Coexistence of many species in random ecosystems

Carlos A. Serván[1,†], José A. Capitán[1,2,†], Jacopo Grilli[1], Kent E. Morrison[3] and Stefano Allesina*[1,4]

[1]Department of Ecology & Evolution, University of Chicago, 1101 E. 57th Chicago, IL 60637, USA
[2]Department of Applied Mathematics, Universidad Politécnica de Madrid
[3]American Institute of Mathematics
[4]Northwestern Institute on Complex Systems, Northwestern University
† Equal contribution

Abstract

Rich ecosystems harbor thousands of species interacting in tangled networks encompassing predation, mutualism and competition. Such widespread biodiversity is puzzling because in ecological models it is exceedingly improbable to obtain the stable coexistence of large communities. One aspect rarely considered in these models, however, is that coexisting species in natural communities are a selected portion of a much larger pool, which has been pruned by population dynamics.

Here we compute the distribution of the number of species that can coexist when we start from a pool of species interacting randomly, and show that even in this case we can observe rich, stable communities. Interestingly, our results show that, once stability conditions are met, network structure has very little influence on the level of biodiversity attained.

Our results identify the main drivers responsible for widespread coexistence in natural communities, providing a baseline for determining which structural aspects of empirical communities promote or hinder coexistence.

Lotka[1] and Volterra[2] first attempted to mathematize the population dynamics of interacting species, and their model has been eviscerated and refined by countless studies[3]. Analyzing models that include more than a handful of interacting populations has however

---

*sallesina@uchicago.edu

proven remarkably difficult, despite the fact that ecosystems harbor hundreds of populations, interacting through complex networks encompassing consumption, competition, and mutualism[4].

In Lotka-Volterra and similar models, it is exceedingly improbable to obtain the coexistence of all species in a large community without fine-tuning the parameters[5–9], and such fine-tuning is questionable at best for biological systems[10]. Consider however that in natural communities the extant species we observe are a selected portion of a much larger pool, which has then been pruned by population dynamics[7,11]. Therefore, to understand the establishment and maintenance of natural communities we need to change our focus: rather than asking what is the probability that all species in a community coexist, here we attempt to predict the number of extant species we obtain when starting from a species pool of n species, and let the dynamics unfold. As a limiting case, we study the behavior of ecological models in which the parameters are randomly drawn from fixed distributions, meaning that species have not had time to co-adapt or co-evolve. While many studies have investigated, numerically[12–19] or analytically[20], the effect of particular parameterizations and network structure on the average number of coexisting species, here we derive the full distribution.

We start by studying coexistence in random ecological communities, and, having derived the behavior of random networks of interacting species, we probe the effect of particular network structures on coexistence. We find that network structure, which has been shown to have strong influence on the stability properties of ecological communities[4,8,21–24], has instead very little effect on coexistence, once stability conditions are met.

The idea of studying random ecological communities was pioneered by May[5], who determined the local stability properties of large ecosystems through an application of random matrix theory. His work was generalized and refined[25,26], so that we can now characterize the stability of ecological networks displaying hierarchical[22] or modular[23] structure. Similarly, "structural stability" (i.e., the range of conditions leading to positive equilibria in ecological systems) has been investigated by letting the growth rate of the species[8,9], or the interactions between species[27] vary randomly. Clearly, to have robust coexistence we need a combination of the two: species densities must be positive, while a stable attractor is needed to allow densities to rebound when perturbed.

Results

Our goal is to compute the probability of observing k species stably coexisting when starting with a pool of n interacting populations and random parameters. For example, take the generalized Lotka-Volterra (GLV) system

$$\frac{dX_i(t)}{dt} = X_i(t)\left(r_i + \sum_j A_{ij}X_j(t)\right) , \tag{1}$$

and sample parameters at random: how many species coexist once the dynamics have elapsed?

We first analyze the case closest to the spirit of May's contribution, which can be taught of as a caricature of a food web: some species can grow in isolation (e.g., producers, with positive intrinsic growth rates), while other species can grow only thanks to their interactions (e.g., consumers, with negative growth rates); all species establish random interactions with each other. More specifically, we sample the intrinsic growth (death) rates (r_i) and the inter-specific interactions (A_{ij}, $i \neq j$) from distributions (not necessarily the same) that are symmetric around zero (such that $P(x) = P(-x)$). For example, we could sample all these entries from a Normal distribution with mean zero. We set the intra-specific interactions (A_{ii}) by summing a mean-zero symmetric random variable and a constant d_i (not necessarily the same for all i). Note that in this way, about half of the species would grow in isolation, while the rest rely on "consumption" for their survival.

We start by presenting a result on the feasibility of equilibria. Under the conditions outlined above, the probability that a system composed of n species has a completely positive equilibrium point (i.e., in which all species have positive density) is $1/2^n$, irrespective of the choice of d_i, and the exact shapes of the distributions (Supplementary Information S2). Our proof extends previously known mathematical results[28], confirming the conjecture put forward by Goh & Jennings forty years ago[13].

Clearly, feasibility is only necessary, but not sufficient for coexistence. To study coexistence, we make the stronger assumption that the matrix $A + A^T$ is negative definite. This property implies Lyapunov diagonal stability, and is a strong form of stability routinely assumed in studies of feasibility[8,9] that can be always attained by choosing suitable large and negative d_i. Under these conditions, a GLV model has a single, globally attractive equilibrium, called the non-invasible solution (also called saturated rest point[29]): k species have positive density at equilibrium, while all the other $n-k$ species cannot invade this community, and will go extinct irrespective of initial conditions. Surprisingly when we sample the parameters at random as specified above, the non-invasibility and feasibility conditions for each subset of species balance out, such that each species has probability $1/2$ of being included in the non-invasible, globally attractive solution. Hence, the probability $P(k|n)$ of finding k species coexisting when we start with n follows the binomial distribution $B(n, 1/2)$ (Fig. 1 and Supplementary Information S2). This beautifully simple result means that if we were to start with a strongly stable (i.e., with $A + A^T$ negative definite) random matrix of interactions and random growth rates, about half of the species would coexist, irrespective of the choice of n. Remarkably, this is exactly what we would expect if species were not to interact with each other at all (i.e., $d_i < 0$ for all i and $A_{ij} = 0$ for all $i \neq j$).

Extending May's results, Allesina & Tang[25] showed how stability is strongly influenced by the correlation between the inter-specific interactions: if we sample interactions in pairs (A_{ij}, A_{ji}) from a bivariate distribution with mean zero and correlation ρ, then stability is enhanced by choosing a negative correlation. When analyzing coexistence, breaking the independence among the inter-specific effects by sampling them in pairs from a bivariate distribution has no effect: we recover the same condition for feasibility, and the same distribution for the number of coexisting species (Fig. 1 and Supplementary Information S2).

So far, we have assumed that every species interacts with every other. To study the effect of network structure, we set most of the interactions to zero, and choose the position of the nonzero coefficients according to the adjacency matrix of a) an Erdős-Rényi random graph, b) a random graph with power-law degree distribution, c) a graph displaying modular, or d) bipartite structure. Irrespective of the choice of network structure, we always recover the same distribution for the number of coexisting species k (Fig. 1 and Supplementary Information S2). This is interesting, because network structure strongly influences stability[22–25]. However, because in our analysis stability is assumed, we find that the exact location of the nonzero interactions has no effect on coexistence.

The results above hold when we sample the growth rates and the inter-specific effects from symmetric distributions with mean zero, meaning that positive effects (e.g., contribution of prey to the growth of predators) on average counterbalance negative ones (e.g., effects of predators on prey). Of course this needs not to be the case in natural communities, and therefore we examine the mathematically much more challenging case in which the entries have mean nonzero.

To this end, we consider a simple model of interacting competitors: we set all inter- and intra-specific interactions to be negative, and consider the case of random growth rates. In this case we assume that all species in the pool are sampled from a common habitat, and therefore have growth rates with a well-defined average value. In particular, we sample the intrinsic growth rates from a Normal distribution with mean γ, and, for simplicity, we construct A by setting all inter-specific interaction to be competitive, $A_{ij} = \mu = \hat{\mu}/n < 0$, and all intra-specific effects to $A_{ii} = d_i = \alpha < 0$. Numerical simulations presented below show that our results well-approximate the case in which the elements of A are variable (e.g., when the nonzero elements are arranged in a network).

Again, we consider matrices for which α is sufficiently strong to yield Lyapunov diagonal stability ($\alpha < \mu < 0$). When we sample the growth rates from a Normal distribution, then the equilibrium point $X = -A^{-1}r$ is described by a multivariate Normal distribution. Exploiting this fact, we are able to express the probability that k species form a non-invasible and feasible subset as a double integral that can be used to compute the size of the non-invasible community (see Supplement). The double integral can be approximated, for large n, via a saddle-point technique to obtain an accurate analytical approximation for the distribution $P(k|n; \alpha, \hat{\mu}, \gamma)$. Note that in this setting, growth rates need to be positive for species to survive, and therefore we only consider the case of $\gamma \geq 0$. We also show (Supplementary Information S7) that the results remain qualitatively unchanged when rates are drawn from a truncated Gaussian distribution, which forces all rates to remain strictly positive.

The results (Fig. 2) show that a nonzero mean γ in growth rates can yield a larger (red area of parameter space) or smaller (blue) number of coexisting species, compared to the mean-zero case. If

$$\frac{\alpha\gamma}{\hat{\mu}} > \frac{1}{\sqrt{2\pi}}, \tag{2}$$

averages are larger than expected in the mean-zero case (and conversely). The distribution $P(k|n; \alpha, \hat{\mu}, \gamma)$ is not binomial anymore, but still retains a strong central tendency. Importantly, the mode of the number of species can be estimated analytically (Supplementary Information S6).

When we repeat the calculation but position the nonzero elements according to a network structure, we find results that are quite similar to the mean-zero case: though not all network structures yield the same exact distribution, the effect is very modest, such that our analytical approximation well-describes coexistence in all cases (Fig. 3).

In summary, we have computed the distribution of the number of coexisting species under the assumptions of random parameters and strong stability. We have two cases: a) when inter-specific interactions have mean zero, the number of coexisting species follows the binomial distribution with probability 1/2, and network structure has no influence whatsoever—in fact, we would recover the same result if species were not to interact at all; b) when the inter-specific interactions have mean nonzero, the distribution is not binomial anymore, and we can expect either a larger or smaller proportion of populations to survive, depending on the choice of parameters. Also in this case, however, network structure has a very modest effect.

Discussion

Our results show that large communities can stably coexist thanks to the selection imposed by the dynamical pruning of a large species pool. In practice, we can attain communities of any size (with no saturation) even when setting parameters at random—all we need is to start with a much larger species pool.

The study of the stability of large ecological communities started by considering completely random matrices of interactions[5]; further studies included more realistic models in which interactions were paired[25] and organized in patterns[22,23]. We believe that our results can be similarly extended, and we see three main directions that need to be explored.

First, we have considered here a "weak" form of network structure: the location of the nonzero elements of the matrix is specified, but other than that the coefficient values are randomly determined. A stronger form of network structure would be one in which also the values of the nonzero coefficients are organized in a pattern. For example, a "cascade" structure in which all the positive (negative) elements of the matrix A are confined to the upper (lower) triangular part has been shown to have a strong stabilizing (or destabilizing) effect[22]. Similarly, arranging the strong/weak competitive interactions in modules or in a nested fashion can greatly influence stability[24]. It would therefore be important to determine whether this "strong" formulation of network structure can indeed influence coexistence as well as stability.

Second, we have determined coexistence under the assumption of strong stability (Lyapunov diagonal stability). Relaxing this constraint will be challenging, but could however shed light on mechanisms of coexistence involving for example limit cycles or chaotic attractors. Recently, Bunin[20] studied coexistence in species pools with random (weak) interactions and identical growth rates, identifying the transitions between systems characterized by a single stable equilibrium, and those displaying multiple attractors. Though this study disregards other types of attractors, it shows that analytical progress in this area is possible.

Third, as pointed out by Sigmund[7], "Mother Nature does not assemble her networks by throwing n species together in one go". Understanding the process of assembly in which communities are built one species at a time is perhaps the greatest challenge ahead for theoretical community ecology[30]. In the Supplementary Information (S9) we show that, although some of our non-invasible communities cannot be built by a sequential assembly, the probability of finding such cases decreases rapidly with the size of the community. We conjecture that, asymptotically, the probability of finding an assembly sequence for communities built in this way converges to one.

In the last few decades, ecologists have compiled ever more detailed interaction networks[31], documenting the intricate relationships occurring in ecosystems[32,33]. These networks display interesting patterns, such as broad degree distributions[34], modular organization of interactions[35], hierarchical structure[36], and nestedness[37,38]. One of the main questions in community ecology is therefore to determine whether these network properties have some bearing for the robust coexistence of ecological communities. In this context, our results provide a baseline for species coexistence under Lotka-Volterra dynamics—one can use these reference points to prove that certain features of empirical communities promote or hinder coexistence.

References

[1] Lotka, A. J. *Elements of physical biology* (Williams & Wilkins Company, Baltimore, MD, 1925).

[2] Volterra, V. Fluctuations in the abundance of a species considered mathematically. *Nature* **118**, 558–560 (1926).

[3] Kingsland, S. Alfred J. Lotka and the origins of theoretical population ecology. *Proceedings of the National Academy of Sciences* **112**, 9493–9495 (2015).

[4] Pascual, M. & Dunne, J. A. *Ecological networks: linking structure to dynamics in food webs* (Oxford University Press, 2006).

[5] May, R. M. Will a large complex system be stable? *Nature* **238**, 413–414 (1972).

[6] Goh, B. S. Global stability in many-species systems. *The American Naturalist* **111**, 135–143 (1977).

[7] Sigmund, K. Darwin's "circles of complexity": Assembling ecological communities. *Complexity* **1**, 40–44 (1995).

[8] Rohr, R. P., Saavedra, S. & Bascompte, J. On the structural stability of mutualistic systems. *Science* **345**, 1253497 (2014).

[9] Grilli, J. *et al.* Feasibility and coexistence of large ecological communities. *Nature Communications* **8** (2017).

[10] Grilli, J., Barabás, G., Michalska-Smith, M. J. & Allesina, S. Higher-order interactions stabilize dynamics in competitive network models. *Nature* **548**, 210–213 (2017).

[11] Levine, J. M., Bascompte, J., Adler, P. B. & Allesina, S. Beyond pairwise mechanisms of species coexistence in complex communities. *Nature* **546**, 56–64 (2017).

[12] Roberts, A. The stability of a feasible random ecosystem. *Nature* **251**, 607–608 (1974).

[13] Goh, B. S. & Jennings, L. S. Feasibility and stability in randomly assembled Lotka-Volterra models. *Ecological Modelling* **3**, 63–71 (1977).

[14] Drake, J. A. The mechanics of community assembly and succession. *Journal of Theoretical Biology* **147**, 213–233 (1990).

[15] Brose, U., Williams, R. J. & Martinez, N. D. Allometric scaling enhances stability in complex food webs. *Ecology Letters* **9**, 1228–1236 (2006).

[16] Otto, S. B., Rall, B. C. & Brose, U. Allometric degree distributions facilitate food-web stability. *Nature* **450**, 1226–1229 (2007).

[17] Williams, R. J. Effects of network and dynamical model structure on species persistence in large model food webs. *Theoretical Ecology* **1**, 141–151 (2008).

[18] Stouffer, D. B. & Bascompte, J. Compartmentalization increases food-web persistence. *Proceedings of the National Academy of Sciences* **108**, 3648–3652 (2011).

[19] James, A., Pitchford, J. W. & Plank, M. J. Disentangling nestedness from models of ecological complexity. *Nature* **487**, 227–230 (2012).

[20] Bunin, G. Ecological communities with Lotka-Volterra dynamics. *Physical Review E* **95**, 042414 (2017).

[21] Yodzis, P. The stability of real ecosystems. *Nature* **289**, 674–676 (1981).

[22] Allesina, S. *et al.* Predicting the stability of large structured food webs. *Nature Communications* **6** (2015).

[23] Grilli, J., Rogers, T. & Allesina, S. Modularity and stability in ecological communities. *Nature Communications* **7** (2016).

[24] Barabás, G., J. Michalska-Smith, M. & Allesina, S. The effect of intra-and interspecific competition on coexistence in multispecies communities. *The American Naturalist* **188**, E1–E12 (2016).

[25] Allesina, S. & Tang, S. Stability criteria for complex ecosystems. *Nature* **483**, 205–208 (2012).

[26] Allesina, S. & Tang, S. The stability–complexity relationship at age 40: a random matrix perspective. *Population Ecology* **57**, 63–75 (2015).

[27] Stone, L. The Google matrix controls the stability of structured ecological and biological networks. *Nature Communications* **7** (2016).

[28] Morrison, K. E. From bocce to positivity: some probabilistic linear algebra. *Mathematics Magazine* **86**, 110–119 (2013).

[29] Hofbauer, J. Saturated equilibria, permanences, and stability for ecological systems. In Gross, L. J., Hallam, T. G. & Levin, S. A. (eds.) *Mathematical Ecology - Proceedings Of The Autumn Course Research Seminars International Ctr For Theoretical Physics* (World Scientific Publishing Company, 1988).

[30] Maynard, D. S., Serván, C. A. & Allesina, S. Network spandrels reflect ecological assembly. *Ecology Letters* **21**, 324–334 (2018).

[31] Dunne, J. A., Labandeira, C. C. & Williams, R. J. Highly resolved early Eocene food webs show development of modern trophic structure after the end-Cretaceous extinction. *Proceedings of the Royal Society of London B: Biological Sciences* **281**, 20133280 (2014).

[32] Kéfi, S. *et al.* Network structure beyond food webs: mapping non-trophic and trophic interactions on chilean rocky shores. *Ecology* **96**, 291–303 (2015).

[33] Sander, E. L., Wootton, J. T. & Allesina, S. What can interaction webs tell us about species roles? *PLoS Computational Biology* **11**, e1004330 (2015).

[34] Dunne, J. A., Williams, R. J. & Martinez, N. D. Food-web structure and network theory: the role of connectance and size. *Proceedings of the National Academy of Sciences* **99**, 12917–12922 (2002).

[35] Olesen, J. M., Bascompte, J., Dupont, Y. L. & Jordano, P. The modularity of pollination networks. *Proceedings of the National Academy of Sciences* **104**, 19891–19896 (2007).

[36] Cohen, J. E., Briand, F. & Newman, C. M. *Community food webs: data and theory*, vol. 20 (Springer Science & Business Media, 1990).

[37] Bascompte, J., Jordano, P., Melián, C. J. & Olesen, J. M. The nested assembly of plant–animal mutualistic networks. *Proceedings of the National Academy of Sciences* **100**, 9383–9387 (2003).

[38] Staniczenko, P. P., Kopp, J. C. & Allesina, S. The ghost of nestedness in ecological networks. *Nature Communications* **4**, 1391 (2013).

References in Supplementary Information.

[39] Johnson, C. R. Positive definite matrices. *The American Mathematical Monthly* **77**, 259–264 (1970).

[40] Eugenius, K. & Amit, B. *Matrix Diagonal Stability in Systems and Computation* (Birkäuser Boston, 2000).

[41] Hofbauer, J. & Sigmund, K. *Evolutionary games and population dynamics* (Cambridge university press, 1998).

[42] Hofbauer, J. On the occurrence of limit cycles in the Volterra-Lotka equation. *Nonlinear Analysis: Theory, Methods & Applications* **5**, 1003–1007 (1981).

[43] 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).

Acknowledgements

We thank D. Maynard and G. Barabás for comments. C.A.S. and S.A. supported by NSF-DEB 1148867; J.G. by the Human Frontier Science Program; J.A.C. by the Spanish Ministerio de Economía y Competitividad project CGL2015-69034-P; a Fulbright Fellowship (program FMECD-ST-2016, grant number CAS16/00096) allowed J.A.C. to visit the U. of Chicago.

Contributions

S.A. and C.A.S. devised the study; C.A.S. and K.E.M. solved the mean-zero case; J.A.C. and J.G. the nonzero-mean case; S.A. wrote the main text; J.A.C., C.A.S. and J.G. the supplement; C.A.S. drew the figures; all authors edited the manuscript.

Competing financial interests

The authors declare no competing financial interests.

Corresponding author

Correspondence to: Stefano Allesina (`sallesina@uchicago.edu`).

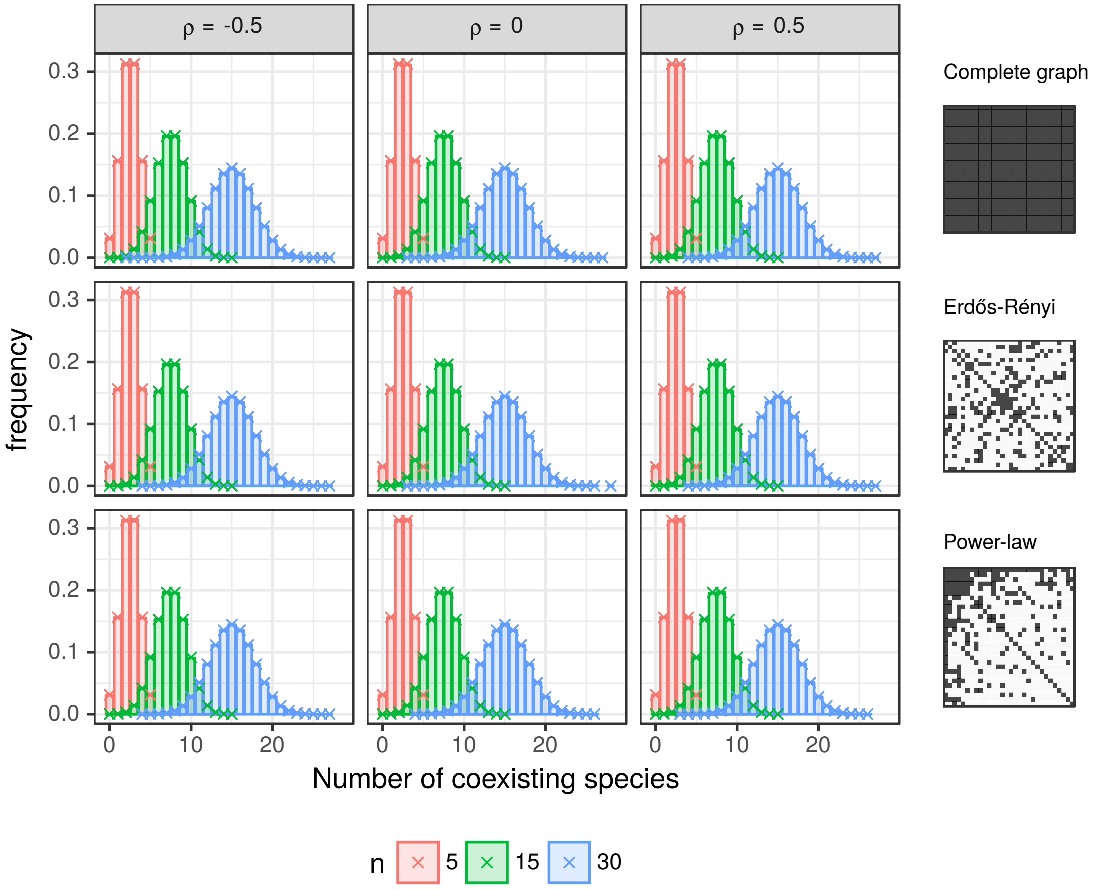

Figure 1. Number of coexisting species when interactions and intrinsic growth rates are randomly sampled from the standard Normal distribution. For each panel, histograms show the number of coexisting species out of $2 \cdot 10^5$ simulations, when starting from a different number of species n (colors) and interaction matrices A that are strongly stable. Binomial distributions $B(n, 1/2)$ are reported as crosses. In the three rows, different network structures are used to set the positions of the nonzero coefficients (as exemplified by the adjacency matrices on the right); top: complete graphs, middle: Erdős-Rényi graphs; bottom: Power-law graphs; the results for other network structures are presented in Fig. S1. Sampling the off-diagonal coefficients of matrix A independently ($\rho = 0$, center), or in correlated pairs (A_{ij}, A_{ji}) ($\rho \neq 0$), has no effect on the expected number of coexisting species.

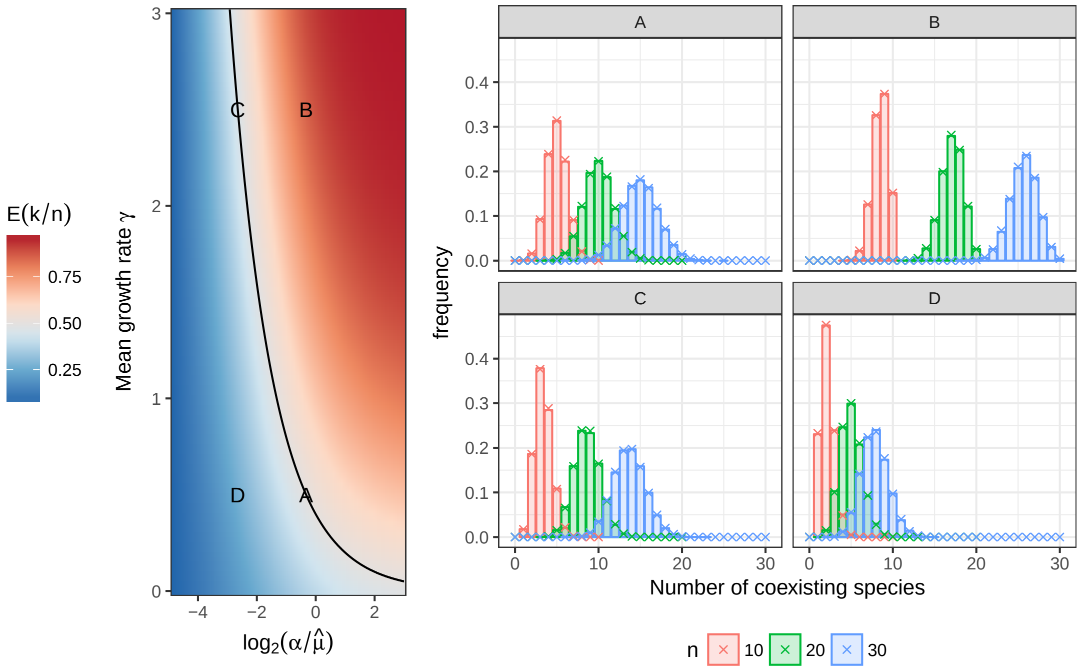

Figure 2. Number of coexisting species for competitive interactions. When species interact competitively, the histograms deviate from the binomial distribution, but can still be computed using a double integral (crosses, see Eq. S38). Here the interactions are set to $A_{ij} = \hat{\mu}/n < 0$, intra-specific competition to $A_{ii} = \alpha$, and intrinsic growth rates are normally-distributed with mean γ. The expected value of the ratio k/n, $E(k/n)$, is drawn on the left in the relevant parameter space: we chose two points (A, C) for which predictions in the nonzero mean case match closely those for mean zero ($E(k/n) = 1/2$); in case B the number of species coexisting exceeds that for the mean-zero case; for point D the expectation is lower. The analytical prediction in equation (2) is also shown (line).

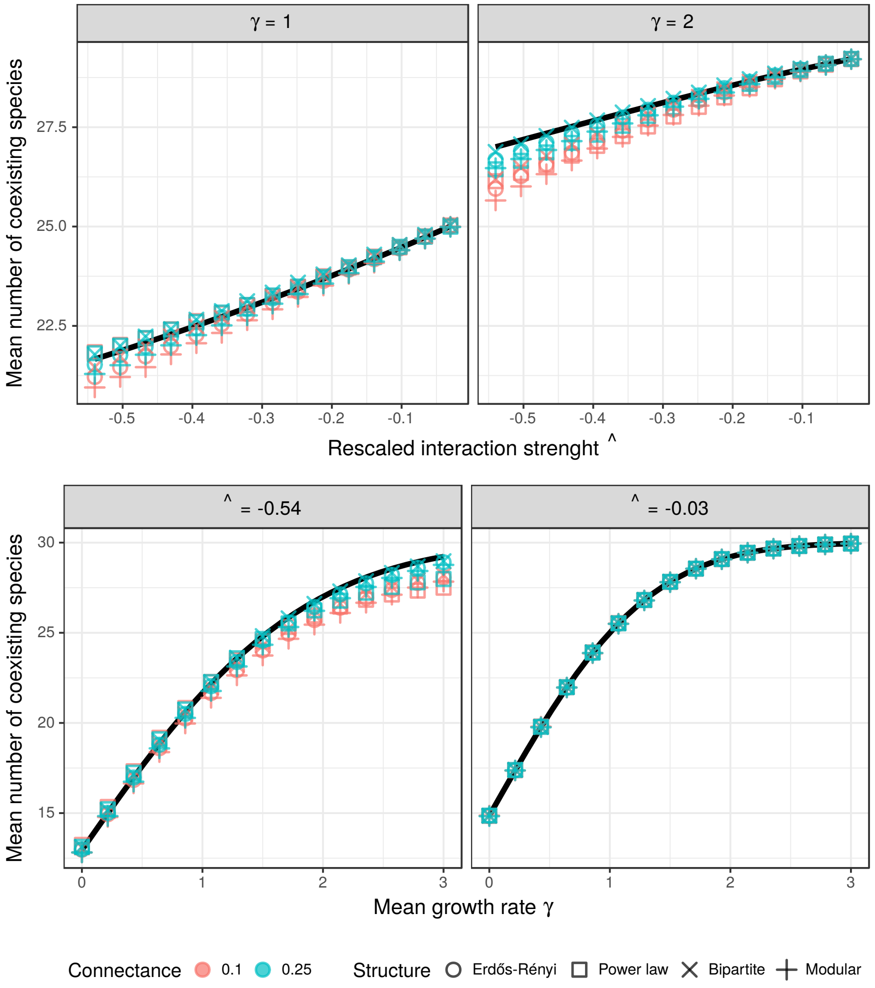

Figure 3. Effect of network structure on coexistence for the case of nonzero means. The position of the nonzero coefficients is chosen according to one of four structures (shape), and for two levels of connectance (proportion of nonzero coefficients, colors). Because most of the coefficients are zero, one needs to calculate a "rescaled" $\hat{\mu}$ (x-axis in upper panels, see Supplement) in order to contrast the results of the simulations (point) with our analytical approximation for the fully-connected case (line, see Supplement, Eq. (S87)). The four panels show that, although both the interaction strength $\hat{\mu}$ and the mean intrinsic growth rate γ interacts with connectance and network structure in nontrivial ways, the overall effect is very modest.

Supplementary Information

S1 Problem statement

We consider n interacting populations, whose dynamics are defined by a system of Generalized Lotka-Volterra (GLV) equations:

$$\frac{dX_i(t)}{dt} = X_i(t)\left(r_i + \sum_j A_{ij} X_j(t)\right) \,, \tag{S1}$$

where $X_i(t)$ is the abundance of population i at time t, r_i is the intrinsic growth rate of species i, and A_{ij} is the per-capita effect of species j on the growth rate of species i. For notational convenience, we collect the coefficients A_{ij} into the interaction matrix $\boldsymbol{A}$, and X_i and r_i into the (column) vectors $\boldsymbol{X}$ and $\boldsymbol{r}$, respectively.

A vector $\boldsymbol{x}^\star$ is a fixed point (equilibrium) of the system if

$$0 = x_i^\star\left(r_i + \sum_j A_{ij} x_j^\star\right) \quad \text{for } i = 1, 2, ..., n. \tag{S2}$$

Since $x_i^\star = 0$ is always a possible solution, the system admits up to 2^n fixed points, corresponding to all the combinations of presence and absence of each species.

A fixed point is feasible if $x_i^\star > 0$ for all i. If a feasible fixed point exists, it is the solution of

$$\boldsymbol{r} = -\boldsymbol{A}\boldsymbol{x}^\star. \tag{S3}$$

If $\boldsymbol{A}$ is invertible, then

$$\boldsymbol{x}^\star = -\boldsymbol{A}^{-1}\boldsymbol{r}. \tag{S4}$$

S1.1 Global stability and non-invasible fixed points.

In the following, we assume that $\boldsymbol{A}$ is negative definite, and in particular that the matrix $\boldsymbol{A} + \boldsymbol{A}^T$ has only negative eigenvalues[39]. A matrix $\boldsymbol{A}$ is Lyapunov diagonally stable if there exists a positive diagonal matrix $\boldsymbol{D}$ such that $\boldsymbol{D}\boldsymbol{A} + \boldsymbol{A}^T\boldsymbol{D}$ is negative definite[40]. Our assumption therefore implies Lyapunov diagonal stability (corresponding to choosing $\boldsymbol{D}$ as the identity matrix).

If $\boldsymbol{A}$ is diagonally stable, then there exists a fixed point of equation (S1) that is globally attractive: irrespective of the (positive) initial conditions, dynamics always converge to the same fixed point[41]. This globally stable fixed point has k positive entries and $n-k$ entries equal to zero. We define the support $\{S\}_k$ as the set of k persistent species (i.e., those for which at equilibrium $x_i^\star > 0$) and $\{N\}_{n-k} = \{S\}_n \setminus \{S\}_k$ as the set of $n-k$ species with zero abundance. The i^{th} entry of the globally stable fixed point $\boldsymbol{x}^\star$ is equal to zero if $i \in \{N\}_{n-k}$ and equal to $x_i > 0$ if $i \in \{S\}_k$, where $\boldsymbol{x} = (x_i)$ is a k-dimensional (column) vector with positive components. We define the $k \times k$ matrix $\boldsymbol{A}^{(s)}$ as the submatrix of $\boldsymbol{A}$ obtained by considering only rows and columns belonging to $\{S\}_k$. Similarly, we define the $(n-k)\times(n-k)$ matrix $\boldsymbol{A}^{(n)}$ by considering rows and columns in $\{N\}_{n-k}$, the $k \times (n-k)$ matrix $\boldsymbol{A}^{(sn)}$ by considering rows in $\{S\}_k$ and columns in $\{N\}_{n-k}$, and the $(n-k) \times k$ matrix $\boldsymbol{A}^{(ns)}$ by considering rows in $\{N\}_{n-k}$ and columns in $\{S\}_k$. Finally, the entries of the intrinsic growth rate vector can be split into two subvectors $\boldsymbol{r}^{(s)}$, a $k-$dimensional (column) vector with same components of $\boldsymbol{r}$ for the entries in $\{S\}_k$, and $\boldsymbol{r}^{(n)}$, a $(n-k)-$dimensional (column) vector with entries corresponding to $\{N\}_{n-k}$.

If we rearrange the indices of the vectors such that the k persistent species occupy the first k entries, the globally stable fixed point $\boldsymbol{x}^\star$ can be written as the vector $\begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{0}_{n-k} \end{pmatrix}$, where $\boldsymbol{0}_{n-k}$ denotes a (column) vector with $n-k$ zero entries, the intrinsic growth rate vector becomes $\boldsymbol{r} = \begin{pmatrix} \boldsymbol{r}^{(s)} \\ \boldsymbol{r}^{(n)} \end{pmatrix}$, and the interaction matrix reads

$$\boldsymbol{A} = \left(\begin{array}{c|c} \boldsymbol{A}^{(s)} & \boldsymbol{A}^{(sn)} \\ \hline \boldsymbol{A}^{(ns)} & \boldsymbol{A}^{(n)} \end{array}\right) . \tag{S5}$$

The abundance of the k persistent species is therefore a solution of the equation

$$\boldsymbol{A}^{(s)}\boldsymbol{x} = -\boldsymbol{r}^{(s)} . \tag{S6}$$

Since we are considering only diagonally stable matrices, this point is also not invasible by any of the remaining $n-k$ species (i.e., none of the species in $\{N\}_{n-k}$ can invade when the system is resting at the equilibrium point)[41]. The condition of non-invasibility can be written by imposing that the growth rate of each of the $n-k$ species is negative for small densities. In the limit of small densities, the per-capita growth rates of the invaders become independent of their densities, and one obtains the following $n-k$ conditions

$$\boldsymbol{r}^{(n)} + \boldsymbol{A}^{(ns)}\boldsymbol{x} < 0 \ . \tag{S7}$$

In the case of diagonally stable matrices, the combination of $\{S\}_k$ and $\boldsymbol{x}$ is unique. It is the only one for which the solution $\boldsymbol{x}$ of equation (S6) has positive components and, simultaneously, equation (S7) holds.

S1.2 Distribution of non invasible fixed points

Provided that $\boldsymbol{A}$ is diagonally stable, the number of coexisting species k is fully and uniquely determined by the vector of intrinsic growth rates $\boldsymbol{r}$. More precisely, only the direction of the vector $\boldsymbol{r}$, and not its norm, determines coexistence. Our goal is to determine $P(k|n)$, the probability of observing k coexisting species out of n, given a distribution for the entries of the matrix $\boldsymbol{A}$ and a distribution for the intrinsic growth rates $\boldsymbol{r}$. In particular, we parameterize the entries of $\boldsymbol{A}$ as the sum of a deterministic and a random matrix:

$$A_{ij} = (\alpha - \mu)\delta_{ij} + \mu + B_{ij} \ , \tag{S8}$$

where $\boldsymbol{B}$ is a random matrix, whose entries are random variables with mean zero, and $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. As such, the entry A_{ii} (self-interaction) has mean α, while the off-diagonal entries have mean μ. Similarly, we consider

$$r_i = \gamma + b_i \ , \tag{S9}$$

where the entries of vector $\boldsymbol{b}$ are random variables with mean zero.

We define $\mathcal{P}(\{S\}_k|\boldsymbol{A})$ as the probability (calculated over the growth rate vectors $\boldsymbol{r}$) that the support of the globally stable fixed points is $\{S\}_k$. By averaging this quantity over the distribution of $\boldsymbol{A}$, we obtain

$$\mathcal{P}(\{S\}_k|n) := \mathbb{E}(\mathcal{P}(\{S\}_k|\boldsymbol{A})) \ . \tag{S10}$$

The probability that the support has cardinality k is simply

$$P(k|n) := \sum_{\{S\}_k} \mathcal{P}(\{S\}_k|n) \ . \tag{S11}$$

In section (S2) we focus on the case $\mu = 0$ and $\gamma = 0$, showing that if the distributions of the entries $\boldsymbol{B}$ and $\boldsymbol{b}$ are symmetric around zero

$$P(k|n) = \binom{n}{k} \frac{1}{2^n} \ . \tag{S12}$$

In section (S3) we provide an integral formula for $\mathcal{P}(\{S\}_k|\boldsymbol{A})$ in case of a arbitrary matrix $\boldsymbol{A}$ and in section (S4) we exploit this results to compute explicitly $P(k|n)$ in the case of $\mu \neq 0$, $\gamma \neq 0$, $\boldsymbol{B} = 0$ and normally distributed entries of $\boldsymbol{b}$.

S2 Mean zero

S2.1 Toy model: uncoupled logistic equations

Suppose that $\boldsymbol{A}$ 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 $\boldsymbol{x}$ 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.

S2.2 Feasibility

In this and the following section, we show that when the entries of matrix $\boldsymbol{B}$ and vector $\boldsymbol{r}$ are random variables whose distribution is symmetric around 0, and that any n element subset of the columns of $\boldsymbol{B}$ and $\boldsymbol{r}$ are linearly independent (which holds almost surely if the entries of $\boldsymbol{B}$ and $\boldsymbol{r}$ 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 (S4), each of them is equally probable.

Let $\boldsymbol{x}^\star$ be an arbitrary solution of equation (S4), and define the matrix $\boldsymbol{D}_k = ((-1)^{\delta_{ik}}\delta_{ij})$. Then, $\boldsymbol{D}_k\boldsymbol{x}^\star$ satisfies $(\boldsymbol{D}_k\boldsymbol{A}\boldsymbol{D}_k)\boldsymbol{D}_k\boldsymbol{x}^\star = -\boldsymbol{D}_k\boldsymbol{r}$. Because of the symmetry assumption, we have that $\boldsymbol{D}_k\boldsymbol{A}\boldsymbol{D}_k$ has the same distribution[1] as $\boldsymbol{A}$, and similarly for $\boldsymbol{D}_k\boldsymbol{r}$ and $\boldsymbol{r}$. Since $\boldsymbol{D}_k$ just flips the sign of the k^{th} component of $\boldsymbol{x}^\star$, by repeating this operation a sufficient number of times we can connect any two sign patterns of solutions to equation (S4), and thus the conclusion follows.

S2.3 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 system[41]. 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 $\boldsymbol{x}$ with k positive components (with support $\{S\}_k$) is the joint probability of the conditions expressed in equations (S6) and (S7), 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 $\boldsymbol{x} > 0$. Let $\boldsymbol{z} = \boldsymbol{r}^{(n)} + \boldsymbol{A}^{(ns)}\boldsymbol{x}$. By following the same procedure illustrated in the previous section (applying the appropriate change of signs to $\boldsymbol{A}$ and $\boldsymbol{r}$), one can show that any sign pattern for $\boldsymbol{z}$ 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 $\boldsymbol{A}$ and a vector of rates $\boldsymbol{r}$, 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.

---

[1] This 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.

S2.4 Adding Structure

Let $\boldsymbol{G}$ be the adjacency matrix of an undirected graph, and consider the matrix $\boldsymbol{M} = \boldsymbol{G} \circ \boldsymbol{A}$, where $\circ$ represents the Hadamard (entry-wise) product between $\boldsymbol{G}$ and $\boldsymbol{A}$. Because this type of product is commutative with respect to the multiplication by a diagonal matrix, i.e., $\boldsymbol{D}(\boldsymbol{G} \circ \boldsymbol{A})\boldsymbol{D} = \boldsymbol{G} \circ (\boldsymbol{D}\boldsymbol{A}\boldsymbol{D})$ for $\boldsymbol{D}$ diagonal, the arguments used in the previous two sections still hold. This means that the distribution of $\boldsymbol{M}$ is invariant to $\boldsymbol{D}_k\boldsymbol{M}\boldsymbol{D}_k$ (even when $\boldsymbol{G}$ 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.

S3 Calculating the distribution of persistent species

If we integrate the GLV dynamics starting from an interaction matrix $\boldsymbol{A}$, a vector of intrinsic growth rates $\boldsymbol{r}$, 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 $\boldsymbol{A}$ is diagonally stable, the end point of the dynamics always correspond to a fixed point $\boldsymbol{x}^\star$, 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 $\boldsymbol{A}$, under the assumption that $\boldsymbol{A}$ is diagonally stable. We assume that the entries of $\boldsymbol{r}$ are drawn from a Normal distribution with mean γ and unit variance. This choice of a variance does not affect the generality of out results, since the coexistence properties of the Generalized Lotka-Volterra equations are independent of the norm of $\boldsymbol{r}$: 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 $\boldsymbol{z}$ with $n-k$ components as

$$\boldsymbol{z} := \boldsymbol{r}^{(n)} + \boldsymbol{A}^{(ns)}\boldsymbol{x}. \tag{S13}$$

On the other hand, we have equation (S6), that defines $\boldsymbol{x}$. By imposing feasibility and non-invasibility —equation (S7)—, it must hold that $\boldsymbol{x} > 0$ and $\boldsymbol{z} < 0$.

Using the probability density of the growth rates,

$$P(\boldsymbol{r}) = \frac{1}{\sqrt{(2\pi)^n}} \exp\left(- \sum_{i=1}^{n} \frac{(r_i - \gamma)^2}{2} \right) = \frac{1}{\sqrt{(2\pi)^n}} \exp\left(-\frac{1}{2} \|\boldsymbol{r}^{(s)} - \gamma \mathbf{1}_k\|^2 - \frac{1}{2}\|\boldsymbol{r}^{(n)} - \gamma \mathbf{1}_{n-k}\|^2 \right) , \tag{S14}$$

where $\mathbf{1}_k$ stands for a $k-$dimensional column vector whose entries are all equal to one. Introducing equation (S6) and (S13), we can write the joint probability density as

$$f(\boldsymbol{x}, \boldsymbol{z}|\boldsymbol{A}) = \frac{|\det \boldsymbol{\Lambda}|}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} \|\boldsymbol{A}^{(s)}\boldsymbol{x} + \gamma \mathbf{1}_k\|^2 - \frac{1}{2} \|\boldsymbol{z} - \boldsymbol{A}^{(ns)}\boldsymbol{x} - \gamma \mathbf{1}_{n-k}\|^2 \right) , \tag{S15}$$

where $\boldsymbol{\Lambda}$ is the Jacobian matrix obtained from the change of variables $\boldsymbol{r} \to (\boldsymbol{x}, \boldsymbol{z})$. According to equations (S6) and (S13), it is simple to observe that $\boldsymbol{\Lambda}$ as the following structure:

$$\boldsymbol{\Lambda} := \left(\begin{array}{c|c} \frac{\partial \boldsymbol{r}^{(s)}}{\partial \boldsymbol{x}} & \frac{\partial \boldsymbol{r}^{(s)}}{\partial \boldsymbol{z}} \\ \hline \frac{\partial \boldsymbol{r}^{(n)}}{\partial \boldsymbol{x}} & \frac{\partial \boldsymbol{r}^{(n)}}{\partial \boldsymbol{z}} \end{array} \right) = \left(\begin{array}{c|c} \boldsymbol{A}^{(s)} & \mathbf{0} \\ \hline \boldsymbol{A}^{(ns)} & \boldsymbol{I}_{n-k} \end{array} \right) , \tag{S16}$$

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

The first term appearing in the exponential in equation (S15) can be written as

$$\|\boldsymbol{A}^{(s)}\boldsymbol{x} + \gamma \mathbf{1}_k\|^2 = (\boldsymbol{x} - \boldsymbol{\xi})^T \boldsymbol{G} (\boldsymbol{x} - \boldsymbol{\xi}) , \tag{S17}$$

where

$$\boldsymbol{\xi} = -\gamma (\boldsymbol{A}^{(s)})^{-1} \mathbf{1}_k , \tag{S18}$$

and

$$\boldsymbol{G} = (\boldsymbol{A}^{(s)})^T \boldsymbol{A}^{(s)} . \tag{S19}$$

We obtain therefore

$$f(\boldsymbol{x}, \boldsymbol{z}|\boldsymbol{A}) = \frac{|\det \boldsymbol{A}^{(s)}|}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} (\boldsymbol{x} - \boldsymbol{\xi})^T \boldsymbol{G} (\boldsymbol{x} - \boldsymbol{\xi}) - \frac{1}{2} \|\boldsymbol{z} - \boldsymbol{A}^{(ns)}\boldsymbol{x} - \gamma \mathbf{1}_{n-k}\|^2 \right) . \tag{S20}$$

The probability $\mathcal{P}(\{S\}_k|\boldsymbol{A})$ of observing the globally stable fixed point with support $\{S\}_k$, can be obtained from the joint probability in equation (S15) by imposing the feasibility condition for the k species ($\boldsymbol{x} > 0$) and the non-invasibility condition for the other $n-k$ species

$(\boldsymbol{z} < 0)$. The equation reads

$$\mathcal{P}(\{S\}_k|\boldsymbol{A}) \equiv \int d^k\boldsymbol{x}\left(\prod_{i=1}^{k}\Theta(x_i)\right)\int d^{n-k}\boldsymbol{z}\left(\prod_{j=k+1}^{n}\Theta(-z_j)\right)f(\boldsymbol{x},\boldsymbol{z}|\boldsymbol{A})\ . \tag{S21}$$

S4 Mean non zero

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

$$\boldsymbol{A} = (\alpha-\mu)\boldsymbol{I}_n + \mu\mathbf{1}_n\mathbf{1}_n^T\ . \tag{S22}$$

Since the matrix $\boldsymbol{A}$ is a deterministic matrix, in this case $\mathcal{P}(\{S\}_k|\boldsymbol{A}) = \mathcal{P}(\{S\}_k|n)$. By introducing equation (S22) in equation (S20) and using equation (S21), we obtain

$$\begin{aligned}\mathcal{P}(\{S\}_k|n) = \frac{|\det\boldsymbol{A}^{(s)}|}{(2\pi)^{n/2}}&\int d^k\boldsymbol{x}\prod_{i=1}^{k}\Theta(x_i)\int d^{n-k}\boldsymbol{z}\prod_{j=k+1}^{n}\Theta(-z_j)\\ &\times\exp\left\{-\frac{1}{2}\left(\boldsymbol{x}-\xi^{(k)}\mathbf{1}_k\right)^T\boldsymbol{G}\left(\boldsymbol{x}-\xi^{(k)}\mathbf{1}_k\right)-\frac{1}{2}\|\boldsymbol{z}-(\mu(\mathbf{1}_k^T\boldsymbol{x})+\gamma)\mathbf{1}_{n-k}\|^2\right\},\end{aligned} \tag{S23}$$

where we used the fact that, with the parameterization of equation (S22), $\boldsymbol{\xi} = \xi^{(k)}\mathbf{1}_k$, where

$$\xi^{(k)} = -\frac{\gamma}{\alpha+(k-1)\mu}\ . \tag{S24}$$

Again, using equation (S22) together with equation (S19), we have

$$\boldsymbol{G} = (\alpha-\mu)^2\boldsymbol{I}_k + \left[k\mu^2 + 2\mu(\alpha-\mu)\right]\mathbf{1}_k\mathbf{1}_k^T\ . \tag{S25}$$

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

$$\begin{aligned}\mathcal{P}(\{S\}_k|n) = \frac{|\det\boldsymbol{A}^{(s)}|}{(2\pi)^{n/2}}&\int d^k\boldsymbol{x}\prod_{i=1}^{k}\Theta(x_i+\xi^{(k)})e^{-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{G}\boldsymbol{x}}\\ &\times\int d^{n-k}\boldsymbol{z}\prod_{j=k+1}^{n}\Theta(-z_j)e^{-\frac{1}{2}\|\boldsymbol{z}-\left[\gamma+k\mu\xi^{(k)}+\mu(\mathbf{1}_k^T\boldsymbol{x})\right]\mathbf{1}_{n-k}\|^2}.\end{aligned} \tag{S26}$$

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

$$\mathcal{P}(\{S\}_k|n) = \frac{|\alpha-\mu|^{k-1}|\alpha+(k-1)\mu|}{(2\pi)^{n/2}} \int d^k\boldsymbol{x} \prod_{i=1}^{k} \Theta(x_i + \xi^{(k)}) e^{-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{G}\boldsymbol{x}} \\ \times \int d^{n-k}\boldsymbol{z} \prod_{j=k+1}^{n} \Theta(-z_j - \gamma - k\mu\xi^{(k)}) e^{-\frac{1}{2}\|\boldsymbol{z}-\mu(\mathbf{1}_k^T\boldsymbol{x})\mathbf{1}_{n-k}\|^2} , \quad \text{(S27)}$$

where we used

$$|\det \boldsymbol{A}^{(s)}| = |\alpha-\mu|^{k-1}|\alpha+(k-1)\mu| \ . \quad \text{(S28)}$$

By introducing the expression for $\boldsymbol{G}$ obtained in equation (S25), we get

$$\mathcal{P}(\{S\}_k|n) = \frac{|\alpha-\mu|^{k-1}|\alpha+(k-1)\mu|}{(2\pi)^{n/2}} \int d^k\boldsymbol{x} \prod_{i=1}^{k} \Theta(x_i + \xi^{(k)}) \\ \times \int d^{n-k}\boldsymbol{z} \prod_{j=k+1}^{n} \Theta(-z_j - \gamma - k\mu\xi^{(k)}) e^{g(\boldsymbol{x},\boldsymbol{z})} \quad \text{(S29)}$$

where

$$g(\boldsymbol{x},\boldsymbol{z}) = -\frac{1}{2}\Big[(\alpha-\mu)^2\boldsymbol{x}^T\boldsymbol{x} + \Big[n\mu^2 + 2\mu(\alpha-\mu)\Big](\mathbf{1}_k^T\boldsymbol{x})^2 - 2\mu(\mathbf{1}_k^T\boldsymbol{x})(\mathbf{1}_{n-k}^T\boldsymbol{z}) + \boldsymbol{z}^T\boldsymbol{z}\Big]. \quad \text{(S30)}$$

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

$$e^{-bd^2/c^2 - de/c} = \frac{c}{2\pi}\int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dw\, e^{-(by^2 + ey + idw - icwy)}. \quad \text{(S31)}$$

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

$$e^{-bd^2/c^2 + de/c} = \frac{c}{2\pi}\int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dw\, e^{-(by^2 + ey + idw + icwy)}. \quad \text{(S32)}$$

In our case [cf. equations (S25) and (S30)], we choose $d = \mathbf{1}_k^T\boldsymbol{x}$ and $e = \mathbf{1}_{n-k}^T\boldsymbol{z}$ and identify the exponents of the l.h.s. of equations (S31) or (S32) with the terms in (S30). If $\mu > 0$, we find $\frac{1}{c} = \mu$ and use equation (S32). If $\mu < 0$, we set $\frac{1}{c} = |\mu|$ and consider equation (S31). In both cases, we set $\frac{b}{c^2} = \frac{1}{2}\left[n\mu^2 + 2\mu(\alpha-\mu)\right]$. In general, we can choose $c = \frac{1}{|\mu|}$ and $b = \frac{1}{2}\Big[n + 2\Big(\frac{\alpha}{\mu} - 1\Big)\Big]$. To ensure diagonal stability, all the eigenvalues of matrix $\boldsymbol{A}$ 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 (S31). 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

$$\begin{aligned}\mathcal{P}(\{S\}_k|n) = \frac{|\alpha-\mu|^{k-1}|\alpha+(k-1)\mu|}{(2\pi)^{n/2+1}|\mu|}\int_{-\infty}^{\infty}dy\int_{-\infty}^{\infty}dw\, e^{-\frac{1}{2}\left[n+2\left(\frac{\alpha}{\mu}-1\right)\right]y^2+i\frac{yw}{|\mu|}}\int d^k\boldsymbol{x}\prod_{i=1}^{k}\Theta(x_i+\xi^{(k)})\\ \times\int d^{n-k}\boldsymbol{z}\prod_{j=k+1}^{n}\Theta(-z_j-\gamma-k\mu\xi^{(k)})e^{-\frac{1}{2}(\alpha-\mu)^2\boldsymbol{x}^T\boldsymbol{x}-i\left(\mathbf{1}_k^T\boldsymbol{x}\right)w}e^{-\frac{1}{2}\boldsymbol{z}^T\boldsymbol{z}-\left(\mathbf{1}_{n-k}^T\boldsymbol{z}\right)y}. \quad \text{(S33)}\end{aligned}$$

We complete squares and obtain

$$\begin{aligned}\mathcal{P}(\{S\}_k|n) = \frac{|\alpha-\mu|^{k-1}|\alpha+(k-1)\mu|}{(2\pi)^{n/2+1}|\mu|}\int_{-\infty}^{\infty}dy\int_{-\infty}^{\infty}dw\, e^{-\frac{1}{2}\left[n+2\left(\frac{\alpha}{\mu}-1\right)\right]y^2+i\frac{yw}{|\mu|}}e^{-\frac{k}{2(\alpha-\mu)^2}w^2+\frac{1}{2}(n-k)y^2}\\ \times\left[\int dx\Theta(x+\xi^{(k)})e^{-\frac{1}{2}(\alpha-\mu)^2\left(x+\frac{iw}{(\alpha-\mu)^2}\right)^2}\right]^k\left[\int dz\Theta(-z-\gamma-k\mu\xi^{(k)})e^{-\frac{1}{2}(z+y)^2}\right]^{n-k}. \quad \text{(S34)}\end{aligned}$$

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

$$\begin{aligned}\mathcal{P}(\{S\}_k|n) = \frac{|\alpha-\mu|^{k-1}|\alpha+(k-1)\mu|}{2\pi|\mu||\alpha-\mu|^k}\int_{-\infty}^{\infty}dy\int_{-\infty}^{\infty}dw\, e^{-\frac{1}{2}\left[k+2\left(\frac{\alpha}{\mu}-1\right)\right]y^2+i\frac{yw}{|\mu|}-\frac{k}{2(\alpha-\mu)^2}w^2}\\ \times\left[1-\Phi\left(\frac{iw}{|\alpha-\mu|}-|\alpha-\mu|\xi^{(k)}\right)\right]^k\left[\Phi\left(y-\gamma-k\mu\xi^{(k)}\right)\right]^{n-k}, \quad \text{(S35)}\end{aligned}$$

and therefore we find

$$\begin{aligned}\mathcal{P}(\{S\}_k|n) = \frac{1}{2\pi}\left|k+\frac{\alpha}{\mu}-1\right|\int_{-\infty}^{\infty}dy\int_{-\infty}^{\infty}dw\, e^{-\frac{1}{2}\left[k+2\left(\frac{\alpha}{\mu}-1\right)\right]y^2+i\left|\frac{\alpha}{\mu}-1\right|yw-\frac{1}{2}kw^2}\\ \times\left[1-\Phi\left(iw-|\alpha-\mu|\xi^{(k)}\right)\right]^k\left[\Phi\left(y-\gamma-k\mu\xi^{(k)}\right)\right]^{n-k}. \quad \text{(S36)}\end{aligned}$$

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

$$v := \frac{\gamma(\alpha-\mu)}{\alpha-\mu+k\mu} = \frac{\gamma s}{k+s}. \tag{S37}$$

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

$$\mathcal{P}(\{S\}_k|n) = \frac{k+s}{2\pi}\int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dw\, e^{-\frac{1}{2}(k+2s)y^2+isyw-\frac{1}{2}kw^2}[1-\Phi(iw-v)]^k[\Phi(y-v)]^{n-k}. \tag{S38}$$

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 $x_0 \to 0$, so that the integral over the imaginary axis is well defined. Therefore, an equivalent form of writing this equation is

$$\mathcal{P}(\{S\}_k|n) = \frac{k+s}{2\pi i}\int_{-\infty}^{\infty} dy \int_{\Gamma} dw\, e^{-\frac{1}{2}(k+2s)y^2+syw+\frac{1}{2}kw^2}[1-\Phi(w-v)]^k[\Phi(y-v)]^{n-k}, \tag{S39}$$

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

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

$$f(\boldsymbol{z}) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}(\boldsymbol{z}-\gamma\boldsymbol{1}_n)^T(\boldsymbol{x}-\gamma\boldsymbol{1}_n)} \tag{S40}$$

and the condition for non-invasibility reduces to

$$\mathcal{P}(\emptyset|n) = P[z_1 < 0, \ldots, z_n < 0] = \frac{1}{(2\pi)^{n/2}}\int d^n\boldsymbol{z} \prod_{i=1}^{n} \Theta(-z_i) e^{-\frac{1}{2}(\boldsymbol{z}-\gamma\boldsymbol{1}_n)^T(\boldsymbol{x}-\gamma\boldsymbol{1}_n)} = [\Phi(-\gamma)]^n. \tag{S41}$$

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

$$\int_{-\infty}^{\infty} dw\, e^{-\frac{1}{2}w^2-iaw}[1-\Phi(iw)] = \sqrt{2\pi}\Theta(-a)e^{-\frac{1}{2}a^2} \tag{S42}$$

we get

$$\mathcal{P}(\{S\}_1|n) = \frac{s+1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dy\, \Theta(sy+v)e^{-\frac{1}{2}(s+1)^2y^2}[\Phi(y-v)]^{n-1}, \tag{S43}$$

or, alternatively,

$$\mathcal{P}(\{S\}_1|n) = \frac{1}{\sqrt{2\pi}} \int_{-\gamma}^{\infty} dy\, e^{-\frac{1}{2}y^2} \left[\Phi\left(\frac{\mu y}{\alpha} - v\right)\right]^{n-1}. \tag{S44}$$

S4.1 Numerical evaluation of the double integral

Equation (S38) can be evaluated numerically via a Fast Fourier Transform (FFT). We can express it as

$$\mathcal{P}(\{S\}_k|n) = \frac{k+s}{2\pi} \int_{-\infty}^{\infty} dy\, e^{-\frac{1}{2}(k+2s)(y^2+2vy)+svy} [\Phi(y)]^{n-k} \widehat{F}(-s(y+v) - kv; k) \tag{S45}$$

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

$$F(w;k) = e^{-\frac{1}{2}kw^2} [1 - \Phi(iw)]^k \tag{S46}$$

and the Fourier Transform of the function $F(t;k)$ is defined as $\widehat{F}(x;k) := \int_{-\infty}^{\infty} dt\, f(t;k) e^{-itx}$. Then we first calculate $\widehat{F}(x;k)$ 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 $t_j = (j - m/2)\delta$, $0 \le j < m$. Then

$$\widehat{F}(x_\ell;k) = \int_{-\infty}^{\infty} dt\, F(t;k) e^{-itx_\ell} \approx \int_{-T/2}^{T/2} dt\, F(t;k) e^{-itx_\ell} \approx \delta \sum_{j=0}^{m-1} F(t_j;k) e^{-itx_\ell}. \tag{S47}$$

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

$$\widehat{F}(x_\ell;k) = \delta e^{i\pi(\ell - m/2)} \sum_{j=0}^{m-1} F(t_j;k) e^{i\pi j(1-2\ell/m)} = (-1)^{\ell - m/2} \delta D_\ell\Big[\{(-1)^j F(t_j;k)\}\Big] \tag{S48}$$

where $0 \le \ell < m$. Once we have calculated $\widehat{F}(x_\ell;k)$ over the set of sampling points, we interpolate to evaluate numerically the transform at an arbitrary point [see equation (S45)]. For numerical evaluation over a finite interval, equation (S45) is more conveniently expressed by changing to the variable $z = \Phi(y)$ as

$$\mathcal{P}(\{S\}_k|n) = \frac{k+s}{\sqrt{2\pi}} \int_0^1 dz\, \widehat{F}\Big(-s\Big[\Phi^{-1}(z) + v\Big] - kv; k\Big) e^{-\frac{1}{2}\left[\Phi^{-1}(z)\right]\left\{2v(k+s)+\left[\Phi^{-1}(z)\right](k-1+2s)\right\}+(n-k)\log z}. \tag{S49}$$

For $k = 1$ from (S44) we derive the expression

$$\mathcal{P}(\{S\}_1|n) = (s+1)\int_{\Phi(-\gamma)}^{1} dz\, e^{-\frac{1}{2}(s+1)^2\left[\Phi^{-1}(z)+v\right]^2+\frac{1}{2}\left[\Phi^{-1}(z)\right]^2+(n-1)\log z}. \tag{S50}$$

S4.2 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

$$P(k|n) = \binom{n}{k}\mathcal{P}(\{S\}_k|n), \tag{S51}$$

with $\mathcal{P}(\{S\}_k|n)$ given by equation (S39). 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 $k^\star$.

We use the saddle point technique from statistical mechanics to evaluate integrals of the form $\int d^n\boldsymbol{u} e^{-nh(\boldsymbol{u})}k(\boldsymbol{u})$ for n large. We define q through $k = qn$ and regard q as a continuous, finite variable such that $0 \le q \le 1$. Then equation (S39) can be written as

$$\mathcal{P}(\{S\}_k|n) = \frac{k+s}{2\pi i}\int_{-\infty}^{\infty} dy \int_\Gamma dw\, e^{-sy^2+syw} e^{-n\hat{h}(y,w;q,v)}, \tag{S52}$$

where

$$\hat{h}(y,w;q) = \frac{q}{2}\Big(y^2 - w^2\Big) - q\log[1-\Phi(w-v)] - (1-q)\log\Phi(y-v). \tag{S53}$$

In the limit $n \to \infty$, we assume q to take a fixed value (which will be associated to any possible value that k can take in the range $0 \le k \le n$). 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 (S52) through the combination $s = \frac{\alpha}{\mu} - 1$, if we do not assume the scaling in the limit for $n \to \infty$ any dependence on interaction strengths will be lost for n large. Therefore we write $s = nu - 1$, where $u := \alpha/\hat{\mu}$ and equation (S52) becomes

$$\mathcal{P}(\{S\}_k|n) = \frac{k + nu - 1}{2\pi i} \int_{-\infty}^{\infty} dy \int_{\Gamma} dw \, e^{y^2 - yw} e^{-nh(y,w;\boldsymbol{\sigma})}, \tag{S54}$$

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

$$h(y, w; \boldsymbol{\sigma}) = \frac{q}{2}\Big(y^2 - w^2\Big) - q \log[1 - \Phi(w - v)] - (1 - q) \log \Phi(y - v) + uy^2 - uyw. \tag{S55}$$

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 $\boldsymbol{\sigma}$:

$$\begin{aligned}
\frac{\partial h}{\partial y} &= qy - (1 - q)\frac{\Phi'(y - v)}{\Phi(y - v)} + 2uy - uw = qy - (1 - q)\frac{e^{-(y-v)^2/2}}{\sqrt{2\pi}\Phi(y - v)} + 2uy - uw = 0,\\
\frac{\partial h}{\partial w} &= -qw + q\frac{\Phi'(w - v)}{1 - \Phi(w - v)} - uy = -qw + q\frac{e^{-(w-v)^2/2}}{\sqrt{2\pi}[1 - \Phi(w - v)]} - uy = 0.
\end{aligned} \tag{S56}$$

This system can be solved numerically for each tuple $\boldsymbol{\sigma} = (q, u, v)$, yielding the functions $y^\star(\boldsymbol{\sigma})$ and $w^\star(\boldsymbol{\sigma})$ 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 (S56), we find

$$\begin{aligned}
\left.\frac{\partial^2 h}{\partial y^2}\right|_{\substack{y=y^\star\\ w=w^\star}} &= 2u+q+(1-q)\left[y-v+\frac{\Phi'(y-v)}{\Phi(y-v)}\right]\left.\frac{\Phi'(y-v)}{\Phi(y-v)}\right|_{\substack{y=y^\star\\ w=w^\star}}\\
&= 2u+q+(2uy^\star+qy^\star-uw^\star)\left(-v+\frac{y^\star-u(w^\star-2y^\star)}{1-q}\right),\\
\left.\frac{\partial^2 h}{\partial w^2}\right|_{\substack{y=y^\star\\ w=w^\star}} &= -q+q\left[-w+v+\frac{\Phi'(w-v)}{1-\Phi(w-v)}\right]\left.\frac{\Phi'(w-v)}{1-\Phi(w-v)}\right|_{\substack{y=y^\star\\ w=w^\star}}\\
&= -q+(uy^\star+qw^\star)\left(v+\frac{uy^\star}{q}\right),\\
\left.\frac{\partial^2 h}{\partial y\partial w}\right|_{\substack{y=y^\star\\ w=w^\star}} &= -u.
\end{aligned}\tag{S57}$$

In Section S5 we show that the critical point obtained by solving the coupled system (S56) is precisely a saddle point, as stated above. Therefore, up to second order around the saddle point,

$$h(y,w;\boldsymbol{\sigma})\approx h(y^\star,w^\star;\boldsymbol{\sigma})+\frac{1}{2}\left.\frac{\partial^2 h}{\partial y^2}\right|_{\substack{y=y^\star\\ w=w^\star}}(y-y^\star)^2+\frac{1}{2}\left.\frac{\partial^2 h}{\partial w^2}\right|_{\substack{y=y^\star\\ w=w^\star}}(w-w^\star)^2+\left.\frac{\partial^2 h}{\partial y\partial w}\right|_{\substack{y=y^\star\\ w=w^\star}}(y-y^\star)(w-w^\star).\tag{S58}$$

Substituting the expansion into equation (S54) 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 $\{S\}_k$:

$$\mathcal{P}(\{S\}_k|n)=\frac{n(q+u)-1}{\sqrt{K(\boldsymbol{\sigma},n)}}e^{-nh(y^\star,w^\star;\boldsymbol{\sigma})+y^\star(y^\star-w^\star)},\tag{S59}$$

with

$$\begin{aligned}
K(\boldsymbol{\sigma},n) := (nu-1)^2+n^2&\left[-q+(uy^\star+qw^\star)\left(v+\frac{uy^\star}{q}\right)\right]\\
&\times\left[\frac{2}{n}-2u-q-(2uy^\star+qy^\star-uw^\star)\left(-v+\frac{y^\star-u(w^\star-2y^\star)}{1-q}\right)\right].
\end{aligned}\tag{S60}$$

We can write equation (S59) as

$$\mathcal{P}(\{S\}_k|n) = \frac{n(q+u)-1}{\sqrt{K(\boldsymbol{\sigma},n)}} e^{nH(\boldsymbol{\sigma})+G(\boldsymbol{\sigma})} \tag{S61}$$

where

$$\begin{aligned} H(\boldsymbol{\sigma}) &:= \frac{q}{2}\Big(w^{\star 2} - y^{\star 2}\Big) + (1-q)\log[\Phi(y^\star - v)] + q\log[1-\Phi(w^\star - v)] - uy^{\star 2} + uy^\star w^\star, \\ G(\boldsymbol{\sigma}) &:= y^\star(y^\star - w^\star). \end{aligned} \tag{S62}$$

We now use the Stirling approximation to get

$$\binom{n}{qn} \approx \frac{e^{-n[q\log q + (1-q)\log(1-q)]}}{\sqrt{2\pi q(1-q)}}. \tag{S63}$$

According to equation (S51), our approximation for the probability of coexistence is

$$P(k|n) = \frac{n(q+u)-1}{\sqrt{2\pi q(1-q)K(\boldsymbol{\sigma},n)}} e^{nF(\boldsymbol{\sigma})+G(\boldsymbol{\sigma})}, \tag{S64}$$

where

$$\begin{aligned} F(\boldsymbol{\sigma}) := \frac{q}{2}\Big(w^{\star 2} - y^{\star 2}\Big) + (1-q)\log[\Phi(y^\star - v)] + q\log[1-\Phi(w^\star - v)] \\ - uy^{\star 2} + uy^\star w^\star - q\log q - (1-q)\log(1-q), \end{aligned} \tag{S65}$$

In the discrete distribution given by equation (S64) 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 $\frac{\alpha}{n\mu}$.

S5 Classification of the critical point

In order to prove that the critical point $(y^\star, w^\star)$ obtained as the solution of Eq. (S56) is a saddle point, we only have to show that the discriminant satisfies

$$D(y^\star, w^\star) = \left(\frac{\partial^2 h}{\partial y^2}\right)\left(\frac{\partial^2 h}{\partial w^2}\right) - \left(\frac{\partial^2 h}{\partial y \partial w}\right)^2 < 0, \tag{S66}$$

where all the derivatives are evaluated at the critical point. From Eq. (S57) we observe that

$$D(y^\star, w^\star) = \left(\frac{\partial^2 h}{\partial y^2}\right)\left(\frac{\partial^2 h}{\partial w^2}\right) - u^2. \tag{S67}$$

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 $\boldsymbol{\sigma} = (q, u, v)$. This will complete the proof.

First, consider the expression in (S57) 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

$$\left.\frac{\partial^2 h}{\partial y^2}\right|_{\substack{y=y^\star \\ w=w^\star}} \geq (2uy^\star + qy^\star - uw^\star)\left(-v + \frac{y^\star - u(w^\star - 2y^\star)}{1-q}\right). \tag{S68}$$

This product is positive or zero. On the one hand, according to (S56),

$$2uy^\star + qy^\star - uw^\star = \frac{1-q}{\sqrt{2\pi}} \frac{e^{-(y^\star - v)^2/2}}{\Phi(y^\star - v)}, \tag{S69}$$

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

$$-v + \frac{y^\star - u(w^\star - 2y^\star)}{1-q} = y^\star - v + \frac{e^{-(y^\star - v)^2/2}}{\sqrt{2\pi}\Phi(y^\star - v)} = f_1(y^\star - v), \tag{S70}$$

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 $x \to -\infty$, $f_1(x) \approx -\frac{1}{x} > 0$. Therefore $f_1(x) > 0$ for all x and we have shown that $\frac{\partial^2 h}{\partial y^2} \geq 0$.

Now, from (S56) we obtain

$$uy^\star + qw^\star = \frac{q}{\sqrt{2\pi}} \frac{e^{-(w^\star - v)^2/2}}{1 - \Phi(w^\star - v)}. \tag{S71}$$

Therefore we can express the term $v + uy^\star/q$ that appears in Eq. (S57) as

$$v + \frac{uy^\star}{q} = -(w^\star - v) + \frac{e^{-(w^\star - v)^2/2}}{\sqrt{2\pi}[1 - \Phi(w^\star - v)]}. \tag{S72}$$

Let us define the function

$$f_2(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}[1-\Phi(x)]}. \tag{S73}$$

Using the three equations above into (S57) we find

$$\left.\frac{\partial^2 h}{\partial w^2}\right|_{\substack{y=y^\star \\ w=w^\star}} = -q\{1 + f_2(w^\star - v)[w^\star - v - f_2(w^\star - v)]\}. \tag{S74}$$

Now we observe that the function

$$f_3(x) := 1 + f_2(x)[x - f_2(x)] \tag{S75}$$

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 $\frac{\partial^2 h}{\partial w^2} = -qf_3(w^\star - v) < 0$). A simple graphical analysis for $f_4(x)$ proves that this is indeed the case. As a consequence,

$$D(y^\star, w^\star) = \left(\frac{\partial^2 h}{\partial y^2}\right)\left(\frac{\partial^2 h}{\partial w^2}\right) - u^2 \leq -u^2 < 0 \tag{S76}$$

and $(y^\star, w^\star)$ is a saddle pont.

In summary, we have shown that the solution $(y^\star, w^\star)$ of Eq. (S56) is a saddle point for the function $h(y, w; \boldsymbol{\sigma})$ defined in Eq. (S55), 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 (S54) of finding the globally stable fixed point with support $\{S\}_k$.

S6 Mode of the distribution for large number of species

For large n, the mode of the distribution (S64) is recovered at a $q^\star$ value such that F takes its maximum value. We now calculate this $q^\star$ 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 (S37)], $v = \frac{\gamma s}{k+s}$. 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 (S65),

$$\frac{\partial F}{\partial q} = \frac{1}{2}\Big(w^{\star 2} - y^{\star 2}\Big) + q\Big(w^{\star}w^{\star\prime} - y^{\star}y^{\star\prime}\Big) - \log\Phi(y^{\star} - v) + \log(1 - \Phi(w^{\star} - v)) + (w^{\star} - 2y^{\star})uy^{\star\prime} \\ + uy^{\star}w^{\star\prime} + (1-q)\Big(y^{\star\prime} - v'\Big)\frac{\Phi'(y^{\star} - v)}{\Phi(y^{\star} - v)} - q\Big(w^{\star\prime} - v'\Big)\frac{\Phi'(w^{\star} - v)}{1 - \Phi(w^{\star} - v)} + \log\frac{1-q}{q}. \quad \text{(S77)}$$

Now, according to equation (S56),

$$\begin{aligned} \frac{\Phi'(y^{\star} - v)}{\Phi(y^{\star} - v)} &= \frac{qy^{\star} + 2uy^{\star} - uw^{\star}}{1-q}, \\ \frac{\Phi'(w^{\star} - v)}{1 - \Phi(w^{\star} - v)} &= \frac{uy^{\star} + qw^{\star}}{q}, \end{aligned} \quad \text{(S78)}$$

so the derivative with respect to q simplifies to

$$\frac{\partial F}{\partial q} = \frac{1}{2}\Big(w^{\star 2} - y^{\star 2}\Big) - v(w^{\star} - y^{\star}) - \log\Phi(y^{\star} - v) + \log(1 - \Phi(w^{\star} - v)) + \log\frac{1-q}{q}. \quad \text{(S79)}$$

Setting the derivative to zero yields the condition

$$(1 - q^{\star})e^{w^{\star 2}/2 - vw^{\star}}[1 - \Phi(w^{\star} - v)] = q^{\star}e^{y^{\star 2}/2 - vy^{\star}}\Phi(y^{\star} - v), \quad \text{(S80)}$$

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

$$\begin{aligned} \frac{\Phi'(y^{\star} - v)}{\Phi(y^{\star} - v)} &= \frac{qy^{\star} + 2uy^{\star} - uw^{\star}}{1-q} = \frac{e^{-(y^{\star} - v)^2/2}}{\sqrt{2\pi}\Phi(y^{\star} - v)}, \\ \frac{\Phi'(w^{\star} - v)}{1 - \Phi(w^{\star} - v)} &= \frac{uy^{\star} + qw^{\star}}{q} = \frac{e^{-(w^{\star} - v)^2/2}}{\sqrt{2\pi}[1 - \Phi(w^{\star} - v)]}, \end{aligned} \quad \text{(S81)}$$

hence

$$\begin{aligned} (1-q)e^{-(y^{\star} - v)^2/2} &= \sqrt{2\pi}\Phi(y^{\star} - v)(qy^{\star} + 2uy^{\star} - uw^{\star}), \\ qe^{-(w^{\star} - v)^2/2} &= \sqrt{2\pi}[1 - \Phi(w^{\star} - v)](uy^{\star} + qw^{\star}). \end{aligned} \quad \text{(S82)}$$

Substituting these expressions into equation (S80) yields, after some algebra, this simple condition for the mode of the distribution, $q^{\star}$:

$$y^{\star}(q^{\star}, u, v(q^{\star})) = w^{\star}(q^{\star}, u, v(q^{\star})). \quad \text{(S83)}$$

Then, if this condition is satisfied, equation (S79) reduces to $\log \frac{1-\Phi(y^\star - v)}{\Phi(y^\star - v)} = \log \frac{q^\star}{1-q^\star}$, which implies

$$\Phi(y^\star - v) = 1 - q^\star. \tag{S84}$$

From this we get

$$y^\star(q^\star, u, v(q^\star)) = v(q^\star) + \sqrt{2}\mathrm{erf}^{-1}(1 - 2q^\star). \tag{S85}$$

Finally we take into account the last expression and use equation (S83) into equation (S56) to obtain

$$\sqrt{2}\gamma u + 2(q^\star + u)\mathrm{erf}^{-1}(1 - 2q^\star) = \frac{e^{-\left[\mathrm{erf}^{-1}(1-2q^\star)\right]^2}}{\sqrt{\pi}} \tag{S86}$$

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

$$\frac{\alpha}{\hat{\mu}} = \frac{e^{-\left[\Phi^{-1}(1-q^\star)\right]^2/2} - \sqrt{2\pi} q^\star \Phi^{-1}(1 - q^\star)}{\sqrt{2\pi}[\Phi^{-1}(1 - q^\star) + \gamma]}, \tag{S87}$$

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

$$\frac{\alpha\gamma}{\hat{\mu}} = \frac{1}{\sqrt{2\pi}}. \tag{S88}$$

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

$$\frac{\alpha\gamma}{\mu} = \frac{n}{\sqrt{2\pi}} \tag{S89}$$

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

In the limit of small interaction strengths ($\hat{\mu} \ll \alpha$) of the mean zero case ($\gamma = 0$), condition (S87) reduces to

$$\frac{k^\star}{n} \approx \frac{1}{2} - \frac{1}{2\pi}\frac{\hat{\mu}}{\alpha} + \frac{1}{4\pi}\left(\frac{\hat{\mu}}{\alpha}\right)^2, \tag{S90}$$

which reproduces the expected (binomial) behavior.

S7 Truncated-Gaussian distributed rates

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

$$P(\boldsymbol{r}) = \frac{1}{Z_n}\exp\biggl(-\sum_{i=1}^{n}\frac{(r_i-\gamma)^2}{2}\biggr)\prod_{j=1}^{n}\Theta(r_j)\ , \tag{S91}$$

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

$$\mathcal{P}_T(\{S\}_k|\boldsymbol{A}) \equiv \int d^k\boldsymbol{x}\biggl(\prod_{i=1}^{k}\Theta(x_i)\biggr)\int d^{n-k}\boldsymbol{z}\biggl(\prod_{j=k+1}^{n}\Theta(-z_j)\biggr)f_T(\boldsymbol{x},\boldsymbol{z}|\boldsymbol{A})\ . \tag{S92}$$

where

$$f_T(\boldsymbol{x},\boldsymbol{z}|\boldsymbol{A}) = \frac{|\det \boldsymbol{A}^{(s)}|}{Z_n}\exp\Bigl(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\xi})^T\boldsymbol{G}(\boldsymbol{x}-\boldsymbol{\xi}) - \frac{1}{2}\|\boldsymbol{z}-\boldsymbol{A}^{(ns)}\boldsymbol{x}-\gamma\boldsymbol{1}_{n-k}\|^2\Bigr) \\ \times\prod_{i=1}^{k}\Theta\Bigl(-(\boldsymbol{A}^{(s)}\boldsymbol{x})_i\Bigr)\prod_{j=k+1}^{n}\Theta\Bigl(z_j-(\boldsymbol{A}^{(ns)}\boldsymbol{x})_j\Bigr)\ . \tag{S93}$$

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

$$(\boldsymbol{A}^{(s)}\boldsymbol{x})_i = (\alpha-\mu)x_i + \mu(\boldsymbol{1}_k^T\boldsymbol{x}) = \alpha x_i + \mu\sum_{\substack{s=1\\ s\neq i}}^{k} x_s. \tag{S94}$$

Since Eq. (S92) forces that $x_i > 0$, and α and μ are both negative, we find that $-(\boldsymbol{A}^{(s)}\boldsymbol{x})_i$ is always positive, i.e., it holds that

$$\Theta\Bigl(-(\boldsymbol{A}^{(s)}\boldsymbol{x})_i\Bigr)\Theta(x_i) = \Theta(x_i). \tag{S95}$$

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

$$\Theta\Bigl(z_j-(\boldsymbol{A}^{(ns)}\boldsymbol{x})_j\Bigr)\Theta(-z_j) = \Theta\Bigl(z_j-\mu(\boldsymbol{1}_k^T\boldsymbol{x})\Bigr) + \Theta(-z_j) - 1. \tag{S96}$$

Now, we apply the same changes of variable leading to Eq. (S27). Then we can write

$$\begin{aligned}\mathcal{P}_T(\{S\}_k|n) = {} & \frac{|\alpha-\mu|^{k-1}|\alpha+(k-1)\mu|}{Z_n}\int d^k\boldsymbol{x}\prod_{i=1}^{k}\Theta(x_i+\xi^{(k)})e^{-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{G}\boldsymbol{x}} \\ & \times\int d^{n-k}\boldsymbol{z}\prod_{j=k+1}^{n}\Big[\Theta(-z_j-\gamma-k\mu\xi^{(k)})+\Theta(z_j-\mu(\mathbf{1}_k^T\boldsymbol{x})+\gamma)-1\Big]e^{-\frac{1}{2}\|\boldsymbol{z}-\mu(\mathbf{1}_k^T\boldsymbol{x})\mathbf{1}_{n-k}\|^2}\ .\end{aligned} \tag{S97}$$

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

$$\begin{aligned}&\prod_{j=k+1}^{n}\Big\{\Theta(-z_j-\gamma-k\mu\xi^{(k)})+\Big[\Theta(z_j-\mu(\mathbf{1}_k^T\boldsymbol{x})+\gamma)-1\Big]\Big\} \\ &\quad=\sum_{\ell=0}^{n-k}\sum_{\substack{p\in C_\ell^{n-k}\\ b=\mathcal{K}\setminus p}}\prod_{j=1}^{\ell}\Theta(-z_{p(j)+k}-\gamma-k\mu\xi^{(k)})\prod_{i=1}^{n-k-\ell}\Big[\Theta(z_{b(i)+k}-\mu(\mathbf{1}_k^T\boldsymbol{x})+\gamma)-1\Big],\end{aligned} \tag{S98}$$

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 $\boldsymbol{z}$, we can decompose

$$\begin{aligned}\mathcal{P}_T(\{S\}_k|n) = {} & \frac{|\alpha-\mu|^{k-1}|\alpha+(k-1)\mu|}{Z_n}\int d^k\boldsymbol{x}\prod_{i=1}^{k}\Theta(x_i+\xi^{(k)})e^{-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{G}\boldsymbol{x}} \\ & \times\sum_{\ell=0}^{n-k}\binom{n-k}{\ell}\prod_{j=k+1}^{\ell+k}\int dz_j\Theta(-z_j-\gamma-k\mu\xi^{(k)})e^{-\frac{1}{2}\left[z_j-\mu(\mathbf{1}_k^T\boldsymbol{x})\right]^2} \\ & \times\prod_{i=\ell+k+1}^{n}\int dz_i\Big[\Theta(z_i-\mu(\mathbf{1}_k^T\boldsymbol{x})+\gamma)-1\Big]e^{-\frac{1}{2}\left[z_i-\mu(\mathbf{1}_k^T\boldsymbol{x})\right]^2}.\end{aligned} \tag{S99}$$

Note now that

$$\int_{-\infty}^{\infty}dz\Big[\Theta(z-\mu(\mathbf{1}_k^T\boldsymbol{x})+\gamma)-1\Big]e^{-\frac{1}{2}\left[z-\mu(\mathbf{1}_k^T\boldsymbol{x})\right]^2}=-\sqrt{2\pi}\Phi(-\gamma). \tag{S100}$$

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

$$\mathcal{P}_T(\{S\}_k|n)=\frac{(2\pi)^{n/2}}{Z_n}\sum_{\ell=0}^{n-k}\binom{n-k}{\ell}\left[-\Phi(-\gamma)\right]^{n-\ell-k}\mathcal{P}(\{S\}_k|\ell+k), \tag{S101}$$

where $\mathcal{P}(\{S\}_k|n)$ is precisely the expression (S27) obtained for the non-truncated Gaussian distribution. According to Eq. (S39),

$$\mathcal{P}(\{S\}_k|\ell{+}k) = \frac{k+s}{2\pi i}\int_{-\infty}^{\infty} dy \int_{\Gamma} dw\, e^{-\frac{1}{2}(k+2s)y^2+syw+\frac{1}{2}kw^2}[1-\Phi(w-v)]^k[\Phi(y-v)]^{\ell}. \quad \text{(S102)}$$

We introduce (S102) into (S101) and use the binomial expansion

$$\sum_{\ell=0}^{n-k}\binom{n-k}{\ell}[-\Phi(-\gamma)]^{n-\ell-k}[\Phi(y-v)]^{\ell} = [\Phi(y-v)-\Phi(-\gamma)]^{n-k} \quad \text{(S103)}$$

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

$$\begin{aligned}\mathcal{P}_T(\{S\}_k|n) = \frac{(2\pi)^{n/2-1}(k+s)}{i\ Z_n}\int_{-\infty}^{\infty} dy \int_{\Gamma} dw\, e^{-\frac{1}{2}(k+2s)y^2+syw+\frac{1}{2}kw^2} \\ \times [1-\Phi(w-v)]^k[\Phi(y-v)-\Phi(-\gamma)]^{n-k}. \quad \text{(S104)}\end{aligned}$$

Note that the only difference with Eq. (S39) 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 $P_T(k|n) \sim e^{nF_T(\boldsymbol{\sigma})}$, where

$$\begin{aligned}F_T(\boldsymbol{\sigma}) := \frac{q}{2}\left(w^{\star 2}-y^{\star 2}\right) + (1-q)\log[\Phi(y^\star-v)-\Phi(-\gamma)] \\ + q\log[1-\Phi(w^\star-v)] - uy^{\star 2} + uy^\star w^\star - q\log q - (1-q)\log(1-q). \quad \text{(S105)}\end{aligned}$$

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

$$\begin{aligned} qy - (1-q)\frac{\Phi'(y-v)}{\Phi(y-v)-\Phi(-\gamma)} + 2uy - uw = 0, \\ qw - q\frac{\Phi'(w-v)}{1-\Phi(w-v)} + uy = 0. \end{aligned} \quad \text{(S106)}$$

As can be easily checked, the condition $\frac{\partial F_T}{\partial q} = 0$ to be satisfied by the mode $q^\star$ leads to the same constraint as in the Gaussian case, $y^\star(q^\star, u, v(q^\star)) = w^\star(q^\star, u, v(q^\star))$, see Eq. (S83). This implies that

$$\Phi(y^\star - v) = 1 - q^\star + q^\star \Phi(-\gamma), \tag{S107}$$

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:

$$[1 - \Phi(-\gamma)]\left[\sqrt{2}\gamma u + 2(q^\star + u)\mathrm{erf}^{-1}(1 - 2q^\star + 2q^\star\Phi(-\gamma))\right] = \frac{1}{\sqrt{\pi}} e^{-\left[\mathrm{erf}^{-1}(1-2q^\star+2q^\star\Phi(-\gamma))\right]^2}. \tag{S108}$$

Figure S1 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 $\gamma \gtrsim 1$.

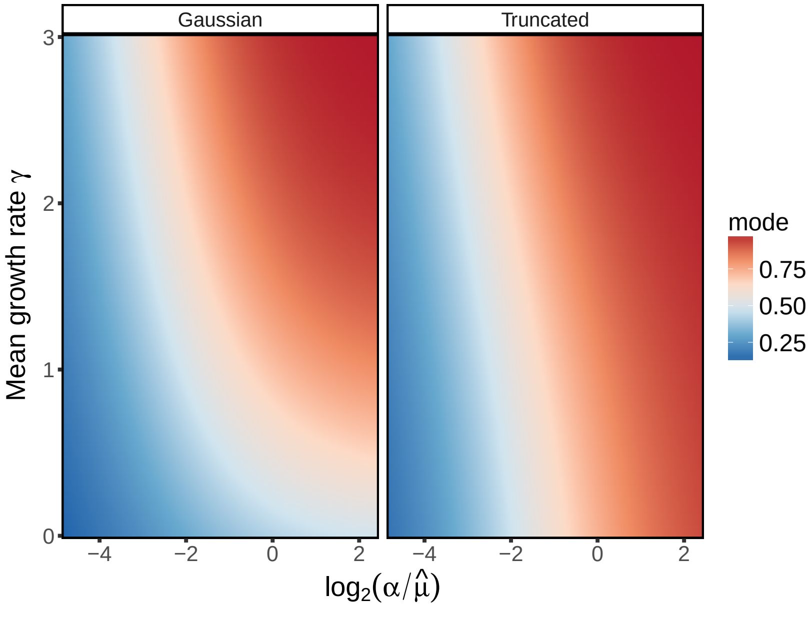

Fig. S1. Comparison between the modes for purely Gaussian and truncated-Gaussian distributed growth rates.

S8 Final communities

Figure S2 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 recently[30], 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 is related to the negative diagonal that we need to add to the matrix in order to ensure stability, pushing the r's values towards the right.

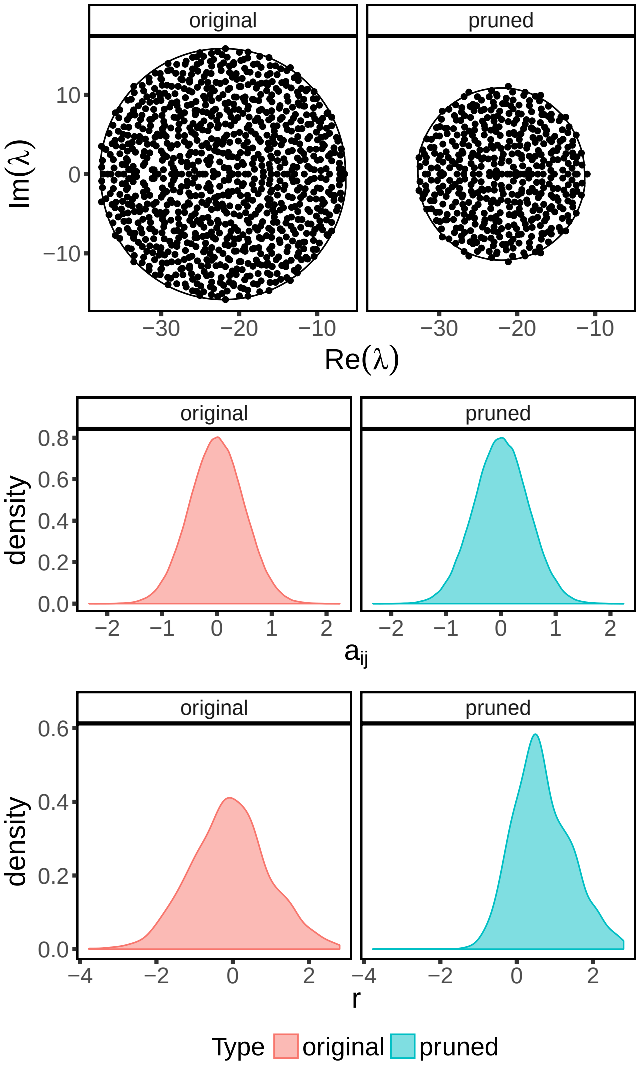

Fig. S2. Comparison between the properties of the original community with $n = 1000$ and the final community, after dynamical pruning, comprising $n = 472$ species. The first row shows the eigenvalue distribution of the matrix of interaction A; for a matrix in which the entries are i.i.d. samples from 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[5,26]. 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.

S9 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.

S9.1 Lemke-Howson algorithm

Because of the equivalence between the Lotka-Volterra and the replicator equation[42], 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 probability[41] (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 algorithm[43] to find such a solution. This algorithm is based on exploring the vertices of the following polytope:

$$P = \{\boldsymbol{z} \in \mathbb{R}^n | \boldsymbol{z} \geq \boldsymbol{0}, \boldsymbol{C}\boldsymbol{z} \leq \boldsymbol{1}\}, \tag{S109}$$

where $\boldsymbol{C}$ 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 $\boldsymbol{z} \in P$ has label k if $z_k = 0$ and label $-k$ if $(\boldsymbol{C}\boldsymbol{z})_k = 1$. 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 $\boldsymbol{z}$ represents strategy k if either it has label k or $-k$, then because of the simplicity assumption any $\boldsymbol{z}$ that represents all strategies is either $\mathbf{0}$ or the normalized vector $\widehat{\boldsymbol{z}} = \boldsymbol{z}/\sum_i z_i$ is a *symmetric Nash equilibrium* for the game.

In order to find the solution we move around the vertices of P starting from $\boldsymbol{v_0} = \mathbf{0}$ using a tableaux $T : \boldsymbol{r} = \mathbf{1} - \boldsymbol{C}\boldsymbol{z}$ with a slack variable $\boldsymbol{r}$. Say that r_k is in the basis for a vertex $\boldsymbol{v} \in P$ if and only if $\boldsymbol{v}$ does not have label $-k$, and z_k is in the basis if and only if $\boldsymbol{v}$ does not have label k. Then $\boldsymbol{v_0}$ has basis $\{r_1, \ldots, r_n\}$, 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 $\boldsymbol{v_1}$. In the next iteration bring z_k to the basis and move to an adjacent vertex $\boldsymbol{v_2}$. We keep repeating this process until we get to a vertex $\boldsymbol{v}$ which represents all strategies, that is, $\boldsymbol{v}$ 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

$$\boldsymbol{A} = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}, \tag{S110}$$

and intrinsic growth rates:

$$\boldsymbol{r} = \begin{pmatrix} -1 \\ 3 \end{pmatrix} \tag{S111}$$

We build the payoff matrix:

$$\boldsymbol{C} = \begin{pmatrix} -2 & 1 & -1 \\ 1 & -2 & 3 \\ 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 4 & 2 \\ 4 & 1 & 6 \\ 3 & 3 & 3 \end{pmatrix}, \tag{S112}$$

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:

$$\begin{aligned} r_1 &= 1 - z_1 - 4z_2 - 2z_3, \\ r_2 &= 1 - 4z_1 - z_2 - 6z_3, \\ r_3 &= 1 - 3z_1 - 3z_2 - 3z_3. \end{aligned} \tag{S113}$$

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:

$$\begin{aligned} r_1 &= \frac{2}{3} + \frac{1}{3}z_1 - \frac{11}{3}z_2 + \frac{1}{3}r_2, \\ z_3 &= \frac{1}{6} - \frac{4}{6}z_1 - \frac{1}{6}z_2 - \frac{1}{6}r_2, \\ r_3 &= \frac{1}{2} - z_1 - \frac{5}{2}z_2 + \frac{1}{2}r_2. \end{aligned} \tag{S114}$$

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

$$\begin{aligned} z_2 &= \frac{2}{11} + \frac{1}{11}z_1 - \frac{3}{11}r_1 + \frac{1}{11}r_2, \\ z_3 &= \frac{3}{22} - \frac{15}{22}z_1 + \frac{1}{22}r_1 - \frac{2}{11}r_2, \\ r_3 &= \frac{1}{22} - \frac{27}{22}z_1 + \frac{15}{22}r_1 + \frac{3}{11}r_2. \end{aligned} \tag{S115}$$

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 :

$$\begin{aligned} z_1 &= \frac{1}{27} + \frac{15}{27}r_1 + \frac{2}{9}r_2 - \frac{22}{27}r_3, \\ z_2 &= \frac{5}{27} - \frac{6}{27}r_1 + \frac{1}{9}r_2 - \frac{2}{27}r_3, \\ z_3 &= \frac{1}{9} - \frac{1}{3}r_1 - \frac{1}{3}r_2 + \frac{15}{27}r_3. \end{aligned} \tag{S116}$$

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.

S9.2 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.

S9.2.1 Mean zero

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

S9.2.2 Mean non zero

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

S9.2.3 Adding Structure

In order to include a network structure, we generate an adjacency matrix $\boldsymbol{G}$ 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 Figure S3.

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: $\hat{\mu} = n\mu C$.

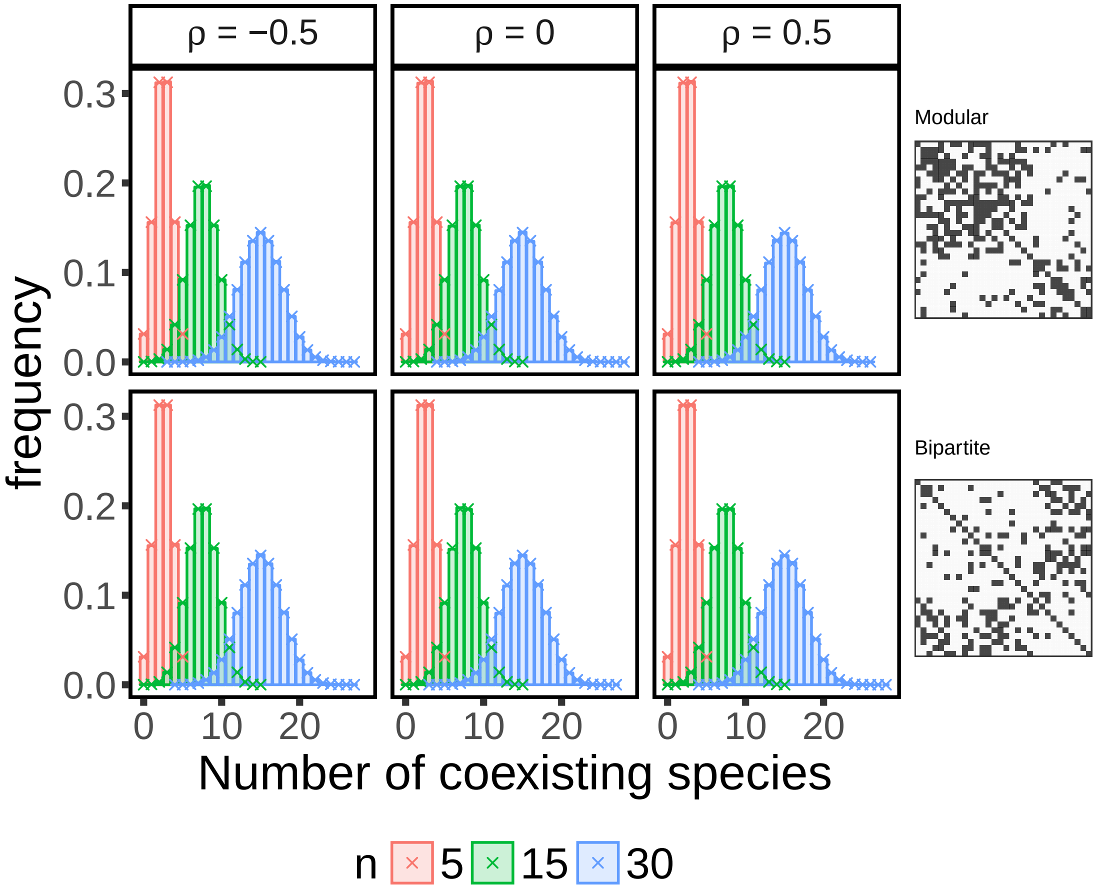

Fig. S3. As Figure 1 of the main text, but with modular (top) or "anti-modular" (i.e., close to bipartite, bottom) structures.

S10 Assembly

So far, we have described the dynamical process associated with equation (S1) 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 Fig. S2). 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 (S7) which cannot be assembled (Fig. S4). 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 (Fig. S5).

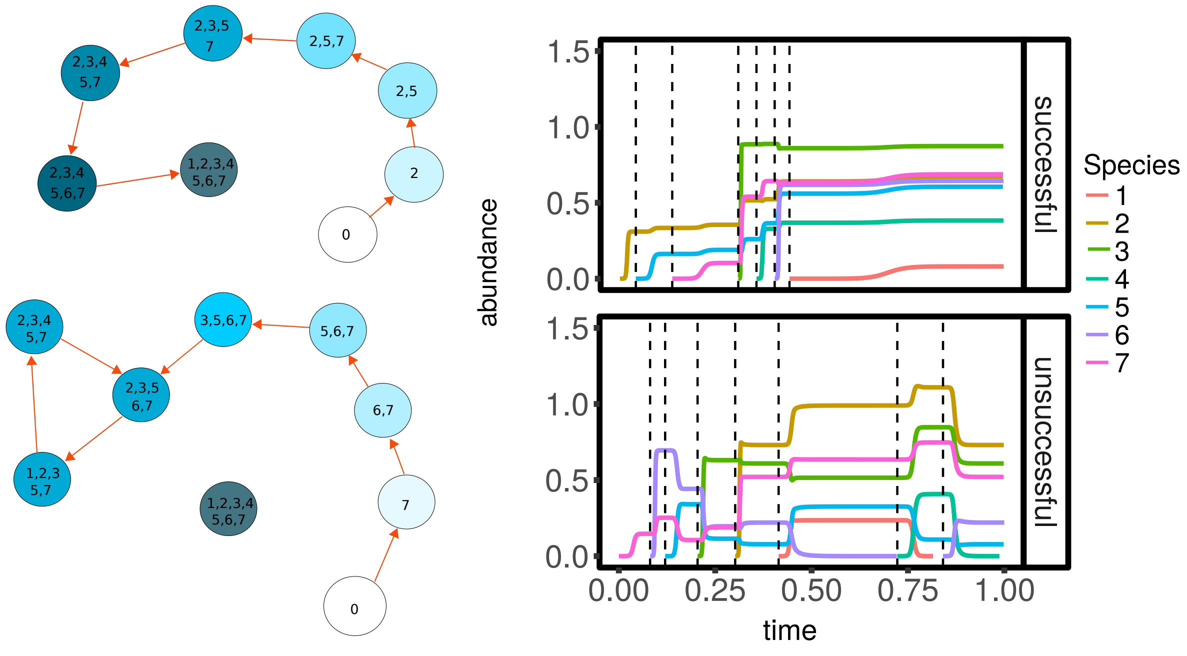

Fig. S4. Assembling communities one species at a time. Top: we want to build the community with species 1, 2, ..., 7 present (darker shades for more speciose communities), by adding a species at a time. Starting from an empty system (state 0), we can try all assembly pathways in which we sequentially 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.

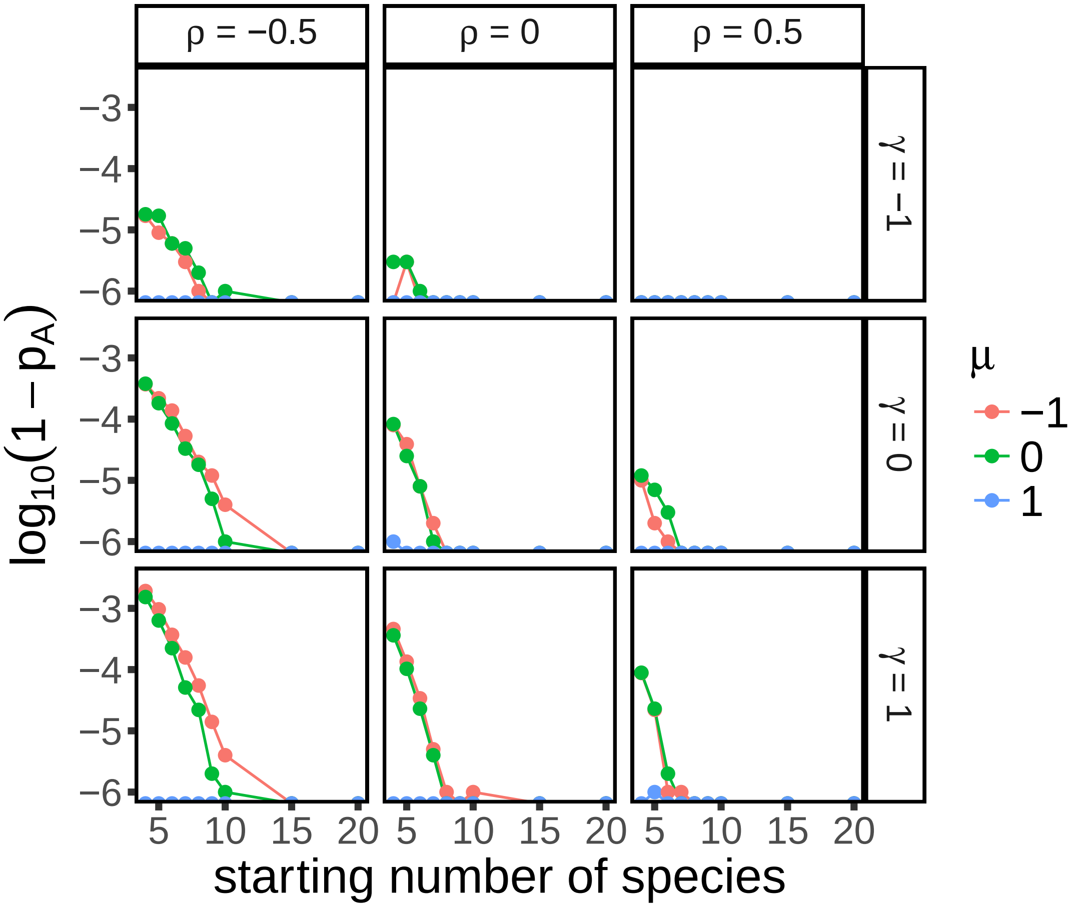

Fig. S5. Probability p_A of finding an assembly path when starting from n species. For different means and correlations of interactions strengths, sampled from a bivariate Normal with mean μ (colors) and correlation ρ (columns), and for different values of mean intrinsic growth rates (γ, rows), we plot 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.