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

Jonas Wickman^{1,*}, Thomas Koffel¹, and Christopher A. Klausmeier¹

1. W. K. Kellogg Biological Station and Department of Plant Biology, Michigan State University, East Lansing, MI 48824, USA;

* Corresponding author; e-mail: jonas.wickman@gmail.com.

Keywords: Moment methods, adaptive dynamics, evolutionarily stable communities.

Abstract

To understand how functional traits shape ecological communities it is necessary to understand both how traits across the community affect its functioning and how eco-evolutionary dynamics within the community change the traits over time. Of particular interest are so-called evolutionarily stable communities (ESCs), since these are the end points of eco-evolutionary dynamics and can persist over long time scales. One theoretical framework that has successfully been used for assembling ESCs is adaptive dynamics. However, this framework cannot account for intraspecific variation—neither locally nor across structured populations. On the other hand, in moment-based approaches, intraspecific variation is accommodated, but community assembly has been neglected. This is unfortunate as some questions regarding for example local adaptation vis-a-vis diversification into multiple species requires both facets. In this paper we develop a general theoretical framework that bridges the gap between these two approaches. We showcase how ESCs can be assembled using the framework, and illustrate various aspects of the framework using two simple models of resource competition. We believe this unifying framework could be of great use to address questions regarding the role of functional traits in communities where population structure, intraspecific variation, and eco-evolutionary dynamics are all important.

Introduction

In recent years, the need to understand intraspecific variation of functional traits has been increasingly recognized as important for understanding the functioning of ecological communities (Albert et al., 2010; Violle et al., 2012). For example, intraspecific variation has been shown to be important in relation to niche differentiation and environmental filtering (Paine et al., 2011) and pairwise species interactions (Kraft et al., 2014). Such intraspecific variation in functional traits has also been shown to often be substantial compared to variation between species (Siefert et al., 2015).

If species have heritable variation in their traits, natural selection will act on that variation. In order to understand these effects, theorists and modelers have used models in quantitative genetics (Lande, 1979; Lande and Arnold, 1983) and moment-equation-based frameworks (Wirtz and Eckhardt, 1996; Norberg et al., 2001; Savage et al., 2007) where this variation and selection is taken into account. These types of models have been used to evaluate how intraspecific variation influences species coexistence (Barabás and D’Andrea, 2016) and how spatial gradients interact with standing variation (Norberg et al., 2012). However, these frameworks typically do not account for mutations, diversification into multiple species is hard to accommodate, and robust methods for determining whether communities are evolutionarily stable with regards to the addition of further species do not exist. On the other hand, another prominent framework for modeling eco-evolutionary dynamics known as adaptive dynamics (Metz et al., 1992; Dieckmann and Law, 1996; Geritz et al., 1998; Dercole and Rinaldi, 2008) readily accommodates questions of diversification and evolutionary stability. However, it relies on the assumption that species have no intraspecific variation in their traits, so it cannot be used to understand the eco-evolutionary consequences of intraspecific trait variation.

Over the last decade, starting with Sasaki and Dieckmann (2011), a new framework has emerged to try to bridge this gap and study problems that require including facets of both moment-based and adaptive-dynamics-based eco-evolutionary frameworks. Although several aspects of the theory such as evolutionary branching (Sasaki and Dieckmann, 2011), and the effects of multivariate traits (Débarre et al., 2014) have been studied, much still remains to be worked out. In particular, general theoretical and modeling tools for assembling evolutionarily stable communities while still incorporating intraspecific variation are not yet available. Such tools would be valuable, as evolutionarily stable communities can serve as good model communities as they can persist over long time scales (Edwards et al., 2018), and incorporating intraspecific diversity would allow us to explore how different ecological conditions affect the total diversity of traits in assembled communities. Furthermore, these eco-evolutionary models taking both multiple species and intraspecific variation into account have so far almost exclusively neglected population structure (but see Débarre et al., 2013, for an example with spatial structure). The combination of multiple species and structured populations is interesting as modeling the joint trait distribution across species and spatial patches could help address questions of how diversity is structured under the influence of evolution in complex communities where both local adaptation within species as well as separate species occupying different parts of trait space are possible.

In this paper we synthesize many recent advances in describing eco-evolutionary dynamics through moment equations taking the total density, mean trait vectors, and variance-covariance matrices into account, while also accounting for phenotypic mutations from first principle. Furthermore, we add some key missing components for using this theory for eco-evolutionary modeling,

in particular techniques for assembling eco-evolutionarily stable communities. We derive generic moment equations for class-structured populations, which can then easily be adapted to specific examples, and we showcase two such example to illustrate various facets of the modeling framework.

Methods and models

In this section we will specify our generic model for class-structured communities, show how moment approximations can be derived, and devise a community-assembly procedure to assemble eco-evolutionarily stable communities. We will illustrate the generic models with a simple two-patch model of resource competition.

Generic class-structured model and moment equations for fixed number of species

We consider a community of clonally reproducing individuals structured in K classes (e.g., spatial patches or age classes), where the rates of the ecological processes for individuals in the community are governed by an n -dimensional trait vector $\xi = (\xi_1, \xi_2, \dots, \xi_n) \in \mathbb{R}^n$. Each component of the trait vector encodes some quantitative trait of the organism, e.g., body mass or degree of resource specialization, with n different traits being tracked in total. The trait-density distribution in each class k is given by $v_k(\xi)$, $k = 1, 2, \dots, K$, which describes how abundant individuals with a given trait vector ξ in class k are. Figure 1 shows examples of trait-density distributions in one- and two-dimensional trait space. We let $\mathbf{v} := (v_1, v_2, \dots, v_K)$ be the vector of density distributions over all the classes.

The trait-density $v_k(\xi)$ in each class changes over time in response to various ecological processes. For each focal class k and source class l we specify N_{kl} ecological processes with per-capita rates $f_{klm}(\xi, \mathbf{v})$ indexed by $m = 1, 2, \dots, N_{kl}$. For within-class processes ($l = k$) these processes may include for example local birth and mortality rates on a spatial patch. For between-class processes ($l \neq k$) they may describe for example how v_k changes in response to immigration from spatial patch l , or from births from adult class l if k is a juvenile class in a stage-structured model. We also consider the possibility that each such process can have replicative infidelity associated with it, so that a phenotype with trait vector η undergoing the process with rate f_{klm} will yield a normal distribution of individuals. Write $\mathcal{N}(\xi, \eta, M_{klm})$ for the probability-density function with argument ξ of a normal distribution with mean vector η and variance-covariance matrix M_{klm} . For simplicity, we shall refer to M_{klm} as the mutation variance-covariance matrix. The instantaneous rate of change to individuals with trait ξ due to the ecological process with rate f_{klm} and mutation variance-covariance matrix M_{klm} is then given by the convolution $\int f_{klm}(\eta) v_l(\eta) \mathcal{N}(\xi, \eta, M_{klm}) d\eta$. For processes with no mutation—which will typically be the case, except for births—we set $M_{klm} = 0$ and interpret it in the delta-dirac sense so that

$$\int f_{klm}(\eta, \mathbf{v}) \mathcal{N}(\xi, \eta, 0) v_l(\eta) d\eta = f_{klm}(\xi, \mathbf{v}) v_l(\xi), \quad (1)$$

Summing over all classes and processes yields the complete equations

$$\frac{dv_k(\xi, t)}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \int f_{klm}(\eta, \mathbf{v}) \mathcal{N}(\xi, \eta, M_{klm}) v_l(\eta) d\eta. \quad (2)$$

We shall refer to these types of equations as *trait-space equations*. As in adaptive dynamics and trait-diffusion approaches (Merico et al., 2014; Le gland et al., 2020; Nordbotten et al., 2020), we are not here concerned with any specific genetic makeup that would result in these dynamics for phenotypic traits, but simply assume that new heritable variation is generated in some ecological process such as births and mutations. For a single class, our setup is however very similar to the one-locus continuum-of-alleles class of models (Kimura, 1965; Bürger, 1986). Our equations also have various things in common with other eco-evolutionary approaches; we discuss these similarities further in the discussion section.

In Fig. 1 we illustrate two example models in the framework of Eq. 2: a two-patch model with a single scalar trait and a stage-structured model with one juvenile and one adult stage with two traits. We will work out the two-patch example in Box 1, and the stage-structured model is treated in section “Additional example: A stage-structured resource-competition model”. In section “Trait-space equations” of Box 1 we show an example of trait-space equations for the two-patch model, and how it relates to the generic trait-space equations (Eq. 2).

In general, Eq. 2 will not be analytically tractable; even the equilibrium of Eq. 2 will in general be hard to ascertain, even for very simple examples with only a single class (e.g., Kimura, 1965; Bürger, 1986). Numerically too, discretizing trait space and solving Eq. 2 requires significant computational power; for multiple traits especially, the feasibility of numerical explorations of the equations is highly limited. Thus, rather than solving the trait-space equations directly, we instead derive moment equations that tracks the total density, mean trait, and variance-covariance of several peaks approximating the trait-density distribution in each class. Our derivations closely follow similar methods in earlier work (Wirtz and Eckhardt, 1996; Norberg et al., 2001; Bruggeman, 2009; Merico et al., 2014; Sasaki and Dieckmann, 2011; Débarre et al., 2014; Nordbotten et al., 2020), but differs in considering class-structured populations, adding new variation through a mutational kernel, and multiple traits simultaneously.

We assume that the densities $v_k(\xi)$ can be usefully split into S peaks (which we will refer to as species) so that $v_k(\xi) = \sum_{i=1}^S v_{ik}(\xi)$, and that each of these peaks can be approximated by a normal distribution so that

$$v_k(\xi, t) = \sum_{i=1}^S v_{ik}(\xi, t) \approx \tilde{v}_k(\xi, t) := \sum_{i=1}^S u_{ik}(t) \mathcal{N}(\xi, \chi_{ik}(t), V_{ik}(t)). \quad (3)$$

Here $u_{ik} = \int v_{ik}(\xi) d\xi$ is the total density of species i in class k , $\chi_{ik} = u_{ik}^{-1} \int \xi v_{ik}(\xi) d\xi$ is the mean-trait vector of species i in class k , and $V_{ik} = u_{ik}^{-1} \int (\xi - \chi_{ik})(\xi - \chi_{ik})^T v_{ik}(\xi) d\xi$ is the variance-covariance matrix of species i in class k . We also let $\tilde{\mathbf{v}} := (\tilde{v}_1, \tilde{v}_2, \dots, \tilde{v}_K)$ be the vector of all approximate trait distributions over all classes. Furthermore, we make the approximating assumption that each species is reproductively isolated, so that all transitions can be computed species-by-species without any trait-density ‘leaking’ between species. For now, we take the number of species S to be arbitrary; we will address the question of how to determine S in section “Eco-evolutionary community assembly”.

Figure 1: Two illustrations of the generic class-structured trait-space equations. Rate functions placed at bases of arrows indicate losses and rate functions placed at heads of arrows indicate gains. All per capita rate functions f_{klm} can depend both on the trait of the individual under consideration, ξ , as well as the trait-density distributions in all classes $\mathbf{v}$. **(A)** A two-patch model with local birth and mortality on each patch and dispersal between the patches with one trait ξ . New individuals are born at (per capita) rates f_{AA1} and f_{BB1} on each patch respectively, and the traits of the newborn individuals are filtered through a convolution with normal distributions $\mathcal{N}(0, M_{AA1})$ and $\mathcal{N}(0, M_{BB1})$ with zero means and variance-covariance M_{AA1} and M_{BB1} due to mutations on each patch respectively. Local mortality on each patch occurs at rates f_{AA2} and f_{BB2} respectively. Individuals disperse out of patch A at rate f_{AA3} and arrive at patch B with rate f_{BA1} . Conversely, individuals disperse out of patch B at rate f_{BB3} and arrive at patch A at rate f_{AB1} . **(B)** A two-stage model with juveniles and adults with two traits $\xi = (\xi_1, \xi_2)$. Juveniles grow at rate f_{JJ1} and suffer mortality at rate f_{JJ2} . Juveniles mature out of the juvenile stage at rate f_{JJ3} and into adults at rate f_{AJ1} . Adults suffer a mortality at rate f_{AA1} and produce new juveniles at rate f_{JA1} . The new juveniles thus produced have their traits filtered through a convolution with a normal distribution $\mathcal{N}(0, M_{JA1})$ with mean zero and variance-covariance matrix M_{JA1} due to mutations.

We define the per capita population rate of process f_{klm} for a population with mean trait ξ and variance-covariance W as

$$\hat{f}_{klm}(\xi, W, \tilde{\mathbf{v}}) := \int f_{klm}(\eta, \tilde{\mathbf{v}}) \mathcal{N}(\eta, \xi, W) d\eta, \quad (4)$$

which averages the trait-dependent per capita rates over the normal phenotype distribution. Then, under these assumptions the moment equations approximating Eqs. 2 for species $i = 1, 2, \dots, S$ are then given by (see Appendix A for the derivation)

$$\frac{du_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \underbrace{\hat{f}_{iklm}}_{(i)} u_{il}, \quad (5a)$$

$$\frac{d\chi_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \underbrace{\frac{u_{il}}{u_{ik}}}_{(ii)} \left(\underbrace{V_{il} \nabla_{\xi} \hat{f}_{iklm}}_{(iii)} + \underbrace{\hat{f}_{iklm} (\chi_{il} - \chi_{ik})}_{(iv)} \right), \quad (5b)$$

$$\begin{aligned} \frac{dV_{ik}}{dt} = & \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \underbrace{\frac{u_{il}}{u_{ik}}}_{(ii)} \left(\underbrace{V_{il} \nabla_{\xi}^2 \hat{f}_{iklm} V_{il}}_{(v)} + \underbrace{\hat{f}_{iklm} (V_{il} - V_{ik})}_{(vi)} + \underbrace{\hat{f}_{iklm} M_{klm}}_{(vii)} \right. \\ & + \underbrace{\hat{f}_{iklm} (\chi_{il} - \chi_{ik}) (\chi_{il} - \chi_{ik})^T}_{(viii a)} \\ & \left. + \underbrace{V_{il} \nabla_{\xi} \hat{f}_{iklm} (\chi_{il} - \chi_{ik})^T + (\chi_{il} - \chi_{ik}) [V_{il} \nabla_{\xi} \hat{f}_{iklm}]^T}_{(viii b)} \right). \end{aligned} \quad (5c)$$

Here, the superscript T denotes vector or matrix transpose. Furthermore, the arguments of all rate functions have been suppressed for notational clarity, and

$$\hat{f}_{iklm} := \hat{f}_{klm}(\chi_{il}, V_{il}, \tilde{\mathbf{v}}) \quad (6a)$$

$$\nabla_{\xi} \hat{f}_{iklm} := \nabla_{\xi} \hat{f}_{klm}(\chi_{il}, V_{il}, \tilde{\mathbf{v}}) \text{ Gradient of } \hat{f}_{klm} \text{ with respect to trait, evaluated at } \chi_{il} \text{ and } V_{il} \quad (6b)$$

$$\nabla_{\xi}^2 \hat{f}_{iklm} := \nabla_{\xi}^2 \hat{f}_{klm}(\chi_{il}, V_{il}, \tilde{\mathbf{v}}) \text{ Hessian of } \hat{f}_{klm} \text{ with respect to trait, evaluated at } \chi_{il} \text{ and } V_{il} \quad (6c)$$

The integral in Eq. 4 is analytically tractable for some functions—for example, exponentials, Gaussians, and polynomials in the scalar-trait case (see e.g., Owen, 1980)—but is in general not solvable directly. In such cases, assuming that the variance-covariances of the species are not too large, we can Taylor-expand the process rate functions f_{klm} around the mean traits to some order. In particular, to third order this yields

$$\hat{f}_{iklm} \approx f_{klm}(\chi_{il}, \tilde{\mathbf{v}}) + \frac{1}{2} \nabla_{\xi}^2 f_{klm}(\chi_{il}, \tilde{\mathbf{v}}) : V_{il}, \quad (7a)$$

$$\nabla_{\xi} \hat{f}_{iklm} \approx \nabla_{\xi} f_{klm}(\chi_{il}, \tilde{\mathbf{v}}) + \nabla_{\xi} \left(\frac{1}{2} \nabla_{\xi}^2 f_{klm} : V_{il} \right) (\chi_{il}, \tilde{\mathbf{v}}), \quad (7b)$$

$$\nabla_{\xi}^2 \hat{f}_{iklm} \approx \nabla_{\xi}^2 f_{klm}(\chi_{il}, \tilde{\mathbf{v}}). \quad (7c)$$

where “:” is the double dot product, which for two square $n \times n$ matrices A and B is given by $A : B = \sum_{i=1}^n \sum_{j=1}^n A_{ij} B_{ij}$. We shall use this third-order approximation for the remainder of this text. See Appendix A for expressions of arbitrary order for the Taylor expansions.

As noted in 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 these equations and their interpretations. Figure 2 illustrates how these terms come about through the dynamics of the trait-space equations and their effects.

Equation 5a describes the rate of change of the total density u_{ik} of species i in class k . Term (i), *per-capita growth*, describes the per-capita rate of process m between class k and class l evaluated at the mean-trait vector of species i in class l but is corrected by the a second-order term emanating from the Taylor expansion (Eq. 7a). This second-order correction is often referred to as the ‘genetic load’ in the quantitative genetics literature (Lande, 1979; Kirkpatrick and Barton, 1997), since if the species is distributed around a fitness optimum, the growth rate suffers a penalty due to the maladapted individuals around the optimum. The equations capture how the total densities of species change both due to within-class processes ($l = k$) such as local birth and death, and between-class processes ($l \neq k$) such as dispersal between patches (c.f., Fig. 1A).

Equation 5b describes the rate of change of the mean-trait vector χ_{ik} of species i in class k . Term (ii), *relative density weight*, weighs contributions by the relative total densities of classes k and l , so that if the density in the focal class k is much greater than that in the source class l , the ecological processes going from l to k will only have a marginal impact on the mean-trait vector in class k . Conversely, if the total density in class l is much greater than in class k , the ecological process going from l to k will have a large impact on the mean-trait vector in class k . The change of the mean-trait vector is then governed by two terms, (iii) and (iv). The first term, (iii), *directional selection*, describes the effect of directional selection in process m in class l , pushing χ_{ik} in the direction of maximum increase of $\hat{f}_{iklm}$, at a rate mediated by the available variance of species i in class l , V_{il} . 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 (Fig. 2A–C). The second term, (iv), *mean-trait flow*, describes how mean traits are homogenized by between-class transitions, where the rate of homogenization is governed by the per capita rate function $\hat{f}_{iklm}$, so that the mean trait of the focal class, χ_{ik} will change in the direction of the mean trait in the source class, χ_{il} (Fig. 2D–F). For within-class processes ($l = k$) we note that the relative density weight (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, the directional selection (iii) would be equal to zero, but mean-trait flow (iv) could still contribute towards changing the mean-trait vector in class k if the mean-trait vectors in l and k are different.

Equation 5c describes the rate of change of the variance-covariance matrix V_{ik} of species i in class k . Like for the mean-trait vectors, the changes are weighted by the relative densities between class k and l , term (ii). The dynamics of the variance-covariance then depends on five terms. The first term, (v), describes *stabilizing/disruptive selection* resulting from the process with rate $\hat{f}_{iklm}$. Roughly speaking, if individuals close to the mean-trait vector χ_{il} of species i in class l contribute more than individuals away from this optimum to process $\hat{f}_{iklm}$ then the curvature as measured by $\nabla_{\xi}^2 \hat{f}_{iklm}$ will be negative, which will contribute to a decrease in variances (Fig. 2G–I). Conversely, if individuals close to the mean contribute less, the curvature will be positive and this will contribute towards increases in variances. The second term, (vi), *variance-covariance flow*, homogenizes variance-covariances between classes, so that having trait-density flow from class l to k

contributes towards driving the variance-covariance V_{ik} of species i in class k closer to the variance-covariance V_{il} in class l (Fig. 2J–L). The third term, (vii), *mutation*, is the contribution to variance-covariance from mutations in process f_{klm} , which will contribute towards increasing variances and aligning the covariances of the species with the covariances of the mutational variance-covariance matrix M_{klm} . Note that typically for most processes under consideration M_{klm} would be zero, as in the two examples depicted in Fig. 1, where only birth processes are assumed to have mutations associated with them. How much mutations contribute towards variance increase also depends on the rate $\hat{f}_{iklm}$. 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 variance-covariance than in a system with low birth and death rates even if net per capita growth were the same in both systems. Terms (viii), *effects of class-local adaptation on variance-covariance*, describe the effects of variability in mean-trait vectors between classes on the variance-covariance 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 population. The fourth term, (viiiia), *between-to-within class variation*, describes how variances are increased and covariances altered by the differences in mean-trait vectors between classes, converting between-class variation into within-class variation (Fig. 2M–O). The final term, (viiiib), *class-local adaptation and directional selection interaction*, comprises two terms which are the transposes of one another so that their sum is symmetric. They describe the interaction between directional selection and between-class differences in mean-trait vectors. This interaction is not trivial when the dimensionality is higher than one, but roughly speaking, variances will decrease when the mean trait difference and the selection gradient point in opposite directions, and increase if they are aligned (Fig. 2M–O). For within-class processes ($l = k$), the relative density weight (ii) will be equal to one, and only stabilizing/disruptive selection (v) and mutations (vii) will be nonzero, making these the only contributing factors. If the process f_{klm} under consideration is trait independent, stabilizing/disruptive selection (v) and class-local adaptation and directional selection interaction (viiiib) will be zero, but mutations (vii), variance-covariance flow (vi), and between-to-within class variation (viiiia) can still contribute to changes in the variance-covariances.

In section B1.2 of Box 1, we showcase a concrete example of moment equations for the two-patch model.

(iii) Directional selection

(iv) Mean-trait flow

(v) Stabilizing selection

(vi) Variance-covariance flow

(viii) Effects of class-local adaptation on variance-covariance

Figure 2: Illustration of terms (iii), (iv), (v), (vi), and (viii) of Eqs. 5. For visual clarity and purposes of explanation, these panels are mock-ups and are not true calculations based on the trait-space equations. **(A–C)** Causes and effects of term (iii) of Eqs. 5. Panel A shows trait-density distributions $v_k(\xi, t)$ and $v_l(\xi, t)$ of focal class k and source class l at time t along with a rate function f_{kl} describing how v_k will change under the influence of v_l . Panel B shows the rate of change over time for v_k as a result of f_{kl} and v_l . Note that the rate of change is offset from v_k so that higher trait values will grow faster than lower trait values. Panel C shows how, over time, this offset will drive the trait-density distribution of the focal class v_k to shift its mean in the direction of the gradient of f_{kl} . Thus, at time $t + \Delta t$, the mean trait χ_k of v_k will have shifted to become greater. **(D–F)** Causes and effects of term (iv) of Eqs. 5. Panel D shows trait-density distributions $v_k(\xi, t)$ and $v_l(\xi, t)$ of focal class k and source class l at time t along with a rate function f_{kl} describing how v_k will change under the influence of v_l . Panel E shows the rate of change over time for v_k as a result of f_{kl} and v_l . Note that the rate of change is offset from v_k so that higher trait values will grow faster than lower trait values. Panel F shows how, over time, this offset will drive the trait-density distribution of the focal class v_k to shift its mean in the direction towards the mean trait χ_l of v_l . Thus, at time $t + \Delta t$, the mean trait χ_k of v_k will have shifted to become closer to χ_l . **(G–I)** Causes and effects of term (v) of Eqs. 5. Panel G shows trait-density distributions $v_k(\xi, t)$ and $v_l(\xi, t)$ of focal class k and source class l at time t along with a rate function f_{kl} describing how v_k will change under the influence of v_l . Panel H shows the rate of change over time for v_k as a result of f_{kl} and v_l . Note that the rate of change is proportionally higher around the mean trait χ_k of v_k so that trait values closer to χ_k grow faster than trait values further away. Panel I shows how, over time, this higher growth around χ_k will drive the trait-density distribution of the focal class v_k to shift its variance V_k to become lower. Thus, at time $t + \Delta t$, the standard deviation $\sqrt{V_k}$ of v_k will have shifted to become smaller. **(J–L)** Causes and effects of term (vi) of Eqs. 5. Panel J shows trait-density distributions $v_k(\xi, t)$ and $v_l(\xi, t)$ of focal class k and source class l at time t along with a rate function f_{kl} describing how v_k will change under the influence of v_l . Panel K shows the rate of change over time for v_k as a result of f_{kl} and v_l . Note that the rate of change is proportionally higher around the mean trait χ_k of v_k so that trait values closer to χ_k grow faster than trait values further away. Panel L shows how, over time, this higher growth around χ_k will drive the trait-density distribution of the focal class v_k to shift its variance V_k towards the variance V_l of v_l . Thus, at time $t + \Delta t$, the standard deviation $\sqrt{V_k}$ of v_k will have shifted to become closer to the standard deviation $\sqrt{V_l}$ of v_l . **(M–O)** Causes and effects of term (viii) of Eqs. 5. Panel M shows trait-density distributions $v_k(\xi, t)$ and $v_l(\xi, t)$ of focal class k and source class l at time t along with rate functions f_{kl1} , f_{kl2} , and f_{kl3} , each describing one way of how v_k will change under the influence of v_l . Note that for f_{kl1} , there is no directional selection, and only term (viiiia) will play a role. Panel N shows the rate of change over time for v_k as a result of each f_{kl} and v_l . The rate of change is offset from v_k so that trait values at low density will grow faster than trait values at high density. Note also that the rate of change is more offset for f_{kl2} and less offset for f_{kl3} compared to f_{kl1} . Panel O shows how, over time, these alterations to the offsets will drive the trait-density distribution of the focal class v_k to shift its variance V_k . This change is more pronounced for f_{kl2} and less for f_{kl3} compared to f_{kl1} due to the positive and negative gradients of f_{kl2} and f_{kl3} respectively. Thus, at time $t + \Delta t$, the standard deviation $\sqrt{V_k}$ of v_k will have shifted more for f_{kl2} and less for f_{kl3} compared to the shift caused by term (viiiia).

Box 1: A two-patch model

In this box we showcase a simple example of our generic class-structured equations in the form of a two-patch model. This example model is used throughout the methods and models section to illustrate various facets of the theory.

Trait-space equations

For the two-patch model, we consider a system where consumers compete for two resources R_1 and R_2 and where the two classes are two patches, A and B . We assume that resources are renewed chemostatically on each patch, but that the supply of R_1 is greater on patch A and that the supply of R_2 is greater on patch B . We assume that the consumers consume resources with a type-I (linear) functional response, where the affinities $a_1(\xi)$ and $a_2(\xi)$ for R_1 and R_2 respectively are in a trade-off parameterized by a scalar trait $\xi \in \mathbb{R}$ that is under selection. The trade-off is such that $\xi = 0$ yields a consumer with equal affinity for both resources, $\xi > 0$ implies a greater affinity for R_1 , and $\xi < 0$ implies a greater affinity for R_2 . We also assume that this trade-off is generalist-favoring (weak) so that $a_1(\xi) + a_2(\xi)$ is maximized for $\xi = 0$.

In a well-mixed system on a single patch this would result in a single generalist consumer being evolutionarily stable (Levins, 1962; MacArthur and Levins, 1964; Rueffler et al., 2006; Wickman et al., 2019). Let $v_A(\xi)$ be the trait-density distribution of consumers on path A and $v_B(\xi)$ be the trait-density distribution on patch B . We then have the equations

$$\frac{dv_A(\xi)}{dt} = \int b(\eta, R_{1A}, R_{2A})v_A(\eta)\mathcal{N}(\xi, \eta, M)d\eta - \mu v_A(\xi) - dv_A(\xi) + dv_B(\xi), \quad (8a)$$

$$\frac{dv_B(\xi)}{dt} = \int b(\eta, R_{1B}, R_{2B})v_B(\eta)\mathcal{N}(\xi, \eta, M)d\eta - \mu v_B(\xi) + dv_A(\xi) - dv_B(\xi), \quad (8b)$$

$$\frac{dR_{jI}}{dt} = r(K_{jI} - R_{jI}) - \int U_j(\xi, R_{1I}, R_{2I})v_I(\xi)d\xi, \quad j = 1, 2, \quad I = A, B, \quad (8c)$$

$$U_j(\xi, R_1, R_2) = a_j(\xi)R_j, \quad (8d)$$

$$b(\xi, R_1, R_2) = c_1 a_1(\xi)R_1 + c_2 a_2(\xi)R_2. \quad (8e)$$

Here, b is the per capita birth rate, which depends on both the trait ξ as well as the densities of resources R_1 and R_2 on a given patch. The consumers suffer a constant per-capita background mortality μ , and disperse between patches at a constant per capita rate d . Figure 1A provides an illustration of the system in the notation used for the generic equations (Eq. 2), where we can identify $f_{AA1} = b(\xi, R_{1A}, R_{2A})$, $f_{AA2} = b(\xi, R_{1B}, R_{2B})$, $f_{AA2} = f_{BB2} = -\mu$, $f_{AB1} = f_{BA1} = d$, and $f_{AA3} = f_{BB3} = -d$. The birth rate has an associated mutational variance M , so that new individuals born from a parent with trait η are normally distributed around this trait with variance M , so that $M_{AA1} = M_{BB1} = M$ in Fig. 1. The resource renewal rate for each resource on each patch is r and the maximal resource supply is K_{jI} for resource $j = 1, 2$ on patch $I = A, B$.

Moment equations

By comparing the rate functions of the two-patch model (Eqs. 8) with the generic trait-space equations (Eq. 2), we can plug in our specific rate functions for the two-patch example into the generic moment equations (Eqs. 5) to get the moment equations for the two-patch system, which yields

$$\frac{du_{iA}}{dt} = \hat{b}_{iA}u_{iA} - \mu u_{iA} - du_{iA} + du_{iB}, \quad (9a)$$

$$\frac{d\chi_{iA}}{dt} = V_{iA}\frac{\partial \hat{b}_{iA}}{\partial \xi} + \frac{u_{iB}}{u_{iA}}d(\chi_{iB} - \chi_{iA}), \quad (9b)$$

$$\frac{dV_{iA}}{dt} = V_{iA}^2\frac{\partial^2 \hat{b}_{iA}}{\partial \xi^2} + \hat{b}_{iA}M + \frac{u_{iB}}{u_{iA}}d(V_{iB} - V_{iA} + (\chi_{iB} - \chi_{iA})^2), \quad (9c)$$

$$\frac{du_{iB}}{dt} = \hat{b}_{iB}u_{iB} - \mu u_{iB} + du_{iA} - du_{iB}, \quad (9d)$$

$$\frac{d\chi_{iB}}{dt} = V_{iB}\frac{\partial \hat{b}_{iB}}{\partial \xi} + \frac{u_{iA}}{u_{iB}}d(\chi_{iA} - \chi_{iB}), \quad (9e)$$

$$\frac{dV_{iB}}{dt} = V_{iB}^2 \frac{\partial^2 \hat{b}_{iB}}{\partial \xi^2} + \hat{b}_{iB} M + \frac{u_{iA}}{u_{iB}} d \left(V_{iA} - V_{iB} + (\chi_{iA} - \chi_{iB})^2 \right), \quad (9f)$$

$$\frac{dR_{jI}}{dt} = r(K_{jI} - R_{jI}) - \sum_{i=1}^S \hat{U}_{jIi} u_{iI}, \quad j = 1, 2, I = A, B. \quad (9g)$$

These equations are made less cumbersome by the fact that the only trait-dependent rates are the birth rates, which are within-class processes. As for the generic equations, we use a hat on a rate function to indicate that it represents a population rate or a Taylor approximation thereof. Thus, to Taylor order 3 we have e.g. that $\hat{b}_{iA} = b(\chi_{iA}, R_{1A}, R_{2A}) + (1/2)\partial^2 b/\partial \xi^2(\chi_{iA}, R_{1A}, R_{2A})V_{iA}$. Equation 9a describes how the total density u_{iA} for each species changes on patch A due to births, deaths, and dispersal to and from patch B . Equation 9b describes how the mean trait χ_{iA} for each species changes on patch A due to directional selection on birth rates and mean-trait flow from patch B . Equation 9c describes how the variance V_{iA} for each species changes on patch A due to stabilizing/disruptive selection in the birth rates, increases in variances due to mutations associated with the birth rate, variance-covariance flow from path B , and finally between-to-within class variation between the patches. Equations 9d–f describe the same processes on patch B .

We have also approximated the resource-uptake integral in Eq. 8c by Taylor-expanding the uptake function U around the mean-trait value of each species to second order. The details of how generic environmental variables such as resources can be incorporated into generic moment equations are given in Appendix A.

Figure 3 shows an example of numerical solutions of Eqs. 9 for two different rates of dispersal, d . For high rates of dispersal, the two-patch model is close to a well-mixed model and the eco-evolutionary equilibrium consists of a single generalist due to the generalist-favoring trade-off. For low rates of dispersal, a two-species equilibrium of two moderate specialists is attained due to the differential supply rate of the two resources on the two patches.

Eco-evolutionary community assembly

In the two-patch model example (Box 1, Fig. 3) we saw that the model could yield either a one-species or two-species equilibrium, depending on the rate of dispersal between patches. A question that naturally arises then is: how can we tell how many species we need? For the relatively simple case of the two-patch model we could numerically solve the trait-space equations directly to determine this, but in more-complicated models the computational costs will in general be prohibitive. Thus, we would like some procedure for assembling an eco-evolutionary equilibrium that matches the equilibrium of the trait-space equations using only information from the moment equations. Here, we devise such a procedure by adapting two concepts from adaptive dynamics. The first criterion is local in trait space and determines when an evolutionary branching (Geritz et al., 1998) can take place. The second is a global invasion criterion to determine when an invader with traits that are not necessarily close to a resident can successfully invade the community. These conditions are then used to form so-called evolutionarily stable communities (Edwards et al., 2018).

We note that our goal here is with respect to the ultimate assembled community in eco-evolutionary equilibrium. Our assembly procedure will in general not be a good approximation

Figure 3: Illustration of numerical solutions of Eqs. 9 of the two-patch model (Box 1) for two parameter sets.

The two left-side columns show solutions with one species for a high rate of dispersal ($d = 10^{-1}$), and the two right-side columns show solutions with two species for a low rate of dispersal ($d = 10^{-3}$). (A–D) Total densities over time on both patches for the two parameter sets. The densities change over time in accordance with Eqs. 9a,d. Note that for panels C–D each patch has one abundant species and one less abundant species, which is reversed between the patches. (E–H) Mean traits and standard deviations over time on both patches for the two parameter sets. The mean traits are shown as solid lines, and the colored area around each mean trait depicts one standard deviation around the mean. For high rates of dispersal, a single generalist evolves (Panels E–F), and for low rates of dispersal, two partial specialists evolve (Panels G–H). $\chi > 0$ implies greater affinity for resource one and $\chi < 0$ implies greater affinity for resource two. (I–L) Trait-density distributions at the final equilibrium on both patches for the two parameter sets. The numerical equilibrium is shown both for the numerical solution of the trait-space equations (Eqs. 8) as well as the species from the moment equations (Eqs. 9) depicted in the other panels. For these parameters, the moment equations approximate the equilibrium of the trait-space equations very well. **Parameter values:** $\mu = 0.1$, $r = 1$, $K_{1A} = 1$, $K_{1B} = 0.4$, $K_{2A} = 0.4$, $K_{2B} = 1$, $c_1 = 1$, $c_2 = 1$, $\alpha = 0.5$, $M = 5 \cdot 10^{-6}$.

of the temporal dynamics of the trait-space equations leading up to the final community. We note also that our results here cannot be directly applied to questions of speciation, which would require the inclusion of genetic details not present in our framework.

Evolutionary branchings

In adaptive dynamics, an evolutionary branching occurs when a species has evolved a trait in accordance 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 be directly translated to our moment equations due to the fact that we include trait variance and local adaptation. Instead, after Eqs. 5 have reached an equilibrium for S species we in turn split each species into two identical copies and examine the linear stability of the resulting $S + 1$ species equilibrium. If such a split-species equilibrium is unstable, we say that the system has undergone an evolutionary branching and keep the split-species pair. The details of the procedure is available in Appendix B, where we also present a concrete example for the two-patch model (Fig. B.1).

Global invasion dynamics

Evolutionary branchings are not by themselves enough to ascertain global evolutionary stability across trait space as there may be invaders with substantially different traits from any resident species able to invade the community (Geritz et al., 1999; Klausmeier et al., 2020). Thus, once the moment equations for some fixed number of species (Eqs. 5) has reached equilibrium we would like to determine whether additional species that are not necessarily similar to any resident species can be added to the community. In adaptive dynamics, this is done by computing the invasion fitness of a invader with a fixed trait, assuming that the invader is so rare that it does not affect the environment it is invading. The invasion fitness is then defined to be the long-term exponential per capita growth rate of the invader, while the invader is still rare (Metz et al. 1992; see Caswell 2001 for structured populations).

Using this idea, we wish to construct a measure for our moment equations that tells us the ultimate fate of an introduced rare invader with some given trait. To do this, we also assume that the invader will remain so rare as to not affect the environment during the invasion process. We introduce an invader with a given mean-trait vector χ^{inv} that is the same in all classes, and with negligible variance-covariance, simulating the appearance of a single invading individual. We then let this invader mean-trait vector evolve until it has reached an equilibrium. We then let the variance-covariance of the invader evolve according to the moment equations until both the mean-trait vector and variance-covariance of the invader are in equilibrium. At this point, we can calculate the long-term exponential per capita growth rate of this invader. We will say that this is the invasion fitness of the invader. If the invasion fitness is positive, the invader will grow in the environment set by the resident community, and if the invasion fitness is negative or zero, it will not. We can thus associate an invasion fitness with any invading trait vector χ^{inv} . The mathematical details of the procedure are available in Appendix B (Kremer and Klausmeier, 2013, used a similar approach in a non-structured model with fixed variance).

We have chosen to carry out the invasion process in steps, where first only the mean-trait vector is allowed to evolve, to avoid situations where the invader starts close to a fitness minimum in trait

space, where the variance of the invader could potentially become too large before the invader has a chance to evolve a more favorable mean trait. The result is that each invading trait will trace a path in trait space until it reaches an equilibrium where its growth rate can be evaluated. However, calculating all such paths is not necessary for the community assembly process. If what we are interested in is to find any invader with positive growth rate, we can sidestep the first step where invaders evolve their mean-trait vectors by calculating the adaptive-dynamics fitness landscape. For rare invaders with no variance, evolving their mean traits then simply amounts to climbing this fitness landscape, and thus we need concern ourselves only with invading mutants with mean-trait vectors corresponding to the local peaks in the fitness landscape. Note that the adaptive-dynamics fitness landscape does not by itself yield sufficient information for our moment equations, as invaders may change their mean-trait vectors, and the population fitness of the invaders can change when variance-covariance is taken into account. An example of this phenomenon for the two-patch model is depicted in Fig. B.2 in Appendix B, where the adaptive dynamics fitness landscape indicates an evolutionary branching into two species, whereas the trait-space and moment equations admit only one.

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 in K classes with total densities u_{ik}^0 , mean trait vectors χ_{ik}^0 and variance-covariance matrices V_{ik}^0 . We then proceed along the following steps.

1. We let the community evolve according to Eqs. 5 until they reach 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. 5 (as in e.g., Fig. B.1).
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. As described above, we first compute the adaptive-dynamics fitness landscape and locate its peaks. We then let an invader with a mean-trait vector corresponding to each peak invade in turn and let its mean-trait vectors and variance-covariances evolve until they all reach an equilibrium, after which we calculate the invasion fitness of each invader. If any invader with positive invasion fitness is found, the invader with maximum invasion fitness 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. 5.

We continue going through these steps until we have reached a community where Eqs. 5 are in equilibrium and no more invaders with positive invasion fitness can be found. This community is thus eco-evolutionarily stable. An illustration for this case is depicted for the two-patch model with asymmetric resource supplies in Fig. 4.

Following this process allows us to find an eco-evolutionarily stable community with the correct number of species in the final community while including trait variance-covariance. This allows us to explore parameter space to determine how different environments and non-evolving species properties may generate different eco-evolutionarily stable communities. In Fig. 5 we show an

Figure 4: Illustration of the community-assembly procedure for the two-patch model (Box 1) with asymmetric resource supplies. (A–B) The species assembly process starts with one species, and the mean traits are evolved according to Eqs.9 until equilibrium. **(C–D)** At equilibrium, the adaptive-dynamics fitness landscape is calculated, and an invader is tested at each local peak in the fitness landscape. Panels C and D show that one invader with $\chi^{\text{inv}} \approx -4.7$ (green arrows) and one with $\chi^{\text{inv}} \approx 2.1$ (gray dots) are tested. The bases of the green arrows show where the invader mean traits start, and the tip of the arrows show where they end up after their local mean traits and variances equilibrate. The invader with $\chi^{\text{inv}} \approx 2.1$ is close to the resident and ultimately ends up with the same mean trait and variance and resident on both patches, and hence neutral (zero) invasion fitness. While the invader close to the resident also adjusts its mean traits in response to trait variance, this movement is so slight it is not depicted. The invader with $\chi^{\text{inv}} \approx -4.7$ however ends up with positive invasion fitness and will be added to the community. **(E–F)** The community is then once again evolved according to Eqs. 8 until the community is once again at equilibrium. At this point, no more invaders with positive invasion fitness can be found (not shown). **Parameter values:** $\mu = 0.1$, $r = 1$, $K_{1A} = 10$, $K_{1B} = 0.1$, $K_{2A} = 1$, $K_{2B} = 1$, $c_1 = 1$, $c_2 = 1$, $\alpha = 0.4$, $M = 10^{-3}$, and $d = 10^{-2}$.

example for the two-patch model where we assemble eco-evolutionarily stable communities for different dispersal rates d and mutation variances M . As surmised, for high rates of dispersal a single generalist makes up the community, whereas low rates of dispersal engender a two-species community of moderate resource specialists.

In Fig. 5E–F we also display eco-evolutionarily stable communities in the adaptive-dynamics sense, which can be considered the limiting behavior of the moment equations when there is no trait variance. Apart from the obvious difference that our moment equations track the variance of each species, the moment-equation eco-evolutionarily stable communities show two notable differences. The first is that non-zero variance slightly delays the transition to a two-species community compared to adaptive dynamics. This is likely due to both the intraspecific variation itself as well as the local adaptation—that the mean traits can differ from patch A and B for the same species—it engenders. The second difference is the local adaptation itself for the sink-population species on the off-patch for each species. Due to the intraspecific variation, each species is slightly more generalist on the patch on which it is least adapted, as opposed to the adaptive-dynamics case where a single trait describes the entire species in all classes.

The price we pay for using the moment equations in our assembly procedure is that the accuracy of the moment equations in representing the trait-space equations can be variable (Fig. 5D). For the cases we have tested, the moment equation solutions at eco-evolutionary equilibrium typically have good quantitative accuracy, nearly always good qualitative accuracy and only rarely outright disagreement with the trait-space solutions. When there is approximation inaccuracy, it is primarily driven by two things. First, our assembly framework performs worse when the number of species in the trait-space solutions is ambiguous, in this example between one and two species. This is likely due to our assumption that each species can be treated as if it were in reproductive isolation. However, around the transition point the trait-space equations tend to converge on trait-density distributions that are multimodal but without well-separated peaks; for this case, the notion of splitting this distribution into two reproductively isolated distributions does not always work well. Second, under high mutation variances, the parameter space generating this ambiguity is increased. The assembled moment solutions can however be surprisingly effective even when mutation variances—and hence the variances of the species—are high, as seen towards the top left of Fig. 5D. In Fig. C.1 (Appendix C) we show two sample solutions for intermediate and high mutation variances. For the larger mutation variances, approximately $M \geq 3 \cdot 10^{-3}$, and low dispersal rates the model undergoes a regime shift as the trait variation of the species become comparable to the distance between the individual patch optima $\chi_{A,\text{opt}} \approx 1$ and $\chi_{B,\text{opt}} \approx -1$, and the trait-space equations yields unimodal distributions on both patches even for low rates of dispersal. Our assembly model yields two species, but the mean traits of the species are close enough and the variances large enough that the sum of the normal distributions is unimodal (Fig. C.1).

In addition to this, for some small slivers of parameter space, the moment equations do not converge on a stable community, but are instead characterized by eco-evolutionary cycling (Fig. C.2). For Fig. 5 the resolution in parameter is coarse-grained enough that we do not hit any of these points. Experiments with using second-order Taylor expansions did yield larger areas in parameter space characterized by cycling (data not shown) compared to the third-order expansions (Eqs. 7) used to generate Fig. 5.

In spite of these shortcomings, the moment equations coupled with our assembly procedure works remarkably well outside of the small band of parameter space where the number of species

Figure 5: Properties of eco-evolutionarily stable communities in the two-patch model (Box 1) for various rates of dispersal and mutation variances. (A) Number of species in the assembled eco-evolutionarily stable community. The vertical white line indicates the dispersal rate for which the adaptive-dynamics version (zero variance) of the two-patch model undergoes an evolutionary branching from one to two species. The horizontal white line indicates the transect displayed in Panels E and F. (B) Mean-trait value of the resource one specialist on patch A, χ_{1A} . (C) Standard deviation of trait distribution of the resource one specialist on patch A, $\sqrt{V_{1A}}$. (D) Relative error ϵ between the moment equations and trait-space equations at the eco-evolutionarily stable state. The error is calculated as $\epsilon = (1/2) \sum_{I=A,B} \int |v_I^{\text{moment}}(\xi) - v_I^{\text{trait-space}}(\xi)| d\xi / \int v_I^{\text{trait-space}}(\xi) d\xi$. Thus, perfect agreement between the distributions depicted in Fig. 3K–L is $\epsilon \approx 8.97 \cdot 10^{-3}$. (E–F) Fuller display of the transect depicted as a horizontal white line in Panels A–D on patches A and B respectively. The mean traits of the species are depicted as blue lines, and one standard deviation is depicted around these mean traits as blue areas. The adaptive-dynamics eco-evolutionarily stable communities are depicted as orange lines. The modes (locations of local maxima) of the trait-space solutions are depicted as broken black lines. For all panels, parameters other than M and d are as in Fig. 3.

is ambiguous (Fig. 5D).

Additional example: A stage-structured resource-competition model

To 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 two-patch model we used to illustrate the generic equations was deliberately simple. First, it involved only a single trait. Second, the transition between classes (dispersal) was constant, and did not depend on traits. However, such simplifications cannot always be made. In stage-structured models where juveniles mature and adults give birth to new juveniles, these rates will often depend on the traits of the species. 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

$$\frac{dv_J(\xi, t)}{dt} = \int r_A(\eta)v_A(\eta)\mathcal{N}(\xi, \eta, M)d\eta + (r_J(\xi) - \mu_J(\xi))v_J(\xi) - \gamma(\xi)v_J(\xi), \quad (10a)$$

$$\frac{dv_A(\xi, t)}{dt} = \gamma(\xi)v_J(\xi) - \mu_A(\xi)v_A(\xi), \quad (10b)$$

$$\frac{dR_1(t)}{dt} = \rho_1(K_1 - R_1) - \int r_{J1}(\eta)v_J(\eta)d\eta - \int r_{A1}(\eta)v_A(\eta)d\eta, \quad (10c)$$

$$\frac{dR_2(t)}{dt} = \rho_2(K_2 - R_2) - \int r_{J2}(\eta)v_J(\eta)d\eta - \int r_{A2}(\eta)v_A(\eta)d\eta. \quad (10d)$$

Here, the distributions $v_J(\xi)$ and $v_A(\xi)$ 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 $j = 1, 2$. A conceptual illustration of the trait-space equations is depicted in Fig. 1B.

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 μ_A . Juveniles grow at a per capita rate $r_J = r_{J1} + r_{J2}$, suffer a mortality rate of μ_J , and transition into adults at a rate γ that depends both on juvenile growth and mortality ([De Roos et al., 2007](#)). The specific forms of these functions are given in Appendix C. In the absence of consumers, the resources are renewed chemostatically at rates ρ_j up to maximal supplies of K_j for resource j .

We let the trait vector $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$. The first trait ξ_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 the same trait component ξ_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. Such differences could come about as foraging behavior changes between life stages. The second trait ξ_2 parameterizes a trade-off between juvenile mortality

μ_J and adult mortality μ_A . Such trade-offs between juvenile and adult mortality could come about through, e.g., parental investment or risk taking on behalf of juveniles. The parameter values used in our numerical solutions can be found in Appendix C.

Following our general procedure, to get our moment equations we now assume that the solutions to Eqs. 10a–b are composed of a set of S discrete peaks for both juveniles and adults, and approximate the solutions by

$$v_J(\xi, t) \approx \tilde{v}_J(\xi, t) = \sum_{i=1}^S u_{Ji}(t) \mathcal{N}(\xi, \chi_{Ji}(t), V_{Ji}(t)), \quad (11a)$$

$$v_A(\xi, t) \approx \tilde{v}_A(\xi, t) = \sum_{i=1}^S u_{Ai}(t) \mathcal{N}(\xi, \chi_{Ai}(t), V_{Ai}(t)). \quad (11b)$$

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, χ_{Ji} , and adults, χ_{Ai} ; and the trait variance-covariance matrices for juveniles, V_{Ji} , and adults, V_{Ai} . We can derive the moment equations for these quantities by plugging in the expressions for the trait-space equations for this model into our derived generic equations (Eq. 5) yielding the moment equations for this stage-structured model that read

$$\frac{du_{Ji}}{dt} = \hat{g}_{Ji}u_{Ji} - \hat{\gamma}_i u_{Ji} + \hat{r}_{Ai}u_{Ai}, \quad (12a)$$

$$\frac{d\chi_{Ji}}{dt} = V_{Ji}(\nabla_\xi \hat{g}_{Ji} - \nabla_\xi \hat{\gamma}_i) + \frac{u_{Ai}}{u_{Ji}} (V_{Ai} \nabla_\xi \hat{r}_{Ai} + \hat{r}_{Ai}(\chi_{Ai} - \chi_{Ji})), \quad (12b)$$

$$\begin{aligned} \frac{dV_{Ji}}{dt} = & V_{Ji}(\nabla_\xi^2 \hat{g}_{Ji} - \nabla_\xi^2 \hat{\gamma}_i) V_{Ji} \\ & + \frac{u_{Ai}}{u_{Ji}} \left(V_{Ai} \nabla_\xi^2 \hat{r}_{Ai} V_{Ai} + \hat{r}_{Ai} (V_{Ai} - V_{Ji}) + \hat{r}_{Ai} M \right. \\ & + \hat{r}_{Ai}(\chi_{Ai} - \chi_{Ji})(\chi_{Ai} - \chi_{Ji})^T \\ & \left. + V_{Ai} \nabla_\xi \hat{r}_{Ai}(\chi_{Ai} - \chi_{Ji})^T + (\chi_{Ai} - \chi_{Ji})(V_{Ai} \nabla_\xi \hat{r}_{Ai})^T \right), \end{aligned} \quad (12c)$$

$$\frac{du_{Ai}}{dt} = \hat{\gamma}_i u_{Ji} - \hat{\mu}_{Ai} u_{Ai}, \quad (12d)$$

$$\frac{d\chi_{Ai}}{dt} = \frac{u_{Ji}}{u_{Ai}} (V_{Ji} \nabla_\xi \hat{\gamma}_i + \hat{\gamma}_i(\chi_{Ji} - \chi_{Ai})) - V_{Ai} \nabla_\xi \hat{\mu}_{Ai}, \quad (12e)$$

$$\begin{aligned} \frac{dV_{Ai}}{dt} = & \frac{u_{Ji}}{u_{Ai}} \left(V_{Ji} \nabla_\xi^2 \hat{\gamma}_i V_{Ji} + \hat{\gamma}_i (V_{Ji} - V_{Ai}) \right. \\ & + \hat{\gamma}_i(\chi_{Ji} - \chi_{Ai})(\chi_{Ji} - \chi_{Ai})^T \\ & + V_{Ji} \nabla_\xi \hat{\gamma}_i(\chi_{Ji} - \chi_{Ai})^T + (\chi_{Ji} - \chi_{Ai})(V_{Ji} \nabla_\xi \hat{\gamma}_i)^T \\ & \left. - V_{Ai} \nabla_\xi^2 \hat{\mu}_{Ai} V_{Ai} \right), \end{aligned} \quad (12f)$$

$$\frac{dR_1}{dt} = \rho_1(K_1 - R_1) - \sum_{i=1}^S \hat{r}_{Ji1} u_{Ji} - \sum_{i=1}^S \hat{r}_{Ai1} u_{Ai}, \quad (12g)$$

$$\frac{dR_2}{dt} = \rho_2(K_2 - R_2) - \sum_{i=1}^S \hat{r}_{Ji2} u_{Ji} - \sum_{i=1}^S \hat{r}_{Ai2} u_{Ai}. \quad (12h)$$

While somewhat involved, the various terms admit to ecological interpretation, and can be understood in the following way. First, recall that we use functions with a hat as a shorthand for the third-order Taylor approximation of the population rates, see Eqs. 7. We also let g_{Ji} be the net per capita growth of juveniles, i.e., $g_{Ji} = r_{Ji} - \mu_{Ji}$. For the purposes of readability we have also suppressed any dependence on function arguments in the equations above.

The terms of Eq. 12a 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. While suppressed for the sake of notation clarity, these quantities depend also on the variance-covariance matrices of the juveniles and adults through the third-order Taylor approximation (Eqs. 7).

The changes to the mean trait vector of the juveniles χ_{Ji} (Eq. 12b) 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 $\nabla_{\xi} \hat{g}_{Ji}$. 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 $\nabla_{\xi} \hat{g}_{Ji}$. The 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, $V_{Ai} \nabla_{\xi} \hat{r}_{Ai}$, 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 $\nabla_{\xi} \hat{r}_{Ai}$. This effect is stronger if the variation among the adults V_{Ai} is larger. The second term, $\hat{r}_{Ai}(\chi_{Ai} - \chi_{Ji})$, 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 i will on average give birth to juveniles with trait χ_{Ai} , 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 of juveniles by adults has on the juvenile mean trait depends on the mass ratio between the two stages. If there are few adult compared to juveniles, then the birth of new juveniles will have little effect on the mean trait of all the juveniles.

Equation 12c describes the changes to the variance-covariance matrix of the juveniles over time. The first term, $V_{Ji}(\nabla_{\xi}^2 g_{Ji} - \nabla_{\xi}^2 \gamma_i) V_{Ji}$, describes the effects of stabilizing or disruptive selection of the net growth and transitions. Thus, for example, if the curvature of the net growth function, $\nabla_{\xi}^2 g_{Ji}$, 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, $V_{Ai} \nabla_{\xi}^2 r_{Ai} V_{Ai}$, reflects the effects of stabilizing or disruptive selection in the birth rates. The second 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 to mutations. Next, the term $\hat{r}_{Ai}(\chi_{Ai} - \chi_{Ji})(\chi_{Ai} - \chi_{Ji})^T$ 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 juveniles will be away from the mean of the extant juveniles, and thus increase variance. The next two terms, $V_{Ai}\nabla_{\xi}r_{Ai}(\chi_{Ai} - \chi_{Ji})^T + (\chi_{Ai} - \chi_{Ji})(V_{Ai}\nabla_{\xi}r_{Ai})^T$, which are the transposes of one another, take into account the interaction of distance between the means of the stages and effects 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 γ now plays the role that the birth rate function r_A 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. 10), we assume that a juvenile with trait ξ matures into an adult with the same trait ξ , and thus no mutation term enters into the equations for adult variance-covariance. Finally, we approximate the integrals in trait space in the resource equations (Eqs. 10c–d) to second order around the mean traits (see Appendix C for details).

Using these equations we let the population go through our evolutionary assembly process (Fig. 6A). The traits ξ_1 and ξ_2 parameterize the trade-offs in resource affinities and mortalities respectively. However, as the traits that actually matter for the ecology are the affinities 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 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 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. 6A stem mostly from their different trade-offs; the mean traits χ_J and χ_A differ little within the same species. The distributions in $\xi_1 \times \xi_2$ space can be seen in Fig. C.4, where we also compare the numerical moment solutions for the eco-evolutionarily stable two-species community to numerical solutions of the trait-space equations.

Discussion

In this paper we have presented a general framework for eco-evolutionary community assembly 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 covariance for each species in a community, and combining it with a procedure for determining whether additional species need to 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 a spatially-structure two-patch model and a stage-structured model with one juvenile and one adult stage.

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 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 variance in the community. Such an

Figure 6: 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.

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. \(2020\)](#), who derived equations for mean traits and 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. 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.

Finally, our moment equations 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.

Our framework for eco-evolutionary community assembly builds on key ideas from other frameworks where the ecological dynamics can be fruitfully described by differential equations and thus inherits some of their key strengths, but also some of their weaknesses. On the one hand, as opposed to adaptive dynamics ([Metz et al., 1992](#); [Dieckmann and Law, 1996](#); [Geritz et al., 1998](#); [Dercole and Rinaldi, 2008](#); [Kremer and Klausmeier, 2013](#)) where every species is entirely monomorphic with a single trait value, our framework incorporates intraspecific variation while still allowing us to perform an invasion analysis to assemble eco-evolutionarily stable communities. Our evolutionary branching criterion closely mirrors that found in adaptive dynamics. However, the global assembly process we have devised is a heuristic as opposed to the situation in adaptive dynamics where, at least in theory, the fitness landscape can always be computed to ascertain that a community is truly closed to further invasions. In practice, as evidenced by the examples we have presented here, our heuristic has proven reliable for community assembly for many scenarios, often covering most of parameter space, at least for modest mutation variances (Fig. 5D). Our approach is thus a good choice when intraspecific variation needs to be taken into account and the assembly of an eco-evolutionarily stable community is desired.

On the other hand, our framework takes into account replicative infidelity (mutations) as opposed to most moment-based frameworks (e.g., [Norberg et al., 2001](#); [Barabás and D’Andrea, 2016](#)),

which is a key part for our community assembly process. However, this too comes at cost, as evaluating the mutation integrals explicitly required us to assume that both species' trait distributions and mutation kernels were normally distributed. In other moment-based frameworks, other so-called moment closures are also possible, which could be more appropriate depending on the problem under study. For example, [Klauschies et al. \(2018\)](#) and [Cropp and Norbury \(2021\)](#) used beta distributions as the approximating distribution to close their moment equations. 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. 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, including both multimodal distributions as well as highly non-normal unimodal distributions (see Fig. C.1C–D 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.

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 multi-modality 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 assembling the correct number of species, at least as long as the number of species is not ambiguous.

One important benefit of incorporating intraspecific variation into an eco-evolutionary community assembly framework is that it makes possible to study a wider array of diversity-related questions. This allows theoretical explorations of e.g., the effects of productivity on diversity ([Adler et al., 2011](#); [Cusens et al., 2012](#)) and the effects of diversity on ecosystem services ([Cardinale et al., 2011](#)) from a trait-based perspective while recognizing the importance of eco-evolutionarily stable communities ([Edwards et al., 2018](#)). In recent years, there has been much interest in the effects of intraspecific variability ([Albert et al., 2010](#); [Violle et al., 2012](#)) and many different metrics have been proposed and used ([Schleuter et al., 2010](#); [Fontana et al., 2016](#)). Our class-based framework tracks not only the intraspecific trait variation, but also the difference in mean traits between species, as well as the difference in trait distribution between classes within each species. For the case when classes are spatial patches, this opens up the potential to study both how intraspecific and interspecific diversity, as well as alpha and beta diversity are shaped by conditions and influences outcomes.

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 eco-evolutionarily stable communities that accounts for trait variation could help address many ecological questions where intraspecific trait variation is thought to be important.

Acknowledgments

This project was funded by NSF grant DEB-1754250 to E. Litchman and CAK. We thank Florence Débarre, Sébastien Lion, Elena Litchman, and Bob Week for their comments on the work.

Appendix A: Derivation of moment equations

Calculating the time derivatives and general notation

Time derivatives of the moments

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

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

$$\chi_\psi = \frac{1}{u} \int \xi_\psi v(\xi) d\xi \quad \text{Mean-trait vector component } \psi \quad (\text{A.1b})$$

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

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

For the total density u :

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

For the mean trait component χ_ψ :

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

$$= -\frac{1}{u} \int \dot{v}(\xi) \chi_\psi d\xi + \frac{1}{u} \int \xi_\psi \dot{v}(\xi) d\xi = \frac{1}{u} \int (\xi - \chi)_\psi \dot{v}(\xi) d\xi \quad (\text{A.3b})$$

$\Longleftrightarrow$

$$u \dot{\chi}_\psi = \int (\xi - \chi)_\psi \dot{v}(\xi) d\xi \quad (\text{A.3c})$$

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

$$V_{\psi\omega} = \frac{1}{u} \int (\xi - \chi)_{\psi} (\xi - \chi)_{\omega} v(\xi) d\xi \quad (\text{A.4a})$$

$$\Longleftrightarrow$$

$$uV_{\psi\omega} = \int (\xi - \chi)_{\psi} (\xi - \chi)_{\omega} v(\xi) d\xi \quad (\text{A.4b})$$

$$\Longrightarrow$$

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

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

Together we thus have

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

$$u\dot{\chi}_{\psi} = \int (\xi - \chi)_{\psi} \dot{v}(\xi) d\xi \quad (\text{A.5b})$$

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

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.

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

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

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

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

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

We will use Greek letters to indicate trait components. For these indices the Einstein summation 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 indices.

Deriving the moment equations

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

First, let $\mathcal{N}(\xi, \chi, V)$ be a normal distribution probability-density function with argument ξ , mean vector χ , and variance-covariance matrix V . We can then calculate the first and second derivatives of this normal distribution with respect to its argument components:

$$\mathcal{N}_{\psi} = \frac{\partial \mathcal{N}}{\partial \xi_{\psi}} = -V_{\psi\alpha}^{-1}(\xi - \chi)_{\alpha} \mathcal{N}(\xi, \chi, V) \quad (\text{A.8a})$$

$$\mathcal{N}_{\psi\omega} = \frac{\partial^2 \mathcal{N}}{\partial \xi_{\psi} \partial \xi_{\omega}} = \left[V_{\psi\alpha}^{-1}(\xi - \chi)_{\alpha} V_{\omega\beta}^{-1}(\xi - \chi)_{\beta} - V_{\psi\omega}^{-1} \right] \mathcal{N}(\xi, \chi, V) \quad (\text{A.8b})$$

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

$$\hat{f}(\xi, W) = \int f(\eta) \mathcal{N}(\eta, \xi, W) d\eta. \quad (\text{A.9})$$

We can now differentiate the population-level rate with respect to the mean-trait components:

$$\hat{f}_{\psi} = \frac{\partial \hat{f}}{\partial \xi_{\psi}} = W_{\psi\alpha}^{-1} \int (\eta - \xi)_{\alpha} f(\eta) \mathcal{N}(\eta, \xi, W) d\eta \quad (\text{A.10a})$$

$$\hat{f}_{\psi\omega} = \frac{\partial^2 \hat{f}}{\partial \xi_{\psi} \partial \xi_{\omega}} = \int \left[W_{\psi\alpha}^{-1}(\eta - \xi)_{\alpha} W_{\omega\beta}^{-1}(\eta - \xi)_{\beta} - W_{\psi\omega}^{-1} \right] f(\eta) \mathcal{N}(\eta, \xi, W) d\eta \quad (\text{A.10b})$$

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

$$\int (\xi - \chi)_{\psi} f(\xi) \mathcal{N}(\xi, \chi, V) d\xi = V_{\psi\alpha} V_{\alpha\beta}^{-1} \int (\xi - \chi)_{\beta} f(\xi) \mathcal{N}(\xi, \chi, V) d\xi \quad (\text{A.11a})$$

$$= V_{\psi\alpha} \hat{f}_{\alpha}(\chi, V) \quad (\text{A.11b})$$

and

$$\int (\xi - \chi)_{\psi} (\xi - \chi)_{\omega} f(\xi) \mathcal{N}(\xi, \chi, V) d\xi \quad (\text{A.12a})$$

$$= V_{\psi\alpha} V_{\omega\beta} \int V_{\alpha\gamma}^{-1} (\xi - \chi)_{\gamma} V_{\beta\epsilon}^{-1} (\xi - \chi)_{\epsilon} f(\xi) \mathcal{N}(\xi, \chi, V) d\xi \quad (\text{A.12b})$$

$$= V_{\psi\alpha} V_{\omega\beta} \int [V_{\alpha\gamma}^{-1} (\xi - \chi)_{\gamma} V_{\beta\epsilon}^{-1} (\xi - \chi)_{\epsilon} - V_{\alpha\beta}^{-1} + V_{\alpha\beta}^{-1}] f(\xi) \mathcal{N}(\xi, \chi, V) d\xi \quad (\text{A.12c})$$

$$= V_{\psi\alpha} \hat{f}_{\alpha\beta}(\chi, V) V_{\beta\omega} + \hat{f}(\chi, V) V_{\psi\omega} \quad (\text{A.12d})$$

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

Assuming that we have a normally distributed trait-density distribution $v(\xi) := u\mathcal{N}(\xi, \chi, V)$ with total density u , mean χ , 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:

$$\int f(\xi)v(\xi)d\xi = \int f(\xi)u\mathcal{N}(\xi, \chi, V)d\xi = \hat{f}(\chi, V)u \quad (\text{A.13})$$

$$\int (\xi - \chi)_\psi f(\xi)v(\xi)d\xi = \int (\xi - \chi)_\psi f(\xi)u\mathcal{N}(\xi, \chi, V)d\xi \quad (\text{A.14a})$$

$$= uV_{\psi\alpha}\hat{f}_\alpha(\chi, V) \quad (\text{A.14b})$$

$$\int (\xi - \chi)_\psi (\xi - \chi)_\omega f(\xi)v(\xi)d\xi = \int (\xi - \chi)_\psi (\xi - \chi)_\omega f(\xi)u\mathcal{N}(\xi, \chi, V)d\xi \quad (\text{A.15a})$$

$$= u \left(V_{\psi\alpha}\hat{f}_{\alpha\beta}(\chi, V)V_{\beta\omega} + \hat{f}(\chi, V)V_{\psi\omega} \right) \quad (\text{A.15b})$$

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:

$$\int \int f(\eta)v(\eta)\mathcal{N}(\eta, \xi, M)d\eta d\xi = \quad (\text{A.16a})$$

$$= \int f(\eta)v(\eta) \int \mathcal{N}(\eta, \xi, M)d\xi d\eta \quad (\text{A.16b})$$

$$= \int f(\eta)v(\eta)d\eta = \{\text{Eq. A.13}\} = \hat{f}(\chi, V)u \quad (\text{A.16c})$$

The second integral is

$$\int (\xi - \chi)_\psi \int f(\eta)v(\eta)\mathcal{N}(\eta, \xi, M)d\eta d\xi = \int f(\eta)v(\eta) \int (\xi - \chi)_\psi \mathcal{N}(\eta, \xi, M)d\xi d\eta \quad (\text{A.17a})$$

$$= \int f(\eta)v(\eta) \int [(\xi - \eta)_\psi + (\eta - \chi)_\psi] \mathcal{N}(\xi, \eta, M)d\xi d\eta \quad (\text{A.17b})$$

$$= \int (\eta - \chi)_\psi f(\eta)v(\eta)d\eta = \{\text{Eqs. A.14}\} = uV_{\psi\alpha}\hat{f}_\alpha(\chi, V) \quad (\text{A.17c})$$

The third integral is

$$\int (\xi - \chi)_\psi (\xi - \chi)_\omega \int f(\eta)v(\eta)\mathcal{N}(\eta, \xi, M)d\eta d\xi \quad (\text{A.18a})$$

$$= \int f(\eta)v(\eta) \int (\xi - \chi)_\psi (\xi - \chi)_\omega \mathcal{N}(\xi, \eta, M)d\xi d\eta \quad (\text{A.18b})$$

$$= \int f(\eta)v(\eta) \int [(\xi - \eta)_\psi (\xi - \eta)_\omega + (\eta - \chi)_\psi (\xi - \eta)_\omega \quad (\text{A.18c})$$

$$+ (\xi - \eta)_\psi (\eta - \chi)_\omega + (\eta - \chi)_\psi (\eta - \chi)_\omega] \mathcal{N}(\xi, \eta, M)d\xi d\eta \quad (\text{A.18d})$$

$$= M_{\psi\omega} \int f(\eta)v(\eta)d\eta + \int (\eta - \chi)_\psi (\eta - \chi)_\omega f(\eta)v(\eta)d\eta \quad (\text{A.18e})$$

$$= \{\text{Eq. A.13 and Eqs. A.15}\} \quad (\text{A.18f})$$

$$= u \left(M_{\psi\omega} \hat{f}(\chi, V) + V_{\psi\alpha} \hat{f}_{\alpha\beta}(\chi, V) V_{\beta\omega} + \hat{f}(\chi, V) V_{\psi\omega} \right) \quad (\text{A.18g})$$

Recall that the trait-space equations are given by

$$\frac{dv_k(\xi)}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \int f_{klm}(\eta) v_l(\eta) \mathcal{N}(\xi, \eta, M_{klm}) d\eta \quad (\text{A.19})$$

Here, the index $i \in \{1, \dots, S\}$ pertains to species and $k, l \in \{1, \dots, K\}$ pertain to classes. The index $m \in \{1, \dots, N_{kl}\}$ pertains to processes with focal class k and source class l . See main text for details. We now make the assumptions that the trait-density distribution in each class $v_k(\xi)$ can be decomposed into a sum of component distributions (species)

$$v_k(\xi) = \sum_{i=1}^S v_{ik}(\xi), \quad (\text{A.20})$$

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

$$v_{ik}(\xi) \approx \tilde{v}_{ik}(\xi) := u_{ik} \mathcal{N}(\xi, \chi_{ik}, V_{ik}) \quad (\text{A.21})$$

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

For u :

$$\frac{du_{ik}}{dt} = \int \dot{v}_{ik}(\xi) d\xi \approx \int \dot{\tilde{v}}_{ik}(\xi) d\xi = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \int \int f_{klm}(\eta) \tilde{v}_{il}(\eta) \mathcal{N}(\xi, \eta, M_{klm}) d\eta d\xi \quad (\text{A.22a})$$

$$= \{\text{Eqs. A.16}\} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \hat{f}_{klm}(\chi_{il}, V_{il}) u_{il} \quad (\text{A.22b})$$

For χ :

$$u_{ik} \frac{d\chi_{ik\psi}}{dt} = \int (\xi - \chi_{ik})_{\psi} \dot{v}_{ik}(\xi) d\xi \approx \int (\xi - \chi_{ik})_{\psi} \dot{\tilde{v}}_{ik}(\xi) d\xi \quad (\text{A.23a})$$

$$= \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \int (\xi - \chi_{ik})_{\psi} \int f_{klm}(\eta) \tilde{v}_{il}(\eta) \mathcal{N}(\xi, \eta, M_{klm}) d\eta d\xi \quad (\text{A.23b})$$

$$= \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \int [(\xi - \chi_{il})_{\psi} + (\chi_{il} - \chi_{ik})_{\psi}] \int f_{klm}(\eta) \tilde{v}_{il}(\eta) \mathcal{N}(\xi, \eta, M_{klm}) d\eta d\xi \quad (\text{A.23c})$$

$$= \{\text{Eqs. A.16 and Eqs. A.17}\}$$

$$= \sum_{l=1}^K \sum_{m=1}^{N_{kl}} u_{il} \left(V_{\psi\alpha} \hat{f}_{klm\alpha}(\chi_{il}, V_{il}) + \hat{f}_{klm}(\chi_{il}, V_{il}) (\chi_{il} - \chi_{ik})_{\psi} \right) \quad (\text{A.23d})$$

$$\begin{aligned} & \Longleftrightarrow \\ \frac{d\chi_{ik\psi}}{dt} &= \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left(V_{\psi\alpha} \hat{f}_{klm\alpha}(\chi_{il}, V_{il}) + \hat{f}_{klm}(\chi_{il}, V_{il})(\chi_{il} - \chi_{ik})_{\psi} \right) \end{aligned} \quad (\text{A.23e})$$

For V :

$$u_{ik} \frac{dV_{ik\psi\omega}}{dt} + V_{ik\psi\omega} \frac{du_{ik}}{dt} = \int (\xi - \chi_{ik})_{\psi} (\xi - \chi_{ik})_{\omega} \dot{v}_{ik}(\xi) d\xi \quad (\text{A.24a})$$

$$\approx \int (\xi - \chi_{ik})_{\psi} (\xi - \chi_{ik})_{\omega} \dot{\tilde{v}}_{ik}(\xi) d\xi \quad (\text{A.24b})$$

$$= \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \int (\xi - \chi_{ik})_{\psi} (\xi - \chi_{ik})_{\omega} \int f_{klm}(\eta) \tilde{v}_{il}(\eta) \mathcal{N}(\xi, \eta, M_{klm}) d\eta d\xi \quad (\text{A.24c})$$

$$\begin{aligned} &= \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \int [(\xi - \chi_{il})_{\psi} (\xi - \chi_{il})_{\omega} + (\chi_{il} - \chi_{ik})_{\psi} (\xi - \chi_{il})_{\omega} \\ &\quad + (\xi - \chi_{il})_{\psi} (\chi_{il} - \chi_{ik})_{\omega} + (\chi_{il} - \chi_{ik})_{\psi} (\chi_{il} - \chi_{ik})_{\omega}] \int f_{klm}(\eta) \tilde{v}_{il}(\eta) \mathcal{N}(\xi, \eta, M_{klm}) d\eta d\xi \end{aligned} \quad (\text{A.24d})$$

$$= \{\text{Eqs. A.16, Eqs. A.17, and Eqs. A.18}\}$$

$$\begin{aligned} &= \sum_{l=1}^K \sum_{m=1}^{N_{kl}} u_{il} \left[M_{klm\psi\omega} \hat{f}_{klm}(\chi_{il}, V_{il}) + V_{il\psi\alpha} \hat{f}_{klm\alpha\beta}(\chi_{il}, V_{il}) V_{il\beta\omega} + \hat{f}_{klm}(\chi_{il}, V_{il}) V_{il\psi\omega} \right. \\ &\quad + V_{il\psi\alpha} \hat{f}_{klm\alpha}(\chi_{il}, V_{il})(\chi_{il} - \chi_{ik})_{\omega} + V_{il\omega\alpha} \hat{f}_{klm\alpha}(\chi_{il}, V_{il})(\chi_{il} - \chi_{ik})_{\psi} \\ &\quad \left. + \hat{f}_{klm}(\chi_{il}, V_{il})(\chi_{il} - \chi_{ik})_{\psi} (\chi_{il} - \chi_{ik})_{\omega} \right] \end{aligned} \quad (\text{A.24e})$$

$$\begin{aligned} & \Longleftrightarrow \\ \frac{dV_{ik\psi\omega}}{dt} &= \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left[M_{klm\psi\omega} \hat{f}_{klm}(\chi_{il}, V_{il}) + V_{il\psi\alpha} \hat{f}_{klm\alpha\beta}(\chi_{il}, V_{il}) V_{il\beta\omega} \right. \\ &\quad + \hat{f}_{klm}(\chi_{il}, V_{il})(V_{il\psi\omega} - V_{ik\psi\omega}) \\ &\quad + V_{il\psi\alpha} \hat{f}_{klm\alpha}(\chi_{il}, V_{il})(\chi_{il} - \chi_{ik})_{\omega} + V_{il\omega\alpha} \hat{f}_{klm\alpha}(\chi_{il}, V_{il})(\chi_{il} - \chi_{ik})_{\psi} \\ &\quad \left. + \hat{f}_{klm}(\chi_{il}, V_{il})(\chi_{il} - \chi_{ik})_{\psi} (\chi_{il} - \chi_{ik})_{\omega} \right] \end{aligned} \quad (\text{A.24f})$$

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

$$\frac{du_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \hat{f}_{iklm} u_{il} \quad (\text{A.25a})$$

$$\frac{d\chi_{ik\psi}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left(V_{\psi\alpha} \hat{f}_{iklm\alpha} + \hat{f}_{iklm}(\chi_{il} - \chi_{ik})_{\psi} \right) \quad (\text{A.25b})$$

$$\frac{dV_{ik\psi\omega}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left[V_{il\psi\alpha} \hat{f}_{iklm\alpha\beta} V_{il\beta\omega} + \hat{f}_{iklm} (V_{il\psi\omega} - V_{ik\psi\omega}) + M_{klm\psi\omega} \hat{f}_{iklm} \right. \quad (\text{A.25c})$$

$$\left. + V_{il\psi\alpha} \hat{f}_{iklm\alpha} (\chi_{il} - \chi_{ik})_{\omega} + V_{il\omega\alpha} \hat{f}_{iklm\alpha} (\chi_{il} - \chi_{ik})_{\psi} \right. \quad (\text{A.25d})$$

$$\left. + \hat{f}_{iklm} (\chi_{il} - \chi_{ik})_{\psi} (\chi_{il} - \chi_{ik})_{\omega} \right] \quad (\text{A.25e})$$

$$\hat{f}_{iklm} := \hat{f}_{klm}(\chi_{il}, V_{il}, \tilde{\mathbf{v}}) \quad (\text{A.25f})$$

$$\hat{f}_{iklm\alpha} := \frac{\partial \hat{f}_{klm}}{\partial \xi_{\alpha}}(\chi_{il}, V_{il}, \tilde{\mathbf{v}}) \quad (\text{A.25g})$$

$$\hat{f}_{iklm\alpha\beta} := \frac{\partial^2 \hat{f}_{klm}}{\partial \xi_{\alpha} \partial \xi_{\beta}}(\chi_{il}, V_{il}, \tilde{\mathbf{v}}) \quad (\text{A.25h})$$

$$(\text{A.25i})$$

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 $\tilde{v}_k(\xi) = \sum_{i=1}^S u_{ik} \mathcal{N}(\xi, \chi_{ik}, V_{ik})$. For convenience we did not include this dependence during the derivation, as their inclusion does not affect the derivation.

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

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

$$\frac{d\chi_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left(V_{il} \partial_{\xi} \hat{f}_{iklm} + \hat{f}_{iklm} (\chi_{il} - \chi_{ik}) \right) \quad (\text{A.26b})$$

$$\frac{dV_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left[V_{il}^2 \partial_{\xi}^2 \hat{f}_{iklm} + \hat{f}_{iklm} (V_{il} - V_{ik}) + M_{klm} \hat{f}_{iklm} \right. \quad (\text{A.26c})$$

$$\left. + 2V_{il} \partial_{\xi} \hat{f}_{iklm} (\chi_{il} - \chi_{ik}) + \hat{f}_{iklm} (\chi_{il} - \chi_{ik})^2 \right] \quad (\text{A.26d})$$

$$\hat{f}_{klm}(\xi, W, \tilde{\mathbf{v}}) = \int f(\eta, \tilde{\mathbf{v}}) \mathcal{N}(\eta, \xi, W) d\eta \quad (\text{A.26e})$$

$$\hat{f}_{iklm} := \hat{f}_{klm}(\chi_{il}, V_{il}, \tilde{\mathbf{v}}) \quad (\text{A.26f})$$

$$\partial_{\xi} \hat{f}_{iklm} := \frac{\partial \hat{f}_{klm}}{\partial \xi}(\chi_{il}, V_{il}, \tilde{\mathbf{v}}) \quad (\text{A.26g})$$

$$\partial_{\xi}^2 \hat{f}_{iklm} := \frac{\partial^2 \hat{f}_{klm}}{\partial \xi^2}(\chi_{il}, V_{il}, \tilde{\mathbf{v}}) \quad (\text{A.26h})$$

Series solutions and Taylor approximations

The population per capita rates

$$\hat{f}(\xi, W) = \int f(\eta) \mathcal{N}(\eta, \xi, W) d\eta \quad (\text{A.27})$$

may not be analytically computable, and in such cases we may Taylor expand the process rates. To this end let α^n denote an ordered set of indices such that

$$\alpha^n := \alpha_1 \alpha_2 \cdots \alpha_n \quad (\text{A.28})$$

where each α_i indexes the trait-components. Thus, for example

$$f_{\alpha^4} = f_{\alpha_1 \alpha_2 \alpha_3 \alpha_4} \quad (\text{A.29})$$

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

$$\xi^{\alpha^n} := \xi_{\alpha_1} \xi_{\alpha_2} \cdots \xi_{\alpha_n} \quad (\text{A.30})$$

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

$$W^{\alpha^{2n}} := W_{\alpha_1 \alpha_2} W_{\alpha_3 \alpha_4} \cdots W_{\alpha_{2n-1} \alpha_{2n}} \quad (\text{A.31})$$

denote the outer product of n copies of the matrices.

We can now Taylor-expand

$$\hat{f}(\xi, W) = \int f(\eta) \mathcal{N}(\eta, \xi, W) d\eta = \int \sum_{n=0}^{\infty} \frac{f_{\alpha^n}(\xi)}{n!} (\eta - \xi)^{\alpha^n} \mathcal{N}(\eta, \xi, W) d\eta \quad (\text{A.32a})$$

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_{α^n} these expressions simplify considerably so that

$$\int f_{\alpha^n}(\xi) (\eta - \xi)^{\alpha^n} \mathcal{N}(\eta, \xi, W) d\eta = \begin{cases} 0 & \text{if } n \text{ is odd} \\ (n-1)!! f_{\alpha^n} W^{\alpha^n} & \text{if } n \text{ is even} \end{cases} \quad (\text{A.33})$$

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

$$n!! = n(n-2)(n-4) \cdots 1, \quad 0!! = 1, \quad (-1)!! = 1 \quad (\text{A.34})$$

For the Taylor expansion we thus have

$$\int \sum_{n=0}^{\infty} \frac{f_{\alpha^n}(\xi)}{n!} (\eta - \xi)^{\alpha^n} \mathcal{N}(\eta, \xi, W) d\eta = \sum_{\substack{n=0 \\ n \text{ even}}}^{\infty} \frac{f_{\alpha^n}(\xi)}{n!} (n-1)!! W^{\alpha^n} \quad (\text{A.35a})$$

$$= \sum_{n=0}^{\infty} \frac{f_{\alpha^{2n}}(\xi)}{(2n)!} (2n-1)!! W^{\alpha^{2n}} = \sum_{n=0}^{\infty} \frac{f_{\alpha^{2n}}(\xi)}{2^n n!} W^{\alpha^{2n}} \quad (\text{A.35b})$$

and hence

$$\hat{f}(\xi, W) = \sum_{n=0}^{\infty} \frac{f_{\alpha^{2n}}(\xi)}{2^n n!} W^{\alpha^{2n}} = f(\xi) + \frac{1}{2} f_{\alpha\beta}(\xi) W_{\alpha\beta} + \frac{1}{8} f_{\alpha\beta\gamma\epsilon}(\xi) W_{\alpha\beta} W_{\gamma\epsilon} + \dots \quad (\text{A.36a})$$

$$\hat{f}_{\psi}(\xi, W) = \sum_{n=0}^{\infty} \frac{f_{\alpha^{2n}\psi}(\xi)}{2^n n!} W^{\alpha^{2n}} = f_{\psi}(\xi) + \frac{1}{2} f_{\psi\alpha\beta}(\xi) W_{\alpha\beta} + \frac{1}{8} f_{\psi\alpha\beta\gamma\epsilon}(\xi) W_{\alpha\beta} W_{\gamma\epsilon} + \dots \quad (\text{A.36b})$$

$$\hat{f}_{\psi\omega}(\xi, W) = \sum_{n=0}^{\infty} \frac{f_{\alpha^{2n}\psi\omega}(\xi)}{2^n n!} W^{\alpha^{2n}} = f_{\psi\omega}(\xi) + \frac{1}{2} f_{\psi\omega\alpha\beta}(\xi) W_{\alpha\beta} + \frac{1}{8} f_{\psi\omega\alpha\beta\gamma\epsilon}(\xi) W_{\alpha\beta} W_{\gamma\epsilon} + \dots \quad (\text{A.36c})$$

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

$$\hat{f}(\xi, W) = \sum_{n=0}^{\infty} \frac{f_{\alpha^{2n}}(\xi)}{2^n n!} W^{\alpha^{2n}} \approx f(\xi) + \frac{1}{2} f_{\alpha\beta}(\xi) W_{\alpha\beta} \quad (\text{A.37a})$$

$$\hat{f}_{\psi}(\xi, W) = \sum_{n=0}^{\infty} \frac{f_{\alpha^{2n}\psi}(\xi)}{2^n n!} W^{\alpha^{2n}} \approx f_{\psi}(\xi) + \frac{1}{2} f_{\psi\alpha\beta}(\xi) W_{\alpha\beta} \quad (\text{A.37b})$$

$$\hat{f}_{\psi\omega}(\xi, W) = \sum_{n=0}^{\infty} \frac{f_{\alpha^{2n}\psi\omega}(\xi)}{2^n n!} W^{\alpha^{2n}} \approx f_{\psi\omega}(\xi) \quad (\text{A.37c})$$

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

$$\hat{f}(\xi, W) = \sum_{n=0}^{\infty} \frac{f^{(2n)}(\xi)}{2^n n!} W^n = f(\xi) + \frac{1}{2} f^{(2)}(\xi) W + \frac{1}{8} f^{(4)}(\xi) W^2 + \dots \quad (\text{A.38a})$$

$$\hat{f}^{(1)}(\xi, W) = \sum_{n=0}^{\infty} \frac{f^{(2n+1)}(\xi)}{2^n n!} W^n = f^{(1)}(\xi) + \frac{1}{2} f^{(3)}(\xi) W + \frac{1}{8} f^{(5)}(\xi) W^2 + \dots \quad (\text{A.38b})$$

$$\hat{f}^{(2)}(\xi, W) = \sum_{n=0}^{\infty} \frac{f^{(2n+2)}(\xi)}{2^n n!} W^n = f^{(2)}(\xi) + \frac{1}{2} f^{(4)}(\xi) W + \frac{1}{8} f^{(6)}(\xi) W^2 + \dots \quad (\text{A.38c})$$

Variable mutation variance-covariance matrices and mutation bias

For completeness we here presents result of more complicated mutation dynamics where the 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 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 $M(\eta)$ is a positive-definite matrix for all η . The mutation kernel is then given by $\mathcal{N}(\xi, \eta, M(\eta))$. Note that for each fixed η , this is just the probability density function of a normal distribution over ξ .

Working through the problem in the same way as we did for constant mutation matrices (Eqs. A.16–A.17) we can then evaluate the three integrals used to account for mutation, the first of which is:

$$\int \int f(\eta) \mathcal{N}(\eta, \chi, V) (\eta, \xi, M(\eta)) d\eta d\xi = \int f(\eta) \mathcal{N}(\eta, \chi, V) \int \mathcal{N}(\eta, \xi, M(\eta)) d\xi d\eta \quad (\text{A.39a})$$

$$= \int f(\eta) \mathcal{N}(\eta, \chi, V) d\eta = \int f(\xi) \mathcal{N}(\xi, \chi, V) d\xi = \hat{f}(\chi, V) \quad (\text{A.39b})$$

The second integral is

$$\int (\xi - \chi)_\psi \int f(\eta) \mathcal{N}(\eta, \chi, V) (\eta, \xi, M(\eta)) d\eta d\xi = \int f(\eta) \mathcal{N}(\eta, \chi, V) \int (\xi - \chi)_\psi \mathcal{N}(\eta, \xi, M(\eta)) d\xi d\eta \quad (\text{A.40a})$$

$$= \int f(\eta) \mathcal{N}(\eta, \chi, V) \int [(\xi - \eta)_\psi + (\eta - \chi)_\psi] \mathcal{N}(\xi, \eta, M(\eta)) d\xi d\eta \quad (\text{A.40b})$$

$$= \int (\eta - \chi)_\psi f(\eta) \mathcal{N}(\eta, \chi, V) d\eta = \int (\xi - \chi)_\psi f(\xi) \mathcal{N}(\xi, \chi, V) d\xi = V_{\psi\alpha} \hat{f}_\alpha(\chi, V) \quad (\text{A.40c})$$

The third integral is

$$\int (\xi - \chi)_\psi (\xi - \chi)_\omega \int f(\eta) \mathcal{N}(\eta, \chi, V) (\eta, \xi, M(\eta)) d\eta d\xi \quad (\text{A.41a})$$

$$= \int f(\eta) \mathcal{N}(\eta, \chi, V) \int (\xi - \chi)_\psi (\xi - \chi)_\omega \mathcal{N}(\xi, \eta, M(\eta)) d\xi d\eta \quad (\text{A.41b})$$

$$= \int f(\eta) \mathcal{N}(\eta, \chi, V) \int [(\xi - \eta)_\psi (\xi - \eta)_\omega + (\eta - \chi)_\psi (\xi - \eta)_\omega \quad (\text{A.41c})$$

$$+ (\xi - \eta)_\psi (\eta - \chi)_\omega + (\eta - \chi)_\psi (\eta - \chi)_\omega] \mathcal{N}(\xi, \eta, M(\eta)) d\xi d\eta \quad (\text{A.41d})$$

$$= \int f(\eta) M_{\psi\omega}(\eta) \mathcal{N}(\eta, \chi, V) d\eta + \int (\eta - \chi)_\psi (\eta - \chi)_\omega f(\eta) \mathcal{N}(\eta, \chi, M) d\eta \quad (\text{A.41e})$$

$$= \widehat{fM}_{\psi\omega}(\chi, V) + V_{\psi\alpha} \hat{f}_{\alpha\beta}(\chi, V) V_{\beta\omega} + \hat{f}(\chi, V) V_{\psi\omega} \quad (\text{A.41f})$$

The mutation term in the moment equations when the mutation variance-covariance matrix is trait-dependent will thus read

$$(\widehat{fM})_{iklm\psi\omega} = \int f_{klm}(\xi) M_{klm\psi\omega}(\xi) \mathcal{N}(\xi, \chi_{il}, V_{il}) d\xi \quad (\text{A.42})$$

To second order we have

$$\int f(\eta) M_{\psi\omega}(\eta) \mathcal{N}(\eta, \chi, V) d\eta \quad (\text{A.43a})$$

$$\approx f(\chi) M_{\psi\omega}(\chi) + \frac{1}{2} V_{\alpha\beta} [f M_{\psi\omega}]_{\alpha\beta}(\chi) \quad (\text{A.43b})$$

$$= f(\chi) M_{\psi\omega}(\chi) + \frac{1}{2} V_{\alpha\beta} \left[f_{\alpha\beta}(\chi) M_{\psi\omega}(\chi) + f_\alpha(\chi) M_{\psi\omega\beta}(\chi) + f_\beta(\chi) M_{\psi\omega\alpha}(\chi) + f(\chi) M_{\psi\omega\alpha\beta}(\chi) \right] \quad (\text{A.43c})$$

Apart from the term in the moment equations that add to variance-covariance through mutation, which is now given by Eq. A.42, the rest of the moment equations remain unchanged under the assumption that the variance-covariance matrix is trait dependent.

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

$$\int f(\eta) v(\eta) \mathcal{N}(\eta, \xi - \mu(\eta), M(\eta)) d\eta \quad (\text{A.44})$$

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

$$\int \int f(\eta) \mathcal{N}(\eta, \chi, V)(\eta, \xi - \mu(\eta), M(\eta)) d\eta d\xi = \int f(\eta) \mathcal{N}(\eta, \chi, V) \int \mathcal{N}(\xi, \eta + \mu(\eta), M(\eta)) d\xi d\eta \quad (\text{A.45a})$$

$$= \int f(\eta) \mathcal{N}(\eta, \chi, V) d\eta = \int f(\xi) \mathcal{N}(\xi, \chi, V) d\xi = \hat{f}(\chi, V) \quad (\text{A.45b})$$

The second mutation integral is

$$\int (\xi - \chi)_\psi \int f(\eta) \mathcal{N}(\eta, \chi, V)(\eta, \xi - \mu(\eta), M(\eta)) d\eta d\xi \quad (\text{A.46a})$$

$$= \int f(\eta) \mathcal{N}(\eta, \chi, V) \int (\xi - \chi)_\psi \mathcal{N}(\xi, \eta + \mu(\eta), M(\eta)) d\xi d\eta \quad (\text{A.46b})$$

$$= \int f(\eta) \mathcal{N}(\eta, \chi, V) \int [(\xi - (\eta + \mu(\eta)))_\psi + (\eta + \mu(\eta) - \chi)_\psi] \mathcal{N}(\xi, \eta + \mu(\eta), M(\eta)) d\xi d\eta \quad (\text{A.46c})$$

$$= \int [(\eta - \chi)_\psi + \mu(\eta)_\psi] f(\eta) \mathcal{N}(\eta, \chi, V) d\eta = V_{\psi\alpha} \hat{f}_\alpha(\chi, V) + \widehat{f\mu}_\psi(\chi, V) \quad (\text{A.46d})$$

where

$$\widehat{f\mu}_\psi(\chi, V) = \int f(\eta) \mu_\psi(\eta) \mathcal{N}(\eta, \chi, V) d\eta \quad (\text{A.47})$$

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

$$\int (\xi - \chi)_\psi (\xi - \chi)_\omega \int f(\eta) \mathcal{N}(\eta, \chi, V)(\eta, \xi - \mu(\eta), M(\eta)) d\eta d\xi \quad (\text{A.48a})$$

$$= \int f(\eta) \mathcal{N}(\eta, \chi, V) \int (\xi - \chi)_\psi (\xi - \chi)_\omega \mathcal{N}(\xi, \tilde{\eta}, M(\eta)) d\xi d\eta \quad (\text{A.48b})$$

$$= \int f(\eta) \mathcal{N}(\eta, \chi, V) \int [(\xi - \tilde{\eta})_\psi (\xi - \tilde{\eta})_\omega + (\tilde{\eta} - \chi)_\psi (\xi - \tilde{\eta})_\omega \quad (\text{A.48c})$$

$$+ (\xi - \tilde{\eta})_\psi (\tilde{\eta} - \chi)_\omega + (\tilde{\eta} - \chi)_\psi (\tilde{\eta} - \chi)_\omega] \mathcal{N}(\xi, \tilde{\eta}, M(\eta)) d\xi d\eta \quad (\text{A.48d})$$

$$= \int f(\eta) M_{\psi\omega}(\eta) \mathcal{N}(\eta, \chi, V) d\eta + \int (\tilde{\eta} - \chi)_\psi (\tilde{\eta} - \chi)_\omega f(\eta) \mathcal{N}(\eta, \chi, V) d\eta \quad (\text{A.48e})$$

$$= \widehat{fM}_{\psi\omega}(\chi, V) + \int (\eta - \chi)_\psi (\eta - \chi)_\omega f(\eta) \mathcal{N}(\eta, \chi, V) d\eta \quad (\text{A.48f})$$

$$+ \int \mu(\eta)_\psi (\eta - \chi)_\omega f(\eta) \mathcal{N}(\eta, \chi, V) d\eta + \int (\eta - \chi)_\psi \mu(\eta)_\omega f(\eta) \mathcal{N}(\eta, \chi, V) d\eta \quad (\text{A.48g})$$

$$+ \int \mu(\eta)_\psi \mu(\eta)_\omega f(\eta) \mathcal{N}(\eta, \chi, V) d\eta \quad (\text{A.48h})$$

$$= \widehat{fM}_{\psi\omega}(\chi, V) + V_{\psi\alpha} \hat{f}_{\alpha\beta}(\chi, V) V_{\beta\omega} + \hat{f}(\chi, V) V_{\psi\omega} \quad (\text{A.48i})$$

$$+ V_{\psi\alpha} (\widehat{\mu_\psi f})_\alpha(\chi, V) + V_{\omega\alpha} (\widehat{\mu_\psi f})_\alpha(\chi, V) + \widehat{\mu_\psi \mu_\omega f}(\chi, V) \quad (\text{A.48j})$$

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

$$\frac{du_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \hat{f}_{iklm} u_{il} \quad (\text{A.49a})$$

$$\frac{d\chi_{ik\psi}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left(V_{il\psi\alpha} \hat{f}_{iklm\alpha} + \hat{f}_{iklm}(\chi_{il} - \chi_{ik})_{\psi} + \widehat{f\mu}_{iklm\psi} \right) \quad (\text{A.49b})$$

$$\frac{dV_{ik\psi\omega}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left(\widehat{fM}_{iklm\psi\omega} + V_{il\psi\alpha} \hat{f}_{iklm\alpha\beta} V_{il\beta\omega} + \hat{f}_{iklm}(V_{il\psi\omega} - V_{ik\psi\omega}) \right) \quad (\text{A.49c})$$

$$+ V_{il\psi\alpha}(\widehat{f\mu}_{iklm\omega})_{\alpha} + V_{il\omega\alpha}(\widehat{f\mu}_{iklm\psi})_{\alpha} + \widehat{\mu\mu} f_{iklm\psi\omega} \quad (\text{A.49d})$$

$$+ (\chi_{il} - \chi_{ik})_{\psi}(V_{il\omega\alpha} \hat{f}_{iklm\alpha} + \widehat{f\mu}_{iklm\omega}) + (\chi_{il} - \chi_{ik})_{\omega}(V_{il\psi\alpha} \hat{f}_{iklm\alpha} + \widehat{f\mu}_{iklm\psi}) \quad (\text{A.49e})$$

$$+ \hat{f}_{iklm}(\chi_{il} - \chi_{ik})_{\psi}(\chi_{il} - \chi_{ik})_{\omega} \quad (\text{A.49f})$$

Here, all functions with subscript $iklm$ are evaluated at χ_{il} and V_{il} .

In vector-matrix form these equations are written as

$$\frac{du_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \hat{f}_{iklm} u_{il} \quad (\text{A.50a})$$

$$\frac{d\chi_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left(V_{il} \nabla_{\xi} \hat{f}_{iklm} + \hat{f}_{iklm}(\chi_{il} - \chi_{ik}) + \widehat{f\mu}_{iklm} \right) \quad (\text{A.50b})$$

$$\frac{dV_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left(\widehat{fM}_{iklm} + V_{il} \nabla_{\xi}^2 \hat{f}_{iklm} V_{il} + \hat{f}_{iklm}(V_{il} - V_{ik}) \right) \quad (\text{A.50c})$$

$$+ V_{il} \nabla_{\xi} \widehat{f\mu}_{iklm} + [V_{il} \nabla_{\xi} \widehat{f\mu}_{iklm}]^T + \widehat{\mu\mu}^T \hat{f}_{iklm} \quad (\text{A.50d})$$

$$+ (\chi_{il} - \chi_{ik})(V_{il} \nabla_{\xi} \hat{f}_{iklm} + \widehat{f\mu}_{iklm})^T + (V_{il} \nabla_{\xi} \hat{f}_{iklm} + \widehat{f\mu}_{iklm})(\chi_{il} - \chi_{ik})^T \quad (\text{A.50e})$$

$$+ \hat{f}_{iklm}(\chi_{il} - \chi_{ik})(\chi_{il} - \chi_{ik})^T \quad (\text{A.50f})$$

For scalar traits, these simplify somewhat to

$$\frac{du_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \hat{f}_{iklm} u_{il} \quad (\text{A.51a})$$

$$\frac{d\chi_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left(V_{il} \partial_{\xi} \hat{f}_{iklm} + \hat{f}_{iklm}(\chi_{il} - \chi_{ik}) + \widehat{f\mu}_{iklm} \right) \quad (\text{A.51b})$$

$$\frac{dV_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} \left(\widehat{fM}_{iklm} + V_{il}^2 \partial_{\xi}^2 \hat{f}_{iklm} + \hat{f}_{iklm}(V_{il} - V_{ik}) \right) \quad (\text{A.51c})$$

$$+ 2V_{il} \partial_{\xi} \widehat{f\mu}_{iklm} + \widehat{\mu^2} \hat{f}_{iklm} \quad (\text{A.51d})$$

$$+ 2(\chi_{il} - \chi_{ik})(V_{il} \partial_{\xi} \hat{f}_{iklm} + \widehat{f\mu}_{iklm}) \quad (\text{A.51e})$$

$$+ \hat{f}_{iklm}(\chi_{il} - \chi_{ik})^2 \quad (\text{A.51f})$$

Environmental variables

In our examples we have included non-evolving environmental variables. 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 modern 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 our two-patch model example, resources follow the same class structure as the species and are distributed among the patches, but in our stage-structured model 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 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 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

$$\frac{dv_k(\xi)}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \int f_{klm}(\eta, \mathbf{v}, \mathbf{E}) v_l(\eta) \mathcal{N}(\xi, \eta, M_{klm}) \quad (\text{A.52a})$$

$$\frac{dE_{jn_j}}{dt} = \rho_{jn_j}(\mathbf{E}) + \sum_{l=1}^K \int I_{jn_jl}(\mathbf{E}, \xi) v_l(\xi) d\xi \quad (\text{A.52b})$$

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

As before, we now make the assumption that the trait-density distribution $v_l(\xi)$ can be approximated by a sum of normal distributions:

$$v_l(\xi) \approx \tilde{v}_l(\xi) = \sum_{i=1}^S u_{il} \mathcal{N}(\xi, \chi_{il}, V_{il}) \quad (\text{A.53})$$

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

$$\hat{I}_{ijn_jl} := \int I_{jn_jl}(\mathbf{E}, \xi) \mathcal{N}(\xi, \chi_{il}, V_{il}) d\xi \quad (\text{A.54})$$

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

$$\hat{I}_{ijn_jl} \approx I_{jn_jl}(\mathbf{E}, \chi_{il}) + \frac{1}{2} \nabla_{\xi}^2 I_{jn_jl}(\mathbf{E}, \chi_{il}) : V_{il} \quad (\text{A.55})$$

Thus, the moment equations including environmental variables read

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

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

$$\begin{aligned} \frac{dV_{ik}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_{il}}{u_{ik}} & \left(V_{il} \nabla_{\xi}^2 \hat{f}_{iklm} V_{il} + \hat{f}_{iklm} (V_{il} - V_{ik}) + \hat{f}_{iklm} M_{klm} \right. \\ & + \hat{f}_{iklm} (\chi_{il} - \chi_{ik}) (\chi_{il} - \chi_{ik})^T \\ & \left. + V_{il} \nabla_{\xi} \hat{f}_{iklm} (\chi_{il} - \chi_{ik})^T + (\chi_{il} - \chi_{ik}) \left[V_{il} \nabla_{\xi} \hat{f}_{iklm} \right]^T \right) \end{aligned} \quad (\text{A.56c})$$

$$\frac{dE_{jn_j}}{dt} = \rho_{jn_j} + \sum_{i=1}^S \sum_{l=1}^K \hat{I}_{jn_jil} u_{il} \quad (\text{A.56d})$$

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

In Appendix C we use this general formulation to derive the rates of change for resources in a two-patch model and a stage-structured model.

Explicit frequency-dependent dynamics

Though not part of any our example models, some models include explicit frequency-dependent dynamics. For example, consider our focal organism with trait-density distribution $v(\xi)$ as the prey of a non-evolving predator with density P consuming the species with a type-II functional response with attack rate a and handling time h . If we now let the attack rates of the predators depend on the trait of the prey we have:

$$\frac{dP}{dt} = c \frac{\int a(\xi) v(\xi) d\xi}{1 + h \int a(\xi) v(\xi) d\xi} P - m_P P \quad (\text{A.57a})$$

$$\frac{dv(\xi)}{dt} = \int b(\eta) v(\eta) \mathcal{N}(\xi, \eta, M) d\eta - \frac{a(\xi) P}{1 + h \int a(\eta) v(\eta) d\eta} v(\xi) \quad (\text{A.57b})$$

here, the mortality rate

$$m(\xi, v) = \frac{a(\xi) P}{1 + h \int a(\eta) v(\eta) d\eta} \quad (\text{A.58})$$

of the prey depends explicitly on the entire trait-density distribution of prey. However, under the assumption that the trait-density distribution $v(\xi)$ can be approximated by a sum of normals, i.e.,

$$v(\xi) \approx \tilde{v}(\xi) = \sum_{j=1}^S u_j \mathcal{N}(\xi, \chi_j, V_j) \quad (\text{A.59})$$

we may approximate the mortality rate by

$$m(\xi, v) \approx m(\xi, \tilde{v}) = \frac{a(\xi) P}{1 + h \sum_{j=1}^S u_j \int a(\eta) \mathcal{N}(\xi, \chi_j, V_j) d\eta} \quad (\text{A.60a})$$

$$= \frac{a(\xi)P}{1 + h \sum_{j=1}^S \hat{a}(\chi_j, V_j)u_j} \quad (\text{A.60b})$$

where

$$\hat{a}(\xi, W) := \int a(\eta) \mathcal{N}(\eta, \xi, W) d\eta \quad (\text{A.61})$$

As for our general derivation, this integral may either be evaluated analytically if possible or approximated with a Taylor-expansion of a . Thus, as long as the explicit frequency dependence of any model comes in the form of functions of linear functionals of v this general recipe will always work. Typically, a linear functional F of v would take the form

$$F[v](\xi) = \int K(\xi, \eta) v(\eta) d\eta \quad (\text{A.62})$$

For example, in the predator-prey system above, $K(\xi, \eta) = a(\eta)$.

Another example would be Lotka-Volterra dynamics (omitting mutations for brevity) where

$$\frac{dv(\xi)}{dt} = (b(\xi) - m(\xi, v))v(\xi) = \left(r(\xi) - \int \alpha(\xi, \eta) v(\eta) d\eta \right) v(\xi) \quad (\text{A.63})$$

If we let

$$\alpha(\xi, \eta) = \mathcal{N}(\xi, \eta, V_\alpha) \quad (\text{A.64})$$

be the competition kernel, then, once again assuming that v is a sum of normal distributions we have

$$\int \alpha(\xi, \eta) \tilde{v}(\eta) d\eta = \sum_{j=1}^S u_j \int \alpha(\xi, \eta) \mathcal{N}(\eta, \chi_j, V_j) d\eta \quad (\text{A.65a})$$

$$= \sum_{j=1}^S u_j \hat{\alpha}(\xi, \chi_j, V_j) = \sum_{j=1}^S u_j \mathcal{N}(\xi, \chi_j, V_j + V_\alpha) = m(\xi, \tilde{v}) \quad (\text{A.65b})$$

In this case we could solve the integral analytically, and furthermore, for our moment equations we can obtain

$$\hat{m}(\xi, W) = \int m(\eta, \tilde{v}) \mathcal{N}(\eta, \xi, W) d\eta = \sum_{j=1}^S u_j \mathcal{N}(\xi, \chi_j, W + V_j + V_\alpha) \quad (\text{A.66})$$

so that for species i we would have

$$\hat{m}(\chi_i, V_i) = \sum_{j=1}^S u_j \mathcal{N}(\chi_i, \chi_j, V_i + V_j + V_\alpha) \quad (\text{A.67})$$

Appendix B: Invasion heuristics

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 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 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 happens either when the number of species is ambiguous (a distribution with several peaks, but where the peaks are not well separated).

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 two different ways of investigating when diversification 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. The second condition treats global invasion analysis where the invader does not necessarily have a trait that is close to any resident species.

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 in their traits.

Consider a community of S species in K classes being in a stable eco-evolutionary equilibrium so that

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

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

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

$$\left. + V_{il} \nabla_{\xi} \hat{f}_{iklm} (\chi_{il} - \chi_{ik})^T + (\chi_{il} - \chi_{ik}) \left[V_{il} \nabla_{\xi} \hat{f}_{iklm} \right]^T + \right. \quad (\text{B.1d})$$

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

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

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

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

$$\chi_{i_1k} = \chi_{i_2k} = \chi_{ik} \quad (\text{B.2b})$$

$$V_{i_1k} = V_{i_2k} = V_{ik} \quad (\text{B.2c})$$

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

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

where $J_F(y^*)$ is the Jacobian matrix of F evaluated at y^* . Normally, in dynamical systems theory, 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 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 model (Box 1, main text). The blue 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 λ_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 ϕ be the eigenvector associated with the dominant eigenvalue λ_d . The components of the eigenvector are

$$u_{1A}^\phi \approx -0.4730 \quad u_{1B}^\phi \approx 0.4730 \quad (\text{B.4a})$$

$$u_{2A}^\phi \approx 0.4730 \quad u_{2B}^\phi \approx -0.4730 \quad (\text{B.4b})$$

$$\chi_{1A}^\phi \approx -0.1612 \quad \chi_{1B}^\phi \approx -0.1612 \quad (\text{B.4c})$$

$$\chi_{2A}^\phi \approx 0.1612 \quad \chi_{2B}^\phi \approx 0.1612 \quad (\text{B.4d})$$

$$V_{1A}^\phi \approx 0.01727 \quad V_{1B}^\phi \approx -0.01727 \quad (\text{B.4e})$$

$$V_{2A}^\phi \approx -0.01727 \quad V_{2B}^\phi \approx 0.01727 \quad (\text{B.4f})$$

In other words, the unstable direction corresponds to one species becoming relatively more abundant on patch A and more resource one specialized, and one species becoming relatively more abundant on patch B and more resource two specialized. We construct the new two-species system after the branching by letting

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

where α is some small number, for this example $\alpha = 0.05$.

Global invasion dynamics

Local branching in trait space will not always be sufficient to ensure that community is globally closed to further invasions. Here, we will work out a criterion for when an additional species with

Figure B.1: Illustration of branching procedure for the two-patch model (Box 1, 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 A and the right two panels depict patch B. Panels (C–D) also depict one standard deviation around the mean as filled areas. The initial species (blue) evolves to an eco-evolutionary equilibrium after which a branching takes place, the species is split, and the two new species (green and orange) 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:** $\mu = 0.1$, $r = 1$, $K_{1A} = 1$, $K_{1B} = 0.4$, $K_{2A} = 0.4$, $K_{2B} = 1$, $c_1 = 1$, $c_2 = 1$, $\alpha = 0.5$, $M = 5 \cdot 10^{-6}$, $d = 8 \cdot 10^{-3}$.

an arbitrary trait can be added to the community in such a way that this added species grows from being initially rare.

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

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

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

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

$$\left. + V_{il} \nabla_{\xi} \hat{f}_{iklm} (\chi_{il} - \chi_{ik})^T + (\chi_{il} - \chi_{ik}) \left[V_{il} \nabla_{\xi} \hat{f}_{iklm} \right]^T + \right. \quad (\text{B.6d})$$

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

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

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 very rare invader enter this community, and try to determine whether an invader with a given trait can grow in the environment set by the resident. However, we also need to account for the internal dynamics of the invader, i.e., its mean traits and variance-covariances across classes. In principle, we would thus need to test for the ultimate invasion success of all invaders with all possible combinations of mean traits and variance-covariance matrices across all classes. This, however, presents two issues. First, the state space of all such combinations grows very quickly as the number of traits or number of classes increases making this approach infeasible. Second, since we often need to employ a Taylor approximation for the various rate functions governing the moment equations, we wish to avoid counting invaders with variances that are so large that their per capita growth rates can be spuriously positive due to a e.g., a positive second derivative. Thus, we take a two-stage approach for determining whether any successful invaders can be found. First, we assume that the variance-covariance of the invader is very small and let its mean traits evolve until equilibrium. Then, we let the variance-covariance matrices of the invader change until equilibrium as well, all while the invader is considered to be so rare as not to affect its environment. Finally, when the invader has reached an equilibrium in terms of mean traits and variance-covariances, we can calculate whether this invader will grow or decline in the environment set by the resident. Below, we will in turn make precise these notions and how to find any potential successful invaders. This invasion procedure was inspired by the invasion procedure used by Kremer and Klausmeier (2013).

For our first step, we assume that the invader is both very rare and that its variance-covariance is negligibly small. In this case, the term $\hat{f}_{iklm}(\chi_{il} - \chi_{ik})$ will dominate over the term $V_{il} \nabla_{\xi} \hat{f}_{iklm}$ in Eq. B.6b, with the result being that the mean trait vector can be considered to be the same in each class (c.f., Wickman et al., 2017), making the situation analogous to adaptive dynamics. We can thus calculate the invasion fitness of the mutant with mean trait χ^{inv} as the dominant eigenvalue

of the matrix F^0 , where the entries of F^0 are given by

$$F_{kl}^0 := \sum_{m=1}^{N_{kl}} f_{klm}(\chi^{\text{inv}}, \tilde{\mathbf{v}}^{\text{res}}), \quad (\text{B.7})$$

see e.g., Caswell (2001). This *adaptive-dynamics fitness landscape* describes how a mutant with negligible density and variance-covariance (and hence negligible local adaptation) would change over time in terms of mean trait, as the mutant would climb the fitness landscape to the peaks. For our second step, where we allow the variance-covariance matrices of the mutant to evolve, we thus need only consider such traits that correspond to the fitness peaks of this landscape (see Fig. 4 in the main text and Fig. B.2 for examples).

There is however a problem that first needs to be overcome: The equations for the mean traits and variance-covariance matrices depend explicitly on the densities u_{ik} . Crucially however, they only depend on the relative densities of the classes u_{il}/u_{ik} for some classes k and l . We will use this fact to formulate a set of invasion equations that allows the density distribution over the classes to vary without the total density of an invader across all classes changing over time.

First, we construct a matrix with entries F_{kl}^{inv} governing the per capita rates of an invader:

$$F_{kl}^{\text{inv}} := \sum_{m=1}^{N_{kl}} \hat{f}_{klm}^{\text{inv}} := \sum_{m=1}^{N_{kl}} \hat{f}_{klm}(\chi_l^{\text{inv}}, V_l^{\text{inv}}, \tilde{\mathbf{v}}^{\text{res}}) \quad (\text{B.8})$$

Given this notation, the density of a mutant would change over time according to

$$\frac{du_k^{\text{inv}}}{dt} = \sum_{l=1}^K F_{kl}^{\text{inv}} u_l^{\text{inv}}. \quad (\text{B.9})$$

However, since we have assumed that the total density of the invader is very low, it will not affect F , and the densities of the invader will either grow towards infinity or decline towards zero, which will cause numerical problems. We can remedy this situation by tracking the relative densities across classes by changing F in such a way that it subtracts the total growth or decline across classes, while still allowing the invading mutant to assume the correct relative distribution among the classes. Thus, we let $\tilde{u}_k^{\text{inv}}$ be the normalized frequency distribution among classes, i.e.,

$$\tilde{u}_k^{\text{inv}} := \frac{u_k^{\text{inv}}}{\sum_{l=1}^K u_l^{\text{inv}}}. \quad (\text{B.10})$$

Furthermore, we let $\tilde{F}_{kl}^{\text{inv}}$ be the relative per capita growth matrix where the total growth is compensated for, i.e.,

$$\tilde{F}_{kl}^{\text{inv}} := F_{kl}^{\text{inv}} - \delta_{kl} \sum_{p=1}^K \sum_{q=1}^K F_{pq}^{\text{inv}} \tilde{u}_q^{\text{inv}}, \quad (\text{B.11})$$

where δ_{kl} is the Kronecker delta that equals one for $k = l$ and zero otherwise. If we then let the densities of the invading mutant change according to

$$\frac{du_k^{\text{inv}}}{dt} = \sum_{l=1}^K \tilde{F}_{kl}^{\text{inv}} u_l^{\text{inv}} \quad (\text{B.12})$$

then the total density of $u_{\text{tot}}^{\text{inv}} = \sum_{k=1}^K u_k^{\text{inv}}$ across classes, will not change over time, as can be seen by calculating

$$\begin{aligned} \frac{du_{\text{tot}}^{\text{inv}}}{dt} &= \frac{d \sum_{k=1}^K u_k^{\text{inv}}}{dt} = \sum_{k=1}^K \frac{du_k^{\text{inv}}}{dt} = \sum_{k=1}^K \sum_{l=1}^K \tilde{F}_{kl}^{\text{inv}} u_l^{\text{inv}} = \\ &= \sum_{k=1}^K \sum_{l=1}^K \left(F_{kl}^{\text{inv}} - \delta_{kl} \sum_{p=1}^K \sum_{q=1}^K F_{pq}^{\text{inv}} \tilde{u}_q^{\text{inv}} \right) u_l^{\text{inv}} = \\ &= \sum_{k=1}^K \sum_{l=1}^K F_{kl}^{\text{inv}} u_l^{\text{inv}} - \sum_{k=1}^K \sum_{l=1}^K \delta_{kl} u_l^{\text{inv}} \sum_{p=1}^K \sum_{q=1}^K F_{pq}^{\text{inv}} \frac{u_q^{\text{inv}}}{\sum_{r=1}^K u_r^{\text{inv}}} = \\ &= \sum_{k=1}^K \sum_{l=1}^K F_{kl}^{\text{inv}} u_l^{\text{inv}} - \frac{\sum_{k=1}^K u_k^{\text{inv}}}{\sum_{r=1}^K u_r^{\text{inv}}} \sum_{p=1}^K \sum_{q=1}^K F_{pq}^{\text{inv}} u_q^{\text{inv}} = \end{aligned} \quad (\text{B.13a})$$

$$= \sum_{k=1}^K \sum_{l=1}^K F_{kl}^{\text{inv}} u_l^{\text{inv}} - \sum_{p=1}^K \sum_{q=1}^K F_{pq}^{\text{inv}} u_q^{\text{inv}} = 0. \quad (\text{B.13b})$$

Having computed how to track the relative densities in each class, we may now carry out the second step of our invasion process, and evaluate the invasion fitness of the mutant at the end. The second step is computed as follows.

We pick an initial mean trait vector that maximizes the fitness landscape as described above, $\chi_k^{\text{inv},0}$, that is the same for each class k , and choose a variance-covariance matrix $V_k^{\text{inv},0}$ that is also the same for each class, without covariances and with some small variances for each trait. We let the initial density-distribution $u_k^{\text{inv},0}$ be the eigenvector associated with the eigenvalue of F^0 at a fitness peak (the stable class distribution for zero variance-covariance), normalized so that $\sum_{k=1}^K u_k^{\text{inv},0} = 1$. We then solve the equations

$$\frac{du_k^{\text{inv}}}{dt} = \sum_{l=1}^K \tilde{F}_{kl}^{\text{inv}} u_l^{\text{inv}} \quad (\text{B.14a})$$

$$\frac{d\chi_k^{\text{inv}}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_l^{\text{inv}}}{u_k^{\text{inv}}} \left(V_l^{\text{inv,fix}} \nabla_{\xi} f_{klm}^{\text{inv}} + \hat{f}_{klm}^{\text{inv}} (\chi_l^{\text{inv}} - \chi_k^{\text{inv}}) \right) \quad (\text{B.14b})$$

$$\frac{dV_k^{\text{inv}}}{dt} = \sum_{l=1}^K \sum_{m=1}^{N_{kl}} \frac{u_l^{\text{inv}}}{u_k^{\text{inv}}} \left(V_l^{\text{inv}} \nabla_{\xi}^2 f_{klm}^{\text{inv}} V_l^{\text{inv}} + \hat{f}_{klm}^{\text{inv}} M_{klm} + \right. \quad (\text{B.14c})$$

$$\left. + V_l^{\text{inv}} \nabla_{\xi} f_{klm}^{\text{inv}} (\chi_l^{\text{inv}} - \chi_k^{\text{inv}})^T + (\chi_l^{\text{inv}} - \chi_k^{\text{inv}}) \left[V_l^{\text{inv}} \nabla_{\xi} f_{klm}^{\text{inv}} \right]^T + \right. \quad (\text{B.14d})$$

$$\left. + \hat{f}_{klm}^{\text{inv}} \left[V_l^{\text{inv}} - V_k^{\text{inv}} + (\chi_l^{\text{inv}} - \chi_k^{\text{inv}})(\chi_l^{\text{inv}} - \chi_k^{\text{inv}})^T \right] \right) \quad (\text{B.14e})$$

$$u_k^{\text{inv}}(0) = u_k^{\text{inv},0} \quad (\text{B.14f})$$

$$\chi_k^{\text{inv}}(0) = \chi_k^{\text{inv},0} \quad (\text{B.14g})$$

$$V_k^{\text{inv}}(0) = V_k^{\text{inv},0} \quad (\text{B.14h})$$

until they too are in equilibrium with $u_k^{\text{inv},*}$, $\chi_k^{\text{inv},*}$, and $V_k^{\text{inv},*}$ being the equilibrium densities, mean trait vectors, and variance-covariance matrices respectively.

Finally, we need to compute invasion fitness (i.e., the long-term exponential growth rate while rare) of the equilibrium invader. To this end, let $\tilde{F}^{\text{inv}}$ be the matrix that has $\tilde{F}_{kl}^{\text{inv}}$ as its elements, and u^{inv} be the vector that has u_k^{inv} as its elements. Examining the equations for densities we get that

$$0 = \tilde{F} u^{\text{inv},*} = \left(F^{\text{inv}} - (\mathbf{1}^T F^{\text{inv}} \tilde{u}^{\text{inv},*}) I \right) u^{\text{inv},*} \quad (\text{B.15a})$$

$$\iff F^{\text{inv}} u^{\text{inv},*} = \left(\mathbf{1}^T F^{\text{inv}} \tilde{u}^{\text{inv},*} \right) u^{\text{inv},*}, \quad (\text{B.15b})$$

where $\mathbf{1}$ is a $K \times 1$ vector having ones as all its elements and I is the $K \times K$ identity matrix.

From this we see that $\mathbf{1}^T F^{\text{inv}} \tilde{u}^{\text{inv}}$ is an eigenvalue of the matrix F^{inv} , and thus again in analogy with adaptive dynamics (see Caswell 2001 and Wickman et al. 2017 for structured populations), we can let

$$\lambda := \mathbf{1}^T F^{\text{inv}} \tilde{u}^{\text{inv}} \quad (\text{B.16})$$

be the invasion fitness of the mutant.

We can now determine the ultimate fate of the mutant by looking at the sign of λ . If $\lambda > 0$, invasion fitness is positive, and the invader will successfully invade the resident community. If $\lambda < 0$, invasion fitness is negative and the invader will fail to invade the community. If $\lambda = 0$, the invasion fitness is neutral, and the invader will fail to invade the community. This last case typically occurs when the invading mutant ends up having the same mean traits and variance-covariances as one of the residents. This allows us to create invasion diagrams for structured populations, tracking the paths in trait space of invading mutants, and determining whether they will ultimately be able to successfully invade a community of residents. The process is illustrated in Fig. 4 in the main text.

We wish to emphasize that the zero-variance (adaptive-dynamics) fitness landscape is, in and of itself, insufficient to determine the invasion success of invaders. Figure B.2 illustrates a scenario where the fitness landscape would indicate a two-species equilibrium, but where the finite-variance case of the moment equations yields only one.

Figure B.2: Illustration of a parameter set for the two-patch model (Box 1, main text) where the adaptive-dynamics fitness landscape does not predict qualitative features of the eco-evolutionarily stable community. (A–B) Trait-density distributions on patches *A* and *B* respectively in eco-evolutionary equilibrium of both the moment equations (black lines) and trait-space equations (blue areas). The moment equations capture the qualitative nature of the equilibrium, but is limited in its quantitative power due to the fact that the trait-space solution becomes platykurtic for this set of parameters. Comparing the moments on patch *A* we have $u_A^{\text{trait-space}} \approx 12.586$, $u_A^{\text{moment}} \approx 12.586$, $\chi_A^{\text{trait-space}} \approx 0.018627$, $\chi_A^{\text{moment}} \approx 0.01453$, $V_A^{\text{trait-space}} \approx 0.064175$, and $V_A^{\text{moment}} \approx 0.04995$. **(C–D)** Adaptive-dynamics fitness landscapes and invasion process on patches *A* and *B* respectively. The gray arrows illustrate the path of two invaders starting with near-zero variance at the fitness peaks and how their mean traits change over time as the variances and mean traits evolve over time for the invaders. The invaders end up with the same mean traits and variances as the resident and thus have neutral invasion fitness. The two-peaked fitness landscape implies that under zero-variance dynamics as in adaptive dynamics, the system would have a two-species eco-evolutionary equilibrium. Parameters are as in Fig. 3 in the main text except for $d = 1.8 \cdot 10^{-2}$.

Appendix C: Additional details of the example models

In this appendix we provide additional details of the example models presented in the main text, including details of how they were adapted from the generic equations, what numerical methods we used, and tables of parameter values used in numerical solutions.

Two-patch resource-competition model

The trait-space equations described in the main text are given by

$$\frac{dv_A(\xi)}{dt} = \int b(\eta, R_{1A}, R_{2A})v_A(\eta)\mathcal{N}(\xi, \eta, M)d\eta - \mu v_A(\xi) - dv_A(\xi) + dv_B(\xi) \quad (\text{C.1a})$$

$$\frac{dv_B(\xi)}{dt} = \int b(\eta, R_{1B}, R_{2B})v_B(\eta)\mathcal{N}(\xi, \eta, M)d\eta - \mu v_B(\xi) + dv_A(\xi) - dv_B(\xi) \quad (\text{C.1b})$$

$$\frac{dR_{jI}}{dt} = r(K_{jI} - R_{jI}) - \int U_j(\xi, R_{1I}, R_{2I})v_I(\xi)d\xi, \quad j = 1, 2, \quad I = A, B \quad (\text{C.1c})$$

$$U_j(\xi, R_1, R_2) = a_j(\xi)R_j \quad (\text{C.1d})$$

$$b(\xi, R_1, R_2) = c_1 a_1(\xi)R_1 + c_2 a_2(\xi)R_2 \quad (\text{C.1e})$$

$$a_1(\xi) = \left(\frac{e^\xi}{1 + e^\xi} \right)^\alpha, \quad (\text{C.1f})$$

$$a_2(\xi) = \left(1 - \frac{e^\xi}{1 + e^\xi} \right)^\alpha. \quad (\text{C.1g})$$

For our numerical solutions, especially for small mutation variances, the numerical solutions would require an inordinately high resolution in trait space, and so we approximate the mutation dynamics by a diffusion approximation (Kimura, 1965; Débarre et al., 2013)

$$\int b(\eta, R_{1A}, R_{2A})v_A(\eta)\mathcal{N}(\xi, \eta, M)d\eta \approx b(\xi, R_{1A}, R_{2A})v_A(\xi) + \frac{M}{2} \frac{\partial^2}{\partial \xi^2} (b(\xi, R_{1A}, R_{2A})v_A(\xi)) \quad (\text{C.2a})$$

$$\int b(\eta, R_{1B}, R_{2B})v_B(\eta)\mathcal{N}(\xi, \eta, M)d\eta \approx b(\xi, R_{1A}, R_{2A})v_B(\xi) + \frac{M}{2} \frac{\partial^2}{\partial \xi^2} (b(\xi, R_{1A}, R_{2A})v_B(\xi)) \quad (\text{C.2b})$$

We then discretize trait space into 512 equidistant points between $\xi = -5$ and $\xi = 5$, and approximate the second derivative by a finite-difference approximation. On the edges, $\xi = -5$ and $\xi = 5$ we assume reflective boundary conditions, but this assumption is highly unlikely to affect results as the trait-density distributions are always very close to zero near the edges.

To get all eco-evolutionarily stable solutions of the trait-space equations depicted in Fig. 5 we run the discretized trait-space equations for 10^7 time units for each combination of d and M using the `DifferentialEquations.jl` library (Rackauckas and Nie 2017) in Julia (Bezanson et al. 2017).

The moment equations were solved numerically using the same differential equation library. Comparisons between trait-space and moment equation solutions were made by evaluating the

moment equation normal distribution sum on the same grid on which the trait-space equations were solved, and the integrals in trait space were then approximated as

$$\int v(\xi)d\xi \approx \Delta\xi \sum_{k=1}^{512} v_k, \quad (\text{C.3})$$

where $\Delta\xi$ was the spacing between grid points and v_k the value of $v(\xi)$ evaluated at the k :th grid point.

To get the moment equations for the resources R_{jI} we use the generic equations for environmental variables derived in Appendix A and identify

Generic	Two-patch model	
$\rho_{1A}(\mathbf{E}) = r(K_{1A} - R_{1A})$		(C.4a)
$\rho_{1B}(\mathbf{E}) = r(K_{1B} - R_{1B})$		(C.4b)
$\rho_{2A}(\mathbf{E}) = r(K_{2A} - R_{2A})$		(C.4c)
$\rho_{2B}(\mathbf{E}) = r(K_{2B} - R_{2B})$		(C.4d)
$I_{1AA}(\xi, \mathbf{E}) = -U_1(\xi, R_{1A}, R_{2A})$		(C.4e)
$I_{1AB}(\xi, \mathbf{E}) = 0$		(C.4f)
$I_{1BA}(\xi, \mathbf{E}) = 0$		(C.4g)
$I_{1BB}(\xi, \mathbf{E}) = -U_1(\xi, R_{1B}, R_{2B})$		(C.4h)
$I_{2AA}(\xi, \mathbf{E}) = -U_2(\xi, R_{1A}, R_{2A})$		(C.4i)
$I_{2AB}(\xi, \mathbf{E}) = 0$		(C.4j)
$I_{2BA}(\xi, \mathbf{E}) = 0$		(C.4k)
$I_{2BB}(\xi, \mathbf{E}) = -U_2(\xi, R_{1B}, R_{2B})$		(C.4l)

Comparison with the generic moment equations including environmental variables (Eq. A.56d) then yields that the resource dynamics in the moment equations will be

$$\frac{dR_{jI}}{dt} = r(K_{jI} - r_{jI}) - \sum_{i=1}^S \hat{U}_{jIi} u_{iI}, \quad j = 1, 2, \quad I = A, B \quad (\text{C.5a})$$

$$\hat{U}_{jIi} = U_j(\chi_{iI}, R_{1I}, R_{2I}) + \frac{1}{2} \frac{\partial^2 U_j}{\partial \xi^2}(\chi_{iI}, R_{1I}, R_{2I}) V_{iI} \quad (\text{C.5b})$$

Figure C.1: Examples of evolutionary cycling and trait-density distributions for high mutation rates in the two-patch model. (A) Reproduction of panel D of Fig. 5 in the main text showing the relative error between trait-space and moment solutions at eco-evolutionary equilibrium. The three dots show where in parameter space the examples shown in panels B, C–D and E–F are taken from. (B) Example of the moment-equation assembly procedure not accurately reproducing the trait-density distribution of the trait-space equations. The equilibrium distribution of the trait-space equations (green) assumes a shape which is not easily represented as a sum of normal distributions and the moment equations (dotted black line), which here converge on a one-species equilibrium (blue area). Only patch A is depicted, as patch B has a near-identical, but reflected around $\xi = 0$, distribution. (C–D) Trait-space and moment solutions for high mutation variance and low dispersal on patches A and B respectively for panels B and C. Trait-space solutions are depicted as solid green lines, and moment solutions are depicted as dotted black lines. The two species that make up the moment solution are depicted as filled blue and orange areas. (E–F) Trait-space and moment solutions for intermediate mutation variance and low dispersal on patches A and B respectively for panels D and E. Trait-space solutions are depicted as solid green lines, and moment solutions are depicted as dotted black lines. The one species that make up the moment solution are depicted as filled blue areas.

Figure C.2: Examples of the two-patch model failing to converge to an eco-evolutionarily stable community. (A–B) Example of an invasion–merging cycle, where the densities (panel A) and mean traits (panel B) of the species on patch A are depicted. After settling onto a one-species equilibrium the resident species (blue) is invaded by a new species (orange). However, after successfully invading, the orange species evolves to the same mean trait as the blue species becoming equivalent. As the two species are equivalent, we merge them into one species and the cycle repeats. Parameter values: $d = 0.017450183677616502$, $M = 3.3451854991893273 \cdot 10^{-6}$. (C–D) Example of a branching–extinction cycle, where the densities (panel C) and mean traits (panel D) of the species on patch A are depicted. After settling onto a two-species equilibrium (blue and orange), one resident species (blue) undergoes a branching into two new species (blue and green). However, over time, one of the branches goes extinct (green) and is removed from the system, after which the cycle repeats. Parameter values: $d = 0.017406558218422463$, $M = 3.3451854991893273 \cdot 10^{-6}$. Other parameter values: $\mu = 0.1$, $r = 1$, $K_{1A} = 1$, $K_{1B} = 0.4$, $K_{2A} = 0.4$, $K_{2B} = 1$, $c_1 = 1$, $c_2 = 1$, and $\alpha = 0.5$.

Stage-structured resource-competition model

The trait-space equations, adapted from De Roos et al. (2007), read

$$\frac{dv_J(\xi, t)}{dt} = \int r_A(\eta)v_A(\eta)\mathcal{N}(\xi, \eta, M)d\eta + (r_J(\xi) - \mu_J(\xi))v_J(\xi) - \gamma(\xi)v_J(\xi) \quad (\text{C.6a})$$

$$\frac{dv_A(\xi, t)}{dt} = \gamma(\xi)v_J(\xi) - \mu_A(\xi)v_A(\xi) \quad (\text{C.6b})$$

$$\frac{dR_1(t)}{dt} = \rho_1(K_1 - R_1) - \int r_{J1}(\eta)v_J(\eta)d\eta - \int r_{A1}(\eta)v_A(\eta)d\eta \quad (\text{C.6c})$$

$$\frac{dR_2(t)}{dt} = \rho_2(K_2 - R_2) - \int r_{J2}(\eta)v_J(\eta)d\eta - \int r_{A2}(\eta)v_A(\eta)d\eta \quad (\text{C.6d})$$

We can now compare these to the generic equations 2 in the main text, and identify the rate functions as

$$f_{JJ1} = r_J \quad f_{JJ2} = -\mu_J \quad f_{JJ3} = -\gamma \quad (\text{C.7a})$$

$$f_{JA1} = r_A \quad (\text{C.7b})$$

$$f_{AJ1} = \gamma \quad (\text{C.7c})$$

$$f_{AA1} = -\mu_A \quad (\text{C.7d})$$

Mutations only occur in the transition from adults to juveniles, so $M_{JA} = M$ and the other mutation matrices are zero.

The necessary ingredients for the moment equations are now available, and we can insert the rate functions into Eqs. 5 in the main text to get the moment equations for the stage-structured model. Before we can write down the moment equations there we also need to identify what the renewal terms and resource impacts of the species will be. To that end, we compare with the trait-space equations including resources (Eqs.A.52) and identify

$$\begin{array}{cc} \text{Generic} & \text{Stage-structured model} \end{array} \quad (\text{C.8a})$$

$$\rho_j(\mathbf{E}) = \rho_j(K_j - R_j), \quad j = 1, 2 \quad (\text{C.8b})$$

$$I_{jl}(\xi, \mathbf{E}) = -r_{lj}(\xi, R_1, R_2), \quad j = 1, 2, \quad l = J, A \quad (\text{C.8c})$$

The moment equations for the resources will thus read

$$\frac{dR_1}{dt} = \rho_1(K_1 - R_1) - \sum_{i=1}^S \hat{r}_{Ji1}u_{Ji} - \sum_{i=1}^S \hat{r}_{Ai1}u_{Ai} \quad (\text{C.9a})$$

$$\frac{dR_2}{dt} = \rho_2(K_2 - R_2) - \sum_{i=1}^S \hat{r}_{Ji2}u_{Ji} - \sum_{i=1}^S \hat{r}_{Ai2}u_{Ai} \quad (\text{C.9b})$$

$$\hat{r}_{Iij} = r_{Ij}(\chi_{iI}, R_1, R_2) + \frac{1}{2} \nabla_{\xi}^2 r_{Ij}(\chi_{iI}, R_1, R_2) : V_{iI}, \quad i = 1, 2, \dots, S, \quad j = 1, 2, \quad I = J, A \quad (\text{C.9c})$$

The resource uptake and stage-transition functions are given by

$$r_I(\xi) = q_I \frac{a_{I1}(\xi)R_1 + a_{I2}(\xi)R_2}{1 + a_{I1}(\xi)R_1 + a_{I2}(\xi)R_2}, \quad (\text{C.10a})$$

$$r_{Ij}(\xi) = q_I \frac{a_{Ij}(\xi) R_j}{1 + a_{I1}(\xi) R_1 + a_{I2}(\xi) R_2} \quad I \in \{J, A\}, j \in \{1, 2\}, \quad (\text{C.10b})$$

$$\gamma = \frac{r_J(\xi) - \mu_J(\xi)}{1 - z^{1 - \mu_J(\xi)/r_J(\xi)}}. \quad (\text{C.10c})$$

The maturation rate γ depends on the growth $r_J(\xi)$, the mortality $\mu_J(\xi)$, and the size ratio between juveniles and adults z , see [De Roos et al. \(2007\)](#) for an explanation of how the specific form comes about. Here, $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$, where ξ_1 parameterizes a trade-off between the resource affinities and ξ_2 parameterizes a trade-off between juvenile and adult mortality. Specifically,

$$\theta_1 = \frac{e^{\xi_1}}{1 + e^{\xi_1}}, \quad \theta_2 = \frac{e^{\xi_2}}{1 + e^{\xi_2}}, \quad (\text{C.11a})$$

$$a_{J1} = \theta_1^{\alpha_J}, \quad a_{J2} = (1 - \theta_1)^{\alpha_J}, \quad (\text{C.11b})$$

$$a_{A1} = \theta_1^{\alpha_A}, \quad a_{A2} = (1 - \theta_1)^{\alpha_A}, \quad (\text{C.11c})$$

$$\mu_J = \mu_{0J} \theta_2^{\alpha_\mu}, \quad \mu_A = \mu_{0A} (1 - \theta_2)^{\alpha_\mu}. \quad (\text{C.11d})$$

See Fig. C.3 for illustrations using the parameter values used in the numerical example. The parameters used in the numerical solutions are given in Table C.1.

Figure C.3: 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. Equations for the trade-offs are given in Appendix C.

Solving the trait-space equations 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 approximate the birth/mutation integral with a diffusion term in the following way ([Kimura, 1965](#); [Débarre et al., 2013](#)):

$$\int r_A(\eta) v_A(\eta) \mathcal{N}(\xi, \eta, M) d\eta \approx r_A(\xi) v_A(\xi) + \frac{M_{11}}{2} \Delta_\xi (r_A(\xi) v_A(\xi)), \quad (\text{C.12})$$

Symbol	Definition/Description	Value
ρ_1	Resource 1 renewal rate	1.0
ρ_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
μ_{0J}	Juvenile base mortality	0.1
μ_{0A}	Adult base mortality	0.1
α_J	Juvenile resource trade-off shape parameter	1.3
α_A	Adult resource trade-off shape parameter	0.7
α_μ	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 C.1: Parameter values used for numerical solutions of the stage-structured example.

where $\Delta_\xi = \frac{\partial^2}{\partial \xi_1^2} + \frac{\partial^2}{\partial \xi_2^2}$ is the Laplacian in trait space. We discretize two-dimensional trait space into an equispaced grid of 256×256 points with $\xi_1, \xi_2 \in [-4, 4]$, and approximate the Laplacian with a central-difference approximation. We then used the `DifferentialEquations.jl` library (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.

Literature Cited

- Adler, P.B., Seabloom, E.W., Borer, E.T., Hillebrand, H., Hautier, Y., Hector, A., Harpole, W.S., O’Halloran, L.R., Grace, J.B., Anderson, T.M., et al., 2011. Productivity is a poor predictor of plant species richness. *Science* 333, 1750–1753. doi:[10.1126/science.1204498](https://doi.org/10.1126/science.1204498).
- Albert, C.H., Thuiller, W., Yoccoz, N.G., Soudant, A., Boucher, F., Saccone, P., Lavorel, S., 2010. Intraspecific functional variability: extent, structure and sources of variation. *Journal of Ecology* 98, 604–613. doi:[10.1111/j.1365-2745.2010.01651.x](https://doi.org/10.1111/j.1365-2745.2010.01651.x).
- Barabás, G., Parent, C., Kraemer, A., Van de Perre, F., De Laender, F., 2020. The evolution of trait variance creates a tension between species diversity and functional diversity. *bioRxiv* doi:[10.1101/2020.07.01.182261](https://doi.org/10.1101/2020.07.01.182261).
- Barabás, G., D’Andrea, R., 2016. The effect of intraspecific variation and heritability on community pattern and robustness. *Ecology Letters* 19, 977–986. doi:[10.1111/ele.12636](https://doi.org/10.1111/ele.12636).
- Bezanson, J., Edelman, A., Karpinski, S., Shah, V.B., 2017. Julia: A fresh approach to numerical computing. *SIAM review* 59, 65–98. doi:[10.1137/141000671](https://doi.org/10.1137/141000671).
- Brännström, Å., Johansson, J., von Festenberg, N., 2013. The hitchhiker’s guide to adaptive dynamics. *Games* 4, 304–328. doi:[10.3390/g4030304](https://doi.org/10.3390/g4030304).

Figure C.4: 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 ξ_1 and ξ_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.

- Bruggeman, J., 2009. Succession in plankton communities: A trait-based perspective. Ph.D. thesis. Vrije Universiteit. Amsterdam. URL: <https://research.vu.nl/en/publications/succession-in-plankton-communities-a-trait-based-perspective>.
- Bürger, R., 1986. On the maintenance of genetic variation: global analysis of kimura's continuum-of-alleles model. *Journal of Mathematical Biology* 24, 341–351. doi:[10.1007/BF00275642](https://doi.org/10.1007/BF00275642).
- Cardinale, B.J., Matulich, K.L., Hooper, D.U., Byrnes, J.E., Duffy, E., Gamfeldt, L., Balvanera, P., O'Connor, M.I., Gonzalez, A., 2011. The functional role of producer diversity in ecosystems. *American Journal of Botany* 98, 572–592. doi:[10.3732/ajb.1000364](https://doi.org/10.3732/ajb.1000364).
- Caswell, H., 2001. *Matrix Population Models*. 2nd ed., Sinauer Associates, Sunderland, MA.
- Chase, J.M., Leibold, M.A., 2003. *Ecological niches*. University of Chicago Press, Chicago, IL.
- Cropp, R., Norbury, J., 2021. Modelling the evolution of naturally bounded traits in a population. *Theoretical Ecology* doi:[10.1007/s12080-020-00495-2](https://doi.org/10.1007/s12080-020-00495-2).
- Cusens, J., Wright, S.D., McBride, P.D., Gillman, L.N., 2012. What is the form of the productivity–animal-species-richness relationship? A critical review and meta-analysis. *Ecology* 93, 2241–2252. doi:[10.1890/11-1861.1](https://doi.org/10.1890/11-1861.1).
- De Roos, A., Schellekens, T., Kooten, T., Wolfshaar, K., Claessen, D., Persson, L., 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. doi:[10.1086/520119](https://doi.org/10.1086/520119).
- Débarre, F., Nuismer, S.L., Doebeli, M., 2014. Multidimensional (co)evolutionary stability. *The American Naturalist* 184, 158–171. doi:[10.1086/677137](https://doi.org/10.1086/677137).
- Débarre, F., Ronce, O., Gandon, S., 2013. Quantifying the effects of migration and mutation on adaptation and demography in spatially heterogeneous environments. *Journal of Evolutionary Biology* 26, 1185–1202. doi:[10.1111/jeb.12132](https://doi.org/10.1111/jeb.12132).
- Dercole, F., Rinaldi, S., 2008. *Analysis of evolutionary processes: the adaptive dynamics approach and its applications*. Princeton Series in Theoretical and Computational Biology, Princeton University Press, Princeton, NJ.
- Dieckmann, U., Law, R., 1996. The dynamical theory of coevolution: a derivation from stochastic ecological processes. *Journal of Mathematical Biology* 34, 579–612. doi:[10.1007/bf02409751](https://doi.org/10.1007/bf02409751).
- Edwards, K.F., Kremer, C.T., Miller, E.T., Osmond, M.M., Litchman, E., Klausmeier, C.A., 2018. Evolutionarily stable communities: a framework for understanding the role of trait evolution in the maintenance of diversity. *Ecology Letters* 21, 1853–1868. doi:[10.1111/ele.13142](https://doi.org/10.1111/ele.13142).
- Fontana, S., Petchey, O.L., Pomati, F., 2016. Individual-level trait diversity concepts and indices to comprehensively describe community change in multidimensional trait space. *Functional Ecology* 30, 808–818. doi:[10.1111/1365-2435.12551](https://doi.org/10.1111/1365-2435.12551).

- Geritz, S.A.H., Kisdi, E., Meszéna, G., Metz, J.A.J., 1998. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology* 12, 35–57. doi:[10.1023/a:1006554906681](https://doi.org/10.1023/a:1006554906681).
- Geritz, S.A.H., van der Meijden, E., Metz, J.A.J., 1999. Evolutionary dynamics of seed size and seedling competitive ability. *Theoretical Population Biology* 55, 324–343. doi:[10.1006/tpbi.1998.1409](https://doi.org/10.1006/tpbi.1998.1409).
- 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. doi:[10.1073/pnas.54.3.731](https://doi.org/10.1073/pnas.54.3.731).
- Kirkpatrick, M., Barton, N.H., 1997. Evolution of a species' range. *The American Naturalist* 150, 1–23. doi:[10.1086/286054](https://doi.org/10.1086/286054).
- Klauschies, T., Coutinho, R.M., Gaedke, U., 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. doi:[10.1016/j.ecolmodel.2018.02.001](https://doi.org/10.1016/j.ecolmodel.2018.02.001).
- Klausmeier, C.A., Kremer, C.T., Koffel, T., 2020. Trait-based ecological and eco-evolutionary theory, in: McCann, K.S., Gellner, G. (Eds.), *Theoretical Ecology: concepts and applications*. Oxford University Press, Oxford, pp. 161–194. doi:[10.1093/oso/9780198824282.001.0001](https://doi.org/10.1093/oso/9780198824282.001.0001).
- Koffel, T., Daufresne, T., Klausmeier, C.A., 2021. From competition to facilitation and mutualism: a general theory of the niche. *Ecological Monographs* 91, e01458. doi:doi.org/10.1002/ecm.1458.
- Kraft, N.J.B., Crutsinger, G.M., Forrestel, E.J., Emery, N.C., 2014. Functional trait differences and the outcome of community assembly: an experimental test with vernal pool annual plants. *Oikos* 123, 1391–1399. doi:[10.1111/oik.01311](https://doi.org/10.1111/oik.01311).
- Kremer, C.T., Klausmeier, C.A., 2013. Coexistence in a variable environment: Eco-evolutionary perspectives. *Journal of Theoretical Biology* 339, 14–25. doi:[10.1016/j.jtbi.2013.05.005](https://doi.org/10.1016/j.jtbi.2013.05.005).
- Lande, R., 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: Body size allometry. *Evolution* 33, 402–416. doi:[10.1111/j.1558-5646.1979.tb04694.x](https://doi.org/10.1111/j.1558-5646.1979.tb04694.x).
- Lande, R., Arnold, S.J., 1983. The measurement of selection on correlated characters. *Evolution* 37, 1210–1226. doi:[10.1111/j.1558-5646.1983.tb00236.x](https://doi.org/10.1111/j.1558-5646.1983.tb00236.x).
- Le gland, G., Vallina, S.M., Smith, S.L., Cermeño, P., 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. doi:[10.5194/gmd-2020-302](https://doi.org/10.5194/gmd-2020-302).
- Levins, R., 1962. Theory of fitness in a heterogeneous environment. i. the fitness set and adaptive function. *The American Naturalist* 96, 361–373. doi:[10.1086/282245](https://doi.org/10.1086/282245).
- MacArthur, R., Levins, R., 1964. Competition, habitat selection, and character displacement in a patchy environment. *Proceedings of the National Academy of Sciences, U.S.A.* 51, 1207–1210. doi:[10.1073/pnas.51.6.1207](https://doi.org/10.1073/pnas.51.6.1207).

- Merico, A., Brandt, G., Smith, S.L., Oliver, M., 2014. Sustaining diversity in trait-based models of phytoplankton communities. *Frontiers in Ecology and Evolution* 2, 59. doi:[10.3389/fevo.2014.00059](https://doi.org/10.3389/fevo.2014.00059).
- Metz, J.A.J., Nisbet, R.M., Geritz, S.A.H., 1992. How should we define 'fitness' for general ecological scenarios? *Trends in Ecology & Evolution* 7, 198–202. doi:[10.1016/0169-5347\(92\)90073-k](https://doi.org/10.1016/0169-5347(92)90073-k).
- Norberg, J., Swaney, D.P., Dushoff, J., Lin, J., Casagrandi, R., Levin, S.A., 2001. Phenotypic diversity and ecosystem functioning in changing environments: A theoretical framework. *Proceedings of the National Academy of Sciences* 98, 11376–11381. doi:[10.1073/pnas.171315998](https://doi.org/10.1073/pnas.171315998).
- Norberg, J., Urban, M.C., Vellend, M., Klausmeier, C.A., Loeuille, N., 2012. Eco-evolutionary responses of biodiversity to climate change. *Nature Climate Change* 2, 747–751. doi:[10.1038/nclimate1588](https://doi.org/10.1038/nclimate1588).
- Nordbotten, J.M., Bokma, F., Hermansen, J.S., Stenseth, N.C., 2020. The dynamics of trait variance in multi-species communities. *Royal Society Open Science* 7, 200321. doi:[10.1098/rsos.200321](https://doi.org/10.1098/rsos.200321).
- Owen, D.B., 1980. A table of normal integrals. *Communications in Statistics - Simulation and Computation* 9, 389–419. doi:[10.1080/03610918008812164](https://doi.org/10.1080/03610918008812164).
- Paine, C.E.T., Baraloto, C., Chave, J., Hérault, B., 2011. Functional traits of individual trees reveal ecological constraints on community assembly in tropical rain forests. *Oikos* 120, 720–727. doi:[10.1111/j.1600-0706.2010.19110.x](https://doi.org/10.1111/j.1600-0706.2010.19110.x).
- Peeters, F., Straile, D., 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. doi:[10.1371/journal.pone.0194076](https://doi.org/10.1371/journal.pone.0194076).
- Rackauckas, C., Nie, Q., 2017. Differentialequations.jl—a performant and feature-rich ecosystem for solving differential equations in julia. *Journal of Open Research Software* 5. doi:[10.5334/jors.151](https://doi.org/10.5334/jors.151).
- Rueffler, C., Van Dooren, T.J.M., Metz, J.A.J., 2006. The evolution of resource specialization through frequency-dependent and frequency-independent mechanisms. *The American Naturalist* 167, 81–93. doi:[10.1086/498275](https://doi.org/10.1086/498275).
- Sasaki, A., Dieckmann, U., 2011. Oligomorphic dynamics for analyzing the quantitative genetics of adaptive speciation. *Journal of Mathematical Biology* 63, 601–635. doi:[10.1007/s00285-010-0380-6](https://doi.org/10.1007/s00285-010-0380-6).
- Savage, V.M., Webb, C.T., Norberg, J., 2007. A general multi-trait-based framework for studying the effects of biodiversity on ecosystem functioning. *Journal of Theoretical Biology* 247, 213 – 229. doi:[10.1016/j.jtbi.2007.03.007](https://doi.org/10.1016/j.jtbi.2007.03.007).
- Schleuter, D., Daufresne, M., Massol, F., Argillier, C., 2010. A user's guide to functional diversity indices. *Ecological Monographs* 80, 469–484. doi:[10.1890/08-2225.1](https://doi.org/10.1890/08-2225.1).

- Siefert, A., Violle, C., Chalmandrier, L., Albert, C.H., Taudiere, A., Fajardo, A., Aarssen, L.W., Baraloto, C., Carlucci, M.B., Cianciaruso, M.V., de L. Dantas, V., de Bello, F., Duarte, L.D.S., Fonseca, C.R., Freschet, G.T., Gaucherand, S., Gross, N., Hikosaka, K., Jackson, B., Jung, V., Kamiyama, C., Katabuchi, M., Kembel, S.W., Kichenin, E., Kraft, N.J.B., Lagerström, A., Bagousse-Pinguet, Y.L., Li, Y., Mason, N., Messier, J., Nakashizuka, T., Overton, J.M., Peltzer, D.A., Pérez-Ramos, I.M., Pillar, V.D., Prentice, H.C., Richardson, S., Sasaki, T., Schamp, B.S., Schöb, C., Shipley, B., Sundqvist, M., Sykes, M.T., Vandewalle, M., Wardle, D.A., 2015. A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. *Ecology Letters* 18, 1406–1419. doi:[10.1111/ele.12508](https://doi.org/10.1111/ele.12508).
- Tilman, D., 1982. *Resource Competition and Community Structure*. Monographs in population biology, Princeton University Press, Princeton, NJ.
- Triantafyllopoulos, K., 2002. Moments and cumulants of the multivariate real and complex gaussian distributions. Department of Mathematics, University of Bristol .
- Violle, C., Enquist, B.J., McGill, B.J., Jiang, L., Albert, C.H., Hulshof, C., Jung, V., Messier, J., 2012. The return of the variance: intraspecific variability in community ecology. *Trends in Ecology & Evolution* 27, 244–252. doi:[10.1016/j.tree.2011.11.014](https://doi.org/10.1016/j.tree.2011.11.014).
- Wickman, J., Diehl, S., Blasius, B., Klausmeier, C.A., Ryabov, A., Brännström, Å., 2017. Determining selection across heterogeneous landscapes: a perturbation-based method and its application to modeling evolution in space. *The American Naturalist* 189, 381–395. doi:[10.1086/690908](https://doi.org/10.1086/690908).
- Wickman, J., Diehl, S., Brännström, Å., 2019. Evolution of resource specialisation in competitive metacommunities. *Ecology Letters* 22, 1746–1756. doi:[10.1111/ele.13338](https://doi.org/10.1111/ele.13338).
- Wirtz, K.W., Eckhardt, B., 1996. Effective variables in ecosystem models with an application to phytoplankton succession. *Ecological Modelling* 92, 33–53. doi:[10.1016/0304-3800\(95\)00196-4](https://doi.org/10.1016/0304-3800(95)00196-4).