (\text{B.18})$$

For each such initial condition we then solve the invader equations to equilibrium and compute that equilibrium's invasion fitness as described above. While not guaranteed to find every invader equilibrium, this method tends to work well in practice for finding the relevant equilibria as the zero-variance populations serve as good first approximations to where positive invasion fitness might be available.

## Online Appendix C: Additional details of the Lotka-Volterra competition models

In this appendix we provide some additional details of the two example models of Lotka-Volterra competition we use in the main text.

### *The unstructured Lotka-Volterra model*

The trait-space equations of the model are given by

$$\frac{dv(x)}{dt} = \int b(y)v(y)\mathcal{N}(x, y, M)dy - \mu(x)v(x) - a(x, v)v(x), \quad (\text{C.1a})$$

$$b(x) = r_0, \quad \mu(x) = \frac{(x - x^{\text{opt}})^2}{V_r}, \quad (\text{C.1b})$$

$$r(x) := b(x) - \mu(x) = r_0 - \frac{(x - x^{\text{opt}})^2}{V_r}, \quad (\text{C.1c})$$

$$a(x, v) = \int \alpha(y, x)v(y)dy, \quad \alpha(y, x) = \exp\left(-\frac{(y - x)^2}{2V_c}\right). \quad (\text{C.1d})$$

Here,  $r(x)$  describes the intrinsic density-independent per capita net growth rate due to the environment. The strength of competition between two individuals with traits  $x$  and  $y$  is given by  $\alpha(y, x)$ , which means that the total per capita mortality experienced by an individual with trait  $x$  given the total trait-density distribution of the community is  $a(x, v)$ . For the unstructured model, the value of  $x^{\text{opt}}$  can be set to be zero without loss of generality as we do in the main text. We include it here to more easily generalize to the two-patch model.

Using our general method, we can plug in these rates to get the moment equations for this system:

$$\frac{du_i}{dt} = (\hat{r}(\bar{x}_i, V_i) - \hat{a}(\bar{x}_i, V_i, \bar{v})) u_i \quad (\text{C.2a})$$

$$\frac{d\bar{x}_i}{dt} = V_i \left( \frac{\partial \hat{r}}{\partial x}(\bar{x}_i, V_i) - \frac{\partial \hat{a}}{\partial x}(\bar{x}_i, V_i, \bar{v}) \right) \quad (\text{C.2b})$$

$$\frac{dV_i}{dt} = V_i^2 \left( \frac{\partial^2 \hat{r}}{\partial x^2}(\bar{x}_i, V_i) - \frac{\partial^2 \hat{a}}{\partial x^2}(\bar{x}_i, V_i, \bar{v}) \right) + \hat{b}(\bar{x}_i, V_i)M \quad (\text{C.2c})$$

1656 From this, we see that we need to compute the population-level rates

$$\hat{r}(x, W) = \int r(y) \mathcal{N}(y, x, W) dy \quad (\text{C.3a})$$

$$\hat{a}(x, W, \tilde{v}) = \int a(y, \tilde{v}) \mathcal{N}(y, x, W) dy \quad (\text{C.3b})$$

1659 
$$\hat{b}(x, W) = \int b(y) \mathcal{N}(y, x, W) dy \quad (\text{C.3c})$$

The notation

$$\frac{\partial \hat{r}}{\partial x}(\bar{x}_i, V_i) \quad (\text{C.4})$$

1662 means that we take the derivative of the function  $\hat{r}(x, W)$  with respect to  $x$  and evaluate the resulting function at  $x = \bar{x}_i$  and  $W = V_i$ .

Here, we compute the population-level rates in turn. First, the population-level rate for  $r$  can  
1665 be computed as

$$\hat{r}(x, W) = \int r(y) \mathcal{N}(y, x, W) dy = \int \left( r_0 - \frac{(y - x^{\text{opt}})^2}{V_r} \right) \mathcal{N}(y, x, W) dy \quad (\text{C.5a})$$

$$= \int \left( r_0 - \frac{(y - x)^2 + 2(y - x)(x - x^{\text{opt}}) + (x - x^{\text{opt}})^2}{V_r} \right) \mathcal{N}(y, x, W) dy \quad (\text{C.5b})$$

1668 
$$= r_0 - \frac{(x - x^{\text{opt}})^2 + W}{V_r} \quad (\text{C.5c})$$

Next, we can compute the population-level rate for  $a$ . To do this, we first rewrite the competition kernel  $\alpha$  as

1671 
$$\alpha(x, y) = \exp\left(-\frac{(x - y)^2}{2V_c}\right) = \sqrt{2\pi V_c} \mathcal{N}(x, y, V_c) \quad (\text{C.6})$$

By a well known result we know that

$$\int \mathcal{N}(x, a, A) \mathcal{N}(x, b, B) dx = \mathcal{N}(a, b, A + B) \quad (\text{C.7})$$

1674 which we can now use to calculate population-level competition. We use that the trait-density distribution  $\tilde{v}(x)$  is a sum of normals  $\tilde{v}(x)$  like

$$\tilde{v}(x) = \sum_{j=1}^S u_j \mathcal{N}(x, \bar{x}_j, V_j). \quad (\text{C.8})$$



1677 We can now compute the rate  $\hat{a}(x, W, \tilde{v})$ :

$$\hat{a}(x, W, \tilde{v}) = \int a(y, \tilde{v}) \mathcal{N}(y, x, W) dy = \int \int \alpha(z, y) \tilde{v}(z) dz \mathcal{N}(y, x, W) dy \quad (\text{C.9a})$$

$$= \int \int \sqrt{2\pi V_c} \mathcal{N}(z, y, V_c) \sum_{j=1}^S u_j \mathcal{N}(z, \bar{x}_j, V_j) dz \mathcal{N}(y, x, W) dy \quad (\text{C.9b})$$

$$1680 = \sqrt{2\pi V_c} \sum_{j=1}^S u_j \int \mathcal{N}(y, \bar{x}_j, V_j + V_c) \mathcal{N}(y, x, W) dy \quad (\text{C.9c})$$

$$= \sqrt{2\pi V_c} \sum_{j=1}^S u_j \mathcal{N}(x, \bar{x}_j, V_j + V_c + W) \quad (\text{C.9d})$$

Finally, we need to compute the population-level birth rate  $\hat{b}$ , which is given by

$$\hat{b}(x, W) = \int b(y) \mathcal{N}(y, x, W) dy = \int r_0 \mathcal{N}(y, x, W) dy = r_0. \quad (\text{C.10})$$

1683 Inserting the expressions for the population-level rates and their derivatives yields the equations for the unstructured Lotka-Volterra model shown in the main text (Eq. 10).

### *Adaptive-dynamics version of the unstructured Lotka-Volterra model*

1686 Under the assumptions of adaptive dynamics, each species  $i$  is characterized by its density  $u_i$  and its trait value  $\bar{x}_i$ , which is assumed to be the same for all individuals within a species. We can then get the rate functions for the density-dynamics for the adaptive-dynamics case by letting the trait-density distribution of each species be given by

$$1689 v_i^{\text{AD}}(x) = u_i \delta(x - \bar{x}_i), \quad (\text{C.11})$$

where  $\delta$  is the Dirac delta distribution. The total distribution will thus be given by summing over all  $S$  species,

$$1692 v^{\text{AD}} = \sum_{i=1}^S u_i \delta(x - \bar{x}_i) \quad (\text{C.12})$$

Now, letting  $\hat{f}^{\text{AD}}(x)$  denote the population-level growth rate of a rare population with trait  $x$ , we have

$$1695 \hat{f}^{\text{AD}}(x, v^{\text{AD}}) = \int_{-\infty}^{\infty} f(y, v^{\text{AD}}) \delta(y - x) dy \quad (\text{C.13})$$

We thus have that the environmental rate for the adaptive-dynamics case are given by

$$\hat{r}^{\text{AD}}(x) = \int_{-\infty}^{\infty} \left( r_0 - \frac{(y - x^{\text{opt}})^2}{V_r} \right) \delta(y - x) dy = r_0 - \frac{(x - x^{\text{opt}})^2}{V_r} \quad (\text{C.14})$$

1698 by the properties of the Dirac delta distribution. The competition can be computed as

$$\begin{aligned} \hat{a}^{\text{AD}}(x, v^{\text{AD}}) &= \int_{-\infty}^{\infty} a(y, v^{\text{AD}}) \delta(y - x) dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} a(z, y) v^{\text{AD}}(z) dz \delta(y - x) dy \\ &= \sqrt{2\pi V_c} \sum_{j=1}^S u_j \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{N}(y, z, V_c) \delta(z - \bar{x}_j) dz \delta(y - x) dy \\ 1701 &= \sqrt{2\pi V_c} \sum_{j=1}^S u_j \int_{-\infty}^{\infty} \mathcal{N}(y, \bar{x}_j, V_c) \delta(y - x) dy = \sqrt{2\pi V_c} \sum_{j=1}^S u_j \mathcal{N}(x, \bar{x}_j, V_c) \end{aligned} \quad (\text{C.15a})$$

This yields the density dynamics

$$\frac{du_i}{dt} = \left[ \hat{r}^{\text{AD}}(\bar{x}_i) - \hat{a}^{\text{AD}}(\bar{x}_i, v^{\text{AD}}) \right] u_i = \left[ r_0 - \frac{\bar{x}_i^2}{V_r} - \sqrt{2\pi V_c} \sum_{j=1}^S u_j \mathcal{N}(\bar{x}_i, \bar{x}_j, V_c) \right] u_i \quad (\text{C.16})$$

1704 Under the assumptions of adaptive dynamics, mutations are rare and of small phenotypic effect.

Under these assumptions one can derive the canonical equations of adaptive dynamics, which describes how the mean traits of the species will change over time (Dieckmann and Law, 1996;

1707 Champagnat, 2003). However, since we are primarily interested in the eco-evolutionary equilibrium, we can here use a simplified version where we introduce a parameter  $\epsilon^{\text{evo}}$  which is a small number that characterizes the separation in time scale between ecological and evolutionary

1710 dynamics. Doing so yields the complete adaptive-dynamics equations

$$\frac{du_i}{dt} = \left[ \hat{r}^{\text{AD}}(\bar{x}_i) - \hat{a}^{\text{AD}}(\bar{x}_i, v^{\text{AD}}) \right] u_i = \left[ r_0 - \frac{\bar{x}_i^2}{V_r} - \sqrt{2\pi V_c} \sum_{j=1}^S u_j \mathcal{N}(\bar{x}_i, \bar{x}_j, V_c) \right] u_i \quad (\text{C.17a})$$

$$\frac{d\bar{x}_i}{dt} = \epsilon^{\text{evo}} \left[ \frac{\partial \hat{r}^{\text{AD}}}{\partial x}(\bar{x}_i) - \frac{\partial \hat{a}^{\text{AD}}}{\partial x}(\bar{x}_i, v^{\text{AD}}) \right] = \epsilon^{\text{evo}} \left[ -\frac{2\bar{x}_i}{V_r} + \sqrt{2\pi V_c} \sum_{j=1}^S u_j \frac{\bar{x}_i - \bar{x}_j}{V_c} \mathcal{N}(\bar{x}_i, \bar{x}_j, V_c) \right] \quad (\text{C.17b})$$

1713 In the adaptive dynamics context, the methods for performing evolutionary-branching and invasion analysis are well established (Metz et al., 1992; Geritz et al., 1998; Edwards et al., 2018; Klausmeier et al., 2020), and we assemble the community one species at a time until the community is uninviable or use continuation to make bifurcation diagrams.

1716

## The two-patch Lotka-Volterra model

For the two-patch model we consider two spatial patches with trait-density distributions  $v_1(x)$

1719 and  $v_2(x)$ , with trait-space equations given by

$$\frac{dv_1(x)}{dt} = \int b_1(y)v_1(y)\mathcal{N}(y, x, M)dy - \mu_1(x)v_1(x) - a(x, v_1)v_1(x) - dv_1(x) + dv_2(x), \quad (\text{C.18a})$$

$$\frac{dv_2(x)}{dt} = \int b_2(y)v_2(y)\mathcal{N}(y, x, M)dy - \mu_2(x)v_2(x) - a(x, v_2)v_2(x) + dv_1(x) - dv_2(x), \quad (\text{C.18b})$$

$$1722 \quad b_k(x) = r_0, \quad \mu_k(x) = \frac{(x - x_k^{\text{opt}})^2}{V_r}, \quad (\text{C.18c})$$

$$r_k(x) := b_k(x) - \mu_k(x) = r_0 - \frac{(x - x_k^{\text{opt}})^2}{V_r}, \quad (\text{C.18d})$$

$$a(x, v_k) = \int \alpha(y, x)v_k(y)dy, \quad \alpha(y, x) = \exp\left(-\frac{(y-x)^2}{2V_c}\right). \quad (\text{C.18e})$$

1725 The two-patch model generalizes the unstructured model in that locally on each patch the same dynamics as in the unstructured model takes place, albeit with possibly different trait optima  $x_1^{\text{opt}}$  and  $x_2^{\text{opt}}$ . The patches are then connected by a symmetric and constant rate of dispersal  $d$ .

1728 We can then get the moment equations for this model by plugging in the rates into our generic moment equations which yields

$$\frac{du_{i1}}{dt} = [\hat{r}_1(\bar{x}_{i1}, V_{i1}) - \hat{a}(\bar{x}_{i1}, V_{i1}, \tilde{v}_1)] u_{i1} - du_{i1} + du_{i2}, \quad (\text{C.19a})$$

$$1731 \quad \frac{d\bar{x}_{i1}}{dt} = V_{i1} \left[ \frac{\partial \hat{r}_1}{\partial x}(\bar{x}_{i1}, V_{i1}) - \frac{\partial \hat{a}}{\partial x}(\bar{x}_{i1}, V_{i1}, \tilde{v}_1) \right] + \frac{u_{i2}}{u_{i1}} d(\bar{x}_{i2} - \bar{x}_{i1}), \quad (\text{C.19b})$$

$$\begin{aligned} \frac{dV_{i1}}{dt} &= V_{i1}^2 \left[ \frac{\partial^2 \hat{r}_1}{\partial x^2}(\bar{x}_{i1}, V_{i1}) - \frac{\partial^2 \hat{a}}{\partial x^2}(\bar{x}_{i1}, V_{i1}, \tilde{v}_1) \right] + \frac{u_{i2}}{u_{i1}} d(V_{i2} - V_{i1}) + \frac{u_{i2}}{u_{i1}} d(\bar{x}_{i2} - \bar{x}_{i1})^2 \\ &\quad + \hat{b}_1(\bar{x}_{i1}, V_{i1})M, \end{aligned} \quad (\text{C.19c})$$

$$1734 \quad \frac{du_{i2}}{dt} = [\hat{r}_2(\bar{x}_{i2}, V_{i2}) - \hat{a}(\bar{x}_{i2}, V_{i2}, \tilde{v}_2)] u_{i2} + du_{i1} - du_{i2}, \quad (\text{C.19d})$$

$$\frac{d\bar{x}_{i2}}{dt} = V_{i2} \left[ \frac{\partial \hat{r}_2}{\partial x}(\bar{x}_{i2}, V_{i2}) - \frac{\partial \hat{a}}{\partial x}(\bar{x}_{i2}, V_{i2}, \tilde{v}_2) \right] + \frac{u_{i1}}{u_{i2}} d(\bar{x}_{i1} - \bar{x}_{i2}), \quad (\text{C.19e})$$

$$\begin{aligned} \frac{dV_{i2}}{dt} &= V_{i2}^2 \left[ \frac{\partial^2 \hat{r}_2}{\partial x^2}(\bar{x}_{i2}, V_{i2}) - \frac{\partial^2 \hat{a}}{\partial x^2}(\bar{x}_{i2}, V_{i2}, \tilde{v}_2) \right] + \frac{u_{i1}}{u_{i2}} d(V_{i1} - V_{i2}) + \frac{u_{i1}}{u_{i2}} d(\bar{x}_{i1} - \bar{x}_{i2})^2 \\ &\quad + \hat{b}_2(\bar{x}_{i2}, V_{i2})M. \end{aligned} \quad (\text{C.19f})$$

1737

Inserting the expressions for the population-level rates yields the equations in the main text (Eqs. 10).

1740 *Adaptive-dynamics version of the two-patch Lotka-Volterra model*

For the two-patch adaptive-dynamics model, we proceed in the same way as for the unstructured adaptive-dynamics model. We let  $u_{ik}$  be the density of species  $i$  on patch  $k$  and let  $\bar{x}_i$  be the trait value of species  $i$ . Note that in adaptive dynamics, since all individuals within a species are assumed to be identical the trait value for a species will be the same on both patches. We let the trait-density distribution  $v_k^{\text{AD}}(x)$  on patch  $k$  be given by

$$1746 \quad v_k^{\text{AD}}(x) = \sum_{i=1}^S u_{ik} \delta(x - \bar{x}_i). \quad (\text{C.20})$$

The expression for the selection gradient needs to take the spatial dynamics into account (see Wickman et al., 2017), and so the equations describing the density and trait dynamics are given by

$$\frac{du_{i1}}{dt} = \left[ \hat{r}_1^{\text{AD}}(\bar{x}_i) - \hat{a}^{\text{AD}}(\bar{x}_i, v_1^{\text{AD}}) \right] u_{i1} - du_{i1} + du_{i2} \quad (\text{C.21a})$$

$$\frac{du_{i2}}{dt} = \left[ \hat{r}_2^{\text{AD}}(\bar{x}_i) - \hat{a}^{\text{AD}}(\bar{x}_i, v_2^{\text{AD}}) \right] u_{i2} + du_{i1} - du_{i2} \quad (\text{C.21b})$$

$$1752 \quad \frac{d\bar{x}_i}{dt} = \epsilon^{\text{evo}} \frac{1}{u_{i1}^2 + u_{i2}^2} \left[ u_{i1}^2 \left( \frac{\partial \hat{r}_1^{\text{AD}}}{\partial x}(\bar{x}_i) - \frac{\partial \hat{a}^{\text{AD}}}{\partial x}(\bar{x}_i, v_1^{\text{AD}}) \right) + u_{i2}^2 \left( \frac{\partial \hat{r}_2^{\text{AD}}}{\partial x}(\bar{x}_i) - \frac{\partial \hat{a}^{\text{AD}}}{\partial x}(\bar{x}_i, v_2^{\text{AD}}) \right) \right] \quad (\text{C.21c})$$

The community assembly then proceeds by solving the equations to equilibrium and checking for evolutionary branchings and invasions, see Wickman et al. (2017).

1755 **Online Appendix D: A stage-structured resource-competition model**

To further illustrate the application and utility of our methods, we present here an additional example of eco-evolutionary community assembly taking intraspecific variation into account. The Lotka-Volterra two-patch model we used to illustrate the generic equations in the main text was deliberately simple. First, it involved only a single trait. Second, the transition between classes (dispersal) was constant, and did not depend on traits. Third, the rate functions were simple enough for us to calculate the population-level rates analytically, and fourth, the model had no dependence on any external variables. However, such simplifications cannot always be made. To demonstrate such a case, we here adapt a model by De Roos et al. (2007) of a population subdivided into two stages, juveniles with trait-density distribution  $v_J$  and adults with trait-density distribution  $v_A$ , competing for two resources  $R_1$  and  $R_2$ . Both the maturation rate and the birth rate depend on how much resources the two stages acquire, as well as the traits associated with resource acquisition.

The trait-space equations read

$$\frac{dv_J(x,t)}{dt} = \int r_A(y)v_A(y)\mathcal{N}(x,y,M)dy + (r_J(x) - \mu_J(x))v_J(x) - \gamma(x)v_J(x), \quad (\text{D.1a})$$

$$\frac{dv_A(x,t)}{dt} = \gamma(x)v_J(x) - \mu_A(x)v_A(x), \quad (\text{D.1b})$$

$$\frac{dR_1(t)}{dt} = \rho_1(K_1 - R_1) - \int r_{J1}(y)v_J(y)dy - \int r_{A1}(y)v_A(y)dy, \quad (\text{D.1c})$$

$$\frac{dR_2(t)}{dt} = \rho_2(K_2 - R_2) - \int r_{J2}(y)v_J(y)dy - \int r_{A2}(y)v_A(y)dy, \quad (\text{D.1d})$$

1773 where the rate functions are given by

$$r_I(x) = q_I \frac{a_{I1}(x)R_1 + a_{I2}(x)R_2}{1 + a_{I1}(x)R_1 + a_{I2}(x)R_2}, \quad (\text{D.2a})$$

$$r_{Ij}(x) = q_I \frac{a_{Ij}(x)R_j}{1 + a_{I1}(x)R_1 + a_{I2}(x)R_2} \quad I \in \{J, A\}, j \in \{1, 2\}, \quad (\text{D.2b})$$

$$\gamma = \frac{r_J(x) - \mu_J(x)}{1 - z^{1 - \mu_J(x)/r_J(x)}}, \quad (\text{D.2c})$$

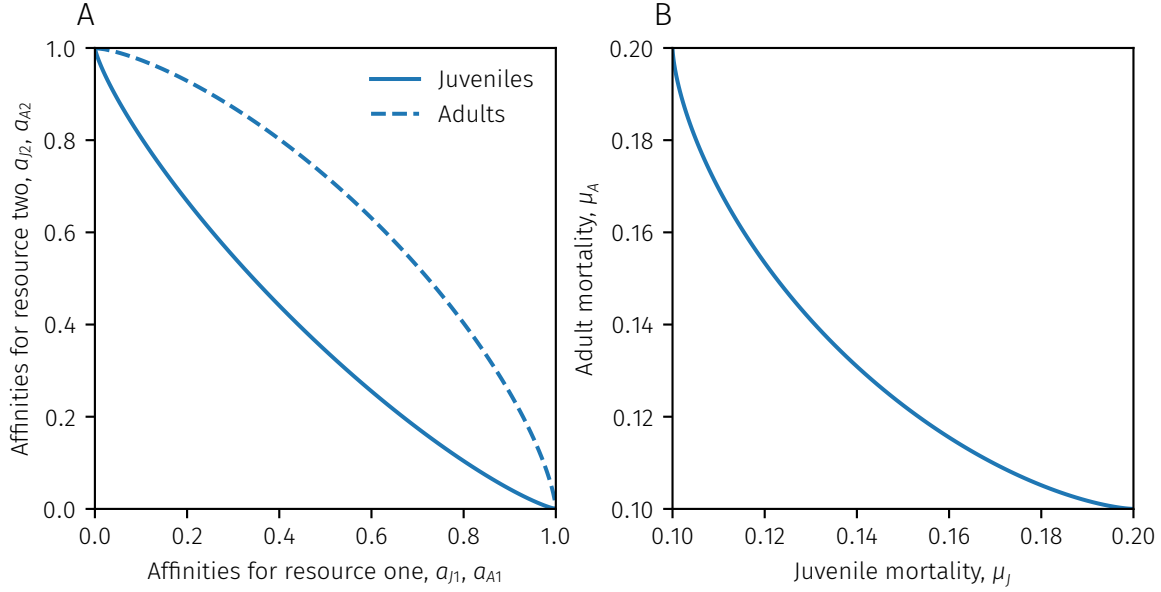
$$a_{J1}(x) = \left( \frac{e^x}{1 + e^x} \right)^{\alpha_J}, \quad a_{J2} = \left( 1 - \frac{e^x}{1 + e^x} \right)^{\alpha_J}, \quad (\text{D.2d})$$

$$a_{A1}(x) = \left( \frac{e^x}{1 + e^x} \right)^{\alpha_A}, \quad a_{A2} = \left( 1 - \frac{e^x}{1 + e^x} \right)^{\alpha_A}, \quad (\text{D.2e})$$

1779 Here, the distributions  $v_J(x)$  and  $v_A(x)$  characterize the distributions of biomass density in trait space for juveniles and adults respectively. Both adults and juveniles consume the substitutable resources  $R_1$  and  $R_2$  with a type-II functional response at rates  $r_{Ij}$  for  $I = A, J$ , and  $j = 1, 2$ .

1782 The adults use all of their intake of resources  $R_1$  and  $R_2$  to give birth to juveniles at a per capita rate of  $r_A = r_{A1} + r_{A2}$ , and suffer a mortality at a per capita rate of  $\mu_A$ . Juveniles grow at a per capita rate  $r_J = r_{J1} + r_{J2}$ , suffer a mortality rate of  $\mu_J$ , and transition into adults at a rate  
1785  $\gamma$  that depends both on juvenile growth and mortality (De Roos et al., 2007). In the absence of consumers, the resources are renewed chemostatically at rates  $\rho_j$  up to maximal supplies of  $K_j$  for resource  $j$ .

1788 We let the trait vector  $x = (x_1, x_2) \in \mathbb{R}^2$ . The first trait  $x_1$  parameterizes trade-offs in resource affinities, so that juveniles experience a trade-off between their affinities  $a_{J1}$  and  $a_{J2}$  for  $R_1$  and  $R_2$  respectively, and adults experience a trade-off between their affinities  $a_{A1}$  and  $a_{A2}$ . While  
1791 the same trait component  $x_1$  parameterizes the trade off in resource affinity for both juveniles and adults, we let the shapes differ so that the juveniles experience a specialist-favoring trade-off while adults experience a generalist-favoring trade-off (Fig. D.1A). Such differences could  
1794 come about as foraging behavior changes between life stages. The second trait  $x_2$  parameterizes a trade-off between juvenile mortality  $\mu_J$  and adult mortality  $\mu_A$  (Fig. D.1B). Such trade-offs between juvenile and adult mortality could come about through, e.g., parental investment or risk  
1797 taking on behalf of juveniles.



**Figure D.1: Trade-offs in the stage-structured model. (A)** Trade-offs in resource affinities  $a_{J1}$  and  $a_{J2}$  for juveniles (solid line) and  $a_{A1}$  and  $a_{A2}$  for adults (dashed line). Juveniles experience a specialist-favoring trade-off, and adults experience a generalist-favoring trade-off. **(B)** Trade-off between juvenile and adult mortality. The trade-off between the mortalities is generalist-favoring.

1800 Following our general procedure, to get our moment equations we now assume that the solutions to Eqs. D.1a–b are composed of a set of  $S$  normally distributed discrete peaks for both juveniles and adults, and approximate the solutions by

1803 
$$v_J(x, t) \approx \tilde{v}_J(x, t) = \sum_{i=1}^S u_{Ji}(t) \mathcal{N}(x, \bar{x}_{Ji}(t), V_{Ji}(t)), \quad (\text{D.3a})$$

$$v_A(x, t) \approx \tilde{v}_A(x, t) = \sum_{i=1}^S u_{Ai}(t) \mathcal{N}(x, \bar{x}_{Ai}(t), V_{Ai}(t)). \quad (\text{D.3b})$$

This means that each species  $i = 1, 2, \dots, S$  is represented by six variables: the total densities of juveniles,  $u_{Ji}$ , and adults,  $u_{Ai}$ ; the mean trait vector for juveniles,  $\bar{x}_{Ji}$ , and adults,  $\bar{x}_{Ai}$ ; and the trait variance-covariance matrices for juveniles,  $V_{Ji}$ , and adults,  $V_{Ai}$ . To be able to write down our moment equations we must first deal with the population-level rate functions. For this example it is not possible to analytically calculate the population-level rates given by, for example,

$$\hat{\gamma}(x, W) = \int \gamma(y) \mathcal{N}(y, x, W) dy \quad (\text{D.4})$$

Instead, we Taylor-approximate the various rates to third order using the expressions derived in

1812 Appendix A, section *Series solutions and Taylor approximations*, so that our various population-level rates and their derivatives will be given by

$$\hat{f}(x, W) \approx f(x) + \frac{1}{2} \nabla_x^2 f(x) : W, \quad (\text{D.5a})$$

$$1815 \quad \nabla_x \hat{f}(x, W) \approx \nabla_x f(x) + \nabla_x \left( \frac{1}{2} \nabla_x^2 f(x) : W \right), \quad (\text{D.5b})$$

$$\nabla_x^2 \hat{f}(x, W) \approx \nabla_x^2 f(x). \quad (\text{D.5c})$$

Here, “:” denotes the double-dot product, which for two square  $N \times N$  matrices  $A$  and  $B$  with entries  $A_{ij}$  and  $B_{ij}$  is given by  $A : B = \sum_{i=1}^N \sum_{j=1}^N A_{ij} B_{ij}$ . To make the moment equations more readable we will also omit the arguments of the rate functions below, and so we shall write for example

$$1821 \quad \hat{\gamma}_i := \hat{\gamma}(\bar{x}_{Ji}, V_{Ji}) + \frac{1}{2} \nabla_x^2 \gamma(\bar{x}_{Ji}) : V_{Ji}, \quad (\text{D.6a})$$

$$\nabla_x \hat{\gamma}_i := \nabla_x \gamma(\bar{x}_{Ji}) + \left[ \nabla_x \left( \frac{1}{2} \nabla_x^2 \gamma(x) : V_{Ji} \right) \right]_{x=\bar{x}_{Ji}}, \quad (\text{D.6b})$$

$$\nabla_x^2 \hat{\gamma}_i := \nabla_x^2 \gamma(\bar{x}_{Ji}), \quad (\text{D.6c})$$

$$1824 \quad \hat{r}_{Ai} := r_A(\bar{x}_{Ai}) + \frac{1}{2} \nabla_x^2 r_A(\bar{x}_{Ai}) : V_{Ai} \quad (\text{D.6d})$$

$$\nabla_x \hat{r}_{Ai} := \nabla_x r_A(\bar{x}_{Ai}) + \left[ \nabla_x \left( \frac{1}{2} \nabla_x^2 r_A(x) : V_{Ai} \right) \right]_{x=\bar{x}_{Ai}}, \quad (\text{D.6e})$$

$$\nabla_x^2 \hat{r}_{Ai} := \nabla_x^2 r_A(\bar{x}_{Ai}), \quad (\text{D.6f})$$

1827 and so on.

We can now derive the moment equations for the stage-structured model by plugging in the expressions for the trait-space equations for this model into our generic equations (Eq. A.28,

1830 Appendix A) yielding

$$\frac{du_{Ji}}{dt} = \hat{g}_{Ji} u_{Ji} - \hat{\gamma}_i u_{Ji} + \hat{r}_{Ai} u_{Ai}, \quad (\text{D.7a})$$

$$\frac{d\bar{x}_{Ji}}{dt} = V_{Ji} (\nabla_x \hat{g}_{Ji} - \nabla_x \hat{\gamma}_i) + \frac{u_{Ai}}{u_{Ji}} (V_{Ai} \nabla_x \hat{r}_{Ai} + \hat{r}_{Ai} (\bar{x}_{Ai} - \bar{x}_{Ji})), \quad (\text{D.7b})$$

$$1833 \quad \frac{dV_{Ji}}{dt} = V_{Ji} (\nabla_x^2 \hat{g}_{Ji} - \nabla_x^2 \hat{\gamma}_i) V_{Ji}$$



$$\begin{aligned}
& + \frac{u_{Ai}}{u_{Ji}} \left( V_{Ai} \nabla_x^2 \hat{r}_{Ai} V_{Ai} + \hat{r}_{Ai} (V_{Ai} - V_{Ji}) + \hat{r}_{Ai} M \right. \\
& + \hat{r}_{Ai} (\bar{x}_{Ai} - \bar{x}_{Ji}) (\bar{x}_{Ai} - \bar{x}_{Ji})^T \\
& \left. + V_{Ai} \nabla_x \hat{r}_{Ai} (\bar{x}_{Ai} - \bar{x}_{Ji})^T + (\bar{x}_{Ai} - \bar{x}_{Ji}) (V_{Ai} \nabla_x \hat{r}_{Ai})^T \right), \tag{D.7c}
\end{aligned}$$

1836

$$\frac{du_{Ai}}{dt} = \hat{\gamma}_i u_{Ji} - \hat{\mu}_{Ai} u_{Ai}, \tag{D.7d}$$

$$\frac{d\bar{x}_{Ai}}{dt} = \frac{u_{Ji}}{u_{Ai}} (V_{Ji} \nabla_x \hat{\gamma}_i + \hat{\gamma}_i (\bar{x}_{Ji} - \bar{x}_{Ai})) - V_{Ai} \nabla_x \hat{\mu}_{Ai}, \tag{D.7e}$$

1839

$$\begin{aligned}
\frac{dV_{Ai}}{dt} &= \frac{u_{Ji}}{u_{Ai}} \left( V_{Ji} \nabla_x^2 \hat{\gamma}_i V_{Ji} + \hat{\gamma}_i (V_{Ji} - V_{Ai}) \right. \\
& + \hat{\gamma}_i (\bar{x}_{Ji} - \bar{x}_{Ai}) (\bar{x}_{Ji} - \bar{x}_{Ai})^T \\
& + V_{Ji} \nabla_x \hat{\gamma}_i (\bar{x}_{Ji} - \bar{x}_{Ai})^T + (\bar{x}_{Ji} - \bar{x}_{Ai}) (V_{Ji} \nabla_x \hat{\gamma}_i)^T \left. \right) \\
& - V_A \nabla_x^2 \hat{\mu}_{Ai} V_A, \tag{D.7f}
\end{aligned}$$

1842

$$\frac{dR_1}{dt} = \rho_1 (K_1 - R_1) - \sum_{i=1}^S \hat{r}_{Ji1} u_{Ji} - \sum_{i=1}^S \hat{r}_{Ai1} u_{Ai}, \tag{D.7g}$$

$$\frac{dR_2}{dt} = \rho_2 (K_2 - R_2) - \sum_{i=1}^S \hat{r}_{Ji2} u_{Ji} - \sum_{i=1}^S \hat{r}_{Ai2} u_{Ai}. \tag{D.7h}$$

1845

Here, we let  $g_{Ji}$  be the net per capita growth of juveniles, i.e.,  $g_{Ji} = r_{Ji} - \mu_{Ji}$ . While somewhat involved, the various terms admit to ecological interpretation, and can be understood in the following way.

1848

The terms of Eq. D.7a describe in turn: net growth in mass of juveniles, the removal of juveniles that transition into adults, and the addition of mass due to birth by adults.

1851

The changes to the mean trait vector of the juveniles  $\bar{x}_{Ji}$  (Eq. D.7b) are divided into two terms. The first, with a matrix factor  $V_{Ji}$ , describes how directional selection on the juvenile processes of net growth and transition to adulthood moves the mean trait of the juveniles. This is because juveniles with certain traits will have higher net growth than others, and this will push the mean trait of juveniles in the direction of the gradient  $\nabla_x \hat{\gamma}_i$ . Conversely, some juveniles with certain traits will mature into adulthood more quickly than others, which removes them from the juveniles and thus moves the mean trait in the opposite direction of the gradient  $\nabla_x \hat{\gamma}_i$ . The

1854

1857 strength of both these effects depends on the how much variation  $V_{Ji}$  there is for a given species  $i$ .  
 The second term, with a scalar factor  $u_{Ai}/u_{Ji}$  describes the influence of births from the adults on  
 the mean trait of the juveniles. This is, in turn, composed of two terms. The first term,  $V_{Ai}\nabla_x\hat{r}_{Ai}$ ,  
 1860 describes how directional selection in the birth rate of juveniles by adults contributes towards  
 change of juvenile mean traits. This effect comes about since adults with different traits give  
 birth at different rates, and more offspring will be produced along the direction of the gradient  
 1863  $\nabla_x\hat{r}_{Ai}$ . This effect is stronger if the variation among the adults  $V_{Ai}$  is larger. The second term,  
 $\hat{r}_{Ai}(\bar{x}_{Ai} - \bar{x}_{Ji})$ , describes how the juvenile mean trait is pushed in the direction of the adult mean  
 trait, homogenizing juvenile and adult mean traits. This is due to the fact that adults of species  
 1866  $i$  will on average give birth to juveniles with trait  $\bar{x}_{iA}$ , and so, if there is a difference between  
 juvenile and adult mean trait, the juvenile trait will become more similar to the adult mean trait  
 over time. The overall factor  $u_{Ai}/u_{Ji}$  comes about through the fact that how much effect birth  
 1869 of juveniles by adults has on the juvenile mean trait depends on the mass ratio between the two  
 stages. If there are few adults compared to juveniles, then the birth of new juveniles will have  
 little effect on the mean trait of all the juveniles.

1872 Equation D.7c describes the changes to the variance-covariance matrix of the juveniles over  
 time. The first term,  $V_{Ji}(\nabla_x^2g_{Ji} - \nabla_x^2\gamma_i)V_{Ji}$ , describes the effects of stabilizing or disruptive se-  
 lection of the net growth and transitions. Thus, for example, if the curvature of the net growth  
 1875 function,  $\nabla_x^2g_{Ji}$ , is negative around a mean trait component, this will reduce the variance of that  
 trait component over time. There are then several terms with a factor  $u_{Ai}/u_{Ji}$ . The first term,  
 $V_{Ai}\nabla_x^2r_{Ai}V_{Ai}$ , reflects the effects of stabilizing or disruptive selection in the birth rates. The sec-  
 1878 ond term,  $\hat{r}_{Ai}(V_{Ai} - V_{Ji})$ , describes the homogenization of variance-covariance matrices between  
 the stages, so that this term changes the juvenile variance-covariance to be more similar to that of  
 the adults over time. The third term,  $\hat{r}_{Ai}M$ , describes the increase to juvenile trait variances due  
 1881 to mutations. Next, the term  $\hat{r}_{Ai}(\bar{x}_{Ai} - \bar{x}_{Ji})(\bar{x}_{Ai} - \bar{x}_{Ji})^T$  describes how variances are increased, and  
 covariances changed over time due to differences in means between the juveniles and adults. This  
 reflects that if the means of the adults are away from the means of the juveniles, then birthed

1884 juveniles will be away from the mean of the extant juveniles, and thus increase variance. The  
 next two terms,  $V_{Ai} \nabla_x r_{Ai} (\bar{x}_{Ai} - \bar{x}_{Ji})^T + (\bar{x}_{Ai} - \bar{x}_{Ji}) (V_{Ai} \nabla_x r_{Ai})^T$ , which are the transposes of one  
 another, take into account the interaction of distance between the means of the stages and effects  
 1887 of directional selection in the birth rates.

The terms in the equations for the adults are mostly analogous to those for the juveniles, with  
 the difference that the maturation function  $\gamma$  now plays the role that the birth rate function  $r_A$   
 1890 did for the juveniles, as this is the function that determines the rate of transition from the juvenile  
 to the adult stage. Additionally, as we can see in the trait-space equations (Eqs. D.1), we assume  
 that a juvenile with trait  $x$  matures into an adult with the same trait  $x$ , and thus no mutation term  
 1893 enters into the equations for adult variance-covariance. Finally, we approximate the integrals in  
 trait space in the resource equations (Eqs. D.1c–d) in the same way as we do for the other rates  
 by Taylor-expanding to third order.

1896 Using these equations we let the population go through our evolutionary assembly process  
 (Fig. D.2A). The traits  $x_1$  and  $x_2$  parameterize the trade-offs in resource affinities and mortal-  
 ities respectively. However, as the traits that actually matter for the ecology are the affini-  
 1899 ties and mortalities themselves, we depict this assembly process in the derived trait space of  
 $(a_{I1} - a_{I2}) \times (\mu_J - \mu_A)$  for  $I = J, A$  to get a better understanding of how the expressed traits  
 that impact the ecology change. The final community of one moderate  $R_1$  specialist and one  
 1902 moderate  $R_2$  specialist comes about through the conflicting trade-offs for juveniles and adults.  
 The specialist-favoring trade-off in resource affinity for the juveniles engenders selective pressure  
 towards specialization into two species, but the generalist-favoring trade-off for the adults will  
 1905 eventually counter balance this disruptive force, and the community will settle on substantial,  
 but not complete, specialization for both species. The differences between juveniles and adults  
 in their resource affinities seen in Fig. D.2A stem mostly from their their different trade-offs; the  
 1908 mean traits  $\bar{x}_J$  and  $\bar{x}_A$  differ little within the same species. The distributions in  $x_1 \times x_2$  space  
 can be seen in Fig. D.3, where we also compare the numerical moment solutions for the eco-  
 evolutionarily stable two-species community to numerical solutions of the trait-space equations.

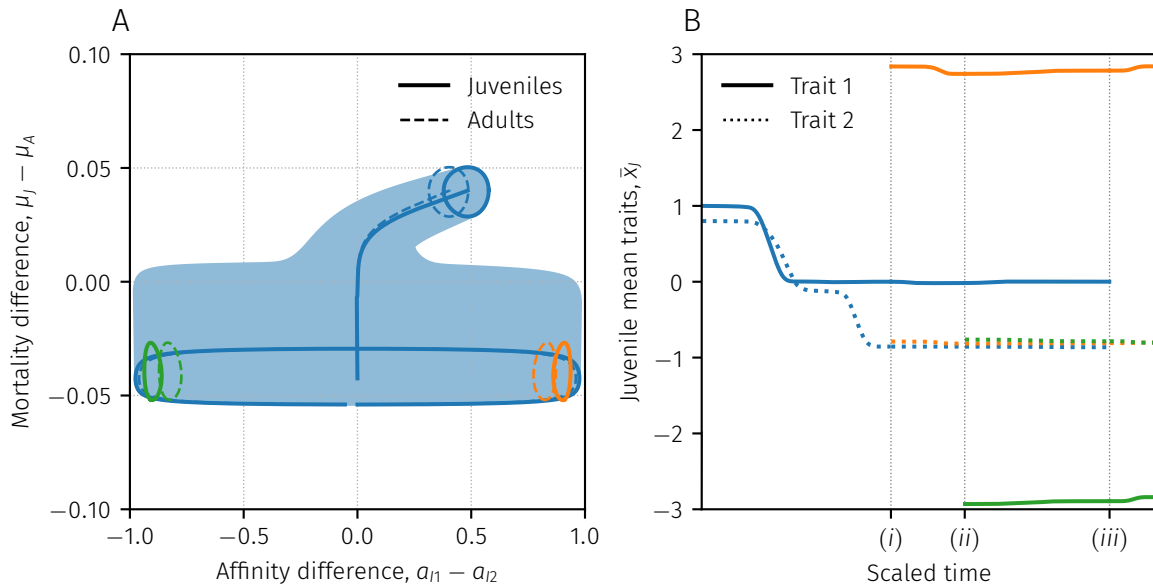
1911 Solving the trait-space equations Eq. D.1 for this system with these given parameters would  
require an unfeasible numerical resolution in trait space. However, since the mutation variance-  
covariance matrix is very small, without covariance, and with the same variances, we can ap-  
1914 proximate the birth/mutation integral with a diffusion term in the following way (Kimura, 1965;  
Débarre et al., 2013):

$$\int r_A(y)v_A(y)\mathcal{N}(x,y,M)dy \approx r_A(x)v_A(x) + \frac{M_{11}}{2}\Delta_x(r_A(x)v_A(x)), \quad (\text{D.8})$$

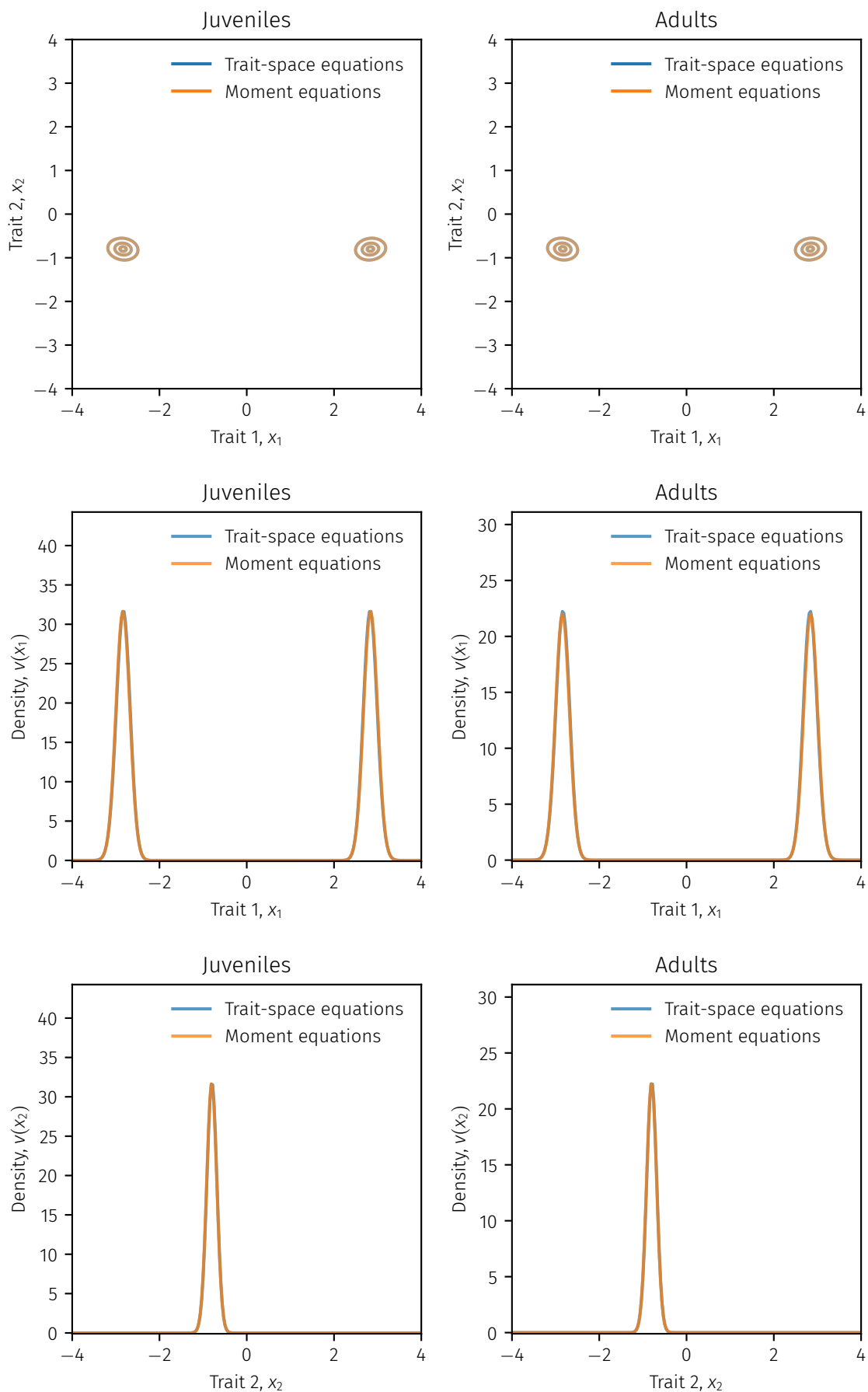
1917 where  $\Delta_x = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}$  is the Laplacian in trait space. We discretize two-dimensional trait space  
into an equispaced grid of  $256 \times 256$  points with  $x_1, x_2 \in [-4, 4]$ , and approximate the Laplacian  
with a central-difference approximation. We then used the `DifferentialEquations.jl` library  
1920 (Rackauckas and Nie 2017) in Julia (Bezanson et al. 2017) to solve the semi-discretized ordinary  
differential equations to equilibrium. Both trait-space equations and moment equations were  
solved to equilibrium.

## 1923 Literature Cited

- Albert, C. H., F. Grassein, F. M. Schurr, G. Vieilledent, and C. Violle. 2011. When and how  
should intraspecific variability be considered in trait-based plant ecology? Perspectives in  
1926 Plant Ecology, Evolution and Systematics 13:217–225.
- Albert, C. H., W. Thuiller, N. G. Yoccoz, A. Soudant, F. Boucher, P. Saccone, and S. Lavorel.  
2010. Intraspecific functional variability: extent, structure and sources of variation. Journal of  
1929 Ecology 98:604–613.
- Araújo, M. S., D. I. Bolnick, and C. A. Layman. 2011. The ecological causes of individual special-  
isation. Ecology Letters 14:948–958.
- 1932 Barabás, G., C. Parent, A. Kraemer, F. Van de Perre, and F. De Laender. 2022. The evolution  
of trait variance creates a tension between species diversity and functional diversity. Nature  
Communications 13:2521.



**Figure D.2: Community assembly in the stage-structured example.** (A) Depiction of the assembly process and final eco-evolutionary equilibrium in the derived trait space  $(a_{I1} - a_{I2}) \times (\mu_J - \mu_A)$  for  $I = J, A$ . The eco-evolutionary process is started towards the top-right corner, where the evolutionary paths of the mean derived traits  $a_{I1} - a_{I2}$  and  $\mu_J - \mu_A$  are depicted as a solid blue line for juveniles and broken blue line for adults. The 95% transformed confidence ellipse for the juveniles is depicted at each time-point around the means in lighter color. The evolutionary process continues until the system reaches a one-species equilibrium. The 95% transformed confidence ellipses are shown as a solid blue ellipse (juvenile) and dashed blue ellipse (adult). This one-species equilibrium is subsequently invaded by two specialist species, and the final two-species equilibrium is depicted in orange and green. (B) Evolution and assembly of juvenile mean traits over time. Trait 1 (solid line) governs the resource-affinity trade-off and trait 2 (dotted line) governs the juvenile-adult mortality trade-off. The process starts with one species reaching an eco-evolutionary equilibrium, but is invaded by species two (orange) at (i). The two-species community subsequently reaches an eco-evolutionary equilibrium, which is in turn invaded by a third species (green) at (ii). With two specialists to compete with, the generalist species (blue) cannot persist and goes extinct at (iii). The new two-species community then evolves to a new equilibrium which is stable to further invasion, and the assembly process is complete. Time has been rescaled non-uniformly to illustrate the process.



**Figure D.3: Comparison of trait-space equations and moment equations for the stage-structured model.** Top row shows contour lines for the trait-space and moment equation solutions at eco-evolutionary equilibrium. The contour lines are very nearly on top of each other. The bottom four panels show slices in the  $x_1$  and  $x_2$  directions, where both slices are made across the global density maximum of the trait-space solution. The curves are very nearly on top of each other.

Symbol	Description	Value
$\rho_1$	Resource 1 renewal rate	1.0
$\rho_2$	Resource 2 renewal rate	1.0
$K_1$	Resource 1 supply	1.0
$K_2$	Resource 2 supply	1.0
$q_J$	Juvenile conversion efficiency	1.0
$q_A$	Adult conversion efficiency	1.0
$z$	Juvenile to adult size ratio	0.3
$\mu_{0J}$	Juvenile base mortality	0.1
$\mu_{0A}$	Adult base mortality	0.1
$\alpha_J$	Juvenile resource trade-off shape parameter	1.3
$\alpha_A$	Adult resource trade-off shape parameter	0.7
$\alpha_\mu$	Juvenile–adult mortality trade-off shape parameter	1.5
$M$	Mutation variance-covariance matrix	$10^{-5} \begin{bmatrix} 1.0 & 0.0 \\ 0.0 & 1.0 \end{bmatrix}$

**Table D.1: Parameter values used for numerical solutions of the stage-structured example.**

- 1935 Barabás, G., and R. D'Andrea. 2016. The effect of intraspecific variation and heritability on  
community pattern and robustness. *Ecology Letters* 19:977–986.
- Barton, N., A. Etheridge, and A. Véber. 2017. The infinitesimal model: Definition, derivation,  
1938 and implications. *Theoretical Population Biology* 118:50–73.
- Bolnick, D. I., P. Amarasekare, M. S. Araújo, R. Bürger, J. M. Levine, M. Novak, V. H. Rudolf, S. J.  
Schreiber, M. C. Urban, and D. A. Vasseur. 2011. Why intraspecific trait variation matters in  
1941 community ecology. *Trends in Ecology & Evolution* 26:183–192.
- Brännström, Å., J. Johansson, and N. von Festenberg. 2013. The hitchhiker's guide to adaptive  
dynamics. *Games* 4:304–328.
- 1944 Bruggeman, J. 2009. Succession in plankton communities: A trait-based perspective. Ph.D. thesis.  
Vrije Universiteit, Amsterdam.
- Bürger, R. 1986. On the maintenance of genetic variation: global analysis of kimura's continuum-  
1947 of-alleles model. *Journal of Mathematical Biology* 24:341–351.
- Champagnat, N. 2003. Convergence of adaptive dynamics n-morphic jump processes to the  
canonical equation and degenerate diffusion approximation. *Prépublication de l'Université de*  
1950 *Nanterre (Paris X) no 3.*
- Champagnat, N., R. Ferrière, and S. Méléard. 2006. Unifying evolutionary dynamics: From  
individual stochastic processes to macroscopic models. *Theoretical Population Biology* 69:297–  
1953 321. *ESS Theory Now.*
- Costa-Pereira, R., V. H. W. Rudolf, F. L. Souza, and M. S. Araújo. 2018. Drivers of individual  
niche variation in coexisting species. *Journal of Animal Ecology* 87:1452–1464.
- 1956 Cropp, R., and J. Norbury. 2021. Modelling the evolution of naturally bounded traits in a popu-  
lation. *Theoretical Ecology* .



- 1959 Débarre, F., and S. Gandon. 2010. Evolution of specialization in a spatially continuous environment. *Journal of Evolutionary Biology* 23:1090–1099.
- Débarre, F., S. L. Nuismer, and M. Doebeli. 2014. Multidimensional (co)evolutionary stability. *The American Naturalist* 184:158–171.
- 1962 Débarre, F., O. Ronce, and S. Gandon. 2013. Quantifying the effects of migration and mutation on adaptation and demography in spatially heterogeneous environments. *Journal of Evolutionary Biology* 26:1185–1202.
- 1965 Dercole, F., and S. Rinaldi. 2008. Analysis of evolutionary processes: the adaptive dynamics approach and its applications. Princeton Series in Theoretical and Computational Biology. Princeton University Press, Princeton, NJ.
- 1968 Dieckmann, U., and M. Doebeli. 1999. On the origin of species by sympatric speciation. *Nature* 400:354.
- Dieckmann, U., and R. Law. 1996. The dynamical theory of coevolution: a derivation from stochastic ecological processes. *Journal of Mathematical Biology* 34:579–612.
- 1971 Edwards, K. F., C. T. Kremer, E. T. Miller, M. M. Osmond, E. Litchman, and C. A. Klausmeier. 2018. Evolutionarily stable communities: a framework for understanding the role of trait evolution in the maintenance of diversity. *Ecology Letters* 21:1853–1868.
- 1974 Fortelius, M., S. A. H. Geritz, M. Gyllenberg, and J. Toivonen. 2015. Adaptive dynamics on an environmental gradient that changes over a geological time-scale. *Journal of Theoretical Biology* 376:91–104.
- 1977 Gaudard, C. A., M. P. Robertson, and T. R. Bishop. 2019. Low levels of intraspecific trait variation in a keystone invertebrate group. *Oecologia* 190:725–735.
- 1980 Geritz, S. A. H., E. Kisdi, G. Meszéna, and J. A. J. Metz. 1998. Evolutionarily singular strategies

and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology* 12:35–57.

1983 Griffiths, H. M., J. Louzada, R. D. Bardgett, and J. Barlow. 2016. Assessing the importance of intraspecific variability in dung beetle functional traits. *PLoS ONE* 11:e0145598.

Gross, N., Y. L. Bagousse-Pinguet, P. Liancourt, M. Berdugo, N. J. Gotelli, and F. T. Maestre. 1986 2017. Functional trait diversity maximizes ecosystem multifunctionality. *Nature Ecology & Evolution* 1:0132.

Hart, S. P., S. J. Schreiber, and J. M. Levine. 2016. How variation between individuals affects 1989 species coexistence. *Ecology Letters* 19:825–838.

Kimura, M. 1965. A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proceedings of the National Academy of Sciences of the United States of America* 54:731–736. 1992

Kirkpatrick, M., and N. H. Barton. 1997. Evolution of a species' range. *The American Naturalist* 150:1–23.

1995 Klauschies, T., R. M. Coutinho, and U. Gaedke. 2018. A beta distribution-based moment closure enhances the reliability of trait-based aggregate models for natural populations and communities. *Ecological Modelling* 381:46–77.

1998 Klausmeier, C. A., C. T. Kremer, and T. Koffel. 2020. Trait-based ecological and eco-evolutionary theory. Pages 161–194 *in* K. S. McCann and G. Gellner, eds. *Theoretical Ecology: concepts and applications*. Oxford University Press, Oxford.

2001 Kraft, N. J. B., G. M. Crutsinger, E. J. Forrestel, and N. C. Emery. 2014. Functional trait differences and the outcome of community assembly: an experimental test with vernal pool annual plants. *Oikos* 123:1391–1399.

- 2004 Kremer, C. T., and C. A. Klausmeier. 2013. Coexistence in a variable environment: Eco-evolutionary perspectives. *Journal of Theoretical Biology* 339:14–25.
- Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: Body  
2007 size allometry. *Evolution* 33:402–416.
- Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. *Evolution* 37:1210–1226.
- 2010 Le Gland, G., S. M. Vallina, S. L. Smith, and P. Cermeño. 2020. Spead 1.0 – a model for simulating plankton evolution with adaptive dynamics in a two-trait continuous fitness landscape applied to the sargasso sea. *Geoscientific Model Development Discussions* 2020:1–54.
- 2013 Levins, R. 1962. Theory of fitness in a heterogeneous environment. i. the fitness set and adaptive function. *The American Naturalist* 96:361–373.
- Lin, W.-H., E. Kussell, L.-S. Young, and C. Jacobs-Wagner. 2020. Origin of exponential growth  
2016 in nonlinear reaction networks. *Proceedings of the National Academy of Sciences* 117:27795–27804.
- Lion, S., M. Boots, and A. Sasaki. 2022. Multimorph eco-evolutionary dynamics in structured  
2019 populations. *The American Naturalist* 200:345–372.
- Lotka, A. J. 1925. *Elements of physical biology*. Williams & Wilkins Company, Baltimore, MD, USA.
- 2022 MacArthur, R., and R. Levins. 1964. Competition, habitat selection, and character displacement in a patchy environment. *Proceedings of the National Academy of Sciences, U.S.A.* 51:1207–1210.
- MacArthur, R. H. 1972. *Geographical ecology*. Harper and Row, New York.
- 2025 Merico, A., G. Brandt, S. L. Smith, and M. Oliver. 2014. Sustaining diversity in trait-based models of phytoplankton communities. *Frontiers in Ecology and Evolution* 2:59.

- 2028 Metz, J. A. J., R. M. Nisbet, and S. A. H. Geritz. 1992. How should we define 'fitness' for general ecological scenarios? *Trends in Ecology & Evolution* 7:198–202.
- Mirrahimi, S., and S. Gandon. 2020. Evolution of specialization in heterogeneous environments: Equilibrium between selection, mutation and migration. *Genetics* 214:479–491.
- 2031 Mouillot, D., S. Villéger, M. Scherer-Lorenzen, and N. W. Mason. 2011. Functional structure of biological communities predicts ecosystem multifunctionality. *PloS ONE* 6:e17476.
- 2034 Norberg, J., D. P. Swaney, J. Dushoff, J. Lin, R. Casagrandi, and S. A. Levin. 2001. Phenotypic diversity and ecosystem functioning in changing environments: A theoretical framework. *Proceedings of the National Academy of Sciences* 98:11376–11381.
- 2037 Norberg, J., M. C. Urban, M. Vellend, C. A. Klausmeier, and N. Loeuille. 2012. Eco-evolutionary responses of biodiversity to climate change. *Nature Climate Change* 2:747–751.
- Nordbotten, J. M., F. Bokma, J. S. Hermansen, and N. C. Stenseth. 2020. The dynamics of trait variance in multi-species communities. *Royal Society Open Science* 7:200321.
- 2040 Paine, C. E. T., C. Baraloto, J. Chave, and B. Hérault. 2011. Functional traits of individual trees reveal ecological constraints on community assembly in tropical rain forests. *Oikos* 120:720–727.
- 2043 Peeters, F., and D. Straile. 2018. Trait selection and co-existence of phytoplankton in partially mixed systems: Trait based modelling and potential of an aggregated approach. *PLoS ONE* 13:e0194076.
- 2046 Raffard, A., F. Santoul, J. Cucherousset, and S. Blanchet. 2019. The community and ecosystem consequences of intraspecific diversity: a meta-analysis. *Biological Reviews* 94:648–661.
- 2049 Ranjan, R., and C. A. Klausmeier. 2022. How the resource supply distribution structures competitive communities. *Journal of Theoretical Biology* 538:111054.

- Rueffler, C., T. J. Van Dooren, O. Leimar, and P. A. Abrams. 2006. Disruptive selection and then what? *Trends in Ecology & Evolution* 21:238–245.
- 2052 Sasaki, A., and U. Dieckmann. 2011. Oligomorphic dynamics for analyzing the quantitative genetics of adaptive speciation. *Journal of Mathematical Biology* 63:601–635.
- Savage, V. M., C. T. Webb, and J. Norberg. 2007. A general multi-trait-based framework for  
2055 studying the effects of biodiversity on ecosystem functioning. *Journal of Theoretical Biology* 247:213 – 229.
- Schreiber, S. J., R. Bürger, and D. I. Bolnick. 2011. The community effects of phenotypic and  
2058 genetic variation within a predator population. *Ecology* 92:1582–1593.
- Siefert, A., C. Violle, L. Chalmandrier, C. H. Albert, A. Taudiere, A. Fajardo, L. W. Aarssen, C. Baraloto, M. B. Carlucci, M. V. Cianciaruso, V. de L. Dantas, F. de Bello, L. D. S. Duarte,  
2061 C. R. Fonseca, G. T. Freschet, S. Gaucherand, N. Gross, K. Hikosaka, B. Jackson, V. Jung, C. Kamiyama, M. Katabuchi, S. W. Kembel, E. Kichenin, N. J. B. Kraft, A. Lagerström, Y. L. Bagousse-Pinguet, Y. Li, N. Mason, J. Messier, T. Nakashizuka, J. M. Overton, D. A. Peltzer,  
2064 I. M. Pérez-Ramos, V. D. Pillar, H. C. Prentice, S. Richardson, T. Sasaki, B. S. Schamp, C. Schöb, B. Shipley, M. Sundqvist, M. T. Sykes, M. Vandewalle, and D. A. Wardle. 2015. A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. *Ecology Letters* 18:1406–1419.  
2067
- Slatkin, M. 1980. Ecological character displacement. *Ecology* 61:163–177.
- Tirok, K., B. Bauer, K. Wirtz, and U. Gaedke. 2011. Predator-prey dynamics driven by feedback  
2070 between functionally diverse trophic levels. *PLOS ONE* 6:e27357.
- Troost, T. A., B. W. Kooi, and S. A. L. M. Kooijman. 2005. Ecological specialization of mixotrophic plankton in a mixed water column. *The American Naturalist* 166:E45–E61.

- 2073 Turcotte, M. M., and J. M. Levine. 2016. Phenotypic plasticity and species coexistence. *Trends in Ecology & Evolution* 31:803–813.
- Turelli, M., and N. H. Barton. 1994. Genetic and statistical analyses of strong selection on poly-  
2076 genic traits: what, me normal? *Genetics* 138:913–941.
- Umaña, M. N., and N. G. Swenson. 2019. Does trait variation within broadly distributed species mirror patterns across species? a case study in puerto rico. *Ecology* 100:e02745.
- 2079 Uriarte, M., and D. Menge. 2018. Variation between individuals fosters regional species coexistence. *Ecology Letters* 21:1496–1504.
- Violle, C., B. J. Enquist, B. J. McGill, L. Jiang, C. H. Albert, C. Hulshof, V. Jung, and J. Messier.  
2082 2012. The return of the variance: intraspecific variability in community ecology. *Trends in Ecology & Evolution* 27:244–252.
- Volterra, V. 1928. Variations and fluctuations of the number of individuals in animal species  
2085 living together. *ICES Journal of Marine Science* 3:3–51.
- Westerband, A. C., J. L. Funk, and K. E. Barton. 2021. Intraspecific trait variation in plants: a renewed focus on its role in ecological processes. *Annals of Botany* 127:397–410.
- 2088 Wickman, J., S. Diehl, B. Blasius, C. A. Klausmeier, A. Ryabov, and Å. Brännström. 2017. Determining selection across heterogeneous landscapes: a perturbation-based method and its application to modeling evolution in space. *The American Naturalist* 189:381–395.
- 2091 Wickman, J., S. Diehl, and Å. Brännström. 2019. Evolution of resource specialisation in competitive metacommunities. *Ecology Letters* 22:1746–1756.
- Wirtz, K.-W., and B. Eckhardt. 1996. Effective variables in ecosystem models with an application  
2094 to phytoplankton succession. *Ecological Modelling* 92:33–53.

Xavier Jordani, M., N. Mouquet, L. Casatti, M. Menin, D. de Cerqueira Rossa-Feres, and C. H. Albert. 2019. Intraspecific and interspecific trait variability in tadpole meta-communities from the brazilian atlantic rainforest. *Ecology and Evolution* 9:4025–4037.

## References cited only in the online enhancements

Bezanson, J., A. Edelman, S. Karpinski, and V. B. Shah. 2017. Julia: A fresh approach to numerical computing. *SIAM review* 59:65–98.

Caswell, H. 2001. *Matrix Population Models*. 2nd ed. Sinauer Associates, Sunderland, MA.

Chase, J. M., and M. A. Leibold. 2003. *Ecological niches*. University of Chicago Press, Chicago, IL.

De Roos, A., T. Schellekens, T. Kooten, K. Wolfshaar, D. Claessen, and L. Persson. 2007. Food-dependent growth leads to overcompensation in stage-specific biomass when mortality increases: The influence of maturation versus reproduction regulation. *The American Naturalist* 170:E59–E76.

Geritz, S. A. H., E. van der Meijden, and J. A. J. Metz. 1999. Evolutionary dynamics of seed size and seedling competitive ability. *Theoretical Population Biology* 55:324–343.

Koffel, T., T. Daufresne, and C. A. Klausmeier. 2021. From competition to facilitation and mutualism: a general theory of the niche. *Ecological Monographs* 91:e01458.

Owen, D. B. 1980. A table of normal integrals. *Communications in Statistics - Simulation and Computation* 9:389–419.

Rackauckas, C., and Q. Nie. 2017. Differentialequations.jl—a performant and feature-rich ecosystem for solving differential equations in julia. *Journal of Open Research Software* 5.

Tilman, D. 1982. *Resource Competition and Community Structure*. Monographs in population biology. Princeton University Press, Princeton, NJ.

<sup>2118</sup> Triantafyllopoulos, K. 2002. Moments and cumulants of the multivariate real and complex gaussian distributions. Department of Mathematics, University of Bristol .