Diverse communities behave like typical random ecosystems

Wenping Cui*
Department of Physics, Boston University, 590 Commonwealth Avenue, Boston, MA 02139 and Department of Physics, Boston College, 140 Commonwealth Ave, Chestnut Hill, MA 02467

Robert Marsland III† and Pankaj Mehta‡
Department of Physics, Boston University, 590 Commonwealth Avenue, Boston, MA 02139
(Dated: April 1, 2019)

With a brief letter to *Nature* in 1972, Robert May triggered a worldwide research program in theoretical ecology and complex systems that continues to this day[1]. Building on powerful mathematical results about large random matrices, he argued that systems with sufficiently large numbers of interacting components are generically unstable. In the ecological context, May's thesis directly contradicted the long-standing ecological intuition that diversity promotes stability[2–4]. In economics and finance, May's work helped to consolidate growing concerns about the fragility of an increasingly interconnected global marketplace[5–7]. In this Letter, we draw on recent theoretical progress in random matrix theory and statistical physics to fundamentally extend and reinterpret May's theorem. We confirm that a wide range of ecological models become unstable at the point predicted by May, even when the models do not strictly follow his assumptions. Surprisingly, increasing the interaction strength or diversity beyond the May threshold results in a reorganization of the ecosystem – through extinction of a fixed fraction of species – into a new stable state whose properties are well described by purely random interactions. This self-organized state remains stable for arbitrarily large ecosystem and suggests a new interpretation of May's original conclusions: when interacting complex systems with many components become sufficiently large, they will generically undergo a transition to a "typical" self-organized, stable state.

For an ecosystem of S species, May's theorem concerns the $S \times S$ community matrix $\mathbf{J}$, whose entries J_{ij} describe how much the growth rate of species i is affected by a small change in the population N_j of species j from its equilibrium value $\bar{N}_j$[1]. The stability of this equilibrium can be quantified in terms of the largest eigenvalue $\lambda_{\max}$ of $\mathbf{J}$. If $\lambda_{\max}$ is positive, the equilibrium is unstable, and a small perturbation will cause the system to flow away from the equilibrium state. In the 1950's, Eugene Wigner derived a mathematical formula for the distribution of eigenvalues in a special class of large random matrices[8]. May pointed out that the resulting estimate of the maximum eigenvalue $\lambda_{\max}$ is actually more general, and applies whenever the J_{ij} are sampled independently from probability distributions with finite means μ and variances σ^2[1]. With this result in mind, May considered a simple ecosystem where each species inhibits itself, with $J_{ii} = -1$, but different species initially do not interact with each other. This ecosystem is guaranteed to be stable for any level of diversity. He then examined how the stability is affected by adding randomly sampled interactions, and found that $\lambda_{\max}$ typically becomes positive when the root-mean-squared total strength $\sigma_M = \sqrt{S\sigma^2}$ of inter-specific interactions reaches parity with the intra-specific interactions, that is, when $\sigma_M = 1$. For a given pairwise interaction strength σ, this relation gives the maximum diversity S compatible with ecosystem stability.

This result has proven to be robust against a wide array of changes in the assumptions, including adding biologically realistic correlation structures to the matrix, or incorporating the dependence of the community matrix on population sizes in the Lotka-Volterra model [9, 10]. In order to be stable at high diversities, a generic ecosystem must have a fine-tuned interaction structure, which can sometimes be justified in terms of biological constraints[11–14]. Very recently, it was noted that the required level of structure may emerge spontaneously in simple models of population dynamics, typically through extinction of some fraction of the initial set of species [15]. Surprisingly, the resulting self-organized states were relatively insensitive to the network structure encoded in the interaction matrix. This numerical work suggests that May's result might admit of a more intuitive interpretation, and that the high-diversity states likely possess interesting generic properties.

To explore these ideas, we devised a more concrete version of May's original thought experiment describing an ecosystem consisting of S non-interacting species where interactions are gradually turned on. May's original argument only considered the local dynamics near a prespecified equilibrium point that eventually becomes unstable. Since we are interested in exploring what happens after the onset of this instability, we must make additional modeling assumptions to arrive at a complete set of nonlinear dynamics. We focus on MacArthur's Consumer

* cuiw@bc.edu
† marsland@bu.edu
‡ pankajm@bu.edu

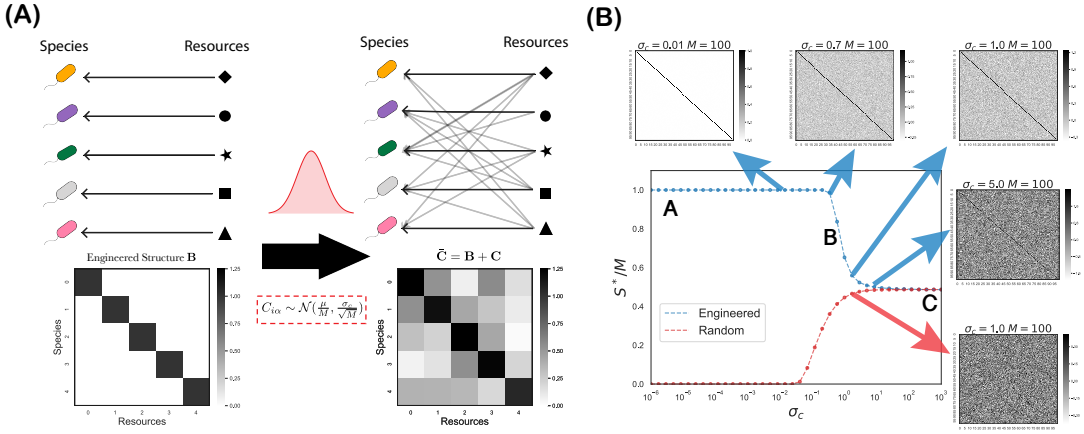

FIG. 1. **Random interactions destabilize an ecosystem of specialist consumers.** **(A)** Left: an ecosystem with system size $M = 5$ starts with specialists consuming only one type of resource, resulting in a consumer preference matrix $\mathbf{B} = \mathbb{1}$. Right: off-target consumption coefficients $\mathbf{C} \sim \mathcal{N}(\frac{\mu}{M}, \frac{\sigma_c}{\sqrt{M}})$ are sampled from a Gaussian distribution, resulting in an overall consumer preference matrix $\bar{\mathbf{C}} = \mathbf{B} + \mathbf{C}$. **(B)** Fraction of surviving species S^*/M vs. σ_c, numerically computed using $M = 100$ as described in the Methods, along with the corresponding results for a completely random ecosystem with $\mathbf{B} = 0$. Also shown are examples of the matrices $\bar{\mathbf{C}}$ employed in the simulations.

Resource Model (CRM) where interactions are mediated by competition for M substitutable resources[16]. For simplicity, we assume that $S = M$ (though this does not affect our main results). MacArthur's model allows for a wide range of exact mathematical results, but assumes that the resources are themselves self-replicating entities. To check the generality of our results, we also numerically analyzed generalizations of the CRM including the case of constant flux of externally supplied resources, and a model of microbial ecology with trophic feedbacks where organisms can feed each other via metabolic byproducts[17, 18]. These extended results can be found in the Supplemental Information.

The identity of each species in these models is determined by its consumption preferences. A set of non-interacting species can be constructed by engineering each species to consume a different resource type, with no overlap between consumption preferences. One can imagine designing strains of *E. coli* where one strain only expresses transporters for lactose, and another only expresses transporters for sucrose, etc., with all other transporters edited out of the genome, as illustrated in Figure 1(A). In such an experiment, horizontal gene transfer would eventually begin distributing transporter genes from one strain to another, so a realistic model would have to allow for some amount of unintended resource consumption. The resulting preference $\bar{C}_{i\alpha}$ of species i for resource α is the sum of the identity matrix $\mathbb{1}$ and a random component $C_{i\alpha}$ with variance σ^2.

Figure 1(B) shows the results of adding non-specific interactions to the CRM. Just as in May's analysis, the appropriate measure of the importance of the random component is the root-mean-squared off-target consumption $\sigma_c = \sqrt{M\sigma^2}$ (recall $M = S$). Since we are expecting some species to go extinct in the self-organization process, we start by plotting the fraction of surviving species S^*/M versus σ_c in the steady state of numerical simulations with $S = 100$ species. At small values of σ_c, all the species survive, as expected for a weakly interacting ecosystem. As high as $\sigma_c = 0.7$, almost all of the original species are still present in the community. But between $\sigma_c = 0.7$ and $\sigma_c = 1$, there is a sharp transi-

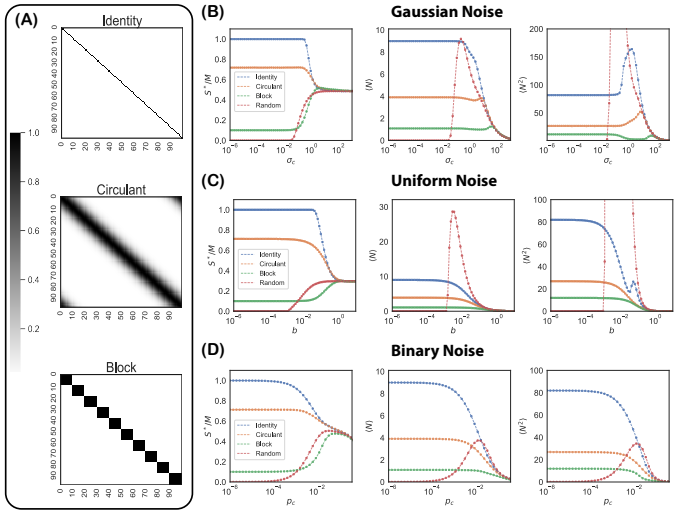

FIG. 2. **Community properties for structured and random ecosystems.** **(A)**: Examples of designed interactions $\mathbf{B}$. Top: an identity matrix; Middle: a Gaussian-type circulant matrix; Bottom: a block matrix. **(B)** Gaussian Noise $\mathcal{N}(0, \frac{\sigma_c}{\sqrt{M}})$, **(C)** Uniform Noise: $\mathcal{U}(0, b)$ and **(D)**: Binary Noise: $Bernoulli(p_c)$ are added to the engineered interactions. Community properties including survival fraction S^*/M, mean abundance $\langle N \rangle$ and mean square abundance $\langle N^2 \rangle$ are shown for the different $\mathbf{B}$ listed above.

tion in community structure, which results in about half of the original species becoming extinct. Remarkably, the survival fraction converges to the same value as for a completely random consumer preference matrix, and remains finite as $\sigma_c \to \infty$. This means that arbitrarily high amounts of diversity can be maintained at a given amount of uncertainty σ in the model parameters by considering a sufficiently large starting ecosystem. These numerical predictions are in excellent agreement with analytic predictions for the limit $S \to \infty$ derived in the Methods.

We proceeded to investigate the properties of these self-organized high-diversity states more closely. In addition to the number of surviving species, we considered two other community-level properties: the mean population size $\langle N \rangle$ over all species, and the second moment of the population size $\langle N^2 \rangle$, which includes information about the range of population sizes. Figure 2 shows that both of these quantities are also well-approximated by the random consumer preference matrix for $\sigma_c > 1$. This convergence to random ecosystem behavior is quite robust, and remains present when a set of designed interactions is added to the original noninteracting community before adding the random component. Figure 2 shows the results for two basic interaction structures: a block structure with pre-defined groups of species exhibiting strong intra-group competition, and a unimodal structure where each species is more likely to consume resources similar to its preferred resource. The only effect of the choice of structure is to adjust the threshold value of σ_c where the transition takes place. The character of the self-organized state is also robust to changes in the sampling scheme for the random component. Gaussian sampling allows the clearest comparison to May's result, but it produces some negative values, while consumer preferences should always be positive. We therefore tested two sampling schemes that always produce positive values: uniform sampling in the interval from 0 to b, and binary sampling with probability p of choosing 1. Changing b or p affects both the mean and the variance of the random component, and so the transition point cannot be directly compared to the Gaussian case. But we still find that the large-scale properties become similar to those of the random ecosystem when the average total off-target consumption capacity over all M resource types becomes greater than the original designed capacity.

Finally, we examined the stability of the ecosystems. For comparison with May's results, we obtained effective competition coefficients A_{ij} between species in a generalized Lotka-Volterra model, by assuming that the resource abundances always remain close to their steady state values, as illustrated in Figure 3(A). This matrix is related to May's community matrix by $J_{ij} = -\bar{N}_i A_{ij}$. For the symmetric interaction matrices arising from the consumer resource model, one can prove that the largest eigenvalue $\lambda_{\max}$ of $\mathbf{J}$ reaches zero from below only when the smallest eigenvalue $\lambda_{\min}$ of $\mathbf{A}$ reaches zero from above (see Methods). Figure 3(B) shows how the eigenvalues

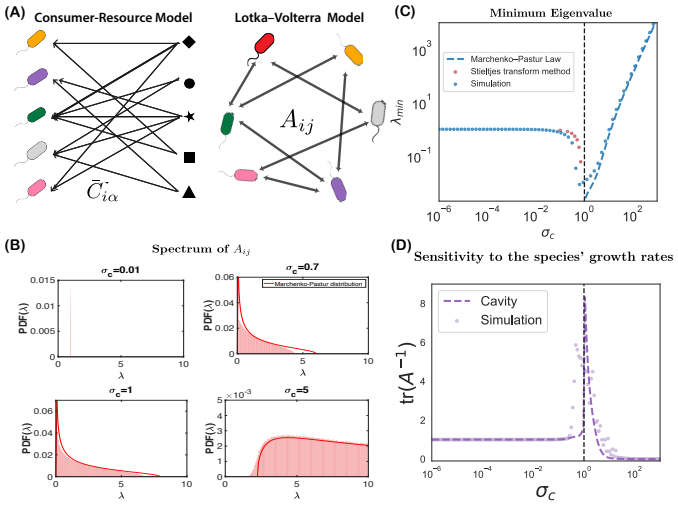

FIG. 3. **Effect of random interactions on ecosystem stability**. **(A)**: The bipartite interactions $\bar{C}_{i\alpha}$ in MacArthur's consumer-resource model can be mapped to pairwise competition coefficients A_{ij} in generalized Lotka-Volterra equations through $A_{ij} = \sum_{\alpha \in \mathbf{M}^*} \bar{C}_{i\alpha} \bar{C}^T_{\alpha j}$. **(B)** Spectra of A_{ij} at different σ_c for $\mathbf{B} = \mathbb{1}$. The red solid line is the Marchenko-Pastur distribution. **(C)**: Comparison between numerical simulations and analytic results for the minimum eigenvalue of $\mathbf{A}$ at different σ_c. **(D)**: Comparison between numerical simulations and analytic solutions for the mean sensitivity ν of steady-state population sizes to changes in species' growth rates. See Methods for detailed calculations.

of $\mathbf{A}$ change as σ_c increases. Initially, all the eigenvalues of the noninteracting community are identical, but with increasing σ_c the distribution spreads out, and reaches the threshold of stability $\lambda_{\min} = 0$ at $\sigma_c \approx 1$. In the Appendix we show analytically using the Cavity method [19–21] that the transition point approaches 1 as $S \to \infty$, in agreement with expectations based on May's analysis.

But for $\sigma_c > 1$, we observe two new phenomena that were not accessible in May's original framework. First, the spectrum is well-described by the eigenvalue distribution for interactions resulting from completely random consumer preference matrices, known as the Marchenko-Pastur Distribution[22]. This is consistent with our earlier observations on community-level observables that all coincided with the random case for sufficiently large σ_c. Secondly, the minimum eigenvalue in the Marchenko-Pastur Distribution is located at $\lambda_{\min} = \sigma_c^2(1 - \sqrt{S^*/M^*})^2$, where M^* is the number of resource types that persist at nonzero abundance in the steady state. As we saw earlier, about half the species go extinct when $\sigma_c > 1$, leading to $S^*/M^* < 1$, so that $\lambda_{\min}$ is always positive. This means that the random ecosystem is stable, for arbitrarily large values of σ_c.

The spectrum of $\mathbf{A}$ also contains quantitative information about the degree of ecosystem stability. Specifically, as shown in the Methods, the sum of the inverse eigenvalues $\sum_i (1/\lambda_i) = \mathrm{tr}(\mathbf{A}^{-1})$ measures the average response of the steady-state population size $\bar{N}_i$ to a given pertur-

bation of the species' growth rate. Figure 3(D) shows that this quantity is initially constant as σ_c is increased from 0, then diverges at $\sigma_c = 1$, and finally rapidly decreases to near zero. In the Methods we provide analytical calculations based on the cavity method confirming that this divergence is a signature of a continuous phase transition, analogous to the increased sensitivity near ecological tipping points known to result in large stochastic fluctuations [23, 24].

The foregoing analysis leads to a reinterpretation of May's theorem as a bound on the feasibility of bottom-up engineering in complex systems. As the number of components increases, small uncertainties in each of the interaction parameters eventually overwhelm the designed interactions, and destabilize the intended system state. But the system generically finds a new stable state, which may be even more stable than the originally engineered one. The cumulative parameter uncertainties preclude a priori prediction of the detailed structure of the new state, but methods from statistical physics and Random Matrix Theory can be employed to predict system-level properties. Further development of these methods and their applications will play an important role in enabling top-down control of systems beyond the May bound and helping to identify assembly rules for microbial communities [25]

ACKNOWLEDGMENTS

We thank Zhenyu Liao, Guangwei Si, Jean Vila and Yu Hu for helpful discussions. The work was supported by NIH NIGMS grant 1R35GM119461, Simons Investigator in the Mathematical Modeling of Living Systems (MMLS). The authors are pleased to acknowledge that the computational work reported on in this paper was performed on the Shared Computing Cluster which is administered by Boston University Research Computing Services.

[1] R. M. May, Nature **238**, 413 (1972).
[2] R. MacArthur, Ecology **36**, 533 (1955).
[3] D. Tilman and J. A. Downing, Nature **367**, 363 (1994).
[4] K. S. McCann, Nature **405**, 228 (2000).
[5] A. G. Haldane and R. M. May, Nature **469**, 351 (2011).
[6] P. Gai and S. Kapadia, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **466**, 2401 (2010).
[7] P. Glasserman and H. P. Young, Journal of Banking & Finance **50**, 383 (2015).
[8] E. P. Wigner, in *The Collected Works of Eugene Paul Wigner* (Springer, 1993) pp. 524–540.
[9] T. Gross, L. Rudolf, S. A. Levin, and U. Dieckmann, Science **325**, 747 (2009).
[10] S. Allesina and S. Tang, Nature **483**, 205 (2012).
[11] P. Yodzis, Nature **289**, 674 (1981).
[12] M. C. Emmerson and D. Raffaelli, Journal of Animal Ecology **73**, 399 (2004).
[13] S. Tang, S. Pawar, and S. Allesina, Ecology letters **17**, 1094 (2014).
[14] J. Grilli, M. Adorisio, S. Suweis, G. Barabás, J. R. Banavar, S. Allesina, and A. Maritan, Nature communications **8**, 14389 (2017).
[15] C. A. Serván, J. A. Capitán, J. Grilli, K. E. Morrison, and S. Allesina, Nature Ecology & Evolution **2**, 1237 (2018).
[16] R. MacArthur and R. Levins, The American Naturalist **101**, 377 (1967).
[17] J. E. Goldford, N. Lu, D. Bajić, S. Estrela, M. Tikhonov, A. Sanchez-Gorostiaga, D. Segrè, P. Mehta, and A. Sanchez, Science **361**, 469 (2018).
[18] R. Marsland III, W. Cui, J. Goldford, A. Sanchez, K. Korolev, and P. Mehta, PLoS computational biology **15**, e1006793 (2019).
[19] G. Bunin, Physical Review E **95**, 042414 (2017).
[20] M. Advani, G. Bunin, and P. Mehta, Journal of Statistical Mechanics: Theory and Experiment **2018**, 033406 (2018).
[21] P. Mehta, W. Cui, C.-H. Wang, and R. Marsland III, arXiv preprint arXiv:1809.04221 (2018).
[22] V. A. Marchenko and L. A. Pastur, Matematicheskii Sbornik **114**, 507 (1967).
[23] M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. Van Nes, M. Rietkerk, and G. Sugihara, Nature **461**, 53 (2009).
[24] L. Dai, D. Vorselen, K. S. Korolev, and J. Gore, Science **336**, 1175 (2012).
[25] J. Friedman, L. M. Higgins, and J. Gore, Nature ecology & evolution **1**, 0109 (2017).
[26] S. Butler and J. P. O'Dwyer, Nature Communications **9**, 2970 (2018).
[27] A. Altieri and S. Franz, Physical Review E **99**, 010401 (2019).
[28] E. Agliari, F. Alemanno, A. Barra, and A. Fachechi, arXiv preprint arXiv:1811.08298 (2018).
[29] R. B. Dozier and J. W. Silverstein, Journal of Multivariate Analysis **98**, 678 (2007).
[30] R. Couillet and M. Debbah, *Random matrix methods for wireless communications* (Cambridge University Press, 2011).

METHODS

A. Model

In this work, we will focus primarily on MacArthur's consumer-resource model(MCRM)[16]. This model consists of S species or consumers with abundances N_i $(i = 1...S)$ that can consume one of M substitutable resources with abundances R_α $(\alpha = 1...M)$

$$\begin{cases} \frac{dN_i}{dt} = N_i(\sum_\beta \bar{C}_{i\beta} R_\beta - m_i) \\ \frac{dR_\alpha}{dt} = R_\alpha(K_\alpha - R_\alpha - \sum_j N_j \bar{C}_{j\alpha}) \end{cases} \tag{1}$$

The consumption rate of species i for resource α is encoded by the element $\bar{C}_{i\alpha}$ in the $S \times M$ matrix $\bar{\mathbf{C}}$. K_α is the carrying capacity of each resource α. m_i is some minimum maintenance cost that species i must harvest from resources in order to grow. Note, both the species and resource abundances N_i and R_α must be strictly non-negative. When the system is in the steady state, some species and resources can vanish. We denote the numbers of surviving species and resources by S^* and M^*, respectively.

To construct our mechanistic version of May's analysis, we decompose the consumer matrix $\bar{\mathbf{C}}$ into two parts: $\bar{\mathbf{C}} = \mathbf{B} + \mathbf{C}$ with $\mathbf{B}$ encoding a pre-designed set of resource-mediated interactions, and $\mathbf{C}$ a random matrix encoding "off-target" consumption. We consider three types of $\mathbf{B}$ (see Figure 2): the identity matrix, a square Gaussian-type circulant matrix $B_{i\alpha} = e^{-\min(i,|M-i|)^2/r^2}$ with $r = 7$[26] and a block matrix with identical 10×10 blocks(all elements are 1 inside the 10×10 block). We also consider three types of random matrices $\mathbf{C}$. In all cases, each element in the matrix is sampled independently from an underlying probability distribution. The three distributions we consider are a normal distribution with mean zero and standard deviation $\sigma_c/\sqrt{M}$, a uniform distribution where each element is sampled uniformly from $[0, b]$, and a Bernoulli distribution where each element can be $+1$ with probability p_c and 0 with probability $1 - p_c$ (i.e Binary Noise).

For all simulations, unless otherwise specified the default choices for parameters are: $M = 100$, $\mu = 0$, $K = 10$, $\sigma_K = 1$, $m = 1$ and $\sigma_m = 0.1$ and each data point is averaged from 4000 independent realizations. For Figure 3 (C, D), each data point is averaged from 8000 independent realizations. All simulations are available on GitHub at https://github.com/Emergent-Behaviors-in-Biology/typical-random-ecosystems.

1. *Alternative Models used in SI*

To test the generality of our results, we also simulated more complicated variants of the consumer resource model (see Figure S1). First, we simulated a consumer resource model with linear resource dynamics. In this case, the second equation in (1) is replaced by

$$\frac{dR_\alpha}{dt} = K_\alpha - R_\alpha - \sum_i N_i C_{i\alpha} R_\alpha. \tag{2}$$

This small change can significantly change the ecosystem properties, because it prevents resources from going extinct in the steady state. In the simulations, we set $M = 100$, $\mu = 1$, $K = 10$, $\sigma_K = 1$, $m = 1$ and $\sigma_m = 0.1$ and each data point is averaged from 4000 independent realizations.

Second, we simulated a generalization of the MacArthur's Consumer Resource model we call the Microbial Consumer Resource Model (MicroCRM). The MicroCRM was introduced in [17] and refined in [18] to simulate microbial communities. In this model, in addition to consuming resources species can produce new resources through cross-feeding. This dramatically changes the resource dynamics through the introduction of trophic feedbacks. Unlike the original CRM and the extensio to linear resource dynamics, the MicroCRM possesses no Lyapunov function. Full details of the model are available in the appendix of [18]. In particular, the dynamics we use are described in equation (17)[18] with the leakage rate $l = 0.4$. The fraction of secretion flux secreted to the same resource type is $f_s = 0.45$, the fraction of secretion flux to 'waste' resource is $f_w = 0.45$ and variability in secretion fluxes among resources is $d_0 = 0.2$. We set $M = 100$, $\mu = 1$, $K = 10$, $\sigma_K = 1$, $m = 1$ and $\sigma_m = 0.1$ and each data point is averaged from 4000 independent realizations.

B. Sensitivity to Parameter Perturbations

We begin by defining four susceptibility matrices that measure how the steady-state resource and species abundances respond to changes in the resource supply and species death(growth) rates:

$$\chi^R_{\alpha\beta} = \frac{\partial \bar{R}_\alpha}{\partial K_\beta}, \chi^N_{i\alpha} = \frac{\partial \bar{N}_i}{\partial K_\alpha}, \nu^R_{\alpha i} = \frac{\partial \bar{R}_\alpha}{\partial m_i}, \nu^N_{ij} = \frac{\partial \bar{N}_i}{\partial m_j} \tag{3}$$

where the bar $\bar{X}$ over the variable X denotes the steady-state (equilibrium) solution.

For the extinct species and resources, by definition the susceptibilities are zero. For this reason, we focus only on the surviving resources and species. At steady-state, equation (1) gives:

$$0 = \sum_{\alpha \in \mathbf{M}^*} \bar{C}_{i\alpha} \bar{R}_\alpha - m_i \tag{4}$$

$$0 = K_\alpha - \bar{R}_\alpha - \sum_{j \in \mathbf{S}^*} \bar{N}_j \bar{C}_{j\alpha} \tag{5}$$

where $\mathbf{M}^*$ and $\mathbf{S}^*$ denote the sets of resources and species, respectively, that survive in the ecosystem at steady-state. Differentiating these equations yields the relations

$$\begin{aligned} 0 &= \sum_{\alpha \in \mathbf{M}^*} \bar{C}_{i\alpha} \frac{\partial \bar{R}_\alpha}{\partial K_\beta}, \quad \delta_{\alpha\beta} = \frac{\partial \bar{R}_\alpha}{\partial K_\beta} + \sum_{j \in \mathbf{S}^*} \frac{\partial \bar{N}_j}{\partial K_\beta} \bar{C}_{j\alpha} \\ \delta_{ij} &= \sum_{\alpha \in \mathbf{M}^*} \bar{C}_{i\alpha} \frac{\partial \bar{R}_\alpha}{\partial m_j}, \quad 0 = \frac{\partial \bar{R}_\alpha}{\partial m_i} + \sum_{j \in \mathbf{S}^*} \frac{\partial \bar{N}_j}{\partial m_i} \bar{C}_{j\alpha}. \end{aligned} \tag{6}$$

Substituting in for the partial derivatives using the susceptibility matrices defined above, we have:

$$\begin{aligned} 0 &= \sum_{\alpha \in \mathbf{M}^*} \bar{C}_{i\alpha} \chi^R_{\alpha\beta}, \quad \delta_{\alpha\beta} = \chi^R_{\alpha\beta} + \sum_{j \in \mathbf{S}^*} \chi^N_{j\beta} \bar{C}_{j\alpha} \\ \delta_{ij} &= \sum_{\alpha \in \mathbf{M}^*} \bar{C}_{i\alpha} \nu^R_{\alpha j}, \quad 0 = \nu^R_{\alpha i} + \sum_{j \in \mathbf{S}^*} \nu^N_{ji} \bar{C}_{j\alpha}. \end{aligned} \tag{7}$$

These two equations can be written as single matrix equation for block matrices:

$$\begin{pmatrix} \bar{\mathbf{C}} & 0 \\ \mathbb{1} & \bar{\mathbf{C}}^T \end{pmatrix} \begin{pmatrix} \nu^R & \chi^R \\ \nu^N & \chi^N \end{pmatrix} = \mathbb{1} \tag{8}$$

To solve this equation, we define a $S^* \times S^*$ matrix: $A_{ij} = \sum_{\alpha \in M^*} \bar{C}_{i\alpha} \bar{C}^T_{\alpha j}$. A straightforward calculation yields

$$\chi^R_{\alpha\beta} = \delta_{\alpha\beta} - \sum_{i \in \mathbf{S}^*} \sum_{j \in \mathbf{S}^*} \bar{C}^T_{\alpha i} A^{-1}_{ij} \bar{C}_{j\beta} \tag{9}$$

$$\chi^N_{i\alpha} = \sum_{j \in \mathbf{S}^*} A^{-1}_{ij} \bar{C}_{j\beta}, \quad \nu^R_{\alpha i} = \sum_{j \in \mathbf{S}^*} \bar{C}^T_{\alpha j} A^{-1}_{ji} \tag{10}$$

$$\nu^N_{ij} = -A^{-1}_{ij}, \quad i, j \in \mathbf{S}^* \text{ and } \alpha, \beta \in \mathbf{M}^* \tag{11}$$

C. Cavity Solution

For an initially non-interacting ecosystem $\mathbf{B} = \mathbb{1}$, the effect of random off-target consumption on system-scale properties can be computed analytically in the $M, S \to \infty$ limit using the cavity method [19–21]. The cavity calculation is straightforward but tedious. For this reason, it is helpful to introduce the notation:

- $\frac{M^*}{M} = \phi_R$, $\langle R \rangle = \frac{1}{M} \sum_\beta R_\beta$ and $q_R = \frac{1}{M} \sum_\beta R^2_\beta = \langle R^2 \rangle$, where M^* is the number of surviving resources.
- $\frac{S^*}{S} = \phi_N$, $\langle N \rangle = \frac{1}{S} \sum_j N_j$ and $q_N = \frac{1}{S} \sum_j N^2_j = \langle N^2 \rangle$, where S^* is the number of surviving species.
- $C_{i\alpha} \equiv \frac{\mu}{M} + \sigma_c d_{i\alpha}$ assuming $\langle d_{i\alpha} \rangle = 0$, $\langle d_{i\alpha} d_{j\beta} \rangle = \frac{\delta_{ij}\delta_{\alpha\beta}}{M}$. with $\langle c_{i\alpha} \rangle = \frac{\mu}{M}$, $\langle c_{i\alpha} c_{j\beta} \rangle = \frac{\sigma_c^2}{M} \delta_{ij}\delta_{\alpha\beta} + \frac{\mu^2}{M^2} \approx \frac{\sigma_c^2}{M} \delta_{ij}\delta_{\alpha\beta}$.

- $K_\alpha = K + \delta K_\alpha$ with $\langle K_\alpha \rangle = \frac{1}{M}\sum_\beta K_\beta = K$, $\langle \delta K_\alpha \delta K_\beta \rangle = \delta_{\alpha\beta}\sigma_K^2$.
- $m_i = m + \delta m_i$ with $\langle m_i \rangle = m$, $\langle \delta m_i \delta m_j \rangle = \delta_{ij}\sigma_m^2$.
- $\gamma = \frac{M}{S}$ and for the identity matrix $\gamma = 1$.

Following similar steps as in [20, 21], we perturb the ecosystem with a new species and resource N_0 and R_0. Ignoring $\mathcal{O}(1/M)$ terms yields the following equations:

$$\frac{dN_i}{dt} = N_i \left[R_i - m + \sum_\beta (\frac{\mu}{M} + \sigma_c d_{i\beta}) R_\beta + (\frac{\mu}{M} + \sigma_c d_{i0}) R_0 - \delta m_i \right] \tag{12}$$

$$\frac{dR_\alpha}{dt} = R_\alpha \left[K + \delta K_\alpha - R_\alpha - N_\alpha - \sum_j (\frac{\mu}{M} + \sigma_c d_{j\alpha}) N_j - (\frac{\mu}{M} + \sigma_c d_{0\alpha}) N_0 \right] \tag{13}$$

$$\frac{dN_0}{dt} = N_0 \left[R_0 - m + \sum_\beta (\frac{\mu}{M} + \sigma_c d_{j\alpha}) R_\beta - \delta m_0 \right] \tag{14}$$

$$\frac{dR_0}{dt} = R_0 \left[K + \delta K_0 - R_0 - N_0 - \sum_j (\frac{\mu}{S} + \sigma_c d_{j0}) N_j \right] \tag{15}$$

Denote by $\bar{N}_{\alpha/0}$, $\bar{R}_{\alpha/0}$ and $\bar{N}_i$, $\bar{R}_\alpha$ the equilibrium values of the species and resources before and after adding the newcomers, respectively. These can be related to each other using the susceptibilities defined above:

$$\bar{N}_i = \bar{N}_{i/0} - \sigma_c \sum_j \nu_{ij}^N d_{j0} R_0 - \sigma_c \sum_\beta \chi_{i\beta}^N d_{0\beta} N_0 \tag{16}$$

$$\bar{R}_\alpha = \bar{R}_{\alpha/0} - \sigma_c \sum_i \nu_{\alpha i}^R d_{i0} R_0 - \sigma_c \sum_\beta \chi_{\alpha\beta}^R d_{0\beta} N_0 \tag{17}$$

It is is helpful to introduce new auxiliary random variables:

$$z_N = \sum_\beta \sigma_c \bar{R}_{\beta/0} d_{0\beta} - \delta m_0 \tag{18}$$

$$z_R = \sum_j \sigma_c \bar{N}_{j/0} d_{j0} - \delta K_0 \tag{19}$$

where $\langle z_N \rangle = 0$, $\sigma_{z_N} = \sqrt{\sigma_c^2 q_R + \sigma_m^2}$ and $\langle z_R \rangle = 0$, $\sigma_{z_R} = \sqrt{\sigma_c^2 q_N + \sigma_K^2}$. Following calculations analogous to [20] and noting that $\gamma = \frac{M}{S} = 1$ yields:

$$\bar{R}_0 = \max \left[0, \frac{\sigma_c^2 \chi (K - \mu \langle N \rangle + z_R) - \mu \langle R \rangle + m - z_N}{(1 - \sigma_c^2 \nu)\sigma_c^2 \chi + 1} \right] \tag{20}$$

$$\bar{N}_0 = \max \left[0, \frac{(1 - \sigma_c^2 \nu)(\mu \langle R \rangle - m + z_N) + K - \mu \langle N \rangle + z_R}{(1 - \sigma_c^2 \nu)\sigma_c^2 \chi + 1} \right] \tag{21}$$

Cavity equations for the susceptibilities can be obtained directly by differentiating these equations:

$$\nu = \frac{1}{M} \sum_i \nu_{ii}^N = \frac{\partial \langle \bar{N}_0 \rangle}{\partial m} = -\frac{\phi_N (1 - \sigma_c^2 \nu)}{(1 - \sigma_c^2 \nu)\sigma_c^2 \chi + 1} \tag{22}$$

$$\chi = \frac{1}{M} \sum_\alpha \chi_{\alpha\alpha}^R = \frac{\partial \langle \bar{R}_0 \rangle}{\partial K} = \frac{\phi_R \sigma_c^2 \chi}{(1 - \sigma_c^2 \nu)\sigma_c^2 \chi + 1} \tag{23}$$

Two solutions are found:

$$\sigma_c < 1, \quad \chi = 0, \quad \nu = \frac{\phi_N}{\sigma_c^2 \phi_N - 1} \tag{24}$$

$$\sigma_c > 1, \chi = \left(\phi_R - \frac{1}{\sigma_c^2} \right) \left(1 - \frac{\phi_N}{\phi_R} \right), \quad \nu = -\frac{\phi_N}{\sigma_c^2 (\phi_R - \phi_N)} \tag{25}$$

The comparison between cavity solutions and numerical simulations are given in Figure S3 and Figure 3(C).

1. Three Regimes of Behavior

To understand these solutions and behaviors better, it is helpful to consider three regimes: Regime A where $\sigma_c \ll 1$, Regime B where $\sigma_c \approx 1$, and Regime C where $\sigma_c \gg 1$. Equation (25) shows the linear response function χ in Regime B: $\sigma_c \sim 1$ and Regime C: $\sigma_c \gg 1$ are only different at order $1/\sigma_c^2$. After the phase transition $\sigma_c^* = 1$, a slight increase of σ_c will induce a transition from Regime B into Regime C. This explains the dramatic drop of the species packing shown in Figure 1(B). In Regime A ($\sigma_c \ll 1$), the equations for the steady-states become

$$R_0 = \max\left[0, m - z_N\right], \quad N_0 = \max\left[0, K + z_R\right] \tag{26}$$

For Regime C ($\sigma_c \gg 1$), the solution is well approximated by

$$R_0 = \max\left[0, \frac{K - \mu\left\langle N\right\rangle + z_R}{1 - \sigma_c^2 \nu}\right] \tag{27}$$

$$N_0 = \max\left[0, \frac{\mu\left\langle R\right\rangle - m + z_N}{\sigma_c^2 \chi}\right], \tag{28}$$

in agreement with the equations obtained in [20] for purely random interactions.

D. Lotka-Volterra Model, Wishart Matrix and Marchenko-Pastur Law

In this section, we show how the generalized Lotka-Volterra model can be related to the CRM, and in particular, the how the steady states of the two models can be made to coincide. Solving for the steady-state values of the non-extinct resources by setting the bottom equation in (1) equal to zero gives:

$$\bar{R}_\alpha = K_\alpha - \sum_i N_i \bar{C}_{i\alpha}$$

Substituting this into the top equation in (1) gives:

$$\frac{dN_i}{dt} = N_i \left(\sum_{\alpha \in \mathbf{M}^*} C_{i\alpha} K_\alpha - m_i - \sum_j A_{ij} N_j \right)$$

where we have defined an interaction matrix $A_{ij} = \sum_{\alpha \in \mathbf{M}^*} \bar{C}_{i\alpha} \bar{C}^T_{\alpha j}$ and $\mathbf{M}^*$ is the set of surviving resources. We can use this equation to solve for the steady-state (equilibrium) abundances of non-extinct species, and arrive at the expression:

$$\bar{N}_i = \sum_{j \in \mathbf{S}^*} A^{-1}_{ij} \left(\sum_{\alpha \in \mathbf{M}^*} C_{j\alpha} K_\alpha - m_j \right)$$

where $\mathbf{S}^*$ is the set of surviving species. In terms of $\bar{N}_i$, the Lotka-Volterra equations become:

$$\frac{dN_i}{dt} = -\bar{N}_i \sum_j A_{ij} (N_j - \bar{N}_j) \tag{29}$$

with community matrix

$$J_{ij} = \left(\frac{\partial}{\partial N_j} \frac{dN_i}{dt} \right)_{\{\bar{N}_j\}} = -\bar{N}_i A_{ij}. \tag{30}$$

In May's work, J_{ij} is assumed to be an i.i.d. random matrix and an extension of Wigner's arguments about Gaussian random matrices is used to compute the leading eigenvalue [1]. Since the $\bar{N}_i$ are not known *a priori*, the stability of Lotka-Volterra type dynamics are more easily studied in terms of the eigenvalues of A_{ij}, using the connection between the leading eigenvalues of $\mathbf{J}$ and $\mathbf{A}$ derived below. Furthermore, for the Lotka-Volterra model obtained from the MCRM, $\mathbf{A}$ is the outer product of a random matrix $\bar{\mathbf{C}}$ with itself, i.e., a Wishart matrix. The underlying reason for this is the bipartite nature of the MCRM resulting from the presence of two types of degrees of freedom: resources and

species [27, 28]. Wishart matrices are well-known to follow a different eigenvalue distribution, the Marchenko-Pastur law[22] given by

$$\rho(x) = \frac{1}{2\pi\sigma_c^2 cx}\sqrt{(b-x)(x-a)} \tag{31}$$

where $c = \frac{S^*}{M^*}$.

For the problem we are studying, the bounds of the Marchenko-Pastur distribution are given by: $a(c) = \sigma_c^2(1-\sqrt{c})^2$ and $b(c) = \sigma_c^2(1+\sqrt{c})^2$. In Regime C: the random phase, the spectrum of $\mathbf{A}$ is extremely well described by $\rho(x)$ (see Figure 3(B) and Figure S2). The Marchenko-Pastur law also allows us to find a analytic expression for the minimum eigenvalue of $\mathbf{A}$:

$$\lambda_{min} = \sigma_c^2(1-\sqrt{c})^2. \tag{32}$$

E. Relating the eigenvalues of A and J

In this section, we prove that the largest eigenvalue $\lambda_{\max}$ of the community matrix $\mathbf{J}$ (which controls the Lyapunov stability of the fixed point) is negative if and only if the smallest eigenvalue $\lambda_{\min}$ of the Lotka-Volterra competition matrix $\mathbf{A}$ is positive. For this stability analysis, we remove the rows and columns corresponding to species that go extinct in the steady state, since allowing $N_i = 0$ trivially generates zero eigenvalues. $\mathbf{J}$ and $\mathbf{A}$ will always refer to the resulting matrices of dimension $S^* \times S^*$.

We start by defining the diagonal matrix $\bar{\mathbf{N}}$, whose nonzero elements are the equilibrium population sizes $\bar{N}_i$. This lets us write

$$\mathbf{J} = -\bar{\mathbf{N}}^{1/2}(\bar{\mathbf{N}}^{1/2}\mathbf{A}\bar{\mathbf{N}}^{1/2})\bar{\mathbf{N}}^{-1/2} \tag{33}$$

where $\bar{\mathbf{N}}^{1/2}$ is the diagonal matrix whose entries are the square roots of the population sizes. This equation says that $\mathbf{J}$ is similar to $-\mathbf{W} \equiv -\bar{\mathbf{N}}^{1/2}\mathbf{A}\bar{\mathbf{N}}^{1/2}$, which implies that they share the same eigenvalues.

Since $\mathbf{W}$ and $\mathbf{A}$ are both symmetric matrices, their eigenvalues are all real, and the positivity of all the eigenvalues is equivalent to the positive-definiteness of the matrix.

Now we note that $\mathbf{W}$ is positive definite if and only if $\mathbf{A}$ is positive definite. For if $\mathbf{A}$ is positive definite, then $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ for all column vectors $\mathbf{x} \neq 0$, including the column vector $\mathbf{x} = \bar{\mathbf{N}}^{1/2}\mathbf{y}$ for any column vector $\mathbf{y} \neq 0$. But this implies that $\mathbf{y}^T\bar{\mathbf{N}}^{1/2}\mathbf{A}\bar{\mathbf{N}}^{1/2}\mathbf{y} > 0$ for all $\mathbf{y} \neq 0$, i.e., that $\mathbf{W}$ is positive definite. Conversely, if $\mathbf{W}$ is positive definite, then $\mathbf{y}^T\bar{\mathbf{N}}^{1/2}\mathbf{A}\bar{\mathbf{N}}^{1/2}\mathbf{y} > 0$ for all $\mathbf{y} \neq 0$, including $\mathbf{y} = \bar{\mathbf{N}}^{-1/2}\mathbf{x}$ for any $\mathbf{x} \neq 0$. But this implies that $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ for all $\mathbf{x} \neq 0$, i.e., that $\mathbf{A}$ is positive definite.

We conclude that the eigenvalues of $\mathbf{W}$ are all positive if and only if the eigenvalues of $\mathbf{A}$ are all positive. Therefore the largest eigenvalue of $\mathbf{J} = -\bar{\mathbf{N}}^{1/2}\mathbf{W}\bar{\mathbf{N}}^{-1/2}$ is negative if and only if the smallest eigenvalue of $\mathbf{A}$ is positive, as claimed in the main text.

F. Correspondence between RMT and cavity solution

Our numerical simulations show that after the transition, our ecosystems are well described by purely random interactions. This suggests that we should be able to derive our cavity results using Random Matrix Theory (RMT). We now show that this is indeed the case. Our starting point are the average susceptibilities which are defined as:

$$\chi = \frac{1}{M}\sum_{\alpha\in\mathbf{M}}\chi^R_{\alpha\alpha} = \frac{1}{M}\sum_{\alpha\in\mathbf{M}^*}\chi^R_{\alpha\alpha} \tag{34}$$

$$\nu = \frac{1}{S}\sum_{i\in\mathbf{S}}\nu^N_{ii} = \frac{1}{S}\sum_{i\in\mathbf{S}^*}\nu^N_{ii}. \tag{35}$$

From the cavity calculations, we only care about $\chi^R_{\alpha\beta}$ and ν^N_{ij}, because the other susceptibilities are lower order in $1/M$.

We can combine these equations with (10) and (11) to obtain

$$\chi = \frac{1}{M}\sum_{\alpha\in\mathbf{M}^*}\chi^R_{\alpha\alpha} = \frac{1}{M}\mathrm{Tr}(\chi^R_{\alpha\beta}) \tag{36}$$

$$= \frac{1}{M}\mathrm{Tr}(\delta_{\alpha\beta}) - \frac{1}{M}\mathrm{Tr}\left(\sum_{i\in\mathbf{S}^*}\sum_{j\in\mathbf{S}^*}\bar{C}^T_{\alpha i}A^{-1}_{ij}\bar{C}_{j\beta}\right)$$

$$= \frac{M^*}{M} - \frac{1}{M}\mathrm{Tr}\left(\sum_{i\in\mathbf{S}^*}\sum_{j\in\mathbf{S}^*}A^{-1}_{ij}\bar{C}_{j\beta}\bar{C}^T_{\beta h}\right)$$

$$= \frac{M^*}{M} - \frac{S^*}{M} = \phi_R - \gamma^{-1}\phi_N \tag{37}$$

We now show that the cavity solutions are consistent with results from RMT using equations (10) and(11) in Regime A and Regime C described in the main text.

1. *Regime A:* $\bar{\mathbf{C}} = \mathbb{1}$

This regime happens when $\sigma_c \ll 1$. Substituting, $\bar{\mathbf{C}} = \mathbb{1}$ into equations (10) and (11) yields

$$\chi = 0, \quad \nu = -1. \tag{38}$$

This is consistent with the cavity solution equation (24) with $\sigma_c = 0$ since in this case $S^* = S = M$.

2. *Regime C:* $\bar{C}_{i\alpha}$ *i.i.d.* $\mathcal{N}(0, \sigma_c/\sqrt{M})$

In this regime, $\sigma_c \gg 1$. In this case, $A_{ij} = \sum_{\alpha\in\mathbf{S}^*}\bar{C}_{i\alpha}\bar{C}^T_{\alpha j}$ takes the form of a Wishart Matrix. We will exploit this to calculate χ and ν. Notice,

$$\nu = \frac{1}{S}\sum_{i\in\mathbf{S}^*}\nu^N_{ii} = -\frac{1}{S}\mathrm{Tr}(A^{-1}_{ij}) = -\frac{1}{S}\sum_{i=1}^{S^*}\lambda_i^{-1} \tag{39}$$

where λ_i is the eigenvalue of A_{ij}. From the Marchenko-Pastur law [22], we know that the eigenvalues of a random Wishart matrix obey the Marchenko-Pastur distribution. Substituting equation (32) into the expression for ν and replacing the sum with an integral yields:

$$\nu = -\frac{S^*}{S}\int_a^b \frac{1}{x}\rho(x)dx \tag{40}$$

$$= -\frac{S^*}{S}\frac{a+b-2\sqrt{ab}}{4\sigma_c^2 y\sqrt{ab}}$$

$$= -\frac{1}{\sigma_c^2}\frac{\phi_N}{\phi_R - \gamma^{-1}\phi_N}$$

The second line of equation (40) is obtained by transferring the integral function to a complex analytic function and applying the residue theorem. This result is the same as the cavity solution equation (25) when $\sigma_c \gg 1$.

3. *Regime B using the Stieltjes transformation*

In Regime B, it hard to estimate the minimum eigenvalue. We can use Stieltjes transformation of information-plus-noise-type matrices which are well studied in wireless communications[29, 30], where $\mathbf{B}$ represents the information encoded in the signal and $\mathbf{C}$ is the noise in wireless communications. In this case, we have

$$\bar{C}_{i\alpha} = \mathbb{1} + C_{i\alpha}, \quad C_{i\alpha} \text{ i.i.d. } \mathcal{N}(0, \sigma_c/\sqrt{M}).$$

$$A_{ij} = \sum_{\alpha \in M^*} \bar{C}_{i\alpha} \bar{C}^T_{\alpha j} = \sum_{\alpha \in M^*} C_{i\alpha} C^T_{\alpha j} + C_{i\alpha} + C^T_{\alpha i} + \mathbb{1} \tag{41}$$

Using **Theorem 1.1** in Dozier and Silverstein[29], the Stieltjes transform $m(z)$ of A_{ij} satisfies

$$\sigma_c^4 z m^3 - 2\sigma_c^2 z m + (\sigma_c^2 + z - 1)m - 1 = 0 \tag{42}$$

The asymptotic spectrum of A_{ij} can be obtained by $m(z)$, the solution of equation (42) with

$$\rho(x) = \lim_{\varepsilon \to 0^+} \frac{m(x - i\varepsilon) - m(x + i\varepsilon)}{2i\pi} \tag{43}$$

The result is shown in Figure S4. The minimum eigenvalue reaches 0 nearly at $\sigma_c^* = 1$, as predicted by the cavity solution.

Supplementary Information

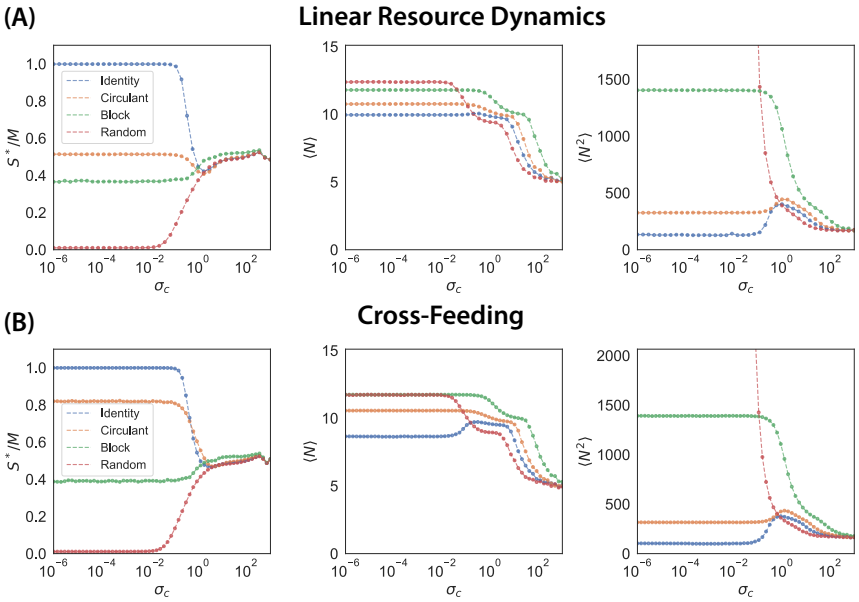

FIG. S1. Community properties for generalized consumer-resource models under Gaussian noise. **(A)** Linear resource dynamics: the resource dynamics is changed to $\frac{dR_\alpha}{dt} = K_\alpha - R_\alpha - \sum_i N_i C_{i\alpha} R_\alpha$. **(B)** With cross-feeding: the dynamics is described in equation (17) in Supplementary Information of [18]. The noise is only applied on the consumption matrix and D is kept the same at different σ_c. In both models, $\mathbf{B} = \mathbb{1}$. See Methods for parameter values. Each data point is averaged over 4000 independent realizations.

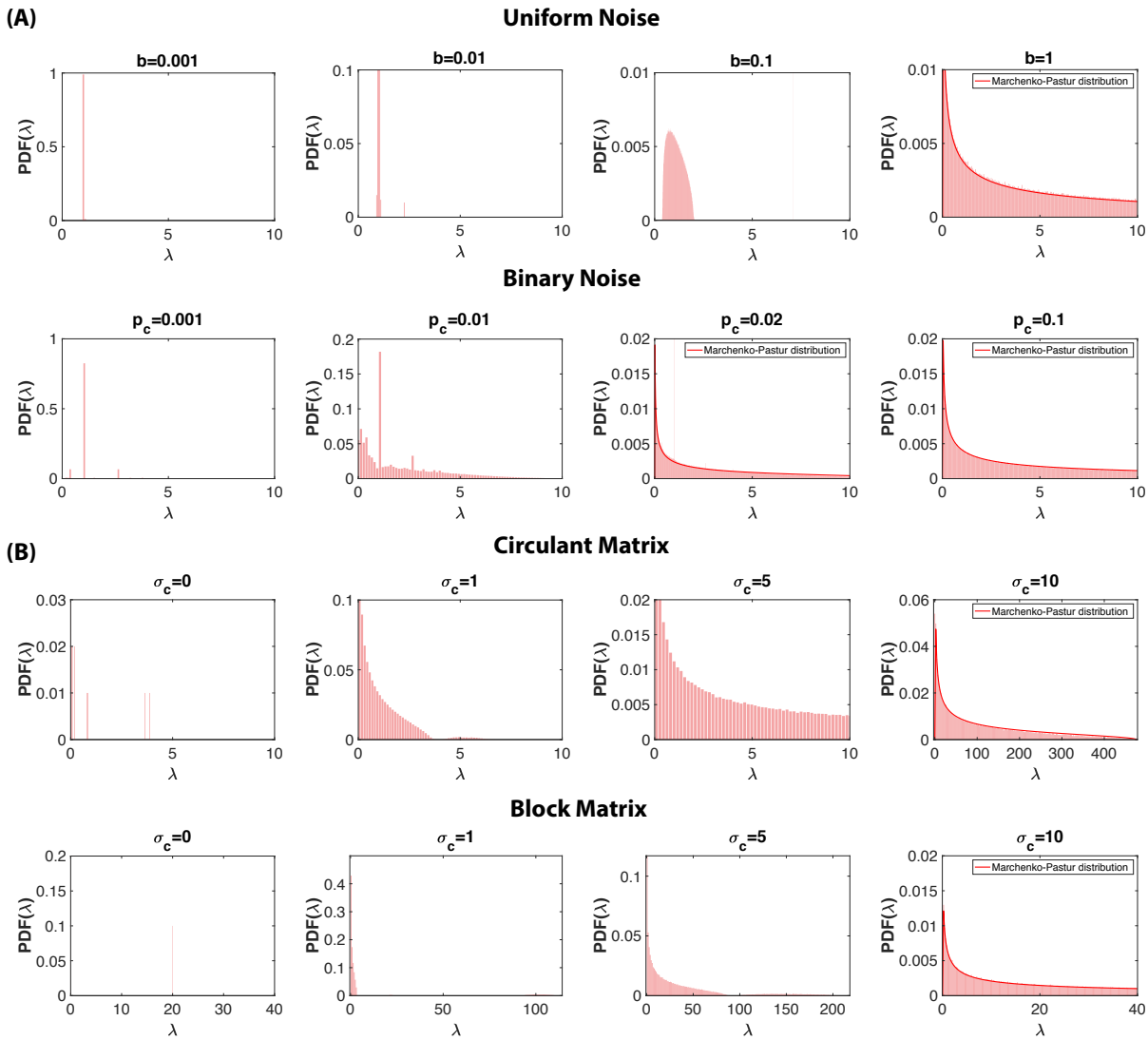

FIG. S2. Spectra of A_{ij} in different cases. **(A)** Uniform Noise: $\mathcal{U}(0,b)$ and Binary Noise: $Bernoulli(p_c)$; **B** is an identity matrix. **(B)** Gaussian noise and **B** is a circulant matrix. **(C)** Gaussian noise and **B** is a block matrix. Note that S^*,M^* in A_{ij} are obtained from numerical simulations.

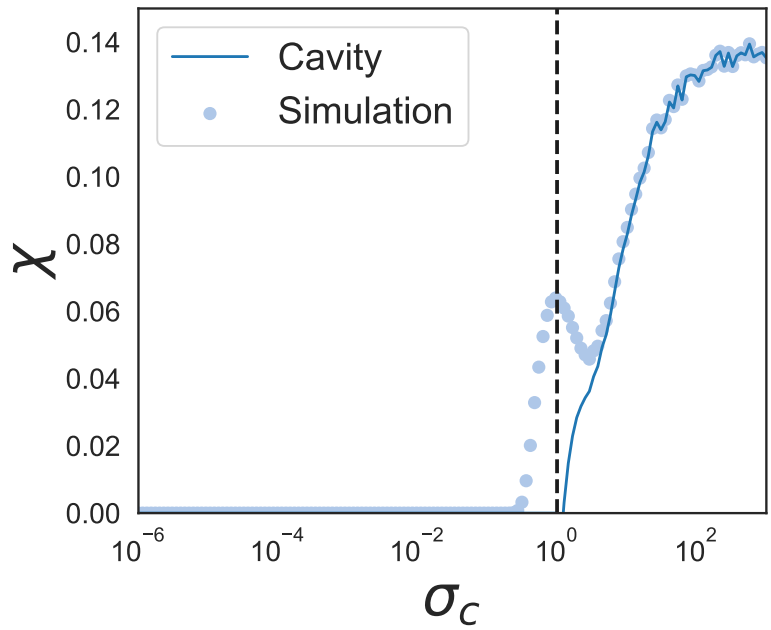

FIG. S3. Comparison between numerical simulations and cavity solutions for χ at different σ_c. Note S^* and M^* are obtained from the numerical simulations, although in principle they could be obtained by solving the cavity equations directly.

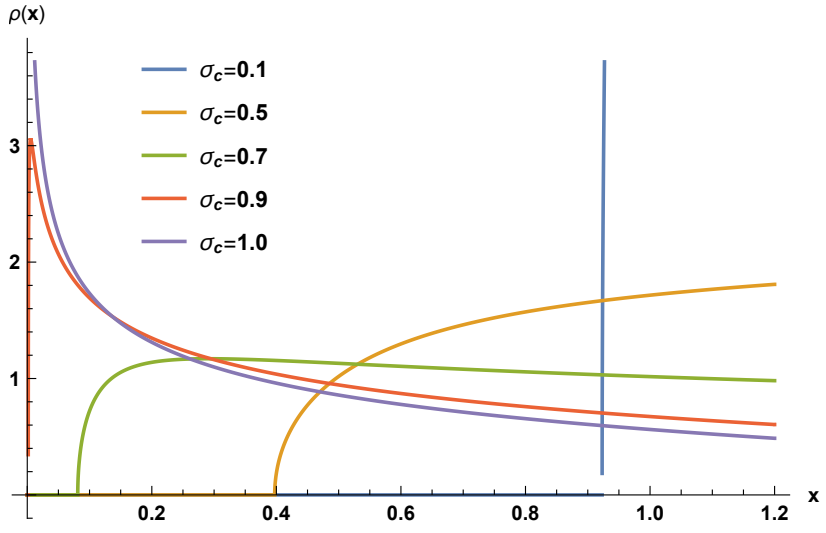

FIG. S4. The asymptotic spectrum of A_{ij} for different values of σ_c by solving equation (43) numerically.