

In the format provided by the authors and unedited.

Coexistence of many species in random ecosystems

Carlos A. Serván^{1,5}, José A. Capitán^{1,2,5}, Jacopo Grilli¹ , Kent E. Morrison³ and Stefano Allesina^{1,4} *

¹Department of Ecology and Evolution, University of Chicago, Chicago, IL, USA. ²Department of Applied Mathematics, Universidad Politécnica de Madrid, Madrid, Spain. ³American Institute of Mathematics, San José, CA, USA. ⁴Northwestern Institute on Complex Systems, Northwestern University, Evanston, IL, USA. ⁵These authors contributed equally: Carlos A. Serván, José A. Capitán. *e-mail: sallesina@uchicago.edu

Coexistence of many species in random ecosystems

Supplementary Information

Carlos A. Serván^{1,†}, José A. Capitán^{1,2,†}, Jacopo Grilli¹, Kent E. Morrison³ and Stefano Allesina^{*1,4}

¹Department of Ecology & Evolution, University of Chicago, 1101 E. 57th Chicago, IL 60637, USA

²Department of Applied Mathematics, Universidad Politécnica de Madrid

³American Institute of Mathematics

⁴Northwestern Institute on Complex Systems, Northwestern University

[†] Equal contribution

S1 Mean zero

S1.1 Toy model: uncoupled logistic equations

Suppose that A is a diagonal matrix, and therefore that species do not interact with each other. For stability, we need $A_{ii} < 0$ for all i (self-regulation). Let p_i be the probability of $r_i > 0$. Then, the probability that a solution \mathbf{x} with k positive components $\{S\}_k$ is non-invasible is $\prod_{i \in \{S\}_k} p_i \prod_{i \notin \{S\}_k} (1 - p_i)$.

When the distribution of r_i is symmetric around zero, $p_i = \frac{1}{2}$ irrespective of the distribution of $A_{ii} < 0$, and thus the probability of non-invasibility is $\frac{1}{2^n}$ for any particular subsystem. Therefore, the binomial distribution with parameters n and $\frac{1}{2}$ describes the the number of persistent species.

S1.2 Feasibility

In this and the following section, we show that when the entries of matrix B and vector \mathbf{r} are random variables whose distribution is symmetric around 0, and that any n element subset of the columns of B and \mathbf{r} are linearly independent (which holds almost surely if the entries of

*sallesina@uchicago.edu

B and \mathbf{r} are sampled from a continuous probability distribution function and are independent of each other), then the probability $P(k|n)$ is still described by the binomial distributions with parameters n and $\frac{1}{2}$ —exactly what we found for non-interacting species. Note that this holds true both for the case in which the coefficients B_{ij} are sampled independently, and for the case in which these coefficients are sampled in pairs (B_{ij}, B_{ji}) , and the pairs are sampled independently from a bivariate distribution symmetric around $(0, 0)$.

First we show that $P(n|n) = \frac{1}{2^n}$. The proof amounts to showing that, of all the possible 2^n sign $(+, -)$ patterns for the entries of a solution to equation 6 of the main text, each of them is equally probable.

Let \mathbf{x}^* be an arbitrary solution of equation 6 of the main text, and define the matrix $D_k = ((-1)^{\delta_{ik}} \delta_{ij})$. Then, $D_k \mathbf{x}^*$ satisfies $(D_k A D_k) D_k \mathbf{x}^* = -D_k \mathbf{r}$. Because of the symmetry assumption, we have that $D_k A D_k$ has the same distribution¹ as A , and similarly for $D_k \mathbf{r}$ and \mathbf{r} . Since D_k just flips the sign of the k^{th} component of \mathbf{x}^* , by repeating this operation a sufficient number of times we can connect any two sign patterns of solutions to equation 6, and thus the conclusion follows.

S1.3 Persistent species

As noted before, in the regime of diagonally stable matrices, the final state of the system is the non-invasible (also called saturated) fixed point of the system¹. With the same assumptions of the previous section the distribution for the number of persistent species follows naturally: the probability of having a non-invasible solution \mathbf{x} with k positive components (with support $\{S\}_k$) is the joint probability of the conditions expressed in equations 8 and 9 of the main text, which can be written as $\mathcal{P}(\{S\}_k|n) = P(k|k)[1 - P_{\text{inv}}(\{S\}_n \setminus \{S\}_k | \{S\}_k)]$, where P_{inv} denotes the probability of being invasible by any of the remaining species given that $\mathbf{x} > 0$. Let $\mathbf{z} = \mathbf{r}^{(n)} + A^{(ns)} \mathbf{x}$. By following the same procedure illustrated in the previous section (applying the appropriate change of signs to A and \mathbf{r}), one can show that any sign pattern for \mathbf{z} is equally likely, therefore $1 - P_{\text{inv}}(\{S\}_n \setminus \{S\}_k | \{S\}_k) = \frac{1}{2^{n-k}}$. As a consequence, $\mathcal{P}(\{S\}_k|n) = \frac{1}{2^n}$. Because of the uniqueness of this type of solution for a given interaction matrix A and a vector

¹This transformation also has the property of preserving the eigenvalues of the matrix, which allows this argument to hold also if we condition on Lyapunov diagonally stable matrices.

of rates \mathbf{r} , the binomial distribution with parameters n and $\frac{1}{2}$ describes the distribution of the number of species having positive density at the globally stable equilibrium.

S1.4 Adding Structure

Let G be the adjacency matrix of an undirected graph, and consider the matrix $M = G \circ A$, where \circ represents the Hadamard (entry-wise) product between G and A . Because this type of product is commutative with respect to the multiplication by a diagonal matrix, i.e., $D(G \circ A)D = G \circ (DAD)$ for D diagonal, the arguments used in the previous two sections still hold. This means that the distribution of M is invariant to $D_k M D_k$ (even when G is also a random matrix) and by restricting ourselves to diagonally stable matrices the linear independence assumption is assured (the matrix is invertible). Consequently, adding a network structure in this way does not change the probability of feasibility nor the distribution of persistent species.

S2 Calculating the distribution of persistent species

If we integrate the GLV dynamics starting from an interaction matrix A , a vector of intrinsic growth rates \mathbf{r} , and an arbitrary (positive) initial condition with n species, we end up with k species with density different from zero and $n - k$ species with density equal to zero. If the matrix A is diagonally stable, the end point of the dynamics always correspond to a fixed point \mathbf{x}^* , irrespective of the initial conditions.

The goal of this section is to provide a formula for the probability $P(k|n)$ of finding k persisting species out of n , for an arbitrary matrix A , under the assumption that A is diagonally stable. We assume that the entries of \mathbf{r} are drawn from a Normal distribution with mean γ and unit variance. This choice of a variance does not affect the generality of our results, since the coexistence properties of the Generalized Lotka-Volterra equations are independent of the norm of \mathbf{r} : rescaling all growth rates by a constant simply rescales all equilibrium abundances by the same constant, with no impact on feasibility or stability.

We define the vector \mathbf{z} with $n - k$ components as

$$\mathbf{z} := \mathbf{r}^{(n)} + A^{(ns)} \mathbf{x}. \quad (\text{S1})$$

On the other hand, we have equation 8 of the main text, defining \mathbf{x} . By imposing feasibility and non-invasibility —equation 9 of the main text—, it must hold that $\mathbf{x} > 0$ and $\mathbf{z} < 0$.

Using the probability density of the growth rates,

$$P(\mathbf{r}) = \frac{1}{\sqrt{(2\pi)^n}} \exp\left(-\sum_{i=1}^n \frac{(r_i - \gamma)^2}{2}\right) = \frac{1}{\sqrt{(2\pi)^n}} \exp\left(-\frac{1}{2}\|\mathbf{r}^{(s)} - \gamma\mathbf{1}_k\|^2 - \frac{1}{2}\|\mathbf{r}^{(n)} - \gamma\mathbf{1}_{n-k}\|^2\right), \quad (\text{S2})$$

where $\mathbf{1}_k$ stands for a k –dimensional column vector whose entries are all equal to one. Introducing equation 8 of the main text and equation (S1), we can write the joint probability density as

$$f(\mathbf{x}, \mathbf{z}|A) = \frac{|\det \Lambda|}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}\|A^{(s)}\mathbf{x} + \gamma\mathbf{1}_k\|^2 - \frac{1}{2}\|\mathbf{z} - A^{(ns)}\mathbf{x} - \gamma\mathbf{1}_{n-k}\|^2\right), \quad (\text{S3})$$

where Λ is the Jacobian matrix obtained from the change of variables $\mathbf{r} \rightarrow (\mathbf{x}, \mathbf{z})$. According to equations 8 and (S1), it is simple to observe that Λ has the following structure:

$$\Lambda := \begin{pmatrix} \frac{\partial \mathbf{r}^{(s)}}{\partial \mathbf{x}} & \frac{\partial \mathbf{r}^{(s)}}{\partial \mathbf{z}} \\ \frac{\partial \mathbf{r}^{(n)}}{\partial \mathbf{x}} & \frac{\partial \mathbf{r}^{(n)}}{\partial \mathbf{z}} \end{pmatrix} = \begin{pmatrix} A^{(s)} & 0 \\ A^{(ns)} & I_{n-k} \end{pmatrix}, \quad (\text{S4})$$

I_{n-k} being the $(n - k)$ –dimensional identity matrix. Therefore $|\det \Lambda| = |\det A^{(s)}|$.

The first term appearing in the exponential in equation (S3) can be written as

$$\|A^{(s)}\mathbf{x} + \gamma\mathbf{1}_k\|^2 = (\mathbf{x} - \boldsymbol{\xi})^T G (\mathbf{x} - \boldsymbol{\xi}), \quad (\text{S5})$$

where

$$\boldsymbol{\xi} = -\gamma(A^{(s)})^{-1}\mathbf{1}_k, \quad (\text{S6})$$

and

$$G = (A^{(s)})^T A^{(s)}. \quad (\text{S7})$$

We obtain therefore

$$f(\mathbf{x}, \mathbf{z}|A) = \frac{|\det A^{(s)}|}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\xi})^T G (\mathbf{x} - \boldsymbol{\xi}) - \frac{1}{2}\|\mathbf{z} - A^{(ns)}\mathbf{x} - \gamma\mathbf{1}_{n-k}\|^2\right). \quad (\text{S8})$$

The probability $\mathcal{P}(\{S\}_k|A)$ of observing the globally stable fixed point with support $\{S\}_k$, can be obtained from the joint probability in equation (S3) by imposing the feasibility condition for the k species ($\mathbf{x} > 0$) and the non-invasibility condition for the other $n - k$ species ($\mathbf{z} < 0$). The equation reads

$$\mathcal{P}(\{S\}_k|A) \equiv \int d^k \mathbf{x} \left(\prod_{i=1}^k \Theta(x_i) \right) \int d^{n-k} \mathbf{z} \left(\prod_{j=k+1}^n \Theta(-z_j) \right) f(\mathbf{x}, \mathbf{z}|A) . \quad (\text{S9})$$

S3 Mean non zero

In this section we consider a simplified interaction matrix A whose diagonal coefficients are all equal to α , and all the off-diagonal elements are set to a fixed value μ :

$$A = (\alpha - \mu)I_n + \mu \mathbf{1}_n \mathbf{1}_n^T . \quad (\text{S10})$$

Since the matrix A is a deterministic matrix, in this case $\mathcal{P}(\{S\}_k|A) = \mathcal{P}(\{S\}_k|n)$. By introducing equation (S10) in equation (S8) and using equation (S9), we obtain

$$\begin{aligned} \mathcal{P}(\{S\}_k|n) &= \frac{|\det A^{(s)}|}{(2\pi)^{n/2}} \int d^k \mathbf{x} \prod_{i=1}^k \Theta(x_i) \int d^{n-k} \mathbf{z} \prod_{j=k+1}^n \Theta(-z_j) \\ &\quad \times \exp \left\{ -\frac{1}{2} \left(\mathbf{x} - \xi^{(k)} \mathbf{1}_k \right)^T G \left(\mathbf{x} - \xi^{(k)} \mathbf{1}_k \right) - \frac{1}{2} \left\| \mathbf{z} - (\mu(\mathbf{1}_k^T \mathbf{x}) + \gamma) \mathbf{1}_{n-k} \right\|^2 \right\}, \end{aligned} \quad (\text{S11})$$

where we used the fact that, with the parameterization of equation (S10), $\boldsymbol{\xi} = \xi^{(k)} \mathbf{1}_k$, where

$$\xi^{(k)} = -\frac{\gamma}{\alpha + (k-1)\mu} . \quad (\text{S12})$$

Again, using equation (S10) together with equation (S7), we have

$$G = (\alpha - \mu)^2 I_k + \left[k\mu^2 + 2\mu(\alpha - \mu) \right] \mathbf{1}_k \mathbf{1}_k^T . \quad (\text{S13})$$

We change variables to $x'_i = x_i - \xi^{(k)}$ to get

$$\begin{aligned} \mathcal{P}(\{S\}_k|n) &= \frac{|\det A^{(s)}|}{(2\pi)^{n/2}} \int d^k \mathbf{x} \prod_{i=1}^k \Theta(x_i + \xi^{(k)}) e^{-\frac{1}{2} \mathbf{x}^T G \mathbf{x}} \\ &\quad \times \int d^{n-k} \mathbf{z} \prod_{j=k+1}^n \Theta(-z_j) e^{-\frac{1}{2} \|\mathbf{z} - [\gamma + k\mu\xi^{(k)} + \mu(\mathbf{1}_k^T \mathbf{x})] \mathbf{1}_{n-k}\|^2}. \end{aligned} \quad (\text{S14})$$

We now write $z'_j = z_j - \gamma - k\mu\xi^{(k)}$ and obtain

$$\begin{aligned} \mathcal{P}(\{S\}_k|n) &= \frac{|\alpha - \mu|^{k-1} |\alpha + (k-1)\mu|}{(2\pi)^{n/2}} \int d^k \mathbf{x} \prod_{i=1}^k \Theta(x_i + \xi^{(k)}) e^{-\frac{1}{2} \mathbf{x}^T G \mathbf{x}} \\ &\quad \times \int d^{n-k} \mathbf{z} \prod_{j=k+1}^n \Theta(-z_j - \gamma - k\mu\xi^{(k)}) e^{-\frac{1}{2} \|\mathbf{z} - \mu(\mathbf{1}_k^T \mathbf{x}) \mathbf{1}_{n-k}\|^2}, \end{aligned} \quad (\text{S15})$$

where we used

$$|\det A^{(s)}| = |\alpha - \mu|^{k-1} |\alpha + (k-1)\mu|. \quad (\text{S16})$$

By introducing the expression for G obtained in equation (S13), we get

$$\begin{aligned} \mathcal{P}(\{S\}_k|n) &= \frac{|\alpha - \mu|^{k-1} |\alpha + (k-1)\mu|}{(2\pi)^{n/2}} \int d^k \mathbf{x} \prod_{i=1}^k \Theta(x_i + \xi^{(k)}) \\ &\quad \times \int d^{n-k} \mathbf{z} \prod_{j=k+1}^n \Theta(-z_j - \gamma - k\mu\xi^{(k)}) e^{g(\mathbf{x}, \mathbf{z})} \end{aligned} \quad (\text{S17})$$

where

$$g(\mathbf{x}, \mathbf{z}) = -\frac{1}{2} \left[(\alpha - \mu)^2 \mathbf{x}^T \mathbf{x} + [n\mu^2 + 2\mu(\alpha - \mu)] (\mathbf{1}_k^T \mathbf{x})^2 - 2\mu(\mathbf{1}_k^T \mathbf{x})(\mathbf{1}_{n-k}^T \mathbf{z}) + \mathbf{z}^T \mathbf{z} \right]. \quad (\text{S18})$$

We can express this probability as a double integral by introducing two new variables thanks to a Hubbard-Stratonovich transformation: if $b > 0$ and $c > 0$, it holds that

$$e^{-bd^2/c^2 - de/c} = \frac{c}{2\pi} \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dw e^{-(by^2 + ey + idw - icwy)}. \quad (\text{S19})$$

for any real d and e numbers. Similarly, for $b > 0$ and $c > 0$,

$$e^{-bd^2/c^2+de/c} = \frac{c}{2\pi} \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dw e^{-(by^2+ey+idw+icwy)}. \quad (\text{S20})$$

In our case [cf. equations (S13) and (S18)], we choose $d = \mathbf{1}_k^T \mathbf{x}$ and $e = \mathbf{1}_{n-k}^T \mathbf{z}$ and identify the exponents of the l.h.s. of equations (S19) or (S20) with the terms in (S18). If $\mu > 0$, we find $\frac{1}{c} = \mu$ and use equation (S20). If $\mu < 0$, we set $\frac{1}{c} = |\mu|$ and consider equation (S19). In both cases, we set $\frac{b}{c^2} = \frac{1}{2}[n\mu^2 + 2\mu(\alpha - \mu)]$. In general, we can choose $c = \frac{1}{|\mu|}$ and $b = \frac{1}{2}\left[n + 2\left(\frac{\alpha}{\mu} - 1\right)\right]$. To ensure diagonal stability, all the eigenvalues of matrix A must be negative. This implies the conditions $\alpha - \mu < 0$ and $\alpha - \mu + n\mu < 0$. If $\mu > 0$, the second restriction can be violated for n sufficiently large. Therefore we limit the discussion to the $\mu < 0$ case (competitive communities) and use equation (S19). In this case we have $\alpha < \mu < 0$ (hence $|\alpha| > |\mu|$) and $\frac{\alpha}{\mu} - 1 + n > 0$ (hence $b > 0$ and we can apply the Hubbard-Stratonovich transformation). Therefore

$$\begin{aligned} \mathcal{P}(\{S\}_k|n) &= \frac{|\alpha - \mu|^{k-1}|\alpha + (k-1)\mu|}{(2\pi)^{n/2+1}|\mu|} \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dw e^{-\frac{1}{2}\left[n+2\left(\frac{\alpha}{\mu}-1\right)\right]y^2+i\frac{yw}{|\mu|}} \int d^k \mathbf{x} \prod_{i=1}^k \Theta(x_i + \xi^{(k)}) \\ &\times \int d^{n-k} \mathbf{z} \prod_{j=k+1}^n \Theta(-z_j - \gamma - k\mu\xi^{(k)}) e^{-\frac{1}{2}(\alpha-\mu)^2 \mathbf{x}^T \mathbf{x} - i(\mathbf{1}_k^T \mathbf{x})w} e^{-\frac{1}{2}\mathbf{z}^T \mathbf{z} - (\mathbf{1}_{n-k}^T \mathbf{z})y}. \quad (\text{S21}) \end{aligned}$$

We complete squares and obtain

$$\begin{aligned} \mathcal{P}(\{S\}_k|n) &= \frac{|\alpha - \mu|^{k-1}|\alpha + (k-1)\mu|}{(2\pi)^{n/2+1}|\mu|} \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dw e^{-\frac{1}{2}\left[k+2\left(\frac{\alpha}{\mu}-1\right)\right]y^2+i\frac{yw}{|\mu|}} e^{-\frac{k}{2(\alpha-\mu)^2}w^2+\frac{1}{2}(n-k)y^2} \\ &\times \left[\int dx \Theta(x + \xi^{(k)}) e^{-\frac{1}{2}(\alpha-\mu)^2\left(x+\frac{iw}{(\alpha-\mu)^2}\right)^2} \right]^k \left[\int dz \Theta(-z - \gamma - k\mu\xi^{(k)}) e^{-\frac{1}{2}(z+y)^2} \right]^{n-k}. \quad (\text{S22}) \end{aligned}$$

Denoting the cumulative distribution function of the standard Normal distribution $N(0, 1)$ as $\Phi(x) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{x}{\sqrt{2}}\right)\right]$ we can write

$$\begin{aligned} \mathcal{P}(\{S\}_k|n) &= \frac{|\alpha - \mu|^{k-1}|\alpha + (k-1)\mu|}{2\pi|\mu||\alpha - \mu|^k} \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dw e^{-\frac{1}{2}\left[k+2\left(\frac{\alpha}{\mu}-1\right)\right]y^2+i\frac{yw}{|\mu|}-\frac{k}{2(\alpha-\mu)^2}w^2} \\ &\times \left[1 - \Phi\left(\frac{iw}{|\alpha - \mu|} - |\alpha - \mu|\xi^{(k)}\right) \right]^k \left[\Phi\left(y - \gamma - k\mu\xi^{(k)}\right) \right]^{n-k}, \quad (\text{S23}) \end{aligned}$$

and therefore we find

$$\begin{aligned} \mathcal{P}(\{S\}_k|n) = \frac{1}{2\pi} \left| k + \frac{\alpha}{\mu} - 1 \right| \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dw e^{-\frac{1}{2}[k+2(\frac{\alpha}{\mu}-1)]y^2 + i|\frac{\alpha}{\mu}-1|yw - \frac{1}{2}kw^2} \\ \times \left[1 - \Phi\left(iw - |\alpha - \mu|\xi^{(k)}\right) \right]^k \left[\Phi\left(y - \gamma - k\mu\xi^{(k)}\right) \right]^{n-k}. \end{aligned} \quad (\text{S24})$$

Note that $\gamma + k\mu\xi^{(k)} = \gamma\left(1 - \frac{k\mu}{\alpha+(k-1)\mu}\right) = \frac{\gamma(\alpha-\mu)}{\alpha+(k-1)\mu}$. We define $s := \frac{\alpha}{\mu} - 1$ (which satisfies $s > 0$ to ensure diagonal stability) and

$$v := \frac{\gamma(\alpha - \mu)}{\alpha - \mu + k\mu} = \frac{\gamma s}{k + s}. \quad (\text{S25})$$

Then, given that $\alpha < \mu$, it holds that $|\alpha - \mu|\xi^{(k)} = -\frac{\gamma|\alpha-\mu|}{\alpha+(k-1)\mu} = v$ and we can express the probability in its final form as

$$\mathcal{P}(\{S\}_k|n) = \frac{k+s}{2\pi} \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dw e^{-\frac{1}{2}(k+2s)y^2 + isyw - \frac{1}{2}kw^2} [1 - \Phi(iw - v)]^k [\Phi(y - v)]^{n-k}. \quad (\text{S26})$$

In this formula, the integration over w must be performed in the complex plane. An alternative way to express it is to consider a path Γ in the complex plane such that $\Gamma = \{w' \in \mathbb{C} | w' = iw + x_0\}$ and then reducing the result to the limit $x_0 \rightarrow 0$, so that the integral over the imaginary axis is well defined. Therefore, an equivalent form of writing this equation is

$$\mathcal{P}(\{S\}_k|n) = \frac{k+s}{2\pi i} \int_{-\infty}^{\infty} dy \int_{\Gamma} dw e^{-\frac{1}{2}(k+2s)y^2 + syw + \frac{1}{2}kw^2} [1 - \Phi(w - v)]^k [\Phi(y - v)]^{n-k}, \quad (\text{S27})$$

where the integral in w has to be evaluated over the contour Γ and then take the limit $x_0 \rightarrow 0$.

Note that for the case $k = 0$ the probability density of $\mathbf{x} = \mathbf{0}$ being non-invasible is simply

$$f(\mathbf{z}) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}(\mathbf{z} - \gamma \mathbf{1}_n)^T (\mathbf{x} - \gamma \mathbf{1}_n)} \quad (\text{S28})$$

and the condition for non-invasibility reduces to

$$\mathcal{P}(\emptyset|n) = P[z_1 < 0, \dots, z_n < 0] = \frac{1}{(2\pi)^{n/2}} \int d^n \mathbf{z} \prod_{i=1}^n \Theta(-z_i) e^{-\frac{1}{2}(\mathbf{z} - \gamma \mathbf{1}_n)^T (\mathbf{x} - \gamma \mathbf{1}_n)} = [\Phi(-\gamma)]^n. \quad (\text{S29})$$

In addition, for $k = 1$ the integral over w can be actually calculated. Using that

$$\int_{-\infty}^{\infty} dw e^{-\frac{1}{2}w^2 - iaw} [1 - \Phi(iw)] = \sqrt{2\pi} \Theta(-a) e^{-\frac{1}{2}a^2} \quad (\text{S30})$$

we get

$$\mathcal{P}(\{S\}_1|n) = \frac{s+1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dy \Theta(sy + v) e^{-\frac{1}{2}(s+1)^2 y^2} [\Phi(y - v)]^{n-1}, \quad (\text{S31})$$

or, alternatively,

$$\mathcal{P}(\{S\}_1|n) = \frac{1}{\sqrt{2\pi}} \int_{-\gamma}^{\infty} dy e^{-\frac{1}{2}y^2} \left[\Phi\left(\frac{\mu y}{\alpha} - v\right) \right]^{n-1}. \quad (\text{S32})$$

S3.1 Numerical evaluation of the double integral

Equation (S26) can be evaluated numerically via a Fast Fourier Transform (FFT). We can express it as

$$\mathcal{P}(\{S\}_k|n) = \frac{k+s}{2\pi} \int_{-\infty}^{\infty} dy e^{-\frac{1}{2}(k+2s)(y^2+2vy)+svy} [\Phi(y)]^{n-k} \hat{F}(-s(y+v) - kv; k) \quad (\text{S33})$$

where $\hat{F}(x; k)$ is the Fourier transform over w of the complex function

$$F(w; k) = e^{-\frac{1}{2}kw^2} [1 - \Phi(iw)]^k \quad (\text{S34})$$

and the Fourier Transform of the function $F(t; k)$ is defined as $\hat{F}(x; k) := \int_{-\infty}^{\infty} dt f(t; k) e^{-itx}$. Then we first calculate $\hat{F}(x; k)$ via a FFT algorithm. For that purpose, we assume that $f(t)$ is approximately equal to zero outside the interval $(-T/2, T/2)$ and sample t at m equally spaced points separated a distance $\delta = T/m$ (m is even), so that $t_j = (j - m/2)\delta$, $0 \leq j < m$. Then

$$\hat{F}(x_\ell; k) = \int_{-\infty}^{\infty} dt F(t; k) e^{-itx_\ell} \approx \int_{-T/2}^{T/2} dt F(t; k) e^{-itx_\ell} \approx \delta \sum_{j=0}^{m-1} F(t_j; k) e^{-itx_\ell}. \quad (\text{S35})$$

If $x_\ell = 2\pi(\ell - m/2)/T = 2\pi(\ell - m/2)/(m\delta)$, the last expression can be written in terms of the Discrete Fourier Transform, $D_\ell(\{z_j\}) = \sum_{j=0}^{m-1} z_j e^{-2\pi i j \ell / m}$, as

$$\hat{F}(x_\ell; k) = \delta e^{i\pi(\ell - m/2)} \sum_{j=0}^{m-1} F(t_j; k) e^{i\pi j(1 - 2\ell/m)} = (-1)^{\ell - m/2} \delta D_\ell[\{(-1)^j F(t_j; k)\}] \quad (\text{S36})$$

where $0 \leq \ell < m$. Once we have calculated $\hat{F}(x_\ell; k)$ over the set of sampling points, we interpolate to evaluate numerically the transform at an arbitrary point [see equation (S33)]. For numerical evaluation over a finite interval, equation (S33) is more conveniently expressed by changing to the variable $z = \Phi(y)$ as

$$\mathcal{P}(\{S\}_k|n) = \frac{k+s}{\sqrt{2\pi}} \int_0^1 dz \hat{F}\left(-s\left[\Phi^{-1}(z) + v\right] - kv; k\right) e^{-\frac{1}{2}[\Phi^{-1}(z)]\{2v(k+s)+[\Phi^{-1}(z)](k-1+2s)\}+(n-k)\log z}. \quad (\text{S37})$$

For $k = 1$ from (S32) we derive the expression

$$\mathcal{P}(\{S\}_1|n) = (s+1) \int_{\Phi(-\gamma)}^1 dz e^{-\frac{1}{2}(s+1)^2[\Phi^{-1}(z)+v]^2 + \frac{1}{2}[\Phi^{-1}(z)]^2 + (n-1)\log z}. \quad (\text{S38})$$

S3.2 Probability of coexistence

Assuming diagonal stability, the probability of observing k species in stable coexistence out of a pool of n species is given by

$$P(k|n) = \binom{n}{k} \mathcal{P}(\{S\}_k|n), \quad (\text{S39})$$

with $\mathcal{P}(\{S\}_k|n)$ given by equation (S27). We now approximate $\mathcal{P}(\{S\}_k|n)$ for large n in order to obtain an analytical formula for the distribution, as well as the mode of the distribution k^* .

We use the saddle point technique from statistical mechanics to evaluate integrals of the form $\int d^n \mathbf{u} e^{-nh(\mathbf{u})} k(\mathbf{u})$ for n large. We define q through $k = qn$ and regard q as a continuous, finite variable such that $0 \leq q \leq 1$. Then equation (S27) can be written as

$$\mathcal{P}(\{S\}_k|n) = \frac{k+s}{2\pi i} \int_{-\infty}^{\infty} dy \int_{\Gamma} dw e^{-sy^2 + syw} e^{-n\hat{h}(y,w;q,v)}, \quad (\text{S40})$$

where

$$\hat{h}(y,w;q) = \frac{q}{2}(y^2 - w^2) - q \log[1 - \Phi(w-v)] - (1-q) \log \Phi(y-v). \quad (\text{S41})$$

In the limit $n \rightarrow \infty$, we assume q to take a fixed value (which will be associated to any possible value that k can take in the range $0 \leq k \leq n$). To calculate the limit correctly, at this point we assume that interactions scale with n as $\mu = \hat{\mu}/n$. In this way, the total

interaction strength for any species is independent of n . Otherwise, since μ only enters in equation (S40) through the combination $s = \frac{\alpha}{\mu} - 1$, if we do not assume the scaling in the limit for $n \rightarrow \infty$ any dependence on interaction strengths will be lost for n large. Therefore we write $s = nu - 1$, where $u := \alpha/\hat{\mu}$ and equation (S40) becomes

$$\mathcal{P}(\{S\}_k|n) = \frac{k + nu - 1}{2\pi i} \int_{-\infty}^{\infty} dy \int_{\Gamma} dw e^{y^2 - yw} e^{-nh(y,w;\boldsymbol{\sigma})}, \quad (\text{S42})$$

where we use the shorthand $\boldsymbol{\sigma} := (q, u, v)$ and

$$h(y, w; \boldsymbol{\sigma}) = \frac{q}{2}(y^2 - w^2) - q \log[1 - \Phi(w - v)] - (1 - q) \log \Phi(y - v) + uy^2 - uyw. \quad (\text{S43})$$

In this limit of large n , the exponential function $e^{-nh(y,w;\boldsymbol{\sigma})}$ is very peaked around the global minimum of the real part of $h(y, w; \boldsymbol{\sigma})$. Then we can evaluate the integral by approximating the exponent up to second order around the minimum. Note also that w is a complex variable and h is an analytic function of w . Then the Cauchy-Riemann condition holds (i.e., the real part of h satisfies the Laplace equation) and the minimum of $\Re(h)$ calculated along the integration path Γ is given by the maximum of $\Re(h)$ when w is regarded as a real variable. Then we expect a saddle point in the real (y, w) plane.

The conditions for the critical point form a coupled system of non-linear equations for y and w as functions of $\boldsymbol{\sigma}$:

$$\begin{aligned} \frac{\partial h}{\partial y} &= qy - (1 - q) \frac{\Phi'(y - v)}{\Phi(y - v)} + 2uy - uw = qy - (1 - q) \frac{e^{-(y-v)^2/2}}{\sqrt{2\pi}\Phi(y - v)} + 2uy - uw = 0, \\ \frac{\partial h}{\partial w} &= -qw + q \frac{\Phi'(w - v)}{1 - \Phi(w - v)} - uy = -qw + q \frac{e^{-(w-v)^2/2}}{\sqrt{2\pi}[1 - \Phi(w - v)]} - uy = 0. \end{aligned} \quad (\text{S44})$$

This system can be solved numerically for each tuple $\boldsymbol{\sigma} = (q, u, v)$, yielding the functions $y^*(\boldsymbol{\sigma})$ and $w^*(\boldsymbol{\sigma})$ as the coordinates of the critical point. We now expand $h(y, w; \boldsymbol{\sigma})$ around these coordinates point up to second order. Using that $\Phi''(y - v) = -(y - v)\Phi'(y - v)$ and

the conditions (S44), we find

$$\begin{aligned}
\left. \frac{\partial^2 h}{\partial y^2} \right|_{\substack{y=y^* \\ w=w^*}} &= 2u + q + (1 - q) \left[y - v + \frac{\Phi'(y - v)}{\Phi(y - v)} \right] \frac{\Phi'(y - v)}{\Phi(y - v)} \Big|_{\substack{y=y^* \\ w=w^*}} \\
&= 2u + q + (2uy^* + qy^* - uw^*) \left(-v + \frac{y^* - u(w^* - 2y^*)}{1 - q} \right), \\
\left. \frac{\partial^2 h}{\partial w^2} \right|_{\substack{y=y^* \\ w=w^*}} &= -q + q \left[-w + v + \frac{\Phi'(w - v)}{1 - \Phi(w - v)} \right] \frac{\Phi'(w - v)}{1 - \Phi(w - v)} \Big|_{\substack{y=y^* \\ w=w^*}} \\
&= -q + (uy^* + qw^*) \left(v + \frac{uy^*}{q} \right), \\
\left. \frac{\partial^2 h}{\partial y \partial w} \right|_{\substack{y=y^* \\ w=w^*}} &= -u.
\end{aligned} \tag{S45}$$

In Section S4 we show that the critical point obtained by solving the coupled system (S44) is precisely a saddle point, as stated above. Therefore, up to second order around the saddle point,

$$h(y, w; \boldsymbol{\sigma}) \approx h(y^*, w^*; \boldsymbol{\sigma}) + \frac{1}{2} \left. \frac{\partial^2 h}{\partial y^2} \right|_{\substack{y=y^* \\ w=w^*}} (y - y^*)^2 + \frac{1}{2} \left. \frac{\partial^2 h}{\partial w^2} \right|_{\substack{y=y^* \\ w=w^*}} (w - w^*)^2 + \left. \frac{\partial^2 h}{\partial y \partial w} \right|_{\substack{y=y^* \\ w=w^*}} (y - y^*)(w - w^*). \tag{S46}$$

Substituting the expansion into equation (S42) and transforming the integral over Γ back into an integral over a real variable yields, up to first order in the asymptotic expansion of the exponent in powers of $1/n$, the following approximation for the probability $\mathcal{P}(\{S\}_k | n)$ that the support of the globally stable fixed point is $\{S\}_k$:

$$\mathcal{P}(\{S\}_k | n) = \frac{n(q + u) - 1}{\sqrt{K(\boldsymbol{\sigma}, n)}} e^{-nh(y^*, w^*; \boldsymbol{\sigma}) + y^*(y^* - w^*)}, \tag{S47}$$

with

$$\begin{aligned}
K(\boldsymbol{\sigma}, n) &:= (nu - 1)^2 + n^2 \left[-q + (uy^* + qw^*) \left(v + \frac{uy^*}{q} \right) \right] \\
&\times \left[\frac{2}{n} - 2u - q - (2uy^* + qy^* - uw^*) \left(-v + \frac{y^* - u(w^* - 2y^*)}{1 - q} \right) \right]. \tag{S48}
\end{aligned}$$

We can write equation (S47) as

$$\mathcal{P}(\{S\}_k|n) = \frac{n(q+u)-1}{\sqrt{K(\boldsymbol{\sigma}, n)}} e^{nH(\boldsymbol{\sigma})+G(\boldsymbol{\sigma})} \quad (\text{S49})$$

where

$$\begin{aligned} H(\boldsymbol{\sigma}) &:= \frac{q}{2}(w^{\star 2} - y^{\star 2}) + (1-q)\log[\Phi(y^{\star} - v)] + q\log[1 - \Phi(w^{\star} - v)] - uy^{\star 2} + uy^{\star}w^{\star}, \\ G(\boldsymbol{\sigma}) &:= y^{\star}(y^{\star} - w^{\star}). \end{aligned} \quad (\text{S50})$$

We now use the Stirling's approximation to get

$$\binom{n}{qn} \approx \frac{e^{-n[q\log q + (1-q)\log(1-q)]}}{\sqrt{2\pi nq(1-q)}}. \quad (\text{S51})$$

According to equation (S39), our approximation for the probability of coexistence is

$$P(k|n) = \frac{n(q+u)-1}{\sqrt{2\pi nq(1-q)}K(\boldsymbol{\sigma}, n)} e^{nF(\boldsymbol{\sigma})+G(\boldsymbol{\sigma})}, \quad (\text{S52})$$

where

$$\begin{aligned} F(\boldsymbol{\sigma}) &:= \frac{q}{2}(w^{\star 2} - y^{\star 2}) + (1-q)\log[\Phi(y^{\star} - v)] + q\log[1 - \Phi(w^{\star} - v)] \\ &\quad - uy^{\star 2} + uy^{\star}w^{\star} - q\log q - (1-q)\log(1-q), \end{aligned} \quad (\text{S53})$$

In the discrete distribution given by equation (S52) we have to set $k = qn$ for $0 \leq q \leq 1$ (i.e., $0 \leq k \leq n$). We can reproduce the original parameterization with non-scaled interspecific interactions (μ) by changing $\hat{\mu}$ back to $n\mu$, i.e, replacing the constant u by $\frac{\alpha}{n\mu}$.

S4 Classification of the critical point

In order to prove that the critical point (y^{\star}, w^{\star}) obtained as the solution of Eq. (S44) is a saddle point, we only have to show that the discriminant satisfies

$$D(y^{\star}, w^{\star}) = \left(\frac{\partial^2 h}{\partial y^2}\right)\left(\frac{\partial^2 h}{\partial w^2}\right) - \left(\frac{\partial^2 h}{\partial y \partial w}\right)^2 < 0, \quad (\text{S54})$$

where all the derivatives are evaluated at the critical point. From Eq. (S45) we observe that

$$D(y^*, w^*) = \left(\frac{\partial^2 h}{\partial y^2} \right) \left(\frac{\partial^2 h}{\partial w^2} \right) - u^2. \quad (\text{S55})$$

We now show that $\frac{\partial^2 h}{\partial y^2} \geq 0$ and $\frac{\partial^2 h}{\partial w^2} \leq 0$ at the critical point for any combination of parameters $\sigma = (q, u, v)$. This will complete the proof.

First, consider the expression in (S45) for $\frac{\partial^2 h}{\partial y^2}$. Since $u > 0$ (recall that we study the case $\alpha < \mu < 0$ and $u = \alpha/\hat{\mu} = \alpha/(n\mu) > 0$) and $0 \leq q \leq 1$, we can write

$$\left. \frac{\partial^2 h}{\partial y^2} \right|_{\substack{y=y^* \\ w=w^*}} \geq (2uy^* + qy^* - uw^*) \left(-v + \frac{y^* - u(w^* - 2y^*)}{1 - q} \right). \quad (\text{S56})$$

This product is positive or zero. On the one hand, according to (S44),

$$2uy^* + qy^* - uw^* = \frac{1 - q}{\sqrt{2\pi}} \frac{e^{-(y^*-v)^2/2}}{\Phi(y^* - v)}, \quad (\text{S57})$$

which is obviously a non-negative quantity. On the other hand, $y^* - u(w^* - 2y^*) = (1 - q)y^* + 2uy^* + qy^* - uw^*$, hence

$$-v + \frac{y^* - u(w^* - 2y^*)}{1 - q} = y^* - v + \frac{e^{-(y^*-v)^2/2}}{\sqrt{2\pi}\Phi(y^* - v)} = f_1(y^* - v), \quad (\text{S58})$$

where we have defined the function $f_1(x) = x + \frac{e^{-x^2/2}}{\sqrt{2\pi}\Phi(x)}$. It increases monotonically and, as $x \rightarrow -\infty$, $f_1(x) \approx -\frac{1}{x} > 0$. Therefore $f_1(x) > 0$ for all x and we have shown that $\frac{\partial^2 h}{\partial y^2} \geq 0$.

Now, from (S44) we obtain

$$uy^* + qw^* = \frac{q}{\sqrt{2\pi}} \frac{e^{-(w^*-v)^2/2}}{1 - \Phi(w^* - v)}. \quad (\text{S59})$$

Therefore we can express the term $v + uy^*/q$ that appears in Eq. (S45) as

$$v + \frac{uy^*}{q} = -(w^* - v) + \frac{e^{-(w^*-v)^2/2}}{\sqrt{2\pi}[1 - \Phi(w^* - v)]}. \quad (\text{S60})$$

Let us define the function

$$f_2(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}[1 - \Phi(x)]}. \quad (\text{S61})$$

Using the three equations above into (S45) we find

$$\left. \frac{\partial^2 h}{\partial w^2} \right|_{\substack{y=y^* \\ w=w^*}} = -q\{1 + f_2(w^* - v)[w^* - v - f_2(w^* - v)]\}. \quad (\text{S62})$$

Now we observe that the function

$$f_3(x) := 1 + f_2(x)[x - f_2(x)] \quad (\text{S63})$$

is equal to the derivative of $f_4(x) = x - f_2(x)$ with respect to x , $f_3(x) = f_4'(x)$. Therefore, to show that $\frac{\partial^2 h}{\partial w^2} \leq 0$ it is sufficient to see that $f_4(x)$ is a monotonically increasing function (hence $f_3(x) > 0$ and $\frac{\partial^2 h}{\partial w^2} = -qf_3(w^* - v) < 0$). A simple graphical analysis for $f_4(x)$ proves that this is indeed the case. As a consequence,

$$D(y^*, w^*) = \left(\frac{\partial^2 h}{\partial y^2} \right) \left(\frac{\partial^2 h}{\partial w^2} \right) - u^2 \leq -u^2 < 0 \quad (\text{S64})$$

and (y^*, w^*) is a saddle pont.

In summary, we have shown that the solution (y^*, w^*) of Eq. (S44) is a saddle point for the function $h(y, w; \sigma)$ defined in Eq. (S43), when w is regarded as a real variable. This implies, by the Cauchy-Riemman condition, that the real part of h has a minimum along the imaginary w axis (i.e, along the integration contour Γ). Since the saddle point is unique, it yields a global minimum for the exponent in the probability (S42) of finding the globally stable fixed point with support $\{S\}_k$.

S5 Mode of the distribution for large number of species

For large n , the mode of the distribution (S52) is recovered at a q^* value such that F takes its maximum value. We now calculate this q^* in the limits $\alpha/\hat{\mu} \gg 1$ (the mode has to be close to 1/2) and the ecological case $\alpha/\hat{\mu} \ll 1$.

First recall that, by definition [cf. equation (S25)], $v = \frac{\gamma s}{k+s}$. In the limit of large n , $v = \frac{\gamma u}{q+u}$ is a function of q , so we have to take into account this implicit dependence on q . We

take the derivative with respect to q on equation (S53),

$$\begin{aligned} \frac{\partial F}{\partial q} = & \frac{1}{2}(w^{*2} - y^{*2}) + q(w^*w^{*'} - y^*y^{*'}) - \log \Phi(y^* - v) + \log(1 - \Phi(w^* - v)) + (w^* - 2y^*)uy^{*'} \\ & + uy^*w^{*'} + (1 - q)(y^{*'} - v') \frac{\Phi'(y^* - v)}{\Phi(y^* - v)} - q(w^{*'} - v') \frac{\Phi'(w^* - v)}{1 - \Phi(w^* - v)} + \log \frac{1 - q}{q}. \end{aligned} \quad (\text{S65})$$

Now, according to equation (S44),

$$\begin{aligned} \frac{\Phi'(y^* - v)}{\Phi(y^* - v)} &= \frac{qy^* + 2uy^* - uw^*}{1 - q}, \\ \frac{\Phi'(w^* - v)}{1 - \Phi(w^* - v)} &= \frac{uy^* + qw^*}{q}, \end{aligned} \quad (\text{S66})$$

so the derivative with respect to q simplifies to

$$\frac{\partial F}{\partial q} = \frac{1}{2}(w^{*2} - y^{*2}) - v(w^* - y^*) - \log \Phi(y^* - v) + \log(1 - \Phi(w^* - v)) + \log \frac{1 - q}{q}. \quad (\text{S67})$$

Setting the derivative to zero yields the condition

$$(1 - q^*)e^{w^{*2}/2 - vw^*}[1 - \Phi(w^* - v)] = q^*e^{y^{*2}/2 - vy^*}\Phi(y^* - v), \quad (\text{S68})$$

where the functions $y^*(\sigma)$, $w^*(\sigma)$ and $v(q)$ are evaluated at $q = q^*$. On the other hand,

$$\begin{aligned} \frac{\Phi'(y^* - v)}{\Phi(y^* - v)} &= \frac{qy^* + 2uy^* - uw^*}{1 - q} = \frac{e^{-(y^* - v)^2/2}}{\sqrt{2\pi}\Phi(y^* - v)}, \\ \frac{\Phi'(w^* - v)}{1 - \Phi(w^* - v)} &= \frac{uy^* + qw^*}{q} = \frac{e^{-(w^* - v)^2/2}}{\sqrt{2\pi}[1 - \Phi(w^* - v)]}, \end{aligned} \quad (\text{S69})$$

hence

$$\begin{aligned} (1 - q)e^{-(y^* - v)^2/2} &= \sqrt{2\pi}\Phi(y^* - v)(qy^* + 2uy^* - uw^*), \\ qe^{-(w^* - v)^2/2} &= \sqrt{2\pi}[1 - \Phi(w^* - v)](uy^* + qw^*). \end{aligned} \quad (\text{S70})$$

Substituting these expressions into equation (S68) yields, after some algebra, this simple condition for the mode of the distribution, q^* :

$$y^*(q^*, u, v(q^*)) = w^*(q^*, u, v(q^*)). \quad (\text{S71})$$

Then, if this condition is satisfied, equation (S67) reduces to $\log \frac{1-\Phi(y^*-v)}{\Phi(y^*-v)} = \log \frac{q^*}{1-q^*}$, which implies

$$\Phi(y^* - v) = 1 - q^*. \quad (\text{S72})$$

From this we get

$$y^*(q^*, u, v(q^*)) = v(q^*) + \sqrt{2}\text{erf}^{-1}(1 - 2q^*). \quad (\text{S73})$$

Finally we take into account the last expression and use equation (S71) into equation (S44) to obtain

$$\sqrt{2}\gamma u + 2(q^* + u)\text{erf}^{-1}(1 - 2q^*) = \frac{e^{-[\text{erf}^{-1}(1-2q^*)]^2}}{\sqrt{\pi}} \quad (\text{S74})$$

which is a transcendental equation that determines the mode of the distribution $q^* = \frac{k^*}{n}$ as a function of interaction strengths and growth rates. Equivalently, the transcendental condition for the mode can be expressed as

$$\frac{\alpha}{\hat{\mu}} = \frac{e^{-[\Phi^{-1}(1-q^*)]^2/2} - \sqrt{2\pi}q^*\Phi^{-1}(1-q^*)}{\sqrt{2\pi}[\Phi^{-1}(1-q^*) + \gamma]}, \quad (\text{S75})$$

with $\Phi^{-1}(q) = \sqrt{2}\text{erf}^{-1}(2q - 1)$. A simple relation arises for the curve that separates left- and right-skewed distributions by choosing the mode to be $q^* = \frac{1}{2}$:

$$\frac{\alpha\gamma}{\hat{\mu}} = \frac{1}{\sqrt{2\pi}}. \quad (\text{S76})$$

In terms of the original (non-scaled) parameterization, this expression becomes

$$\frac{\alpha\gamma}{\mu} = \frac{n}{\sqrt{2\pi}} \quad (\text{S77})$$

via the substitution $\hat{\mu} \rightarrow n\mu$.

In the limit of small interaction strengths ($\hat{\mu} \ll \alpha$) of the mean zero case ($\gamma = 0$), condition (S75) reduces to

$$\frac{k^*}{n} \approx \frac{1}{2} - \frac{1}{2\pi} \frac{\hat{\mu}}{\alpha} + \frac{1}{4\pi} \left(\frac{\hat{\mu}}{\alpha} \right)^2, \quad (\text{S78})$$

which reproduces the expected (binomial) behavior.

S6 Truncated-Gaussian distributed rates

In this section we analyze the case that growth rates are drawn from a truncated Gaussian distribution,

$$P(\mathbf{r}) = \frac{1}{Z_n} \exp\left(-\sum_{i=1}^n \frac{(r_i - \gamma)^2}{2}\right) \prod_{j=1}^n \Theta(r_j), \quad (\text{S79})$$

so that every rate $r_j > 0$ for $j = 1, \dots, n$ (Z_n is a suitable normalization constant). Then we can express the probability $\mathcal{P}(\{S\}_k|A)$ of observing the globally stable fixed point with support $\{S\}_k$ in a simple form:

$$\mathcal{P}_T(\{S\}_k|A) \equiv \int d^k \mathbf{x} \left(\prod_{i=1}^k \Theta(x_i) \right) \int d^{n-k} \mathbf{z} \left(\prod_{j=k+1}^n \Theta(-z_j) \right) f_T(\mathbf{x}, \mathbf{z}|A). \quad (\text{S80})$$

where

$$f_T(\mathbf{x}, \mathbf{z}|A) = \frac{|\det A^{(s)}|}{Z_n} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\xi})^T G(\mathbf{x} - \boldsymbol{\xi}) - \frac{1}{2}\|\mathbf{z} - A^{(ns)}\mathbf{x} - \gamma\mathbf{1}_{n-k}\|^2\right) \\ \times \prod_{i=1}^k \Theta(-(A^{(s)}\mathbf{x})_i) \prod_{j=k+1}^n \Theta(z_j - (A^{(ns)}\mathbf{x})_j). \quad (\text{S81})$$

We focus on the rank-one competitive case: $A^{(s)} = (\alpha - \mu)I_k + \mu\mathbf{1}_k\mathbf{1}_k^T$, $A^{(ns)} = \mu\mathbf{1}_{n-k}\mathbf{1}_k^T$ for $\alpha < \mu < 0$. Then

$$(A^{(s)}\mathbf{x})_i = (\alpha - \mu)x_i + \mu(\mathbf{1}_k^T \mathbf{x}) = \alpha x_i + \mu \sum_{\substack{s=1 \\ s \neq i}}^k x_s. \quad (\text{S82})$$

Since Eq. (S80) forces that $x_i > 0$, and α and μ are both negative, we find that $-(A^{(s)}\mathbf{x})_i$ is always positive, i.e., it holds that

$$\Theta(-(A^{(s)}\mathbf{x})_i)\Theta(x_i) = \Theta(x_i). \quad (\text{S83})$$

On the other hand, $\mu(\mathbf{1}_k^T \mathbf{x}) < 0$ and we can express

$$\Theta(z_j - (A^{(ns)}\mathbf{x})_j)\Theta(-z_j) = \Theta(z_j - \mu(\mathbf{1}_k^T \mathbf{x})) + \Theta(-z_j) - 1. \quad (\text{S84})$$

Now, we apply the same changes of variable leading to Eq. (S15). Then we can write

$$\begin{aligned} \mathcal{P}_T(\{S\}_k|n) &= \frac{|\alpha - \mu|^{k-1}|\alpha + (k-1)\mu|}{Z_n} \int d^k \mathbf{x} \prod_{i=1}^k \Theta(x_i + \xi^{(k)}) e^{-\frac{1}{2} \mathbf{x}^T G \mathbf{x}} \\ &\times \int d^{n-k} \mathbf{z} \prod_{j=k+1}^n \left[\Theta(-z_j - \gamma - k\mu\xi^{(k)}) + \Theta(z_j - \mu(\mathbf{1}_k^T \mathbf{x}) + \gamma) - 1 \right] e^{-\frac{1}{2} \|\mathbf{z} - \mu(\mathbf{1}_k^T \mathbf{x}) \mathbf{1}_{n-k}\|^2}. \end{aligned} \quad (\text{S85})$$

Let $\mathcal{K} = \{1, \dots, n-k\}$. Expanding the product we get

$$\begin{aligned} &\prod_{j=k+1}^n \left\{ \Theta(-z_j - \gamma - k\mu\xi^{(k)}) + \left[\Theta(z_j - \mu(\mathbf{1}_k^T \mathbf{x}) + \gamma) - 1 \right] \right\} \\ &= \sum_{\ell=0}^{n-k} \sum_{\substack{p \in C_\ell^{n-k} \\ b = \mathcal{K} \setminus p}} \prod_{j=1}^{\ell} \Theta(-z_{p(j)+k} - \gamma - k\mu\xi^{(k)}) \prod_{i=1}^{n-k-\ell} \left[\Theta(z_{b(i)+k} - \mu(\mathbf{1}_k^T \mathbf{x}) + \gamma) - 1 \right], \end{aligned} \quad (\text{S86})$$

where $p = (p(1), \dots, p(\ell))$ is a combination of ℓ elements taken from \mathcal{K} , $p \in C_\ell^{n-k}$, and b is formed by the remaining elements of the set, $b = \{1, \dots, n-k\} \setminus p$. Without loss of generality, since integrals are invariant under changes of indices in variable \mathbf{z} , we can decompose

$$\begin{aligned} \mathcal{P}_T(\{S\}_k|n) &= \frac{|\alpha - \mu|^{k-1}|\alpha + (k-1)\mu|}{Z_n} \int d^k \mathbf{x} \prod_{i=1}^k \Theta(x_i + \xi^{(k)}) e^{-\frac{1}{2} \mathbf{x}^T G \mathbf{x}} \\ &\times \sum_{\ell=0}^{n-k} \binom{n-k}{\ell} \prod_{j=k+1}^{\ell+k} \int dz_j \Theta(-z_j - \gamma - k\mu\xi^{(k)}) e^{-\frac{1}{2} [z_j - \mu(\mathbf{1}_k^T \mathbf{x})]^2} \\ &\times \prod_{i=\ell+k+1}^n \int dz_i \left[\Theta(z_i - \mu(\mathbf{1}_k^T \mathbf{x}) + \gamma) - 1 \right] e^{-\frac{1}{2} [z_i - \mu(\mathbf{1}_k^T \mathbf{x})]^2}. \end{aligned} \quad (\text{S87})$$

Note now that

$$\int_{-\infty}^{\infty} dz \left[\Theta(z - \mu(\mathbf{1}_k^T \mathbf{x}) + \gamma) - 1 \right] e^{-\frac{1}{2} [z - \mu(\mathbf{1}_k^T \mathbf{x})]^2} = -\sqrt{2\pi} \Phi(-\gamma). \quad (\text{S88})$$

Therefore we can decompose $\mathcal{P}(\{S\}_k|n)$ as the sum

$$\mathcal{P}_T(\{S\}_k|n) = \frac{(2\pi)^{n/2}}{Z_n} \sum_{\ell=0}^{n-k} \binom{n-k}{\ell} [-\Phi(-\gamma)]^{n-\ell-k} \mathcal{P}(\{S\}_k|\ell+k), \quad (\text{S89})$$

where $\mathcal{P}(\{S\}_k|n)$ is precisely the expression (S15) obtained for the non-truncated Gaussian distribution. According to Eq. (S27),

$$\mathcal{P}(\{S\}_k|\ell+k) = \frac{k+s}{2\pi i} \int_{-\infty}^{\infty} dy \int_{\Gamma} dw e^{-\frac{1}{2}(k+2s)y^2 + syw + \frac{1}{2}kw^2} [1 - \Phi(w-v)]^k [\Phi(y-v)]^\ell. \quad (\text{S90})$$

We introduce (S90) into (S89) and use the binomial expansion

$$\sum_{\ell=0}^{n-k} \binom{n-k}{\ell} [-\Phi(-\gamma)]^{n-\ell-k} [\Phi(y-v)]^\ell = [\Phi(y-v) - \Phi(-\gamma)]^{n-k} \quad (\text{S91})$$

to get the probability $\mathcal{P}_T(\{S\}_k|n)$ expressed as a double integral,

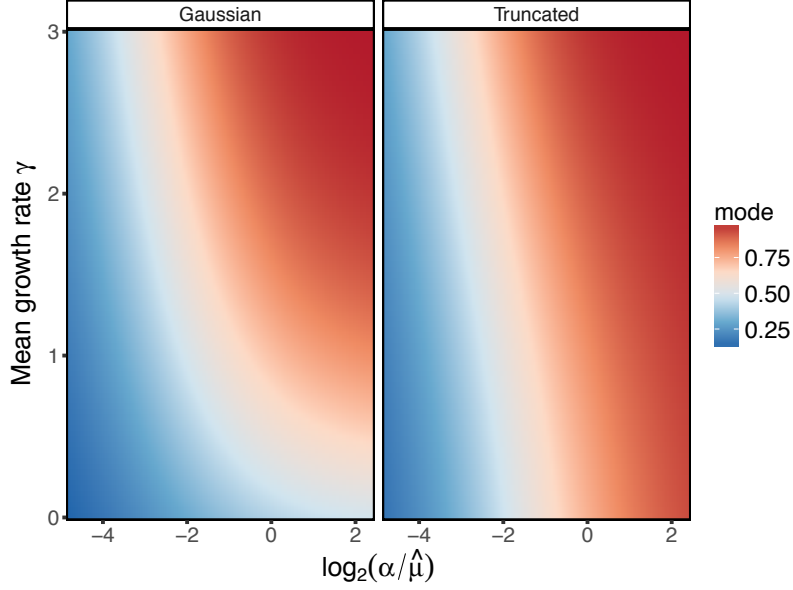
$$\begin{aligned} \mathcal{P}_T(\{S\}_k|n) &= \frac{(2\pi)^{n/2-1}(k+s)}{i Z_n} \int_{-\infty}^{\infty} dy \int_{\Gamma} dw e^{-\frac{1}{2}(k+2s)y^2 + syw + \frac{1}{2}kw^2} \\ &\quad \times [1 - \Phi(w-v)]^k [\Phi(y-v) - \Phi(-\gamma)]^{n-k}. \end{aligned} \quad (\text{S92})$$

Note that the only difference with Eq. (S27) is the term $\Phi(-\gamma)$ that appears in the last factor of the integrand. Hence we can easily extend the saddle-point calculation for the truncated Gaussian case. The probability $P_T(k|n) = \binom{n}{k} \mathcal{P}_T(\{S\}_k|n)$ that the support has cardinality k in this case can be written, up to a normalization factor and sub-leading corrections, as $P_T(k|n) \sim e^{nF_T(\sigma)}$, where

$$\begin{aligned} F_T(\sigma) &:= \frac{q}{2} (w^{\star 2} - y^{\star 2}) + (1-q) \log[\Phi(y^\star - v) - \Phi(-\gamma)] \\ &\quad + q \log[1 - \Phi(w^\star - v)] - uy^{\star 2} + uy^\star w^\star - q \log q - (1-q) \log(1-q). \end{aligned} \quad (\text{S93})$$

We can compare the mode of the distribution for the truncated and the purely Gaussian cases. The calculation of the mode follows the same steps of the Gaussian case. The equations for the saddle point (y^\star, w^\star) are now

$$\begin{aligned} qy - (1-q) \frac{\Phi'(y-v)}{\Phi(y-v) - \Phi(-\gamma)} + 2uy - uw &= 0, \\ qw - q \frac{\Phi'(w-v)}{1 - \Phi(w-v)} + uy &= 0. \end{aligned} \quad (\text{S94})$$



Supplementary Figure 1: Comparison between the modes for purely Gaussian and truncated-Gaussian distributed growth rates.

As can be easily checked, the condition $\frac{\partial F_T}{\partial q} = 0$ to be satisfied by the mode q^* leads to the same constraint as in the Gaussian case, $y^*(q^*, u, v(q^*)) = w^*(q^*, u, v(q^*))$, see Eq. (S71). This implies that

$$\Phi(y^* - v) = 1 - q^* + q^* \Phi(-\gamma), \quad (\text{S95})$$

which reduces to the Gaussian-case condition for the mode in the limit of large γ , where both the truncated and the Gaussian distributions tend to almost overlap. Finally, after the same algebraic manipulations in the condition above we obtain the following non-linear equation that determines the mode in the truncated-Gaussian case:

$$[1 - \Phi(-\gamma)] \left[\sqrt{2}\gamma u + 2(q^* + u) \text{erf}^{-1}(1 - 2q^* + 2q^* \Phi(-\gamma)) \right] = \frac{1}{\sqrt{\pi}} e^{-[\text{erf}^{-1}(1 - 2q^* + 2q^* \Phi(-\gamma))]^2}. \quad (\text{S96})$$

Supplementary Figure 1 shows the most probable number of coexisting species obtained for the Gaussian and the truncated Gaussian distributions as function of the parameters γ and $\alpha/\hat{\mu}$. We observe that the expected values for both cases are roughly the same for $\gamma \gtrsim 1$.

S7 Final communities

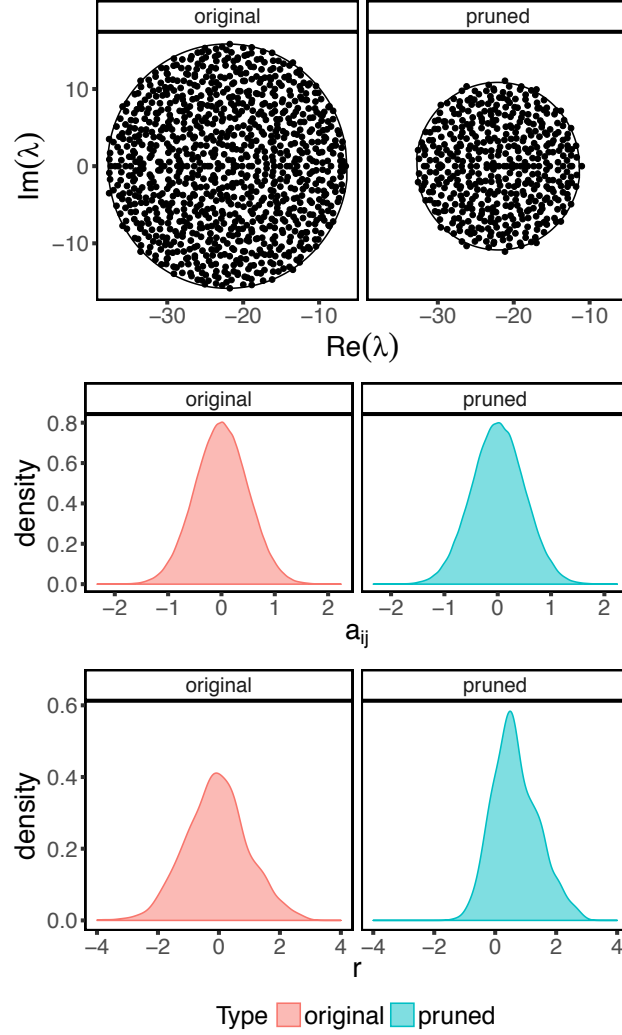
Supplementary Figure 2 shows the properties of the parameters of the communities found after the dynamical pruning, for an starting community of 1000 species, and a final community comprising 472 species. As proposed recently², the matrix of interactions in the pruned community is a random subset of the original. On the other hand, the distribution of growth rates changes in a nontrivial way, with a larger mean and positive skewness. This change, as explained in the discussion, is related to the interplay between the negative diagonal that we need to add to the matrix in order to ensure stability and the sign and magnitude of the interspecific interactions that each species is assigned, which in the end pushes the growth rates towards the right. As a limiting case to this behavior one can see that if the matrix is purely diagonal, then because we require stability all its entries are negative and the distribution of growth rates will be a truncated version of the original one.

S8 Numerical simulations

In this section, we detail the numerical simulations we used to corroborate our argument, and extend it to cases in which a direct analytic computation is unfeasible. We start by illustrating the Lemke-Howson algorithm that can be used to efficiently search for the non-invasible solution. Applying this algorithm, we were able to determine the non-invasible solution of a system without the need to integrate the dynamics numerically. Then, we detail the parameters for the numerical simulations—how are the matrices constructed, and how the network structure is introduced.

S8.1 Lemke-Howson algorithm

Because of the equivalence between the Lotka-Volterra and the replicator equation⁵, the non-invasible solution in the diagonally stable regime *is* the *unique* symmetric Nash equilibrium for the replicator dynamics in which the last element of the solution is played with non-zero probability¹ (this last element can be interpreted as “the environment” when moving from LV with n equations to a replicator system with $n + 1$ equations). We use the Lemke-Howson algorithm⁶ to find such a solution. This algorithm is based on exploring the vertices of the



Supplementary Figure 2: Comparison between the properties of the original community with $n = 1000$ and the final community, after dynamical pruning, comprising $n = 472$ species. The first row shows the eigenvalue distribution of the matrix of interaction A ; for a matrix in which the entries are i.i.d. samples from a distribution, we expect the eigenvalues to be approximately uniformly distributed in a circle in the complex plane, whose radius depends on the size of the system and the variance of the distribution^{3,4}. In the second panel, we show that indeed the distribution of the off-diagonal elements of A is the same before/after dynamics. Finally, in the third panel we show that instead the distribution of growth rates changes non trivially.

following polytope:

$$P = \{\mathbf{z} \in \mathbb{R}^n | \mathbf{z} \geq \mathbf{0}, C\mathbf{z} \leq \mathbf{1}\}, \quad (\text{S97})$$

where C is a positive payoff matrix of an $n \times n$ symmetric game—the positivity of the payoffs can be assumed without loss of generality, because adding a suitable constant to all the elements of the payoff matrix does not affect the dynamics.

We say that $\mathbf{z} \in P$ has label k if $z_k = 0$ and label $-k$ if $(C\mathbf{z})_k = 1$. Let us assume that P is simple (which holds almost surely in the cases we explore), that is, each vertex is adjacent to exactly n facets—a facet is defined by setting to equality one of the inequalities defining the polytope. Say that \mathbf{z} represents strategy k if either it has label k or $-k$, then because of the simplicity assumption any \mathbf{z} that represents all strategies is either $\mathbf{0}$ or the normalized vector $\hat{\mathbf{z}} = \mathbf{z} / \sum_i z_i$ is a *symmetric Nash equilibrium* for the game.

In order to find the solution we move around the vertices of P starting from $\mathbf{v}_0 = \mathbf{0}$ using a tableaux $T : \mathbf{r} = \mathbf{1} - C\mathbf{z}$ with a slack variable \mathbf{r} . Say that r_k is in the basis for a vertex $\mathbf{v} \in P$ if and only if \mathbf{v} does not have label $-k$, and z_k is in the basis if and only if \mathbf{v} does not have label k . Then \mathbf{v}_0 has basis $\{r_1, \dots, r_n\}$, bring z_n to the basis and by the min. ratio rule—i.e., by looking at the ratio between the free variable (in this case 1) and the coefficients of z_n in the tableaux—choose r_k to leave the basis and proceed to an adjacent vertex \mathbf{v}_1 . In the next iteration bring z_k to the basis and move to an adjacent vertex \mathbf{v}_2 . We keep repeating this process until we get to a vertex \mathbf{v} which represents all strategies, that is, \mathbf{v} is a Nash equilibrium which moreover will have z_n in the basis (since by construction the process will stop when the element leaving the basis is r_n). Because of the simplicity assumption the process is going to terminate, having to do in the worst case 2^n iterations. As it often happens, this worst-case scenario is never found in practice, making the algorithm efficient.

Let us illustrate this ideas by a simple example. Take the Lotka-Volterra system with interactions

$$A = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}, \quad (\text{S98})$$

and intrinsic growth rates:

$$\mathbf{r} = \begin{pmatrix} -1 \\ 3 \end{pmatrix} \quad (\text{S99})$$

We build the payoff matrix:

$$C = \begin{pmatrix} -2 & 1 & -1 \\ 1 & -2 & 3 \\ 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 4 & 2 \\ 4 & 1 & 6 \\ 3 & 3 & 3 \end{pmatrix}, \quad (\text{S100})$$

where we have added a constant to all entries to make them all positive. At the beginning of the algorithm we have the following tableaux:

$$\begin{aligned} r_1 &= 1 - z_1 - 4z_2 - 2z_3, \\ r_2 &= 1 - 4z_1 - z_2 - 6z_3, \\ r_3 &= 1 - 3z_1 - 3z_2 - 3z_3. \end{aligned} \quad (\text{S101})$$

We now bring z_3 into the basis, and by the min. ratio rule: the ratio of 1 and the coefficients of z_3 , r_2 should leave the basis and the updated tableaux is:

$$\begin{aligned} r_1 &= \frac{2}{3} + \frac{1}{3}z_1 - \frac{11}{3}z_2 + \frac{1}{3}r_2, \\ z_3 &= \frac{1}{6} - \frac{4}{6}z_1 - \frac{1}{6}z_2 - \frac{1}{6}r_2, \\ r_3 &= \frac{1}{2} - z_1 - \frac{5}{2}z_2 + \frac{1}{2}r_2. \end{aligned} \quad (\text{S102})$$

Now z_2 enters the basis, and in this case r_1 leaves from the basis:

$$\begin{aligned} z_2 &= \frac{2}{11} + \frac{1}{11}z_1 - \frac{3}{11}r_1 + \frac{1}{11}r_2, \\ z_3 &= \frac{3}{22} - \frac{15}{22}z_1 + \frac{1}{22}r_1 - \frac{2}{11}r_2, \\ r_3 &= \frac{1}{22} - \frac{27}{22}z_1 + \frac{15}{22}r_1 + \frac{3}{11}r_2. \end{aligned} \quad (\text{S103})$$

We bring z_1 into the basis and then we are done because r_3 leaves the basis in this case. So the Nash equilibrium for this game has full support. The final state of the tableaux is :

$$\begin{aligned} z_1 &= \frac{1}{27} + \frac{15}{27}r_1 + \frac{2}{9}r_2 - \frac{22}{27}r_3, \\ z_2 &= \frac{5}{27} - \frac{6}{27}r_1 + \frac{1}{9}r_2 - \frac{2}{27}r_3, \\ z_3 &= \frac{1}{9} - \frac{1}{3}r_1 - \frac{1}{3}r_2 + \frac{15}{27}r_3. \end{aligned} \tag{S104}$$

By normalizing the free elements in the final tableaux we also get the values at equilibrium, which in this case is $(1/9, 5/9, 3/9)$. Because the last element is positive, then the two species coexist, the second with an equilibrium value that is five times as large as the first.

S8.2 Sampling the matrices and growth rates

In the following we give the details of the construction of the matrices and growth rates for the cases we explored. For each case we repeat the process 2×10^5 times.

S8.2.1 Mean zero

We sample the entries of B in pairs, (B_{ij}, B_{ji}) for $j \neq i$ from a bivariate Normal distribution $N(\mathbf{0}, \Sigma)$ where Σ is a covariance matrix with diagonal 1 and off-diagonal ρ . The diagonal elements B_{ii} are chosen from a standard Normal distribution $N(0, 1)$. We then calculate the leading eigenvalue of $B + B^T$: $\lambda_M = \max_{\lambda}(\Re(\lambda(B + B^T)))$. We define $A = B - dI$, where d is a constant sufficient to make $A + A^T$ negative definite. More precisely, we choose $d = -\lambda_M - 10^{-6}$ (so that the matrix $A + A^T$ is barely stable). The entries of \mathbf{r} are sampled from a standard Normal distribution $N(0, 1)$.

S8.2.2 Mean non zero

In this case, the entries of the matrix are fixed and we choose each entry of \mathbf{r} from a Normal distribution $N(\gamma, 1)$.

S8.2.3 Adding Structure

In order to include a network structure, we generate an adjacency matrix G with a desired connectance level C (we used $C = 0.1$ and $C = 0.25$) and all diagonal elements set to one. In

the case of a power-law structure, we use the `sample_fitness_pl` function from the `igraph` package in R with an exponent of 2. For the modular and bipartite structures we split the matrix in two blocks, and arrange the connectance levels within and among them such that one is higher than the other—in particular we require two parameters b_r and c_r that determine the ratio of the size among the blocks and the ratio of the connectance within and among blocks (e.g. $c_r > 1$ for a modular structure). The values used were $b_r = 1/3$ for both cases, with $c_r = 3$ for modular, and $c_r = 1/3$ for bipartite. This adjacency matrix is then multiplied element-wise to our original matrix. The results are presented in Figure 1 in the main text as well as in Supplementary Figure 3.

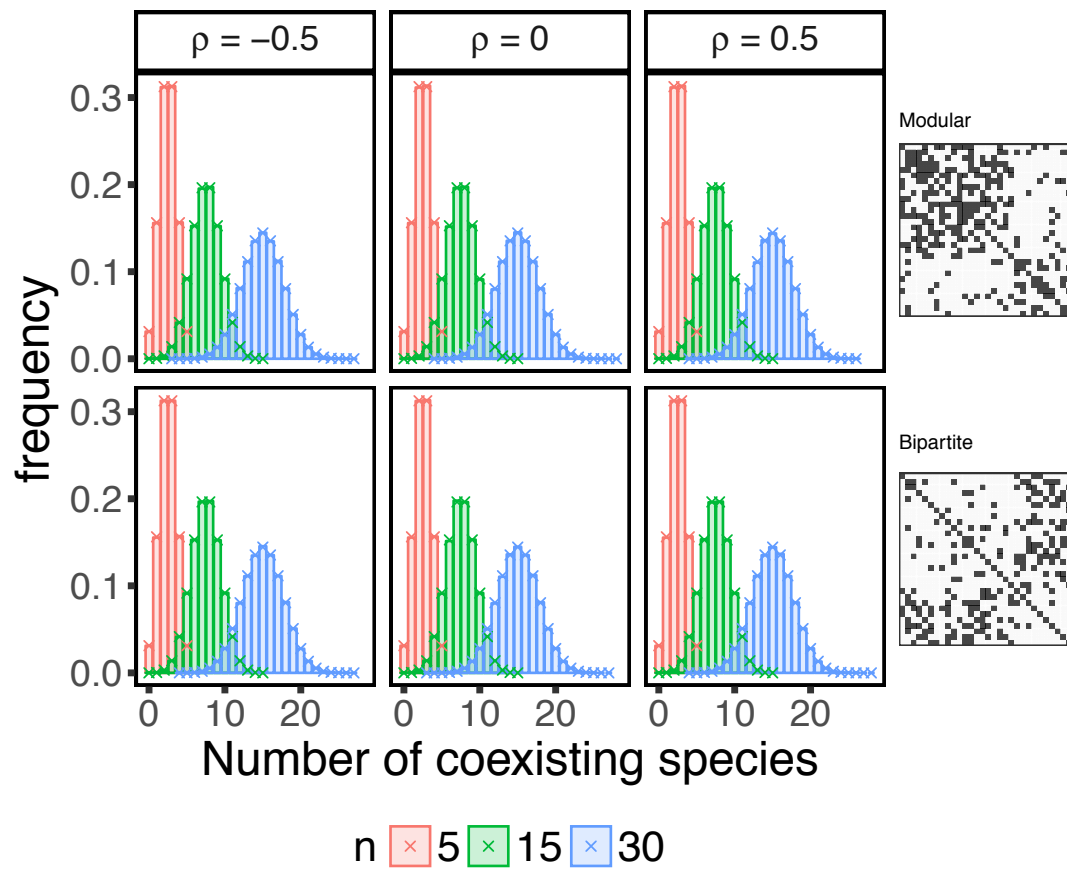
In the mean-zero case the matrix is made negative definite by the same process described above.

In the mean non-zero case the fully connected matrix is by construction negative definite ($\alpha < \mu < 0$) but when we add structure we need to restrict the values of μ that keep the negative definiteness.

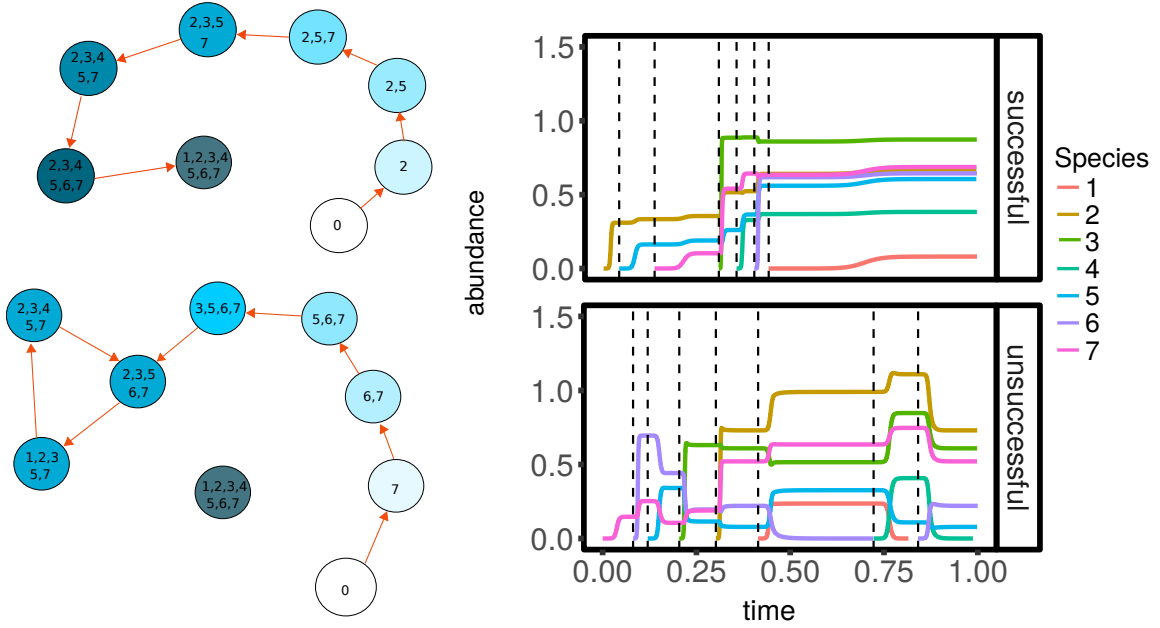
The prediction shown in Figure 3 of the main text is the mode of a fully connected system using the rescaled μ : $\hat{\mu} = n\mu C$.

S9 Assembly

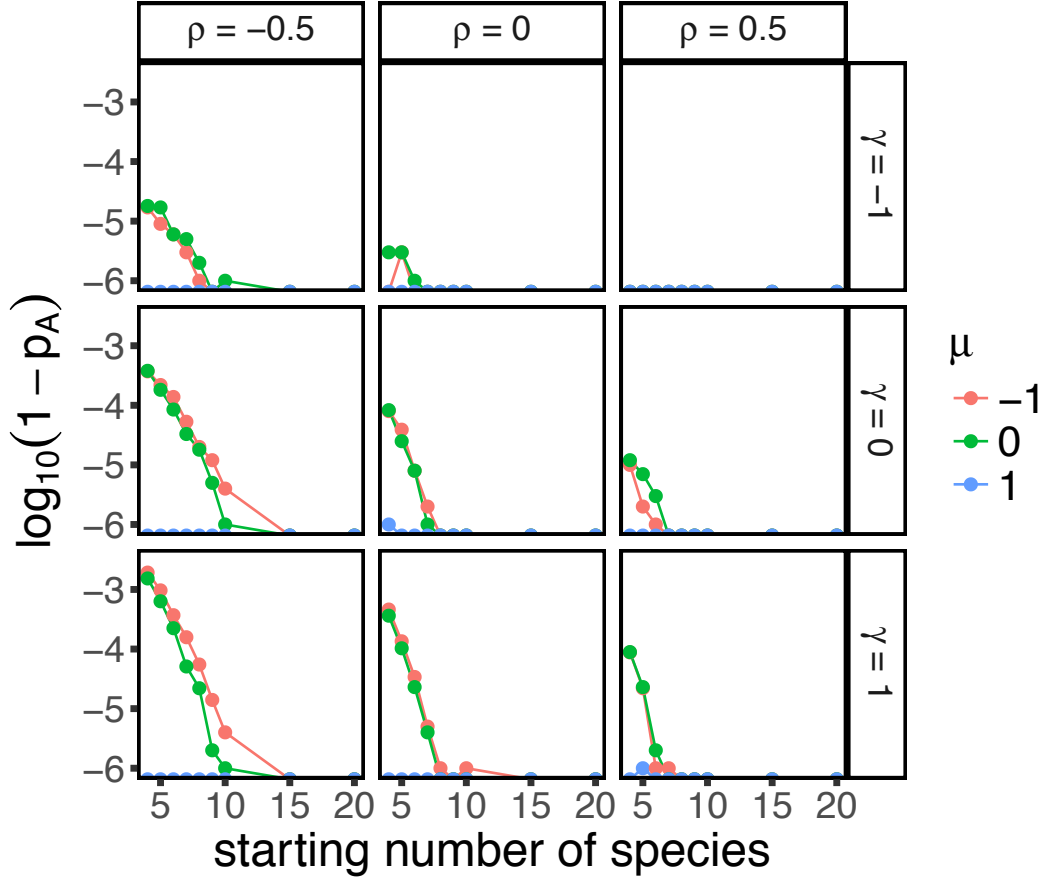
So far, we have described the dynamical process associated with equation 1 of the main text when starting with all n species present. A different view of the problem is to take our original n -dimensional system as a *species pool*, and from that derive the possible states to which one can arrive by adding *one species at a time*. This define a directed graph in which the nodes are the feasible states, and the edges represent invasion events connecting the two states (a subset of which is shown for example in Supplementary Figure 2). In this section we present numerical evidence that suggests that, in the regime of diagonal stability, one can find sets of persistent species satisfying equation 8 of the main text which cannot be assembled (Supplementary Figure 4). In such cases, our end-state with k species cannot be built by adding a species at a time. The probability of finding such a case, however, decreases rapidly with k : when our final community has many species, the probability of finding at least one assembly pathway to build the community approaches one (Supplementary Figure 5).



Supplementary Figure 3: As Figure 1 of the main text, but with modular (top) or “anti-modular” (i.e., close to bipartite, bottom) structures.



Supplementary Figure 4: Assembling communities one species at a time. Top: we want to build the community with species 1, 2, ..., 7 present (darker shades for more speciose communities), by adding a species at a time. Starting from an empty system (state 0), we can try all assembly pathways in which we sequentially add one species at a time, let the dynamics unfold, and reach a new state. In this case, an assembly path exists: by adding species 2, 5, 7, 3, 4, 6 and 1 one at a time, we always recover a feasible and stable community (dynamics are shown on the right). Bottom: again, we would like to build the community with all seven species present. In this case, no assembly path exist. For example, we can add sequentially 7, 6, 5, 3, and 2, reaching a stable community with five species. At this point, however, whenever we add one of the remaining species, we lose another—the state with all species present is unreachable, even when considering all possible assembly paths.



Supplementary Figure 5: Probability p_A of finding an assembly path when starting from n species. For different means and correlations of interactions strengths, sampled from a bivariate Normal with mean μ (colors) and correlation ρ (columns), and for different values of mean intrinsic growth rates (γ , rows), we plot the probability of not finding an assembly path out of 10^6 simulations. While there is a nontrivial effect of all parameters (for example, for $\rho = 0.5$ and $\gamma = 1$ we found an assembly path for all simulations), in all cases we found that for sufficiently large n , all communities could be built by sequential invasions.

References

- [1] Hofbauer, J. & Sigmund, K. *Evolutionary games and population dynamics* (Cambridge University Press, 1998).
- [2] Maynard, D. S., Serván, C. A. & Allesina, S. Network spandrels reflect ecological assembly. *Ecology Letters* **21**, 324–334 (2018).
- [3] May, R. M. Will a large complex system be stable? *Nature* **238**, 413–414 (1972).
- [4] Allesina, S. & Tang, S. The stability–complexity relationship at age 40: a random matrix perspective. *Population Ecology* **57**, 63–75 (2015).
- [5] Hofbauer, J. On the occurrence of limit cycles in the Volterra-Lotka equation. *Nonlinear Analysis: Theory, Methods & Applications* **5**, 1003–1007 (1981).
- [6] Lemke, C. E. & Howson, J. T., Jr. Equilibrium points of bimatrix games. *Journal of the Society for Industrial and Applied Mathematics* **12**, 413–423 (1964).