Robust coexistence in ecological competitive communities

Pablo Lechón-Alonso[1], Srilena Kundu[1,2], Paula Lemos-Costa[1], José A. Capitán[3], Stefano Allesina[1,4*]

[1]Department of Ecology & Evolution, University of Chicago, 1101 E. 57th, Chicago, 60637, IL, USA.
[2]Helmholtz Institute for Functional Marine Biodiversity at the University of Oldenburg (HIFMB), Im Technologiepark 5, Oldenburg, 26129, Germany.
[3]Complex Systems Group, Department of Applied Mathematics, Universidad Politécnica de Madrid, Avenida Juan de Herrera, 6, Madrid, E28040, Spain.
[4]Northwestern Institute on Complex Systems, Northwestern University, 600 Foster Street, Evanston, 60208, IL, USA.

*Corresponding author(s). E-mail(s): sallesina@uchicago.edu;
Contributing authors: plechon@uchicago.edu; srilena.kundu@hifmb.de; plemos@uchicago.edu; ja.capitan@upm.es;

Abstract

The notions of intraspecific competition and population self-regulation play a central role in ecological theory, leading to foundational concepts such as limiting similarity and niche differentiation[1–6]. They also are crucial for coexistence: for example, May[7] showed that ecological dynamics can be stabilized by imposing sufficiently strong intraspecific competition, and that for large systems with random interactions the transition from instability to stability is sharp[7–9], and achieved at a critical value of intraspecific competition.

Here we consider the existence of an equilibrium and its stability as the strength of intraspecific competition $\boldsymbol{\alpha}$ is increased. We show that an ecological system can gain or lose a coexistence equilibrium as intraspecific competition becomes more prevalent, but that for any competitive system there exists a critical value $\boldsymbol{\alpha_F}$ after which the existence of an equilibrium is guaranteed, in analogy to stability.

Having derived results on general systems, we turn to ecological systems with random interactions, and compute the probability that a system would display a coexistence equilibrium, given the level of intraspecific interaction, the size of the system $\boldsymbol{n}$, and the statistics on the distribution of interactions—thereby extending, generalizing and refining results that have been accumulating for decades[10–16]. We show that the transition to feasibility is smooth, contrary to what is found for stability, and that the existence of an equilibrium almost invariably guarantees its stability[10, 11, 14] This has strong consequences for the dynamics of such large ecological systems: coexistence via limit cycles or chaos[17, 18] is precluded, and dynamics always converge to an equilibrium. The surviving competitors coexist extremely robustly; the ecological community will return to equilibrium after any perturbation, can be assembled from the ground-up, and is resistant to invasions[19, 20].

Keywords: Feasibility, Stability, Lotka-Volterra, Random matrices

Our goal is to study the effect of intraspecific competition on the existence and stability of an equilibrium where all populations are present, in ecological communities in which competition is the primary mode of interaction. To this end, we consider the Generalized Lotka-Volterra[21–24] (GLV) model for n interacting populations:

$$\dot{x}_i = x_i r_i \left(1 - \sum_j A_{ij} x_j\right) , \tag{1}$$

where intrinsic growth rates are organized in the vector r, and interactions in the matrix A. Following May[7], we model A as the sum of two matrices $A = \alpha I + B$, where B collects the inter-specific interactions, as well as the baseline levels of intraspecific interactions, I is the identity matrix, and $\alpha \geq 0$ can be interpreted as the "excess" intraspecific competition in the system. We call these systems *competitive* whenever $r > \mathbb{0}$ (i.e., all populations can grow in isolation), and $\sum_j B_{ij} > 0$ for every population i (i.e., each population has, on average, a negative effect on the growth of the other populations). Now we consider increasing α from zero to an arbitrary large positive value, and track the effects of this change on coexistence. We say that Eq. (1) has a *feasible* (i.e., biologically attainable) equilibrium whenever there is a $x^\star > \mathbb{0}$ such that $\dot{x}_i = 0$ for all populations, and that such a feasible equilibrium is *stable* if dynamics return to it following a perturbation. May[7] showed that, whenever a feasible equilibrium exists, a sufficiently strong level of intraspecific competition guarantees its stability. Here we denote α_S as the level of intraspecific competition that is sufficient to guarantee the stability of a feasible equilibrium, for any choice of $r > \mathbb{0}$ (SI, Sec. 1.3).

To study the effect of α on the existence of an equilibrium, we start by showing (Methods) that the feasibility of the system in Eq. (1) is unchanged when we subtract to each column of B its mean, obtaining matrix C with columns summing to zero. This shows that the actual strength of competition does not influence feasibility—we can strengthen the competitive effect of certain populations on all others without affecting the feasibility of the equilibrium. Because infinitely many competitive matrices B will map into the same C, we focus on the feasibility of the following *canonical* competitive system.

$$\dot{y}_i = y_i r_i \left(1 - \alpha y_i - \sum_j C_{ij} y_j\right) . \tag{2}$$

Analyzing Eq. (2) we find that, in general, increasing α could either make the system gain or lose the positive equilibrium, and in fact that the system can cycle in and out of feasibility several times as α increases (Fig. 1, SI Sec. 2.3). However, for any competitive system, there exists a critical level of $\alpha = \alpha_F$ such that the equilibrium $y^\star > \mathbb{0}$ exists, and feasibility is maintained if α is increased further (SI, Sec. 2.2). Interestingly, the critical value α_F is an eigenvalue of a matrix derived from that of interaction strengths, and this allows us to devise an algorithm to determine α_F for arbitrary systems (Methods). These findings parallel what is observed for stability, where a critical level $\alpha = \alpha_S$, which is also an eigenvalue of a matrix derived from B[7–9], ensures the global stability of dynamics.

Having derived results that apply to arbitrary competitive GLV systems, we turn to the class of systems with random interactions in which populations are *statistically equivalent* (i.e., there is no *a priori* structure making some populations behave differently from the rest). Here we consider the simplest model in this class, in which each pair of interactions (B_{ij}, B_{ji}) is sampled from a bivariate distribution with mean $\mu > 0$, unit variance, and correlation ρ (SI, Sec. 3.1). The stability properties of these models are well understood[8, 9], and results on their feasibility have been accumulating in the literature for decades [25–28] (SI, Sec. 3.2).

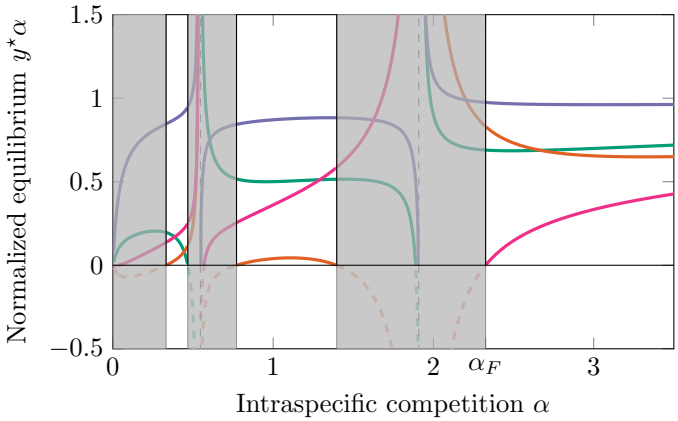

Fig. 1 Components of $y^\star$ (y-axis) for the system in Eq. (2) when α (x-axis) is varied. Areas where a feasible equilibrium does not exist are shaded in gray, and the colors report the densities of the populations at equilibrium. Increasing α can make the system gain or lose the feasible equilibrium: for this example, values of α belonging to three disjoint intervals yield a feasible equilibrium. Moreover, some particular values of α (two, in this case) can lead the values $y_i^\star$ to diverge to $\pm\infty$. For any competitive system, however, there is a critical value of $\alpha = \alpha_F$ such that the equilibrium is feasible, and remains feasible for any value $\alpha > \alpha_F$.

The first contribution to this problem can be traced back to a response to May's work on stability, in which Roberts[10] showed that the probability of stability, conditioned on the presence of a feasible equilibrium, would increase with system's size. Stone[11, 12] derived an approximation for the probability of observing a feasible equilibrium for a given level of intraspecific competition α when $\mu = 0$ and $\rho = 0$, and Bizeul & Najim[13] found that in this case there is a transition from low to high probability of feasibility around $\alpha \approx \sqrt{n}\sqrt{2\log n}$.

The probability that a feasible equilibrium exists when $\alpha = 0$ decays very rapidly as the size of the system increases (SI Sec. 3.3); however, we have shown that a sufficiently large $\alpha = \alpha_F$ ensures feasibility—we therefore track how the probability of feasibility varies as we increase α. Thanks to the results we have derived for arbitrary systems, we can extend and generalize previous findings, and show that in large systems with random interactions, feasibility is typically gained only at the critical level α_F, and thus there is a single transition to feasibility (Fig. 2). This is in contrast with smaller or arbitrary systems, for which several transitions can be observed (Fig. 1). We then approximate the probability of feasibility, $p_F(\alpha)$, given the level of intraspecific competition, the size of the system, and the statistics of the distribution of interaction strengths: we consider three distinct approaches, each revealing a key feature of the problem.

We start by deriving an approximation for the limit of many interacting populations, based on the theory of random matrices perturbed by small-rank matrices[29] (SI, Sec. 3.6). While this approximation is not very accurate (especially for low α), it shows that, contrary to current expectations, the effect of the pairwise correlation ρ on the probability of feasibility remains significant for large systems; interestingly, exactly as a negative correlation between interaction strengths improves the chance of stability[8], it also increases the probability of feasibility[15]. This derivation also shows that the details of the distribution of interaction strengths, once accounted for the mean, variance and correlation, are irrelevant when the system is large—asymptotically (i.e., as $n \to \infty$), the curve describing the probability of feasibility is *universal*. Using this approach, we also recover the critical value of Bizeul & Najim[13], and extend it to the case of correlated entries (SI, Sec. 3.6).

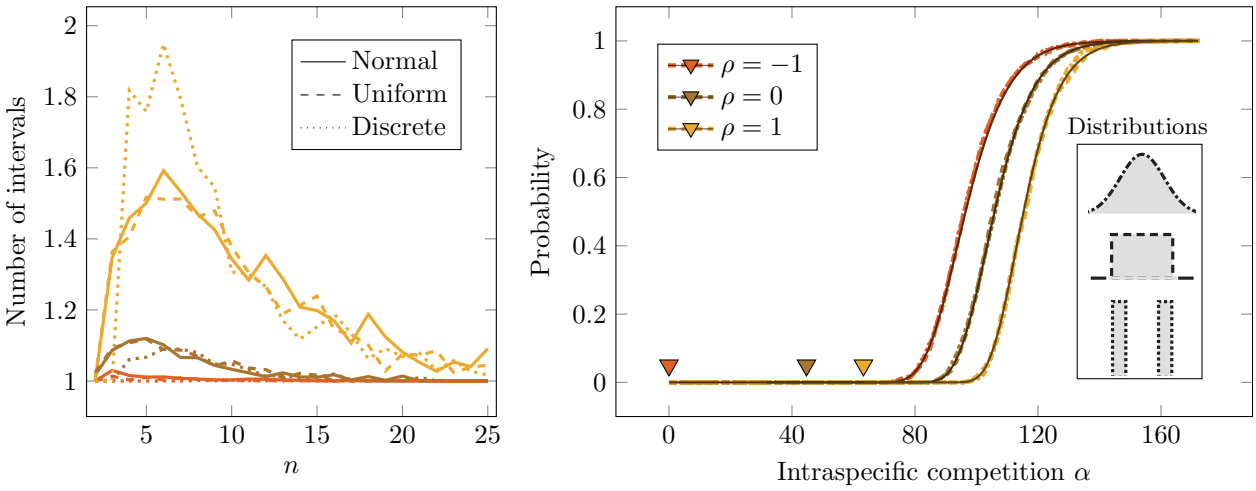

Fig. 2 Transition to feasibility for the GLV competitive system in Eq. (1) with random interactions. Left: while for small or arbitrary systems a feasible equilibrium can appear and disappear several times as we increase α, in random systems the number of intervals containing values of α that yield feasibility converges to one as n increases (contrast with Fig. 1, displaying three such intervals). Hence, these systems typically gain feasibility only at α_F, and maintain it for any larger α. The panel shows simulation results in which the interactions are sampled in pairs from a bivariate distribution with positive mean, unit variance, and correlation ρ (colors), and the number of disjoint intervals of α yielding feasibility are averaged (y-axis) for a given value of n (x-axis). To show that results are asymptotically the same irrespective of the choice of distribution, we sample entries from distributions whose marginals are either normal, uniform, or take a positive or negative value with equal probability ("discrete", SI Appendix A). The line type marks the choice of the distribution used to sample interactions. Right: probability of feasibility for a given level of intraspecific competition (α, x-axis) for systems of size $n = 1000$, and interactions sampled from different distributions (line types) and correlation values (colors). The curves for different distributions collapse to the same curve, showing that the probability of feasibility is asymptotically *universal* (i.e., the higher moments of the distribution have no effect). However, even for such large matrices, the effect of the correlation is noticeable, with negative ρ increasing the probability of feasibility for any α, and a positive ρ decreasing it—remarkably, the same is found for stability[8]. For each choice of correlation, the value $\alpha = \alpha_S$ at which the stability of a feasible equilibrium is guaranteed is marked by a triangle: note that the order of these points is the same as that of the curves describing the probability of feasibility, and that moreover the probability of feasibility at α_S is negligible. The black solid lines represent the analytical expectations obtained by approximating the variance using resolvents —Eqs. (3) and (4).

Next, we show that, for any model in which populations are statistically equivalent, the distribution of the equilibrium values has always the same mean and correlation structure (Methods). In particular, considering Eq. (2), one can see that the average of $y_i^\star$ is fixed (because $\sum_j C_{ij} = 0$), and equal to $1/\alpha$. Because populations are statistically equivalent, they all have the same expectation and variance; but then, given that their sum is fixed, their correlation is necessarily $\mathbb{C}\mathrm{or}(y_i^\star, y_j^\star) = -1/(n-1)$. Hence, different models of random interactions (e.g., with and without correlated interaction strengths, with higher-order correlations, with cyclical structure, etc.) will yield the same mean and correlation for the components of $y^\star$, but different variances. Moreover, if we assume that the components of $y^\star$ are approximately normally distributed[11], and we neglect the negative correlation when n is large, computing the variance $\sigma^2_{y^\star}$ (which in general will depend on n, α, ρ, etc.) is sufficient to approximate the probability of feasibility (black solid lines in Fig. 2) as:

$$p_F(\alpha) \approx \Phi\left(\frac{1}{\alpha\sigma_{y^\star}}\right)^n \quad , \tag{3}$$

where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal.

The variance $\sigma^2_{y^\star}$ can be approximated in different ways, and we provide two approaches. We start by showing that one can use the theory of resolvents to estimate the variance for any ρ (SI, Sec. 3.8), producing an accurate approximation (Fig. 2) that does not depend on the higher moments of the distribution of

interaction strengths—again pointing at the asymptotic universality of the results. We obtain (SI, Sec. 3.8):

$$\frac{1}{(\alpha\sigma_{y^\star})^2} \approx \frac{\alpha^2 + \alpha\sqrt{\alpha^2 - 4\rho(n-1)}}{2(n-1)} - (1+\rho) \quad , \tag{4}$$

showing that the variance grows (and hence the probability of feasibility decreases) with ρ.

Finally, we use the theory of iterative solutions of linear systems to provide an alternative approximation of $\sigma^2_{y^\star}$ (SI, Sec. 3.9), which now depends on all the moments of the distribution of B_{ij} (though the contribution of higher-order moments is negligible for large n). In this way, one can derive an accurate approximation for smaller systems, and show that for example the (transient) effect of skewed interaction strengths is similar to that of correlation, with a positive skew reducing the probability of feasibility, and a negative skew increasing it (Sec. 3.9).

In all cases, we find that the probability of feasibility at the critical level for stability $p_F(\alpha_S)$ is negligible when n is large (as argued before[11, 13, 14]). Then, for a large system, whenever feasibility is attained, stability is almost surely guaranteed, i.e. $\alpha_S < \alpha_F$, and thus $p_F(\alpha)$ measures the *probability of coexistence* for a large, random competitive system. These findings reinforce the notion that, in competitive systems, feasibility is key for coexistence[5, 10, 16, 30]. This is in analogy with consumer-resource models[31], as well as in GLV models with symmetric interactions produced by niche overlap or species' similarity[32], for which stability is guaranteed, and thus feasibility determines coexistence[20, 33].

These results have profound implications for dynamics. Competitive GLV systems can give rise to any type of dynamics, including limit cycles (for three or more populations) and chaos (when $n \geq 4$)[17, 18]. Importantly, however, non-equilibrium coexistence still requires the existence of a feasible equilibrium[19]. We have seen that in random competitive systems feasibility generically implies global stability, and therefore non-equilibrium coexistence is precluded: for the systems analyzed here, dynamics will almost surely converge to an equilibrium. This observation is in line with empirical results: experimenting with organisms at the same trophic level (for which competition is believed to be the dominant mode of interaction) often results in equilibrium dynamics [34–38], while sustained cycles and chaos are typically associated with predator-prey systems and structured populations[39–44].

One can go a step further, and imagine starting a large competitive system defined by $r > \mathbb{0}$, random interactions B, and a given level of intraspecific competition α (Figs. 3 and 4). If $\alpha_S < \alpha_F < \alpha$, then both the feasibility and global stability of the coexistence equilibrium are guaranteed, and thus all populations coexist robustly, returning to equilibrium after any perturbation short of extinction (Fig. 3, c). When $\alpha_S < \alpha < \alpha_F$, on the other hand, the lack of feasibility implies that some populations will go extinct (Fig. 3, b); however, because $\alpha_S < \alpha$, i) the same populations will go extinct irrespective of the (positive) initial conditions, and ii) those that go extinct cannot re-invade the system (i.e., the equilibrium is *saturated*[19], SI Sec. 1.3). The most interesting case is obtained when $\alpha < \alpha_S < \alpha_F$: because there is no feasible equilibrium, extinctions are guaranteed; however, because $\alpha < \alpha_S$, the identities of the surviving populations depend on the initial

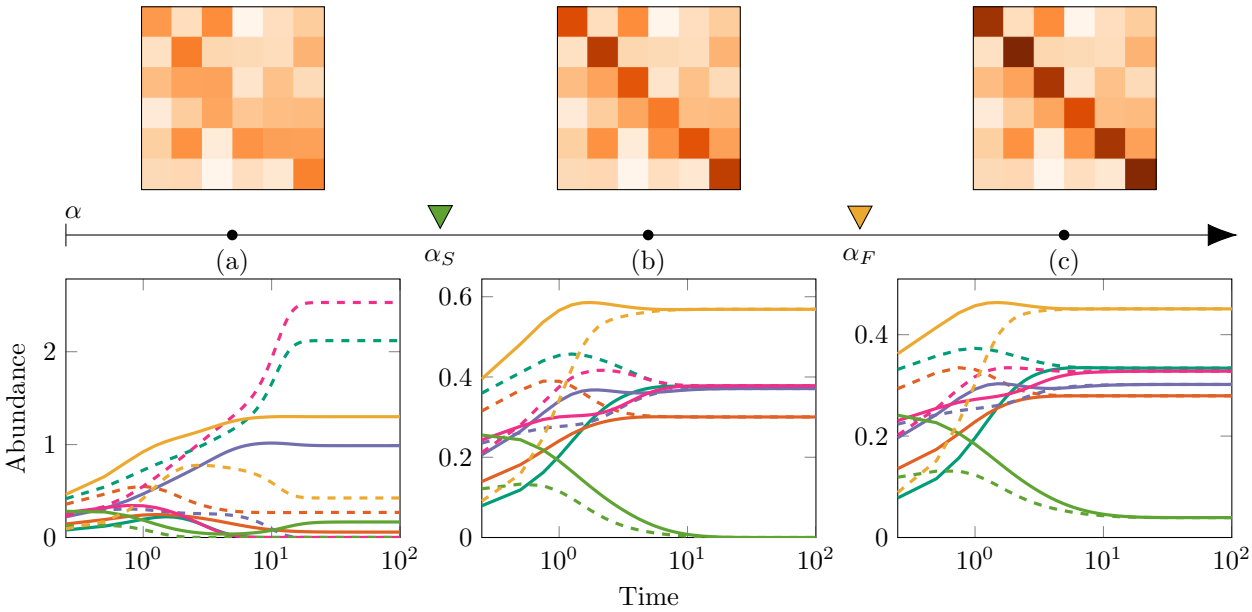

Fig. 3 Three scenarios for dynamics in random competitive communities. Top: We consider the dynamics for three different values of α, when $A = \alpha I + B$ (Eq. 1), and B is a small random matrix with independent entries sampled from a distribution with positive mean (the three matrices, with the same off-diagonal coefficients and increasingly strong diagonal coefficients, are drawn above the line). We compute two critical values for B: α_S guarantees stability of a feasible equilibrium, and α_F guarantees the existence of a feasible equilibrium. c) When $\alpha_S < \alpha_F < \alpha$, a feasible equilibrium exists, and is globally stable; when initializing the population densities (colors) at different values (e.g., dashed vs. solid lines), dynamics converge to the equilibrium. b) When $\alpha_S < \alpha < \alpha_F$, the equilibrium is not feasible, but any feasible equilibrium for a sub-community is stable, and dynamics always converge to a saturated equilibrium (SI, Sec. 1.3); then, the same populations go extinct irrespective of initial conditions (in this case, the green population), while the remaining populations converge to a stable equilibrium. a) Finally, when $\alpha < \alpha_S < \alpha_F$, depending on the initial conditions, the set of coexisting populations might differ (priority effects or alternative stable states; contrast the dashed and solid lines); when extinctions ensue, however, both the values of α_F and α_S decrease, such that, after a sufficiently large number of extinctions, the sub-system goes back to the dynamics found for case b).

conditions (Fig. 3, a); we thus can observe priority effects and alternative stable states. Once a population goes extinct, however, both α_S and α_F are reduced for the surviving sub-community, while α is largely unaltered (because all populations have, in expectation, the same α). Hence, as the number of extinctions mounts, eventually α will cross α_S, and the system will settle at a (reduced) saturated equilibrium. Thus, as long as the relationship $\alpha_S < \alpha_F$ is maintained as populations go extinct (i.e., as long as the number of surviving populations is large enough), the outcome is always qualitatively the same: a sub-set of the species pool will coexist robustly at an equilibrium that is characterized by $\alpha_S < \alpha_F < \alpha$; the equilibrium is stable, and the last few populations to go extinct (in case a), or all the extinct populations (in case b) cannot re-invade the system.

These three scenarios are portrayed in Fig. 4, where we sample a pool of n populations, setting $\alpha < \alpha_S < \alpha_F$ (thus, starting the system at scenario a) of Fig. 3), and consider two different assembly methods. First, we introduce all populations at the same time ("top-down assembly" [20]; Fig. 4, left panels). As time increases and dynamics unfold, we track the values of α_F and α_S, which change after each extinction. We can observe how the community traverses scenarios (a), (b), and (c) of Fig. 3 during its assembly, with dynamics converging to a stable equilibrium characteristic of scenario (c). Further, we consider mainland-island dynamics (Fig. 4, right panels) and ask whether the community resulting from "top-down assembly" can be built from the bottom-up, via successive invasions. Again, we track α_F and α_S as we add populations to the community, and observe that i) we can choose the order of invasions such that there are no extinctions ($\alpha_F < \alpha$), ii) at

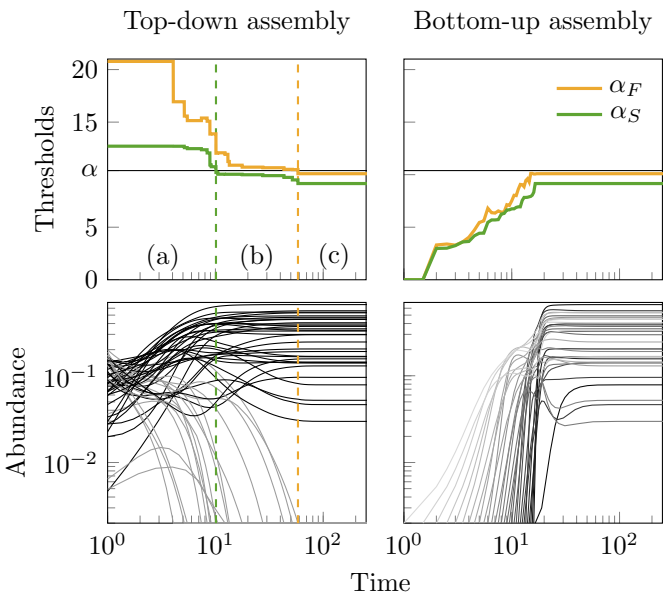

Fig. 4 Top-down and bottom-up assembly of a competitive community. Left: we initialize a community of 50 populations with random interactions, setting $\alpha = \alpha_F/2$; then, there is no feasible equilibrium, and extinctions are guaranteed; note that for this simulation we also have $\alpha_S > \alpha$, and thus we are in scenario a) of Fig. 3. As populations are lost to extinction (marked in grey in bottom-left panel), the values of α_F and α_S (yellow and green lines respectively, top left panel) for the surviving sub-community change. As populations are lost, eventually α_S crosses α (dashed green line, top left), and thus the system transitions to case b) of Fig. 3. Extinctions continue until α_F falls below α (yellow dashed line, top left), and dynamics follow case c) of Fig. 3. The extant populations coexist at a globally stable equilibrium, and those that went extinct since α_S crossed α cannot re-invade the system. Right: we assemble the final community, found via "top-down" assembly, by sequentially invading the system, one population at a time ("bottom-up" assembly). The simulation shows a case in which at each invasion (bottom-right, color corresponding to the time of invasion) the system grows in size, without extinctions, since α_F remains below α. Additionally, we have $\alpha_S < \alpha_F$ throughout assembly, such that scenario c) always applies. As populations are added to the community, both α_F and α_S grow, eventually reaching the values found for top-down assembly.

each successive invasion, the community settles at an equilibrium ($\alpha_S < \alpha$), and iii) both α_F and α_S grow as we add species to the community, eventually reaching the same values yielded by the top-down assembly. This is in fact a general feature of large random systems: the probability that a feasible community can be built in this way rapidly converges to one as the number of populations increases (SI, Sec. 3.11).

In summary, we have shown that, for large competitive communities with random coefficients, the probability of non-equilibrium coexistence is negligible, and thus chaos or limit cycles will not be observed; moreover, the space for priority effects is greatly reduced: when considering the set of coexisting populations, we find that each population rebounds from perturbations, and that the community can be built through successive invasions—the hallmark of robust coexistence [4, 6].

Acknowledgements. A. Skwara, Z.R. Miller, and C. Serván for discussion. JAC acknowledges financial support from grant PRIORITY (PID2021-127202NB-C22), funded by MCIN/AEI/10.13039/501100011033 and "ERDF. A way of making Europe". We acknowledge the University of Chicago's Research Computing Center for their support of this work.

Supplementary Information. Supplementary Information accompanies this work.

Author contributions. SA conceived and designed the study, and drafted the manuscript. PLA drafted the code and the figures. SA and JAC performed the derivations; JAC derived the approximation using resolvents; SA those based on iterative solutions and small-rank perturbations. All Authors contributed to

the mathematical derivations, the implementation of the code, edited the manuscript, and discussed and interpreted the results.

Competing interests. The authors declare no competing interests.

Code availability. Code to reproduce all findings is available at `https://github.com/StefanoAllesina/robust_coexistence`

Methods

A canonical form for competitive systems

We consider the competitive GLV system in Eq. (1), with $r > \mathbb{0}$, $A = \alpha I + B$ and $B^T \mathbb{1} > \mathbb{0}$. We can further decompose the matrix $A = \alpha I + C + \mathbb{1} m^T$, where C is the matrix obtained by subtracting from each column of B its (positive) mean, and m is a vector containing the column means. We can write the solution of $(\alpha I + C + \mathbb{1} m^T) x^\star = \mathbb{1}$ as a function of the solution of the simpler system $(\alpha I + C) y^\star = \mathbb{1}$ (SI, Sec. 2.1), finding:

$$x^\star = \frac{1}{1 + m^T y^\star} y^\star \quad .$$

Then, $x^\star$ and $y^\star$ are proportional to each other, and thus $x^\star > \mathbb{0}$ requires all the components of $y^\star$ to have the same sign. However, because the sum of all the components of $y^\star$ is positive (as seen when summing all the equations: $\mathbb{1}^T (\alpha I + C) y^\star = \alpha \mathbb{1}^T y^\star = n$), and $m > \mathbb{0}$ (because the system is competitive), then $x^\star$ is feasible if and only if $y^\star$ is feasible. Because infinitely many competitive matrices B will map into the same C, we say that this is the *canonical* matrix of interactions for the competitive system.

Intraspecific competition guaranteeing feasibility

One can show that, in competitive systems, there exists a level of intraspecific competition at which the feasibility of the equilibrium is guaranteed, and further strengthening intraspecific competition retains feasibility (SI, Sec. 2.2). We denote this critical level as α_F. To find this value for a given matrix of interactions strengths B, we consider the eigenvalues of several matrices derived from B (or, equivalently, C, SI Sec 2.3). In particular, consider the real, positive eigenvalues of $-C = \mathbb{1} m^T - B$, and call α_0 the largest value such that the components of the solution $(\alpha_0 I + C) y^\star = \mathbb{1}$ diverge to $\pm\infty$ (if it exists, otherwise $\alpha_0 = 0$). Then, we form n matrices Q_i, each of size $n - 1$. Each Q_i is obtained as the sum of two matrices, $Q_i = \mathbb{1} (b^{(i)})^T - B^{(i)}$, where $b^{(i)}$ is the ith row of matrix B with the ith element removed, and $B^{(i)}$ is the matrix B with the ith row and the ith column removed. Denote as α_i the maximum real eigenvalue of Q_i (if it exists, otherwise $\alpha_i = 0$). Finally, we have that the value α_F that guarantees the feasibility of the system is $\alpha_F = \max\{\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_n\}$.

Interestingly, the value of intraspecific competition sufficient to guarantee the global stability of any feasible equilibrium is also an eigenvalue, this time the largest positive eigenvalue of $-(B + B^T)/2$ (SI, Sec. 1.3).

Distribution of $y^\star$ for systems with random parameters

For random models in which populations are *statistically equivalent* (SI, Sec. 3.1), we have that the expectations $\mathbb{E}(y_i^\star) = \mathbb{E}(y_j^\star)$, the variances $\mathbb{V}(y_i^\star) = \mathbb{V}(y_j^\star) = \sigma^2_{y^\star}$, and the correlation $\mathbb{C}\mathrm{or}(y_i^\star, y_j^\star)$ is the same for all pairs. Note that the sum of the components of $y^\star$ is fixed: $\mathbb{1}^T y^\star = n/\alpha$, and therefore the mean correlation must be negative; because it has to be the same for all pairs, it is exactly $\mathbb{C}\mathrm{or}(y_i^\star, y_j^\star) = -1/(n-1)$, as shown also by calculating it explicitly (SI, Sec. 3.7). Moreover, if one approximates the distribution of $y^\star$ as a multivariate normal distribution, finding the unknown variance (which will depend on the exact definition of the model, the statistics of the distribution, etc.) is sufficient to approximate the probability, by either ignoring

the correlation between the components (because for large n it is negligible), or by computing an integral over the positive orthant (SI, Sec. 3.7).

Supplementary Information

1 Preliminaries: competitive Generalized Lotka-Volterra models

We consider the dynamics of ecological communities composed on n interacting populations, and following the Generalized Lotka-Volterra (GLV) model[21, 22, 24]:

$$\dot{x}_i = x_i r_i \left(1 - \sum_j A_{ij} x_j\right), \tag{1}$$

where $x_i(t) \equiv x_i$ is the density of population i at time t, $\dot{x}_i \equiv dx_i(t)/dt$ its time derivative, r_i is the intrinsic growth/death rate of population i, achieved when growing in isolation at low abundance, and A_{ij} is the amount by which population j depresses/enhances the growth of population i.

We analyze cases in which all populations can grow in isolation ($r > \mathbb{0}$), and, moreover, interactions are "mostly competitive" (defined below). Following May[7], we write A as the sum of two matrices:

$$A = \alpha I + B \quad ,$$

with $\alpha \geq 0$. Then, B collects all the inter-specific interactions, as well as the baseline levels of intraspecific interactions, and α can be interpreted as the "excess" intraspecific competition in the system.

We call the system *competitive* whenever $r > \mathbb{0}$, and $\frac{1}{n} B^T \mathbb{1} = m > \mathbb{0}$ (i.e., the column means of B are all positive, and thus the columns of B sum to positive values). While not *every* interaction needs to be competitive, each population must, on average, depress the growth of the others.

An equilibrium $x^\star$ for this system is a point at which all the derivatives in Eq. (1) vanish. We say that the equilibrium is feasible if $x^\star > \mathbb{0}$. In this work, we analyze the effect of α on the existence of a feasible equilibrium for competitive GLV systems.

1.1 Local stability

The goal of this section is to show that there exists a level of $\alpha = \alpha_L$, corresponding to the celebrated May's bound[7], that guarantees that all the eigenvalues of A have positive real part, and discuss its implication for local stability.

The eigenvalues of A are related to the eigenvalues of B, which are generically complex numbers of the form $\lambda_j(B) = \gamma_j + i\beta_j$, with $i = \sqrt{-1}$. Because the matrix B is composed of real numbers, whenever $\lambda_j(B) = \gamma_j + i\beta_j$ is an eigenvalue, then $\lambda_k(B) = \gamma_j - i\beta_j = \lambda_j^*(B)$, i.e., its conjugate, is also an eigenvalue. Adding a constant diagonal αI to B results in a shift in the real part of all the eigenvalues: $\lambda_j(A) = \lambda_j(\alpha I + B) = \alpha + \gamma_j + i\beta_j$. Hence, whenever $\alpha > \alpha_L = \min\{0, \min_j \mathcal{R}e(\lambda_j)\}$, all eigenvalues of $\alpha I + B$ are contained in the right half of the complex plane, and have thus positive real part.

This is important for the local stability of an equilibrium of the GLV system in Eq.(1). An equilibrium $x^\star$ is said to be *locally asymptotically stable* if trajectories originating arbitrarily close to it eventually converge to it. An equilibrium is locally stable if all the eigenvalues of the Jacobian matrix of the system of differential

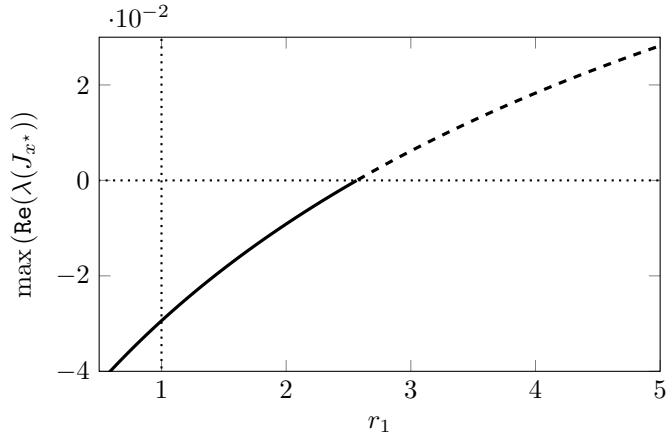

Fig. 5 Bifurcation diagram obtained by varying the growth rates for the three species system in Eq. (2). When the growth rates are chosen as $r = (r_1, 1, 1)^T$, the system experiences a transition from stability to instability as the maximum of $\mathcal{R}e(\lambda(J|_{x^\star}))$ changes sign from negative to positive. The bifurcation occurs at about $r_1 = 2.57$ as r_1 is varied in the range $[1/2, 5]$, but the system remains feasible throughout.

equations, evaluated at the equilibrium, have negative real part. The Jacobian for Eq.(1) evaluated at $x^\star > \mathbb{0}$ is simply:

$$J|_{x^\star} = -D(x^\star \circ r)A \quad ,$$

where $D(z)$ is a matrix with the vector z on the diagonal, and $\circ$ represents the Hadamard (element by element) product. Hence, if all the eigenvalues of A have positive real part, and we were to choose r such that $r_i = 1/x_i^\star$, the equilibrium would be locally asymptotically stable.

Whenever $A = \alpha I + B$, we have that both A and $x^\star$ depend on α, while the vector r does not. We can write:

$$J|_{x^\star} = D(r)(-D(x^\star)A) \quad ,$$

to stress the fact that only the part in parenthesis is influenced by α. Importantly, we could have that $-D(x^\star)A$ has only eigenvalues with a negative real part, while $J|_{x^\star}$ does not. For example, take the competitive GLV system with:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 0 \\ 0 & 3 & 1 \end{pmatrix} \qquad x^\star = \frac{1}{17}\begin{pmatrix} 1 \\ 5 \\ 2 \end{pmatrix} \quad . \tag{2}$$

The system is feasible (for any choice of $r > \mathbb{0}$). However, the local stability of the equilibrium depends on the choice of r. For $r = \mathbb{1}$, we have that the eigenvalues of $J|_{x^\star}$ are $(-1, -\frac{1}{34}(1 \pm i\sqrt{39}))$ and the equilibrium is thus locally stable. However, by increasing the value of r_1 we observe a bifurcation: the real part of the maximum eigenvalue becomes positive and the system loses stability, as shown in Figure 5.

Thus, in general, the stability of a feasible equilibrium $x^\star > \mathbb{0}$ might depend on the choice of r. However, if A is "sufficiently" stable, then the stability of any *feasible* equilibrium is guaranteed for *any* choice of $r > \mathbb{0}$, as discussed below.

1.2 Types of dynamics

Despite its simplicity, the competitive GLV model can give rise to any type of dynamics, including limit cycles and chaos. For purely competitive systems with positive growth rates, when there is a single population, dynamics will always converge to an equilibrium; for two populations, dynamics will generically converge to an equilibrium at which point either population, or both, are present; for three or more populations, one can observe limit cycles, and chaos for four or more populations[17, 18].

Importantly, a necessary condition for the coexistence of all the populations is the existence of a feasible equilibrium $x^\star > \mathbb{0}$, which is also the time-average of the orbits[19]. When such a feasible equilibrium does not exist, trajectories can either reach the boundaries of $\mathbb{R}^n_+$ (i.e., populations undergo extinction), or escape to infinity. When interactions are purely competitive, populations limit each others' growth and divergences to infinity are precluded; therefore, in the absence of a feasible equilibrium, some populations will necessarily go extinct.

1.3 Global stability and saturated equilibria

The goal of this section is to show that there exists a further bound $\alpha = \alpha_S \geq \alpha_L$ that guarantees that a) any feasible equilibrium is locally asymptotically stable, i.e., whenever the densities are slightly perturbed away from the equilibrium, they will eventually return to it; b) the equilibrium is in fact *globally* stable, i.e., any perturbation short of extinction will eventually subside; c) dynamics are guaranteed to result in an equilibrium; d) this equilibrium is *saturated*: some of the populations coexist at a feasible, globally stable equilibrium, and the rest of the populations are extinct and cannot re-invade the system when introduced at low abundance.

In the section discussing local stability, we have shown that choosing $\alpha = \alpha_L$ guarantees local stability for a specific choice of r. However, whenever $A + A^T$ is positive definite (i.e., has only positive eigenvalues), then a) dynamics will always converge to an equilibrium, $\bar{x}$; b) the equilibrium $\bar{x}$ is saturated, meaning that if any population goes extinct, it cannot re-invade the system when starting from low abundance; c) $\bar{x}$ is globally stable (i.e., for any choice of $r > \mathbb{0}$ and initial conditions $x > \mathbb{0}$, dynamics asymptotically converge to $\bar{x}$).

We adapt the proof found for example in Hofbauer & Sigmund[19], because this result is important for our discussion on the dynamics of large competitive communities with random interactions.

We order the vector of densities of the populations in Eq. (1) as $x = (y, z)^T$, where y are the densities of the (unknown) set of the k populations that coexist at the saturated equilibrium, and z the densities of the $n - k$ populations that are extinct at the saturated equilibrium. Therefore $\bar{x} = (y^\star, \mathbb{0})^T$. We also consider a vector of positive weights $w > \mathbb{0}$, which we will use to prove stability. We partition the vector of growth rates, the vector of weights w, and the matrix of interactions according to whether the populations belong to the

set that coexists, or that that goes extinct:

$$w = \begin{pmatrix} u \\ v \end{pmatrix} > \mathbb{0} \quad r = \begin{pmatrix} a \\ b \end{pmatrix} > \mathbb{0} \quad A = \begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{pmatrix} \quad .$$

Next, we consider the candidate Lyapunov function[45]:

$$V(y,z) = \sum_i u_i \left(y_i - {y_i}^\star - {y_i}^\star \log \frac{y_i}{{y_i}^\star} \right) + \sum_l v_l z_l \quad .$$

The function is positive for any $y > \mathbb{0}$ and $z > \mathbb{0}$ (i.e., for any positive initial conditions), given that $u, v, y^\star$ are all positive. Moreover, the function is exactly zero at the saturated equilibrium $\bar{x} = (y^\star, \mathbb{0})^T$. We choose $u_i = 1/a_i$, and $v_l = 1/b_l$, and then differentiate with respect to time:

$$\frac{dV(y,z)}{dt} = \sum_i (y_i - {y_i}^\star) \left(1 - \sum_j \mathcal{A}_{ij} y_j - \sum_k \mathcal{B}_{ik} z_k \right) + \sum_l z_l \left(1 - \sum_j \mathcal{C}_{lj} y_j - \sum_k \mathcal{D}_{lk} z_k \right) \quad .$$

We define $\Delta y_i = y_i - {y_i}^\star$; noticing that $1 = \sum_j \mathcal{A}_{ij} {y_j}^\star$ we can write:

$$\frac{dV(y,z)}{dt} = -\sum_{i,j} \Delta y_i \mathcal{A}_{ij} \Delta y_j - \sum_{i,k} \Delta y_i \mathcal{B}_{ik} z_k - \sum_{l,k} z_l \mathcal{D}_{lk} z_k + \sum_l z_l \left(1 - \sum_j \mathcal{C}_{lj} y_j \right) \quad .$$

Now we add and subtract $\sum_l z_l \sum_j \mathcal{C}_{lj} {y_j}^\star$, thereby obtaining:

$$\begin{aligned} \frac{dV(y,z)}{dt} = & -\sum_{i,j} \Delta y_i \mathcal{A}_{ij} \Delta y_j - \sum_{i,k} \Delta y_i \mathcal{B}_{ik} z_k - \sum_{l,j} z_l \mathcal{C}_{lj} \Delta y_j \\ & - \sum_{l,k} z_l \mathcal{D}_{lk} z_k + \sum_l z_l \left(1 - \sum_j \mathcal{C}_{lj} {y_j}^\star \right) \quad . \end{aligned}$$

Thus, $dV(y,z)/dt$ is the sum of two parts. The first part can be rewritten as a quadratic form:

$$\begin{aligned} -\sum_{i,j} \Delta y_i \mathcal{A}_{ij} \Delta y_j - \sum_{i,k} \Delta y_i \mathcal{B}_{ik} z_k - \sum_{l,j} z_l \mathcal{C}_{lj} \Delta y_j - \sum_{l,k} z_l \mathcal{D}_{lk} z_k = -\sum_{i,j} (x_i - \bar{x}_i) A_{ij} (x_j - \bar{x}_j) = \\ - (x - \bar{x})^T A (x - \bar{x}) = -\frac{1}{2} (x - \bar{x})^T \left(A + A^T \right) (x - \bar{x}) \quad , \end{aligned}$$

and is thus negative whenever $A + A^T$ is positive definite.

Finally, $(1 - \mathcal{C} y^\star)_j$ is positive only when population j can grow (and thus invade the system) when the system is resting at the saturated equilibrium. Hence, by the definition of a saturated equilibrium, this term is also negative, and the proof is complete.

Note that the real parts of the eigenvalues of A are bounded from above and below by those of $H = \frac{1}{2}(A + A^T)$, as stated by Bendixson's inequality; in particular, if we order the eigenvalues of H in ascending order, $\lambda_1(H) \leq \lambda_2(H) \leq \ldots \leq \lambda_n(H)$, we have

$$\lambda_1(H) \leq \mathcal{R}e\lambda_i(A) \leq \lambda_n(H) \quad ,$$

with equality guaranteed for symmetric matrices A (in which case we have $H = A$). Thus, we have that $\alpha_L \leq \alpha_S$ (as expected, given that a globally stable equilibrium must be locally stable as well).

In summary, we have shown that, provided that $r > \mathbb{0}$ and that we chose $\alpha > \alpha_S$ such that the symmetric part of A, H, is positive definite, dynamics will always converge to an (unspecified) saturated equilibrium. This is a *sufficient* condition for dynamics to converge to a saturated equilibrium, which can be made slightly more general. For example, whenever one can choose a positive vector $h > \mathbb{0}$ such that $D(h)A + A^T D(h)$ is positive definite, then the stability of dynamics is also ensured[45]. Because we are interested in an upper bound that is easy to calculate, we just analyze the case in which H is positive definite.

1.4 Assembly

This section discusses the assembly of systems for which $\alpha > \alpha_S$, i.e., whether the systems can be built from the bottom up through successive invasions. We can devise an "assembly graph" in which the nodes represent the subset of populations that can coexist, and the edges invasions moving the system from one state to another. Naturally, this graph representation is a good choice for systems in which the dynamics of any feasible sub-community settle at an equilibrium, invasions are rare, and spaced apart in time[20].

When $\alpha > \alpha_S$, then this relationship is maintained for any sub-system (because all the sub-matrices of H obtained by removing some rows and their corresponding columns are also positive definite). Hence, any feasible sub-community would also be (globally) stable, and might or might not be invasible by other populations. Take two sub-communities x_k, composed of k populations and x_{k+1} composed of the same populations, but with another population added. If both x_k and x_{k+1} are feasible, and both are stable because $\alpha > \alpha_S$, then the population that is added can attach itself to the resident community without causing extinctions.

Suppose that a GLV system has $\alpha > \alpha_S$ and a feasible equilibrium $x^\star > \mathbb{0}$. Then, determining whether there is a way to build the community from the ground up amounts to determining whether there is a path in the graph such that, at each step, a different population attaches itself to the community without causing extinctions.

2 General systems

We start our discussion of the effect of the strength of intraspecific competition α on coexistence by considering the case in which matrix B is a generic matrix. In the next section, we analyze the case in which B is a large random matrix instead.

2.1 Effect of average interaction strength

Here we show that the feasibility of the equilibrium $x^\star$ in Eq. (1) is related to that of the equilibrium of a simpler system, $y^\star$, thereby obtaining a canonical form for the study of feasibility in competitive GLV systems.

The system in Eq. (1) is feasible whenever $x^\star = A^{-1}\mathbb{1} = (\alpha I + B)^{-1}\mathbb{1} > \mathbb{0}$. We can write:

$$A = \alpha I + B = \alpha I + C + \mathbb{1}m^T \quad ,$$

where the columns of C sum to zero ($C^T\mathbb{1} = \mathbb{0}$), the vector $m \geq \mathbb{0}$ (because the system is competitive), I is the identity matrix of dimension n, and $\alpha \geq 0$. The rank-one matrix $\mathbb{1}m^T$ has constant columns with nonnegative coefficients, corresponding to the column means of B. Now consider the system:

$$(\alpha I + C)y^\star = \mathbb{1} \quad . \tag{3}$$

Note that the sum of all the $y_i^\star$ is positive:

$$\begin{aligned}
\mathbb{1}^T(\alpha I + C)y^\star &= \mathbb{1}^T\mathbb{1} \\
\alpha\mathbb{1}^T y^\star + \mathbb{1}^T C y^\star &= n \\
\frac{1}{n}\mathbb{1}^T y^\star + 0 &= \frac{1}{\alpha} > 0 \quad .
\end{aligned}$$

As such, the mean of the components of $y^\star$ is exactly $1/\alpha$, and is thus positive for any positive $\alpha > 0$. We can write $x^\star$, the solution of $Ax^\star = \mathbb{1}$, as a function of $y^\star$ using the Sherman-Morrison formula:

$$\begin{aligned}
(\alpha I + B)x^\star &= \mathbb{1} \\
(\alpha I + C + \mathbb{1}m^T)x^\star &= \mathbb{1} \\
x^\star &= (\alpha I + C + \mathbb{1}m^T)^{-1}\mathbb{1} \\
x^\star &= \left((\alpha I + C)^{-1} - \frac{1}{1 + m^T(\alpha I + C)^{-1}\mathbb{1}}(\alpha I + C)^{-1}\mathbb{1}m^T(\alpha I + C)^{-1}\right)\mathbb{1} \\
x^\star &= y^\star - \frac{1}{1 + m^T y^\star} y^\star m^T y^\star \\
x^\star &= \left(1 - \frac{m^T y^\star}{1 + m^T y^\star}\right) y^\star \\
x^\star &= \frac{1}{1 + m^T y^\star} y^\star \quad .
\end{aligned}$$

Thus, $x^\star$ is proportional to $y^\star$, with constant of proportionality $1/(1 + m^T y^\star)$. As such, we have a feasible equilibrium $x^\star > \mathbb{0}$ whenever a) $y^\star > \mathbb{0}$ (i.e., $y^\star$ is feasible, and because $m^T y^\star = \sum_j m_j y_j^\star$ is the sum of products of nonnegative values, then the constant of proportionality is also positive), or, b) $y^\star < \mathbb{0}$ and $m^T y^\star < -1$ (i.e., all the components of $y^\star$ are negative, and sufficiently large to make the constant of proportionality negative as well). But we have just seen that the average of $y^\star$ is positive, thereby making case b) impossible. We conclude that $x^\star$ is feasible if and only if $y^\star$ is feasible.

Because infinitely many competitive systems in Eq. (1) map into the same Eq. (3), we say that the equation in C is the *canonical* form for studying the feasibility of the competitive systems that can be formed by adding positive columns to C.

2.2 A sufficiently large α ensures feasibility

Here we use the well-known Perron-Frobenius property for nonnegative matrices to show that for sufficiently large $\alpha = \alpha_\infty$ the feasibility of $y^\star$ (and thus of $x^\star$) is guaranteed, and is maintained for any $\alpha > \alpha_\infty$. To this end, we divide both sides of Eq. (3) by $\alpha > 0$:

$$\left(I + \frac{1}{\alpha}C\right) y^\star = \frac{1}{\alpha}\mathbb{1} \quad .$$

We have shown above that $\mathbb{1}^T y^\star = n/\alpha$, but then $\frac{\alpha}{n}\mathbb{1}^T y^\star = 1$. Multiplying the right hand side by this value, we get:

$$\left(I + \frac{1}{\alpha}C\right) y^\star = \frac{1}{n}\mathbb{1}\mathbb{1}^T y^\star \quad .$$

Bringing the right-hand side to the left:

$$\left(I + \frac{1}{\alpha}C - \frac{1}{n}\mathbb{1}\mathbb{1}^T\right) y^\star = \mathbb{0}$$

and, finally,

$$\begin{aligned}\left(\frac{1}{\alpha}C - \frac{1}{n}\mathbb{1}\mathbb{1}^T\right) y^\star &= -y^\star \\ \left(\frac{1}{n}\mathbb{1}\mathbb{1}^T - \frac{1}{\alpha}C\right) y^\star &= y^\star \quad ,\end{aligned}$$

which shows that $y^\star$ is a right eigenvector of $\frac{1}{n}\mathbb{1}\mathbb{1}^T - \frac{1}{\alpha}C$, with associated eigenvalue $\lambda = 1$. Note that the matrix $\frac{1}{n}\mathbb{1}\mathbb{1}^T - \frac{1}{\alpha}C$ has left eigenvector $\mathbb{1}$, also associated with $\lambda = 1$. Now suppose we choose $\alpha = \alpha_\infty > n \max_{i,j} C_{ij}$; then, we have that a) the matrix $\frac{1}{n}\mathbb{1}\mathbb{1}^T - \frac{1}{\alpha_\infty}C$ has only positive entries; b) it is column stochastic (i.e., it is nonnegative with columns summing to 1), and therefore has spectral radius 1. But then $\frac{1}{n}\mathbb{1}\mathbb{1}^T - \frac{1}{\alpha_\infty}C$ has the Perron-Frobenius property, and thus $y^\star$ is positive, and the positivity of $y^\star$ is guaranteed for any $\alpha \geq \alpha_\infty$. Naturally, this choice of $\alpha = \alpha_\infty$ is overly conservative, but it is sufficient to show that there must exists some critical value $\alpha_F \leq \alpha_\infty$ such that: a) $y^\star$ is positive, and b) it remains positive for any $\alpha > \alpha_F$. Computing such a value is the main goal of the next two subsections.

2.3 Transitions to and from feasibility

We can track the value of $y_i^\star(\alpha)$, as we vary α, using Cramer's rule:

$$y_i^\star(\alpha) = \frac{\begin{vmatrix} \alpha + C_{1,1} & C_{1,2} & \cdots & 1 & \cdots & C_{1,n} \\ C_{2,1} & \alpha + C_{2,2} & \cdots & 1 & \cdots & C_{2,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ C_{n,1} & C_{n,2} & \cdots & 1 & \cdots & \alpha + C_{n,n} \end{vmatrix}}{|\alpha I + C|} = \frac{|Q_i(\alpha)|}{|Q(\alpha)|} \quad ,$$

where $Q(\alpha) = \alpha I + C$, and $Q_i(\alpha)$ is obtained substituting the ith column of $Q(\alpha)$ with the vector $\mathbb{1}$.

When $|Q(\alpha)| = 0$, the matrix is singular; this means that C has a real, negative eigenvalue λ_i such that by setting $\alpha = -\lambda_i$ we place an eigenvalue of $Q(\alpha)$ exactly at zero. We have two cases: if, as $\alpha \to -\lambda_i$, the determinant of the numerator $|Q_i(\alpha)|$ tends to a nonzero value, then $y_i^\star(\alpha)$ diverges as $\alpha \to -\lambda_i$; if instead $|Q_i(\alpha)| \to 0$ as $\alpha \to -\lambda_i$, then the solution $y_i^\star(\alpha)$ does not diverge around $\alpha = -\lambda_i$. Importantly, given that the mean of the solution is fixed and positive, whenever some solutions diverge, some components will tend to $+\infty$, while others to $-\infty$ around these singularities. Hence, the system cannot be feasible around any singularity that gives rise to divergences. We denote the largest value of α leading to a divergence as α_0. If no value of α leads to a divergence, then $\alpha_0 = 0$.

Now take α_L to be the minimum value of α such that $\min_j \mathcal{R}e\, \lambda_j(\alpha_L I + C) = 0$ (as seen in Sec. 1.1, this critical value is important for local stability); for any value $\alpha > \alpha_L$ the determinant $|Q(\alpha)|$ is positive, and thus the existence of a solution is guaranteed. This can be easily proven by considering that the determinant of a matrix is the product of its eigenvalues; then the eigenvalues of $Q(\alpha)$ for any $\alpha > \alpha_L$ are either real and positive, or complex and multiplied by their conjugate ($\lambda_j \lambda_j^* > 0$), yielding therefore a positive product, and thus determinant. Therefore, any divergence can only appear when $\alpha \leq \alpha_L$, and thus $\alpha_0 \leq \alpha_L$.

Having considered the denominator, we turn to the numerator $|Q_i(\alpha)|$: whenever this determinant is zero, $y_i^\star(\alpha) = 0$, and thus this component of the solution transitions from negative to positive, or vice versa (because $|Q_i(\alpha)|$ is a polynomial in α and, as such, is a continuous function everywhere). We can characterize this determinant by performing some simplifications. The derivation is easier to write when the focal population $i = n$ is in last position—this can be accomplished for any i by reordering the rows and columns of the matrix. We have (for $i = n$):

$$|Q_i(\alpha)| = \begin{vmatrix} \alpha + C_{1,1} & C_{1,2} & \cdots & 1 \\ C_{2,1} & \alpha + C_{2,2} & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ C_{n,1} & C_{n,2} & \cdots & 1 \end{vmatrix} = \begin{vmatrix} \alpha \widetilde{I} + \widetilde{C} & \mathbb{1} \\ c^T & 1 \end{vmatrix} \quad ,$$

where $\widetilde{I}$ and $\widetilde{C}$ are the corresponding matrices with the last row and column removed, and c^T is the last row of C with the last element removed. We can subtract the last row from all other rows without altering the determinant:

$$\begin{vmatrix} \alpha\widetilde{I} + \widetilde{C} & \mathbb{1} \\ c^T & 1 \end{vmatrix} = \begin{vmatrix} \alpha\widetilde{I} + \widetilde{C} - \mathbb{1}c^T & \mathbb{0} \\ c^T & 1 \end{vmatrix} \quad ,$$

which is a block-structured matrix, with determinant $|\alpha\widetilde{I} + \widetilde{C} - \mathbb{1}c^T|$. Finally, we note that $\tilde{C} - \mathbb{1}c^T = \tilde{B} - \mathbb{1}\tilde{m}^T - (\mathbb{1}b^T - \mathbb{1}\tilde{m}^T) = \tilde{B} - \mathbb{1}b^T$, and thus we can work either with matrix C or B (naturally, if $y_i^\star(\alpha) = 0$, necessarily $x_i^\star(\alpha) = 0$).

This means that $y_i^\star(\alpha) = 0$ whenever $\alpha \geq 0$ matches precisely a real, negative eigenvalue of $\tilde{B} - \mathbb{1}b^T$. Consider the eigenvalues of $\tilde{B} - \mathbb{1}b^T$, and call λ_i the most negative, real eigenvalue (if it exists); then, if $-\lambda_i > \alpha_0$, at $\alpha_i = -\lambda_i$ necessarily $y_i^\star(\alpha)$ crosses from negative to positive as α increases (because we have established that when $\alpha \geq \alpha_F$, $y_i^\star(\alpha)$ must be positive).

Now repeat the procedure for all i, and collect all the $\alpha_i > \alpha_0$ representing the last transition to positive values for each i (with $\alpha_i = 0$ if such a transition does not exist, or it appears before α_0); then necessarily for any $\alpha > \alpha_F = \max\{\alpha_0, \alpha_1, \ldots, \alpha_n\}$ the solution $y^\star(\alpha)$ is positive, and further increasing α maintains positivity.

To summarise these results, we analyze a simple example, showing the location of α_0 and α_i ($i = 1, \ldots, n$). Take the matrix C:

$$C = \begin{pmatrix} -4 & 2 & 3 \\ -4 & -4 & 2 \\ 8 & 2 & -5 \end{pmatrix},$$

with eigenvalues $\lambda(C) = \{-9, -4, 0\}$. Then, $Q(\alpha)$ is singular when $\alpha \in \{4, 9\}$; at these points the three possible $Q_i(\alpha)$ are not singular, and therefore the solution $y^\star(\alpha)$ diverges around these points; we thus set $\alpha_0 = 9$.

The eigenvalues of $\widetilde{C}^{(1)} - \mathbb{1}(c^{(1)})^T$:

$$\widetilde{C}^{(1)} - \mathbb{1}(c^{(1)})^T = \begin{pmatrix} -4 & 2 \\ 2 & -5 \end{pmatrix} - \begin{pmatrix} 2 & 3 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} -6 & -1 \\ 0 & -8 \end{pmatrix}$$

are -6 and -8 and thus $y_1^\star(\alpha) = 0$ when we choose $\alpha = 6$ or $\alpha = 8$. Both transitions appear before the last divergence at $\alpha_0 = 9$, and hence the equilibrium of population 1 will be positive after the last divergence (and we set $\alpha_1 = 0$).

Repeating for population two, we have:

$$\widetilde{C}^{(2)} - \mathbb{1}(c^{(2)})^T = \begin{pmatrix} -4 & 3 \\ 8 & -5 \end{pmatrix} - \begin{pmatrix} -4 & 2 \\ -4 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 12 & -7 \end{pmatrix} \quad ,$$

with eigenvalues $-\frac{1}{2}(7 \pm \sqrt{97})$; thus the equilibrium for population 2 will transition only at $\alpha = \frac{1}{2}(7 + \sqrt{97})$, which is smaller than 9; hence the solution for the second population will be positive after the last divergence (and we set $\alpha_2 = 0$ as well).

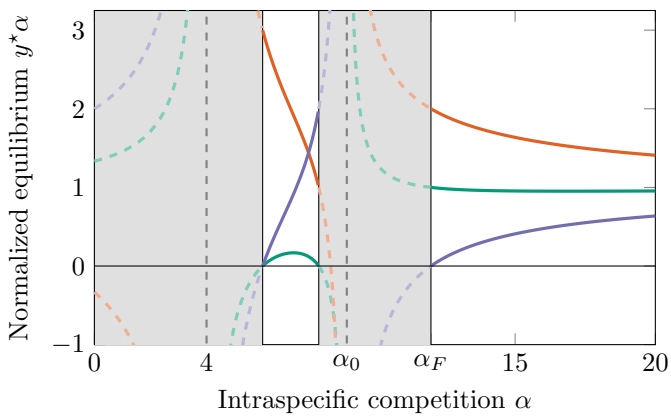

Fig. 6 The components of $y^\star(\alpha)$ as a function of $\alpha \geq 0$. Grey areas denote the values of α for which the equilibrium is not feasible, and the white regions correspond to values of α for which all components $y_i^\star(\alpha)$ (colors) are positive. As explained in the text, for two particular values of $\alpha = 4$ and $\alpha = 9$ (vertical dashed lines), the components of the solution diverge to $\pm\infty$, and thus feasibility is precluded around these points. We denote the largest value of α that leads to a divergence as α_0. After the last divergence, the component associated with population 3 (purple) is still negative, and transitions to a positive value at exactly $\alpha_3 = \alpha_F = 12$. The solution remains positive for any $\alpha > \alpha_F$.

Finally, considering population 3, we have:

$$\widetilde{C}^{(3)} - \mathbb{1}(c^{(3)})^T = \begin{pmatrix} -4 & 2 \\ -4 & -4 \end{pmatrix} - \begin{pmatrix} 8 & 2 \\ 8 & 2 \end{pmatrix} = \begin{pmatrix} -12 & 0 \\ -12 & -6 \end{pmatrix} \quad ,$$

yielding a transition at $\alpha = 6$ and $\alpha = \alpha_3 = 12$.

Therefore, $\alpha_F = \max\{9, 0, 0, 12\} = 12$. Thus, for this system, all components of $y^\star$ will be positive for any $\alpha > 12$. This behavior is illustrated in Figure 6.

2.4 Finding α_F via eigendecomposition

The results in the previous section are sufficient to devise an exact algorithm to determine α_F for a given system specified by the matrix C (or B). It consists of three steps:

1. Compute the eigenvalues of $-C$, and store the largest positive eigenvalue leading to a divergence in α_0; if none exist, $\alpha_0 = 0$.
2. For each i in $1, \ldots, n$:
   (a) Form the matrix $\widetilde{C}^{(i)}$ (by removing the ith row and column from C) and the corresponding $(c^{(i)})^T$ (i.e., the ith row of C with the ith coefficient removed);
   (b) Compute the eigenvalues of the matrix $\mathbb{1}(c^{(i)})^T - \widetilde{C}^{(i)}$;
   (c) Store the largest positive, real eigenvalue in α_i; if no eigenvalues are real and positive, or if all are smaller than α_0, then $\alpha_i = 0$.
3. $\alpha_F = \max\{\alpha_0, \alpha_1, \ldots, \alpha_n\}$.

From a computational standpoint, we note that the eigenvalue with the largest real part can be computed efficiently, for example via the Arnoldi iteration[46]. One could also utilize well-known bounds on eigenvalues to compute α_i only for the matrices that are more likely to produce a maximum.

For the example above, we find $\alpha_0 = 9$, $\alpha_1 = \alpha_2 = 0$, and $\alpha_3 = 12 = \alpha_F$.

3 Random systems

Having derived some general results in the previous section, we now turn to the case in which the matrix of interactions B is a large random matrix. This case has been studied for decades, producing a wealth of results. Here we show that previous results can be extended, generalized and refined in light of the derivations above. This section has two main goals: a) compute the probability that the equilibrium $x^\star$ of Eq. (1) is positive, given a value of α and the statistics of the random matrix B; b) show how these derivations have profound implications for the type of dynamics that can be observed for these systems.

3.1 Random systems with statistically equivalent populations

We start by defining what we mean by random systems. Naturally, dynamical systems in which some (all) of the parameters are sampled from random distributions are indeed random systems; here we focus on a subclass in which the model definition is closed under relabeling of the populations: by permuting the indices x_i of Eq. (1) we obtain the same random model. In practice, this simply means that there is no special "structure" making any of the populations behave differently from the others—the populations are *statistically equivalent*.

Among this subclass of models, the simplest to study is one in which the matrix of interactions is given by the sum of a random matrix B and a positive constant μ.

$$B' = B + \mu \mathbb{1}\mathbb{1}^T \quad ,$$

where we choose B such that: $B_{ii} = 0$, and $\mathbb{E}(B_{ij}) = 0$, $\mathbb{E}(B_{ij}^2) = 1$, $\mathbb{E}(B_{ij}B_{ji}) = \rho$ for every $i \neq j$. Thus the spectrum of B follows the elliptic law[47]: the eigenvalues of $B/\sqrt{n}$ are approximately uniformly distributed in an ellipse in the complex plane, with center $(0, 0i)$, horizontal semi-axis $1 + \rho$, and vertical semi-axis $1 - \rho$.

As we have shown in Section 2.1, as long as the columns of B' sum to positive values (i.e., when the systems are competitive), the value of μ has no effect on feasibility. The value of μ also has no effect on the stability of the system[8]: the eigenvalues of B' will be distributed in an ellipse, possibly with an outlier eigenvalue on the right of the bulk, and thus not influencing stability. Hence, α_F (the value at which the equilibrium is feasible, and feasibility is maintained when we increase α), α_L (the value at which all eigenvalues have positive real part), and α_S (the value at which the symmetric part of the matrix is positive definite) are the same for both B and B'.

Thus, without loss of generality, we can consider the case in which $\mu = 0$, and therefore the matrix of interactions follows the elliptic law. This also allows us to compare our results directly with other studies that have considered random matrices with coefficients sampled from distributions with mean zero[11, 13, 14].

3.2 Previous results

The oldest result in this area goes back to a response to May's article, by Alan Roberts[10]. Roberts used computer simulations to measure the probability that a feasible equilibrium would be locally stable for a GLV model, Eq. (1), in which populations were self–regulating ($A_{ii} = 1$), and interactions were randomly

chosen to be $A_{ij} = \theta$ with probability 1/2 and $A_{ij} = -\theta$ otherwise. He found that the probability of stability conditional on feasibility was very large, and increasing with the size of the system.

The next main contribution was obtained by Lewi Stone as part of his PhD thesis[11], developed under the supervision of Roberts. Using techniques similar to those used here, Stone was able to approximate the probability of feasibility for matrices with random interactions sampled independently from an arbitrary distribution with mean zero. For example, a formula reported in Stone[12], and adapted to the notation employed in this work, reads:

$$p_F(\alpha) = \Phi\left(\frac{\alpha}{\sqrt{n-1}(1+\frac{n-1}{\alpha^2})^{\frac{1}{4}}}\right)^n \quad ,$$

where $p_F(\alpha) = P(x^\star > \mathbb{0}|\alpha)$, and $\Phi(\cdot)$ is the cumulative distribution function for the standard normal distribution. This formula was derived in[11] by approximating the solution of the linear system using the Neumann series (see Sec. 3.7). Stone's thesis also reported larger computer simulations, in the spirit of Roberts, again confirming that feasible systems are almost invariably locally stable.

More recently, two articles[13, 14] used different techniques to show that there exist a critical value below which the probability of feasibility is negligible, and above which it is very high. In particular, using the notation adopted in this work, they found that the probability of feasibility is close to one for $\alpha \gg \sqrt{n}\sqrt{2\log n}$, and is close to zero for $\alpha \ll \sqrt{n}\sqrt{2\log n}$. Clenet *et al.*[14] considered matrices with correlated entries as well, finding that the threshold value for feasibility would not change substantially, and, importantly, noting that the transition would appear after that for global stability in these GLV systems. Also important for the work developed here, Liu *et al.*[15] showed, using yet a different technique, that correlation between the entries would have an effect on feasibility (increasing the probability of feasibility when the correlation is negative), but that the effect would be negligible for large systems.

Finally, Marcus *et al.*[16] considered stable, feasible GLV systems and increased the interaction strengths (equivalently, one could have decreased the strength of intraspecific interactions, as done here) until either the system becomes unfeasible, or unstable. They found that in general, the system loses feasibility before stability, consistently with the previous findings outlined above. Interestingly for our argument, they notice that in symmetric systems, feasibility would be often lost due to the components of the solution diverging to $\pm\infty$, a fact that will be key for the developments below. A similar argument was put forward by Stone[12].

Figure 7 shows simulations measuring the probability of feasibility when varying both α (x-axis) and the correlation between entries (colors); the approximation by Stone[11] for $\rho = 0$, as well as the critical value by Bizeul & Najim[13] are also reported.

The goal of the next few sections is to generalize and improve these results in the light of what we have derived for general matrices. In particular, we will extend both the formula by Stone and the critical value of Bizeul & Najim for the case in which entries are correlated.

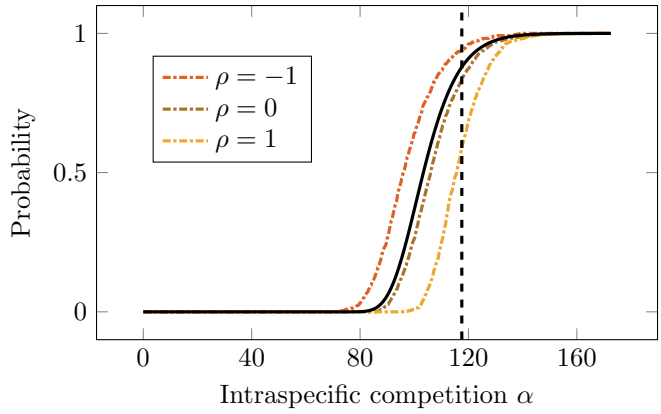

Fig. 7 Probability of feasibility (i.e., $P(x^\star > \mathbb{0})$, y-axis) for a given level of intraspecific interaction α (x-axis). The probability was computed by considering 1000 realizations of systems of size $n = 1000$ with interaction matrices B following the elliptic law (with the colors representing the off-diagonal correlation). The black line represents the approximation by Stone[12] for the case $\rho = 0$. The vertical dashed line is the critical value of Bizeul & Najim[13].

3.3 Feasibility when $\alpha = 0$

For large random systems, the probability of feasibility when $\alpha = 0$ is negligible. We can compute this probability exactly whenever the coefficients of B are sampled independently from a distribution that is symmetric about 0.

We want to compute the probability that $(\alpha I + C)y^\star = \mathbb{1}$ has a positive solution when $\alpha = 0$. Since the columns of C sum to 0 (see Section 2.1), the right-hand side of the equation vanishes when we multiply the equations by the transpose of C:

$$\begin{aligned} Cy^\star &= \mathbb{1} \\ C^T C y^\star &= C^T \mathbb{1} \\ C^T C y^\star &= \mathbb{0} \quad . \end{aligned}$$

The matrix $C^T C$ is symmetric, positive semi-definite, and has an eigenvalue 0; all its eigenvectors are orthogonal. Then, whenever a feasible solution exists, solving the system of equations above is equivalent to solving $Cy^\star = \mathbb{0}$, i.e., $y^\star$ is a right eigenvector of C associated with the zero eigenvalue. Thus, $y^\star$ is defined up to a constant of proportionality, allowing us to set one of the components arbitrarily. Here we set $y^\star_n = 1$, and we partition C and $y^\star$ as:

$$C = \begin{pmatrix} \widetilde{C} & a \\ c^T & b \end{pmatrix} \quad y^\star = \begin{pmatrix} \widetilde{y} \\ 1 \end{pmatrix} \quad ,$$

where $b = -\sum_{j=1}^{n-1} a_j$. We can therefore write $n-1$ independent equations:

$$\begin{aligned} \widetilde{C}\widetilde{y} &= -a \\ \widetilde{y} &= -\widetilde{C}^{-1} a \quad . \end{aligned}$$

For feasibility, we need all $\widetilde{y} > \mathbb{0}$ (because we have chosen a positive $y^\star_n$). To compute the probability that $\widetilde{y} > \mathbb{0}$ exactly, we exploit the assumptions that we have made: because of symmetry, the probability of

observing a given C_{ij} (or the density of the distribution at C_{ij}, when the distribution is continuous), satisfies the equality $P(C_{ij}) = P(-C_{ij})$.

Then, we can follow the approach of Serván *et al.*[33], and define the diagonal matrices $K_i = D(s_i)$, where the s_i are the 2^n distinct vectors of length n containing either 1 or -1: $s_1 = (1, 1, 1, \ldots, 1)^T$, $s_2 = (1, -1, 1, \ldots, 1)^T$, $s_3 = (-1, -1, 1, \ldots, 1)^T$, $s_4 = (1, 1, -1, \ldots, 1)^T$, $\ldots$, $s_{2^n} = (-1, -1, -1, \ldots, -1)^T$. We partition these matrices as well:

$$K_i = \begin{pmatrix} \widetilde{K_i} & \mathbb{0} \\ \mathbb{0}^T & \sigma_i \end{pmatrix} \quad .$$

Because of symmetry and independence, we have $P(C) = P(CK_i)$, given that multiplying C by $K_i = D(s_i)$ simply changes the signs of all the coefficients in the columns corresponding to negative signs in s_i. Finally, we write:

$$\begin{aligned} \widetilde{C}\widetilde{K}_1\widetilde{y} &= -\sigma_1 a \\ \widetilde{y} &= -\sigma_1 \widetilde{K}_1 \widetilde{C}^{-1} a \quad , \end{aligned}$$

where we have used the fact $K_1 = I$. Now notice that by exchanging K_1 with K_i, we can change the signs of the components of $\widetilde{y}$; then, for each possible pattern of signs of $-\widetilde{C}^{-1}a$ there are exactly two choices of K_i that make the vector $\widetilde{y}$ all positive (one solution with $K_j = D(s_j)$, and the other with $K_l = D(-s_j)$). Given that $P(CK_i) = P(CK_j)$, and that there are 2^n matrices K_i, the probability of feasibility is exactly $2/2^n = 1/2^{n-1}$.

Thus, for a random system with B_{ij} sampled independently from a distribution that is symmetric about its mean, for two populations we have fifty percent probability of obtaining a feasible equilibrium, and the probability rapidly decreases, so that we we have less than one in a thousand chance to find a feasible equilibrium for $n = 11$ and less than one in a million for $n = 21$.

3.4 Feasibility when $\alpha < \alpha_L$

We have shown above that the probability of feasibility is vanishingly small for $\alpha = 0$; here we show that the same holds true for any $\alpha < \alpha_L$. In this case, the support for the bulk of the spectrum of $\alpha I + B$ contains the point $(0, 0i)$, because the ellipse is not completely contained in the right half of the complex plane. Because the matrix is large, and eigenvalues are approximately uniformly distributed inside the ellipse, we can always find an eigenvalue that is arbitrarily close to the origin, thereby causing some solutions to reach arbitrarily large absolute values—given that the determinant of matrix $Q(\alpha) = \alpha I + B$ (defined in Sec. 2.3) must be arbitrarily close to zero. Because the mean of $y^\star$ is positive, components reaching large positive values must be balanced by some other components with negative values, and thus feasibility is precluded. Hence, the probability of feasibility at $\alpha < \alpha_L$ is close to zero. A similar argument is found (in the context of symmetric matrices), in Marcus *et al.*[16]. Therefore, we can assume that $\alpha_0 = \alpha_L$.

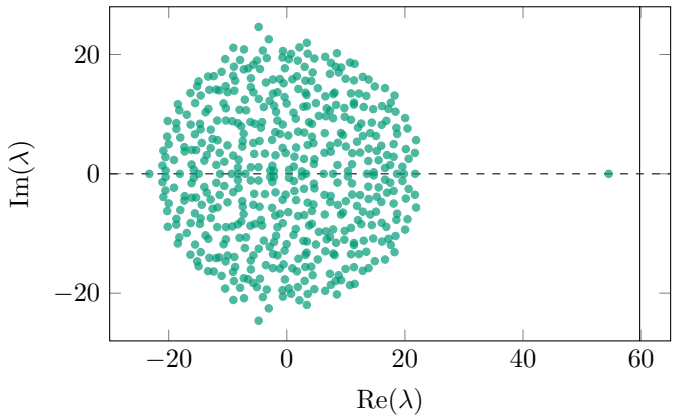

Fig. 8 Spectrum of $\widetilde{Q}_i = \mathbb{1}(b^{(i)})^T - \widetilde{B}^{(i)}$ for a given i, when B is a large random matrix following the circular law. When a large random matrix ($-\widetilde{B}^{(i)}$ in this case) is perturbed by summing a matrix with small rank (in this case, matrix $\mathbb{1}(b^{(i)})^T$ has rank 1), only a few of the original eigenvalues change substantially (one in the case of a rank-one matrix). Moreover, the eigenvalues are expected to be close to those of the small-rank matrix (vertical line).

3.5 A single transition to feasibility when $\alpha > \alpha_L$

Having shown that, with high probability, the system is not feasible for $\alpha < \alpha_L$, we consider the dependence of each component of $y^\star((\alpha)$ on α, as it increases for $\alpha > \alpha_L$. We have seen in Section 2.3 that $y_i^\star(\alpha)$ transitions (either from positive to negative, or vice versa) whenever $\alpha = \lambda_i$, and λ_i is a real, positive eigenvalue of the matrix

$$\widetilde{Q}_i = \mathbb{1}(b^{(i)})^T - \widetilde{B}^{(i)} \quad ,$$

where $\widetilde{B}^{(i)}$ is the matrix B with the row and column corresponding to i removed, and $(b^{(i)})^T$ is the ith row of B with coefficient i removed. If $\widetilde{B}^{(i)}$ is a large random matrix (i.e., if B is a large random matrix following the elliptic law), then perturbing it by adding a rank-one matrix has a simple effect on its spectrum: at most one eigenvalue moves outside of the bulk, while the remaining eigenvalues are fundamentally unchanged[29] (an example is provided in Fig. 8). Thus, if the outlier eigenvalue is on the left of the bulk, then λ_i is negative and $y_i^\star(\alpha)$ cannot cross zero when choosing $\alpha > 0$. If, on the other hand, the outlier eigenvalue $\lambda_i > \alpha_L$ is positive, we can make $y_i^\star(\alpha) = 0$ by choosing $\alpha = \lambda_i$. Then, this transition must bring $y_i^\star(\alpha)$ from negative to positive values: because we know that, for sufficiently large values of α, $y_i^\star(\alpha) > 0$, and that there are no divergences after α_L, then necessarily $y_i^\star(\alpha)$ becomes zero at this value, and remains positive when α is increased further (observe that $y_i^\star(\alpha)$ is a rational function —a quotient of two polynomials— in α, which is a continuous function above the largest point at which it diverge, i.e., for $\alpha > \alpha_0$; therefore, the change of sign follows by continuity of $y_i^\star(\alpha)$ for $\alpha > \alpha_L$). Hence, we can set $\alpha_i = \lambda_i$.

We have just shown that each matrix $Q_i(\alpha)$ has at most a single real eigenvalue to the right of the bulk, with value $\lambda_i = \alpha_i$. Moreover, when this positive outlier eigenvalue exists, component i transitions from negative to positive at $\alpha = \alpha_i$. But then we can order the α_i from smallest to largest, and at $\alpha_F = \max_i \alpha_i$, all components have crossed into positive values. Because for large random matrices feasibility is precluded (with very high probability) when $\alpha < \alpha_L$, we conclude that a) positivity is gained only at α_F, and b) it is never lost for larger values of α. Thus, large random matrices produce a much simpler picture of the dependence on

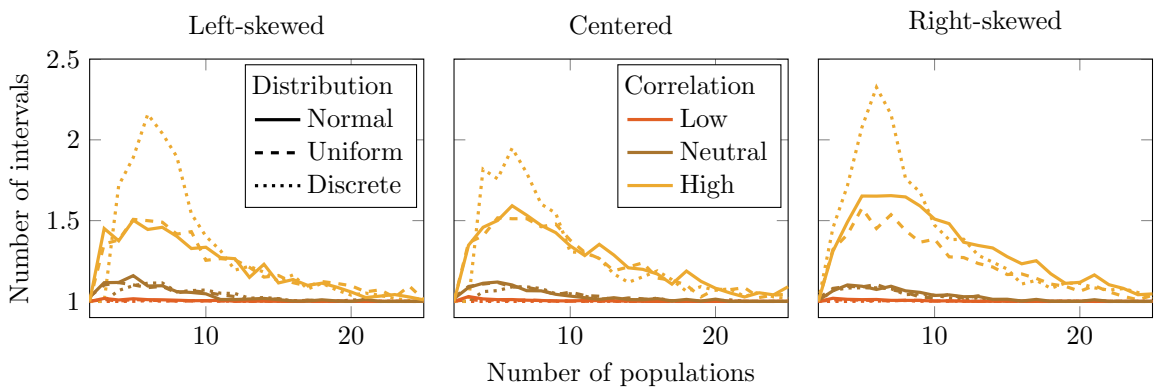

Fig. 9 For random systems of increasing size (n, x-axis), we count the number of disjoint intervals for α yield a feasile equilibrium. As we increase n, the number of intervals converges to one, irrespective of correlation ρ or choice of distribution. Then, large random matrices will yield a single transition, occurring at α_F.

α than small, or arbitrary matrices, in which, when we grow α, the equilibrium can transition from unfeasible to feasible several times (Sec. 2.3). These findings are confirmed by simulations: when we count the number of transitions at which the solution gains or loses feasibility for random systems of increasing size, we see a rapid convergence to a single transition (Fig. 9).

3.6 Approximating the probability of feasibility using the theory of small-rank perturbations of random matrices

We can use the argument above on the perturbation of random matrices with small-rank matrices to provide an estimate of the probability of feasibility. We write $P(x^\star > \mathbb{0}|\alpha) = p_F(\alpha)$. In light of the previous paragraph, this amounts to $P(\lambda_i < \alpha)$ for all n choices of i, where λ_i is either the largest real eigenvalue of $\mathbb{1}(b^{(i)})^T - \widetilde{B}^{(i)}$, when it exists, or $\lambda_i = 0$, when either the matrix has no positive eigenvalues, or all are smaller than α_L.

To approximate the location of λ_i, we follow O'Rourke and Renfrew[29], who have proven that, when A is a large random matrix following the elliptic law, uv^T is a rank-one matrix with real eigenvalue γ, and $|\gamma| > 1 + \rho$, then the spectrum of $A/\sqrt{n} + uv^T$ has an outlier, real eigenvalue, λ_o, with:

$$\lambda_o = \gamma + \frac{\rho}{\gamma} + o(1) \quad .$$

Let λ_i be the outlier of matrix $\mathbb{1}(b^{(i)})^T - \widetilde{B}^{(i)}$. We want to find its location. Define $\gamma_i = \sum_j b_j^{(i)}$, i.e., the only nonzero eigenvalue of $\mathbb{1}(b^{(i)})^T$. Then, the outlier of $\frac{1}{\sqrt{n-1}}(\mathbb{1}(b^{(i)})^T - \widetilde{B}^{(i)})$ is located approximately at

$$\frac{\lambda_i}{\sqrt{n-1}} = \frac{\gamma_i}{\sqrt{n-1}} + \frac{\rho\sqrt{n-1}}{\gamma_i} \quad , \tag{4}$$

if $\left|\frac{\gamma_i}{\sqrt{n-1}}\right| > 1 + \rho$ holds. Otherwise, the outlier does not exists because it is contained in the bulk. For $\gamma_i > \sqrt{n-1}$, the right-hand side of Eq. (4) is maximized when γ_i is maximized, for any value of ρ; hence, the maximum $\lambda_i = \hat{\lambda}$ tends to be associated with the largest $\gamma_i = \hat{\gamma}$, which in turn is simply the maximum row sum of B.

Now we assume that the γ_is are independent and normally distributed, with mean zero and variance $n-1$; an assumption that is correct when $\rho = 0$ and n is sufficiently large, and is met only approximately for nonzero values of ρ (though the effect of the correlation on row sums must be negligible for large n). Thus, the variables $\tilde{\gamma}_i = \gamma_i/\sqrt{n-1}$ follow the standard normal distribution. The distribution of the maximum $\tilde{\gamma}$ therefore converges asymptotically (albeit quite slowly[48]) to the Gumbel distribution[13]. In particular, asymptotically we have:

$$P(\tilde{\gamma} < x) = \exp\left(-\exp\left(-((x - b_n)b_n)\right)\right) \quad ,$$

where b_n is a constant that can be chosen to guarantee a faster convergence to the Gumbel distribution. Here we follow Hall[48], and use $b_n = \sqrt{\mathcal{W}_0\left(\frac{n^2}{2\pi}\right)}$, where $\mathcal{W}_0(z)$ is the principal branch of the Lambert function. One can also Taylor-expand the function for large n, obtaining:

$$\mathcal{W}_0\left(\frac{n^2}{2\pi}\right) \approx \log\frac{n^2}{2\pi} - \log\log\frac{n^2}{2\pi} + \frac{\log\log\frac{n^2}{2\pi}}{\log\frac{n^2}{2\pi}} + \cdots \quad .$$

We want to derive an approximate distribution for $\alpha_F = \sqrt{n-1}\,\tilde{\lambda}$, where, according to the small-rank perturbations, we have

$$\tilde{\lambda} \approx \tilde{\gamma} + \frac{\rho}{\tilde{\gamma}} \quad .$$

We now solve this equation for $\tilde{\gamma}$, obtaining:

$$\tilde{\gamma} \approx \frac{\tilde{\lambda} + \sqrt{\tilde{\lambda}^2 - 4\rho}}{2} \quad .$$

Substituting this value in the definition of the Gumbel distribution above, we have:

$$P(\tilde{\lambda} < x) = \exp\left(-\exp\left(-\left(\frac{x + \sqrt{x^2 - 4\rho}}{2} - b_n\right)b_n\right)\right) \quad .$$

Concluding, the probability that the system is feasible at α is the probability that $\alpha_F < \alpha$, and thus $\alpha_F/\sqrt{n-1} = \tilde{\lambda} < \alpha/\sqrt{n-1}$:

$$p_F(\alpha) = P\left(\frac{\alpha_F}{\sqrt{n-1}} < \frac{\alpha}{\sqrt{n-1}}\right) = P(\tilde{\lambda} < \tilde{\alpha}) = \exp\left(-\exp\left(-\left(\frac{\tilde{\alpha} + \sqrt{\tilde{\alpha}^2 - 4\rho}}{2} - b_n\right)b_n\right)\right) \quad , \tag{5}$$

where $\tilde{\alpha} = \alpha/\sqrt{n-1}$.

The Gumbel cumulative density function $\exp(-\exp(-x))$ describes a random variable with mode 0, median $-\log\log 2 \approx 0.366$ and mean ≈ 0.57721 (the Euler-Mascheroni constant). Thus, we expect the α_F for a given n and ρ to have the following statistics:

Mode of α_F***.***

We solve:

$$\left(\frac{\tilde{\alpha}+\sqrt{\tilde{\alpha}^2-4\rho}}{2}-b_n\right)b_n=0$$

$$\tilde{\alpha}=b_n+\frac{\rho}{b_n}$$

$$\alpha=\sqrt{n-1}\left(b_n+\frac{\rho}{b_n}\right) \quad .$$

Choosing $b_n \approx \sqrt{2\log n}$ and $\rho=0$, we recover the critical value of Bizeul & Najim[13].

Median and mean of α_F

We now solve for $x=k$, obtaining:

$$\left(\frac{\tilde{\alpha}+\sqrt{\tilde{\alpha}^2-4\rho}}{2}-b_n\right)b_n=k$$

$$\tilde{\alpha}=b_n+\frac{k}{b_n}\frac{b_n\rho}{b_n^2+k}$$

$$\alpha=\sqrt{n-1}\left(b_n+\frac{k}{b_n}\frac{b_n\rho}{b_n^2+k}\right) \quad .$$

Choosing $k=-\log\log 2$ we obtain the approximate median, and using the Euler-Mascheroni constant, the mean.

While Eq. (5) is relatively simple and appealing, it tends to overestimate the probability of feasibility (i.e., underestimate the magnitude of the maximum λ_i) for low α, even when the matrices have considerable size. In fact, the equation is derived under the assumption that the maximum $\hat{\lambda}=\lambda_i$ is associated with the index i corresponding to the row with maximum row sum. While this is often the case, it is not universally true: the exact location of λ_i depends in fact on all of the eigenvalues and eigenvectors of both $\widetilde{B}^{(i)}$ and $\mathbb{1}(b^{(i)})^T$—accordingly, the theorem by O'Rourke and Renfrew[29] contains an unknown term of order one, describing the fluctuations in the outlier eigenvalue. Now, consider the index j associated with the row with the second largest row sum; for a large random matrix, γ_j will be extremely close to γ_i, and the corresponding λ_j might well be lager than λ_i, because of the fluctuations captured in the term in $o(1)$. The same argument can be made for the third-largest γ_k, and so forth, at least up to a certain proportion of the γ_k. Overall, this means that the estimate of $\hat{\lambda}$ tends to be conservative (especially for low α), and that therefore the probability of feasibility is typically overestimated in this region. Fig. 10 exemplifies this issue.

Despite these limitations, Eq. (5) still captures qualitatively the results of simulations: the ordering and spacing of the curves when varying ρ is correct, and the formula, despite its simplicity, describes the shape of the transition to feasibility quite well. It also shows that these curves are *universal*—as long as the matrix B follows the elliptic law, the formula holds asymptotically, irrespective of the fine details of the distribution of B_{ij}. Figure 11 shows the approximation along with the results of simulations for random systems with varying correlation parameter ρ.

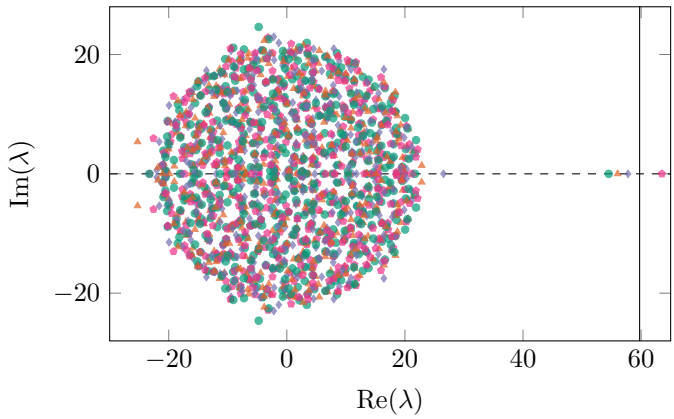

Fig. 10 Spectra of several $Q_i = \mathbb{1}(b^{(i)})^T - \widetilde{B}^{(i)}$ matrices. In green we have the spectrum of the matrix Q_i for the i associated with the largest row sum of B (i.e., $\gamma_i = \hat{\gamma}$), in orange that for the i associated with the second largest row sum, in purple that for the third largest row sum, and in pink that for the fourth largest row sum. Even though we would expect the outlier associated with the largest row sum to be the largest (and thus $\hat{\lambda} \approx \hat{\gamma}$, given that $\rho = 0$), we find that in this case the largest outlier is that associated with the fourth largest row sum. Because of this, we underestimate the value of α_F (vertical line).

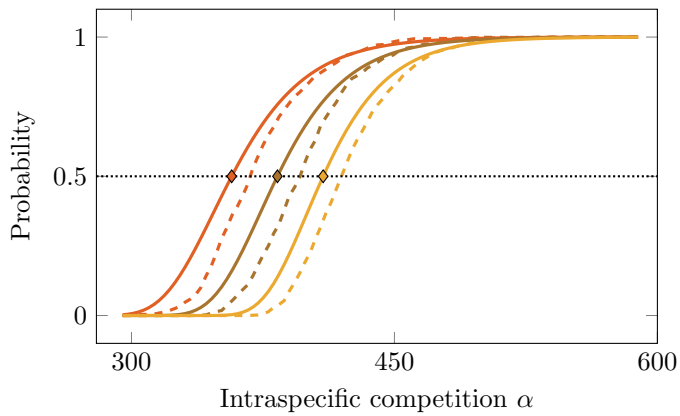

Fig. 11 Simulation results (dashed lines, colors representing the off-diagonal correlation, $n = 10000$), vs. the approximation in Eq. (5) for the probability of feasibility (y-axis) when α is varied (x-axis). The diamonds mark the median of the approximate distribution—i.e., the value of α for which the approximation returns $p_F = 1/2$. Note that the approximation captures the spacing caused by the correlation, but tends to overestimate the probability of feasibility, especially when α is low.

3.7 More accurate approximations of the probability of feasibility

In the section above, we have derived a coarse approximation for the probability of feasibility $p_F(\alpha)$, producing a simple and intuitive equation. The approximation is however inaccurate, and tends to overestimate the probability of feasibility for low values of α. To derive a more accurate approximation, we first state a general result: for any random model with statistically equivalent populations, the mean and correlation between the equilibrium values are fixed; the difference between the models is captured by the variance of the equilibrium values. Having established this result, we provide two methods to approximate the variance of the solution's components, one using the theory of iterative solutions of linear systems, and the other using the theory of resolvents.

When populations are statistical equivalent, then some measures of the distribution of the components of $y^\star$ are especially simple to derive. We have already seen that the expectation of $y_i^\star$, $\mathbb{E}(y_i^\star) = 1/\alpha$, and we denote the variance as $\mathbb{V}(y_i^\star) = \mathbb{E}((y_i^\star)^2) - \mathbb{E}(y_i^\star)^2$, the covariance between $y_i^\star$ and $y_j^\star$ $\mathbb{C}\mathrm{ov}(y_i^\star, y_j^\star) = \mathbb{E}(y_i^\star y_j^\star) - \mathbb{E}(y_i^\star)\mathbb{E}(y_j^\star)$, and their correlation $\mathbb{C}\mathrm{or}(y_i^\star, y_j^\star)$.

Given that the sum of $y_i^\star$ is exactly n/α, then, $(\sum_j y_j^\star)^2 = n^2/\alpha^2$. Expanding, we obtain:

$$
\begin{aligned}
\left(\sum_j y_j^\star\right)^2 &= \frac{n^2}{\alpha^2} \\
\sum_j (y_j^\star)^2 + \sum_i \sum_{j\neq i} y_i^\star y_j^\star &= \frac{n^2}{\alpha^2} \\
n\mathbb{E}((y_i^\star)^2) + n(n-1)\mathbb{E}(y_i^\star y_j^\star) &= \frac{n^2}{\alpha^2} \\
\mathbb{V}(y_i^\star) + \mathbb{E}(y_i^\star)^2 + (n-1)\left(\mathbb{C}\mathrm{ov}(y_i^\star, y_j^\star) + \mathbb{E}(y_i^\star)\mathbb{E}(y_j^\star)\right) &= n\mathbb{E}(y_i^\star)^2 \\
\mathbb{V}(y_i^\star) + \mathbb{E}(y_i^\star)^2 + (n-1)\left(\mathbb{C}\mathrm{ov}(y_i^\star, y_j^\star) + \mathbb{E}(y_i^\star)^2\right) &= n\mathbb{E}(y_i^\star)^2 \\
\mathbb{V}(y_i^\star) + (n-1)\mathbb{C}\mathrm{ov}(y_i^\star, y_j^\star) &= 0 \quad ,
\end{aligned}
$$

where we have exploited the fact that, for a model with statistically equivalent populations, $\mathbb{E}(y_i^\star)\mathbb{E}(y_j^\star) = \mathbb{E}(y_i^\star)^2$, and that the covariance $\mathbb{C}\mathrm{ov}(y_i^\star, y_j^\star)$ must be the same for all i, j. Thus, for *any* unstructured model, we have that the expected covariance is negative. Moreover, for a given variance $\mathbb{V}(y_i^\star) = \sigma^2_{y^\star}$, we have: $\sigma^2_{y^\star} + (n-1)\sigma^2_{y^\star}\mathbb{C}\mathrm{or}(y_i^\star, y_j^\star) = 0$ and thus $\mathbb{C}\mathrm{or}(y_i^\star, y_j^\star) = -1/(n-1)$ for all i and j. Note that this is the maximally negative correlation one can have for equicorrelated random variables. Thus, $y^\star$ has expectation $(1/\alpha)\mathbb{1}$, and covariance matrix $\sigma^2_{y^\star}\left(\frac{n}{n-1}I - \frac{1}{n-1}\mathbb{1}\mathbb{1}^T\right)$. Importantly, the only difference between different unstructured models with random parameters is the value of $\sigma^2_{y^\star}$, while the rest is unchanged. Analyzing increasingly complex models therefore reduces to the problem of determining the variance of $y^\star$.

The joint distribution of $y^\star$ is well-approximated by a multivariate normal[11] in the limit of large n, because of the central limit theorem. To show this, we expand the solution using the Neumann series (see Section 3.7). Whenever $\alpha > \alpha_L$, the solution can be written as:

$$
y^\star = \frac{1}{\alpha}\left(I + \frac{1}{\alpha}C\right)^{-1}\mathbb{1} = \frac{1}{\alpha}\sum_{k=0}^{\infty}\left(-\frac{1}{\alpha}\right)^k C^k\mathbb{1} \quad .
$$

It is well known that any vector b follows a multivariate normal distribution if and only if, for an arbitrary vector w of the same dimension, the scalar product $w^T b$ is distributed according to a univariate normal distribution. Then, for an arbitrary vector w, we consider the scalar variable

$$
\xi = w^T y^\star = \frac{1}{\alpha}\sum_{k=0}^{\infty}\left(-\frac{1}{\alpha}\right)^k w^T C^k\mathbb{1} \quad .
$$

We want to show that ξ is normally distributed. Using the spectral decomposition $C = P\Lambda P^{-1}$, where Λ stands for the diagonal matrix of eigenvalues, and the columns of P contain eigenvectors, we can write

$$
w^T C^k\mathbb{1} = \sum_{\ell=1}^{n}\lambda_\ell^k (P^T w)_\ell (P^{-1}\mathbb{1})_\ell \quad .
$$

Then, $w^T C^k\mathbb{1}$ is a sum of n variables that are equally distributed. Thus, by the central limit theorem, in the limit $n \to \infty$, the distribution of $w^T C^k\mathbb{1}$ can be well-approximated by a univariate normal. The normality of ξ follows from the properties of linear combinations of normal random variables. Therefore, $y^\star$ is asymptotically distributed as

$$y^\star \sim \mathcal{N}\left(\frac{1}{\alpha}\mathbb{1}, \sigma^2_{y^\star}\left(\frac{n}{n-1}I - \frac{1}{n-1}\mathbb{1}\mathbb{1}^T\right)\right) \quad .$$

Because the joint distribution of $y^\star$ is well-approximated by a multivariate normal, we can calculate the probability of feasibility as a function of α, n, as well as $\sigma^2_{y^\star}$ by integrating the probability density function in the positive orthant of $\mathbb{R}^n$—what is called an "orthant probability":

$$p_F(n, \alpha, \sigma^2_{y^\star}) = \int_0^\infty \int_0^\infty \cdots \int_0^\infty \phi(y^\star) dy_1^\star dy_2^\star \cdots dy_n^\star \quad ,$$

where $\phi(y^\star)$ is the probability density function of the normal distribution above. For equicorrelated random variables, the multiple integral above can be reduced to a single integral (albeit involving complex variables when the correlation is negative, such as in this case[49]). Equivalently, one can use the theory of order statistics to write the density function of the first order statistics (i.e., the $\min_i(y_1^\star, y_2^\star, \ldots, y_n^\star)$) and integrate this density function for all positive values of the minimum.

When n is large, however, one can safely ignore the negative correlation between the random variables $y_i^\star$, and pretend they are independent. In this case, we want to determine the probability that all $y_i^\star > 0$. If we subtract the mean from both sides, and call $\zeta_i = y_i^\star - 1/\alpha$, we have that the system is feasible when all $\zeta_i > -1/\alpha$. Dividing both sides by $\sigma_{y^\star}$ and calling $\xi_i = \zeta_i/\sigma_{y^\star}$ we have that feasibility is ensured whenever all the ξ_i, following the standard normal distribution, are larger than $-1/(\alpha\sigma_{y^\star})$; because the standard normal is symmetric about zero, we can finally write:

$$p_F(n, \alpha, \sigma^2_{y^\star}) \approx \Phi\left(\frac{1}{\alpha\sigma_{y^\star}}\right)^n \quad , \tag{6}$$

where $\Phi(\cdot)$ is the cumulative distribution function for a standard normal distribution. The approximation is going to be accurate when n is large. Again, we have reduced the problem of computing the probability of feasibility for a system of a known size n and given level of self-regulation α to the problem of computing $\sigma_{y^\star}$ for the specific model and parameters.

3.8 Computing $\sigma_{y^\star}$ using resolvents

As described in the section above, one can approximate the probability of feasibility using Eq. (6) (cf., Eq. (3) of the main text). To this end, one just needs to compute the variance of a single component of the equilibrium solution $y^\star$. Here, we provide an approximation for the standard deviation $\sigma_{y^\star}$, obtained by computing the expectation of the resolvent of the linear system

$$(\alpha I + B)x^\star = \mathbb{1}\,.$$

As a result of this calculation, we obtain an approximated formula for the deviation $\sigma_{y^\star}$ which, together with Eq. (6), is able to capture accurately the behavior of the probability of feasibility as a function of α, even

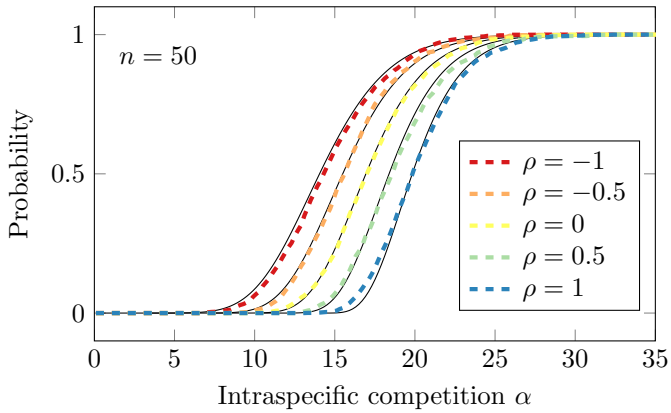

Fig. 12 Approximation using the resolvent method, cf. Eqs. (6) and (27), for matrices B of size $n = 50$, whose entries are sampled from normal distribution with off-diagonal correlation ρ (color). The approximations (solid lines) match the simulations (dashed lines) quite perfectly.

for small sizes. Importantly, this approximation is valid for arbitrary values of the correlation, $\rho \in [-1, 1]$. For example, Fig. 12 shows the quality of the approximation for $n = 50$, and different correlation values.

Because B is a random matrix, the solution of the linear system (if it exists) is given by the random vector

$$x^\star = (\alpha I + B)^{-1}\mathbb{1}\,.$$

To calculate $\sigma_{y^\star}$: we first compute, in the limit of large system sizes ($n \to \infty$), the expectations of the equilibrium solution, $\mathbb{E}(x^\star)$, and the expectation of the scalar product ${x^\star}^T x^\star$, $\mathbb{E}({x^\star}^T x^\star)$. With these two results, it is straightforward to evaluate asymptotically the variance $\sigma^2_{x^\star}$ of a single population as $\sigma^2_{x^\star} = \frac{1}{n}(\mathbb{E}({x^\star}^T x^\star) - \mathbb{E}(x^\star)^T\mathbb{E}(x^\star))$. Then, we use the proportionality relation between $x^\star$ and $y^\star$ to approximate the deviation $\sigma_{y^\star}$: because $y^\star = (\alpha I + C)^{-1}\mathbb{1}$ and $C = B - \mathbb{1}m^T$, one can use the Sherman-Morrison's formula to obtain

$$y^\star = \frac{1}{1 - m^T x^\star} x^\star.$$

Ignoring the correlation between $x^\star$ and $1 - m^T x^\star$, and approximating the expectation of the ratio by the ratio of expectations, we can write

$$y^\star = \frac{1}{1 - m^T x^\star} x^\star \approx \mathbb{E}\left(\frac{1}{1 - m^T x^\star}\right) x^\star \approx \frac{1}{\mathbb{E}\left(1 - m^T x^\star\right)} x^\star,$$

and therefore

$$\sigma_{y^\star} \approx \frac{\sigma_{x^\star}}{\mathbb{E}(1 - m^T x^\star)}, \tag{7}$$

where the standard deviations in this equation refer to the deviation of a single component of vectors $x^\star$ and $y^\star$ —all the components have the same expected deviation.

Because $\sigma^2_{x^\star} = \frac{1}{n}(\mathbb{E}({x^\star}^T x^\star) - \mathbb{E}(x^\star)^T\mathbb{E}(x^\star))$, approximating $\sigma_{y^\star}$ simply requires evaluating the expectations $\mathbb{E}(x^\star)$, $\mathbb{E}({x^\star}^T x^\star)$, and $\mathbb{E}(1 - m^T x^\star)$. We start with the cases $\rho = \pm 1$, which are easier to derive, and allow us to devise a route to obtain the general result for arbitrary ρ. In all the cases, it will be critical to employ a

sort of "symmetrization" of matrix B (which is only ensured for $\rho = 1$, in which case the matrix is indeed symmetric).

Symmetric case, $\rho = 1$***.***

First, we consider the case $\rho = 1$, i.e., the fully symmetric case ($B_{ij} = B_{ji}$). We have to compute the aforementioned expectations.

Computation of $\mathbb{E}(x^\star)$ *using resolvents.* To begin with, we will focus on evaluating the expectation $\mathbb{E}(x^\star)$:

$$\mathbb{E}(x^\star) = \mathbb{E}((\alpha I + B)^{-1})\mathbb{1}.$$

Let $z \in \mathbb{C}$ be a complex number with non-zero imaginary part. Then the resolvent of matrix B is defined as the inverse $(zI + B)^{-1}$, and we want to evaluate its expectation

$$G(z) = \mathbb{E}((zI + B)^{-1}),$$

computed over the ensemble of random matrices to which B belongs. $G(z)$ can be expressed formally via the corresponding Neumann series expansion (which is always convergent if $\mathcal{I}m(z) > 0$). Our aim in this section is to obtain an asymptotic expression for $G(z)$, valid when $n \to \infty$, and then compute the expectations $\mathbb{E}(x^\star)$, $\mathbb{E}({x^\star}^T x^\star)$ and $\mathbb{E}(1 - m^T x^\star)$ in terms of $G(z)$. Once we have all the expressions, we will write $z = \alpha + i\epsilon$, and take the limit $\epsilon \to 0$ to obtain the final result.

In the limit $n \to \infty$, the expectation $G(z)$ of the resolvent satisfies the so-called matrix Dyson equation [50]:

$$G(z) = \frac{1}{z} I + \frac{1}{z} \mathbb{E}(BG(z)B)G(z). \tag{8}$$

This equation is exact for normally-distributed matrix entries B_{ij}, and is asymptotically true for arbitrary distributions (see Appendix B for a derivation of this equation).

In order to solve this nonlinear equation, we make a guess for the expectation $\Sigma := \mathbb{E}(BG(z)B)$. Let G_{ij} be the elements of matrix $G(z)$. Since $\mathbb{E}(B_{ij}B_{kl}) = \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}$ (where δ_{ij} stands for the Kronecker symbol) holds for $\rho = 1$, then

$$\Sigma_{il} = \mathbb{E}\left(\sum_{j,k} B_{ij} G_{jk} B_{kl}\right) = \sum_{j,k} \mathbb{E}(B_{ij}B_{kl}) G_{ij},$$

because G_{ij} is an expectation itself. Hence

$$\Sigma_{il} = G_{li} + \delta_{il} \sum_j G_{jj}.$$

Therefore, in the limit $n \to \infty$, it is expected that the diagonal of matrix Σ dominates the off-diagonal entries. Thus, we adopt the guess $\Sigma = a(z)I$, and look for solutions for the function $a(z)$ —observe that, given that populations are statistically equivalent, we expect all the diagonal elements of the expectation $\mathbb{E}(BG(z)B)$ to be the same.

We substitute this assumption into Eq. (8) and solve for $G(z)$:

$$G(z) = \frac{1}{z}I + \frac{a(z)}{z}G(z); \qquad G(z) = (z - a(z))^{-1}I.$$

Now we can compute Σ in a different way:

$$\Sigma = \mathbb{E}(BG(z)B) = (z - a(z))^{-1}\mathbb{E}(B^2) = \frac{n-1}{z - a(z)}I = a(z)I,$$

where we have used that $\mathbb{E}((B^2)_{ii}) = \sum_j \mathbb{E}(B_{ij}B_{ji}) = \sum_j \mathbb{E}(B_{ij}^2) = n - 1$, and the last step follows because the sum over j only contains $n-1$ terms (given that $B_{ii} = 0$). In addition, if $i \neq k$, we have that $\mathbb{E}((B^2)_{ik}) = \sum_j \mathbb{E}(B_{ij}B_{jk}) = \sum_j(\delta_{ij}\delta_{jk} + \delta_{ik}\delta_{jj}) = \delta_{ik} + \delta_{ik}\sum_j \delta_{jj} = 0$. Thus, $\mathbb{E}(B^2) = (n-1)I$ and the last equation implies that the unknown function $a(z)$ satisfies the quadratic equation

$$a(z)(z - a(z)) = n - 1.$$

Solving for $a(z)$, we obtain

$$a(z) = \frac{z \pm \sqrt{z^2 - 4(n-1)}}{2}.$$

According to the definition of the resolvent, $(zI + B)^{-1}$, it must behave as $z^{-1}I$ as $z \to \infty$. Therefore, Dyson equation (8) implies that $z^{-1}\Sigma \to 0$ as $z \to \infty$, which means that we have to choose the minus sign in $a(z)$. Consequently, the resolvent expectation is expressed as

$$G(z) = (z - a(z))^{-1}I = \frac{2}{z + \sqrt{z^2 - 4(n-1)}}I = \frac{z - \sqrt{z^2 - 4(n-1)}}{2(n-1)}I.$$

Having computed the resolvent expectation in the asymptotic limit $n \to \infty$, we immediately find an expression for the expected value of $x^\star$:

$$\mathbb{E}(x^\star) = \lim_{\epsilon \to 0} \mathbb{E}(((\alpha + i\epsilon)I + B)^{-1})\mathbb{1} = \lim_{\epsilon \to 0} G(\alpha + i\epsilon)\mathbb{1} = \frac{\alpha - \sqrt{\alpha^2 - 4(n-1)}}{2(n-1)}\mathbb{1}. \tag{9}$$

Computation of $\mathbb{E}({x^\star}^T x^\star)$. The second expectation to be computed can be written as

$$\mathbb{E}({x^\star}^T x^\star) = \mathbb{1}^T\mathbb{E}(((\alpha I + B)^{-1})^T(\alpha I + B)^{-1})\mathbb{1} = \mathbb{1}^T\mathbb{E}((\alpha I + B)^{-2})\mathbb{1},$$

because $((\alpha I + B)^{-1})^T = ((\alpha I + B)^T)^{-1} = (\alpha I + B)^{-1}$ (since $B^T = B$). Here we denote by $(\alpha I + B)^{-2}$ the square power of $(\alpha I + B)^{-1}$. Thus, we need to evaluate the expectation of $(\alpha I + B)^{-2}$. Let us define it as

$$K(z) = \mathbb{E}((\alpha I + B)^{-2}).$$

We show in the Appendix that $K(z)$ can be written in a simple way in terms of the derivative of $G(z)$, namely:

$$K(z) = -G'(z) = \frac{1}{2(n-1)} \left(\frac{z}{\sqrt{z^2 - 4(n-1)}} - 1 \right) I, \tag{10}$$

which implies that

$$\mathbb{E}(x^{\star T} x^\star) = \lim_{\epsilon \to 0} \mathbb{1}^T K(\alpha + i\epsilon) \mathbb{1} = \frac{n}{2(n-1)} \left(\frac{\alpha}{\sqrt{\alpha^2 - 4(n-1)}} - 1 \right). \tag{11}$$

Using the expression obtained above for $\mathbb{E}(x^\star)$, we find, after simplifications, the following asymptotic limit for the variance $\sigma^2_{x^\star}$:

$$\sigma^2_{x^\star} = \frac{1}{n}(\mathbb{E}(x^{\star T} x^\star) - \mathbb{E}(x^\star)^2) = \frac{1}{2(n-1)^2} \left(n - 1 - \alpha^2 + \frac{\alpha(\alpha^2 - 3(n-1))}{\sqrt{\alpha^2 - 4(n-1)}} \right). \tag{12}$$

Computation of $\mathbb{E}(1 - m^T x^\star)$. To evaluate the last expectation we need, we first observe that $(\alpha I + B)x^\star = \mathbb{1}$ implies, by left-multiplying by $\mathbb{1}^T$, that

$$\alpha \mathbb{1}^T x^\star + \mathbb{1}^T B x^\star = n.$$

Equivalently,

$$\frac{\alpha}{n} \mathbb{1}^T x^\star + m^T x^\star = 1,$$

because, by definition, $m = \frac{1}{n} B^T \mathbb{1}$. Therefore

$$1 - m^T x^\star = \frac{\alpha}{n} \mathbb{1}^T x^\star.$$

The expectation $\mathbb{E}(\mathbb{1}^T x^\star) = \mathbb{1}^T \mathbb{E}(x^\star)$ follows from $\mathbb{E}(x^\star)$: using Eq. (9), we finally obtain

$$\mathbb{E}(1 - m^T x^\star) = \frac{\alpha}{n} \mathbb{1}^T \mathbb{E}(x^\star) = \frac{\alpha^2 - \alpha\sqrt{\alpha^2 - 4(n-1)}}{2(n-1)}.$$

Substituting this result and Eq. (12) into Eq. (7), and after some simplifications, the approximation of $\sigma_{y^\star}$ finally reads

$$\sigma_{y^\star} \approx \frac{1}{\alpha} \sqrt{\frac{\alpha}{2\sqrt{\alpha^2 - 4(n-1)}} - \frac{1}{2}}. \tag{13}$$

This result is valid for $\rho = 1$, large n, and is well defined for $\alpha > 2\sqrt{n-1}$. In particular, this condition is satisfied for $\alpha \geq 2\sqrt{n} = \alpha_S$, i.e., above the global stability threshold for $\rho = 1$. The choice of the distribution for B_{ij} has no effect (asymptotically). This approximation compares very well with simulated data for the distributions considered here (Fig. 12), even for small n.

Skew-symmetric case, $\rho = -1$***.***

In order to proceed as in the symmetric case, we have to find a Dyson equation satisfied by the resolvent expectation $\mathbb{E}((\alpha I + B)^{-1})$. The formalism of the matrix Dyson equation can only be applied to Hermitian

matrices [50]; however, for $\rho = -1$ matrix B is skew-symmetric ($B^T = -B$). In order for the Dyson equation to hold, we transform the original matrix in order to make it Hermitian: let $Q = iB$, where i is the imaginary unit. Then $Q^\dagger = (iB)^\dagger = -iB^T = iB = Q$, and hence Q is Hermitian (here, † means complex conjugate transpose).

To solve $(\alpha I + B)x^\star = 1$ we focus on the transformed linear system

$$(i\zeta I + iB)x^\star = i\mathbb{1},$$

for $\zeta \in \mathbb{C}$. Then let $z = i\zeta$ (with positive imaginary part) and write the system to solve as $(zI + Q)x^\star = i\mathbb{1}$, with Q Hermitian. We just need to compute the expectation

$$G(z) = \mathbb{E}((zI + Q)^{-1})$$

The same matrix Dyson equation holds, see Eq. (8):

$$G(z) = \frac{1}{z}I + \frac{1}{z}\mathbb{E}(QG(z)Q)G(z). \tag{14}$$

Following the same arguments as before, we look for solutions such that $\mathbb{E}(BG(z)B) = a(z)I$. Therefore,

$$\mathbb{E}(QG(z)Q) = i^2\mathbb{E}(BG(z)B) = -a(z)I.$$

We introduce this expression into the Dyson equation and solve for $G(z)$:

$$G(z) = (z + a(z))^{-1}I.$$

Now, given that $G(z)$ is a multiple of the identity, we can evaluate $\mathbb{E}(QG(z)Q)$ in a different way:

$$\mathbb{E}(QG(z)Q) = -\mathbb{E}(BG(z)B) = -(z + a(z))^{-1}\mathbb{E}(B^2).$$

As for the symmetric case, one can check that $\mathbb{E}(B_{ij}^2) = 0$ for $i \neq j$. Only diagonal expectations remain: $\mathbb{E}(B_{ii}^2) = \sum_j \mathbb{E}(B_{ij}B_{ji}) = -\sum_j \mathbb{E}(B_{ij}^2) = -(n-1)$, because $B_{ii} = 0$ by definition. Therefore,

$$\mathbb{E}(QG(z)Q) = \frac{n-1}{z + a(z)}I.$$

As a consequence, $a(z)$ must satisfy the quadratic equation

$$a(z)(z + a(z)) + n - 1 = 0,$$

which yields

$$a(z) = \frac{-z \pm \sqrt{z^2 - 4(n-1)}}{2},$$

and

$$G(z) = \frac{2}{z \pm \sqrt{z^2 - 4(n-1)}} I.$$

In order for $G(z) \sim \frac{1}{z} I$ as $z \to \infty$, we have to choose the positive sign. Then, we can write

$$G(z) = \frac{z - \sqrt{z^2 - 4(n-1)}}{2(n-1)} I.$$

Now we go back to the original variable $z = i\zeta$ and matrix $Q = iB$:

$$G(i\zeta) = \mathbb{E}((i\zeta I + iB)^{-1}) = i^{-1}\mathbb{E}((\zeta I + B)^{-1}) = -i\mathbb{E}((\zeta I + B)^{-1}) = -iG(\zeta),$$

where $G(\zeta) = \mathbb{E}((\zeta I + B)^{-1})$ is the expectation we want to evaluate. On the other hand,

$$G(i\zeta) = -i\frac{\sqrt{\zeta^2 + 4(n-1)} - \zeta}{2(n-1)} = -iG(\zeta).$$

Thus, by substituting $\zeta = \alpha + i\epsilon$, and taking the limit $\epsilon \to 0$, we find the following expression for the resolvent expectation in the skew-symmetric case:

$$G(\alpha) = \lim_{\epsilon \to 0} \mathbb{E}(((\alpha + i\epsilon)I + B)^{-1}) = \frac{\sqrt{\alpha^2 + 4(n-1)} - \alpha}{2(n-1)} I. \tag{15}$$

This result automatically implies that $\mathbb{E}(x^\star) = g(\alpha)\mathbb{1}$, where we have defined

$$g(\alpha) = \frac{\sqrt{\alpha^2 + 4(n-1)} - \alpha}{2(n-1)}. \tag{16}$$

Computation of $\mathbb{E}(x^{\star T} x^\star)$. Because $x^\star = (\alpha I + B)^{-1}\mathbb{1}$ and $x^{\star T} = \mathbb{1}^T(\alpha I - B)^{-1}$ (we have used, in the last calculation, that $B^T = -B$), we can write the expectation $\mathbb{E}(x^{\star T} x^\star)$ as

$$\mathbb{E}(x^{\star T} x^\star) = \mathbb{1}^T \mathbb{E}((\alpha I - B)^{-1}(\alpha I + B)^{-1})\mathbb{1}.$$

Now we resort to the identity

$$(\alpha I - B)^{-1}(\alpha I + B)^{-1} = \frac{1}{2\alpha}(\alpha I + B)^{-1} + \frac{1}{2\alpha}(\alpha I - B)^{-1},$$

which can be easily proven by left- and right-multiplying both sides by $\alpha I - B$ and $\alpha I + B$, respectively. By taking expectations, we obtain

$$\mathbb{E}((\alpha I - B)^{-1}(\alpha I + B)^{-1}) = \frac{1}{2\alpha}\mathbb{E}((\alpha I + B)^{-1}) + \frac{1}{2\alpha}\mathbb{E}((\alpha I - B)^{-1}).$$

It holds that $\mathbb{E}((\alpha I - B)^{-1}) = \mathbb{E}((\alpha I + B)^{-1})$ —this can be verified by solving Eq. (14) for $Q = iB^T = -iB$. Therefore,

$$\mathbb{E}((\alpha I - B)^{-1}(\alpha I + B)^{-1}) = \frac{1}{\alpha}\mathbb{E}((\alpha I + B)^{-1}) = \frac{g(\alpha)}{\alpha} I, \tag{17}$$

and $\mathbb{E}(x^{\star T} x^\star) = ng(\alpha)/\alpha$. Notice that $\mathbb{E}(x^\star)^T \mathbb{E}(x^\star) = ng(\alpha)^2$, and thus

$$\sigma_{x^\star}^2 = \frac{1}{n}(\mathbb{E}(x^{\star T}x^\star) - \mathbb{E}(x^\star)^T\mathbb{E}(x^\star)) = \frac{g(\alpha)}{\alpha} - g(\alpha)^2. \tag{18}$$

Final approximation for $\sigma_{y^\star}$ in the skew-symmetric case. The calculation we did for the expectation $\mathbb{E}(1 - m^T x^\star)$ in the symmetric case is still valid in the skew-symmetric case,

$$\mathbb{E}(1 - m^T x^\star) = \frac{\alpha}{n}\mathbb{1}^T\mathbb{E}(x^\star) = \alpha g(\alpha).$$

Finally, substituting the last expression and Eq. (18) into Eq. (7), we find

$$\sigma_{y^\star} \approx \frac{\sqrt{g(\alpha)(\alpha^{-1} - g(\alpha))}}{\alpha g(\alpha)} = \frac{1}{\alpha}\sqrt{\frac{1}{\alpha g(\alpha)} - 1} = \frac{1}{\alpha}\sqrt{\frac{\sqrt{\alpha^2 + 4(n-1)}}{2\alpha} - \frac{1}{2}}, \tag{19}$$

where we have used that $1/g(\alpha) = (\sqrt{\alpha^2 + 4(n-1)} + \alpha)/2$ and performed some straightforward manipulations. Again, this expression is universal and does not depend (in the limit of large n) on the distribution chosen to draw matrix elements. It compares very well with simulations (Fig. 12).

General case, arbitrary correlation.

We take further advantage of the theory of resolvents to compute an approximation for $\sigma_{y^\star}$ for arbitrary values of the correlation coefficient ρ. We will recover the two limits studied above ($\rho = 1$ and $\rho = -1$) as particular cases of this more general calculation. The rationale of the proof in the general case is based also on a matrix Dyson equation, which is slightly different to that used for $\rho = 1$ and $\rho = -1$: studying these limiting cases, for which the derivation is simpler, allows one to intuit the general route one can follow to tackle the general case of arbitrary correlation.

One critical feature of the Dyson equation used in both cases is that it applies for Hermitian matrices. For $\rho = 1$, matrix B is symmetric by definition, and for $\rho = -1$ we applied the same matrix Dyson equation to matrix $Q = iB$, which is Hermitian and satisfies $Q^\dagger = Q$. For arbitrary ρ we resort to the "Hermitization trick" due to Girko [51]: consider, for $z \in \mathbb{C}$, the $(2n) \times (2n)$ matrix

$$Q(z) = \begin{pmatrix} \mathbb{0} & zI + B \\ z^\star I + B^T & \mathbb{0} \end{pmatrix},$$

where $z^\star$ denotes the conjugate of complex number z, and $\mathbb{0}$ stands for the null matrix of dimension n. Matrix $Q(z)$ is called the Hermitization of matrix $zI + B$, for B real. One can see that Q is Hermitian, i.e., $Q^\dagger = Q$.

$Q(z)$ belongs to a general class of self-adjoint (Hermitian) random matrices with correlated entries, for which Erdős, Krüger and Schröder proved in [50] that the resolvent of matrix Q,

$$S(z, w) := (iwI_2 + Q(z))^{-1},$$

satisfies, in the limit of large n, the following non-linear matrix Dyson equation:

$$-\begin{pmatrix} I & \mathbb{0} \\ \mathbb{0} & I \end{pmatrix} = \left(\begin{pmatrix} iwI & -zI \\ -z^\star I & iwI \end{pmatrix} + \Lambda(S(z,w)) \right) S(z,w), \tag{20}$$

where I is the identity matrix of dimension n, I_2 is the $(2n)\times(2n)$ identity matrix, and Λ is a matrix function defined as

$$\Lambda(A) = (n-1) \begin{pmatrix} A_{22} & \rho A_{21} \\ \rho A_{12} & A_{11} \end{pmatrix}$$

for an arbitrary complex, square matrix $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ with dimension $2n$ and square $(n \times n)$ block structure. Observe that the expectation $\mathbb{E}((iwI_2 + Q(z))^{-1})$ also satisfies the same matrix Dyson equation.

The definition of Λ differs slightly from the one reported by Erdős, Krüger and Schröder because, in reference [50], the authors re-scaled variances and correlations as n^{-1}, and we are not adopting such re-scaling. This explains why the factor $n-1$ appears in front of the matrix in the last equation. Recall that, in our case, $\mathbb{E}(B_{ij}^2) = 1$ and $\mathbb{E}(B_{ij}B_{ji}) = \rho$ for $j \neq i$, which implies that $\mathbb{E}(B^T B) = \mathbb{E}(BB^T) = (n-1)I$ and $\mathbb{E}(B^2) = \mathbb{E}((B^T)^2) = \rho(n-1)I$.

Can we extract the resolvent $(zI+B)^{-1}$ from the resolvent $S(z,w)$ of the Hermitized matrix? Let us partition S in square blocks as

$$S(z,w) = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}.$$

Then, matrix S satisfies

$$\begin{pmatrix} iwI & zI+B \\ z^\star I + B^T & iwI \end{pmatrix} \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} = \begin{pmatrix} I & \mathbb{0} \\ \mathbb{0} & I \end{pmatrix},$$

which implies

$$iwS_{11} + (zI+B)S_{21} = I.$$

Therefore, the resolvent $(zI+B)^{-1}$ we want to calculate is the lower-left block of the augmented resolvent $S(z,w)$ evaluated at $w=0$. Our goal is to compute an asymptotic approximation to $S(z,w)$ via the solution of the matrix Dyson equation (20).

As shown in [52], in the limit $n \to \infty$, the solution to Eq. (20) is of the form

$$S(z,w) = \begin{pmatrix} iv(z,w)I & b(z,w)^\star I \\ b(z,w)I & iv(z,w)I \end{pmatrix}, \tag{21}$$

where the image of v is positive, b takes values on the complex plane, and S satisfies $\mathcal{I}m(S) = \frac{1}{2i}(S - S^\dagger)$, which is positive definite. Since we are interested in the particular value $w = 0$, we can compute explicit expressions for functions v and b in terms of z. From now on, we omit the dependence on z and w in v and b.

We obtain the non-linear system of equations satisfied by v and b by substituting Eq. (21) into the matrix Dyson equation (20):

$$\begin{aligned} -1 &= -wv - zb + (n-1)(-v^2 + \rho b^2), \\ 0 &= wb^\star - zv + (n-1)(vb^\star + \rho vb). \end{aligned} \tag{22}$$

The two remaining equations are simply the complex conjugate of the two equations above. From the second equation we can solve for v:

$$v = -\frac{wb^\star}{(n-1)b^\star + (n-1)\rho b - z}, \tag{23}$$

and substitute into the first equation to get a fourth-order equation to be satisfied by b:

$$((n-1)\rho b^2 - zb + 1)((n-1)b^\star + (n-1)\rho b - z)^2 + w^2 b^\star((n-1)b^\star + (n-1)\rho b - z) + (n-1)w^2 b^{\star 2} = 0.$$

Since the solution of this polynomial equation will be a continuous function of w at $w = 0$, we can set $w = 0$ to get

$$((n-1)\rho b^2 - zb + 1)((n-1)b^\star + (n-1)\rho b - z)^2 = 0.$$

For v to be well defined, the only solution for b is the one that cancels the first factor:

$$b(z, 0) = \frac{z - \sqrt{z^2 - 4\rho(n-1)}}{2\rho(n-1)}, \tag{24}$$

where we have chosen the minus sign so that the resolvent $(zI + B)^{-1}$ behaves as $z^{-1}I$ in the limit $z \to \infty$ (recall that $(zI + B)^{-1} \approx S_{21}|_{w=0} = b(z, 0)I$ in the limit of large n). Therefore, we can obtain the expectation $\mathbb{E}(x^\star)$ by substituting $z = \alpha + i\epsilon$ in the expression for $b(z, 0)$ and letting $\epsilon \to 0$:

$$\mathbb{E}(x^\star) = \lim_{\epsilon \to 0} b(\alpha + i\epsilon, 0)\mathbb{1} = \frac{\alpha - \sqrt{\alpha^2 - 4\rho(n-1)}}{2\rho(n-1)}\mathbb{1}. \tag{25}$$

Note that this expectation is always a positive real number for $\alpha > 2\sqrt{\rho(n-1)}$, even for negative ρ. By taking the limit $\rho \to 0$ in the last equation we obtain $\mathbb{E}(x^\star) = \alpha^{-1}\mathbb{1}$ for $\rho = 0$. One can easily check that Eq. (25) reduces to Eq. (9) for $\rho = 1$ and Eq. (16) for $\rho = -1$.

Computation of $\mathbb{E}(x^{\star T} x^\star)$. As for the derivations above, we need to calculate the expectation $\mathbb{E}(x^{\star T} x^\star)$ in order to approximate σ_y according to Eq. (7). Here we follow an approach similar to the $\rho = -1$ case. Let us focus on the matrix product

$$W := (iwI_2 + Q(z))^{-1\dagger}(iwI_2 + Q(z))^{-1} = (-iw^\star I_2 + Q(z))^{-1}(iwI_2 + Q(z))^{-1},$$

where in the last step we have used $Q^\dagger = Q$. This is in analogy with the product $(zI - B)^{-1}(zI + B)^{-1}$, which was used to compute the expectation $\mathbb{E}(x^{\star T}x^\star)$ in the skew-symmetric case ($\rho = -1$). For arbitrary correlation we use the resolvent of the Hermitization instead. By writing W as a block matrix, it satisfies

$$(iwI_2 + Q(z))^\dagger W(iwI_2 + Q(z)) = \begin{pmatrix} -iw^\star I & zI + B \\ z^\star I + B^T & -iw^\star I \end{pmatrix} \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix} \begin{pmatrix} iwI & zI + B \\ z^\star I + B^T & iwI \end{pmatrix} = \begin{pmatrix} I & \mathbb{0} \\ \mathbb{0} & I \end{pmatrix}.$$

Setting $w = 0$, the lower-right block of this equality implies

$$(z^\star I + B^T)W_{11}(zI + B) = I, \qquad W_{11} = (z^\star I + B^T)^{-1}(zI + B)^{-1}.$$

Therefore, the expectation of the upper-left block of matrix W is precisely related to the expectation we want to compute:

$$\mathbb{E}(x^{\star T}x^\star) = \mathbb{1}^T\mathbb{E}((\alpha I + B^T)^{-1}(\alpha I + B)^{-1})\mathbb{1}.$$

Thus, our next goal is to compute W. To this end, as we did in the skew-symmetric case, we write

$$W = c_1(-iw^\star I_2 + Q(z))^{-1} + c_2(iwI_2 + Q(z))^{-1}.$$

Substituting into $(iwI_2 + Q(z))^\dagger W(iwI_2 + Q(z)) = I_2$ we obtain the condition

$$c_1(iwI_2 + Q(z)) + c_2(-iw^\star I_2 + Q(z)) = I_2,$$

which implies that $c_1 = -c_2$ and $i(c_1 w - c_2 w^\star) = 1$. Therefore, $c_2 = \frac{i}{w+w^\star} = -c_1$ and we can write

$$W = \frac{i}{w + w^\star}\left((iwI_2 + Q(z))^{-1} - (-iw^\star I_2 + Q(z))^{-1}\right) = \frac{i}{w + w^\star}(S(z, w) - S(z, -w^\star)).$$

We now use the asymptotic limit of matrix S, given by Eq. (21), and Eq. (23) to obtain

$$W_{11} = \frac{v(z, -w^\star) - v(z, w)}{w + w^\star} I = \frac{b^\star}{(n-1)b^\star + (n-1)\rho b - z} I,$$

which turns out to be independent of w (this is true in the limit $n \to \infty$). Therefore,

$$\mathbb{E}((\alpha I + B^T)^{-1}(\alpha I + B)^{-1}) = \lim_{\epsilon \to 0} \frac{b^\star(\alpha + i\epsilon, 0)}{(n-1)b^\star(\alpha + i\epsilon, 0) + (n-1)\rho b(\alpha + i\epsilon, 0) - (\alpha + i\epsilon)} I.$$

Using Eq. (24), and after some simplifications, we finally get the asymptotic formula

$$\mathbb{E}((\alpha I + B^T)^{-1}(\alpha I + B)^{-1}) = \frac{2}{\alpha(\alpha + \sqrt{\alpha^2 - 4\rho(n-1)}) - 2(1+\rho)(n-1)} I. \tag{26}$$

One can use this expression to recover Eq. (11) for $\rho = 1$, and Eq. (17) for $\rho = -1$.

Final result for the deviation $\sigma_{y^\star}$. Let us define

$$g(\alpha) := \frac{\alpha - \sqrt{\alpha^2 - 4\rho(n-1)}}{2\rho(n-1)}, \qquad g_2(\alpha) := \frac{2}{\alpha^2 + \alpha\sqrt{\alpha^2 - 4\rho(n-1)}) - 2(1+\rho)(n-1)}.$$

Hence, $\mathbb{E}(x^\star) = g(\alpha)\mathbb{1}$ and $\mathbb{E}(x^{\star T} x^\star) = n g_2(\alpha)$. This allows to approximate the variance of $x^\star$ as

$$\sigma_{x^\star}^2 = \frac{1}{n}(\mathbb{E}(x^{\star T} x^\star) - \mathbb{E}(x^\star)^T \mathbb{E}(x^\star)) = g_2(\alpha) - g(\alpha)^2.$$

As in the particular cases $\rho = \pm 1$, the expectation $\mathbb{E}(1 - m^T x^\star)$ can be computed as

$$\mathbb{E}(1 - m^T x^\star) = \frac{\alpha}{n}\mathbb{1}^T \mathbb{E}(x^\star) = \alpha g(\alpha).$$

Finally,

$$\sigma_{y^\star} \approx \frac{\sqrt{g_2(\alpha) - g(\alpha)^2}}{\alpha g(\alpha)} = \frac{1}{\alpha}\sqrt{\frac{g_2(\alpha)}{g(\alpha)^2} - 1}.$$

After some simplifications, we obtain

$$\frac{1}{\alpha\sigma_{y^\star}} \approx \sqrt{\frac{\alpha^2 + \alpha\sqrt{\alpha^2 - 4\rho(n-1)}}{2(n-1)} - (1+\rho)}, \tag{27}$$

which is Eq. (4) of the main text. For $\rho \leq 0$ this expression always yields a real number, and for $\rho > 0$ does so when $\alpha \geq 2\sqrt{\rho(n-1)}$; in both cases, the stability threshold $\alpha_S = \sqrt{2n(1+\rho)}$ is contained within the interval in which Eq. (27) is valid. The formula reduces to Eqs. (13) and (19) for $\rho = 1$ and $\rho = -1$, respectively. The agreement with the probability of feasibility obtained by sampling matrices is remarkable (Fig. 12), even for moderate sizes.

Observe that $(\alpha\sigma_{y^\star})^{-1}$, i.e., the argument of the standardized normal cumulative density function in Eq. (6), can be written in terms of the re-scaled intraspecific competition $\tilde{\alpha} := \alpha/\sqrt{n-1}$, thereby being independent of n:

$$\frac{1}{\alpha\sigma_{y^\star}} \approx \sqrt{\frac{\tilde{\alpha}^2 + \tilde{\alpha}\sqrt{\tilde{\alpha}^2 - 4\rho}}{2} - (1+\rho)}.$$

However, this does not make all the probability of feasibility curves to collapse for different system sizes, because the probability of feasibility is precisely $\Phi((\alpha\sigma_{y^\star})^{-1})^n$, and rising the cumulative density function to the power of n introduces an intrinsic separation between the curves as n increases (see Fig. 13, which provides a comparison between the resolvent approximation and the approximation via Richardson iteration).

Finally, we use this approximation to estimate the probability of feasibility at the stability threshold, $\alpha_S \approx \sqrt{2n(1+\rho)}$, obtaining:

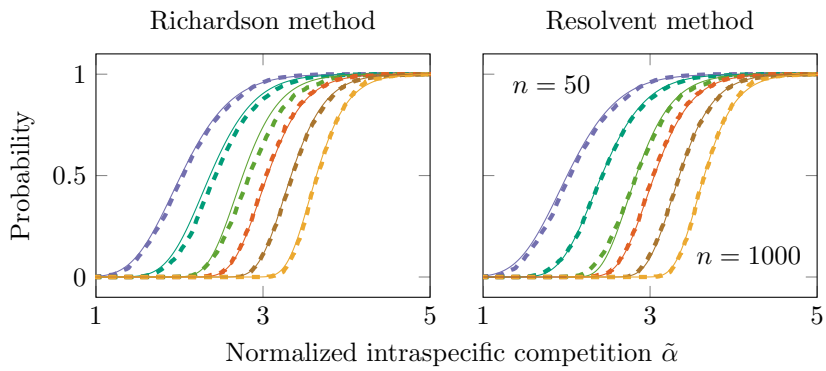

Fig. 13 Approximate probability of feasibility for the two approximations presented in the text. Left panel: approximation based on Richardson's iteration; Right panel: approximation based on resolvents. Intraspecific competition is normalized ($\tilde{\alpha} = \alpha/\sqrt{n-1}$) so that the argument of $\Phi(\cdot)$, i.e., the CDF of the standard normal distribution that appears in Eq. (3) of the main text, does not depend on n. We observe a very good agreement with simulations for both methods, even for small sizes: colder colors illustrate the case of $n = 50$ (with purple for $\rho = -1$, teal for $\rho = 0$ and green for $\rho = 1$) and warmer colors the case of $n = 1000$ (with orange, brown and yellow marking $\rho = -1, 0, 1$, respectively). Dashed lines report simulated data. While Richardson's method tends to overestimate p_F for small sizes, the resolvent method overestimates (underestimates) it for negative (positive) correlations and small sizes.

$$p_F(\alpha_S) \approx \Phi\left(\sqrt{\frac{1+\rho+\sqrt{n(1+\rho)(n-(n-2)\rho)}}{n-1}}\right)^n .$$

The probability of feasibility at the stability threshold tends to zero as n increases. For example, for $n = 50$, we find $p_F(\alpha_S) \approx 2.4 \times 10^{-4}$ for $\rho = 0$, $p_F(\alpha_S) \approx 6 \times 10^{-8}$ for $\rho = 1$ and $p_F(\alpha_S) \approx 8.9 \times 10^{-16}$ for $\rho = -1$.

3.9 Computing $\sigma_{y^\star}$ by iteration

In this section, we approximate $\sigma_{y^\star}$ by approximating $y^\star$ for a given α. To this end, we use two well-known results.

Neumann series.

Whenever the spectral radius of A, $\xi(A) < 1$, the inverse of $(I - A)$ can be written as:

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k \quad ,$$

where $A^0 = I$.

Take the system in $y^\star$, and divide both sides by $\alpha > 0$:

$$\left(I + \frac{1}{\alpha}C\right) y^\star = \frac{1}{\alpha}\mathbb{1} \quad .$$

Whenever the spectral radius of C, $\xi(C) < \alpha$, we can write:

$$
\begin{aligned}
y^\star &= \frac{1}{\alpha}\left(I + \frac{1}{\alpha}C\right)^{-1}\mathbb{1} \\
&= \frac{1}{\alpha}\sum_{k=0}^{\infty}\left(-\frac{1}{\alpha}C\right)^{k}\mathbb{1} \\
&= \frac{1}{\alpha}\left(I - \frac{1}{\alpha}C + \frac{1}{\alpha^2}C^2 - \cdots\right)\mathbb{1} \\
&= \frac{1}{\alpha}\mathbb{1} - \frac{1}{\alpha^2}C\mathbb{1} + \frac{1}{\alpha^3}C^2\mathbb{1} - \cdots \quad .
\end{aligned}
$$

We note that whenever $\alpha > \alpha_L$, the Neumann series converges.

Richardson iteration.

The system $Ax = b$ can be solved iteratively, starting from an initial guess $x(0)$ as:

$$
x(k+1) = (I - \omega A)x(k) + \omega b \quad ,
$$

where ω is an iteration parameter that can always be chosen such that the iteration converges to the solution of $Ax = b$. When $A = A^T$ and is positive definite, the optimal iteration parameter is simply:

$$
\omega = \frac{2}{\lambda_1(A) + \lambda_n(A)} \quad ,
$$

i.e., the reciprocal of the average between the smallest and the largest eigenvalue of A.

If we multiply the equations we want to solve by $(I + \frac{\kappa}{\alpha}C^T)$, the right-hand side of the equations is unchanged, because $C^T\mathbb{1} = \mathbb{0}$:

$$
\left(I + \frac{\kappa}{\alpha}C^T\right)\left(I + \frac{1}{\alpha}C\right)y^\star = \frac{1}{\alpha}\mathbb{1} \quad .
$$

However, the left-hand side changes, and thus the speed at which an iterative method would converge to its solution depends on κ.

Based on numerical simulations, to speed up convergence we choose three different values of κ, depending on the correlation between the entries of B; in particular, we choose $\kappa = 1$ when the correlation is $\rho = -1$ and B is skew-symmetric; $\kappa = 0$ when $\rho = 0$, and $\kappa = -1$ when $\rho = 1$ and B is symmetric.

We start with $\rho = -1$, and multiply both sides of the system of equations by $\left(I + \frac{1}{\alpha}C^T\right)$:

$$
\begin{aligned}
\left(I + \frac{1}{\alpha}C^T\right)\left(I + \frac{1}{\alpha}C\right)y^\star &= \frac{1}{\alpha}\left(I + \frac{1}{\alpha}C^T\right)\mathbb{1} \\
\left(I + \frac{1}{\alpha}(C + C^T) + \frac{1}{\alpha^2}C^TC\right)y^\star &= \frac{1}{\alpha}\mathbb{1} \\
\left(I + \frac{1}{\alpha}(B - \mathbb{1}m^T + B^T - m\mathbb{1}^T) + \frac{1}{\alpha^2}(B^T - m\mathbb{1}^T)(B - \mathbb{1}m^T)\right)y^\star &= \frac{1}{\alpha}\mathbb{1} \\
\left(I + \frac{1}{\alpha}(B - \mathbb{1}m^T - B - m\mathbb{1}^T) + \frac{1}{\alpha^2}(-B - m\mathbb{1}^T)(B - \mathbb{1}m^T)\right)y^\star &= \frac{1}{\alpha}\mathbb{1} \\
\left(I + \frac{1}{\alpha}(-\mathbb{1}m^T - m\mathbb{1}^T) + \frac{1}{\alpha^2}(-B^2 - n\,mm^T)\right)y^\star &= \frac{1}{\alpha}\mathbb{1} \\
Sy^\star &= \frac{1}{\alpha}\mathbb{1} \quad .
\end{aligned}
$$

Note that, a) the multiplication does not alter the right-hand side of the equations; b) S is positive semi-definite, because it is of the form $G^T G$; c) when $\rho = -1$, B is skew-symmetric, and thus $B^T = -B$; d) when $\rho = -1$, necessarily $m^T \mathbb{1} = 0$ (hence the two matrices $\mathbb{1} m^T$ and $m \mathbb{1}^T$ are nilpotent—have all eigenvalues zero).

Next, we choose the iteration parameter, by approximating $S \approx (I - \frac{1}{\alpha^2} B^2)$. Given that the eigenvalues of B^2 are the squared eigenvalues of B, we have $\lambda_1(B^2) \approx -4n$ and $\lambda_n(B^2) \approx 0$. Hence:

$$\omega = \frac{2}{(1 - \frac{1}{\alpha^2} 0) + (1 + \frac{1}{\alpha^2} 4n)} = \frac{\alpha^2}{\alpha^2 + 2n} \quad .$$

Finally, we take $y(0) = \frac{1}{\alpha}\mathbb{1} - \frac{1}{\alpha^2} C \mathbb{1}$ (as given by the first two terms of the Neumann series), $y(1) = (I - \omega S) y(0) + \frac{1}{\alpha}\omega \mathbb{1}$, and compute the approximate variance as:

$$\mathbb{V}(y_i^\star) = \mathbb{E}((y_i^\star)^2) - \mathbb{E}(y_i^\star)^2 \approx \frac{1}{n}\mathbb{E}(y(1)^T y(1)) - \frac{1}{\alpha^2} \quad .$$

Note that this approximation will depend on the first few moments of the distribution of B_{ij} (and not only on their mean and variance, as in the approximation using small-rank perturbations, or that using resolvents). The expectation of $y(1)^T y(1)$ is presented in Appendix B.

For $\rho = 1$, we multiply both sides of the system of equations by $\left(I - \frac{1}{\alpha} C^T\right)$, instead:

$$\begin{aligned}
\left(I - \frac{1}{\alpha} C^T\right)\left(I + \frac{1}{\alpha} C\right) y^\star &= \frac{1}{\alpha}\left(I - \frac{1}{\alpha} C^T\right)\mathbb{1} \\
\left(I + \frac{1}{\alpha}(C - C^T) - \frac{1}{\alpha^2} C^T C\right) y^\star &= \frac{1}{\alpha}\mathbb{1} \\
\left(I + \frac{1}{\alpha}(B - \mathbb{1} m^T - B^T + m \mathbb{1}^T) - \frac{1}{\alpha^2}(B^T - m \mathbb{1}^T)(B - \mathbb{1} m^T)\right) y^\star &= \frac{1}{\alpha}\mathbb{1} \\
\left(I + \frac{1}{\alpha}(B - \mathbb{1} m^T - B + m \mathbb{1}^T) - \frac{1}{\alpha^2}(B - m \mathbb{1}^T)(B - \mathbb{1} m^T)\right) y^\star &= \frac{1}{\alpha}\mathbb{1} \\
\left(I + \frac{1}{\alpha}(m \mathbb{1}^T - \mathbb{1} m^T) - \frac{1}{\alpha^2}(B^2 - n\, m m^T)\right) y^\star &= \frac{1}{\alpha}\mathbb{1} \\
H y^\star &= \frac{1}{\alpha}\mathbb{1} \quad .
\end{aligned}$$

Note that, a) the multiplication again does not alter the right-hand side of the equations; b) H is positive semi-definite as long as $\alpha > \alpha_L$; c) when $\rho = 1$, B is symmetric, and thus $B^T = B$.

Next, we choose the iteration parameter, by approximating $H \approx (I - \frac{1}{\alpha^2} B^2)$. Given that the eigenvalues of B^2 are the squared eigenvalues of B, we have $\lambda_n(B^2) \approx 4n$ and $\lambda_1(B^2) \approx 0$. Hence:

$$\omega = \frac{2}{(1 - \frac{1}{\alpha^2} 0) + (1 - \frac{1}{\alpha^2} 4n)} = \frac{\alpha^2}{\alpha^2 - 2n} \quad .$$

Finally, we take $y(0) = \frac{1}{\alpha}\mathbb{1} - \frac{1}{\alpha^2} C \mathbb{1}$, $y(1) = (I - \omega H) y(0) + \frac{1}{\alpha}\omega \mathbb{1}$, and compute the approximate variance as before. Also for this case, the details of the calculation are presented in Appendix B.

For $\rho = 0$ we choose $\kappa = 0$, leaving the system unchanged. Given that we can safely assume that $\alpha > \alpha_L$, we have that $I - \omega(I + \frac{1}{\alpha} C) = (1 - \omega) I - \frac{\omega}{\alpha} C$ has spectral radius less than one (and thus the iteration converges) when we choose $\omega = 1$, such that the Richardson iteration matches what one would obtain by

truncating the Neumann series. Finally, we set $y(0) = \frac{1}{\alpha}\mathbb{1} - \frac{1}{\alpha^2}C\mathbb{1}$ and compute the variance as done for the other cases. The details of the calculation are also presented in Appendix B.

We conclude by stressing that we a) have started with a guess given by the truncation of the Neumann series that includes only the first two terms, and b) have performed only one step of the Richardson iteration. While the calculations are already quite elaborate for these choices, including either more terms of the Neumann series, or performing more than one step for the Richardson iteration (possibly, updating the parameter ω) would almost certainly result in an even more accurate approximation—allowing one to model smaller systems, or systems in which the distribution of coefficients has extreme skewness/kurtosis, etc. Appendix B shows that the approximation above is however quite precise for moderate n and the distributions we have been analyzing in this work.

3.10 Consequences for dynamics

In Sections 3.6 and 3.8, we have shown that, for large n, a random competitive system that is feasible is also stable. The competitive GLV system can give rise to all types of dynamics, including limit cycles (for $n \geq 3$) and chaos (for $n \geq 4$) [17, 18]. Non-equilibrium coexistence via limit cycles or chaos requires a *feasible* equilibrium[19], but we have seen in Sec. 1.3 that whenever $\alpha > \alpha_F$, then $\alpha > \alpha_S \geq \alpha_L$, and thus the stability of the equilibrium is ensured for any choice of $r > \mathbb{0}$. Hence, we conclude that limit cycles or chaos are going to be almost impossible to observe for large random competitive systems.

If the relationship $\alpha_F > \alpha_S$ is maintained for moderate sizes n as well (as suggested by the calculations in Sec. 3.8), then the dynamical consequences are even more far-reaching. In a competitive GLV system, lack of feasibility guarantees that one or more populations will go extinct[19]. Suppose that we initialize a random competitive system with n populations at arbitrary initial conditions; this system will have a certain α. If $\alpha > \alpha_F$, the system is feasible and stable; trajectories will converge to the equilibrium, and if the equilibrium is perturbed, dynamics will return to it. If, on the other hand, $\alpha < \alpha_F$, then the system is not feasible, and therefore some populations will go extinct. There are two distinct cases:

Case i): $\alpha_S < \alpha$: the system will converge to a saturated equilibrium (as shown by the Lyapunov function in Section 1.3); the same populations will go extinct irrespective of the (positive) initial conditions; the populations that go extinct cannot re-invade the system; the populations that coexist do so stably, and when their densities are perturbed, the system will asymptotically return to the saturated equilibrium.

Case ii): $\alpha < \alpha_S$: this is the most interesting case. Some populations will go extinct, because there is no feasible equilibrium. However, which populations will go extinct *depends on the initial conditions*. The system can therefore display priority effects and alternative stable states. Every time a population goes extinct, the system size shrinks to $n-1$, $n-2$, etc. Accordingly, the threshold values α_S and α_F are also decreased, while the value of α is unchanged. As such, provided that the α_S for the reduced system is still smaller than the corresponding α_F, the system will keep losing populations until $\alpha > \alpha_S$, thus eventually attaining Case i).

Therefore, for large n (and provided that $\alpha_S < \alpha_F$ when populations are lost), the ultimate result is always the same: the surviving populations coexist robustly at a globally-attractive equilibrium. Moreover, while the populations that went extinct at the onset of the dynamics could potentially re-invade the system (only when we start in Case ii), the last few populations (i.e., those that went extinct once $\alpha > \alpha_S$) cannot.

3.11 Consequences for assembly

We can go a step further, and consider whether these systems, obtained by collapsing a larger system, can instead be built from the ground up.

Take a system with n populations coexisting at a feasible equilibrium $x^\star$ (i.e., we have $\alpha > \alpha_F$), possibly obtained by collapsing a larger system, as detailed in the previous section. If we have $\alpha > \alpha_F > \alpha_S$, which is almost invariably the case for a large system, then any feasible sub-community is automatically stable, and might or might not be invaded by the remaining populations.

In such a case, determining whether the equilibrium $x^\star$ can be built by sequential introduction of population can be accomplished by considering an assembly graph where the nodes are the feasible sub-communities that can be formed by selecting some of the n population (out of 2^n possibilities, including the "empty" state in which all populations are absent), and an edge connects two nodes if an invasion can move the system from one state to the other[20]. If the system can be built from the ground up, then there exists a path in the graph connecting the empty state all the way up to the state with n populations (Fig. 14).

We performed simulations in which a) we built a feasible system of n populations, b) built the corresponding assembly graph, and c) tallied the probability of finding such a path in the graph as we increased n. Figure 15 shows that, even for moderate n, the probability is very high, and rapidly converges to one.

Taken collectively, our results show that *robust, stable coexistence* (of a sub-community) *is the norm* for large random competitive communities. For large system with random parameters, dynamics will not display cycles or chaos. Moreover, the final community (i.e., that obtained once the dynamics have elapsed, and extinctions have occurred) is such that priority effects are excluded. The final community can be assembled from the ground-up by introducing one population at a time.

Appendices

A Distribution of the coefficients B_{ij}

Here we detail the distributions we have used to sample interaction coefficients, including their probability distribution/mass functions, and the first six moments (needed for the Richardson approximation). All the distributions are shifted and re-scaled in order to have mean zero and unit variance. We chose a distribution that is discrete and has finite support; a distribution that is continuous and has finite support, and one that is continuous and has infinite support. For each distribution, we include a parameter β that can be varied to regulate the skewness of the distributions. The three distributions are drawn in Fig. 16.

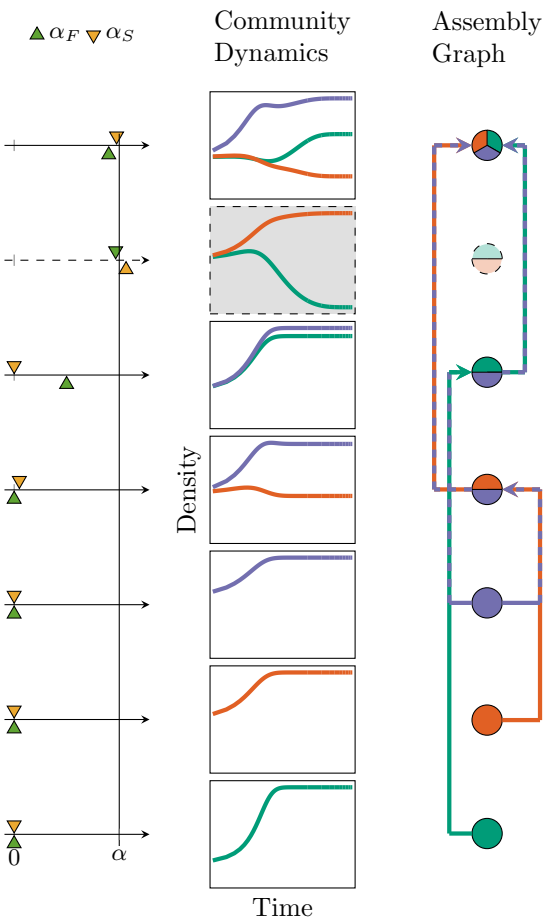

Fig. 14 Assembly graph for a simple competitive community. We show how the interplay between the values of α, α_S and α_F (left) determine dynamics (center), and therefore the existence and connectivity of nodes (right) in the assembly graph associated with the system. Starting from the top, we consider a three-species community with $\alpha > \alpha_F > \alpha_S$; thus, the three populations coexist stably (center, showing typical dynamics) and this community is a node in the assembly graph (right). For the three two-species communities that we can form, we have that $\alpha > \alpha_F$ for two of them, while for one community $\alpha_F > \alpha$, and therefore dynamics will lead to extinctions; consequently, the corresponding node is not present in the assembly graph. Finally, by definition, all single-species communities are feasible and stable. We connect two nodes in the assembly graph if an invasion can transition the system from a sub-community (node in the graph) to another. We say that the full community can be built from the ground up via sequential invasions if there is a path starting at a node encoding a single-species community and ending with the full community.

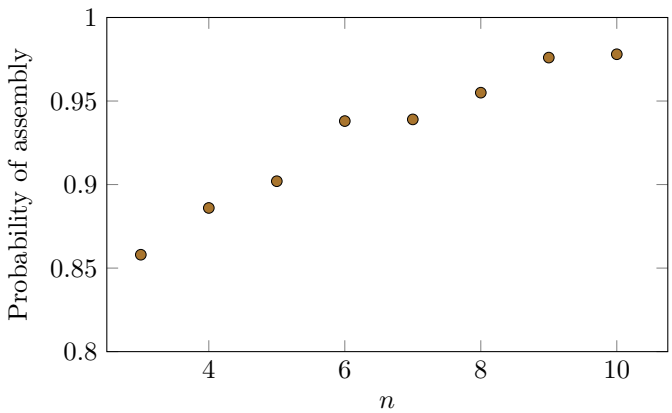

Fig. 15 As we vary n (x-axis), we calculate the probability that a random, feasible system can be built from the ground up via successive invasions (y-axis). We performed 1000 simulations for each n, constructed the corresponding assembly graph, and, for each graph, we determined whether a path starting from the empty state reaches the node corresponding to the full community. The probability that such a path exists, and therefore the system can be built from the ground up, rapidly converges to one as n increases.

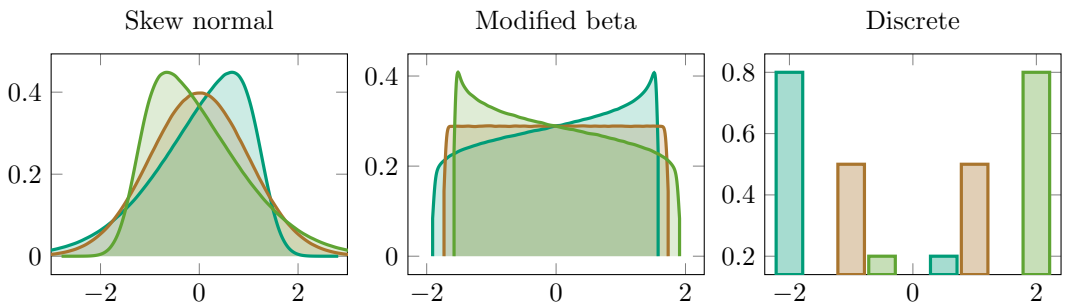

Fig. 16 For our simulations, we sampled the coefficients B_{ij} from distributions with mean zero and unit variance. We choose a continuous distribution with infinite support (Skew normal), a continuous distribution with finite support (Modified beta) and a discrete distribution with finite support (Discrete) to probe the generality of our results. For each distribution, we include a parameter (β) that controls the skewness (colors).

A.1 Skew normal distribution

The PDF is given by

$$\Phi(x) = \frac{2}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x-\mu}{2\sigma^2}\right) \int_{-\infty}^{\beta\left(\frac{x-\mu}{\sigma}\right)} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)$$

with expectation

$$\mathbb{E}(x) = \mu + \frac{\sqrt{\frac{2}{\pi}}\beta\sigma}{\sqrt{1+\beta^2}}$$

and variance

$$\mathbb{V}(x) = \sigma^2\left(1 - \frac{2\beta^2}{\pi(1+\beta^2)}\right)$$

We then solve the system of equations $\mathbb{E}(x) = 0$ and $\mathbb{V}(x) = 1$ for μ and σ, obtaining

$$\mu^\star = -\frac{\sqrt{2}\beta}{\sqrt{\pi + \beta^2(\pi-2)}} \qquad \sigma^\star = \frac{\sqrt{\pi(1+\beta^2)}}{\sqrt{\pi + \beta^2(\pi-2)}}$$

Finally, we sample interactions from the skew normal distribution with $\mu = \mu^\star$ and $\sigma = \sigma^\star$. The moments of the resulting distribution are given in the table below. Note that for $\beta = 0$ we recover the standard normal distribution.

$$\begin{aligned}
\mu_1 &= 0 \\
\mu_2 &= 1 \\
\mu_3 &= -\frac{\sqrt{2}(\pi-4)\beta^3}{(\pi + (\pi-2)\beta^2)^{3/2}} \\
\mu_4 &= \frac{-12\beta^4 + 3\pi^2(1+\beta^2)^2 - 4\pi\beta^2(3+\beta^2)}{(\pi + (\pi-2)\beta^2)^2} \\
\mu_5 &= \frac{\sqrt{2}\beta^3(-10(\pi-4)\pi + (16 + (20-7\pi)\pi)\beta^2)}{(\pi + (\pi-2)\beta^2)^{5/2}} \\
\mu_6 &= \frac{-40\beta^6 + 15\pi^3(1+\beta^2) - 20\pi\beta^4(9+5\beta^2) - 6\pi^2\beta^2(155 + 10\beta^2 + \beta^4)}{(\pi + (\pi-2)\beta^2)^3}
\end{aligned}$$

A.2 Discrete distribution

The PMF is given by

$$\Phi(x) = \begin{cases} \xi_1 & \text{with probability } p \\ \xi_2 & \text{with probability } 1-p \end{cases}$$

The expectation is

$$\mathbb{E}(x) = \xi_1 p + \xi_2(1-p)$$

and the variance is

$$\mathbb{V}(x) = \xi_1^2 p + \xi_2^2(1-p) - E(x)^2$$

We solve the system of equations imposing that $\mathbb{E}(x) = 0$ and $\mathbb{V}(x) = 1$ for ξ_2 and p, obtaining

$$p = \frac{1}{1+\xi_1^2} \qquad \beta = -\frac{1}{\xi_1}$$

We call $\beta = \xi_1$; this parameter controls the skewness of the distribution. For $\beta = 1$ we obtain the Bernoulli distribution with outcomes ± 1 and skewness 0. The first six moments for this distribution are

$$\begin{aligned}
\mu_1 &= 0 \\
\mu_2 &= 1 \\
\mu_3 &= \beta - \frac{1}{\beta} \\
\mu_4 &= \beta^2 - 1 + \frac{1}{\beta^2} \\
\mu_5 &= \frac{1}{\beta^3}(\beta^2-1)(\beta^4+1) \\
\mu_6 &= \beta^4 - \beta^2 + 1 - \frac{1}{\beta^2} + \frac{1}{\beta^4}
\end{aligned}$$

A.3 Beta distribution

The PDF is given by

$$\Phi(x) = \frac{x^{\xi_1-1}(1-x)^{\xi_2-1}}{\mathtt{B}(\xi_1,\xi_2)}$$

where $\mathtt{B}(\xi_1,\xi_2) = \frac{\Gamma(\xi_1)\Gamma(\xi_2)}{\Gamma(\xi_1+\xi_2)}$, and Γ is the Gamma function. The expectation is

$$\mathbb{E}(x) = \frac{\xi_1}{\xi_1+\xi_2}$$

and the variance

$$\mathbb{V}(x) = \frac{\xi_1\xi_2}{(\xi_1+\xi_2)^2(\xi_1+\xi_2+1)}$$

The beta distribution has support between 0 and 1, and therefore it cannot have mean zero and unit variance. For this reason, we perform a change of variables to shift it and re-scale it. That is, we sample x

from a beta distribution and then construct another random variable z by subtracting the mean (shift) and dividing by the square root of the variance (re-scale). The result is a random variable following the beta distribution with mean 0 and variance 1. We take $\xi_1 = 1/\beta$, $\xi_2 = \beta$, and define:

$$z = \sqrt{\frac{1+\beta+\beta^2}{\beta^3}}(-1 + x + x\beta^2)$$

The first six moments of the distribution defined by the random variable z are

$$\begin{aligned}
\mu_1 &= 0 \\
\mu_2 &= 1 \\
\mu_3 &= \frac{2(\beta-1)\sqrt{1+\frac{1}{\beta}+\beta}}{1+\beta} \\
\mu_4 &= -21 + \frac{6}{\beta} + \beta\left(6 + \frac{54}{1+\beta(3+\beta)}\right) \\
\mu_5 &= \frac{4(\beta-1)(1+\frac{1}{\beta}+\beta)^{3/2}(6+5\beta+5\beta^3+6\beta^4)}{(1+\beta)(1+\beta(3+\beta))(1+\beta(4+\beta))} \\
\mu_6 &= \frac{5(1+\beta+\beta^2)^2(24+\beta(-22+\beta(16+\beta(-1-22\beta+24\beta^2))))}{\beta^2(1+\beta(3+\beta))(1+\beta(4+\beta))(1+\beta(5+\beta))}
\end{aligned}$$

When $\beta = 1$, we recover the uniform distribution $\mathcal{U}[-\sqrt{3}, \sqrt{3}]$, which has mean zero, unit variance, and zero skewness.

B Calculation of $\sigma_{y^\star}$

Here we present the details of the calculations for the approximation of $\sigma_{y^\star}$.

B.1 Through iterative solutions of linear systems

We have discussed the approach to approximating the variance $\sigma^2_{y^\star}$ as $\sigma^2_{y^\star} \approx \frac{1}{n} y(1)^T y(1) - \frac{1}{\alpha^2}$, where $y(1)$ is given by the first step of the Richardson iteration:

$$y(1) = y(0) - \omega M y(0) + \omega b$$

We have chosen: a) $y(0)$ as given by the first two terms of the Neumann series, $y(0) = \frac{1}{\alpha}\mathbb{1} - \frac{1}{\alpha^2}C\mathbb{1}$; b) the right-hand side $b = \frac{1}{\alpha}\mathbb{1}$; c) the choice of ω and M depend on the correlation, as follows:

$$\begin{array}{lll}
\rho = -1 & \rho = 0 & \rho = 1 \\
\omega = \dfrac{\alpha^2}{\alpha^2 + 2n} & \omega = 1 & \omega = \dfrac{\alpha^2}{\alpha^2 - 2n} \\
M = \left(I + \dfrac{1}{\alpha}C^T\right)\left(I + \dfrac{1}{\alpha}C\right) & M = \left(I + \dfrac{1}{\alpha}C\right) & M = \left(I - \dfrac{1}{\alpha}C^T\right)\left(I + \dfrac{1}{\alpha}C\right)
\end{array}$$

To compute the approximate variance, we take a generic matrix $B = (B_{ij})$ of size n, compute the Richardson iteration, and then the approximate variance:

$$
\begin{aligned}
\sigma_{y^\star}^2 &\approx \frac{1}{n} y(1)^T y(1) - \frac{1}{\alpha^2} \\
&= \frac{1}{n}\left((y(0) - \omega M y(0) + \omega b)^T (y(0) - \omega M y(0) + \omega b)\right) - \frac{1}{\alpha^2} \\
&= \frac{1}{n}\left(y(0)^T y(0) + \omega^2 y(0)^T M^T M y(0) + \omega^2 b^T b - 2\omega y(0)^T M y(0) + 2\omega b^T y(0) - 2\omega^2 b^T M y(0)\right) - \frac{1}{\alpha^2}
\end{aligned} \tag{28}
$$

Each term of the expansion ($y(0)^T y(0)$, $\omega^2 y(0)^T M^T M y(0)$, etc.) is a polynomial that depends on some of the coefficients B_{ij}, as well as n, α and ω. To compute the expected variance, we then take the expectation of each of the terms; in particular, we need to determine the expectation of polynomials in B_{ij}, such as $B_{ij}^p B_{kl}^q$. When the coefficients are independent (e.g., $\rho = 0$), only the terms in which all the coefficients are raised to a power greater than one will yield a non-zero expectation. In particular, we write:

$$
\begin{aligned}
\mathbb{E}(B_{ij}) &= 0 \\
\mathbb{E}(B_{ij}^2) &= \mu_2 \\
\mathbb{E}(B_{ij}^3) &= \mu_3 \\
\mathbb{E}(B_{ij}^4) &= \mu_4 \\
&\dots
\end{aligned}
$$

Such that for example $\mathbb{E}(B_{ij}^2 B_{kl}^3) = \mu_2 \mu_3$. Note that two distributions might have both $\mathbb{E}(B_{ij}) = 0$ and $\mathbb{E}(B_{ij}^2) = 1$, but different higher moments (for example, the standard normal distribution yields $(\mu_2, \mu_3, \mu_4, \mu_5, \mu_6) = (1, 0, 3, 0, 15)$, while when we assign $B_{ij} = 1$ or $B_{ij} = -1$ with equal probability, we obtain $(\mu_2, \mu_3, \mu_4, \mu_5, \mu_6) = (1, 0, 1, 0, 1)$); hence, the approximations will take the first few moments (up to μ_6, for the choice of starting conditions and Richardson iteration) of the distribution of coefficients into account.

When entries are correlated, we have a more difficult task at hand, because for example $\mathbb{E}(B_{ij} B_{ji}) = \rho$, and this relationship needs to be accounted for when taking higher powers (such that $\mathbb{E}(B_{ij}^2 B_{ji}^2) \neq 0$). However, the problem is simplified when we consider either $\rho = 1$ or $\rho = -1$, in which cases $\mathbb{E}(B_{ij} B_{ji}) = \mathbb{E}(B_{ij}^2) = \mu_2$ (when $\rho = 1$) or $\mathbb{E}(B_{ij} B_{ji}) = \mathbb{E}(-B_{ij}^2) = -\mu_2$ (when $\rho = -1$), and thus we just need to take the powers of B_{ij} into account. For simplicity, we thus consider only the cases of $\rho = 0$, $\rho = -1$, and $\rho = 1$. The calculations can be extended to the case of a general ρ, the cost being to have to keep track of many more combinations of coefficients.

Computing these expectations becomes increasingly complex. We start from the terms that do not depend on the choice of M (and therefore are the same irrespective of the correlation).

The simplest terms in Eq. (28) are $b^T b$:

$$
b^T b = \frac{1}{\alpha^2} \mathbb{1}^T \mathbb{1} = \frac{n}{\alpha^2}
$$

and

$$b^T y(0) = \frac{n}{\alpha^2}$$

A term that is slightly more complicated to compute is $y(0)^T y(0)$, which will depend on the correlation, but not on the choice of M. We report the detailed calculation, because this is the approach we followed to compute all the remaining terms. We can write $y(0)^T y(0)$ as:

$$\begin{aligned}
y(0)^T y(0) &= \left(\left(\frac{1}{\alpha}\mathbb{1}^T - \frac{1}{\alpha^2}\mathbb{1}^T C^T\right)\left(\frac{1}{\alpha}\mathbb{1} - \frac{1}{\alpha^2}C\mathbb{1}\right)\right) \\
&= \mathbb{1}^T\left(\frac{1}{\alpha^2}I + \frac{1}{\alpha^4}C^T C\right)\mathbb{1} \\
&= \frac{n}{\alpha^2} + \frac{1}{\alpha^4}\mathbb{1}^T\left((B^T - m\mathbb{1}^T)(B - \mathbb{1}m^T)\right)\mathbb{1} \\
&= \frac{n}{\alpha^2} + \frac{1}{\alpha^4}\mathbb{1}^T\left(B^T B - n\, mm^T\right)\mathbb{1} \\
&= \frac{n}{\alpha^2} + \frac{1}{\alpha^4}\left(\mathbb{1}^T B^T B\mathbb{1} - n\,\mathbb{1}^T mm^T\mathbb{1}\right) \\
&= \frac{n}{\alpha^2} + \frac{1}{\alpha^4}\left(\mathbb{1}^T B^T B\mathbb{1} - n\,(m^T\mathbb{1})^2\right) \\
&= \frac{n}{\alpha^2} + \frac{1}{\alpha^4}\left(\sum_i\sum_j\sum_k B_{ik}B_{jk} - n\left(\frac{1}{n}\sum_i\sum_j B_{ij}\right)^2\right) \\
&= \frac{n}{\alpha^2} + \frac{1}{\alpha^4}\left(\sum_i\sum_j\sum_k B_{ik}B_{jk} - \frac{1}{n}\left(\sum_i\sum_j B_{ij}\right)^2\right)
\end{aligned}$$

Next, we need to take the expectations for these sums:

$$\mathbb{E}(y(0)^T y(0)) = \frac{n}{\alpha^2} + \frac{1}{\alpha^4}\left(\mathbb{E}\left(\sum_i\sum_j\sum_k B_{ik}B_{jk}\right) - \frac{1}{n}\mathbb{E}\left(\left(\sum_i\sum_j B_{ij}\right)^2\right)\right)$$

These expectations depend on the second moment $\mu_2 = 1$ as well as the correlation ρ; in particular, we have:

$$\mathbb{E}\left(\sum_i\sum_j\sum_k B_{ik}B_{jk}\right) = \mathbb{E}\left(\sum_i\sum_k B_{ik}B_{ik}\right) + \mathbb{E}\left(\sum_i\sum_{j\neq i}\sum_k B_{ik}B_{jk}\right) = n(n-1)\mu_2 + 0$$

and, using that $\mu_2 = 1$:

$$\begin{aligned}
E\left(\left(\sum_i\sum_j B_{ij}\right)^2\right) &= E\left(\sum_i\sum_j B_{ij}^2\right) + E\left(\sum_i\sum_j\sum_{k\neq i}\sum_{l\neq j} B_{ij}B_{kl}\right) \\
&= n(n-1)\mu_2 + n(n-1)\rho = n(n-1)(1+\rho)\mu_2
\end{aligned}$$

Thus yielding:

$$\mathbb{E}(y(0)^T y(0)) = \frac{n}{\alpha^2} + \frac{1}{\alpha^4}\mu_2(n-1)(n-1-\rho).$$

Independent entries, $\rho = 0$

For the remaining four terms in Eq. (28), we need to perform calculations that differ for $\rho = 0$, $\rho = -1$ and $\rho = 0$, because in each case we have a different matrix M. We start with the simplest case, which is obtained when the coefficients of B are sampled independently, and $\rho = 0$. We have:

$$
\begin{aligned}
M &= I + \frac{1}{\alpha} C \\
\mathbb{E}(b^T b) &= \frac{n}{\alpha^2} \\
\mathbb{E}(b^T y(0)) &= \frac{n}{\alpha^2} \\
\mathbb{E}(y(0)^T y(0)) &= \frac{n}{\alpha^2} + \frac{1}{\alpha^4}(n-1)^2 \mu_2
\end{aligned}
$$

We need to compute the remaining terms. $b^T M y(0)$ is very easy to compute:

$$
b^T M y(0) = b^T y(0) = \frac{n}{\alpha^2}
$$

Computing $y(0)^T M y(0)$ is more involved, but can be accomplished using the same approach used for $y(0)^T y(0)$, and noting that only the symmetric part of M matters. We obtain:

$$
\mathbb{E}(y(0)^T M y(0)) = \frac{n}{\alpha^2} - \frac{(n-1)^2}{n\alpha^5} \mu_3
$$

Finally, the most complicated term, $y(0)^T M^T y(0)$, for which we obtain:

$$
\mathbb{E}(y(0)^T M^T M y(0)) = \frac{n}{\alpha^2} + \frac{(n-1)(n^4 - 3n^3 + 4n^2 - 6n + 3)}{n^2 \alpha^6} \mu_2^2 + \frac{(n-1)^2}{n^2 \alpha^6} \mu_4
$$

Combining the terms according to Eq. (28), we find that the expected variance for $\rho = 0$ is:

$$
\mathbb{E}(\sigma_{y^\star}^2) \approx \frac{\mu_2 (n-1)^2}{\alpha^4 n} + \frac{2\mu_3 (n-1)^2}{\alpha^5 n^2} + \frac{\mu_4 (n-1)^2}{\alpha^6 n^3} + \frac{\mu_2^2 \left(n^4 - 3n^3 + 4n^2 - 6n + 3\right)(n-1)}{\alpha^6 n^3}
$$

A few notable features of this equation are: a) the approximation depends on the first few moments of the distribution, and not only on the mean and variance of B_{ij}; b) the moment μ_3, summarizing the skewness of the distribution, is multiplied by a positive coefficient, hence a positive skew will increase the variance (and hence decrease the probability of feasibility, compared to a negative skew); c) however, the terms associated with the higher moments vanish when n is large and $\alpha > \sqrt{n}$, again confirming that for sufficiently large n the behavior will be universal.

Symmetric entries, $\rho = 1$

We repeat the calculation for the case of symmetric matrices B, for which we choose a different matrix M and a different parameter ω, in order to speed up the convergence. We have:

$$
\begin{aligned}
M &= \left(I - \frac{1}{\alpha}C^T\right)\left(I + \frac{1}{\alpha}C\right) \\
\mathbb{E}(b^T b) &= \frac{n}{\alpha^2} \\
\mathbb{E}(b^T y(0)) &= \frac{n}{\alpha^2} \\
\mathbb{E}(y(0)^T y(0)) &= \frac{n}{\alpha^2} + \frac{1}{\alpha^4}(n-1)(n-2)\mu_2
\end{aligned}
$$

We need to compute the remaining terms:

$$
b^T M y(0) = \frac{n}{\alpha^2} + \mu_3 \frac{(n-1)(n-2)^2}{n\alpha^5}
$$

$$
\begin{aligned}
\mathbb{E}(y(0)^T M y(0)) &= \frac{n}{\alpha^2} - \frac{(n-1)(n-2)(n^3-5n^2+18n-12)}{n^2\alpha^6}\mu_2^2 \\
&\quad + \frac{2(n-1)(n-2)^2}{n\alpha^5}\mu_3 - \frac{(n-1)(n-2)^3}{n^2\alpha^6}\mu_4 \\
\mathbb{E}(y(0)^T M^T M y(0)) &= \frac{n}{\alpha^2} + \frac{(n-1)(n-2)(n^3-5n^2+18n-12)}{n^2\alpha^6}\mu_2^2 \\
&\quad + \frac{(n-1)(n-2)(n^5-17n^4+105n^3-392n^2+724n-432)}{n^3\alpha^8}\mu_2^3 \\
&\quad - \frac{4(n-1)(n-2)(n^4-12n^3+49n^2-88n+40)}{n^3\alpha^7}\mu_3\mu_2 \\
&\quad + \frac{2(n-1)(n-2)(n^4-14n^3+56n^2-92n+64)}{n^3\alpha^8}\mu_3^2 \\
&\quad + \frac{2(n-1)(n-2)(2n^4-17n^3+76n^2-140n+92)}{n^3\alpha^8}\mu_4\mu_2 \\
&\quad + \frac{2(n-1)(n-2)^2}{n\alpha^5}\mu_3 - \frac{2(n-1)(n-2)^4}{n^3\alpha^7}\mu_5 + \frac{(n-1)(n-2)^4}{n^3\alpha^7}\mu_6
\end{aligned}
$$

Combining all terms, we find that the expected variance for $\rho = 1$ is:

$$
\begin{aligned}
\mathbb{E}(\sigma_{y^\star}^2) \approx\ & \frac{\mu_2(n-1)(n-2)}{\alpha^4 n} - \frac{4\mu_3(n-1)(n-2)^2\omega}{\alpha^5 n^2} \\
& - \frac{4\mu_2\mu_3(n-1)\left(n^4-12n^3+49n^2-88n+40\right)(n-2)\omega^2}{\alpha^7 n^4} \\
& + \frac{\mu_6(n-1)(n-2)^4\omega^2}{\alpha^8 n^4} - \frac{2\mu_5(n-1)(n-2)^4\omega^2}{\alpha^7 n^4} \\
& + \frac{\mu_2^2(n-2)(n-1)\left(n^3-5n^2+18n-12\right)\omega^2}{\alpha^6 n^3} \\
& + \frac{2\mu_2^2(n-2)(n-1)\left(n^3-5n^2+18n-12\right)\omega}{\alpha^6 n^3} \\
& + \frac{2\mu_3^2(n-1)\left(n^4-14n^3+56n^2-92n+64\right)(n-2)\omega^2}{\alpha^8 n^4} \\
& + \mu_4\left(\frac{(n-1)(n-2)^3\omega^2}{\alpha^6 n^3} + \frac{2(n-1)(n-2)^3\omega}{\alpha^6 n^3}\right) \\
& + \frac{2\mu_2\mu_4(n-1)\left(2n^4-17n^3+76n^2-140n+92\right)(n-2)\omega^2}{\alpha^8 n^4} \\
& + \frac{\mu_2^3(n-1)\left(2n^5-17n^4+105n^3-392n^2+724n-432\right)(n-2)\omega^2}{\alpha^8 n^4}
\end{aligned}
$$

Skew-symmetric entries, $\rho = -1$

Finally, we compute the expectation for all the terms when the matrix B is skew-symmetric, again choosing a different matrix M and a different parameter ω, in order to speed up the convergence. We have:

$$
\begin{aligned}
M &= \left(I + \frac{1}{\alpha}C^T\right)\left(I + \frac{1}{\alpha}C\right) \\
\mathbb{E}(b^T b) &= \frac{n}{\alpha^2} \\
\mathbb{E}(b^T y(0)) &= \frac{n}{\alpha^2} \\
\mathbb{E}(y(0)^T y(0)) &= \frac{n}{\alpha^2} + \frac{1}{\alpha^4}(n-1)n\mu 2
\end{aligned}
$$

We find:

$$
b^T M y(0) = \frac{n}{\alpha^2}
$$

$$
\mathbb{E}(y(0)^T M y(0)) = \frac{n}{\alpha^2} + \frac{(n-1)(n-2)(n-3)}{\alpha^6}\mu_2^2 - \frac{(n-1)(n-2)}{\alpha^6}\mu_4
$$

$$
\begin{aligned}
\mathbb{E}(y(0)^T M^T M y(0)) = &\frac{n}{\alpha^2}\frac{(n-1)(n-2)(n-3)}{\alpha^6}\mu_2^2 \\
&+ \frac{(n-1)(n-2)(n^3 - 15n^2 + 45n - 60)}{n\alpha^8}\mu_2^3 \\
&+ \frac{2(n-1)(n-2)(n^2 - 8n + 14)}{3n\alpha^8}\mu_3^2 \\
&+ \frac{(n-1)(n-2)}{\alpha^6}\mu_4 \\
&+ \frac{2(n-1)(n-2)(2n^2 - 10n + 15)}{n\alpha^8}\mu_4\mu_2 \\
&+ \frac{(n-1)(n-2)^2}{n\alpha^8}\mu_6
\end{aligned}
$$

Putting all these terms together, we find that the expected variance for $\rho = -1$ is:

$$
\begin{aligned}
\mathbb{E}(\sigma_{y^\star}^2) \approx &\frac{2\mu_3^2(n-2)(n-1)((n-8)n+14)\omega^2}{3\alpha^8 n^2} \\
&+ \frac{\mu_6(n-2)^2(n-1)\omega^2}{\alpha^8 n^2} + \frac{\mu_4(n-2)(n-1)\omega\left(30\mu_2\omega + 4\mu_2 n^2\omega + \alpha^2 n(\omega-2)\right)}{\alpha^8 n^2} \\
&+ \frac{\mu_2(n-1)\left(\alpha^4 n^2 + \mu_2(n-2)\omega^2\left(\alpha^2(n-3)n + \mu_2(n(n(2n-15)+45)-60)\right)\right)}{\alpha^8 n^2} \\
&- \frac{2\mu_2^2\alpha^2(n-3)(n-2)(n-1)\omega}{\alpha^8 n} - \frac{20\mu_2\mu_4(n-2)(n-1)\omega^2}{\alpha^8 n}
\end{aligned}
$$

where $\omega = \frac{\alpha^2}{\alpha^2+2n}$.

B.2 Using resolvents

Dyson equation derivation.

In this section we present the derivation of the Dyson equation (8) for symmetric random matrices. We start from the resolvent of matrix B,

$$H(z) = (zI + B)^{-1},$$

i.e., the inverse of matrix $zI + B$. The Neumann series expansion for $H(z)$ is convergent for any $z \in \mathbb{C}$ such that $\mathcal{I}m(z) > 0$. Recall that the expectation of the resolvent is denoted as $G(z) = \mathbb{E}(H(z))$. The equation above can be rewritten as

$$zH(z) - I = -BH(z).$$

We subtract, from both sides, the matrix product $\Sigma H(z)$, with $\Sigma = \mathbb{E}(B\mathbb{E}(H(z))B) = \mathbb{E}(BG(z)B)$. Thus, we get

$$-I + (zI - \Sigma)H(z) = -BH(z) - \Sigma H(z).$$

Taking expectations on both sides of this equation, we obtain

$$\mathbb{E}(-I + (zI - \Sigma)H(z)) = -\mathbb{E}(BH(z) + \Sigma H(z)).$$

Because Σ is an expectation, this expression reduces to

$$-I + (zI - \Sigma)G(z) = -\mathbb{E}(BH(z) + \Sigma H(z)).$$

It can be shown in general that the right-hand side of this equation is negligible up to leading order, as $n \to \infty$, for arbitrary distributions [50]. Hence the resolvent satisfies the matrix equation

$$(zI - \mathbb{E}(BG(z)B))G(z) = I.$$

After elementary manipulations, this equation can be rewritten in the usual form, see Eq. (8).

Proof of Eq. (10)***.***

Let $K(z) = \mathbb{E}((zI+B)^{-2})$. Here we show that $K(z) = -G'(z)$, where $G(z)$ is the expectation of the resolvent, $G(z) = \mathbb{E}((zI + B)^{-1})$, and the prime stands for the derivative with respect to z.

To compute the derivative, let us introduce a real parameter $q \neq 0$ that we can use to take derivatives with respect to it. We define $H_q(z) = (zI + qB)^{-1}$ and $L_q(z) = (zI + qB)^{-2}$, so $H_1(z) = H(z)$, with $H(z) = (zI + B)^{-1}$, and $L_1(z) = L(z)$, with $L(z) = (zI + B)^{-2}$. Observe that $\mathbb{E}(H(z)) = G(z)$ and $\mathbb{E}(L(z)) = K(z)$. It holds that

$$H_q(z) = \left(q\left(\frac{z}{q}I + B\right)\right)^{-1} = \frac{H(z/q)}{q}.$$

On the other hand, because $\mathcal{I}m(z) > 0$, the Neumann series for $L_q(z)$ converges. We compute it using the Neumann series for $H_q(z)$:

$$H_q(z) = \left(z\left(I + \frac{q}{z}B\right)\right)^{-1} = \frac{1}{z}\sum_{k=0}^{\infty}\left(-\frac{q}{z}\right)^k B^k.$$

Hence

$$L_q(z) = \frac{1}{z^2}\sum_{k=0}^{\infty}\left(-\frac{q}{z}\right)^k B^k \sum_{l=0}^{\infty}\left(-\frac{q}{z}\right)^l B^l.$$

We write the product of the two series using the Cauchy product, i.e.,

$$L_q(z) = \frac{1}{z^2}\sum_{s=0}^{\infty} c_s(z,q),$$

with $c_s(z,q)$ defined as

$$c_s(z,q) = \sum_{j=0}^{s}\left(-\frac{q}{z}\right)^j B^j \left(-\frac{q}{z}\right)^{s-j} B^{s-j} = (s+1)\left(-\frac{q}{z}\right)^s B^s.$$

Therefore

$$L_q(z) = \frac{1}{z^2}\sum_{s=0}^{\infty}(s+1)\left(-\frac{q}{z}\right)^s B^s = \frac{1}{z}H_q(z) + \frac{1}{z^2}\sum_{s=1}^{\infty} s\left(-\frac{q}{z}\right)^s B^s.$$

The second term in the last equation can be written in terms of a derivative of $H_q(z)$ with respect to q:

$$\begin{aligned}\frac{\partial}{\partial q}H_q(z) &= \frac{1}{z}\frac{\partial}{\partial q}\sum_{k=0}^{\infty}\left(-\frac{q}{z}\right)^k B^k = -\frac{1}{z^2}\sum_{k=1}^{\infty} k\left(-\frac{q}{z}\right)^{k-1} B^k \\ &= -\frac{1}{z^2}\left(-\frac{z}{q}\right)\sum_{k=1}^{\infty} k\left(-\frac{q}{z}\right)^k B^k = \frac{1}{qz}\sum_{k=1}^{\infty} k\left(-\frac{q}{z}\right)^k B^k\end{aligned}$$

Therefore,

$$\sum_{s=1}^{\infty} s\left(-\frac{q}{z}\right)^s B^s = qz\frac{\partial}{\partial q}H_q(z),$$

and it holds that

$$L_q(z) = \frac{1}{z}H_q(z) + \frac{q}{z}\frac{\partial}{\partial q}H_q(z).$$

Now we use that $H_q(z) = H(z/q)/q$ to compute the partial derivative:

$$\frac{\partial}{\partial q}H_q(z) = -\frac{1}{q^2}H(z/q) - \frac{z}{q^3}H'(z/q).$$

Upon substitution of this result into the equation for $L_q(z)$, and using the expression of $H_q(z)$ in terms of $H(z)$, we find

$$L_q(z) = \frac{1}{qz} H(z/q) + \frac{q}{z}\left(-\frac{1}{q^2} H(z/q) - \frac{z}{q^3} H'(z/q)\right) = -\frac{1}{q^2} H'(z/q).$$

Making the substitution $q = 1$ we get $L(z) = -H'(z)$.

To conclude the proof, we take expectations on both sides of the equation,

$$K(z) = \mathbb{E}(L(z)) = -\mathbb{E}(H'(z)).$$

By differentiating the Neumann series, we obtain

$$H'(z) = -\frac{1}{z^2} \sum_{k=0}^{\infty} (k+1) \left(-\frac{1}{z}\right)^k B^k,$$

and taking expectations

$$\mathbb{E}(H'(z)) = -\frac{1}{z^2} \sum_{k=0}^{\infty} (k+1) \left(-\frac{1}{z}\right)^k \mathbb{E}(B^k).$$

On the other hand, by definition

$$G(z) = \mathbb{E}((zI + B)^{-1}) = \frac{1}{z} \sum_{k=0}^{\infty} \left(-\frac{1}{z}\right)^k \mathbb{E}(B^k).$$

By taking the derivative of $G(z)$ with respect to z we finally obtain

$$G'(z) = -\frac{1}{z^2} \sum_{k=0}^{\infty} (k+1) \left(-\frac{1}{z}\right)^k \mathbb{E}(B^k) = \mathbb{E}(H'(z))$$

and the result $K(z) = -G'(z)$ is proved.

References

[1] MacArthur, R. & Levins, R. The limiting similarity, convergence, and divergence of coexisting species. *The american naturalist* **101**, 377–385 (1967).

[2] Abrams, P. The theory of limiting similarity. *Annual review of ecology and systematics* **14**, 359–376 (1983).

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

[4] Chesson, P. General theory of competitive coexistence in spatially-varying environments. *Theoretical population biology* **58**, 211–237 (2000).

[5] Barabás, G., Michalska-Smith, M. J. & Allesina, S. Self-regulation and the stability of large ecological networks. *Nature ecology & evolution* **1**, 1870–1875 (2017).

[6] Barabás, G., D'Andrea, R. & Stump, S. M. Chesson's coexistence theory. *Ecological monographs* **88**, 277–303 (2018).

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

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

[9] Tang, S. & Allesina, S. Reactivity and stability of large ecosystems. *Frontiers in Ecology and Evolution* **2**, 21 (2014).

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

[11] Stone, L. *Some problems of community ecology: processes, patterns and species persistence in ecosystems.* Ph.D. thesis, Monash University (1988).

[12] Stone, L. The Google matrix controls the stability of structured ecological and biological networks. *Nature communications* **7**, 12857 (2016).

[13] Bizeul, P. & Najim, J. Positive solutions for large random linear systems. *Proceedings of the American Mathematical Society* **149**, 2333–2348 (2021).

[14] Clenet, M., El Ferchichi, H. & Najim, J. Equilibrium in a large Lotka–Volterra system with pairwise correlated interactions. *Stochastic Processes and their Applications* **153**, 423–444 (2022).

[15] Liu, X., Constable, G. W. & Pitchford, J. W. Feasibility and stability in large Lotka Volterra systems with interaction structure. *Physical Review E* **107**, 054301 (2023).

[16] Marcus, S., Turner, A. M. & Bunin, G. Extinctions as a vestige of instability: the geometry of stability and feasibility. *arXiv preprint arXiv:2405.11360* (2024).

[17] Smale, S. On the differential equations of species in competition. *Journal of Mathematical Biology* **3**, 5–7 (1976).

[18] Vano, J., Wildenberg, J., Anderson, M., Noel, J. & Sprott, J. Chaos in low-dimensional Lotka–Volterra models of competition. *Nonlinearity* **19**, 2391 (2006).

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

[20] Serván, C. A. & Allesina, S. Tractable models of ecological assembly. *Ecology Letters* **24**, 1029–1037 (2021).

[21] Lotka, A. J. Analytical note on certain rhythmic relations in organic systems. *Proceedings of the National Academy of Sciences* **6**, 410–415 (1920).

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

[23] Lotka, A. J. *Elements of physical biology* (Williams & Wilkins, 1925).

[24] Volterra, V. Leçons sur la théorie mathématique de la Lutte pour la Vie. *Gauthier-Villars, Paris* (1931).

[25] Gilpin, M. E. Stability of feasible predator-prey systems. *Nature* **254**, 137–139 (1975).

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

[27] Song, C., Rohr, R. P. & Saavedra, S. A guideline to study the feasibility domain of multi-trophic and changing ecological communities. *Journal of theoretical biology* **450**, 30–36 (2018).

[28] AlAdwani, M. & Saavedra, S. Feasibility conditions of ecological models: Unfolding links between model parameters. *Ecological Modelling* **466**, 109900 (2022).

[29] O'Rourke, S. & Renfrew, D. Low rank perturbations of large elliptic random matrices. *Electron. J. Probab* **19**, 1–65 (2014).

[30] Rozdilsky, I. D. & Stone, L. Complexity can enhance stability in competitive systems. *Ecology Letters* **4**, 397–400 (2001).

[31] MacArthur, R. Species packing and competitive equilibrium for many species. *Theoretical population biology* **1**, 1–11 (1970).

[32] Serván, C. A., Capitán, J. A., Miller, Z. R. & Allesina, S. Effects of phylogeny on coexistence in model communities. *The American Naturalist* (2024).

[33] Serván, C. A., Capitán, J. A., Grilli, J., Morrison, K. E. & Allesina, S. Coexistence of many species in random ecosystems. *Nature ecology & evolution* **2**, 1237–1242 (2018).

[34] Tilman, D. Resource competition between plankton algae: an experimental and theoretical approach. *Ecology* **58**, 338–348 (1977).

[35] Dybzinski, R. & Tilman, D. Resource use patterns predict long-term outcomes of plant competition for nutrients and light. *The American Naturalist* **170**, 305–318 (2007).

[36] Burson, A., Stomp, M., Mekkes, L. & Huisman, J. Stable coexistence of equivalent nutrient competitors through niche differentiation in the light spectrum. *Ecology* **100**, e02873 (2019).

[37] Bittleston, L. S., Gralka, M., Leventhal, G. E., Mizrahi, I. & Cordero, O. X. Context-dependent dynamics lead to the assembly of functionally distinct microbial communities. *Nature communications* **11**, 1440 (2020).

[38] Chang, C.-Y., Bajić, D., Vila, J. C., Estrela, S. & Sanchez, A. Emergent coexistence in multispecies microbial communities. *Science* **381**, 343–348 (2023).

[39] Fussmann, G. F., Ellner, S. P., Shertzer, K. W. & Hairston Jr, N. G. Crossing the Hopf bifurcation in a live predator-prey system. *Science* **290**, 1358–1360 (2000).

[40] Yoshida, T., Jones, L. E., Ellner, S. P., Fussmann, G. F. & Hairston Jr, N. G. Rapid evolution drives ecological dynamics in a predator–prey system. *Nature* **424**, 303–306 (2003).

[41] Becks, L., Hilker, F. M., Malchow, H., Jürgens, K. & Arndt, H. Experimental demonstration of chaos in a microbial food web. *Nature* **435**, 1226–1229 (2005).

[42] Becks, L. & Arndt, H. Transitions from stable equilibria to chaos, and back, in an experimental food web. *Ecology* **89**, 3222–3226 (2008).

[43] Benincà, E. *et al.* Chaos in a long-term experiment with a plankton community. *Nature* **451**, 822–825 (2008).

[44] Blasius, B., Rudolf, L., Weithoff, G., Gaedke, U. & Fussmann, G. F. Long-term cyclic persistence in an experimental predator–prey system. *Nature* **577**, 226–230 (2020).

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

[46] Arnoldi, W. E. The principle of minimized iterations in the solution of the matrix eigenvalue problem. *Quarterly of applied mathematics* **9**, 17–29 (1951).

[47] Nguyen, H. H. & O'Rourke, S. The elliptic law. *International Mathematics Research Notices* **2015**, 7620–7689 (2015).

[48] Hall, P. On the rate of convergence of normal extremes. *Journal of Applied Probability* **16**, 433–439 (1979).

[49] Genz, A. & Kwong, K.-S. Numerical evaluation of singular multivariate normal distributions. *Journal of Statistical Computation and Simulation* **68**, 1–21 (2000).

[50] Erdős, L., Krüger, T. & Schröder, D. Random matrices with slow correlation decay. *Forum of Mathematics, Sigma* **7**, e8 (2019).

[51] Girko, V. Elliptic law. *Theory of Probability & Its Applications* **30**, 677–690 (1986).

[52] Alt, J. & Krüger, T. Local elliptic law. *Bernoulli* **28**, 886–909 (2022).