The effect of intraspecific variation and heritability on community pattern and robustness

Supporting Information

György Barabás & Rafael D’Andrea

Contents

1 Deriving the basic quantitative genetic model 2
  1.1 The model in discrete time ................................................. 2
  1.2 Continuous-time dynamics in the weak selection limit .................... 4

2 A quantitative genetic Lotka–Volterra model 5
  2.1 Dynamical equations .......................................................... 5
  2.2 The Jacobian of the quantitative genetic Lotka–Volterra model .......... 6
  2.3 Parameterizing the quantitative genetic Lotka–Volterra model ............ 7

3 Two-species results 9
  3.1 The ecological equilibrium and its stability .................................. 9
  3.2 Robustness and feasibility .................................................... 10
  3.3 The eco-evolutionary dynamics ............................................... 14
  3.4 Robustness of the two-species results to an asymmetry in intrinsic growth 18

4 Multispecies results 21
  4.1 Multiple species with equal intraspecific variances and no heritability .... 21
  4.2 Measuring community robustness ............................................. 22
  4.3 The effect of intraspecific variation and heritability on species richness, trait pattern, and community robustness ..................... 23
  4.4 Trait convergence in multispecies communities ............................. 27

5 Relaxing the assumptions of the extreme quantitative genetic limit: the Shpak-Kondrashov model 28
  5.1 Community model with no environmental effect on phenotype ............ 29
  5.2 Community model with environmental effects on phenotype ............... 30

6 Competition along two independent trait dimensions 32

7 Code 34
1 Deriving the basic quantitative genetic model

1.1 The model in discrete time

We model the dynamics of \( S \) species differing in some ecologically relevant trait \( z \). Each individual can be described by its trait, but each species is comprised of individuals with several different trait values. The number of individuals within species \( i \) at time \( t \) is \( N_i(t) \), while the distribution of traits within the species is given by the function \( p_i(z, t) \). By definition, this function satisfies

\[
\int p_i(z, t) \, dz = 1 \tag{1.1}
\]

at every moment of time \( t \); the limits of integration encompass the whole trait axis, which for simplicity we take to go between minus and plus infinity unless otherwise noted. \( N_i(t)p_i(z, t) \, dz \) is then the population density of species \( i \)'s individuals with phenotype value between \( z \) and \( z + dz \).


To do this, we first assume species with nonverlapping generations undergoing selection and then reproduction:

\[
N_i(t)p_i(z, t) \xrightarrow{\text{selection}} N'_i(t)p'_i(z, t) \xrightarrow{\text{reproduction}} N_i(t + T_i)p_i(z, t + T_i), \tag{1.2}
\]

where the prime denotes state after selection but before reproduction, \( T_i \) is the generation time of species \( i \), and \( W_i(\vec{N}, \vec{p}, z, t) \) is the absolute fitness of species \( i \)'s phenotype \( z \) at time \( t \):

\[
N'_i(t)p'_i(z, t) = W_i(\vec{N}, \vec{p}, z, t)N_i(t)p_i(z, t). \tag{1.3}
\]

Note that \( W_i(\vec{N}, \vec{p}, z, t) \) will in general depend on the abundances and trait distributions of all interacting species (denoted by the vectors \( \vec{N} \) and \( \vec{p} \)). Here we apply the standard sleight of hand of assuming that \( W_i(\vec{N}, \vec{p}, z, t) \) describes both birth and death processes. In principle, the population first undergoes viability selection and then reproduction as in Eq. (1.2). Having \( W_i(\vec{N}, \vec{p}, z, t) \) include both birth and death, it is tacitly assumed that both processes happen relative to the original \( N_i(t)p_i(z, t) \), instead of births happening relative to the primed quantities. This also means that the only role of the reproduction phase is to change the trait distributions \( p_i(z, t) \) but not the population densities: \( N'_i(t) = N_i(t + T_i) \). However, the error made this way is negligible if the difference between \( N_i(t)p_i(z, t) \) and \( N'_i(t)p'_i(z, t) \) is small, which will be the case if selection is sufficiently weak. We will therefore assume we are in the weak selection limit (Bürger 2011), i.e.,

\[
W_i(\vec{N}, \vec{p}, z, t) = 1 + s r_i(\vec{N}, \vec{p}, z, t), \tag{1.4}
\]

where \( s \ll 1 \) is a small parameter with units of inverse time, and the function \( r_i(\vec{N}, \vec{p}, z, t) \) is the rate of births minus deaths at time \( t \). It is therefore the per capita growth rate of species \( i \), determined by the ecological interactions within the community.

We work in the quantitative genetic limit, i.e., the trait in question is determined by the action of many independent loci, each contributing a small additive effect to the trait. In this case, the following important results hold (Bulmer 1980, Falconer 1981):

- The trait distribution is always normal, so one can write

\[
p_i(z, t) = \frac{1}{2\pi\sigma_i^2} \exp\left(-\frac{(z - \mu_i(t))^2}{2\sigma_i^2}\right), \tag{1.5}
\]

where \( \mu_i(t) \) is the mean trait of species \( i \) at time \( t \), and \( \sigma_i^2 \) is its intraspecific trait variance.
We can obtain where we switched to To obtain the time evolution of the trait distributions, note that only their means can change in $N$ Using $h$ is equal to the heritability times the selection differential: according to which the change in mean trait from one generation to the next (response to selection) the quantitative genetics approximation, as discussed above. Therefore we only need to track population densities from one generation to the next, we integrate Eq. (1.3) over $z$ to get

$$N_i(t) \int p_i'(z,t) \, dz = N_i(t) \int W_i(\vec{N}, \vec{p}, z, t) p_i(z,t) \, dz. \quad (1.6)$$

Using $N_i'(t) = N_i(t + T_i)$ and Eq. (1.1) on the left hand side:

$$N_i(t + T_i) = N_i(t) \int W_i(\vec{N}, \vec{p}, z, t) p_i(z,t) \, dz. \quad (1.7)$$

To obtain the time evolution of the trait distributions, note that only their means can change in the quantitative genetics approximation, as discussed above. Therefore we only need to track $\mu_i(t)$ instead of $p_i(z,t)$. By definition, the mean trait at time $t$ is given by

$$\mu_i(t) = \int z p_i(z,t) \, dz. \quad (1.8)$$

The equation for the trait means can be written using the breeder’s equation (Falconer 1981), according to which the change in mean trait from one generation to the next (response to selection) is equal to the heritability times the selection differential:

$$\mu_i(t + T_i) - \mu_i(t) = h_i^2 [\mu_i'(t) - \mu_i(t)]. \quad (1.9)$$

Here $h_i^2$ is the heritability, and $\mu_i'(t)$ is the mean of the trait after selection but before reproduction. We can obtain $\mu_i'(t)$ by first rearranging Eq. (1.3) and using $N_i'(t) = N_i(t + T_i)$:

$$p_i'(z,t) = \frac{W_i(\vec{N}, \vec{p}, z, t) N_i(t) p_i(z,t)}{N_i(t + T_i)}. \quad (1.10)$$

Using Eq. (1.7) in the denominator:

$$p_i'(z,t) = \frac{W_i(\vec{N}, \vec{p}, z, t) N_i(t) p_i(z,t)}{\int W_i(\vec{N}, \vec{p}, z', t) p_i(z',t) \, dz'} = \frac{W_i(\vec{N}, \vec{p}, z, t) p_i(z,t)}{\int W_i(\vec{N}, \vec{p}, z', t) p_i(z',t) \, dz'}, \quad (1.11)$$

where we switched to $z'$ for the integration variable to distinguish it from $z$ in the numerator. Multiplying both sides of Eq. (1.11) by $z$ and integrating, we get

$$\int z p_i'(z,t) \, dz = \frac{\int z W_i(\vec{N}, \vec{p}, z, t) p_i(z,t) \, dz}{\int W_i(\vec{N}, \vec{p}, z', t) p_i(z',t) \, dz'}, \quad (1.12)$$
or, using Eq. (1.8) on the left hand side,

$$\nu_i'(t) = \frac{\int zW_i(\tilde{N}, \tilde{p}, z, t)p_i(z, t) \, dz}{\int W_i(\tilde{N}, \tilde{p}, z', t)p_i(z', t) \, dz'}.$$  \hspace{1cm}(1.13)

Substituting this into the breeder’s equation Eq. (1.9), we finally have

$$\mu_i(t + T_i) - \mu_i(t) = h_i^2 \left( \frac{\int zW_i(\tilde{N}, \tilde{p}, z, t)p_i(z, t) \, dz}{\int W_i(\tilde{N}, \tilde{p}, z', t)p_i(z', t) \, dz'} - \mu_i(t) \right). \hspace{1cm}(1.14)$$

### 1.2 Continuous-time dynamics in the weak selection limit

We can obtain a differential equation approximation to Eqs. (1.7) and (1.14) in the weak selection limit of Eq. (1.4). Starting with Eq. (1.7), it reads

$$N_i(t + T_i) = N_i(t) \int [1 + s r_i(\tilde{N}, \tilde{p}, z, t)]p_i(z, t) \, dz,$$  \hspace{1cm}(1.15)

or, using Eq. (1.1) on the right hand side,

$$N_i(t + T_i) = N_i(t) + sN_i(t) \int r_i(\tilde{N}, \tilde{p}, z, t)p_i(z, t) \, dz.$$  \hspace{1cm}(1.16)

Rearranging, we get

$$\frac{N_i(t + T_i) - N_i(t)}{s} = N_i(t) \int r_i(\tilde{N}, \tilde{p}, z, t)p_i(z, t) \, dz.$$  \hspace{1cm}(1.17)

We define the dimensionless measure of time $\tau = ts$ and the new density measure $\tilde{N}_i(\tau) = N_i(t)$ (Nagylaki 1992, p. 99; Bürger 2011). The left hand side above then reads $[\tilde{N}_i(\tau + T\tau) - \tilde{N}_i(\tau)]/s$, which in the limit of $s \to 0$ converges to $d\tilde{N}_i(\tau)/d\tau$. Understanding that this change of units has taken place but reverting to our original symbols $t$ and $N_i(t)$ for notational convenience, Eq. (1.17) finally becomes

$$\frac{dN_i(t)}{dt} = N_i(t) \int r_i(\tilde{N}, \tilde{p}, z, t)p_i(z, t) \, dz.$$  \hspace{1cm}(1.18)

Turning to the trait means, Eq. (1.14) in the weak selection limit of Eq. (1.4) reads

$$\mu_i(t + T_i) - \mu_i(t) = h_i^2 \left( \frac{\int z \left[ 1 + s r_i(\tilde{N}, \tilde{p}, z, t) \right]p_i(z, t) \, dz}{\int \left[ 1 + s r_i(\tilde{N}, \tilde{p}, z', t) \right]p_i(z', t) \, dz'} - \mu_i(t) \right).$$  \hspace{1cm}(1.19)

Using Eqs. (1.1) and (1.8) on the right hand side:

$$\mu_i(t + T_i) - \mu_i(t) = h_i^2 \left( \frac{\mu_i(t) + s \int zr_i(\tilde{N}, \tilde{p}, z, t)p_i(z, t) \, dz}{1 + s \int r_i(\tilde{N}, \tilde{p}, z', t)p_i(z', t) \, dz'} - \mu_i(t) \right).$$  \hspace{1cm}(1.20)

Taylor expanding in $s \ll 1$ and neglecting terms of order $s^2$ or higher:

$$\mu_i(t + T_i) - \mu_i(t) \approx h_i^2 \left( s \int zr_i(\tilde{N}, \tilde{p}, z, t)p_i(z, t) \, dz - \mu_i(t)s \int r_i(\tilde{N}, \tilde{p}, z, t)p_i(z, t) \, dz \right).$$  \hspace{1cm}(1.21)
Rearranging, we get
\[
\frac{\mu_i(t + T_i) - \mu_i(t)}{s} = h_i^2 \int (z - \mu_i(t)) r_i(\vec{N}, \vec{p}, z, t)p_i(z, t) \, dz,
\]
(1.22)
at which point we can rescale time and take the \( s \to 0 \) limit exactly as before, obtaining
\[
\frac{d\mu_i(t)}{dt} = h_i^2 \int (z - \mu_i(t)) r_i(\vec{N}, \vec{p}, z, t)p_i(z, t) \, dz.
\]
(1.23)
Eqs. (1.18) and (1.23) govern the eco-evolutionary dynamics of the community given the ecology of the system, specified by the function \( r_i(\vec{N}, \vec{p}, z, t) \).

Although we obtained the governing differential equations in the weak selection limit, note that there are alternative ways of arriving at Eqs. (1.18) and (1.23) which do not rely on this approximation—see, e.g., Lande (1982) for a derivation which assumes that the populations are at their stable age structure. The results are therefore expected to hold even when selection is not very weak.

2 A quantitative genetic Lotka–Volterra model

2.1 Dynamical equations

The Lotka–Volterra model for \( S \) species is defined by the per capita growth rates
\[
r(\vec{N}, \vec{p}, z, t) = b(z) - \sum_{j=1}^{S} N_j(t) \int a(z, z') p_j(z', t) \, dz',
\]
(2.1)
where \( b(z) \) and \( a(z, z') \) are appropriately chosen functions for the intrinsic growth rate of phenotype \( z \) and the interaction coefficient between phenotypes \( z \) and \( z' \), respectively. The summation over species and the integral across all phenotypes means that every individual with phenotype \( z \) interacts with all other individuals in the community. This particular formulation of the Lotka–Volterra model is special in the sense that growth and interactions only depend on the phenotype \( z \) but not on species identity. In other words, two individuals of two different species are exchangeable without altering the growth rates. Note that this is only true in the context of the interaction structure defined by Eq. (2.1): species identity does matter in the context of reproduction in the quantitative genetic limit, where the shape of the trait distributions must be maintained to match Eq. (1.5).

We can substitute Eq. (2.1) into Eqs. (1.18) and (1.23) to obtain the dynamics. Starting with Eq. (1.18), we have
\[
\frac{dN_i(t)}{dt} = N_i(t) \int r_i(\vec{N}, \vec{p}, z, t)p_i(z, t) \, dz
\]
\[
= N_i(t) \int p_i(z, t) \left( b(z) - \sum_{j=1}^{S} N_j(t) \int a(z, z') p_j(z', t) \, dz' \right) \, dz.
\]
(2.2)
Introducing the notation
\[
b_i(t) = \int p_i(z, t)b(z) \, dz,
\]
(2.3)
\[
\alpha_{ij}(t) = \iint p_i(z, t)a(z, z')p_j(z', t) \, dz' \, dz,
\]
(2.4)
we can write
\[
\frac{dN_i(t)}{dt} = N_i(t) \left( \int_{b_i(t)}^{p_i(z, t)b(z) \, dz} - \sum_{j=1}^{S} N_j(t) \int_{\alpha_{ij}(t)}^{p_i(z, t)a(z, z', t) \, dz'} \right)
\]
\[
= N_i(t) \left( b_i(t) - \sum_{j=1}^{S} \alpha_{ij}(t)N_j(t) \right).
\]
(2.5)

The interpretation of \(b_i(t)\) is the population-level intrinsic growth rate of species \(i\), while \(\alpha_{ij}(t)\) is the population-level competitive effect of species \(j\) on species \(i\).

Similarly, Eq. (1.23) will read
\[
\frac{d\mu_i(t)}{dt} = h_i^2 \int (z - \mu_i(t))b(z) \, dz
\]
\[
= h_i^2 \int (z - \mu_i(t))p_i(z, t) \left( b(z) - \sum_{j=1}^{S} N_j(t) \int a(z, z')p_j(z', t) \, dz' \right) \, dz.
\]
(2.6)

Again, after introducing
\[
\overline{b}_i(t) = \int (z - \mu_i(t))p_i(z, t) \, dz,
\]
\[
\beta_{ij}(t) = \int \int (z - \mu_i(t))p_i(z, t)a(z, z')p_j(z', t) \, dz' \, dz,
\]
we get
\[
\frac{d\mu_i(t)}{dt} = h_i^2 \left( \overline{b}_i(t) - \sum_{j=1}^{S} \beta_{ij}(t)N_j(t) \right).
\]
(2.9)

Here \(\overline{b}_i(t)\) quantifies the selective pressure on the trait mean of species \(i\) caused by growth, and \(\beta_{ij}(t)\) the selective pressure on the trait mean of species \(i\) caused by competition with species \(j\).

Writing out the final results, the following system of differential equations describes the eco-evolutionary dynamics under Lotka–Volterra competition in the quantitative genetic limit:
\[
\frac{dN_i(t)}{dt} = N_i(t) \left( b_i(t) - \sum_{j=1}^{S} \alpha_{ij}(t)N_j(t) \right),
\]
(2.10)
\[
\frac{d\mu_i(t)}{dt} = h_i^2 \left( \overline{b}_i(t) - \sum_{j=1}^{S} \beta_{ij}(t)N_j(t) \right).
\]
(2.11)

### 2.2 The Jacobian of the quantitative genetic Lotka–Volterra model

The Jacobian of the dynamical equations Eqs. (2.10) and (2.11) is
\[
J_{ik} = \frac{\partial \left( \frac{dN_i}{dt}, \frac{d\mu_i}{dt} \right)}{\partial \left( N_k, \mu_k \right)}.
\]
(2.12)
We can calculate its entries block-by-block:

\[
\frac{\partial (dN_i/dt)}{\partial N_k} = \delta_{ik} \left( b_i - \sum_{j=1}^{S} \alpha_{ij} N_j \right) - N_i \alpha_{ik},
\]

\[
\frac{\partial (dN_i/dt)}{\partial \mu_k} = N_i \left( \frac{\partial b_i}{\partial \mu_k} - \sum_{j=1}^{S} \frac{\partial \alpha_{ij}}{\partial \mu_k} N_j \right) = N_i \left( \frac{\partial b_i}{\partial \mu_k} \delta_{ik} - \sum_{j=1}^{S} (\delta_{ik} + \delta_{jk}) \frac{\partial \alpha_{ij}}{\partial \mu_k} N_j \right),
\]  

(2.13)

\[
\frac{\partial (d\mu_i/dt)}{\partial N_k} = -h_i^2 \beta_{ik},
\]

\[
\frac{\partial (d\mu_i/dt)}{\partial \mu_k} = h_i^2 \left( \frac{\partial b_i}{\partial \mu_k} - \sum_{j=1}^{S} \frac{\partial \beta_{ij}}{\partial \mu_k} N_j \right) = h_i^2 \left( \frac{\partial b_i}{\partial \mu_k} \delta_{ik} - \sum_{j=1}^{S} (\delta_{ik} + \delta_{jk}) \frac{\partial \beta_{ij}}{\partial \mu_k} N_j \right),
\]

where \(\delta_{ij}\) is the Kronecker symbol, equal to 1 if \(i = j\) and to 0 otherwise. Here we used the fact that \(b_i\) and \(\tilde{b}_i\) only depend on \(\mu_i\) but not on \(\mu_k \neq i\); similarly, \(\alpha_{ij}\) and \(\beta_{ij}\) only depend on \(\mu_i\) and \(\mu_j\) but not on \(\mu_k \neq i,j\).

Collecting everything in a block-matrix, we get

\[
J_{ik} = \begin{pmatrix}
\delta_{ik} \left( b_i - \sum_{j=1}^{S} \alpha_{ij} N_j \right) - N_i \alpha_{ik} & N_i \left( \frac{\partial b_i}{\partial \mu_k} \delta_{ik} - \sum_{j=1}^{S} (\delta_{ik} + \delta_{jk}) \frac{\partial \alpha_{ij}}{\partial \mu_k} N_j \right) \\
-h_i^2 \beta_{ik} & h_i^2 \left( \frac{\partial b_i}{\partial \mu_k} \delta_{ik} - \sum_{j=1}^{S} (\delta_{ik} + \delta_{jk}) \frac{\partial \beta_{ij}}{\partial \mu_k} N_j \right)
\end{pmatrix},
\]

(2.14)

For \(S\) coexisting species, the equilibrium condition for the abundances can be expressed from Eq. (2.10) as the solution to

\[
\frac{\partial (dN_i/dt)}{\partial N_k} \bigg|_{\hat{N}} = -\hat{N} \alpha_{ik}.
\]

(2.16)

This is the community matrix (Levins 1968, May 1973), which we will denote with \(M\):

\[
M_{ik} = -\hat{N} \alpha_{ik}.
\]

(2.17)

### 2.3 Parameterizing the quantitative genetic Lotka–Volterra model

In this section, we choose the functional forms of the quantitative genetic Lotka–Volterra model’s ingredient functions in three alternative ways. First, as a common point across all three parameterizations, we choose a Gaussian interaction kernel that depends only on trait difference and not on the actual trait values themselves (homogeneous function):

\[
a(z, z') = \exp \left( -\frac{(z - z')^2}{\omega^2} \right).
\]

(2.18)
Here \( \omega \) determines the width of the kernel. We can now calculate \( \alpha_{ij}(t) \) and \( \beta_{ij}(t) \) by direct integration, using Eqs. (2.4) and (2.8):

\[
\alpha_{ij}(t) = \frac{\omega}{\sqrt{2\sigma_i^2 + 2\sigma_j^2 + \omega^2}} \exp\left(-\frac{(\mu_i(t) - \mu_j(t))^2}{2\sigma_i^2 + 2\sigma_j^2 + \omega^2}\right),
\]

\[(2.19)\]

\[
\beta_{ij}(t) = -\frac{2\omega \sigma_i^2 (\mu_i(t) - \mu_j(t))}{(2\sigma_i^2 + 2\sigma_j^2 + \omega^2)^{3/2}} \exp\left(-\frac{(\mu_i(t) - \mu_j(t))^2}{2\sigma_i^2 + 2\sigma_j^2 + \omega^2}\right).
\]

\[(2.20)\]

For \( b(z) \), we use three alternative forms, each expressing different biological scenarios.

I. \( b(z) \) rectangular:

\[
b(z) = \begin{cases} 
1 & \text{if } -\theta \leq z \leq \theta, \\
0 & \text{otherwise}. 
\end{cases}
\]

\[(2.21)\]

This function equals one in \([-\theta, \theta]\) and zero outside. Therefore, an individual with phenotype \( z \) can thrive equally efficiently within those limits but is suddenly unable to thrive at all outside. With \( b(z) \) specified, the integrals for \( b_i(t) \) and \( \tilde{b}_i(t) \), given in Eqs. (2.3) and (2.7), evaluate to

\[
b_i(t) = \frac{1}{2} \left[ \text{erf}\left(\frac{\theta - \mu_i(t)}{\sqrt{2\sigma_i}}\right) + \text{erf}\left(\frac{\theta + \mu_i(t)}{\sqrt{2\sigma_i}}\right) \right],
\]

\[(2.22)\]

\[
\tilde{b}_i(t) = \frac{\sigma_i}{\sqrt{2\pi}} \exp\left(-\frac{(\theta + \mu_i(t))^2}{2\sigma_i^2}\right) \left[ 1 - \exp\left(\frac{2\theta \mu_i(t)}{\sigma_i^2}\right) \right].
\]

\[(2.23)\]

II. \( b(z) \) quadratic:

\[
b(z) = \begin{cases} 
1 - z^2/\theta^2 & \text{if } -\theta \leq z \leq \theta, \\
0 & \text{otherwise}. 
\end{cases}
\]

\[(2.24)\]

This function decreases quadratically from its maximum at \( z = 0 \) and becomes zero at \( z = \pm \theta \). This means that there is a tradeoff between trait magnitude and performance: more extreme trait values have lower growth rates. Integrating for \( b_i(t) \) and \( \tilde{b}_i(t) \) using Eqs. (2.3) and (2.7), we get

\[
b_i(t) = \frac{1}{2\theta^2} \exp\left(-\frac{\theta^2 + \mu_i^2(t)}{\sigma_i^2}\right) \left\{ \exp\left(\frac{\theta^2 + \mu_i^2(t)}{\sigma_i^2}\right) \left[ \theta^2 - \sigma_i^2 - \mu_i^2(t) \right] \right\}
\times \left[ \text{erf}\left(\frac{\theta - \mu_i(t)}{\sqrt{2\sigma_i}}\right) + \text{erf}\left(\frac{\theta + \mu_i(t)}{\sqrt{2\sigma_i}}\right) \right]
\times \sqrt{\frac{2}{\pi}} \sigma_i \exp\left(\frac{(\theta^2 + \mu_i(t))}{2\sigma_i^2}\right) \left[ \theta + (\theta + \mu_i(t)) \exp\left(\frac{2\theta \mu_i(t)}{\sigma_i^2}\right) - \mu_i(t) \right],
\]

\[(2.25)\]

\[
\tilde{b}_i(t) = -\frac{\sigma_i^2}{\theta^2} \exp\left(-\frac{\theta^2 + \mu_i^2(t)}{\sigma_i^2}\right) \left\{ \mu_i(t) \exp\left(\frac{\theta^2 + \mu_i^2(t)}{\sigma_i^2}\right) \right\}
\times \left[ \text{erf}\left(\frac{\theta - \mu_i(t)}{\sqrt{2\sigma_i}}\right) + \text{erf}\left(\frac{\theta + \mu_i(t)}{\sqrt{2\sigma_i}}\right) \right]
\times \sqrt{\frac{2}{\pi}} \sigma_i \left\{ \exp\left(\frac{(\theta - \mu_i(t))^2}{2\sigma_i^2}\right) - \exp\left(\frac{(\theta + \mu_i(t))^2}{2\sigma_i^2}\right) \right\},
\]

\[(2.26)\]
III. $b(z)$ triangular:

$$b(z) = \begin{cases} 
\frac{z + \theta}{2\theta} & \text{if } -\theta \leq z \leq \theta, \\
0 & \text{otherwise.} 
\end{cases}$$  \tag{2.27}

This function increases linearly from zero to one in the interval $[-\theta, \theta]$ and is zero outside. As in the quadratic case, there is a tradeoff between trait value and performance, but it is asymmetric: smaller trait values are always less viable than larger ones in the region $[-\theta, \theta]$. Using Eqs. (2.3) and (2.7), $b_i(t)$ and $\bar{b}_i(t)$ evaluate to

$$b_i(t) = \frac{\theta + \mu_i(t)}{4\theta} \left[ \text{erf} \left( \frac{\theta - \mu_i(t)}{\sqrt{2}\sigma_i} \right) + \text{erf} \left( \frac{\theta + \mu_i(t)}{\sqrt{2}\sigma_i} \right) \right] - \frac{\sigma_i}{4\theta} \sqrt{\frac{2}{\pi}} \exp \left( -\frac{(\theta + \mu_i(t))^2}{2\sigma_i^2} \right) \left[ \exp \left( \frac{2\theta\mu_i(t)}{\sigma_i^2} \right) - 1 \right],$$

$$b_i(t) = \frac{\theta + \mu_i(t)}{4\theta} \left[ \text{erf} \left( \frac{\theta - \mu_i(t)}{\sqrt{2}\sigma_i} \right) + \text{erf} \left( \frac{\theta + \mu_i(t)}{\sqrt{2}\sigma_i} \right) \right] - \frac{\sigma_i}{4\theta} \sqrt{\frac{2}{\pi}} \exp \left( -\frac{(\theta - \mu_i(t))^2}{2\sigma_i^2} \right).$$

Substituting the results of either parameterization into Eqs. (2.10) and (2.11), the model is fully specified up to choosing numerical values for the parameters.

3 Two-species results

Here we use the quantitative genetic Lotka–Volterra model as described above in Section 2, except we do not specify $b(z)$ in any particular way unless noted otherwise. We set the number of species $S$ to two.

3.1 The ecological equilibrium and its stability

Let us consider just the ecological part of the dynamics. The equilibrium condition is given by setting Eq. (2.10) to zero. Requiring that the equilibrium densities be nonzero, we have, for $S = 2$,

$$b_i = \sum_{j=1}^{2} \alpha_{ij} \hat{N}_j, \tag{3.1}$$

where the hat denotes equilibrium values. The solution for the $\hat{N}_i$ is obtained by multiplying both sides with the inverse of the matrix $\alpha_{ij}$:

$$\hat{N}_i = \sum_{j=1}^{2} (\alpha^{-1})_{ij} b_j. \tag{3.2}$$

Assuming that the equilibrium described by Eq.(3.2) is feasible (i.e., all equilibrium abundances are positive), its global stability can be established. Writing out $\alpha_{ij}$ using Eq. (2.19), we get

$$\alpha_{ij} = \omega \begin{pmatrix} 
\frac{1}{\sqrt{4\sigma_1^2 + \omega^2}} & \exp \left( -d^2 \left( 2\sigma_1^2 + 2\sigma_2^2 + \omega^2 \right)^{-1} \right) \\
\exp \left( -d^2 \left( 2\sigma_1^2 + 2\sigma_2^2 + \omega^2 \right)^{-1} \right) & \frac{1}{\sqrt{4\sigma_2^2 + \omega^2}} 
\end{pmatrix}, \tag{3.3}$$

where $d = \frac{\theta - \mu_i(t)}{\sqrt{2}\sigma_i}.$
where \( d = |\mu_1 - \mu_2| \) is the distance between the species’ mean traits. This matrix is symmetric (implying that both eigenvalues are real), and purely competitive. Then, due to the classic theorem of MacArthur (1970), global stability of the equilibrium Eq. (3.2), if it exists, is guaranteed as long as \(-\alpha_{ij}\) has strictly negative eigenvalues. This in turn will hold if the trace and determinant of \(-\alpha_{ij}\) are negative and positive, respectively (Routh–Hurwitz criteria; Edelstein-Keshet 1988). The trace of \(-\alpha_{ij}\) is trivially negative. To show that the determinant \( \Delta = \text{det}(-\alpha_{ij}) = (-1)^2 \text{det}(\alpha_{ij}) = \text{det}(\alpha_{ij}) \) is positive, we calculate it using Eq. (3.3):

\[
\Delta = \alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21} = \omega^2 \left( \frac{1}{(4\sigma_1^2 + \omega^2)(4\sigma_2^2 + \omega^2)} - \exp\left(-2d^2(2\sigma_1^2 + 2\sigma_2^2 + \omega^2)^{-1}\right) \right). \tag{3.4}
\]

The exponential expression in the numerator of the second term is less than or equal to one. Notice that the larger this exponent is, the smaller the determinant will be. Assuming therefore, as a worst-case scenario, that it is precisely equal to one (i.e., \( d = 0 \)), the determinant \( \Delta \) will be positive as long as the denominator of the first term is smaller than the denominator of the second:

\[
\sqrt{(4\sigma_1^2 + \omega^2)(4\sigma_2^2 + \omega^2)} < 2\sigma_1^2 + 2\sigma_2^2 + \omega^2. \tag{3.5}
\]

Taking the square of both sides, rearranging, and simplifying, we get \( 0 < (\sigma_1^2 - \sigma_2^2)^2 \) as the condition for the positivity of the determinant, which is always satisfied except when \( \sigma_1 = \sigma_2 \). This establishes that any feasible equilibrium given by Eq. (3.2) will be globally stable unless \( d = 0 \) and \( \sigma_1 = \sigma_2 \) simultaneously.

### 3.2 Robustness and feasibility

The sensitivity of the equilibrium densities in Eq. (3.2) to perturbations of the \( b_i \) and \( \alpha_{ij} \) depends on the determinant \( \Delta \). Since the \( \hat{N}_i \) are proportional to the inverse of \( \alpha_{ij} \), and the inverse of a matrix is in turn proportional to the inverse of its determinant, a small \( \Delta \) means that a large number will multiply any perturbation of the parameters, causing a sudden, large shift in the equilibrium densities, increasing the risk of extinction (Meszéna et al. 2006, Barabás et al. 2014b). We should therefore see what the effect of intraspecific variation is on the determinant of \( \alpha_{ij} \).

In the limit of no intraspecific variation (\( \sigma_1 = \sigma_2 = 0 \)), the determinant in Eq. (3.4) reduces to

\[
\Delta_0 = 1 - \exp\left(-2d^2/\omega^2\right). \tag{3.6}
\]

We can compare the determinant \( \Delta \) of the two-species system with that of the exact same system, except that we set \( \sigma_1 \) and \( \sigma_2 \) to zero. This is effectively a comparison of the system’s robustness with- and without intraspecific variability.

Figure 1 shows both \( \Delta_0 \) and \( \Delta \) separately (left) as well as the ratio of \( \Delta \) to \( \Delta_0 \) (right) as a function of the mean trait difference \( d \), for various values of the intraspecific variance \( \sigma_2^2 \). What we see is that for large \( d \) the \( \Delta/\Delta_0 \) ratios asymptote, at smaller values for larger levels of intraspecific variation and always below one. This means that for \( d \to \infty \), intraspecific variation makes the two-species system more sensitive to parameter perturbations than it would be in the absence of such variation. This can be understood in the following way: as the distance between the species grows very large, the interspecific species interactions become negligible. Indeed, from Eq. (3.3) we see that \( \alpha_{12} \) and \( \alpha_{21} \) approach zero as \( d \to \infty \). Then, the two species decouple into two independent species following logistic growth, whose sensitivity depends on the strength of their intraspecific interactions \( \alpha_{11} \) and \( \alpha_{22} \): the stronger the self-regulation, the smaller the sensitivity due to an
Figure 1: The dependence of robustness on mean trait distance and intraspecific variability. For all curves, $\omega = 1.1$. For all curves except the black dashed line on the left, $\sigma_1 = 1$. Left: robustness, as measured by $\Delta = \det(\alpha_{ij})$. The dashed black line is the reference case with no intraspecific variation: $\sigma_1 = \sigma_2 = 0$. For all solid curves, $\sigma_1 = 1$ and colors denote different values of $\sigma_2$ (see legend). Except for very small differences in mean trait, the case with no intraspecific variation leads to the greatest robustness. Intraspecific variation therefore promotes the coexistence of tightly packed species, but only if their $\sigma$s are sufficiently different. Right: the ratio of the determinants with- and without intraspecific variation, $\Delta / \Delta_0$. The ratio is approximately constant but lower than one for $d$ large, decreases slightly for $d$ commensurable with the intraspecific standard deviations, and approaches infinity as $d \to 0$.

In summary, the sensitivity of coexistence to parameter perturbations is in general heightened by intraspecific variation, except when the mean traits of the two species are very close. In that case, stable and robust coexistence is possible even when the mean traits are identical ($d = 0$).

There is an alternative way of characterizing the robustness of coexistence and its dependence on intraspecific variation. We could ask: for two species with given trait means and variances, what...
Figure 2: Left: The intensity of interspecific competition between two species with the same mean trait is lowered if their trait variances are different. The reason is that the species with the higher variance has individuals that do not overlap with individuals of the other species. Right: when two species with the same mean trait have very similar intraspecific variances, the overlap in individuals is not reduced substantially, leading to nonrobust coexistence. The dashed gray line indicates the species’ mean trait position.

is the fraction of different growth rate vectors \((b_1, b_2)\) out of all possible combinations leading to a stable and feasible (all-positive) equilibrium? Since it was established in Section 3.1 that any feasible equilibrium will automatically be globally stable, we can concentrate just on feasibility. First, note that in Lotka–Volterra models only the direction of the vector of intrinsic growth rates matters, but not its magnitude. Indeed, from the equilibrium equations Eq. (3.2), multiplying all \(b_i\) values by a positive constant \(\eta\) simply rescales the equilibrium densities by \(\eta\) as well, therefore feasibility is unaffected. Second, the well-known conditions for feasibility in two-species competitive Lotka–Volterra systems read

\[ \frac{\alpha_{12}}{\alpha_{22}} \frac{b_1}{b_2} < \frac{\alpha_{11}}{\alpha_{21}} \]  

(e.g., Vandermeer 1975, Mallet 2012). These conditions can be written in a slightly different, more geometric form: since only the direction of the intrinsic rates matters, we can parameterize them as \(b_1 = \cos \phi, b_2 = \sin \phi\). Then the vector \((b_1, b_2)\) has length 1, and \(0 < \phi < \pi/2\) is the angle measured counterclockwise from the abscissa of the plane of possible intrinsic growth rate vectors. Substituting this into Eq. (3.7) and simplifying, we get

\[ \arctan \left( \frac{\alpha_{21}}{\alpha_{11}} \right) < \phi < \arctan \left( \frac{\alpha_{22}}{\alpha_{12}} \right). \]  

Using Eq. (3.3), this reads

\[ \arctan \left( e^{-\frac{d^2}{2\sigma_1^2 + 2\sigma_2^2 + \omega^2}} \sqrt{\frac{4\sigma_1^2 + \omega^2}{2\sigma_1^2 + 2\sigma_2^2 + \omega^2}} \right) < \phi < \arctan \left( e^{\frac{d^2}{2\sigma_1^2 + 2\sigma_2^2 + \omega^2}} \sqrt{\frac{2\sigma_1^2 + 2\sigma_2^2 + \omega^2}{4\sigma_2^2 + \omega^2}} \right). \]  

fully characterizing the set of parameters compatible with coexistence as a function of the trait means and variances (Figure 3).
Figure 3: The range of angles of the intrinsic growth rate vector \((b_1, b_2)\) (measured counterclockwise from the abscissa) allowing for two-species coexistence as a function of mean trait distance, for various values of intraspecific variability \(\sigma_2\) (see legend). Here \(\sigma_1 = 1\) and \(\omega = 1.1\). Lines denote the lower and upper boundaries of the shaded coexistence region. The black region with dashed boundaries is the reference case with no intraspecific variability: \(\sigma_1 = \sigma_2 = 0\). Unless mean trait differences are very small, the case without intraspecific variation always leads to a larger feasibility domain.

Figure 4: As Figure 1, but plotting the fraction of feasible parameter space \(\Xi\) (left) and the ratios \(\Xi/\Xi_0\) (right). The qualitative picture is the same as before, except all feasibility ratios converge to one for \(d\) large. This is because as \(d \to \infty\), the interspecific interaction coefficients approach zero, and the system decouples into two independent, logistically growing species. Any single species following logistic growth is viable as long as its intrinsic growth rate is positive, regardless of whether there is any intraspecific variability. Therefore, the feasibility ratios with- and without intraspecific variation converge to one as \(d\) increases.

We can also express the fraction of \(\phi\) values leading to feasibility, out of all possible values \(0 < \phi < \pi/2\), by rearranging the above inequalities. Calling this fraction \(\Xi\), we get

\[
\Xi = \frac{2}{\pi} \left[ \arctan \left( e^{\frac{d^2}{2\sigma_1^2 + 2\sigma_2^2 + \omega^2}} \sqrt{\frac{2\sigma_1^2 + 2\sigma_2^2 + \omega^2}{4\sigma_2^4 + \omega^2}} \right) - \arctan \left( e^{-\frac{d^2}{2\sigma_1^2 + 2\sigma_2^2 + \omega^2}} \sqrt{\frac{4\sigma_1^2 + \omega^2}{2\sigma_1^2 + 2\sigma_2^2 + \omega^2}} \right) \right].
\] (3.10)
In the absence of intraspecific variation \((\sigma_1 = \sigma_2 = 0)\), this simplifies to

\[
\Xi_0 = \frac{2}{\pi} \left[ \arctan \left( e^{d^2/w^2} \right) - \arctan \left( e^{-d^2/w^2} \right) \right]
\]

(3.11)

The feasibility domains \(\Xi\) and \(\Xi_0\) as well as their ratio may be plotted just as the determinants were (Figure 4). The exact same qualitative picture emerges as before. A quantitative difference is that \(\Xi/\Xi_0\) always asymptotes at one for \(d \to \infty\), whereas the determinant ratios do not. The reason is that when the two species are sufficiently far apart they decouple into two independent logistic species, with feasibility ensured for any positive intrinsic growth rate, regardless of intraspecific variation. The ratio of the two feasibility domains is therefore equal to one for \(d\) large.

### 3.3 The eco-evolutionary dynamics

For a single species, the eco-evolutionary dynamics given by Eqs. (2.10) and (2.11) is simple. The equilibrium density is the solution to \(dN/dt = N(b - \alpha N) = 0\), yielding \(\hat{N} = b/\alpha = b\sqrt{1 + (2\sigma/w)^2}\), where Eq. (2.19) was used to express \(\alpha = \alpha_{ii}\). This solution is feasible as long as \(b\) is positive. The evolutionary dynamics reads \(d\mu/dt = h^2(\bar{b} - \beta N)\). Unless \(h^2 = 0\), the evolutionary equilibrium is given by \(\bar{b} = \beta \hat{N}\), but since \(\beta = \beta_{ii} = 0\) from Eq. (2.20), we are left with \(\bar{b} = 0\). That is, any of the roots of \(\bar{b}\) may serve as evolutionary fixed points depending on initial conditions; local stability is ensured by \(\bar{b}\) having a negative slope at the corresponding root as a function of the mean trait \(\mu\).

For two species, Figures 5 and 6 show different scenarios with stable coexistence. In the first example the two species experience character divergence, while in the second the evolutionarily stable strategy involves both species having equal mean trait values. Of course, whether we see the convergence or divergence of the mean traits also depends on initial conditions (imagine Figure 5 with initial conditions \(\mu_1(0) = -5, \mu_2(0) = 5\)). However, it is of special interest to see when one can expect the perfect convergence of the two species’ mean traits.

**Figure 5:** Left panel: population densities of the two species in time. Right panel: the mean trait of the two species in time; the shaded regions mark the \(\pm \sigma_i\) regions around the mean. The two species diverge in their mean traits, ending up with coexistence in line with classical limiting similarity. We use the rectangular growth function (Section 2.3) with parameter values \(\theta = 0.5, \omega = 0.1, \tilde{h}_1^2 = \tilde{h}_2^2 = 0.5, \sigma_1 = 0.1, \sigma_2 = 0.15,\) and initial conditions \(N_1(0) = 1, N_2(0) = 1, \mu_1(0) = -0.2, \mu_2(0) = 0.1\).

To this end, let us assume that the two species have an evolutionary equilibrium at \(\mu_1 = \mu_2 = 0\); this can be achieved by choosing \(b(z)\) to be symmetric around \(z = 0\). In that case, \(\bar{b}_i\) is zero because
Eq. (2.7) reduces to the integral of the product of symmetric and antisymmetric functions, and $\beta_{ij}$ is also zero by Eq. (2.20). Then, from Eq. (2.11), we see that $\mu_1 = \mu_2 = 0$ is indeed an evolutionary equilibrium. We will also assume equal heritabilities ($h_1^2 = h_2^2 = h^2$), and that $b(z)$ is nonnegative.

We can write out the Jacobian $J$ for two species, using Eq. (2.14):

$$J = \begin{pmatrix} b_1 - 2\alpha_{11}N_1 - \alpha_{12}N_2 & -\alpha_{12}N_1 & N_1 \frac{\partial \mu_1}{\partial N_1} - N_1 N_2 \frac{\partial \alpha_{12}}{\partial N_1} & -N_1 N_2 \frac{\partial \alpha_{12}}{\partial N_2} \\ -\alpha_{21}N_2 & b_2 - 2\alpha_{21}N_1 - 2\alpha_{22}N_2 & -N_1 N_2 \frac{\partial \alpha_{21}}{\partial N_1} & N_2 \frac{\partial \alpha_{21}}{\partial N_2} \\ -h^2\beta_{11} & -h^2\beta_{12} & N_2 \frac{\partial \alpha_{11}}{\partial \mu_1} - 2N_1 \frac{\partial \alpha_{11}}{\partial \mu_2} - N_2 \frac{\partial \beta_{12}}{\partial \mu_1} & -N_1 N_2 \frac{\partial \beta_{12}}{\partial N_2} \\ -h^2\beta_{21} & -h^2\beta_{22} & -N_2 \frac{\partial \alpha_{21}}{\partial \mu_2} - 2N_1 \frac{\partial \alpha_{21}}{\partial \mu_2} - N_2 \frac{\partial \beta_{22}}{\partial \mu_2} & N_2 \frac{\partial \beta_{22}}{\partial \mu_2} \end{pmatrix}.$$  \hspace{2cm} (3.12)

At the (stable or unstable) evolutionary equilibrium $\mu_1 = \mu_2 = 0$, further simplifications are available: $\beta_{11} = \beta_{22} = 0,$ and $\partial \alpha_{ij}/\partial \mu_k = 0$ for all $i, j, k = 1, 2$. Also, since $b(z)$ is symmetric, it has an extremum around $z = 0$. Then, by Eq. (2.3), so does $b_i$ around $\mu_i = 0$:

$$\frac{\partial b_i}{\partial \mu_i} = \frac{\partial}{\partial \mu_i} \int p_i(z)b(z)\,dz = \int \frac{\partial p_i(z)}{\partial \mu_i} b(z)\,dz,$$  \hspace{2cm} (3.13)

which, for $\mu_i = 0$, is the integral of the product of an antisymmetric and a symmetric function, evaluating to zero. Therefore, $\partial b_i/\partial \mu_j = 0$ for all $i, j = 1, 2$. We then have

$$J = \begin{pmatrix} b_1 - 2\alpha_{11}N_1 - \alpha_{12}N_2 & -\alpha_{12}N_1 & 0 & 0 \\ -\alpha_{21}N_2 & b_2 - 2\alpha_{21}N_1 - 2\alpha_{22}N_2 & 0 & 0 \\ 0 & 0 & N_2 \frac{\partial \alpha_{11}}{\partial \mu_2} - 2N_1 \frac{\partial \alpha_{11}}{\partial \mu_2} - N_2 \frac{\partial \beta_{12}}{\partial \mu_1} & -N_1 N_2 \frac{\partial \beta_{12}}{\partial N_2} \\ 0 & 0 & -N_2 \frac{\partial \alpha_{21}}{\partial \mu_2} - 2N_1 \frac{\partial \alpha_{21}}{\partial \mu_2} - N_2 \frac{\partial \beta_{22}}{\partial \mu_2} & N_2 \frac{\partial \beta_{22}}{\partial \mu_2} \end{pmatrix}.$$  \hspace{2cm} (3.14)
which is the direct sum of two independent \( 2 \times 2 \) blocks—the top block corresponds to the dynamics of the densities, the bottom one to the trait means. Therefore, in this case the dynamics of the local abundances and trait means are decoupled, making the analysis much simpler. As we have shown in Section 3.1, the top \( 2 \times 2 \) block of the Jacobian is necessarily stable when \(-\alpha_{ij}\) is symmetric and all its entries are negative; this block describes meaningful coexistence when the feasibility condition Eq. (3.7) is also satisfied.

Local stability of the bottom \( 2 \times 2 \) block of Eq. (3.14) would mean that the two species experience convergent evolution, at least in the vicinity of \( \mu_1 = \mu_2 = 0 \). Using Eq. (3.2) for the equilibrium densities, one can calculate the two eigenvalues to determine whether the \( \mu_1 = \mu_2 = 0 \) evolutionary equilibrium is stable; both eigenvalues having negative real parts means stability. Equivalently, one may use the Routh–Hurwitz criteria (Edelstein-Keshet 1988): any equilibrium is stable; both eigenvalues having negative real parts means stability. Equivalently, one can use the Routh–Hurwitz criteria (Edelstein-Keshet 1988): any \( 2 \times 2 \) matrix is stable if and only if its trace is negative while its determinant is positive.

We introduce some simplifying notation. Let the bottom \( 2 \times 2 \) block of \( J \) be \( J^{(\mu)} \). Also, for any quantity \( x \), let \( \bar{x} = x/\omega \), i.e., we measure quantities in units of \( \omega \) whenever appropriate, thus eliminating one of the variables. Let \( \bar{b}_i = \partial b_i / \partial \mu_i \) evaluated at \( \mu_i = 0 \), and let

\[
\begin{align*}
    s_i &= \sqrt{1 + 4\hat{\sigma}_i^2}, \\
    s_{ij} &= \sqrt{1 + 2\hat{\sigma}_i^2 + 2\hat{\sigma}_j^2}.
\end{align*}
\]  

Writing \( \alpha_{ij} \) and \( \beta_{ij} \) using Eqs. (2.19) and (2.20) at \( \mu_1 = \mu_2 = 0 \), we can evaluate \( J^{(\mu)} \) at the equilibrium densities \( \hat{N}_1 \) and \( \hat{N}_2 \):

\[
J^{(\mu)} = h^2 \begin{pmatrix}
    \bar{b}_1' + 2\hat{N}_2\hat{\sigma}_2^2/s_{12}^3 & -2\hat{N}_2\hat{\sigma}_2^2/s_{12}^3 \\
    -2\hat{N}_1\hat{\sigma}_1^2/s_{12}^3 & \bar{b}_2' + 2\hat{N}_1\hat{\sigma}_1^2/s_{12}^3
\end{pmatrix}
\]  

The first Routh–Hurwitz criterion is that the trace of this matrix should be negative:

\[
\text{tr}(J^{(\mu)}) = h^2 \left( \bar{b}_1' + \bar{b}_2' + \frac{\hat{N}_2\hat{\sigma}_2^2 + \hat{N}_1\hat{\sigma}_1^2}{s_{12}^3} \right) < 0.
\]  

Rearranging, we get

\[
\frac{\hat{N}_2\hat{\sigma}_2^2 + \hat{N}_1\hat{\sigma}_1^2}{s_{12}^3} + \frac{\bar{b}_1' + \bar{b}_2'}{2} < 0.
\]  

Similarly, the second Routh–Hurwitz criterion is that the determinant should be positive:

\[
\det(J^{(\mu)}) = h^4 \left( \bar{b}_1' + 2\hat{N}_2\hat{\sigma}_2^2/s_{12}^3 \right) \left( \bar{b}_2' + 2\hat{N}_1\hat{\sigma}_1^2/s_{12}^3 \right) - \frac{4\hat{N}_1\hat{N}_2\hat{\sigma}_1^2\hat{\sigma}_2^2}{s_{12}^6} > 0,
\]  

or

\[
\frac{\bar{b}_1'\bar{b}_2'}{2} + \frac{\hat{\sigma}_1^2\hat{N}_2\bar{b}_2' + \hat{\sigma}_2^2\hat{N}_1\bar{b}_1'}{s_{12}^3} > 0.
\]

Note that \( \bar{b}_i' \) is necessarily negative: \( \bar{b}_i' = -|\bar{b}_i'| \). Indeed, using the definition in Eq. (2.7),

\[
\frac{\partial \bar{b}_i}{\partial \mu_i} = \frac{\partial}{\partial \mu_i} \int (z - \mu_i)p_i(z)b(z) \, dz = \int \left[ \frac{\partial (z - \mu_i)}{\partial \mu_i} p_i(z)b(z) + (z - \mu_i) \frac{\partial p_i(z)}{\partial \mu_i} b(z) \right] \, dz,
\]  

which, using Eq. (1.5) and evaluating at \( \mu_i = 0 \), reads

\[
\bar{b}_i' = \frac{\partial \bar{b}_i}{\partial \mu_i} \bigg|_{\mu_i=0} = -\int p_i(z)b(z) \, dz - \frac{1}{\sigma_i^2} \int z^2 p_i(z)b(z) \, dz.
\]

16
Both integrals are over products of positive functions; with the negative sign, the expression as a whole is negative. This shows that $b_i$ is necessarily negative.

The equilibrium densities $\hat{N}_1$ and $\hat{N}_2$ can be calculated from Eq. (3.2). The $\alpha_{ij}$ are given by Eq. (3.3) with $d = |\mu_1 - \mu_2| = 0$ which, using Eqs. (3.15) and (3.16), can be written

$$\alpha_{ij} = \begin{pmatrix} 1/s_1 & 1/s_2 \\ 1/s_1 & 1/s_2 \end{pmatrix}. \tag{3.24}$$

Inverting this matrix, we get

$$(\alpha^{-1})_{ij} = \frac{s_1 s_2 z^2}{s^2 z - s_1 s_2} \begin{pmatrix} 1/s_2 & -1/s_1 \\ -1/s_2 & 1/s_1 \end{pmatrix}. \tag{3.25}$$

Substituting this inverse into Eq. (3.2), we can express the two equilibrium densities:

$$\hat{N}_1 = \frac{s_1 s_2 (b_1 s_2 - b_2 s_1)}{s^2 z - s_1 s_2}, \tag{3.26}$$

$$\hat{N}_2 = \frac{s_2 s_2 z (b_2 s_1 - b_1 s_2)}{s^2 z - s_1 s_2}. \tag{3.27}$$

Substituting these into the Routh–Hurwitz criteria in Eqs. (3.19) and (3.21), we get

$$\frac{|b_1' + b_2'|}{2} > \frac{\sigma_1^2 s_2 (b_2 s_1 - b_1 s_2) + \sigma_2^2 s_1 (b_1 s_2 - b_2 s_1)}{s^2 z - s_1 s_2}, \tag{3.28}$$

$$\frac{|b_1'| |b_2'|}{2} > \frac{\sigma_1^2 s_2 (b_2 s_1 - b_1 s_2) |b_2'| + \sigma_2^2 s_1 (b_1 s_2 - b_2 s_1) |b_1'|}{s^2 z - s_1 s_2}. \tag{3.29}$$

Eq. (3.28) has the biological interpretation that the average slope of the $\bar{b}_i$ near $\mu = 0$ has to be sufficiently large in magnitude for character convergence to happen. However, the slopes cannot be large when the intraspecific standard deviations are either too small or too large. This is easily verified: for $\sigma_i \to \infty$ the integral in Eq. (2.7) converges to zero for all $\mu_i$ because the density of an infinitely wide Gaussian $p_i(z)$ is vanishing at all points (and so the integral’s derivative with respect to $\mu_i$ is also zero), while for $\sigma_i \to 0$ the trait distribution $p_i(z)$ converges to the Dirac delta distribution $\delta(z - \mu_i)$ and so, using Eq. (2.7),

$$\bar{b}_i = \frac{\partial}{\partial \mu_i} \int (z - \mu_i) \delta(z - \mu_i) b(z) \, dz \bigg|_{\mu_i = 0} = \frac{\partial}{\partial \mu_i} (0 \times b(\mu_i)) = 0. \tag{3.30}$$

Then, due to continuity in $\sigma_i$, very small or very large phenotypic standard deviations imply small values of $|\bar{b}_i|$. Its maximum is therefore obtained somewhere in between, when $\sigma_i \approx \bar{\theta}$. Biologically, this means that the intraspecific standard deviations can be neither simultaneously too small or simultaneously too large compared with the width of $b(z)$, otherwise we do not observe the convergence of the mean traits. Using Eqs. (3.28) and (3.29), Figure 7 shows the precise parameter region where one obtains convergent evolution, for the case of the rectangular growth function (Section 2.3).

Of course, evolutionary convergence does not guarantee that both species can maintain positive densities at the evolutionary equilibrium $\mu_1 = \mu_2 = 0$. For that, the ecological equilibrium must be locally stable and feasible. As discussed before, the top 2 × 2 block of $J$ is necessarily stable,
Figure 7: Region of convergent evolution (blue strip) as a function of the scaled intraspecific standard deviations $\tilde{\sigma}_1$ and $\tilde{\sigma}_2$. The growth function is rectangular (Section 2.3); the value of $\tilde{\theta}$ is set to 5. The blue region shows where the Routh–Hurwitz criteria of Eqs. (3.28) and (3.29) are satisfied; i.e., in this region the species evolve identical mean traits. In general, if one standard deviation is high, then the other must be intermediate to get the evolution of the mean traits, resulting in a generalist and a specialist species.

Therefore the only question is that of feasibility. The two-species feasibility condition is given by Eq. (3.7); for the rectangular growth function (Section 2.3), this reads

$$\frac{1 + 4\tilde{\sigma}_2^2}{1 + 2\tilde{\sigma}_1^2 + 2\tilde{\sigma}_2^2} < \frac{\text{erf}^2\left(\tilde{\theta}/(\sqrt{2}\tilde{\sigma}_1)\right)}{\text{erf}^2\left(\tilde{\theta}/(\sqrt{2}\tilde{\sigma}_2)\right)} < \frac{1 + 2\tilde{\sigma}_1^2 + 2\tilde{\sigma}_2^2}{1 + 4\tilde{\sigma}_1^2}.$$  \hspace{1cm} (3.31)

In particular, these inequalities are very difficult to satisfy when $\tilde{\sigma}_1 \approx \tilde{\sigma}_2$, and become impossible to do so for $\tilde{\sigma}_1 = \tilde{\sigma}_2$. In order to observe trait convergence which also ends up in ecological coexistence, one therefore needs sufficient segregation of the species in their intraspecific standard deviations. Roughly speaking, one species must be a generalist and the other a specialist consumer.

The robust possibility of convergent trait evolution points to the fact that measuring species similarity purely on the basis of trait means can be misleading. Coexistence can result from segregation along the trait axis (Figure 5), but also from having similar mean traits but different intraspecific standard deviations (Figure 6). Limiting similarity in the classical sense (MacArthur and Levins 1967) therefore has to be extended in the presence of intraspecific trait variation: its naive application to the trait means without taking the variances into account may lead to the incorrect conclusion that two species coexist without niche segregation, when they are in fact niche-segregated.

3.4 Robustness of the two-species results to an asymmetry in intrinsic growth

How sensitive are our results to breaking the perfect symmetry of the intrinsic growth function $b(z)$? To examine this, let us introduce an asymmetry in $b(z)$ to see if we can still get character convergence. The difficulty with doing this lies in the fact that, after perturbing $b(z)$, one has to
recalculate $b_i(t)$ and $\bar{b}_i(t)$ via Eqs. (2.3) and (2.7). Unfortunately, these integrals cannot in general be evaluated in explicit form, except in certain very simple cases.

The idea for introducing asymmetry in a tractable way is to assume that it is small, so one can expand it to linear order. Instead of the original $b_i(t)$ given by Eq. (2.3), we use the perturbed

$$b_i^*(t) = \int b(z)(1 + \eta(z))p_i(z,t) \, dz.$$  \hspace{1cm} (3.32)

If the perturbing function $\eta(z)$ is small and has a nonvanishing first moment, then we can Taylor expand it as $\eta(z) \approx cz$, where $c$ is a constant (we choose $\eta$ such that the 0th-order term in the expansion is zero). We then can write

$$b_i^*(t) \approx \int b(z)(1 + cz)p_i(z,t) \, dz = b_i(t) + c \int z b(z)p_i(z,t) \, dz = b_i(t) + c \int (z - \mu_i(t)) b(z)p_i(z,t) \, dz + c \mu_i(t) \int b(z)p_i(z,t) \, dz$$  \hspace{1cm} (3.33)

Similarly, instead of $\bar{b}_i(t)$ given by Eq. (2.7), we have

$$\bar{b}_i^*(t) \approx \int (z - \mu_i(t))b(z)(1 + cz)p_i(z,t) \, dz = \bar{b}_i(t) + c \int (z - \mu_i(t)) b(z)p_i(z,t) \, dz$$  \hspace{1cm} (3.34)

Defining $\hat{b}_i(t)$ as

$$\hat{b}_i(t) = \int (z - \mu_i(t))^2 b(z)p_i(z,t) \, dz$$  \hspace{1cm} (3.35)

(i.e., the second central moment), we can continue writing

$$\bar{b}_i^*(t) \approx \bar{b}_i(t) + c \int [(z - \mu_i(t))^2 + \mu_i(t)(z - \mu_i(t))] b(z)p_i(z,t) \, dz$$  \hspace{1cm} (3.36)

The advantage of this approach is highlighted by the fact that $\hat{b}_i(t)$ can be explicitly integrated, depending on parameterization (Section 2.3):

I. $b(z)$ rectangular:

$$\hat{b}_i(t) = \frac{\sigma_i^2}{2} \left[ \text{erf}\left(\frac{\theta - \mu_i(t)}{\sqrt{2\sigma_i}}\right) + \text{erf}\left(\frac{\theta + \mu_i(t)}{\sqrt{2\sigma_i}}\right) \right] + \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-(\theta + \mu_i(t))^2}{2\sigma_i^2}\right) \left[ (\mu_i(t) - \theta) \sigma_i \exp\left(\frac{2\theta \mu_i(t)}{\sigma_i^2}\right) - (\mu_i(t) + \theta) \sigma_i \right].$$  \hspace{1cm} (3.37)

II. $b(z)$ quadratic:

$$\hat{b}_i(t) = \frac{\sigma_i^2}{2\theta^2} \left\{ \left( \theta^2 - \mu_i(t)^2 - 3\sigma_i^2 \right) \left[ \text{erf}\left(\frac{\theta - \mu_i(t)}{\sqrt{2\sigma_i}}\right) + \text{erf}\left(\frac{\theta + \mu_i(t)}{\sqrt{2\sigma_i}}\right) \right] \right. + \left. \sqrt{\frac{2}{\pi}} \sigma_i \exp\left(\frac{-(\theta + \mu_i(t))^2}{2\sigma_i^2}\right) \left[ (3\theta + \mu_i(t)) \exp\left(\frac{2\theta \mu_i(t)}{\sigma_i^2}\right) + 3\theta - \mu_i(t) \right] \right\}.$$  \hspace{1cm} (3.38)
Figure 8: As Figure 5, except with an asymmetry in the intrinsic growth function \( b(z)(1 + cz) \), with \( c = 0.1 \).

III. \( b(z) \) triangular:

\[
\dot{b}_i(t) = \frac{\sigma_i^2}{4\theta} (\theta + \mu_i(t)) \left[ \text{erf}\left(\frac{\theta - \mu_i(t)}{\sqrt{2}\sigma_i}\right) + \text{erf}\left(\frac{\theta + \mu_i(t)}{\sqrt{2}\sigma_i}\right) \right] \\
+ \frac{\sigma_i}{\sqrt{2\pi\theta}} \exp\left(-\frac{(\theta + \mu_i(t))^2}{2\sigma_i^2}\right) \left[ \sigma_i^2 - \exp\left(\frac{2\theta\mu_i(t)}{\sigma_i^2}\right) \left(\theta^2 + \sigma_i^2 - \theta\mu_i(t)\right) \right].
\] (3.39)

Figure 9: As Figure 6, except with an asymmetry in the intrinsic growth function \( b(z)(1 + cz) \), with \( c = 0.1 \).

When the intraspecific standard deviations are low, then predictably, asymmetry does not alter qualitative behavior, and we get the same result as before (Figure 8, which is much the same as Figure 5 despite the added asymmetry). However, on Figure 9, which is set up just like Figure 6 plus the asymmetry, the two species no longer evolve the exact same mean trait. In and of itself, this is not very surprising, because for \( b(z) \) asymmetric, \( \mu_1 = \mu_2 = 0 \) can no longer in general be expected to be a fixed point of the evolutionary dynamics. However, the difference \( d = |\mu_1 - \mu_2| \) in mean traits is slight, especially compared with the two intraspecific standard deviations. This holds even though the amount of introduced asymmetry is in fact not very small (Figure 10).
Figure 10: Equilibrium state of two-species communities, with- and without asymmetry in intrinsic growth (left and right panels, respectively). Both panels used the rectangular growth function (Section 2.3), and parameter values $\theta = 5$, $\omega = 1$, $h_1^2 = h_2^2 = 0.5$, $\sigma_1 = 3$, $\sigma_2 = 4.5$, starting from initial conditions $N_1(0) = 1$, $N_2(0) = 1$, $\mu_1(0) = -2$, $\mu_2(0) = 1$. The asymmetry parameter is set to $c = 0$ (left panel) and $c = 0.1$ (right panel). The shaded regions are the density distributions of the two species at equilibrium as a function of trait value. The dashed red line in each panel (not drawn to scale) is the intrinsic growth function $b(z)(1 + cz)$, where $b(z)$ is given by Eq. (2.21). On the left, the two species evolve the same mean trait, but on the right their alignment is not perfect due to the asymmetry.

In summary, though the species no longer evolve the exact same mean trait as a stable evolutionary outcome in the presence of asymmetry, it is nevertheless still true that, for sufficiently large intraspecific standard deviations, the difference in their mean traits will be small compared to their $\sigma$s, and so it is still true that the two species represent a specialist and a generalist type.

4 Multispecies results

4.1 Multiple species with equal intraspecific variances and no heritability

Here we study a simple reference case for which an exact answer can be given to the question of how intraspecific variability affects expected species richness. To this end, we make two assumptions. First, all heritabilities are zero: intraspecific variation is purely plastic. Second, each species has the same intraspecific standard deviation $\sigma$.

The competition coefficients $\alpha_{ij}$ are given by Eq. (2.19), which reduces to

$$
\alpha_{ij} = \frac{\omega}{\sqrt{4\sigma^2 + \omega^2}} \exp \left( -\frac{(\mu_i - \mu_j)^2}{4\sigma^2 + \omega^2} \right)
$$

when all the $\sigma$s are equal. In the absence of intraspecific variation, the competition coefficients are instead given by Eq. (2.18):

$$
\alpha_{ij} = \exp \left( -\frac{(\mu_i - \mu_j)^2}{\omega^2} \right),
$$

which can also be obtained by taking the $\sigma \to 0$ limit in Eq. (4.1). If the heritabilities are zero, then species do not evolve, and the only difference between the model with- and without intraspecific variation is that the competition coefficients have a different form, given by Eqs. (4.1) and (4.2).
The number of coexisting species in a Lotka–Volterra trait axis model without intraspecific trait variation is proportional to the inverse of the competition width \( \omega \) in Eq. (4.2) (MacArthur and Levins 1967, Szabó and Meszéna 2006, Scheffer and van Nes 2006, Barabás and Meszéna 2009), a rule of thumb that holds unless parameters are carefully (and unbiologically) fine-tuned (Barabás et al. 2012). However, when all \( \sigma \)s are equal, one can think of the model with intraspecific variation as an effective model for a community without such variation, but with the modified competition coefficients given by Eq. (4.1). The competition width of these effective coefficients is \( \sqrt{4\sigma^2 + \omega^2} \). The ratio of predicted species richnesses with- and without intraspecific variation is therefore \( \omega/\sqrt{4\sigma^2 + \omega^2} \). Equivalently, we can also write \( 1/\sqrt{4\tilde{\sigma}^2 + 1} \) for the ratio, where \( \tilde{\sigma} = \sigma/\omega \). Figure 11 compares this analytical prediction with simulation results, starting from the same initial conditions and varying \( \sigma \) for two different values of \( \omega \).

![Figure 11: Comparison of the analytically predicted ratio of species richnesses with- and without intraspecific variation, \( \omega/\sqrt{4\sigma^2 + \omega^2} \) (blue curves) to simulated ones (red dots; obtained with the rectangular \( b(z) \) of Section 2.3). Left panel: \( \omega = 0.1 \); right panel: \( \omega = 0.2 \). There is more scatter on the right because the average number of species with \( \omega = 0.2 \) is half the number with \( \omega = 0.1 \). Since the number of coexisting species must be an integer, rounding errors are expected to be larger when the number of species is small to begin with.

The important conclusions about species richness are as follows. First, intraspecific variability has an overall negative effect on species diversity. Second, this negative effect increases monotonically with \( \sigma \). Third, for very small amounts of intraspecific variability, there is practically no effect, since \( 1/\sqrt{4\tilde{\sigma}^2 + 1} \approx 1 \) for \( \tilde{\sigma} \) small. Fourth, for large \( \tilde{\sigma} \), \( 1/\sqrt{4\tilde{\sigma}^2 + 1} \approx (2\tilde{\sigma})^{-1} \), therefore the decline in species richness will be roughly inversely proportional to \( 2\sigma/\omega \).

### 4.2 Measuring community robustness

Local stability of an ecological community in the final state is guaranteed if the community matrix \( M_{ij} = -\hat{N}_i\alpha_{ij} \) given in Eq. (2.17) has all its eigenvalues lying in the left half of the complex plane. Here \( \hat{N}_i \) is the equilibrium density of species \( i \), and the \( \alpha_{ij} \) are given by Eq. (2.19). The real part of an eigenvalue measures the rate of decay (or amplification, if the real part is positive) of a dynamical perturbation along the corresponding eigendirection; small negative real parts indicate slow return times.

The eigenvalues of the community matrix are also related to the robustness of the community. A robust system is one that does not change its qualitative behavior in response to parameter perturbations. Intuitively, the reason the eigenvalues measure robustness is that their position in
the complex plane is a continuous function of model parameters. If an eigenvalue is close to the imaginary axis, then even a small parameter perturbation may push it over to the right half plane, destabilizing the system. More precisely, the magnitude of the community matrix’s determinant is a measure of the local robustness of a deterministic dynamical system (Levins 1968, May 1973, Levins 1974, Bender et al. 1984, Yodzis 1988, 2000, Dambacher et al. 2002, Meszéna et al. 2006, Novak et al. 2011, Yeakel et al. 2011, Barabás et al. 2012, Aufderheide et al. 2013, Barabás et al. 2014a,b). The intuition about the eigenvalues measuring robustness is then justified by recognizing that the determinant is the product of all eigenvalues. The $S$th root of its absolute value is therefore the (geometric) mean of the eigenvalues’ moduli: denoting the $i$th eigenvalue of $M$ by $\lambda_i$, we have

$$\sqrt[\frac{S}{S}]{|\det(M)|} = \sqrt[\frac{S}{S}]{|\lambda_1 \cdot \ldots \cdot \lambda_S|} = \sqrt[\frac{S}{S}]{|\prod_{i=1}^{S} |\lambda_i||} = \exp\left(\frac{1}{S} \log\left(\prod_{i=1}^{S} |\lambda_i|\right)\right) = \exp\left(\frac{1}{S} \sum_{i=1}^{S} \log(|\lambda_i|)\right) = \exp\left(\log(|\lambda|)\right),$$

where the overbar denotes (arithmetic) averaging. This quantity can be interpreted as the geometric mean of the return times along each eigendirection, which is especially useful for comparing the robustness of systems with different numbers of species (Barabás et al. 2014b). The determinant itself cannot be used for such comparisons, because its units are [time]$^{-S}$, which will be different for communities with different numbers of species. In contrast, the units of the geometric mean are always [time]$^{-1}$, comparable across systems. We call this quantity the “average community robustness”.

There is another aspect to robustness. Imagine a two-species community whose community matrix has eigenvalues $\lambda_1 = \lambda_2 = -1$. The geometric mean of their magnitudes is 1. Now take another two-species system with eigenvalues $\lambda_1 = -100$ and $\lambda_2 = -0.01$. The geometric mean is still equal to 1, but the situation is now very different. In the first example, the dynamics are equally robust along both eigendirections. In the second, the first eigendirection is very robust, but the second one is not robust at all. This indicates that this second system will be completely insensitive to some types of parameter perturbations, but extremely sensitive to others. It is therefore also of interest to see not just how robust a system is to the average environmental disturbance, but whether it reacts strongly to specific disturbances. This can be measured by the geometric standard deviation of the eigenvalues’ magnitudes:

$$\exp\left(\frac{1}{S} \sum_{i=1}^{S} \log^2(|\lambda_i|) - \left(\frac{1}{S} \sum_{i=1}^{S} \log(|\lambda_i|)\right)^2\right).$$

We dub this quantity “robustness heterogeneity”, i.e., the degree to which the system’s eigendirections vary in their individual robustness.

### 4.3 The effect of intraspecific variation and heritability on species richness, trait pattern, and community robustness

Here we show all our results generated using the rectangular $b(z)$ given in Eq. (2.21), and the triangular $b(z)$ given in Eq. (2.27). The plots are analogous to Figures 4, 5, and 6 in the main text (where the quadratic $b(z)$, given by Eq. (2.24), was used). The quantitative details are different, but the qualitative results are unaffected by the choice of $b(z)$. 

23
Figure 12: As Figure 4 in the main text, except showing results with $b(z)$ rectangular (Eq. 2.21).

Figure 13: As Figure 5 in the main text, except showing results with $b(z)$ rectangular (Eq. 2.21).
Figure 14: As Figure 6 in the main text, except showing results with $b(z)$ rectangular (Eq. 2.21).

Figure 15: As Figure 4 in the main text, except showing results with $b(z)$ triangular (Eq. 2.27).
Figure 16: As Figure 5 in the main text, except showing results with $b(z)$ triangular (Eq. 2.27).

Figure 17: As Figure 6 in the main text, except showing results with $b(z)$ triangular (Eq. 2.27).
4.4 Trait convergence in multispecies communities

In Section 3.3, the possibility of the evolutionary convergence of the trait means of two species was discussed. One important question is whether we observe such convergent evolution in our multispecies simulated communities. The answer is: yes, but only sporadically. Assuming $h^2 > 0$, we counted the total number of simulations where the minimum trait difference between two coexisting species was less than 0.03: it was 60 cases altogether (a little over one percent of all cases with $h^2 > 0$). We can do the same for simulations where the minimum difference of trait means is 0.02 (29 cases, or half a percent), and 0.01 (13 cases, or quarter of a percent). Figure 18 shows examples from our data where pairs of species evolved very similar trait means.

![Figure 18](image)

**Figure 18:** Simulation results with $h^2 > 0$ where the final community has species with mean trait difference less than 0.01 units along the trait axis. The dashed lines (not drawn to scale) depict the shape of the intrinsic growth function.

The biological interpretation of the inequalities (3.28) and (3.29) necessary for two-species trait convergence were that trait variances must be sufficiently different, and the fitness landscape at the point of convergent evolution should be sufficiently sharply changing. The second of these conditions is most easily achieved when $\sigma \approx \theta$, and indeed, that was when we observed trait convergence in the two-species scenarios (Section 3.3). However, in our multispecies communities, the trait variances are always considerably smaller than $\theta$ (compare Figure 18 with Figure 10). The only way to achieve the required sharply changing fitness landscape is through frequency-dependent effects,
such as interactions with flanking species from either side. The resulting landscape will likely not be perfectly symmetric, but the effect of a small asymmetry is simply to shift the trait means slightly away from strict equality (Section 3.4).

Since to get trait convergence the two trait variances need to be sufficiently different, one would think that most simulations where convergence was observed had mixed levels of intraspecific variation (intraspecific standard deviations uniformly sampled from [0.01, 0.3]; this distribution has the largest variance of the $\sigma$s). This is indeed the case: out of the 60 cases where the minimum trait difference was less than 0.03, 41 had mixed levels of variation (68%); out of the 29 with minimum difference less than 0.02, 24 (83%); and out of the 13 with minimum difference below 0.01, 10 (77%).

Since the conditions for trait convergence are somewhat restrictive, we do not expect it to happen very often. Our simulation results show that, while convergence is definitely a rare event, it still happens occasionally. However, measured by its effect on our community metrics (Section 4.3), trait convergence does not appear to be a very strong force in shaping community structure.

5 Relaxing the assumptions of the extreme quantitative genetic limit: the Shpak-Kondrashov model

So far all our results were obtained under the assumption that the ecologically relevant trait $z$ in question is determined by the sum of a random, normally distributed environmental noise, and a genetic component with infinitely many loci contributing to the trait, each locus having an infinitesimal additive effect. Under these assumptions, the intraspecific trait distribution is normal and selection can only change its mean, not the higher moments (see Section 1). Here we check the robustness of our results to relaxing the assumption of infinitely many loci.

One way of doing this would be to do a purely individual-based simulation. The disadvantage to this is that such simulations are very expensive computationally. A convenient middle ground is provided by the Shpak-Kondrashov hypergeometric model (Shpak and Kondrashov 1999, Yamauchi and Miki 2009). In this model the trait of interest is coded by $n$ additive loci in a diploid, sexually reproducing population. The probability $R_{vw}^u$ that two parents with genotypes $v$ and $w$ give birth to an offspring with genotype $u$ is modeled via

$$R_{vw}^u = \sum_{s=0}^{v} H(v, \max(v, s + v - 2n)) H(w, \max(w, u + w - s - 2n)),$$

where

$$H(u, v) = \binom{u}{v} \binom{u}{v - u} \binom{2u}{u}$$

is the hypergeometric distribution. This model is accurate as long as the allelic composition of every phenotype is in linkage equilibrium. When recombination is free, mating is random, and selection only depends on phenotype, no within-phenotype linkage disequilibrium will evolve as long as there was none initially (Shpak and Kondrashov 1999). Since in our Lotka–Volterra ecological models selection is a function of phenotype, the hypergeometric model’s assumptions hold as long as there is no initial linkage disequilibrium within phenotypes.
5.1 Community model with no environmental effect on phenotype

At first let us assume that genotype fully determines phenotype, i.e., there is no environmental component to the total phenotypic variance, making heritability equal to 1. Let the population density of species $i$’s genotype $u$ at time $t$ be $N_{iu}(t)$. Assuming a short time slice $\Delta t$, let the same densities be $N'_{iu}(t)$ after selection, and then $N_{iu}(t + \Delta t)$ after reproduction. Selection happens via Lotka–Volterra dynamics:

$$N'_{iu}(t) = N_{iu}(t) - \Delta t N_{iu}(t) \sum_v \sum_j a_{uv} N_{jv}(t),$$

(5.3)

where $a_{uv}$ is the competitive effect of individuals with genotype $v$ on those with genotype $u$. Given the trait values $z$ and $z'$ associated with these genotypes, the competition coefficients are given by Eq. (2.18). Since we assume that all species have the same genetic structure, these coefficients are independent of species identity.

For reproduction, if the density-independent growth rate of species $i$’s genotype $u$ is $b_{iu}$, and the offspring genotype frequencies are given by the hypergeometric model, we can write

$$N_{iu}(t + \Delta t) = N'_{iu}(t) + \Delta t b_{iu} c_i^{-1} \sum_v \sum_w R_{vw}^u N_{iv}(t) N_{iw}(t)$$

(5.4)

for the densities after the time slice $\Delta t$. Here $c_i^{-1}$ simply ensures proper normalization:

$$c_i = \sum_u \sum_v \sum_w R_{vw}^u N_{iv}(t) N_{iw}(t).$$

(5.5)

Putting Eqs. (5.3) and (5.4) together, we get

$$N_{iu}(t + \Delta t) = \left( N_{iu}(t) + \Delta t b_{iu} c_i^{-1} \sum_v \sum_w R_{vw}^u N_{iv}(t) N_{iw}(t) \right) \left( 1 - \Delta t \sum_v \sum_j a_{uv} N_{jv}(t) \right).$$

(5.6)

Neglecting higher order terms, subtracting $N_{iu}(t)$ from both sides, dividing by $\Delta t$, and taking the $\Delta t \to 0$ limit, we arrive at

$$\frac{dN_{iu}(t)}{dt} = b_{iu} c_i^{-1} \sum_v \sum_w R_{vw}^u N_{iv}(t) N_{iw}(t) - N_{iu}(t) \sum_v \sum_j a_{uv} N_{jv}(t),$$

(5.7)

the differential equation governing genotype $u$ of species $i$, assuming that genotype perfectly determines phenotype. Note that the number of equations is the number of species times the number of possible genotypes; for instance, with $S = 51$ species and $n = 25$ loci (leading to $2n + 1 = 51$ distinct genotypes), the number of differential equations is $51 \times 51 = 2601$. This is computationally much more expensive than simulating the quantitative genetic limit. For this reason, we do not perform the same extensive simulations as before. Instead, we simulate a few scenarios to see whether the results are qualitatively the same as in the quantitative genetic limit.

Figure 19 shows a sample run of this model with 51 diploid species, each with $n = 25$ additive loci. Note that approximately normal trait distributions are maintained even under strong, constant selection. The qualitative result is the same as in the extreme quantitative genetic limit; the only difference is that, since genotype is the only determinant of phenotype, and all species share the same genetic system, species may evolve to have identical phenotypic distributions, thus effectively morphing into the same species. This convergent evolution is enhanced by the effective discretization of the trait space, since for $n$ additive loci only $2n + 1$ different trait values are possible. Once two species achieve the same genotype (and therefore phenotype) distribution, they are, for all intents and purposes, the same species, the differences in their phylogenetic histories completely erased.
Figure 19: Final state of species’ trait distributions (shaded regions) from simulating Eq. (5.7) with \( \omega = 0.1 \), \( S = 51 \) diploid species, and random initial conditions. Each species has \( n = 25 \) additive loci, each with two possible alleles (“0” and “+”), for a total of \( 2n + 1 = 51 \) different genotypic values. We assume a species has trait \(-0.6\) if all its loci are equal to “0”, and every “+” locus contributes 0.024 units to the trait value; i.e., the maximum possible trait value is \( 50 \times 0.024 - 0.6 = 0.6 \). We assume \( b_{iu} \) depends only on the genotypic value \( u \): \( b_{iu} = 1 \) if \( u \) is such that the phenotypic value is between \( \pm 0.5 \) and zero otherwise (red dashed line; not drawn to scale). Species’ trait distributions are approximately normal even close to the edges of the trait axis, where there is strong, externally imposed selection acting. Since the finite number of loci effectively discretizes the trait axis into 51 bins, and all species have the exact same genetic structure, once two species evolve the same phenotypic distribution, they become completely identical with no competitive exclusion between them. This is exactly what we see on the figure: there are five well-defined regions along the axis where species can persist, but each region is occupied by several identical, neutrally coexisting “species”.

5.2 Community model with environmental effects on phenotype

To allow for heritabilities different from one, we assume that there is species-specific Gaussian environmental noise on top of the trait values determined by the genotypes. The distribution \( p_{iu}(z) \) of the phenotype values of species \( i \)'s genotype \( u \) then reads

\[
p_{iu}(z) = \sqrt{\frac{1}{2\pi \sigma_{E,i}^2}} \exp\left( -\frac{(\mu_{iu} - z)^2}{2\sigma_{E,i}^2} \right),
\]

(5.8)

where \( \mu_{iu} \) is the mean phenotypic trait of species \( i \)'s genotype \( u \) (i.e., what it would always be in the absence of environmental noise), which does not depend on time. The variance \( \sigma_{E,i}^2 \) is species-specific. Instead of \( N_{iu}(t) \), we now have \( N_{iu}(z, t) \), which is the trait distribution of species \( i \)'s genotype \( u \):

\[
N_{iu}(z, t) = N_{iu}(t)p_{iu}(z) = N_{iu}(t)\sqrt{\frac{1}{2\pi \sigma_{E,i}^2}} \exp\left( -\frac{(\mu_{iu} - z)^2}{2\sigma_{E,i}^2} \right).
\]

(5.9)

The integral of this distribution along the whole trait axis is

\[
\int N_{iu}(z, t) \, dz = N_{iu}(t) \int p_{iu}(z) \, dz = N_{iu}(t),
\]

(5.10)

as it should be.
We integrate both sides across \( z \). The left hand side is simply
\[
\int \frac{dN_{iu}(z,t)}{dt} \, dz = \frac{d}{dt} \int N_{iu}(z,t) \, dz = \frac{dN_{iu}(t)}{dt},
\]
(5.12)
while the right hand side is
\[
\int b(z)p_{iu}(z) \, dz c_i^{-1} \sum_v \sum_w R_{vw}^u N_{iw}(t) N_{iw}(t) - N_{iu}(t) \sum_v \sum_j N_{jw}(t) \int p_{iu}(z)a(z,z')p_{jw}(z') \, dz' \, dz.
\]
(5.13)
Defining
\[
b_{iu} = \int b(z)p_{iu}(z) \, dz
\]
(5.14)
and
\[
a_{iu,jv} = \int \int p_{iu}(z)a(z,z')p_{jw}(z') \, dz' \, dz,
\]
(5.15)
the final form of the equation reads
\[
\frac{dN_{iu}(t)}{dt} = b_{iu}c_i^{-1} \sum_v \sum_w R_{vw}^u N_{iw}(t) N_{iw}(t) - N_{iu}(t) \sum_v \sum_j a_{iu,jv} N_{jw}(t).
\]
(5.16)
This has the same form as Eq. (5.7), except the two-index \( a_{uw} \) is replaced with the four-index \( a_{iu,jv} \), which depends not only on the genotypes, but also the environmental variances of species \( i \) and \( j \). These ingredient functions can be integrated just like in Sections 2.3 and 2.3. Assuming a rectangular \( b(z) \) given by Eq. (2.21) and a Gaussian \( a(z,z') \) given by Eq. (2.18), we get
\[
b_{iu} = \frac{1}{2} \left[ \text{erf} \left( \frac{\theta - \mu_{iu}}{\sqrt{2}\sigma_{E,i}} \right) + \text{erf} \left( \frac{\theta + \mu_{iu}}{\sqrt{2}\sigma_{E,i}} \right) \right],
\]
(5.17)
\[
a_{iu,jv} = \frac{\omega}{\sqrt{2\sigma_{E,i}^2 + 2\sigma_{E,j}^2 + \omega^2}} \exp \left( -\frac{(\mu_{iu} - \mu_{jv})^2}{2\sigma_{E,i}^2 + 2\sigma_{E,j}^2 + \omega^2} \right).
\]
(5.18)
Note also that Eq. (5.16) collapses into Eq. (5.7) in the limit of all \( \sigma_{E,i}^2 \) going to zero.

Adding an environmental component to species’ phenotypic variances breaks the perfect symmetry of the species in that now even if two species evolve the exact same genotype distribution, their trait distributions will be different (unless their environmental variances are precisely equal). Because of this, species retain their identity even if they undergo convergent evolution, and cannot become identical. In this scenario therefore, we end up with communities that look very much like those obtained under the quantitative genetic limit (Figure 20, left panel). This is despite the fact that, since the amount of within-species genetic variation changes through time, species’ heritabilities are also no longer constant, and are no longer equal across species.

The quantitative genetic assumption of infinitely many loci can be violated even more and one still gets similar results. Reducing the number of loci to \( n = 10 \), or even to a mere \( n = 5 \) (Figure 20, center and right panels), the trait distributions are still close to being normal. Since the possible
**Figure 20:** As Figure 19, except all species have an environmental component to their phenotypic variance. Only the genotypic trait values are plotted. The number of loci is $n = 25$ on the left, $n = 10$ in the center, and $n = 5$ on the right panel. The species-specific environmental standard deviations $\sigma_{E,i}$ were sampled uniformly and independently from the interval $[0.01, 0.05]$. Their values for the six surviving species on the left panel are: 0.029 (light blue), 0.034 (yellow), 0.017 (red), 0.046 (blue), 0.027 (green), 0.028 (purple). The $\sigma_{E,i}$ in the center are: 0.010 (yellow), 0.012 (green), and 0.013 (blue). The $\sigma_{E,i}$ on the right: 0.010 (yellow), and 0.014 (blue). Species now cannot evolve the exact same phenotype distribution, because even if their genetic makeup is identical, environmental noise will affect their traits differently (except in the nongeneric case of precisely equal environmental variances). Due to this fact, convergent evolution as in Figure 19 is only possible if there is a generalist-specialist distinction between the species (Section 3.2). Therefore, only a handful of species survive out of the initial 51, in line with model results in the quantitative genetic limit.

In conclusion, relaxing the assumption of infinitely many loci with infinitesimal additive effects does not seem to fundamentally alter the results obtained under the quantitative genetic limit, even when the number of loci is small and their phenotypic effects are large.

### 6 Competition along two independent trait dimensions

Up to this point we have assumed that a single ecologically relevant trait governs the interactions between individuals of the community. Here we perform a preliminary exploration of the effects of having multiple independent trait axes.

Let us assume there are two independent quantitative traits $z_1$ and $z_2$. For simplicity, we also assume that species’ trait distributions have the same variance $\sigma_i^2$ along both directions. The normalized bivariate trait distribution then reads

$$p_i(\vec{z}) = \frac{1}{2\pi\sigma_i^2} \exp\left(-\frac{||\vec{z} - \vec{\mu}_i||^2}{2\sigma_i^2}\right) = \frac{1}{2\pi\sigma_i^2} \exp\left(-\frac{(z_1 - \mu_{i1})^2 + (z_2 - \mu_{i2})^2}{2\sigma_i^2}\right), \quad (6.1)$$
where \( \vec{z} = (z_1, z_2) \) and \( \vec{\mu}_i = (\mu_{i1}, \mu_{i2}) \) are vectors in trait space, and \( \| \cdot \| \) is the Euclidean norm. Competition is assumed to depend only on distance, and so is given by

\[
a(\vec{z}, \vec{z}') = \exp\left(-\frac{\|\vec{z} - \vec{z}'\|^2}{\omega^2}\right) = \exp\left(-\frac{(z_1 - z'_1)^2 + (z_2 - z'_2)^2}{\omega^2}\right). \tag{6.2}
\]

We use the rectangular growth function in two dimensions:

\[
b(\vec{z}) = \begin{cases} 
1 & \text{if } \max(|z_1|, |z_2|) \leq \theta, \\
0 & \text{otherwise.} 
\end{cases} \tag{6.3}
\]

The fitness of an individual with phenotype \( \vec{z} \) is given by

\[
r(\vec{N}, \vec{p}, \vec{z}) = b(\vec{z}) - \sum_{j=1}^{S} N_j \int a(\vec{z}, \vec{z}') p_j(\vec{z}') d\vec{z}', \tag{6.4}
\]

the multidimensional generalization of Eq. 2.1 (here and onwards, we suppress the time-dependence of quantities for notational convenience).

The effective competition coefficients are still given by Eq. (2.4), but integrated over the whole trait space. Evaluating the integral yields

\[
\alpha_{ij} = \int \int p_i(\vec{z}) a(\vec{z}, \vec{z}') p_j(\vec{z}') d\vec{z} d\vec{z}' = \frac{\omega^2}{\omega^2 + 2\sigma_i^2 + 2\sigma_j^2} \exp\left(-\frac{(\mu_{i1} - \mu_{j1})^2 + (\mu_{i2} - \mu_{j2})^2}{\omega^2 + 2\sigma_i^2 + 2\sigma_j^2}\right). \tag{6.5}
\]

The effective growth rates are similarly given by Eq. (2.3) but integrated over \( \vec{z} \). With the rectangular growth function in Eq. (6.3), the integral gives

\[
b_i = \int b(\vec{z}) p_i(\vec{z}) d\vec{z} = \frac{1}{4} \left[ \text{erf}\left(\frac{\theta - \mu_{i1}}{\sqrt{2} \sigma_i}\right) + \text{erf}\left(\frac{\theta + \mu_{i1}}{\sqrt{2} \sigma_i}\right)\right] \left[ \text{erf}\left(\frac{\theta - \mu_{i2}}{\sqrt{2} \sigma_i}\right) + \text{erf}\left(\frac{\theta + \mu_{i2}}{\sqrt{2} \sigma_i}\right)\right]. \tag{6.6}
\]

The time evolution of the abundances follows the multidimensional generalization of Eq. (1.18):

\[
\frac{dN_i}{dt} = N_i \int r_i(\vec{N}, \vec{p}, \vec{z}) p_i(\vec{z}) d\vec{z} = N_i \left( b_i - \sum_{j=1}^{S} \alpha_{ij} N_j \right). \tag{6.7}
\]

We can obtain the multidimensional versions of \( \bar{b}_i \) (Eq. 2.7) and \( \bar{\beta}_{ij} \) (Eq. 2.8) analogously:

\[
\bar{b}_i = \int (\vec{z} - \vec{\mu}_i) p_i(\vec{z}) b(\vec{z}) d\vec{z} \]

\[
\bar{b}_i = \left( \frac{\sigma_i}{2\sqrt{2\pi}} \right) \exp\left(-\frac{(\theta + \mu_{i1})^2}{2\sigma_i^2}\right) \left[ 1 - \exp\left(\frac{2\theta \mu_{i1}}{\sigma_i^2}\right) \right] \left[ \text{erf}\left(\frac{\theta - \mu_{i2}}{\sqrt{2} \sigma_i}\right) + \text{erf}\left(\frac{\theta + \mu_{i2}}{\sqrt{2} \sigma_i}\right)\right], \tag{6.8}
\]

and

\[
\bar{\beta}_{ij} = \int (\vec{z} - \vec{\mu}_i) p_i(\vec{z}) a(\vec{z}, \vec{z}') p_j(\vec{z}') d\vec{z} d\vec{z}' \]

\[
\bar{\beta}_{ij} = -\frac{2\omega^2 \sigma_j^2 (\vec{\mu}_i - \vec{\mu}_j)}{\left(\omega^2 + 2\sigma_i^2 + 2\sigma_j^2\right)^2} \exp\left(-\frac{(\mu_{i1} - \mu_{j1})^2 + (\mu_{i2} - \mu_{j2})^2}{\omega^2 + 2\sigma_i^2 + 2\sigma_j^2}\right). \tag{6.9}
\]
The time evolution of the trait means $\bar{\mu}_i$ is given by the generalization of Eq. (1.23):

$$\frac{d\bar{\mu}_i}{dt} = G \int (\bar{\zeta} - \bar{\mu}_i) r_i(N, \bar{\mu}, \bar{\zeta}) p_i(\bar{\zeta}) d\bar{\zeta} = G \left( \bar{\zeta}_i - \sum_{j=1}^S \bar{\beta}_{ij} N_j \right), \quad (6.10)$$

where $G$ is the genetic variance-covariance matrix (Lande 1982). Since there are two traits, here $G$ is a $2 \times 2$ matrix.

This two-dimensional model can be simulated just like the original one. We use the simplest assumption of ecologically and genetically independent traits with equal heritability $h^2$ between species and along both trait axes. This means that the variance-covariance matrix is simply $G = h^2 I$ for all species, where $I$ is the $2 \times 2$ identity matrix. The initial number of species is $S = 51$, their initial population densities are all equal to one, and both coordinates of the initial trait positions are uniformly sampled from $[-\theta, \theta]$. We set $\omega = 0.2$, $\theta = 1/2$, and $h^2 = 0.1$ when species can evolve or $h^2 = 0$ when they cannot. Intraspecific standard deviations are uniformly sampled from $[0.05, 0.2]$ (in the presence of intraspecific variation), or are all set to $\sigma_i = 0.005$ (much smaller than $\theta$ and the competition width, effectively meaning the lack of any intraspecific variation).

Results of sample simulations are seen on Figures 21 (no intraspecific variation), 22 (nonheritable variation), and 23 (heritable variation). Results are very much in line with what we have seen for a single trait dimension. Species richness declines with increasing intraspecific variability, and further so with nonzero heritability. Mean traits in the zero heritability cases seem to be more evenly spaced than random, though this needs to be tested rigorously. However, there is a very clear even pattern when heritability is positive. Importantly, these patterns are only visible when the entire two-dimensional trait space is considered. Projecting the trait distributions onto just one of the axes, one sees lots of between-species overlap, and no pattern of even spacing at all. In fact, the opposite may be argued: species appear to be more clustered than random, which makes sense in light of the full two-dimensional trait distributions. Whether this signature of extra clustering is detectable in empirical data, or whether it would still be present if the number of trait axes was much larger than two, are open questions.

In summary, our results appear to be fundamentally unchanged by including multiple trait dimensions. As long as one looks for patterns in the whole trait space and not just a projection onto a single axis, the same expectations of reduced species richness and even spacing are found as before.

7 Code

The code we used to generate our results, along with documentation, can be accessed and downloaded either from the Ecology Letters website, or from https://github.com/dysordys/intravar.
Figure 21: Final state of community with two trait axes instead of just one. This is the case with no intraspecific variation; $h^2 = 0$. Out of the initial 51 species, 28 survive. Their traits are more or less evenly distributed in phenotypic space (top, with the full trait distributions on the left and mean trait positions from a bird's eye view, without abundance information, on the right). When this gets projected onto just a single trait axis at a time (bottom, with projection onto the first axis on the left and onto the second on the right), the patterns look much more random. Red dashed lines (not to scale) show the growth function $b(z)$. 
Figure 22: As Figure 21, but with species’ intraspecific standard deviations uniformly sampled from [0.05, 0.2]; $h^2 = 0$. Transparent disks on the top right mark one standard deviation of species’ trait distributions. Out of the initial 51 species, 18 survive. Again, a pattern that looks more even than random (top) is obscured by projecting onto a single trait axis (bottom).
Figure 23: As Figure 22, but with $h^2 = 0.1$. Out of the initial 51 species, 13 survive. This time the regular pattern is obvious in the whole trait space (top). However, this is still obscured when projected onto any single axis (bottom), where species appear to have substantial trait overlap.
References


