

The effect of resource dynamics on species packing in diverse 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: November 6, 2019)

The competitive exclusion principle asserts that coexisting species must occupy distinct ecological niches (i.e. the number of surviving species can not exceed the number of resources). An open question is to understand if and how different resource dynamics affect this bound. Here, we analyze a generalized consumer resource model with externally supplied resources and show that – in contrast to self-renewing resources – species can occupy only half of all available environmental niches. This motivates us to construct a new schema for classifying ecosystems based on species packing properties.

One of the most stunning aspects of the natural world is the incredible diversity of species present in many environments [1, 2]. A major goal of community ecology is to understand the rules governing community structure and species coexistence patterns in these complex ecosystems. One promising approach that has recently emerged for tackling this challenge is to use ideas from statistical mechanics inspired by spin glass physics [3, 4]. In such an approach, ecosystems are viewed as large interacting disordered systems, allowing for the identification of universal, collective properties [5, 6]. Such statistical physics inspired models are also able to reproduce many experimental observations, especially in the context of microbial ecosystems [7–9].

Much of this work has focused on generalized Lotka-Volterra models where species directly interact with each other in a pair-wise fashion [5, 10–16]. While such models have led to deep ecological insights [17] and have allowed for the identification of interesting ecological phases and phase transitions [10–12], a major drawback of Lotka-Volterra models are that they do not explicitly model the resources present in the ecosystem. Instead, resource dynamics are implicitly represented through the choice of species-species interactions making it difficult to understand the relationship between resource dynamics and community structure.

In contrast, generalized consumer-resource models (GCRMs), first introduced by MacArthur and Levins in a series of seminal papers [18–20], explicitly incorporate both species and resource dynamics. In GCRMs, ecosystems are described by species that can consume and deplete resources according to a set of consumer preferences. Interactions between species arise because species with similar consumer preferences occupy similar environmental niches and hence compete for common resources. An important theoretical and conceptual result that follows from GCRMs is that the number of species that can coexist in an ecosystem is limited by the number of resources that are present. In other words, if we denote

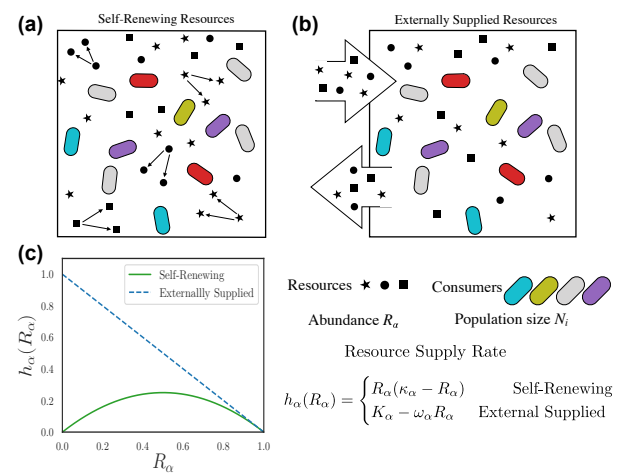


FIG. 1. Schematic description for two types of resources. (a) Self-renewing resources (e.g. plants), which are replenished through organic reproduction; (b) Externally supplied resources (e.g. nutrients that sustain gut microbiota), which are replenished by a constant flux from some external source, and diluted at a constant rate; (c) The supply rate as a function of resource abundance for both choices, with $\kappa = \omega_\alpha = K_\alpha = 1$.

the number of species that can survive in an ecosystem by S^* and the number of supplied resources as M , the competitive exclusion principle yields an upper bound for the amount of species that can be packed into the ecosystem: $\frac{S^*}{M} \leq 1$ [21].

The basic intuition behind this bound is that the growth rates $g_i(\mathbf{R})$ of all coexisting species $i = 1, 2, \dots$ must simultaneously vanish, and since the space of resource concentrations $\mathbf{R}$ is M -dimensional, at most M of these equations can be simultaneously solved (see Appendix for discussion of non-generic phenomena where the bound is violated). While this result gives an upper bound, it is not clear when and if it will be saturated. In particular, we show below that the choice of resource dynamics fundamentally alters species-packing properties.

To show this, we analyze GCRMs with two different resource dynamics: *self-renewing* resources where resources grow logistically in the absence of consumers [18, 19] and *externally supplied* resources that are supplied and degraded at a constant rate [22–24] (see Fig. 1). We derive species packing bounds for both choices of dynamics by analyzing the susceptibilities of a new cavity solution for GCRMs with externally supplied resources and combining it with the previously derived cavity solution for GCRMs with self-renewing resources [6, 25, 26]. Surprisingly, in the absence of metabolic tradeoffs we find that, for externally-supplied resources, species can occupy only half of all available resource niches: $\frac{S^*}{M} < \frac{1}{2}$. Motivated by these results, we suggest a new schema for classifying ecosystems based on their species packing properties.

Model: GCRMs describe the ecological dynamics of S species of consumers N_i ($i = 1, 2, \dots, S$) that can consume M distinct resources R_α ($\alpha = 1, 2, \dots, M$). The rate at which species N_i consumes and depletes resource R_β is encoded in a matrix of consumer preferences $C_{i\beta}$. In order to survive, species have a minimum maintenance cost m_i . Equivalently, m_i can also be thought of as the death rate of species i in the absence of resources. These dynamics can be described using a coupled set of $M + S$ ordinary differential equations of the form

$$\begin{cases} \frac{dN_i}{dt} = N_i \sum_\beta C_{i\beta} R_\beta - N_i m_i \\ \frac{dR_\alpha}{dt} = h_\alpha(R_\alpha) - \sum_j N_j C_{j\alpha} R_\alpha, \end{cases} \quad (1)$$

where $h_\alpha(R_\alpha)$ a function that describes the dynamics of the resources in the absence of any consumers (see Fig. 1).

For self-renewing resources (e.g. plants, animals), the dynamics can be described using logistic growth of the form

$$h_\alpha(R_\alpha) = R_\alpha(\kappa_\alpha - R_\alpha), \quad (2)$$

with κ the carrying capacity. While such resource dynamics is reasonable for biotic resources, abiotic resources such as minerals and small molecules cannot self-replicate and are usually supplied externally to the ecosystem (Fig. 1(b)). A common way to model this scenario is by using linearized resource dynamics of the form

$$h_\alpha(R_\alpha) = K_\alpha - \omega_\alpha R_\alpha. \quad (3)$$

Fig. 1(c) shows a plot of these two choices. Notice that the two resource dynamics behave very differently at low resource levels. The self-renewing resources can go extinct and eventually disappear from the ecosystem while this is not true of externally supplied resources.

Recent research has shown some unexpected and interesting non-generic phenomena can appear in GCRMs in the presence of additional constraints on parameter values. A common choice of such constraints is the imposition of a “metabolic budget” on the consumer preference

matrix [22, 27] tying the maintenance cost m_i to the total consumption capacity $\sum_\beta C_{i\beta}$ [23, 28]. These metabolic tradeoffs can be readily incorporated into the cavity calculations and have significant impacts on species packing as will be discussed below.

Cavity solution: Recently, we derived a mean-field cavity solution for steady-state dynamics of the the GCRM with self-renewing resource dynamics in the high-dimensional limit where the number of resources and species in the regional species pool is large ($S, M \gg 1$) [6, 25, 26]. The overall procedure for deriving the cavity equations for GCRM with externally supplied resource is similar to that for GCRMs with self-renewing resources and is shown in Fig. 5 in the Appendix. We assume the K_α and m_i are independent random normal variables with means K and m and variances σ_K^2 and σ_m^2 , respectively. We also assume ω_α are independent normal variables with mean ω and variance σ_ω^2 . The elements of the consumption matrix $C_{i\alpha}$ are drawn independently from a normal distribution with mean μ/M and variance σ_c^2/M . This scaling with M is necessary to guarantee that $\langle N \rangle$, $\langle R \rangle$ do not vanish when $S, M \gg 1$ with $M/S = \gamma$ fixed. Later, we will consider a slightly modified scenario where the maintenance costs are correlated with the consumption matrix in order to implement the metabolic tradeoffs discussed above.

The basic idea behind the cavity method is to derive self-consistency equations relating an ecosystem with M resources and S species to an ecosystem with $M + 1$ resources and $S + 1$ species. This is done by adding a new “cavity” species 0 and a new “cavity” resource 0 to the original ecosystem. When $S, M \gg 1$, the effect of the new cavity species/resource is small and can be treated using perturbation theory. The cavity solution further exploits the fact that since the $C_{i\alpha}$ are random variables, when $M \gg 1$ the sum $\sum_\alpha C_{i\alpha} R_\alpha$ will be well described by a by a normal distribution with mean $\mu \langle R \rangle$ and variance $\sigma_c^2 q_R$ where $q_R = \langle R^2 \rangle = 1/M \sum_\alpha R_\alpha^2$ (see Appendix for details). Combining this with the non-negativity constraint, the species distribution can be expressed as a truncated normal distribution,

$$\bar{N} = \max \left[0, \frac{\mu \langle R \rangle - m + \sqrt{\sigma_c^2 q_R + \sigma_m^2} z_N}{\sigma_c^2 \chi} \right] \quad (4)$$

where $\chi = -\left\langle \frac{\partial \bar{R}_\alpha}{\partial \omega_\alpha} \right\rangle = -M^{-1} \sum_\alpha \frac{\partial \bar{R}_\alpha}{\partial \omega_\alpha}$ and z_N is a standard normal variable. This equation describes GCRMs with both externally supplied and self-renewing resource dynamics [25].

The steady-state cavity equations for externally supplied resources are significantly more complicated and technically difficult to work with than the corresponding equations for self-renewing resources. To see this, notice that the steady-state abundance of resource α can be found by plugging in Eq. 3 into Eq 1 and setting the

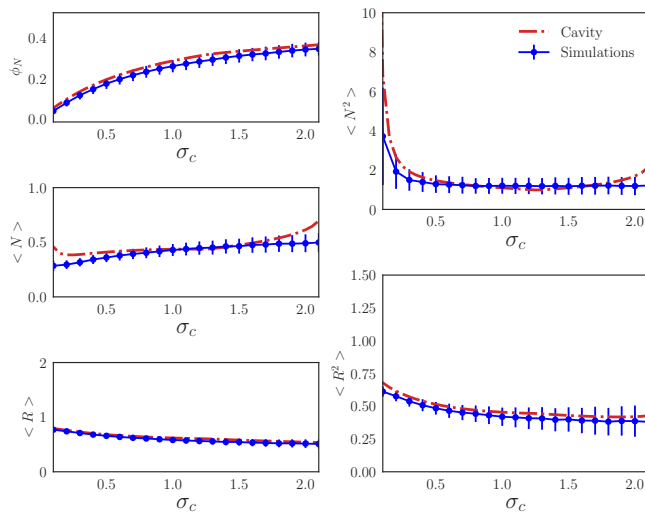


FIG. 2. Comparison between cavity solutions (see main text for definition) and simulations for the fraction of surviving species $\phi_N = \frac{S^*}{S}$ and the first and second moments of the species and resources distributions as a function of σ_c . The error bar shows the standard deviation from 1000 numerical simulations with $M = S = 100$ and all other parameters are defined in the Appendix. Simulations were run using the CVXPY package [30].

left hand side to zero to get

$$\bar{R}_\alpha = K_\alpha / (\omega_\alpha + \sum_j \bar{N}_j C_{j\alpha}) = \frac{K_\alpha}{\omega_\alpha^{\text{eff}}}, \quad (5)$$

where we have defined $\omega_\alpha^{\text{eff}} = \omega_\alpha + \sum_j \bar{N}_j C_{j\alpha}$. When $S \gg 1$, both the denominator $\omega_\alpha^{\text{eff}}$ and the numerator K_α can be modeled by independent normal random variables. This implies that the steady-state resource abundance is described by a ratio of normal variables (i.e. the Normal Ratio Distribution) instead of a truncated Gaussian as in the self-renewing case [29]. At large σ_c , this makes solving the cavity equations analytically intractable. Luckily, if the variance of the denominator $\omega_\alpha^{\text{eff}}$ is small compared with the mean – which is true when σ_c not too large – we can still obtain an approximate replica-symmetric solution by expanding in powers of the standard deviation over the mean of $\omega_\alpha^{\text{eff}}$ (see Appendix). We consider expansions to the cavity solutions where the denominator in Eq. 5 is expanded to 1st order. In general, the backreaction correction is quite involved since resources and species form loopy interactions resulting in non-trivial correlation between $C_{i\alpha}$ and N_i that must be properly accounted for (see Appendix).

Comparison with numerics: The full derivation of 1st order expansions of the mean-field equations are given in the Appendix. The resulting self-consistency equations can be solved numerically in Mathematica. Fig. 2 shows a comparison between the cavity solution and 1000 in-

dependent numerical simulations for various ecosystem properties such as the fraction of surviving species S^*/S and the first and second moment of the species and resource distributions (simulation details are in the Appendix). As can be seen in the figure, our analytic expressions agree remarkably well over a large range of σ_c . However, at very large σ_c (not shown), the cavity solutions start deviating from the numerical simulations because the Ratio Normal Distribution can no longer be described using the 1st order expansion to the full cavity equations.

As a further check on our analytic solution, we ran simulations where the $C_{i\alpha}$ were drawn from different distributions. One pathology of choosing $C_{i\alpha}$ from a Gaussian distribution is that when σ_c is large, many of consumption coefficients are negative. To test whether our cavity solution still describes ecosystems when $C_{i\alpha}$ are strictly positive, we compare our cavity solution to simulations where the $C_{i\alpha}$ are drawn from a Bernoulli or uniform distribution. As before, there is remarkable agreement between analytics and numerics (see Fig. 6)

Species packing without metabolic tradeoffs: The essential ingredients needed to derive species packing bounds for GCRMS are the cavity equations for the average local susceptibilities $\nu = \left\langle \frac{\partial \bar{N}_i}{\partial m_i} \right\rangle = S^{-1} \sum_j \frac{\partial \bar{N}_i}{\partial m_i}$ and $\chi = \left\langle \frac{\partial \bar{R}_\alpha}{\partial X_\alpha} \right\rangle = M^{-1} \frac{\partial \bar{R}_\alpha}{\partial X_\alpha}$, with $X_\alpha = K_\alpha$ for externally supplied resources and $X_\alpha = -\omega_\alpha$ for self-renewing resources. These two susceptibilities measure how the mean species abundance and mean resource abundance respond to changes in the species death rate and the resource supply/depletion rate, respectively. They play an essential role in the cavity equation and can be used for distinguishing different phases in complex systems[6, 31].

For the self-renewing case, the susceptibilities χ_s and ν_s are given by eq. (59, 60) in [26]

$$\nu_s = -\frac{\phi_N}{\sigma_c^2 \chi_s}, \quad \chi_s = \frac{\phi_R}{1 - \gamma^{-1} \sigma_c^2 \nu_s}, \quad (6)$$

and can be reduced to $\chi_s = \phi_R - \gamma^{-1} \phi_N$, where $\phi_R = M^*/M$, with M^* equal to the number of non-extinct resources in the ecosystem. In order to guarantee the positivity of $\langle N \rangle$, we must have $\chi_s = \phi_R - \gamma^{-1} \phi_N > 0$, resulting in an upper bound

$$1 \geq \frac{M^*}{M} > \frac{S^*}{M} \quad (7)$$

which states that the number of surviving resources must be smaller than the number of surviving species.

For the externally supplied case, the corresponding equations take the form

$$\nu = -\frac{\phi_N}{\sigma_c^2 \chi}, \quad \chi = -\frac{1}{2\gamma^{-1} \nu \sigma_c^2} (1 - \langle \dots \rangle), \quad (8)$$

where the full expression of $\langle \dots \rangle$ can be found in eq. (72) in the appendix. For our purposes, the most important

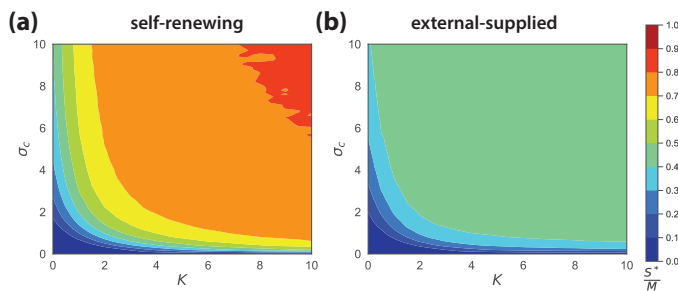


FIG. 3. Comparison of the species packing ratio S^*/M at various σ_c and K for self-renewing and externally supplied resource dynamics. The simulations represent averages from 1000 independent realizations with the system size $M = 100$, $S = 500$ (parameters in Appendix).

property is that in the absence of metabolic tradeoffs, the expression $\langle \dots \rangle$ is always *positive*. Combining this observation with the equations above gives the upper bound

$$\frac{1}{2} > \frac{S^*}{M} = \phi_N \gamma^{-1}. \quad (9)$$

Thus, for externally supplied resources, at most *half of all potential niches* are occupied. Fig. 3 shows numerical simulations confirming the species packing bound for various choices of K and σ_c .

Species packing with metabolic tradeoffs: We also find that metabolic tradeoffs modify the cavity equations in such a way that the expression in brackets $\langle \dots \rangle$ in Equation (8) can become negative (see Appendix). However, it still remains greater than -1, allowing us to derive a species packing bound of the form $S^* < M$ even in the presence of soft metabolic constraints. In Figure 4, we simulated various ecosystems where the maintenance costs of species were chosen to obey metabolic tradeoffs of the form $m_i = \sum_{\alpha} C_{i\alpha} + \delta m_i$, where δm_i are i.i.d. normal variables with variance σ_m^2 . Note that a larger σ_m corresponds to ecosystems with softer metabolic constraints. We found that when $\sigma_m/\sigma_c > 1$, these ecosystems obey the 1/2 species packing bound derived above. This can also be analytically shown using the modified cavity equations derived in the appendix. Finally, we show in the appendix that when the metabolic tradeoffs take the form of hard constraints on the consumer preferences as in [22, 23, 27, 28], the cavity equations allow for interesting non-generic behavior with $S^* \geq M$, consistent with these previous works. Importantly, we find that even modest modifications of the tradeoff equation $m_i \propto \sum_{\alpha} C_{i\alpha}$ results in ecosystems that satisfy the 1/2 species packing bound.

Classifying ecosystems using species packing: Recently, it has become clear that there is a deep relationship between ecosystem and constraint satisfaction problems [23, 24, 26, 28]. In particular, each species can be thought of as a constraint on possible resource abundances [24, 26]. Inspired by jamming [32], this

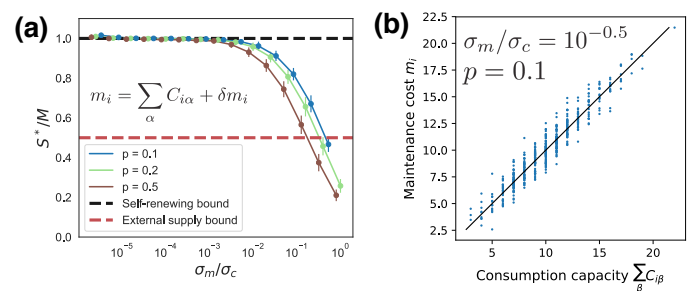


FIG. 4. Species packing bounds in the presence of metabolic tradeoffs. (a) The species packing ratio S^*/M as a function of σ_m/σ_c , where σ_m is the standard deviation of the δm_i and $\sigma_c/\sqrt{M}$ is the standard deviation of $C_{i\alpha}$. Simulations are for binary consumer preference matrix $C_{i\alpha}$ drawn from a Bernoulli distribution with probability p . (b) m_i versus $\sum_{\alpha} C_{i\alpha}$ for $p = 0.1$ and $\sigma_m/\sigma_c = 10^{-0.5}$. See Appendix for all parameters.

suggests that we can separate ecosystems into qualitatively distinct classes depending on whether the competitive exclusion bound is saturated. We designate ecosystems where $S^* \rightarrow M$ (like GCRMs with self-renewing resources) as *isostatic species packings*, and ecosystems where the upper bound $S_{\max}$ on the number of surviving species is strictly less than the number of resources $S^* < S_{\max} < M$ (like GCRMs with externally supplied resources without metabolic tradeoffs) as *hypostatic species packings*. Ecosystems with $S^* \geq M$ (like GCRMs with hard metabolic constraints) are designated as *non-generic species packings* because of the presence of a macroscopic number of additional hard constraints (i.e. the number of additional constraints that are imposed scales with S and M in the limit $S, M \rightarrow \infty$). This basic schema suggests a way of refining the competitive exclusion principle and may help shed light on controversies surrounding the validity of basic species packing bounds.

Discussion: In this paper, we examine the effect of resource dynamics on community structure and large-scale ecosystem level properties. To do so, we analyzed generalized Consumer Resource Models (GCRMs) with two different resource dynamics: externally supplied resources that are supplied and degraded at a constant rate and self-replicating resources whose behavior in the absence of consumers is well described by a logistic growth law. Using a new cavity solution for GCRMs with externally supplied resources and a previously found cavity solution of the GCRM with self-renewing resources, we show that the community structure is fundamentally altered by the choice of resource dynamics. In particular, for externally supplied resources, we find that species generically can only occupy *half* of all available niches whereas for self-renewing resources all environmental niches can be filled. We confirm this surprising bound using numerical simulations.

These results show how resource dynamics, which are neglected in commonly used Lotka-Volterra models, can fundamentally alter the properties of ecosystems. Much work still needs to be done to see if and how our results must be modified to account for other ecological processes such as demographic stochasticity, spatial structure, and microbe-specific interactions such as cross-feeding [7, 8]. It will also be necessary to move beyond steady-states and consider the dynamical properties of these ecosystems. More generally, it will be interesting to further explore the idea that we can classify ecosystems based on species-packing properties and see if such a schema can help us better understand the origins of the incredible diversity we observe in real-world ecosystems.

Acknowledgments: The work was supported by NIH NIGMS grant 1R35GM119461, Simons Investigator in the Mathematical Modeling of Living Systems (MMLS). The authors also acknowledge support from the SSC computing cluster at BU.

* cuiw@bu.edu

† marsland@bu.edu

‡ pankajm@bu.edu

- [1] C. Huttenhower, D. Gevers, R. Knight, S. Abubucker, J. H. Badger, A. T. Chinwalla, H. H. Creasy, A. M. Earl, M. G. FitzGerald, R. S. Fulton, *et al.*, *Nature* **486**, 207 (2012).
- [2] A. H. Gentry, *Proceedings of the National Academy of Sciences* **85**, 156 (1988).
- [3] M. Mézard, G. Parisi, and M. Virasoro, *Spin glass theory and beyond: An Introduction to the Replica Method and Its Applications*, Vol. 9 (World Scientific Publishing Company, 1987).
- [4] H. Nishimori, *Statistical physics of spin glasses and information processing: an introduction*, 111 (Clarendon Press, 2001).
- [5] M. Barbier, J.-F. Arnoldi, G. Bunin, and M. Loreau, *Proceedings of the National Academy of Sciences* **115**, 2156 (2018).
- [6] W. Cui, R. Marsland III, and P. Mehta, arXiv preprint arXiv:1904.02610 (2019).
- [7] 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).
- [8] R. Marsland III, W. Cui, J. Goldford, A. Sanchez, K. Korolev, and P. Mehta, *PLoS computational biology* **15**, e1006793 (2019).
- [9] R. Marsland III, W. Cui, and P. Mehta, arXiv preprint arXiv:1904.12914 (2019).
- [10] C. K. Fisher and P. Mehta, *Proceedings of the National Academy of Sciences* **111**, 13111 (2014).
- [11] D. A. Kessler and N. M. Shnerb, *Physical Review E* **91**, 042705 (2015).
- [12] G. Bunin, *Physical Review E* **95**, 042414 (2017).
- [13] M. Barbier and J.-F. Arnoldi, [bioRxiv preprint bioRxiv: 10.1101/147728](https://doi.org/10.1101/147728) (2017).
- [14] G. Biroli, G. Bunin, and C. Cammarota, *New Journal of Physics* **20**, 083051 (2018).
- [15] F. Roy, G. Biroli, G. Bunin, and C. Cammarota, *Journal of Physics A: Mathematical and Theoretical* (2019).
- [16] M. T. Pearce, A. Agarwala, and D. S. Fisher, *bioRxiv*, 736215 (2019).
- [17] P. Chesson, *Annual review of Ecology and Systematics* **31**, 343 (2000).
- [18] R. MacArthur and R. Levins, *The American Naturalist* **101**, 377 (1967).
- [19] P. Chesson, *Theoretical Population Biology* **37**, 26 (1990).
- [20] R. MacArthur, *Theoretical population biology* **1**, 1 (1970).
- [21] R. McGehee and R. A. Armstrong, *Journal of Differential Equations* **23**, 30 (1977).
- [22] A. Posfai, T. TAILLEFUMIER, and N. S. Wingreen, *Physical review letters* **118**, 028103 (2017).
- [23] M. Tikhonov and R. Monasson, *Physical review letters* **118**, 048103 (2017).
- [24] R. Marsland III, W. Cui, and P. Mehta, arXiv preprint arXiv:1901.09673 (2019).
- [25] M. Advani, G. Bunin, and P. Mehta, *Journal of Statistical Mechanics: Theory and Experiment*, 033406 (2018).
- [26] P. Mehta, W. Cui, C.-H. Wang, and R. Marsland III, *Physical Review E* **99**, 052111 (2019).
- [27] Z. Li, B. Liu, S. H.-J. Li, C. G. King, Z. Gitai, and N. S. Wingreen, *bioRxiv*, 664698 (2019).
- [28] A. Altieri and S. Franz, *Physical Review E* **99**, 010401 (2019).
- [29] G. Marsaglia *et al.*, *Journal of Statistical Software* **16**, 1 (2006).
- [30] A. Agrawal, R. Verschueren, S. Diamond, and S. Boyd, *Journal of Control and Decision* **5**, 42 (2018).
- [31] M. Ramezanali, P. P. Mitra, and A. M. Sengupta, arXiv preprint arXiv:1501.03194 (2015).
- [32] A. J. Liu and S. R. Nagel, *Annu. Rev. Condens. Matter Phys.* **1**, 347 (2010).

Derivation of cavity solution

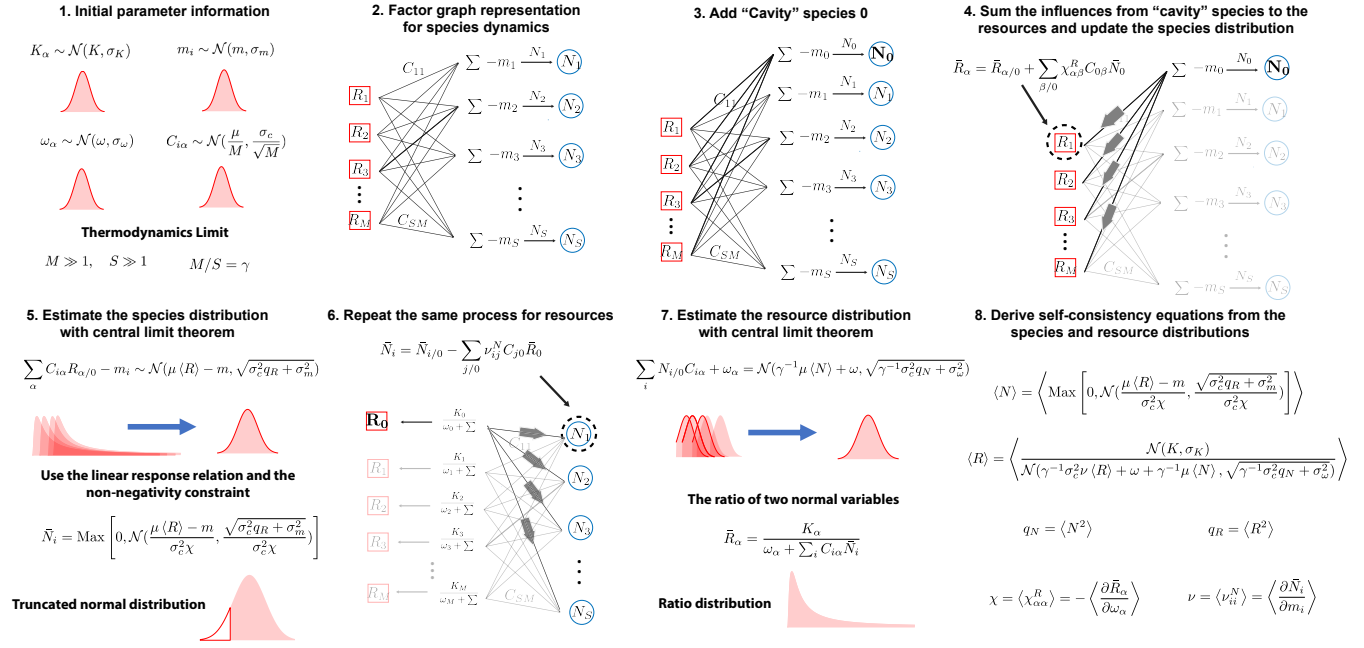


FIG. 5. Schematic outlining steps in cavity solution. **1.** The initial parameter information consists of the probability distributions for the mechanistic parameters: K_α , m_i and $C_{i\alpha}$. We assume they can be described by their first and second moments. **2.** The species dynamics $N_i(\sum_\alpha C_{i\alpha} R_\alpha - m_i)$ in eqs. (10) are expressed as a factor graph. **3.** Add the "Cavity" species 0 as the perturbation. **4.** Sum the resource abundance perturbations from the "Cavity" species 0 at steady state and update the species abundance distribution to reflect the new steady state. **5.** Employing the central limit theorem, the backreaction contribution from the "cavity" species 0 and the non-negativity constraint, the species distribution is expressed as a truncated normal distribution. **6.** Repeat **Step 2-4** for the resources. **7.** The resource distribution is the ratio distribution from the ratio of two normal variables K_α and $\omega_\alpha + \sum_i N_i C_{i\alpha}$. **8.** The self-consistency equations are obtained from the species and resource distributions. Note that $\gamma^{-1} \sigma_c^2 \nu \langle R \rangle$ in the dominator of $\langle R \rangle$ is from the correlation between N_i and $C_{i\alpha}$ in $\sum_i N_i C_{i\alpha}$.

Model setup

In this section, we derive the cavity solution to the linear resource dynamics (eq. (1) in the main text)

$$\begin{cases} \frac{dN_i}{dt} = N_i \left(\sum_\beta C_{i\beta} R_\beta - m_i \right) \\ \frac{dR_\alpha}{dt} = K_\alpha - \omega_\alpha R_\alpha - \sum_j N_j C_{j\alpha} R_\alpha \end{cases} \quad (10)$$

Note that here we follow closely our derivation in [25, 26]. The main difference is that here we consider linear resource dynamics, which as we will see below, makes the problem much more technically challenging.

Consumer preference $C_{i\alpha}$ are random variables drawn from a Gaussian distribution with mean μ/M and variance σ_c^2/M . They can be deposed into $C_{i\alpha} = \mu/M + \sigma_c d_{i\alpha}$, where the fluctuating part $d_{i\alpha}$ obeys

$$\langle d_{i\alpha} \rangle = 0 \quad (11)$$

$$\langle d_{i\alpha} d_{j\beta} \rangle = \frac{\delta_{ij} \delta_{\alpha\beta}}{M}. \quad (12)$$

We also assume that both the carrying capacity K_α and the minimum maintenance cost m_i are independent Gaussian

random variables with mean and covariance given by

$$\langle K_\alpha \rangle = K \quad (13)$$

$$\text{Cov}(K_\alpha, K_\beta) = \delta_{\alpha\beta} \sigma_K^2 \quad (14)$$

$$\langle m_i \rangle = m \quad (15)$$

$$\text{Cov}(m_i, m_j) = \delta_{ij} \sigma_m^2 \quad (16)$$

Let $\langle R \rangle = \frac{1}{M} \sum_\beta R_\beta$ and $\langle N \rangle = \frac{1}{S} \sum_j N_j$ be the average resource and average species abundance, respectively. With all these defined, we can re-write eqs. (10) as

$$\frac{dN_i}{dt} = N_i \left\{ \mu \langle R \rangle - m + \sum_\beta \sigma_c d_{i\beta} R_\beta - \delta m_i \right\} \quad (17)$$

$$\frac{dR_\alpha}{dt} = K + \delta K_\alpha - \left[\omega_\alpha + \gamma^{-1} \mu \langle N \rangle + \sum_j \sigma_c d_{j\alpha} N_j \right] R_\alpha \quad (18)$$

where $\delta K_\alpha = K_\alpha - K$, $\delta m_i = m_i - m$ and $\gamma = M/S$. As noted in the main text, the basic idea of cavity method is to relate an ecosystem with $M + 1$ resources (variables) and $S + 1$ species (inequality constraints) to that with M resources and S species. Following eq. (17) and eq. (18), one can write down the ecological model for the $(M+1, S+1)$ system where resource R_0 and species N_0 are introduced to the (M, S) system as:

$$\frac{dN_0}{dt} = N_0 \left\{ \mu \langle R \rangle - m + \sum_\beta \sigma_c d_{0\beta} R_\beta - \delta m_0 \right\} \quad (19)$$

$$\frac{dR_0}{dt} = K + \delta K_0 - \left[\omega_0 + \gamma^{-1} \mu \langle N \rangle + \sum_j \sigma_c d_{j0} N_j \right] R_0 \quad (20)$$

Perturbations in cavity solution

Following the same procedure as in [25], we introduce the following susceptibilities:

$$\chi_{\alpha\beta}^R = - \frac{\partial \bar{R}_\alpha}{\partial \omega_\beta} \quad (21)$$

$$\chi_{i\alpha}^N = - \frac{\partial \bar{N}_i}{\partial \omega_\alpha} \quad (22)$$

$$\nu_{\alpha i}^R = \frac{\partial \bar{R}_\alpha}{\partial m_i} \quad (23)$$

$$\nu_{ij}^N = \frac{\partial \bar{N}_i}{\partial m_j} \quad (24)$$

where we denote $\bar{X}$ as the steady-state value of X . Recall that the goal is to derive a set of self-consistency equations that relates the ecological system characterized by $M + 1$ resources (variables) and $S + 1$ species (constraints) to that with the new species and new resources removed: $(S + 1, M + 1) \rightarrow (S, M)$. To simplify notation, let $\bar{X}_{/0}$ denote the steady-state value of quantity X in the absence of the new resource and new species. Since the introduction of a new species and resource represents only a small (order $1/M$) perturbation to the original ecological system, we can express the steady-state species and resource abundances in the $(S + 1, M + 1)$ system with a first-order Taylor expansion around the (S, M) values. We note that the new terms $\sigma_c d_{i0} R_0$ in Eq. eq. (18) and $\sigma_c d_{0\alpha} N_0$ in eq. (17) can be treated as perturbations to m_i , and K_α , respectively, yielding:

$$\bar{N}_i = \bar{N}_{i/0} - \sigma_c \sum_{\beta/0} \chi_{i\beta}^N d_{0\beta} \bar{N}_0 - \sigma_c \sum_{j/0} \nu_{ij}^N d_{j0} \bar{R}_0 \quad (25)$$

$$\bar{R}_\alpha = \bar{R}_{\alpha/0} - \sigma_c \sum_{\beta/0} \chi_{\alpha\beta}^R d_{0\beta} \bar{N}_0 - \sigma_c \sum_{j/0} \nu_{\alpha j}^R d_{j0} \bar{R}_0 \quad (26)$$

Note $\sum_{j/0}$ and $\sum_{\beta/0}$ mean the sum excludes the new species 0 and the new resource 0. The next step is to plug eq. (25) and eq. (26) into eq. (19) and eq. (20) and solve for the steady-state value of N_0 and R_0 .

Self-consistency equations for species

For the new cavity species, the steady equation takes the form

$$0 = \bar{N}_0 \left[\mu \langle R \rangle - m - \sigma_c^2 \bar{N}_0 \sum_{\alpha/0, \beta/0} \chi_{\alpha\beta}^R d_{0\alpha} d_{0\beta} - \sigma_c^2 \bar{R}_0 \sum_{\beta/0, j/0} \nu_{\beta j}^R d_{0\beta} d_{0j} + \sum_{\beta/0} \sigma_c d_{0\beta} \bar{R}_{\beta/0} + \sigma_c d_{00} \bar{R}_0 - \delta m_0 \right] \quad (27)$$

Notice that each of the sums in this equation is the sum over a large number of weak correlated random variables, and can therefore be well approximated by Gaussian random variables for large enough M and S . We can calculate the sum of the random variables:

$$\sum_{\beta/0, j/0} \nu_{\beta j}^R d_{0\beta} d_{0j} = \frac{1}{M} \sum_{\beta/0, j/0} \nu_{\beta j}^R \delta_{j0} \delta_{\beta 0} = 0 \quad (28)$$

$$\sum_{\alpha/0, \beta/0} \chi_{\alpha\beta}^R d_{0\alpha} d_{0\beta} = \frac{1}{M} \sum_{\alpha/0, \beta/0} \chi_{\alpha\beta}^R \delta_{\alpha\beta} = \frac{1}{M} \sum_{\alpha} \chi_{\alpha\alpha}^R = \frac{1}{M} \text{Tr}(\chi_{\alpha\beta}^R) = \chi \quad (29)$$

where χ is the average susceptibility. Using these observations about above sums, we obtain

$$0 = \bar{N}_0 \left[\mu \langle R \rangle - m - \sigma_c^2 \chi \bar{N}_0 + \sum_{\beta/0} \sigma_c d_{0\beta} \bar{R}_{\beta/0} + \sigma_c d_{00} \bar{R}_0 - \delta m_0 \right] + \mathcal{O}(M^{-1/2}), \quad (30)$$

Employing the Central Limit Theorem, we introduce an auxiliary Gaussian variable z_N with zero mean and unit variance and rewrite this as

$$\sum_{\beta/0} \sigma_c d_{0\beta} \bar{R}_{\beta/0} + \sigma_c d_{00} \bar{R}_0 - \delta m_0 = z_N \sqrt{\sigma_c^2 q_R + \sigma_m^2}, \quad (31)$$

where q_R is the second moment of the resource distribution,

$$q_R = \frac{1}{M} \sum_{\beta} R_{\beta}^2.$$

We can solve eq. (30) in terms of the quantities just defined:

$$\mu \langle R \rangle - m - \sigma_c^2 \chi \bar{N}_0 + \sqrt{\sigma_c^2 q_R + \sigma_m^2} z_N \leq 0 \quad (32)$$

Inverting this equation one gets the steady state of species

$$\bar{N}_0 = \max \left[0, \frac{\mu \langle R \rangle - m + \sqrt{\sigma_c^2 q_R + \sigma_m^2} z_N}{\sigma_c^2 \chi} \right] \quad (33)$$

which is a truncated Gaussian.

Let $y = \max(0, \frac{a}{b} + \frac{c}{b} z)$, with z being a Gaussian random variable with zero mean and unit variance. Then its j -th moment is given by

$$\langle y^j \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\frac{a}{c}}^{\infty} e^{-\frac{x^2}{2}} \left(\frac{c}{b} x + \frac{a}{b} \right)^j dx \quad (34)$$

$$= \left(\frac{c}{b} \right)^j \frac{1}{\sqrt{2\pi}} \int_{-\frac{a}{c}}^{\infty} e^{-\frac{x^2}{2}} \left(x + \frac{a}{c} \right)^j dx \quad (35)$$

$$= \left(\frac{c}{b} \right)^j w_j \left(\frac{a}{c} \right) \quad (36)$$

here we define $w_j(\frac{a}{c}) = \frac{1}{\sqrt{2\pi}} \int_{-\frac{a}{c}}^{\infty} e^{-\frac{x^2}{2}} \left(x + \frac{a}{c}\right)^j dx$

With this we can easily write down the self-consistency equations for the fraction of non-zero species and resources as well as the moments of their abundances at the steady state:

$$\phi_N = \frac{S^*}{S} = w_0 \left(\frac{\mu \langle R \rangle - m}{\sqrt{\sigma_c^2 q_R + \sigma_m^2}} \right) \quad (37)$$

$$\langle N \rangle = \frac{1}{S} \sum_j N_j = \left(\frac{\sqrt{\sigma_c^2 q_R + \sigma_m^2}}{\sigma_c^2 \chi} \right) w_1 \left(\frac{\mu \langle R \rangle - m}{\sqrt{\sigma_c^2 q_R + \sigma_m^2}} \right) \quad (38)$$

$$q_N = \frac{1}{S} \sum_j N_j^2 = \left(\frac{\sqrt{\sigma_c^2 q_R + \sigma_m^2}}{\sigma_c^2 \chi} \right)^2 w_2 \left(\frac{\mu \langle R \rangle - m}{\sqrt{\sigma_c^2 q_R + \sigma_m^2}} \right) \quad (39)$$

Note that S^* is the number of surviving species at the steady state.

Self-consistency equations for resources

We now derive the equations for the steady-state of the resource dynamics. Inserting eq. (26) into eq. (20) gives:

$$0 = K + \delta K_0 - \bar{R}_0 \left[\omega + \gamma^{-1} \mu \langle N \rangle - \sigma_c^2 \bar{N}_0 \sum_{\beta/0, j/0} \chi_{j\beta}^N d_{j0} d_{0\beta} - \sigma_c^2 \bar{R}_0 \sum_{i/0, j/0} \nu_{ij}^N d_{0i} d_{0j} + \sum_{j/0} \sigma_c d_{j0} \bar{N}_{j/0} + \sigma_c d_{00} \bar{N}_0 + \delta \omega_0 \right] \quad (40)$$

We can simplify the sums by averaging over the random variables:

$$\sum_{\beta/0, j/0} \chi_{j\beta}^N d_{j0} d_{0\beta} = \frac{1}{M} \sum_{\beta/0, j/0} \chi_{j\beta}^N \delta_{j0} \delta_{\beta 0} = 0 \quad (41)$$

$$\sum_{i/0, j/0} \nu_{ij}^N d_{0i} d_{0j} = \frac{1}{M} \sum_{i/0, j/0} \nu_{ij}^N \delta_{ij} = \frac{1}{M} \sum_i \nu_{ii}^N = \frac{1}{M} \text{Tr}(\nu_{ij}^N) = \gamma^{-1} \nu \quad (42)$$

where ν is the average susceptibility. Finally, note that we can write

$$\delta \omega_0 + \sum_j \sigma_c d_{j0} N_j = z_R \sqrt{\gamma^{-1} \sigma_c^2 q_N + \sigma_\omega^2}, \quad (43)$$

where we have introduced another auxiliary Gaussian variable z_R with zero mean and unit variance and q_N is the second moment of the resource distribution defined in eq. (65). Using these observations, we obtain a quadratic expression for the resource.

$$K + \delta K_0 - (\omega_0 + \gamma^{-1} \mu \langle N \rangle + \sqrt{\gamma^{-1} \sigma_c^2 q_N + \sigma_\omega^2} z_R) \bar{R}_0 + \gamma^{-1} \sigma_c^2 \nu \bar{R}_0^2 = 0 \quad (44)$$

Cavity solution: without backreaction

As discussed in the main text, we cannot solve the full resource equations exactly. For this reason, we perform an expansion, as a start, we calculate this equation by setting $\nu = 0$ in the resource equation. This is equivalent in the TAP language of ignoring the backreaction term.

Under this assumption, the quadratic equation for the resource, simply becomes a linear equation that can be re-arranged to give

$$\bar{R}_\alpha = \frac{K + \delta K_\alpha}{\omega + \gamma^{-1} \mu \langle N \rangle + z_R \sqrt{\gamma^{-1} \sigma_c^2 q_N + \sigma_\omega^2}} \quad (45)$$

Assuming the fluctuations in the denominator is small, *i.e.* $\sqrt{\gamma^{-1} \sigma_c^2 q_N + \sigma_\omega^2} \ll \omega + \gamma^{-1} \mu \langle N \rangle$, we can do a first-order Taylor expansion around the mean value and also ignore the coupling term between δK_α and z_R :

$$\bar{R}_\alpha = \frac{K + \delta K_\alpha}{\omega + \gamma^{-1} \mu \langle N \rangle} - \frac{K \sqrt{\gamma^{-1} \sigma_c^2 q_N + \sigma_\omega^2}}{(\omega + \gamma^{-1} \mu \langle N \rangle)^2} z_R \quad (46)$$

With all these approximations, we get the first two moments of the steady-state resource abundance distribution:

$$\langle R \rangle = \frac{K}{\omega + \gamma^{-1}\mu \langle N \rangle} \quad (47)$$

$$q_R = \langle R \rangle^2 + \frac{\sigma_K^2}{(\omega + \gamma^{-1}\mu \langle N \rangle)^2} + \frac{K^2(\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2)}{(\omega + \gamma^{-1}\mu \langle N \rangle)^4} \quad (48)$$

The susceptibility is given by:

$$\chi = - \left\langle \frac{\partial \bar{R}_\alpha}{\partial w_\alpha} \right\rangle = \left\langle \frac{K_\alpha}{(\omega_\alpha + \sum_j c_{j\alpha} \bar{N}_j)^2} + \frac{2K \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2}}{(\omega + \gamma^{-1}\mu \langle N \rangle)^3} z_R \right\rangle = \frac{K}{(\omega + \gamma^{-1}\mu \langle N \rangle)^2} \quad (49)$$

Combined with self-consistency equations for species, we get the full set of :

$$\phi_N = w_0 \left(\frac{\mu \langle R \rangle - m}{\sqrt{\sigma_c^2 q_R + \sigma_m^2}} \right), \quad \chi = \frac{K}{(\omega + \gamma^{-1}\mu \langle N \rangle)^2} \quad (50)$$

$$\langle N \rangle = \left(\frac{\sqrt{\sigma_c^2 q_R + \sigma_m^2}}{\sigma_c^2 \chi} \right) w_1 \left(\frac{\mu \langle R \rangle - m}{\sqrt{\sigma_c^2 q_R + \sigma_m^2}} \right), \quad \langle R \rangle = \frac{K}{\omega + \gamma^{-1}\mu \langle N \rangle} \quad (51)$$

$$q_N = \left(\frac{\sqrt{\sigma_c^2 q_R + \sigma_m^2}}{\sigma_c^2 \chi} \right)^2 w_2 \left(\frac{\mu \langle R \rangle - m}{\sqrt{\sigma_c^2 q_R + \sigma_m^2}} \right), \quad q_R = \langle R \rangle^2 + \frac{\sigma_K^2}{(\omega + \gamma^{-1}\mu \langle N \rangle)^2} + \frac{K^2(\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2)}{(\omega + \gamma^{-1}\mu \langle N \rangle)^4} \quad (52)$$

Cavity solution: with backreaction correction

We start again with the full resource equation:

$$K + \delta K_0 - (\omega_0 + \gamma^{-1}\mu \langle N \rangle + \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2} z_R) \bar{R}_0 + \gamma^{-1}\sigma_c^2 \nu \bar{R}_0^2 = 0 \quad (53)$$

Since $R_0 > 0$ and $\nu < 0$, the solution of eq. (53) gives:

$$R_0 = \frac{\omega + \gamma^{-1}\mu \langle N \rangle + \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2} z_R}{2\gamma^{-1}\sigma_c^2 \nu} - \frac{\sqrt{(\omega + \gamma^{-1}\mu \langle N \rangle + \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2} z_R)^2 - 4\gamma^{-1}\nu\sigma_c^2(K + \delta K_0)}}{2\gamma^{-1}\sigma_c^2 \nu} \quad (54)$$

For the 1st order expansion, we assume $4\gamma^{-1}\nu\sigma_c^2\delta K_0 + 2\sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2} z_R + (\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2) z_R^2 \ll (\omega + \gamma^{-1}\mu \langle N \rangle)^2 - 4\gamma^{-1}\nu\sigma_c^2 K$ and do a 1st order expansion around the mean of the form:

$$\begin{aligned} & \sqrt{(\omega + \gamma^{-1}\mu \langle N \rangle + \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2} z_R)^2 - 4\gamma^{-1}\nu\sigma_c^2(K + \delta K_0)} \\ &= \sqrt{(\omega + \gamma^{-1}\mu \langle N \rangle)^2 - 4\gamma^{-1}\nu\sigma_c^2 K} + \frac{(\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2) z_R^2 + 2(\omega + \gamma^{-1}\mu \langle N \rangle) \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2} z_R - 4\gamma^{-1}\nu\sigma_c^2 \delta K_0}{2\sqrt{(\omega + \gamma^{-1}\mu \langle N \rangle)^2 - 4\gamma^{-1}\nu\sigma_c^2 K}} \end{aligned} \quad (55)$$

Using these expressions, the moments of their abundances at steady state can be calculated yielding:

$$\langle R \rangle = \frac{\omega + \gamma^{-1}\mu \langle N \rangle}{2\gamma^{-1}\sigma_c^2 \nu} - \frac{\sqrt{(\omega + \gamma^{-1}\mu \langle N \rangle)^2 - 4\gamma^{-1}\nu\sigma_c^2 K}}{2\gamma^{-1}\sigma_c^2 \nu} - \frac{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2}{4\gamma^{-1}\sigma_c^2 \nu \sqrt{(\omega + \gamma^{-1}\mu \langle N \rangle)^2 - 4\gamma^{-1}\nu\sigma_c^2 K}} \quad (56)$$

$$\begin{aligned} q_R &= \langle R \rangle^2 + \frac{(\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2)^2 + 8(\gamma^{-1}\nu\sigma_c^2 K)}{2(2\gamma^{-1}\sigma_c^2 \nu)^2 [(\omega + \gamma^{-1}\mu \langle N \rangle)^2 - 4\gamma^{-1}\nu\sigma_c^2 K]} \\ &+ \frac{(\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2) [\sqrt{(\omega + \gamma^{-1}\mu \langle N \rangle)^2 - 4\gamma^{-1}\nu\sigma_c^2 K} - (\omega + \gamma^{-1}\mu \langle N \rangle)]^2}{(2\gamma^{-1}\sigma_c^2 \nu)^2 [(\omega + \gamma^{-1}\mu \langle N \rangle)^2 - 4\gamma^{-1}\nu\sigma_c^2 K]} \end{aligned} \quad (57)$$

From eq. (54),

$$\frac{\partial R_0}{\partial \omega} = \frac{1}{2\gamma^{-1}\sigma_c^2\nu} \left\{ 1 - \frac{\omega + \gamma^{-1}\mu \langle N \rangle + \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2} z_R}{\sqrt{(\omega + \gamma^{-1}\mu \langle N \rangle + \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2} z_R)^2 - 4\gamma^{-1}\nu\sigma_c^2(K + \delta K_0)}} \right\} \quad (58)$$

The term inside the bracket can be expanded as:

$$\begin{aligned} & \frac{\omega + \gamma^{-1}\mu \langle N \rangle + \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2} z_R}{\sqrt{(\omega + \gamma^{-1}\mu \langle N \rangle + \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2} z_R)^2 - 4\gamma^{-1}\nu\sigma_c^2(K + \delta K_0)}} \\ & \approx \frac{\omega + \gamma^{-1}\mu \langle N \rangle + \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2} z_R}{\sqrt{(\omega + \gamma^{-1}\mu \langle N \rangle)^2 - 4\gamma^{-1}\nu\sigma_c^2 K}} \left[1 - \frac{(\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2) z_R^2 + 2(\omega + \gamma^{-1}\mu \langle N \rangle) \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2} z_R - 4\gamma^{-1}\nu\sigma_c^2 \delta K_0}{2(\omega + \gamma^{-1}\mu \langle N \rangle)^2 - 4\gamma^{-1}\nu\sigma_c^2 K} \right] \end{aligned} \quad (59)$$

The susceptibilities are given by averaging eq. (58)

$$\chi = - \left\langle \frac{\partial R}{\partial \omega} \right\rangle \quad (60)$$

$$= - \frac{1}{2\gamma^{-1}\nu\sigma_c^2} \left\{ 1 - \frac{\omega + \gamma^{-1}\mu \langle N \rangle}{\sqrt{(\omega + \gamma^{-1}\mu \langle N \rangle)^2 - 4\gamma^{-1}\nu\sigma_c^2 K}} + \frac{3(\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2)(\omega + \gamma^{-1}\mu \langle N \rangle)}{2[(\omega + \gamma^{-1}\mu \langle N \rangle)^2 - 4\gamma^{-1}\nu\sigma_c^2 K]^{3/2}} \right\} \quad (61)$$

$$\nu = \left\langle \frac{\partial N}{\partial m} \right\rangle = - \frac{\phi_N}{\sigma_c^2 \chi} \quad (62)$$

Combined with self-consistency equations for species, get the full set of 1st order self-consistency equations:

$$\phi_N = w_0 \left(\frac{\mu \langle R \rangle - m}{\sqrt{\sigma_c^2 q_R + \sigma_m^2}} \right) \quad (63)$$

$$\langle N \rangle = \left(\frac{\sqrt{\sigma_c^2 q_R + \sigma_m^2}}{\sigma_c^2 \chi} \right) w_1 \left(\frac{\mu \langle R \rangle - m}{\sqrt{\sigma_c^2 q_R + \sigma_m^2}} \right) \quad (64)$$

$$q_N = \left(\frac{\sqrt{\sigma_c^2 q_R + \sigma_m^2}}{\sigma_c^2 \chi} \right)^2 w_2 \left(\frac{\mu \langle R \rangle - m}{\sqrt{\sigma_c^2 q_R + \sigma_m^2}} \right) \quad (65)$$

$$\langle R \rangle = \frac{\omega + \gamma^{-1}\mu \langle N \rangle}{2\gamma^{-1}\sigma_c^2\nu} - \frac{\sqrt{(\omega + \gamma^{-1}\mu \langle N \rangle)^2 - 4\gamma^{-1}\nu\sigma_c^2 K}}{2\gamma^{-1}\sigma_c^2\nu} - \frac{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2}{4\gamma^{-1}\sigma_c^2\nu\sqrt{(\omega + \gamma^{-1}\mu \langle N \rangle)^2 - 4\gamma^{-1}\nu\sigma_c^2 K}} \quad (66)$$

$$\begin{aligned} q_R &= \langle R \rangle^2 + \frac{(\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2)^2 + 8(\gamma^{-1}\nu\sigma_c^2 \sigma_K)^2}{2(2\gamma^{-1}\sigma_c^2\nu)^2[(\omega + \gamma^{-1}\mu \langle N \rangle)^2 - 4\gamma^{-1}\nu\sigma_c^2 K]} \\ &+ \frac{(\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2)[\sqrt{(\omega + \gamma^{-1}\mu \langle N \rangle)^2 - 4\gamma^{-1}\nu\sigma_c^2 K} - (\omega + \gamma^{-1}\mu \langle N \rangle)]^2}{(2\gamma^{-1}\sigma_c^2\nu)^2[(\omega + \gamma^{-1}\mu \langle N \rangle)^2 - 4\gamma^{-1}\nu\sigma_c^2 K]} \end{aligned} \quad (67)$$

$$\chi = - \frac{1}{2\gamma^{-1}\nu\sigma_c^2} \left\{ 1 - \frac{\omega + \gamma^{-1}\mu \langle N \rangle}{\sqrt{(\omega + \gamma^{-1}\mu \langle N \rangle)^2 - 4\gamma^{-1}\nu\sigma_c^2 K}} + \frac{3(\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2)(\omega + \gamma^{-1}\mu \langle N \rangle)}{2[(\omega + \gamma^{-1}\mu \langle N \rangle)^2 - 4\gamma^{-1}\nu\sigma_c^2 K]^{3/2}} \right\} \quad (68)$$

$$\nu = - \frac{\phi_N}{\sigma_c^2 \chi} \quad (69)$$

Comparison between with and without backreaction

We can reduce the cavity solution with backreaction to the simpler one when σ_c is large. In fact all the complexity of cavity solution with backreaction comes from the expression for eq. (54):

$$R_0 = \frac{\omega + \gamma^{-1}\mu \langle N \rangle + \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2} z_R}{2\gamma^{-1}\sigma_c^2\nu} - \frac{\sqrt{(\omega + \gamma^{-1}\mu \langle N \rangle + \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2} z_R)^2 - 4\gamma^{-1}\nu\sigma_c^2(K + \delta K_0)}}{2\gamma^{-1}\sigma_c^2\nu} \quad (70)$$

However, if we assume $(\omega + \gamma^{-1}\mu\langle N \rangle + \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2 z_R})^2 \gg -4\gamma^{-1}\nu\sigma_c^2(K + \delta K_0)$, we can expand the second term following $\sqrt{1-x} \approx 1 - \frac{x}{2} - \frac{x^2}{8} + \mathcal{O}(x^3)$.

$$R_0 = \frac{K + \delta K_0}{\omega + \gamma^{-1}\mu\langle N \rangle + \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2 z_R}} + \frac{\gamma^{-1}\sigma_c^2 \nu (K + \delta K_0)^2}{(\omega + \gamma^{-1}\mu\langle N \rangle + \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2 z_R})^3} \quad (71)$$

The first term of above equation is the cavity solution without backreaction.

Comparing the cavity solutions to numerical simulations

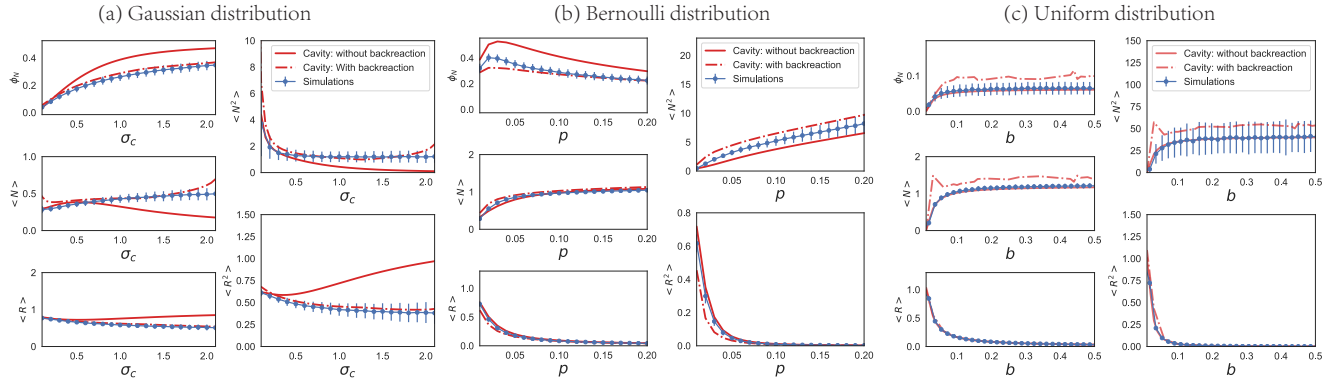


FIG. 6. Comparison of numerics and cavity solutions with and without the backreaction term as a function of σ_c . $\phi_N = \frac{S^*}{S}$ is the fraction of surviving species. $\langle N \rangle$, $\langle N^2 \rangle$, $\langle R \rangle$ and $\langle R^2 \rangle$ are the first and second moments of the species and resources distribution respectively. The simulations details can be found at the Appendix . $\mathbf{C}$ is sampled either from a Gaussian, Bernoulli, or uniform distribution as indicated.

We show a comparison between theoretical and numerical results for different choices of how to sample the consumption matrix in Fig. 2 and Fig. 6. These figures show that the cavity solution with backreaction performs better for the Gaussian and Bernoulli cases. However, in the uniform case, the cavity solution without backreaction matches with numerical simulations perfectly, while the cavity solution with backreaction performs worse than without backreaction. In the section , we have shown the cavity solution with backreaction can be reduced to the cavity solution without backreaction and hence should be a more robust solution. So why does it perform badly in the uniform case? The reason is that in the uniform case $\mu = Mb/2 \gg 1$ when the system size M is large, leading to $|\chi| \sim \frac{1}{(\omega + \gamma^{-1}\mu\langle N \rangle)^2} \ll 1$. From eqs. (64, 65), we see that both $\langle N \rangle$ and $\langle N^2 \rangle$ depends on $\frac{1}{\chi} \gg 1$ and the numerical solver becomes unstable.

Simulation Details

Parameters

All simulations are done with the CVXPY package[30] in PYTHON 3. All codes are available on GitHub at <https://github.com/Emergent-Behaviors-in-Biology/species-packing-bound>.

- Fig. 2: the consumer matrix $\mathbf{C}$ is sampled from the Gaussian distribution $\mathcal{N}(\frac{\mu}{M}, \frac{\sigma_c}{\sqrt{M}})$. $S = 100$, $M = 100$, $\mu = 1$, $K = 1$, $\sigma_K = 0.1$, $m = 1$, $\sigma_m = 0.1$, $\omega = 1$, $\sigma_\omega = 0$ and each data point is averaged from 1000 independent realizations. We only provide the cavity solution with backreaction here.
- Fig. 3 and Fig. 7: the consumer matrix $\mathbf{C}$ is sampled from the Gaussian distribution $\mathcal{N}(\frac{\mu}{M}, \frac{\sigma_c}{\sqrt{M}})$. $S = 500$, $M = 100$, $\mu = 1$, $\sigma_K = 0.1$, $m = 1$, $\sigma_m = 0.1$, $\omega = 1$, $\sigma_\omega = 0$ for externally supplied resource dynamics and $S = 500$, $M = 100$, $\mu = 1$, $\sigma_K = 0.1$, $m = 1$, $\sigma_m = 0.1$, $\tau = 1$, $\sigma_\tau = 0$ for the self-renewing one. Each data point is averaged from 1000 independent realizations. For Fig. 7, $K = 10$.

- Fig. 4: the consumer matrix $\mathbf{C}$ is sampled from the Bernoulli distribution $Bernoulli(p)$ and p are fixed to 0.1, 0.2 and 0.1. m_i follows metabolic tradeoffs Eq. (79) with $\sigma_\epsilon = 0$, $\tilde{m} = 1$. We also set $S = 500$, $M = 100$, $K = 10$, $\sigma_K = 0.1$. Each data point is averaged from 100 independent realizations.
- Fig. 6(a): the simulation is the same as Fig. 2. We show both the cavity solutions with and without reaction here.
- Fig. 6(b): the consumer matrix $\mathbf{C}$ is sampled from the Bernoulli distribution $Bernoulli(p)$. $S = 100$, $M = 100$, $K = 1$, $\sigma_K = 0.1$, $m = 1$, $\sigma_m = 0.1$, $\omega = 1$, $\sigma_\omega = 0$ and each data point is averaged from 1000 independent realizations. The cavity solution is obtained by approximating the Bernoulli distribution to the corresponding Gaussian distribution *i.e.* $\mu = pM$, $\sigma_c = \sqrt{Mp(1-p)}$
- Fig. 6(c): the consumer matrix $\mathbf{C}$ is sampled from the uniform distribution $\mathcal{U}(0, b)$. $S = 100$, $M = 100$, $K = 1$, $\sigma_K = 0.1$, $m = 1$, $\sigma_m = 0.1$, $\omega = 1$, $\sigma_\omega = 0$ and each data point is averaged from 1000 independent realizations. The cavity solution is obtained by approximating the uniform distribution to the corresponding Gaussian distribution, *i.e.* $\mu = bM/2$, $\sigma_c = b\sqrt{M/12}$.
- Fig. 8(a): the consumer matrix $\mathbf{C}$ is sampled from the Bernoulli distribution $Bernoulli(p)$. $S = 500$, $M = 100$, $K = 1$, $\sigma_K = 0.1$, $m = 1$, $\sigma_m = 0.1$, $\omega = 1$, $\sigma_\omega = 0$ and each data point is averaged from 1000 independent realizations. The cavity solution is obtained by approximating the Bernoulli distribution to the corresponding Gaussian distribution *i.e.* $\mu = pM$, $\sigma_c = \sqrt{Mp(1-p)}$
- Fig. 8(b): the consumer matrix $\mathbf{C}$ is sampled from the uniform distribution $\mathcal{U}(0, b)$. $S = 500$, $M = 100$, $K = 1$, $\sigma_K = 0.1$, $m = 1$, $\sigma_m = 0.1$, $\omega = 1$, $\sigma_\omega = 0$ and each data point is averaged from 1000 independent realizations. The cavity solution is obtained by approximating the uniform distribution to the corresponding Gaussian distribution, *i.e.* $\mu = bM/2$, $\sigma_c = b\sqrt{M/12}$.

Distinction between extinct and surviving species

In the main text, we show that the value of species packing $\frac{S^*}{M}$ for the externally supplied resources must be smaller than 0.5. However, in numerical simulations, even for the extinct species the abundance is never exactly equal 0 due to numerical errors. As a result, we must choose a threshold to distinguish extinct and surviving species in order to calculate S^* . Since we are using the equivalence with convex optimization to solve the generalized consumer-resource models[24, 26], we can easily choose a reasonable threshold (e.g. 10^{-2} in Fig. 7) since the extinct and surviving species are well separated in two peaks (see Fig. 7).

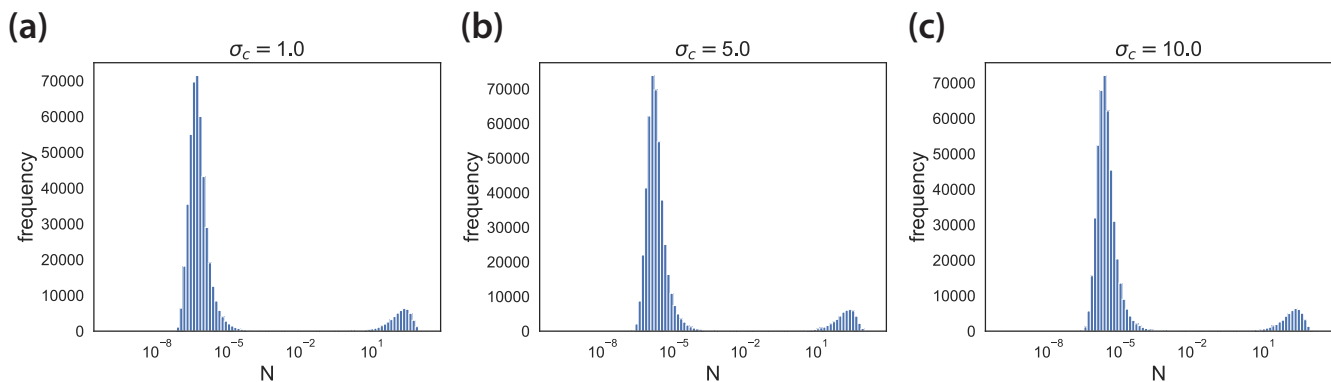


FIG. 7. Species abundance N in equilibrium at different σ_c for externally supplied resource dynamics at $K = 10$. The simulations parameters can be found at the Appendix: Simulation Details.

An upper bound for species packing

By analyzing the susceptibilities in the full Cavity solutions, an upper bound for species packing can be derived for both resource dynamics in GCRMs. The derivations can also be extended to the case where metabolic tradeoffs

impose hard or soft constraints on the parameter values.

Externally supplied resource dynamics

The response functions χ and ν can be written as:

$$\chi = -\frac{1}{2\gamma^{-1}\sigma_c^2\nu} \left\{ 1 - \left\langle \frac{\omega + \gamma^{-1}\mu \langle N \rangle + \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2} z_R}{\sqrt{(\omega + \gamma^{-1}\mu \langle N \rangle + \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2} z_R)^2 - 4\gamma^{-1}\nu\sigma_c^2(K + \delta K_0)}} \right\rangle \right\} \quad (72)$$

$$\nu = -\frac{\phi_N}{\sigma_c^2\chi} \quad (73)$$

Substituting eq. (73) into eq. (72) and rearranging yields

$$\gamma^{-1}\phi_N = \frac{1}{2} \left\{ 1 - \left\langle \frac{\omega + \gamma^{-1}\mu \langle N \rangle + \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2} z_R}{\sqrt{(\omega + \gamma^{-1}\mu \langle N \rangle + \sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_\omega^2} z_R)^2 - 4\gamma^{-1}\nu\sigma_c^2(K + \delta K_0)}} \right\rangle \right\}. \quad (74)$$

The numerator of the term in angle brackets is the total depletion rate for a given resource when it is first added to the system. Depletion rates are always positive in this model, so the right-hand side is always less than 1/2. Noticing $\gamma = \frac{M}{S}$, $\phi_N = S^*/S$, $\chi > 0$, we immediately obtain an upper bound on $\frac{S^*}{M}$:

$$\frac{1}{2} > \frac{S^*}{M}. \quad (75)$$

Self-renewing(MacArthur's) resource dynamics

Using the analytical expressions χ , ν and self-consistent equations in ref. [6], we can derive the following expressions:

$$\langle N \rangle = \left(\frac{\sqrt{\sigma_c^2 q_R + \sigma_m^2}}{\sigma_c^2(\phi_R - \gamma^{-1}\phi_N)} \right) w_1 \left(\frac{\mu \langle R \rangle - m}{\sqrt{\sigma_c^2 q_R + \sigma_m^2}} \right), \quad \langle R \rangle = \left(\frac{\sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_K^2}}{\phi_R(\phi_R - \gamma^{-1}\phi_N)^{-1}} \right) w_1 \left(\frac{\kappa - \gamma^{-1}\mu \langle N \rangle}{\sqrt{\gamma^{-1}\sigma_c^2 q_N + \sigma_K^2}} \right) \quad (76)$$

To derive bounds, we consider various limits of these expressions. First, consider the case were we put many species $S \rightarrow \infty$ into the ecosystem with fixed number of resources M , (i.e $\gamma = \frac{M}{S} \rightarrow 0$). In order to keep $\langle N \rangle$ positive, we must have $\phi_R - \gamma^{-1}\phi_N > 0$, giving an upper bound:

$$1 \geq \frac{M^*}{M} > \frac{S^*}{M} \quad (77)$$

Externally supplied resources with metabolic tradeoffs

Here we consider two kinds of constraints on the parameters, encoding metabolic tradeoffs. In the first, the maintenance cost $m_i = m$ is the same for all species, and the sum of the consumption preferences is constrained to equal some fixed “enzyme budget” E that is nearly the same for all species:

$$\sum_{\alpha} C_{i\alpha} = E + \delta E_i \quad (78)$$

where δE_i is a small random variable with mean zero and variance σ_E^2 . A hard constraint can be generated by taking $\sigma_E = 0$.

The second kind of constraint does not make any assumptions about $C_{i\alpha}$, but assigns a cost $\tilde{m}$ to every unit of consumption capacity, so that

$$m_i = (1 + \epsilon_i)\tilde{m} \sum_{\alpha} C_{i\alpha} + \delta m_i \quad (79)$$

where ϵ_i and δm_i are small random variables with mean zero and variances σ_ϵ^2 and σ_m^2 , respectively. A hard constraint can be generated by taking $\sigma_\epsilon = \sigma_m = 0$.

In the simplest way of setting up the first constraint, the equilibrium equations actually reduce to the same form as the second. Specifically, one usually generates a consumer preference matrix satisfying the constraint by first generating an i.i.d. matrix $\tilde{C}_{i\alpha}$, and then setting $C_{i\alpha} = (E + \delta E_i)\tilde{C}_{i\alpha}/\sum_\beta \tilde{C}_{i\beta}$. The resulting dynamics can be written as:

$$\frac{dN_i}{dt} = N_i \left[\sum_\alpha (E + \delta E_i) \frac{\tilde{C}_{i\alpha}}{\sum_\beta \tilde{C}_{i\beta}} R_\alpha - m_i \right] \quad (80)$$

$$= \frac{N_i(E + \delta E_i)}{\sum_\beta \tilde{C}_{i\beta}} \left[\sum_\alpha \tilde{C}_{i\alpha} R_\alpha - m \frac{\sum_\beta \tilde{C}_{i\beta}}{E + \delta E_i} \right]. \quad (81)$$

Dropping the tilde's, we can write the equilibrium condition in the same form that results from the second kind of constraint:

$$0 = N_i \left\{ \sum_\alpha C_{i\alpha} [R_\alpha - (1 + \epsilon_i)\tilde{m}] - \delta m_i \right\} \quad (82)$$

with

$$\tilde{m} = \frac{m}{E} \quad (83)$$

$$\epsilon_i = -\frac{\delta E_i}{E} \quad (84)$$

$$\delta m_i = 0. \quad (85)$$

Inspection of Equation 82 immediately reveals an important novelty: now when we add a new resource as part of the cavity protocol, the perturbation to the growth rate can either be positive or negative, depending on the sign of $[R_\alpha - (1 + \epsilon_i)\tilde{m}]$. This turns out to be the crucial factor that prevents the proof of the $S^*/M < 1/2$ bound from going through, regardless of the size of σ_ϵ or σ_m .

Following the same steps as above, we arrive at the following set of equilibrium conditions for the new species N_0 and resource R_0 :

$$0 = \bar{N}_0 [\mu\langle R \rangle - \mu\tilde{m} + \sigma_N z_N - \sigma_c^2 \chi \bar{N}_0] \quad (86)$$

$$0 = K + \delta K_0 - (\omega + \gamma^{-1}\mu\langle N \rangle + \sigma_R z_R + \gamma^{-1}\sigma_c^2 \nu \tilde{m}) \bar{R}_0 + \gamma^{-1}\sigma_c^2 \nu \bar{R}_0^2 \quad (87)$$

where

$$\sigma_N^2 = \sigma_m^2 + \sigma_c^2 [q_R - 2\tilde{m}\langle R \rangle + \tilde{m}^2(1 + \sigma_\epsilon^2)] \quad (88)$$

$$\sigma_R^2 = \sigma_\omega^2 + \gamma^{-1}\sigma_c^2 q_N + \gamma^{-2}\sigma_c^4 \nu^2 \tilde{m}^2 \sigma_\epsilon^2. \quad (89)$$

These are nearly identical to the equations we had before. The two key changes are the presence of a term with a negative sign inside the coefficient σ_N of the random variable z_N , and the $\gamma^{-1}\sigma_c^2 \nu \tilde{m}$ term inside the parentheses in the equation for the resources.

We can now proceed in the same way as before, solving for $\bar{N}_0$ and $\bar{R}_0$ and taking derivatives to compute the susceptibilities. We find:

$$\chi = -\frac{1}{2\gamma^{-1}\sigma_c^2 \nu} \left\{ 1 - \left\langle \frac{\omega + \gamma^{-1}\mu\langle N \rangle + \sigma_R z_R + \gamma^{-1}\sigma_c^2 \nu \tilde{m}}{\sqrt{(\omega + \gamma^{-1}\mu\langle N \rangle + \sigma_R z_R + \gamma^{-1}\sigma_c^2 \nu \tilde{m})^2 - 4\gamma^{-1}\sigma_c^2 \nu (K + \delta K_0)}} \right\rangle \right\} \quad (90)$$

$$\nu = -\frac{\phi_N}{\sigma_c^2 \chi} \quad (91)$$

This is almost the same as the expression in Equation (72) obtained in the absence of constraints, except for the extra term $\gamma^{-1}\sigma_c^2 \nu \tilde{m}$ in the numerator and denominator. This term is significant because ν is a negative number, and if its absolute value is large enough, it can make the whole term in angle brackets negative. Inserting the second equation into the first, we obtain a formula for S^*/M :

$$\frac{S^*}{M} = \gamma^{-1}\phi_N = \frac{1}{2} \left\{ 1 - \left\langle \frac{\omega + \gamma^{-1}\mu\langle N \rangle + \sigma_R z_R + \gamma^{-1}\sigma_c^2 \nu \tilde{m}}{\sqrt{(\omega + \gamma^{-1}\mu\langle N \rangle + \sigma_R z_R + \gamma^{-1}\sigma_c^2 \nu \tilde{m})^2 - 4\gamma^{-1}\sigma_c^2 \nu (K + \delta K_0)}} \right\rangle \right\} \quad (92)$$

The term in brackets can now be negative, but is always greater than -1. We thus obtain the bound:

$$\frac{S^*}{M} < 1. \quad (93)$$

The term approaches -1 in the limit $\nu \rightarrow -\infty$, which is the same limit required to saturate the bound in the model with self-renewing resources. As in that case, the limit cannot actually be achieved, because $\nu \rightarrow -\infty$ implies $\chi \rightarrow 0$ (Equation (91)), and χ appears in the denominator of the final expression for $\tilde{N}_0$ (Equation (33)), while the numerator always remains finite.

The only way to achieve the limit $\frac{S^*}{M} = 1$ is to make the numerator vanish in the same way as the denominator, which can only happen in the presence of hard constraints $\sigma_m = \sigma_e = 0$. In this case, it is easy to see that setting $R_\alpha = \tilde{m}$ for all α and $\chi \rightarrow 0, \nu \rightarrow -\infty$ solves both the steady state equations, regardless of the value of $\tilde{N}_0$. In Equation (86) for $\tilde{N}_0$, the mean and the fluctuating part inside the brackets both vanish individually ($\mu\langle R \rangle - \mu\tilde{m} = 0$, $\sigma_N = 0$), and the back-reaction term also vanishes ($\sigma_e^2 \chi \tilde{N}_0 = 0$), leaving the equation trivially satisfied. In Equation (87) only the terms with ν are significant in this limit, and they cancel each other perfectly. This is the “shielded phase” discussed in [23].

Note also that if we take the $\chi \rightarrow 0, \nu \rightarrow -\infty$ limit first, before performing any substitutions, Equations (90) and (91) are satisfied independently of the choice of ϕ_N . This means that $\gamma^{-1}\phi_N = S^*/M$ can be greater than 1, as observed in the simulations of [22].

Numerical evidence

We show a comparison between the cavity solution and numerical results in Fig. 3 and Fig. 8 for three different distributions of the consumption matrix $\mathbf{C}$. For the Gaussian and Bernoulli distributions, $\frac{S^*}{M}$ can reach the upper bound we derived for two different resource dynamics. For externally supplied resource dynamics, $\frac{S^*}{M}$ never exceeds 0.5. For the uniform case, since the fluctuation of consumption matrix is small, the niche overlap is large and there is fierce competitions among species and these ecosystems live very far from the upper bounds we derive. However, even for the uniform case, the species packing fraction is significantly larger for self-renewing resource dynamics than externally supplied resource dynamics.

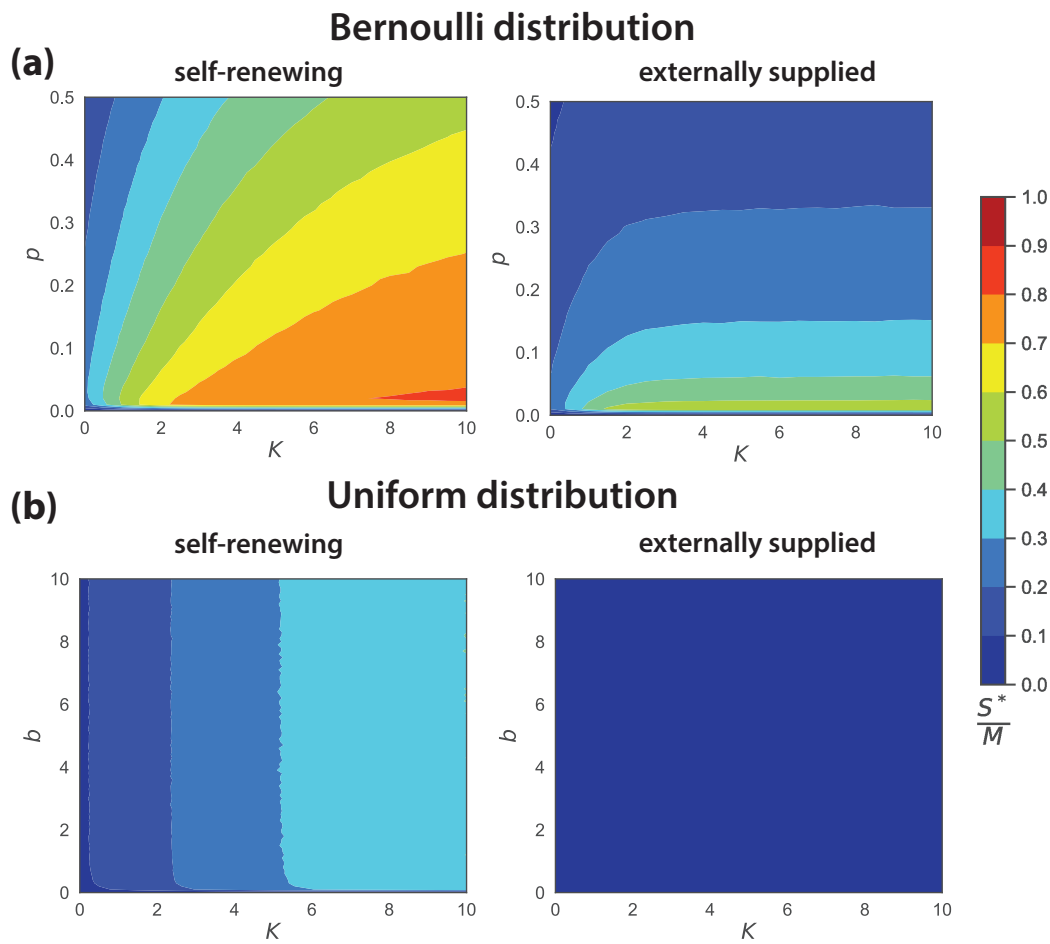


FIG. 8. Comparison of species packing $\frac{S^*}{M}$ for different distributions of consumption matrices $\mathbf{C}$ with self-renewing and externally-supplied resource dynamics. The simulations represent averages from 1000 independent realizations with the system size $M = 100$, $S = 500$ and parameters at the Appendix .