Path-integral solution of MacArthur's resource-competition model for large ecosystems with random species-resources couplings

(Article begins on next page)

12 June 2023

Path-integral solution of MacArthur's resource-competition model for large ecosystems with random species-resources couplings

A. R. Batista-Tomás, Andrea De Martino and Roberto Mulet

View Online   **Export Citation**   **CrossMark**

ARTICLES YOU MAY BE INTERESTED IN

Noise-driven topological changes in chaotic dynamics
Chaos: An Interdisciplinary Journal of Nonlinear Science **31**, 103115 (2021); https://doi.org/10.1063/5.0059461
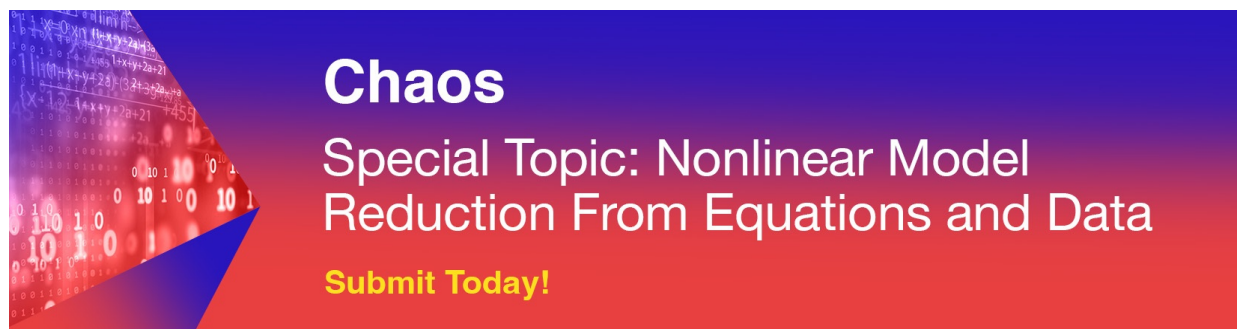
Chaos–hyperchaos transition in three identical quorum-sensing mean-field coupled ring oscillators
Chaos: An Interdisciplinary Journal of Nonlinear Science **31**, 103112 (2021); https://doi.org/10.1063/5.0056907

Study of system dynamics through recurrence analysis of regular windows
Chaos: An Interdisciplinary Journal of Nonlinear Science **31**, 103116 (2021); https://doi.org/10.1063/5.0036505

Chaos
Special Topic: Nonlinear Model Reduction From Equations and Data
Submit Today!

31, 103113

AIP Publishing

Path-integral solution of MacArthur's resource-competition model for large ecosystems with random species-resources couplings

A. R. Batista-Tomás,[1] (ID) Andrea De Martino,[2,3,a)] (ID) and Roberto Mulet[1] (ID)

AFFILIATIONS

[1] Group of Complex Systems and Statistical Physics, Department of Applied Physics, Physics Faculty, University of Havana, La Habana 10400, Cuba
[2] Soft and Living Matter Lab, Institute of Nanotechnology (CNR-NANOTEC), Consiglio Nazionale delle Ricerche, Rome 00185, Italy
[3] Italian Institute for Genomic Medicine, Turin 10060, Italy

[a)] Author to whom correspondence should be addressed: andrea.demartino@roma1.infn.it

ABSTRACT

We solve MacArthur's resource-competition model with random species-resource couplings in the "thermodynamic" limit of infinitely many species and resources using dynamical path integrals à la De Domincis. We analyze how the steady state picture changes upon modifying several parameters, including the degree of heterogeneity of metabolic strategies (encoding the preferences of species) and of maximal resource levels (carrying capacities), and discuss its stability. Ultimately, the scenario obtained by other approaches is recovered by analyzing an effective one-species-one-resource ecosystem that is fully equivalent to the original multi-species one. The technique used here can be applied for the analysis of other model ecosystems related to the version of MacArthur's model considered here.

Mathematical models of ecosystems have repeatedly proved useful to understand how species survival, global stability, and responses to perturbations are controlled by the parameters governing the interactions between species and resources. Following Wigner's recipe, large ecosystems with extended, complicated, and unknown interaction networks can be usefully modeled by assuming quenched random species-resources couplings. In such cases, the statistical mechanics of disordered systems provides tools to calculate macroscopic properties as averages over network realizations. Here, we apply one such method, based on dynamical generating functionals, to MacArthur's resource-competition model, a classic model of an ecosystem in which different species compete for a pool of resources.

I. INTRODUCTION

Despite the complicated structure and dynamics of underlying interactions, microbial ecosystems generate robust statistical outcomes that can now be experimentally probed at genomic resolution.[1–4] To understand the origins and richness of these features and to derive testable predictions, a variety of mathematical models have been considered over time, ranging from coarse-grained ones based on simple differential equations to metabolically realistic schemes capable of accounting for time- and space-dependencies.[5–9] Statistical mechanics approaches are particularly well suited to analyze large instances of such models and derive macroscopic laws, especially when the complex web of intraspecific and trophic interactions can be modeled by quenched random variables.

Among the schemes that have attracted attention, MacArthur's resource-competition model plays a central role as a theoretical reference frame in view of its flexibility, phenomenological richness, and the possibility of studying highly diverse communities with quantitative detail.[10,11] The emergent statistical properties of several versions of MacArthur's model have recently been studied by approaches rooted in the theory of disordered systems, including replica theory and the cavity method (see, e.g., Refs. 12–21), to address features as diverse as the role of heterogeneity of species and resources for the stability of large ecosystems,[12,13] the emergence of robust population structures,[15,20] the impact of different kinds of resource supply dynamics,[19] or the success of

different eco-evolutionary strategies of invasion in small vs large ecosystems.[18]

In this work, we contribute to this line of studies by solving MacArthur's model using dynamical mean field theory (DMFT).[22,23] The key advantage of this method over replicas lies in the fact that it does not rely on specific properties of interactions (e.g., symmetry) or on the existence of a Lyapunov function of the dynamics. On the other hand, dynamical path integrals allow to treat such cases in a more straightforward manner compared to the cavity method. This makes DMFT more broadly applicable, at least in principle. In addition, DMFT reduces the dynamics of the multi-species, multi-resources system to a pair of processes involving one effective species and one effective resource, whose structure renders explicit the role of parameters that only impact the original dynamics implicitly.

Besides retrieving the picture derived by other means, more precisely by the cavity method,[17] we will analyze the stability properties, the role of resource heterogeneity for the survival and numerosity of species, and how a species' survival probability is modulated by its prior fitness (i.e., by the fitness it would face in that environment in absence of other competing species).

II. MACARTHUR'S MODEL

For this paper, MacArthur's model is defined as, e.g., in Ref. 17. We consider a system in which N species interact through M shared resources. If n_i denotes the population size of species i ($i = 1, \ldots, N$) and r^μ denotes the level of resource μ ($\mu = 1, \ldots, M$), its time evolution is described by the coupled equations for their respective growth rates g_i and g^μ ($\dot{x} = \frac{dx}{dt}$),

$$g_i \equiv \frac{\dot{n}_i}{n_i} = \sum_\mu c_i^\mu r^\mu - m_i, \tag{1}$$

$$g^\mu \equiv \frac{\dot{r}^\mu}{r^\mu} = a^\mu (K^\mu - r^\mu) - \sum_j c_j^\mu n_j, \tag{2}$$

where c_i^μ measures the contribution of resource μ to the growth rate (fitness) of species i, m_i is a maintenance level for species i (growth only occurs if the benefit from resources exceeds this threshold), a^μ is the intrinsic growth rate of resource μ and K^μ is the carrying capacity of resource μ. In short, in such a setting, each resource is externally supplied so that in the absence of species, it undergoes logistic growth with rate a^μ and saturation level K^μ. When species are present, they predate on resources with fitness benefits described by the coefficients c_i^μ.

Some basic properties of MacArthur's model are easily derived from (1) and (2) under the simplifying assumption that resources equilibrate much faster than populations. Indeed, in this case, one can replace r^μ in (1) with its steady state value obtained by imposing $g^\mu = 0$. Equation (1) can then be re-cast as

$$g_i = -\frac{\partial L}{\partial n_i}, \tag{3}$$

with

$$L = \frac{1}{2} \sum_\mu' \frac{1}{a^\mu} \left(\sum_j n_j c_j^\mu - K^\mu a^\mu \right)^2 + \sum_i m_i n_i, \tag{4}$$

where the prime indicates that the sum runs over resources whose level is not zero ("non-depleted" for short) at stationarity. As $\dot{L} = -\sum_i (\dot{n}_i)^2 / n_i \leq 0$, L is a (convex) Lyapunov function of the dynamics and its unique minimum describes the steady state population sizes (and in turn resource levels). At stationarity, in particular, the population sizes of surviving species obey the conditions

$$\sum_j J_{ij} n_j + h_i = 0, \tag{5}$$

where $J_{ij} = \sum_\mu' c_i^\mu c_j^\mu / a^\mu$ and $h_i = m_i - \sum_\mu' c_i^\mu K^\mu$. Because the rank of the matrix with elements J_{ij} is at most equal to the number M_s of non-depleted resources, system (5) contains at most M_s independent equations. The number N_s of surviving species [i.e., of variables in (5)], therefore, satisfies $N_s \leq M_s$. This implies that the fractions ϕ_s and ψ_s of surviving species and resources, respectively, are related by

$$\phi_s \leq \frac{\psi_s}{\nu}, \quad \nu = \frac{N}{M}. \tag{6}$$

We shall see that this bound provides a quantitatively accurate description of the relationship between ϕ_s and ψ_s even in the general case in which the timescales of (1) and (2) are not widely separate.

To model complex interdependencies of species on resources in extended ecosystems, it is normally assumed that N and M are large and that c_i^μs are quenched iid random variables. More specifically, the statistical mechanics approach studies the statistical properties emerging when $N, M \to \infty$ at fixed ν [see (6)] and c_i^μs are quenched iid random variables with mean c/N and variance σ_c^2/N. σ_c^2 is an especially important parameter here, as it provides a proxy for the metabolic heterogeneity of species. In the linear approximation and within the same assumptions leading to (6), setting $a^\mu = 1$ for simplicity, one easily sees that a perturbation δn_i of a steady state population $\bar{n}_i > 0$ for the MacArthur model evolves as

$$\dot{\delta n_i} = -\bar{n}_i \sum_j \mathbf{K}_{ij} \delta n_j, \quad \mathbf{K}_{ij} = \sum_\mu' c_i^\mu c_j^\mu. \tag{7}$$

Stability for the above system is governed by the smallest eigenvalue of the random matrix $\{\mathbf{K}_{ij}\}$, which can be calculated using results from Ref. 24. One finds

$$\lambda_{\min} = \sigma_c \left(\sqrt{\frac{\psi_s}{\nu}} - \sqrt{\phi_s} \right). \tag{8}$$

This formula can provide qualitative information about the linear stability of the ecosystem's dynamics. In agreement with (6), $\lambda_{\min}$ vanishes when $\nu = \psi_s / \phi_s$. Moreover, as both ϕ_s and ψ_s are bound to change with σ_c, (8) implies the possibility of a non-trivial dependence of $\lambda_{\min}$ on species heterogeneity. This issue will be explored more thoroughly in what follows.

Besides that encoded in the coefficients c_i^μ, other sources of heterogeneity can be accounted for via randomness in other parameters. For instance, to model diverse availability of resources, K^μs can be taken to be quenched iid random variables with mean K and variance σ_K^2. An important role in our results will be played by the

"prior fitness" of species i, defined as

$$f_{i,0} \equiv \sum_{\mu=1}^{M} c_i^\mu K^\mu - m_i, \qquad (9)$$

in terms of which (1) takes the form

$$g_i = f_{i,0} - \sum_\mu c_i^\mu (K^\mu - r^\mu). \qquad (10)$$

This clarifies how the participation in resource competition affects the *a priori* viability of species i. $f_{i,0}$ measures the fitness that a species would have if it was introduced in the environment at maximal resource levels and in the absence of other species. It is reasonable to expect that $f_{i,0}$ correlates positively with the survival probability also in a complex ecosystem with N species. This was shown numerically in Ref. 17. We shall further explore this issue here by deriving an approximate quantitative relationship linking these quantities.

III. DYNAMICAL GENERATING FUNCTION AND EFFECTIVE PROCESSES

The path-integral approach to the system formed by (1) and (2) is based on the computation of the dynamical generating function

$$Z[\boldsymbol{\psi}, \boldsymbol{\phi}] \equiv \overline{\left\langle e^{i \sum_i \int dt\, \psi_i(t) n_i(t) + i \sum_\mu \int dt\, \phi^\mu(t) r^\mu(t)} \right\rangle_{\text{paths}}}, \qquad (11)$$

where the quantities $\boldsymbol{\psi} = \{\psi_i\}$ and $\boldsymbol{\phi} = \{\phi^\mu\}$ represent auxiliary fields associated, respectively, to the variables $\mathbf{n} = \{n_i\}$ and $\mathbf{r} = \{r^\mu\}$, the over-bar denotes an average over quenched disorder (most notably the coefficients c_i^μ), while the brackets denote an average over the realizations ("paths") of the dynamics of the system at fixed disorder, which includes possibly random initial conditions. The key advantage of Z lies in the fact that the moments of the variables n_i and r^μ can be written in terms of derivatives of Z over the auxiliary fields. For instance, the mean values of n_i and r^μ (overpaths and disorder) are given by

$$\overline{\langle n_i(t) \rangle_{\text{paths}}} = -i \lim_{\substack{\boldsymbol{\psi} \to 0 \\ \boldsymbol{\phi} \to 0}} \frac{\partial Z}{\partial \psi_i(t)}, \qquad (12)$$

$$\overline{\langle r^\mu(t) \rangle_{\text{paths}}} = -i \lim_{\substack{\boldsymbol{\psi} \to 0 \\ \boldsymbol{\phi} \to 0}} \frac{\partial Z}{\partial \phi^\mu(t)}. \qquad (13)$$

Notice that these expressions are valid at any time t. As shown in the Appendix, the explicit calculation of the average over disorder leads to the identification of a set of macroscopic order parameters whose averages provide a full characterization of the N-species, M-resources system. These parameters include the mean population sizes and resource levels, i.e.,

$$\rho_n(t) = \frac{1}{N} \sum_{i=1}^{N} n_i(t), \qquad (14)$$

$$\rho_r(t) = \frac{1}{N} \sum_{\mu=1}^{M} r^\mu(t), \qquad (15)$$

and the two-time correlation functions

$$Q_n(t,t') = \frac{1}{N} \sum_{i=1}^{N} n_i(t) n_i(t'), \qquad (16)$$

$$Q_r(t,t') = \frac{1}{N} \sum_{\mu=1}^{M} r^\mu(t) r^\mu(t'). \qquad (17)$$

In view of the connection between Z and moments of n_i and r^μ, it is clear that Z allows for a full statistical description of the dynamics.

Ultimately, Z can be evaluated in the limit $N, M \to \infty$ (with fixed $\nu = N/M$) via a saddle-point method following which the full dynamics is re-cast in terms of a pair of effective stochastic processes, one for an effective species with population size $n(t)$, the other for an effective resource with level $r(t)$ (see the Appendix for the derivation). Such processes read

$$\frac{\dot{n}(t)}{n(t)} = c \rho_r(t) - m - \frac{\sigma_c^2}{\nu} \int dt'\, G_r(t,t') n(t') + \theta(t) + \xi_n(t), \quad (18)$$

$$\frac{\dot{r}(t)}{r(t)} = a(K - r(t)) - c \rho_n(t) - \sigma_c^2 \int dt'\, G_n(t,t') r(t') + \eta(t) + \xi_r(t), \qquad (19)$$

where m is the maintenance level of the effective species, a and K denote, respectively, the rate of growth and the carrying capacity of the effective resource, ξ_n and ξ_r are zero-average Gaussian random variables with

$$\langle \xi_n(t) \xi_n(t') \rangle = \sigma_c^2 Q_r(t,t'), \qquad (20)$$

$$\langle \xi_r(t) \xi_r(t') \rangle = \sigma_c^2 Q_n(t,t') + a^2 \sigma_K^2, \qquad (21)$$

$\theta(t)$ and $\eta(t)$ are auxiliary probing fields, and G_r and G_n are interaction kernels describing the response functions, i.e.,

$$G_n(t,t') = \left\langle \frac{\partial n(t)}{\partial \theta(t')} \right\rangle_\star, \qquad (22)$$

$$G_r(t,t') = \left\langle \frac{\partial r(t)}{\partial \eta(t')} \right\rangle_\star, \qquad (23)$$

with $\langle \cdot \rangle_\star$ denoting an average over realizations of (18) and (19).

The above equations clarify how disorder affects the dynamics in the limit of large ecosystems. Heterogeneity in metabolic strategies (i.e., $\sigma_c^2 > 0$) translates into a term that couples species and resources along the entire dynamical trajectory. Hence, quenched random c_i^μs generate long-term memory in the dynamics despite the fact that the model defined by (1) and (2) is Markovian. For $\sigma_K^2 = 0$ and large enough σ_c^2, moreover, the memory term dominates over the noise term, as the strength of the latter is proportional to σ_c. On the other hand, heterogeneity in carrying capacities (i.e., σ_K^2) only affects the strength of the noise term in (19). Therefore, when the resource dynamics is dominated by noise in Ks, we should expect to see that resource levels change in an effectively random way. As we shall see below in detail, such a variability bears a non-trivial impact on the abundance of species at steady state. Notice that, as σ_c increases, negative values of c_i^μ, corresponding to resources having a

detrimental effect on species i (e.g., toxic compounds), will become more frequent. Likewise, as σ_K increases some resources will have a "negative carrying capacity," corresponding in effect to a sink that drains those resources from the environment.

It is instructive to study the behavior of (18) at small times. Noting that $G_r(0,0) = 0$ by causality and that $Q_r(0,0) = d$ (constant), one sees that the quantity $g_n(0) = \dot{n}(0)/n(0)$ is a Gaussian variable with mean $c\rho_r(0) - m$ and variance $d\sigma_c^2$. Moreover, by virtue of (15), one has $\rho_r(0) = r_0/v$, where r_0 denotes the mean of $r(0)$ over the distribution of initial conditions of (19). Therefore, on average, the effective species initially increases its population if

$$r_0 > \frac{mv}{c}. \tag{24}$$

As might have been expected, the initial growth rate is more likely positive in less competitive ecosystems, for species with less demanding maintenance requirements and/or when metabolic strategies are more efficient. More precisely, the probability that $g_n(0) > 0$ is given by $\pi_0 \equiv \text{Prob}\left\{\xi_n(0) > m - \frac{cr_0}{v}\right\}$, i.e.,

$$\pi_0 = \frac{1}{2}\,\text{erfc}\,\frac{vm - cr_0}{v\sigma_c\sqrt{2d}}. \tag{25}$$

Hence, in highly competitive ecosystems ($v \gg 1$), $\pi_0 \simeq \frac{1}{2}\,\text{erfc}\,\frac{m}{\sigma_c\sqrt{2d}}$, implying that a stronger metabolic heterogeneity yields a higher initial fitness.

IV. STATIONARY STATE

In the steady state defined by $\dot{n} = \dot{r} = 0$, we have $n(t) \to \underline{n}$ and $r(t) \to \underline{r}$ while two-time quantities are bound to be time-translation invariant, e.g.,

$$\lim_{t \to \infty} Q_n(t + \tau, \tau) = Q_n(\tau), \tag{26}$$

with finite means, average correlations, and integrated responses, i.e. ($x = n, r$),

$$\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau dt\, \rho_x(t) = \bar{\rho}_x, \tag{27}$$

$$\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau dt\, Q_x(t) = q_x, \tag{28}$$

$$\lim_{\tau \to \infty} \int_0^\tau dt\, G_x(t) = \chi_x. \tag{29}$$

Using these properties and definitions, one shows, for instance, that

$$\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau dt \int_0^t dt'\, G_r(t - t')n(t')$$
$$= \int_0^\infty ds\, G_r(s) \lim_{\tau \to \infty} \frac{1}{\tau} \int_s^\tau dt\, n(t - s)$$
$$= \int_0^\infty ds\, G_r(s) \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^{\tau - s} ds'\, n(s')$$
$$= \int_0^\infty ds\, G_r(s) \lim_{\tau \to \infty} \frac{\tau - s}{\tau}\, \frac{1}{\tau - s} \int_0^{\tau - s} ds'\, n(s')$$
$$= \chi_r \underline{n}. \tag{30}$$

Hence, at stationarity, (18) and (19) are easily seen to imply

$$c\bar{\rho}_r - \frac{\sigma_c^2}{v}\chi_r \underline{n} + \sigma_c\sqrt{q_r}\, z = m, \tag{31}$$

$$\underline{r}\left(a + \sigma_c^2\chi_n\right) = aK - c\bar{\rho}_n + \sqrt{\sigma_c^2 q_n + a^2\sigma_K^2}\, z, \tag{32}$$

with z a Gaussian random variable with zero mean and unit variance. (The probing fields θ and η can be set to zero, noting that differentiating with respect to them is equivalent to differentiating with respect to z, modulo constant factors.) The above equations coincide with those obtained in Ref. 17 by the cavity method. Since $\underline{n}, \underline{r} \geq 0$, the steady state population size and resource level are given by

$$\underline{n} \equiv \underline{n}(z) = \frac{c\bar{\rho}_r - m + \sigma_c\sqrt{q_r}z}{\sigma_c^2\chi_r/v}\,\Theta(z - z_n), \tag{33}$$

$$\underline{r} \equiv \underline{r}(z) = \frac{aK - c\bar{\rho}_n + \sqrt{\sigma_c^2 q_n + a^2\sigma_K^2}\,z}{a + \sigma_c^2\chi_n}\,\Theta(z - z_r), \tag{34}$$

where $\Theta(x)$ is the step function while

$$z_n \equiv \frac{m - c\bar{\rho}_r}{\sigma_c\sqrt{q_r}}, \tag{35}$$

$$z_r \equiv \frac{c\bar{\rho}_n - aK}{\sqrt{\sigma_c^2 q_n + a^2\sigma_K^2}}. \tag{36}$$

Equations (33) and (34) finally allow to compute macroscopic parameters from

$$\bar{\rho}_n \equiv \langle n \rangle_\star = \langle \underline{n} \rangle_z, \tag{37}$$

$$\bar{\rho}_r \equiv \frac{1}{v}\langle r \rangle_\star = \frac{1}{v}\langle \underline{r} \rangle_z, \tag{38}$$

$$q_n \equiv \langle n^2 \rangle_\star = \langle \underline{n}^2 \rangle_z, \tag{39}$$

$$q_r \equiv \frac{1}{v}\langle r^2 \rangle_\star = \frac{1}{v}\langle \underline{r}^2 \rangle_z, \tag{40}$$

$$\chi_n = \frac{1}{\sigma_c\sqrt{q_r}}\left\langle \frac{\partial \underline{n}}{\partial z} \right\rangle_z, \tag{41}$$

$$\chi_r = \frac{1}{\sigma_c\sqrt{q_n}}\left\langle \frac{\partial \underline{r}}{\partial z} \right\rangle_z. \tag{42}$$

After some algebra, one ends up with the set of equations

$$\bar{\rho}_n = \frac{v\sqrt{q_r}}{\sigma_c\chi_r}\left[\frac{e^{-\frac{z_n^2}{2}}}{\sqrt{2\pi}} - \frac{z_n}{2}\,\text{erfc}\left(\frac{z_n}{\sqrt{2}}\right)\right], \tag{43}$$

$$\bar{\rho}_r = \frac{\sqrt{\sigma_c^2 q_n + a^2\sigma_K^2}}{v(a + \sigma_c^2\chi_n)}\left[\frac{e^{-\frac{z_r^2}{2}}}{\sqrt{2\pi}} - \frac{z_r}{2}\,\text{erfc}\left(\frac{z_r}{\sqrt{2}}\right)\right], \tag{44}$$

$$q_n = \frac{\nu^2 q_r}{\sigma_c^2 \chi_r^2} \left[\frac{1}{2}(1 + z_n^2)\, \mathrm{erfc}\left(\frac{z_n}{\sqrt{2}}\right) - \frac{z_n e^{-\frac{z_n^2}{2}}}{\sqrt{2\pi}} \right], \qquad (45)$$

$$q_r = \frac{\sigma_c^2 q_n + a^2 \sigma_K^2}{\nu(a + \sigma_c^2 \chi_n)^2} \left[\frac{1}{2}(1 + z_r^2)\, \mathrm{erfc}\left(\frac{z_r}{\sqrt{2}}\right) - \frac{z_r e^{-\frac{z_r^2}{2}}}{\sqrt{2\pi}} \right], \quad (46)$$

$$\chi_n = \frac{\nu}{2\sigma_c^2 \chi_r}\, \mathrm{erfc}\left(\frac{z_n}{\sqrt{2}}\right), \qquad (47)$$

$$\chi_r = \frac{1}{2(a + \sigma_c^2 \chi_n)}\, \mathrm{erfc}\left(\frac{z_r}{\sqrt{2}}\right), \qquad (48)$$

which can be solved numerically for any choice of the parameters ν, c, a, σ_c, and σ_K. In turn, one can obtain the fractions of surviving species and non-depleted resources from

$$\phi_s = \frac{1}{2}\, \mathrm{erfc}\left(\frac{z_n}{\sqrt{2}}\right), \qquad (49)$$

$$\psi_s = \frac{1}{2}\, \mathrm{erfc}\left(\frac{z_r}{\sqrt{2}}\right). \qquad (50)$$

The distributions of population sizes ($p(n)$) and resource levels ($p(r)$) can instead be found by noting that $\underline{n} = \bar{n} + \nu_n z$ and $\underline{r} = \bar{r} + \nu_r z$, with

$$\bar{n} = \frac{c\bar{\rho}_r - m}{\sigma_c^2 \chi_r / \nu}, \qquad \nu_n = \frac{\nu \sqrt{q_r}}{\sigma_c \chi_r}, \qquad (51)$$

$$\bar{r} = \frac{aK - c\bar{\rho}_n}{a + \sigma_c^2 \chi_n}, \qquad \nu_r = \frac{\sqrt{\sigma_c^2 q_n + a^2 \sigma_K^2}}{a + \sigma_c^2 \chi_n}. \qquad (52)$$

This implies that

$$p(n) = \mathcal{G}(\bar{n}, \nu_n^2) \langle \Theta(z - z_n) \rangle_z + \delta(n) \langle \theta(z_n - z) \rangle_z, \qquad (53)$$

$$p(r) = \mathcal{G}(\bar{r}, \nu_r^2) \langle \Theta(z - z_r) \rangle_z + \delta(r) \langle \theta(z_r - z) \rangle_z, \qquad (54)$$

where $\mathcal{G}(x, y)$ denotes the normal distribution with mean x and variance y (see Fig. 1; notice that abundance distributions quantitatively different from truncated Gaussians emerge when factors that are absent in our model, like space and dispersal, are explicitly accounted for; we refer the reader to Ref. 25 for a specific discussion of some of these aspects). In the following, we shall explore the solutions obtained with different choices for the various parameters. First, we shall set $\sigma_K = 0$ (i.e., $K^\mu = K$ for all μ) to focus on the role of metabolic diversity. Next, we shall consider the case of heterogeneous resources ($\sigma_K > 0$). In all cases, $a = 1$ and $m = 1$ for simplicity.

V. RESULTS

A. Survival probability

Numerical solutions of the saddle-point equations can be directly compared with results from computer simulations of (1)
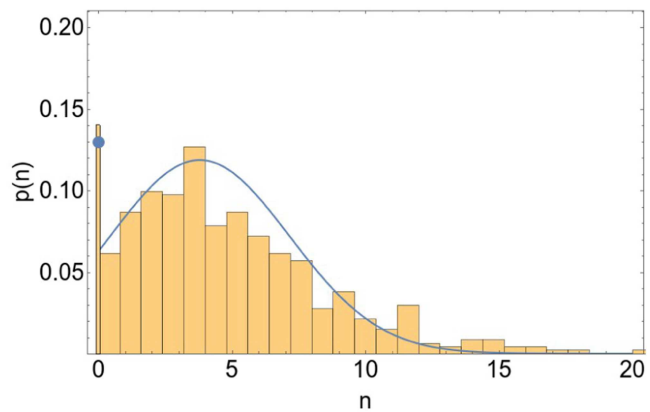

FIG. 1. Representative example of the distribution $p(n)$ of population sizes for $\nu = 0.22$, $K = 5$, $\sigma_c^2 = a = m = 1$, and $\sigma_K^2 = 0$. The solid line is the prediction obtained from (53), the blue marker denoting the weight of the peak at $n = 0$ (i.e., the fraction of extinct species). Bars represent instead the histogram of the population sizes of 10^3 species obtained by averaging the steady state populations of 100 independent realizations of MacArthur's model (1) and (2).

and (2). Figures 2(a)–2(d) show how the fraction of surviving species varies with $\nu = N/M$ for different choices of σ_c^2 and K, with all other parameters fixed. Expectedly, ϕ_s generically decreases as the species-to-resources ratio ν increases and the ecosystem gets more competitive. The bound to ϕ_s given in (6) is, however, more efficiently saturated for larger values of ν, K, and σ_c. In other terms, for any given maximal resource capacity, a higher metabolic diversity allows for a more efficient packing of species into the ecosystem. (The naïve bound $\phi_s = 1/\nu$, where the number of surviving species equals that of resources, is also shown for comparison.)

Notice that metabolic diversity (σ_c^2) appears to impact ϕ_s differently at high and low ν, as a higher σ_c seems to confer higher survival probability only in more competitive ecosystems. This can also be seen in Fig. 3(a). In less competitive ecosystems (smaller ν), the survival probability decreases as σ_c increases due to the stronger impact of inefficient metabolic strategies (for any fixed c, larger values of σ_c imply that negative values of c_i^μ are sampled more and more often). In other terms, the dynamics for small ν favors species that can receive a (possibly small) fitness benefit from all resources (low values of σ) over ones that receive a (possibly large) fitness benefit only from a subset of resources (higher σ). In more competitive systems (high ν), such a trend is reversed, indicating that the ability to extract higher benefits from some resources provides a competitive advantage despite the costs imposed by metabolic inefficiencies (negative values of c_i^μ).

Based on these results, one expects that ϕ_s will increase upon increasing c at fixed σ_c^2 at low ν, i.e., when metabolic strategies become more efficient but more similar. This is indeed the case, as shown in Fig. 3(b) (red markers). When, however, the increase in c is accompanied by an increase of σ_c so that σ_c/c remains fixed (implying that strategies become more diverse as they get more efficient on average), one observes an increase of ϕ_s followed by a decrease at larger c [Fig. 3(b), green markers]. In other words, in less competitive scenarios and for any fixed relative
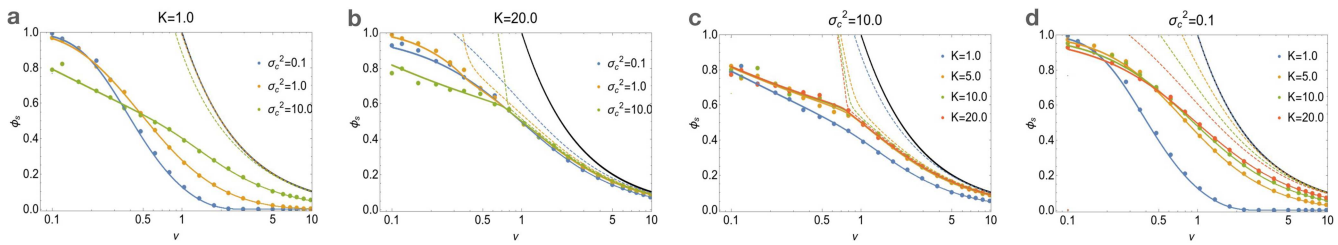
FIG. 2. Fraction of surviving species ϕ_s (survival probability) as a function of ν for constant $c = 1, a = 1, m = 1$ and fixed $K = 1$ (panel a), $K = 20$ (b), $\sigma_c^2 = 10$ (c), and $\sigma_c^2 = 0.1$ (d) and different values of σ_c^2 or K as indicated. Markers are averages over the steady states of 100 independent realizations of Eqs. (1) and (2), continuous lines correspond to theoretical predictions, while dashed lines represent the bounds (6) in each case. Black line is the naïve bound $\phi_s = 1/\nu$.

variability of metabolic strategies, there exists a value of the average efficiency that maximizes the survival probability. Increases in efficiency at fixed σ_c/c, therefore, have opposite effects at low and high c. In the latter regime (metabolic strategies very efficient on average), the detrimental effect of larger diversity dominates the behavior of ϕ_s. By contrast, when metabolic efficiency is small, improvements in this specific quantity are the key determinants of ϕ_s. In competitive ecosystems (large ν), instead, an increase in

efficiency yields a survival advantage at fixed diversity σ_c, so that species can improve their survival probability by either increasing diversity [Fig. 3(a)] or increasing efficiency at fixed σ_c [Fig. 3(b), orange markers]. When σ_c/c is kept fixed, though, increased efficiency does not substantially improve ϕ_s [Fig. 3(b), blue markers].

B. Population sizes

The more efficient species packing found at large ν however comes at a cost for populations. Mean population sizes indeed decrease as ν increases when packing is near optimal [Figs. 4(a) and 4(b)]. This implies that, in such a regime, the introduction of new species feeds back negatively on the typical species population size. By contrast, in less competitive ecosystems (smaller ν), surviving species benefit from the introduction of new species, as their mean population sizes increase with ν. Population sizes also generically increase with increased resource capacity [larger K, Fig. 4(a)], while increased heterogeneity tends to reduce population sizes both in the less competitive regime (low ν) and when competition is stronger and survival probability is lowest [Fig. 4(b)]. This contrasts with the fact that higher σ_c yields a positive impact (albeit weak) on the survival probability in a crowded ecosystem. To summarize, in large random instances of MacArthur's model, population sizes benefit from metabolic innovation (i.e., from the introduction of new species) when competition is weaker, while they benefit from the discovery of new resources (i.e., from a decrease in ν) when the competition is strongest (larger ν).
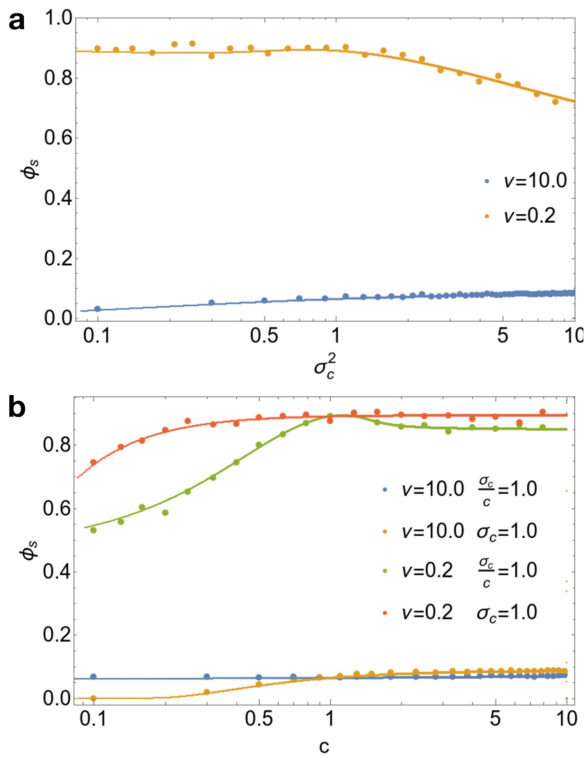

FIG. 3. Fraction of surviving species ϕ_s for different values of ν as a function of σ_c^2 with $c = 1$ (panel a) and as a function of c for different choices of σ_c (panel b). In both panels, $K = 5, a = 1, m = 1$. Markers are averages over the steady states of 100 independent realizations of Eqs. (1) and (2), while continuous lines correspond to theoretical predictions.

FIG. 4. Mean population size ρ_n as a function of ν for different values of K and $\sigma_c^2 = 10.0$ (panel a) and different values of σ_c^2 and $K = 20$ (panel b). In both figures, $c = 1, a = 1, m = 1$. Markers are averages over the steady states of 100 independent realizations of Eqs. (1) and (2), while continuous lines correspond to theoretical predictions.
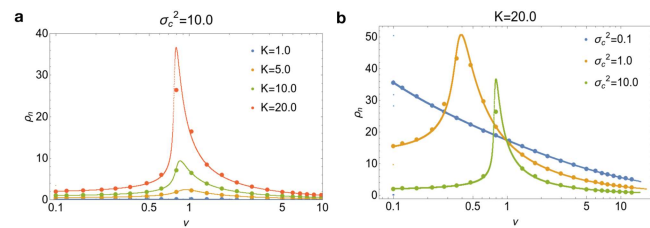
C. Role of the prior fitness of species

The dynamical approach employed for (1) and (2) can be easily applied to the system defined by (10) and (2). The effective process now takes the form

$$\frac{\dot{n}}{n} = f_0 - c\overline{\rho}_r(t) - \frac{\sigma_c^2}{\nu} \int dt' G_r(t,t')n(t') + \theta(t) + \xi_n(t), \quad (55)$$

$$\frac{\dot{r}}{r} = a(K - r(t)) - c\rho_n(t) - \sigma_c^2 \int dt' G_n(t,t') \left[r(t') - K\right]$$
$$+ \eta(t) + \xi_r(t), \quad (56)$$

where f_0 is a Gaussian random variable with mean $\frac{cK}{\nu} - m$ and variance $\frac{\sigma_c^2 K^2}{\nu}$, while

$$\left\langle \xi_n(t)\xi_n(t')\right\rangle = \sigma_c^2 Q_r(t,t'), \quad (57)$$

$$\left\langle \xi_r(t)\xi_r(t')\right\rangle = \sigma_c^2 Q_r(t,t') + \sigma_K^2. \quad (58)$$

To distinguish results for this version of the model from those of Secs. I–IV, we shall denote the steady state quantities by an extra index $^\circ$ (e.g., $\underline{n}^\circ$ instead of $\underline{n}$). One easily sees that the equivalent of (33) and (34) is now

$$\underline{n}^\circ(z) = \frac{f_0 - c\overline{\rho}_r^\circ + \sqrt{\sigma_c^2 q_r^\circ + \sigma_K^2}\, z}{\sigma_c^2 \chi_r^\circ / \nu} \Theta(z - z_n^\circ), \quad (59)$$

$$\underline{r}^\circ(z) = \frac{K(a + \sigma_c^2 \chi_n) - c\rho_n + \sqrt{\sigma^2 q_n + \sigma_K^2}\, z}{a + \sigma_c^2 \chi_n} \Theta(z - z_r^\circ), \quad (60)$$

where

$$z_n^\circ = \frac{c\overline{\rho}_r^\circ - f_0}{\sigma_c \sqrt{q_r^\circ}}, \quad (61)$$

$$z_r^\circ = \frac{c\overline{\rho}_n^\circ - K(a + \sigma_c^2 \chi_n^\circ)}{\sqrt{\sigma_c^2 q_n^\circ + \sigma_K^2}}. \quad (62)$$

In turn, saddle-point equations read

$$\overline{\rho}_n^\circ = \left\langle \underline{n}^\circ(z)\right\rangle_z, \quad (63)$$

$$\overline{\rho}_r^\circ = \frac{1}{\nu}\left\langle K - \underline{r}^\circ(z)\right\rangle_z, \quad (64)$$

$$q_n^\circ = \left\langle [\underline{n}^\circ(z)]^2\right\rangle_z, \quad (65)$$

$$q_r^\circ = \frac{1}{\nu}\left\langle (K - [\underline{r}^\circ(z)])^2\right\rangle_z, \quad (66)$$

$$\chi_n^\circ = \frac{1}{\sigma_c \sqrt{q_r^\circ}}\left\langle \frac{\partial \underline{n}^\circ(z)}{\partial z}\right\rangle_z, \quad (67)$$

$$\chi_r^\circ = \frac{1}{\sigma_c \sqrt{q_n^\circ}}\left\langle \frac{\partial \underline{r}^\circ(z)}{\partial z}\right\rangle_z. \quad (68)$$

Because these changes do not alter the actual dynamics, the mean resource levels and population sizes must be the same, i.e., we must have

$$\langle \underline{n}\rangle = \langle \underline{n}^\circ\rangle, \quad (69)$$

$$\langle \underline{n}^2\rangle = \langle [\underline{n}^\circ]^2\rangle, \quad (70)$$

$$\langle \underline{r}\rangle = \langle \underline{r}^\circ\rangle, \quad (71)$$

$$\langle \underline{r}^2\rangle = \langle [\underline{r}^\circ]^2\rangle, \quad (72)$$

so that

$$\overline{\rho}_r^\circ = \frac{K}{\nu} - \overline{\rho}_r, \quad (73)$$

$$q_r^\circ = \frac{K^2}{\nu} - 2K\overline{\rho}_r + q_r. \quad (74)$$

Having this in mind, we can write the survival probability as a function of the prior fitness f_0 as

$$\phi_s(f_0) = \frac{1}{2}\,\text{erfc}\left[\frac{1}{\sqrt{2}}\frac{c\left(\frac{K}{\nu} - \overline{\rho}_r\right) - f_0}{\sigma_c \sqrt{\frac{K^2}{\nu} - 2K\overline{\rho}_r + q_r}}\right]. \quad (75)$$

These results are compared with numerical simulations in Fig. 5. Notice that ϕ_s tends to become more and more step-like as the ecosystem becomes more competitive [larger ν, Fig. 5(a)], implying that the fate of species is determined to a greater extent by their prior fitnesses. In less competitive cases, instead, lower values of f_0 are more easily overcome through dynamics. The carrying capacity plays a similar role [Fig. 5(b)]: as K increases and resources become available at higher levels, species that would be less fit *a priori* can face better survival odds; on the other hand, the prior fitness becomes more relevant for the final fate of a species as carrying capacities decrease. Likewise [Fig. 5(c)], larger heterogeneities in metabolic strategies improve the survival probabilities of species with lower prior fitnesses. The dependences on σ_c^2 and ν are represented more completely in the heat maps shown in Figs. 5(d) and 5(e).

D. Stability

Figure 6(a) displays the eigenvalue $\lambda_{\min}$ given in (8) as a function of the metabolic heterogeneity σ_c^2 and for various ν, with all other parameters fixed. While this quantity strictly speaking controls stability when the characteristic timescales of species and resources are widely separated, it can still provide useful indications about how different parameters may affect the stability of MacArthur's model. We focus, in particular, on the metabolic heterogeneity σ_c^2. At low species-to-resources ratios, $\lambda_{\min}$ generically increases with σ_c, indicating that higher metabolic heterogeneity improves stability. At high ν, however, i.e., in more competitive situations, $\lambda_{\min}$ displays a maximum at relatively small values of σ_c^2, indicating that the system becomes less stable when metabolic heterogeneity is too large. One, therefore, understands that $\lambda_{\min}$ also decreases systematically as c increases at fixed σ_c^2, as strategies become more homogeneous (not shown). Likewise, one easily sees from (8) that the introduction
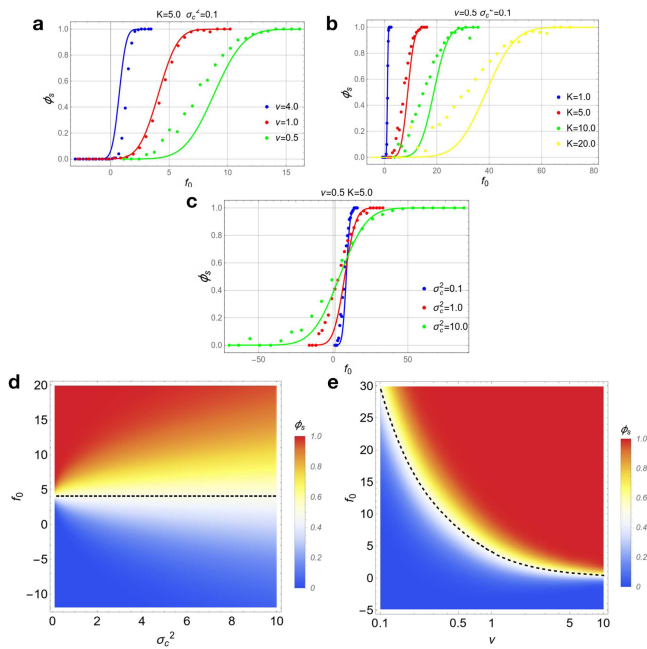
FIG. 5. Panels (a)–(c): Survival probability ϕ_s as a function of f_0 for different values of ν (panel a), K (panel b), and σ_c^2 (panel c), with remaining parameters fixed as indicated. (d) Survival probability ϕ_s as a function of f_0 and σ_c^2 for fixed $K = 5$, $c = 1$, $a = 1$, $m = 1$, and $\nu = 1$. The dashed line at $f_0 = 4$ corresponds to $\phi_s = 1/2$. (e) Survival probability ϕ_s as a function of f_0 and ν for fixed $K = 5$, $c = 1$, $a = 1$, $m = 1$, and $\sigma_c^2 = 1$. The dashed line corresponds to $\phi_s = 1/2$.

FIG. 6. (a) Behavior of $\lambda_{\min}$, Eq. (8), vs σ_c^2 for fixed $K = 5$, $a = 1$, $m = 1$, and different ν. (b) Behavior of $\lambda_{\min}$ (left) and ρ_n (right) vs ν for different values of K (increasing in the direction of the arrow).

of new species (increase ν) always decreases stability, in agreement with the classical results of Ref. 26.

Figure 6(b) links instead the behavior of the mean population size ρ_n for increasing values of K to that of $\lambda_{\min}$. In particular, the value of ν at which ρ_n tends to peak for $K \to \infty$ corresponds to the point where the ecosystem becomes marginally stable ($\lambda_{\min} \to 0$) in the same limit. In addition, it shows that, for a sufficiently large carrying capacity and in presence of strong competition between species (large ν), when surviving species saturate the achievable limit more efficiently, the ecosystem gets as close as possible to becoming unstable. Likewise, at the species-to-resources ratio that allows to sustain the largest overall populations, the ecosystem gets closer and closer to an instability as K increases. Despite the crude assumptions under which it was derived, Eq. (8), therefore, does appear to provide key insight into the properties of the stationary state of the full system.

E. Heterogeneous carrying capacities

To understand the role of diversity in the availability of resources, we studied the effect of disorder in the distribution of K^μ. Specifically, we considered a Gaussian distribution for K with fixed mean and variance σ_K^2 changing in the interval $[0, 20]$. The Gaussian distribution keeps theoretical calculations feasible while introducing negative values of K. The latter can be interpreted as an outflow of resources from the environment (sinks). To simplify the discussion,
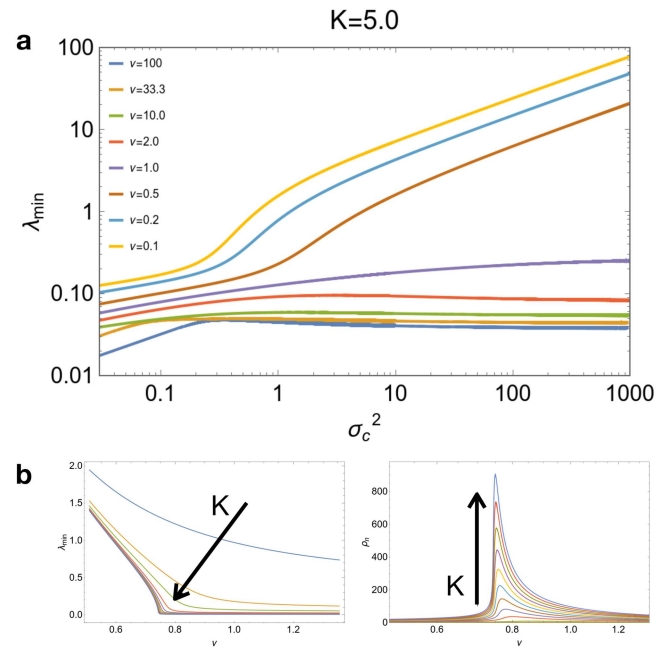
the rest of the parameters of the model were kept fixed at $\sigma_c = a = m = 1$ and we focused on $\nu = 10$ and $\nu = 0.2$ to explore the emerging picture in highly competitive vs noncompetitive situations. Results are summarized in Fig. 7. The survival probability decreases as σ_K^2 increases [Fig. 7(a)], as a consequence of the enhancement of resource depletion [see also Eqs. (36) and (50): as σ_K^2 increases, z_r and, in turn, ψ_s decrease rapidly]. On the other hand, mean population sizes increase markedly as σ_K increases [Fig. 7(b)]. In other words, stronger variability in resource levels (including sinks) has a (small) negative impact on species survival probability but allows for the sustainment of a larger number of individuals. Both of these features follow directly from the analytical solution. According to Eq. (59), the mean population level increases as σ_K increases. On the other hand, so does the mean resource level [see Eq. (60)], which, in turn, causes z_n° [Eq. (61)] to increase as q_r° [Eq. (66)] gets smaller. An increase of z_n° finally implies a worsening of the survival chances, as $\phi_s = \left\langle \Theta(z - z_n^\circ) \right\rangle_z$. The difference between systems with and without large variability in the level of resources is further highlighted in Figs. 7(c) and 7(d). One sees that, for all other parameters fixed, in the presence of heterogeneities in the carrying capacities, prior fitnesses span a much broader range. However, species with the same prior fitness can achieve much larger population sizes in a heterogeneous environment than in a homogeneous one.

VI. DISCUSSION

MacArthur's model provides a versatile theoretical framework to study emergent properties in large complex ecosystems, with
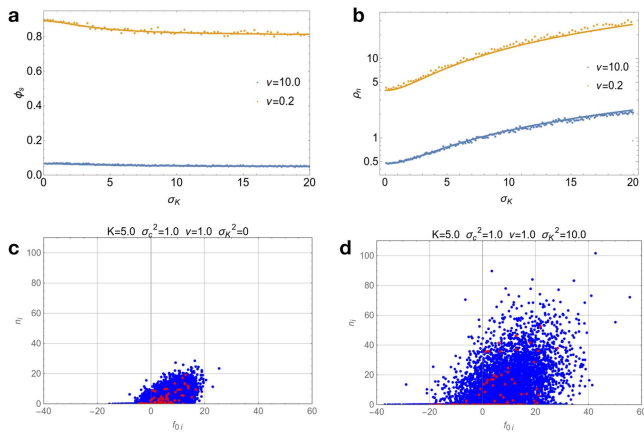
FIG. 7. Dependence of survival probability ϕ_s (panel a) and of the mean population size ρ_n (panel b) on σ_K for $\nu = 10$ (orange) and $\nu = 0.2$ (blue) and $K = 5.0, \sigma_c^2 = a = m = 1$. Markers are averages over the steady states of 100 independent realizations of Eqs. (1) and (2), while continuous lines correspond to theoretical predictions. (c) species population sizes (n_i) vs prior fitness of species ($f_{0,i}$) in the steady state of 100 realizations of Eqs. (1) and (2), with parameters $K = 5, \sigma_c^2 = m = a = c = 1$ and $N = 100, M = 100$, and $\sigma_K^2 = 0$. A single realization is highlighted in red. (d) Same as in (c) but with $\sigma_K^2 = 10$.

much room for improvements that bring the model closer to reality and allow for testable predictions. In analyzing the simplest version, as done here, one mostly aims at defining the set of allowed behaviors and identifying the elementary mechanisms driving the response of the ecosystem to perturbations both at the level of species (e.g., the introduction of new species) or resources (e.g., the disappearance of a viable resource). Understanding how such perturbations affect global features like stability, survival probability, etc. is indeed the focus point of statistical mechanics approaches. To summarize, within MacArthur's model, metabolic heterogeneity far from generically providing an advantage in terms of stability or survival, bears different effects in more vs less competitive situations. In a very competitive scenario, metabolic heterogeneity favors the survival probability, on the contrary, in a less competitive scenario, the survival probability decrease, because metabolic heterogeneity translates essentially into species with lower fitness. Likewise, heterogeneity at the level of resources leads to higher population sizes for smaller numbers of species rather than favoring higher survival probabilities. Still, diversity at steady state in this system is limited by competitive exclusion,[27,28] which in our case takes the form of Eq. (6).

In our view, the key defining aspect of MacArthur's model is given by the fact that the growth rate of individual species is a linear superposition of contributions due to different resources. Such an assumption is unrealistic in many cases, as, e.g., when some resources are essential and the growth rate of a species is positive only when that species has access to at least one of these resources[29] or in presence of a bound in the quantities and quality of the resources that can be consumed.[30] Experimental studies of microbes under co-utilization of different resources have likewise

displayed a more complicated picture.[31] Alternatively, other models capable of overcoming the "niche" scenario exploit dynamical effects,[32,33] and higher-order interactions have also been shown to impact ecosystem diversity in a non-trial way.[34,35] The emergent phenomenology of consumer-resource models in such cases is markedly more complex. To our knowledge, a statistical mechanics treatment of fully disordered cases is still lacking, and techniques like the path-integral method employed here seem to be a proper starting point for the analysis of these models in the limit of large system sizes.

On another level, a more realistic theory would require the integration of a more detailed descriptions of metabolism into the model, possibly along with features like microbial growth laws[21] and cross-feeding[36] or considering directly how the by-products of metabolism shape the environment, leading to secondary interactions between the species.[37] Work along these lines is likely to shed new light about the origin, stability, and reproducibility of species compositions in extended ecosystems.

ACKNOWLEDGMENTS

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

APPENDIX: DERIVATION OF THE EFFECTIVE PROCESSES

We begin by re-writing the dynamics (1) and (2) as

$$\frac{\dot{n}_i}{n_i} - \sum_\mu c_i^\mu r^\mu + m_i - \theta_i = 0, \tag{A1}$$

$$\frac{\dot{r}^\mu}{r^\mu} - a^\mu(K^\mu - r^\mu) + \sum_j c_j^\mu n_j - \theta^\mu = 0, \tag{A2}$$

where we added auxiliary time-dependent probing fields θ_i and θ^μ. Our goal here is to evaluate the dynamical generating function

$$Z[\boldsymbol{\psi}, \boldsymbol{\phi}] \equiv \overline{\left\langle e^{i\sum_i \int dt \psi_i(t) n_i(t) + i\sum_\mu \int dt \phi^\mu(t) r^\mu(t)} \right\rangle_{\text{paths}}} \tag{A3}$$

by explicitly carrying out the averages over trajectories of (A1) and (A2) ("paths") and over the quenched disorder, represented by the coefficients c_i^μ and by the carrying capacities K^μ. Correlations and response functions for population sizes are linked to Z by

$$-i \lim_{\substack{\boldsymbol{\psi}\to 0 \\ \boldsymbol{\phi}\to 0}} \frac{\partial Z}{\partial \psi_i(t)} = \overline{\langle n_i(t)\rangle_{\text{paths}}}, \tag{A4}$$

$$-\mathrm{i}\lim_{\substack{\psi\to 0\\ \phi\to 0}}\frac{\partial^2 Z}{\partial\psi_i(t)\partial\psi_j(t')}=\overline{\langle n_i(t)n_j(t')\rangle_{\text{paths}}}, \qquad (A5)$$

$$-\mathrm{i}\lim_{\substack{\psi\to 0\\ \phi\to 0}}\frac{\partial^2 Z}{\partial\psi_i(t)\partial\theta_j(t')}=\frac{\partial}{\partial\theta_j(t')}\overline{\langle n_i(t)\rangle_{\text{paths}}} \qquad (A6)$$

(with similar formulas for means, correlations and response functions of resource levels using auxiliary fields ϕ and θ^μ). The average over paths can be written explicitly in terms of the equations of motion (A1) and (A2) using appropriate δ-distributions, namely,

$$\langle(\cdots)\rangle_{\text{paths}}=\int D\mathbf{n}\,D\mathbf{r}\,(\cdots)\prod_{i,t}\delta\left(\frac{\dot{n}_i}{n_i}-\sum_\mu c_i^\mu r^\mu+m_i-\theta_i\right)$$

$$\times\prod_{\mu,t}\delta\left(\frac{\dot{r}^\mu}{r^\mu}-a^\mu(K^\mu-r^\mu)+\sum_j c_j^\mu n_j-\theta^\mu\right),$$

where the product over i and t indicates that the condition has to be imposed for all i at all times (and likewise for the product over μ and t). In turn, the δ-functions can be expressed using the Fourier representations $\delta(y)\propto\int_{-\infty}^\infty\exp(\mathrm{i}y\widehat{y})\,d\widehat{y}$. After simple rearrangements, this yields

$$Z[\psi,\phi]=\int D(\mathbf{n},\widehat{\mathbf{n}})D(\mathbf{r},\widehat{\mathbf{r}})\,\underbrace{e^{\mathrm{i}\sum_i\int dt\,\psi_i(t)n_i(t)+\mathrm{i}\sum_i\int dt\,\widehat{n_i}(t)\left(\frac{\dot{n}_i}{n_i}+m_i-\theta_i\right)}}_{\mathcal{A}(\mathbf{n},\widehat{\mathbf{n}})}$$

$$\times\underbrace{e^{\mathrm{i}\sum_\mu\int dt\,\phi^\mu(t)r^\mu(t)+\mathrm{i}\sum_\mu\int dt\,\widehat{r^\mu}(t)\left(\frac{\dot{r}^\mu}{r^\mu}+a^\mu r^\mu-\eta^\mu\right)}}_{\mathcal{B}(\mathbf{r},\widehat{\mathbf{r}})}$$

$$\times\underbrace{\prod_{i,\mu}\overline{e^{\mathrm{i}c_i^\mu\int dt\left[\widehat{r^\mu}(t)n_i(t)-r^\mu(t)\widehat{n_i}(t)\right]}}\prod_\mu\overline{e^{-\mathrm{i}K^\mu\int dt\,a^\mu\widehat{r^\mu}}}}_{\Delta(\mathbf{n},\widehat{\mathbf{n}},\mathbf{r},\widehat{\mathbf{r}})}, \qquad (A7)$$

where we have separated the term $\Delta(\mathbf{n},\widehat{\mathbf{n}},\mathbf{r},\widehat{\mathbf{r}})$ that depends on the quenched random variables c_i^μ and K^μ, which conveniently factorizes. Average are easily computed since the c_i^μs (respectively, K^μs) are iid Gaussian random variables with distributions $\mathcal{G}\left(c/N,\sigma_c^2/N\right)$ (respectively, $\mathcal{G}\left(K,\sigma_K^2\right)$). One gets

$$\Delta(\mathbf{n},\widehat{\mathbf{n}},\mathbf{r},\widehat{\mathbf{r}})=\Delta_1(\mathbf{n},\widehat{\mathbf{n}},\mathbf{r},\widehat{\mathbf{r}})\times e^{-\mathrm{i}\sum_\mu a^\mu K\int dt\,\widehat{r^\mu}(t)}, \qquad (A8)$$

$$\Delta_1(\mathbf{n},\widehat{\mathbf{n}},\mathbf{r},\widehat{\mathbf{r}})=e^{\mathrm{i}cN\int dt[\lambda_r(t)\rho_n(t)-\rho_r(t)\lambda_n(t)]}$$

$$\times\,e^{-\frac{\sigma_c^2 N}{2}\int dt\,dt'\left[L_r(t,t')Q_n(t,t')+Q_r(t,t')L_n(t,t')\right]}$$

$$\times\,e^{\sigma_c^2 N\int dt\,dt'K_r(t,t')K_n(t,t')}\times e^{-\frac{\sigma_K^2 N}{2}(a^\mu)^2\int dt\,dt'L_r(t,t')}, \qquad (A9)$$

where we defined the macroscopic quantities,

$$\rho_n(t)=\frac{1}{N}\sum_i n_i(t),\quad \rho_r(t)=\frac{1}{N}\sum_\mu r^\mu(t),$$

$$\lambda_n(t)=\frac{1}{N}\sum_i\widehat{n_i}(t),\quad \lambda_r(t)=\frac{1}{N}\sum_\mu\widehat{r^\mu}(t),$$

$$Q_n(t,t')=\frac{1}{N}\sum_i n_i(t)n_i(t'),\quad Q_r(t,t')=\frac{1}{N}\sum_\mu r^\mu(t)r^\mu(t'),$$

$$\qquad\qquad (A10)$$

$$L_n(t,t')=\frac{1}{N}\sum_i\widehat{n_i}(t)\widehat{n_i}(t'),\quad L_r(t,t')=\frac{1}{N}\sum_\mu\widehat{r^\mu}(t)\widehat{r^\mu}(t'),$$

$$K_n(t,t')=\frac{1}{N}\sum_i n_i(t)\widehat{n_i}(t'),\quad K_r(t,t')=\frac{1}{N}\sum_\mu r^\mu(t)\widehat{r^\mu}(t').$$

To simplify the notation, we shall henceforth write n_i for $n_i(t)$ and n_i' for $n_i(t')$ (and likewise for all time-dependent quantities). Defining the vectors $\pi=(\rho,\lambda,Q,L,K)$ and $\widehat{\pi}=(\widehat{\rho},\widehat{\lambda},\widehat{Q},\widehat{L},\widehat{K})$, we can insert each of the above definitions into Z again using appropriate δ-functions, obtaining

$$Z=\int D(\pi,\widehat{\pi})\,e^{N\Psi(\pi,\widehat{\pi})+N\Phi(\pi)}$$

$$\times\int D(\mathbf{n},\widehat{\mathbf{n}})\,\mathcal{A}(\mathbf{n},\widehat{\mathbf{n}})\,e^{-\mathrm{i}\sum_i\int dt\left(\widehat{\rho_n}n_i+\widehat{\lambda_n}\widehat{n_i}\right)}\,e^{-\mathrm{i}\sum_i\int dt\,dt'\left[\widehat{Q_n}n_i n_i'+\widehat{L_n}\widehat{n_i}\widehat{n_i'}+\widehat{K_n}n_i\widehat{n_i'}\right]}$$

$$\times\int D(\mathbf{r},\widehat{\mathbf{r}})\left[\mathcal{B}(\mathbf{r},\widehat{\mathbf{r}})\times e^{-\mathrm{i}\sum_\mu a^\mu K\int dt\,\widehat{r^\mu}(t)}\right]e^{-\mathrm{i}\sum_\mu\int dt(\widehat{\lambda_r}\widehat{r^\mu}+\widehat{\rho_r}r^\mu)}\,e^{-\mathrm{i}\sum_\mu\int dt\,dt'\left[\widehat{Q_r}r^\mu(r^\mu)'+\widehat{L_r}\widehat{r^\mu}\widehat{r^\mu{}'}+\widehat{K_r}r^\mu\widehat{r^\mu{}'}\right]}, \qquad (A11)$$

where

$$\Psi(\pi,\widehat{\pi})=\mathrm{i}\int dt\left[\rho_n\widehat{\rho_n}+\rho_r\widehat{\rho_r}+\lambda_n\widehat{\lambda_n}+\lambda_r\widehat{\lambda_r}\right]$$

$$+\mathrm{i}\int dt\,dt'\left[Q_n\widehat{Q_n}+L_n\widehat{L_n}+K_n\widehat{K_n}+Q_r\widehat{Q_r}+L_r\widehat{L_r}+K_r\widehat{K_r}\right], \qquad (A12)$$

$$\Phi(\pi)\equiv\frac{1}{N}\ln\Delta_1=\mathrm{i}c\int dt\left[\lambda_r\rho_n-\rho_r\lambda_n\right]-\frac{\sigma_c^2}{2}\int dt\,dt'\left[Q_n L_r+Q_r L_n-2K_n K_r\right]-\frac{a^2\sigma_K^2}{2}\int dt\,dt'L_r. \qquad (A13)$$

Finally, upon substituting expressions for $\mathcal{A}(\mathbf{n}, \widehat{\mathbf{n}})$ and $\mathcal{B}(\mathbf{r}, \widehat{\mathbf{r}})$ from (A7) into (A11), we arrive at

$$Z = \int D(\boldsymbol{\pi}, \widehat{\boldsymbol{\pi}})\, e^{N[\Psi(\boldsymbol{\pi}, \widehat{\boldsymbol{\pi}}) + \Phi(\boldsymbol{\pi}) + \Omega_n(\widehat{\boldsymbol{\pi}}) + \Omega_r(\widehat{\boldsymbol{\pi}})]}, \tag{A14}$$

with

$$\Omega_n(\widehat{\boldsymbol{\pi}}) = \frac{1}{N} \sum_i \ln \int D(n_i, \widehat{n}_i)\, e^{i\int dt\, \psi_i(t) n_i(t)}\, e^{i\int dt \widehat{n}_i(t)\left(\frac{\dot{n}_i}{n_i} + m_i - \theta_i\right)}$$
$$\times\, e^{-i\int dt\left[\widehat{\rho_n} n_i + \widehat{\lambda_n} \widehat{n}_i\right]}\, e^{-i\int dt\, dt'\left[\widehat{Q_n} n_i n_i' + \widehat{L_n} \widehat{n}_i \widehat{n}_i' + \widehat{K_n} n_i \widehat{n}_i'\right]} \tag{A15}$$

$$\equiv \ln \int D(n, \widehat{n})\, e^{i\int dt \psi(t) n(t)}\, e^{i\int dt \widehat{n}(t)\left(\frac{\dot{n}}{n} + m - \theta\right)}$$
$$\times\, e^{-i\int dt\left[\widehat{\rho_n} n + \widehat{\lambda_n} \widehat{n}\right]}\, e^{-i\int dt\, dt'\left[\widehat{Q_n} n n' + \widehat{L_n} \widehat{n} \widehat{n}' + \widehat{K_n} n \widehat{n}'\right]} \tag{A16}$$

and

$$\Omega_r(\widehat{\boldsymbol{\pi}}) = \frac{1}{N} \sum_\mu \ln \int D(r^\mu, \widehat{r^\mu})\, e^{i\int dt \phi^\mu(t) r^\mu(t)}$$
$$\times\, e^{i\int dt \widehat{r^\mu}(t)\left(\frac{\dot{r}^\mu}{r^\mu} - a^\mu(K^\mu - r^\mu) - \eta^\mu\right)}$$
$$\times\, e^{-i\int dt\left(\widehat{\rho_r} r^\mu + \widehat{\lambda_r} \widehat{r^\mu}\right)}\, e^{-i\int dt\, dt'\left[\widehat{Q_r} r^\mu(r^\mu)' + \widehat{L_r} \widehat{r^\mu} \widehat{r^\mu}' + \widehat{K_r} r^\mu \widehat{r^\mu}'\right]} \tag{A17}$$

$$\equiv \frac{1}{\nu} \ln \int D(r, \widehat{r})\, e^{i\int dt \phi(t) r(t)}\, e^{i\int dt \widehat{r}(t)\left(\frac{\dot{r}}{r} - a(K - r) - \eta\right)}$$
$$\times\, e^{-i\int dt\left(\widehat{\rho_r} r + \widehat{\lambda_r} \widehat{r}\right)}\, e^{-i\int dt\, dt'\left[\widehat{Q_r} r r' + \widehat{L_r} \widehat{r} \widehat{r}' + \widehat{K_r} r \widehat{r}'\right]}. \tag{A18}$$

In the limit $N \to \infty$, integrals like (A14) can be evaluated by saddle-point integration, implying

$$Z = \int D(\boldsymbol{\pi}, \widehat{\boldsymbol{\pi}})\, e^{N[\Psi(\boldsymbol{\pi}, \widehat{\boldsymbol{\pi}}) + \Phi(\boldsymbol{\pi}) + \Omega_n(\widehat{\boldsymbol{\pi}}) + \Omega_r(\widehat{\boldsymbol{\pi}})]} \sim e^{N[\Psi^\star + \Phi^\star + \Omega_n^\star + \Omega_r^\star]}, \tag{A19}$$

where the star indicates that the functions are extremized. The corresponding values of the order parameters are, therefore, to be obtained from the saddle-point conditions

$$\frac{\partial \Psi}{\partial \boldsymbol{\pi}} + \frac{\partial \Phi}{\partial \boldsymbol{\pi}} = 0, \tag{A20}$$

$$\frac{\partial \Psi}{\partial \widehat{\boldsymbol{\pi}}} + \frac{\partial \Omega_N}{\partial \widehat{\boldsymbol{\pi}}} + \frac{\partial \Omega_R}{\partial \widehat{\boldsymbol{\pi}}} = 0. \tag{A21}$$

Computing derivatives explicitly, (A20) takes the form

$$\widehat{\rho_n} = -c\lambda_r, \quad \widehat{\rho_r} = c\lambda_n, \quad \widehat{\lambda_n} = c\rho_r, \quad \widehat{\lambda_r} = -c\rho_n,$$

$$i\widehat{Q_n} = \frac{\sigma_c^2}{2} L_r, \quad i\widehat{Q_r} = \frac{\sigma_c^2}{2} L_n, \quad i\widehat{L_n} = \frac{\sigma_c^2}{2} Q_r, \quad i\widehat{L_r} = \frac{a^2 \sigma_K^2}{2} + \frac{\sigma_c^2}{2} Q_n,$$
$$\tag{A22}$$

$$i\widehat{K_n} = -\sigma_c^2 K_r, \quad i\widehat{K_r} = -\sigma_c^2 K_n,$$

while for (A21), we get

$$\rho_n = \langle n \rangle_{\star,n}, \quad \rho_r = \frac{1}{\nu} \langle r \rangle_{\star,r}, \quad \lambda_n = \langle \widehat{n} \rangle_{\star,n}, \quad \lambda_r = \frac{1}{\nu} \langle \widehat{r} \rangle_{\star,r},$$

$$Q_n = \langle n n' \rangle_{\star,n}, \quad Q_r = \frac{1}{\nu} \langle r r' \rangle_{\star,r}, \quad L_n = \langle \widehat{n} \widehat{n}' \rangle_{\star,n}, \quad L_r = \frac{1}{\nu} \langle \widehat{r} \widehat{r}' \rangle_{\star,r},$$
$$\tag{A23}$$

$$K_n = \langle n \widehat{n}' \rangle_{\star,n}, \quad K_r = \frac{1}{\nu} \langle r \widehat{r}' \rangle_{\star,r},$$

where we used the shorthands

$$\langle [\cdots] \rangle_{\star,n} = \frac{\int D(n, \widehat{n})\, [\cdots]\, e^{i\int dt \psi(t) n(t)}\, e^{i\int dt \widehat{n}(t)\left(\frac{\dot{n}}{n} + m - \theta\right)} \times e^{-i\int dt\left[\widehat{\rho_n} n + \widehat{\lambda_n} \widehat{n}\right]}\, e^{-i\int dt dt'\left[\widehat{Q_n} n n' + \widehat{L_n} \widehat{n} \widehat{n}' + \widehat{K_n} n n'\right]}}{\int D(n, \widehat{n})\, e^{i\int dt \psi(t) n(t)}\, e^{i\int dt \widehat{n}(t)\left(\frac{\dot{n}}{n} + m - \theta\right)} \times e^{-i\int dt\left[\widehat{\rho_n} n + \widehat{\lambda_n} \widehat{n}\right]}\, e^{-i\int dt dt'\left[\widehat{Q_n} n n' + \widehat{L_n} \widehat{n} \widehat{n}' + \widehat{K_n} n n'\right]}}, \tag{A24}$$

$$\langle [\cdots] \rangle_{\star,r} = \frac{\int D(r, \widehat{r})\, [\cdots]\, e^{i\int dt \phi(t) r(t)}\, e^{i\int dt \widehat{r}(t)\left(\frac{\dot{r}}{r} - a(K - r) - \eta\right)} \times e^{-i\int dt\left(\widehat{\rho_r} r + \widehat{\lambda_r} \widehat{r}\right)}\, e^{-i\int dt dt'\left[\widehat{Q_r} r r' + \widehat{L_r} \widehat{r} \widehat{r}' + \widehat{K_r} r \widehat{r}'\right]}}{\int D(r, \widehat{r})\, e^{i\int dt \phi(t) r(t)}\, e^{i\int dt \widehat{r}(t)\left(\frac{\dot{r}}{r} - a(K - r) - \eta\right)} \times e^{-i\int dt\left(\widehat{\rho_r} r + \widehat{\lambda_r} \widehat{r}\right)}\, e^{-i\int dt dt'\left[\widehat{Q_r} r r' + \widehat{L_r} \widehat{r} \widehat{r}' + \widehat{K_r} r \widehat{r}'\right]}}. \tag{A25}$$

Order parameters involving only $\widehat{n}$ or $\widehat{r}$ can be dealt with by noting, for instance, that

$$L_n(t, t') \equiv \frac{1}{N} \sum_i \widehat{n}_i(t) \widehat{n}_i(t') = -\frac{1}{N} \sum_i \frac{\partial^2 Z[\mathbf{0}, \mathbf{0}]}{\partial \theta_i(t) \partial \theta_i(t')} = 0, \tag{A26}$$

since by definition $Z[\mathbf{0}, \mathbf{0}] = 1$ [see (A3)]. One, therefore, easily finds that $\lambda_r = L_r = 0$, implying $\widehat{\rho_r} = \widehat{Q_n} = 0$, and $\lambda_n = L_n = 0$,

implying $\widehat{\rho_r} = \widehat{Q_r} = 0$. Ω_r then reduces to

$$\Omega_r = \frac{1}{\nu} \ln \int D(r, \widehat{r})\, e^{i\int dt \phi(t) r(t)}\, e^{i\int dt \widehat{r}(t)\left(\frac{\dot{r}}{r} - a(K - r) - \eta + c\rho_n - i\sigma_c^2 \int dt' K_n(t,t') r(t')\right)}$$
$$\times\, e^{-\frac{\sigma_c^2}{2} \int dt dt' \widehat{r} Q_n \widehat{r}'}\, \times\, e^{-\frac{a^2 \sigma_K^2}{2} \int dt dt' \widehat{r} \widehat{r}'}, \tag{A27}$$

which corresponds to the effective resources dynamics

$$\frac{\dot{r}(t)}{r(t)} = a(K - r(t)) - c\rho_n(t) - \sigma_c^2 \int dt' G_n(t, t') r(t') + \eta(t) + \xi_r(t),$$
$$\times\, \langle \xi_r(t) \xi_r(t') \rangle = \sigma_c^2 Q_n(t, t') + a^2 \sigma_K^2, \tag{A28}$$

where we re-defined the response function as $-iK_n = G_n$ consistently with the definition of Z. Likewise, Ω_n takes the form

$$\Omega_n = \ln \int D(n,\widehat{n}) \, e^{i \int dt\psi(t)n(t)} \, e^{i \int dt\widehat{n}(t)\left(\frac{\dot{n}}{n}+m-\theta-c\rho_r-i\sigma_c^2 \int dt' K_r(t,t')n(t')\right)}$$
$$\times \, e^{-\frac{\sigma_c^2}{2} \int dt\,dt'\widehat{n}Q_r\widehat{n'}}, \tag{A29}$$

leading to

$$\frac{\dot{n}(t)}{n(t)} = \theta(t) - m + c\rho_r(t) - \frac{\sigma_c^2}{\nu} \int dt' G_r(t,t')n(t')$$
$$+ \xi_n(t), \quad \langle \xi_n(t)\xi_n(t')\rangle = \sigma_c^2 Q_r(t,t'), \tag{A30}$$

where $-iK_r = \frac{1}{\nu}G_r$. Equations (A28) and (A30) are the effective single-resource and single-species processes describing the dynamics of the full system in the limit $N, M \to \infty$ for fixed $\nu = N/M$. We have, therefore, linked the original (Markovian) system described by (A1) and (A2) to a new (non-Markovian) system involving a single effective species and a single effective resource.

REFERENCES

[1] D. R. Hekstra and S. Leibler, "Contingency and statistical laws in replicate microbial closed ecosystems," Cell 149(5), 1164–1173 (2012).
[2] C. Huttenhower, D. Gevers, R. Knight, S. Abubucker, J. H. Badger, A. T. Chinwalla, and M. G. Giglio, "Structure, function and diversity of the healthy human microbiome," Nature 486(7402), 207 (2012).
[3] J. Friedman, L. M. Higgins, and J. Gore, "Community structure follows simple assembly rules in microbial microcosms," Nat. Ecol. Evol. 1(5), 1–7 (2017).
[4] J. E. Goldford, N. Lu, D. Bajic, S. Estrela, M. Tikhonov, A. Sanchez-Gorostiaga, D. Segre, P. Mehta, and A. Sanchez, "Emergent simplicity in microbial community assembly," Science 361(6401), 469–474 (2018).
[5] W. R. Harcombe, W. J. Riehl, I. Dukovski, B. R. Granger, A. Betts, A. H. Lang, and C. J. Marx, "Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics," Cell. Rep. 7(4), 1104–1115 (2014).
[6] M. J. Wade, J. Harmand, B. Benyahia, T. Bouchez, S. Chaillou, B. Cloez, and R. Arditi, "Perspectives in mathematical modelling for microbial ecology," Ecol. Modell. 321, 64–74 (2016).
[7] A. Succurro and O. Ebenhöh, "Review and perspective on mathematical modeling of microbial ecosystems," Biochem. Soc. Trans. 46(2), 403–412 (2018).
[8] G. Biroli, G. Bunin, and C. Cammarota, "Marginally stable equilibria in critical ecosystems," New. J. Phys. 20(8), 083051 (2018).
[9] A. Goyal and S. Maslov, "Diversity, stability, and reproducibility in stochastically assembled microbial ecosystems," Phys. Rev. Lett. 120(15), 158102 (2018).
[10] R. MacArthur, "Species packing and competitive equilibrium for many species," Theor. Popul. Biol. 1(1), 1–11 (1970).
[11] P. Chesson, "MacArthur's consumer-resource model," Theor. Popul. Biol. 37(1), 26–38 (1990).
[12] A. De Martino and M. Marsili, "Statistical mechanics of socio-economic systems with heterogeneous agents," J. Phys. A 39(43), R465 (2006).
[13] Y. Yoshino, T. Galla, and K. Tokita, "Statistical mechanics and stability of a model eco-system," J. Stat. Mech.: Theory Exp. 2007(09), P09003 (2007).
[14] M. Tikhonov, "Community-level cohesion without cooperation," eLife 5, e15747 (2016).
[15] M. Tikhonov, "Theoretical microbial ecology without species," Phys. Rev E 96, 032410 (2017).
[16] M. Tikhonov and R. Monasson, "Collective phase in resource competition in a highly diverse ecosystem," Phys. Rev. Lett. 118(4), 048103 (2017).
[17] M. Advani, G. Bunin, and P. Mehta, "Statistical physics of community ecology: A cavity solution to MacArthur's consumer resource model," J. Stat. Mech.: Theory Exp. 2018(3), 033406 (2018).
[18] M. Tikhonov and R. Monasson, "Innovation rather than improvement: A solvable high-dimensional model highlights the limitations of scalar fitness," J. Stat. Phys. 172(1), 74–104 (2018).
[19] W. Cui, R. Marsland, and P. Mehta, "Effect of resource dynamics on species packing in diverse ecosystems," Phys. Rev. Lett. 125(4), 048101 (2020).
[20] R. Marsland, W. Cui, and P. Mehta, "A minimal model for microbial biodiversity can reproduce experimentally observed ecological patterns," Sci. Rep. 10(1), 1–17 (2020).
[21] L. Pacciani-Mori, A. Giometto, S. Suweis, and A. Maritan, "Dynamic metabolic adaptation can promote species coexistence in competitive communities," PLoS Comput. Biol. 16(5), e1007896 (2020).
[22] C. De Dominicis, "Dynamics as a substitute for replicas in systems with quenched random impurities," Phys. Rev. B 18(9), 4913 (1978).
[23] H. Sompolinsky and A. Zippelius, "Relaxational dynamics of the Edwards-Anderson model and the mean-field theory of spin-glasses," Phys. Rev. B 25(11), 6860 (1982).
[24] A. M. Sengupta and P. P. Mitra, "Distributions of singular values for some random matrices," Phys. Rev. E 60(3), 3389 (1999).
[25] R. S. Etienne and D. Alonso, "A dispersal-limited sampling theory for species and alleles," Ecol. Lett. 8(11), 1147 (2005).
[26] R. M. May, "Will a large complex system be stable?," Nature 238, 413 (1972).
[27] G. Hardin, "The competitive exclusion principle," Science 131, 1292 (1960).
[28] R. MacArthur and R. Levins, "Competition, habitat selection, and character displacement in a patchy environment," Proc. Natl. Acad. Sci. U.S.A. 51, 1207 (1964).
[29] V. Dubinkina, Y. Fridman, P. P. Pandey, and S. Maslov, "Multistability and regime shifts in microbial communities explained by competition for essential nutrients," eLife 8, e49720 (2019).
[30] A. Posfai, T. Taillefumier, and N. S. Wingreen, "Metabolic trade-offs promote diversity in a model ecosystem," Phys. Rev. Lett. 118(2), 028103 (2017).
[31] R. Hermsen, H. Okano, C. You, N. Werner, and T. Hwa, "A growth-rate composition formula for the growth of E.coli on co-utilized carbon substrates," Mol. Syst. Biol. 11(4), 801 (2015).
[32] R. Levins, "Coexistence in a variable environment," Am. Nat. 114(6), 765–783 (1979).
[33] B. Descamps-Julien and A. Gonzalez, "Stable coexistence in a fluctuating environment: An experimental demonstration," Ecology 86(10), 2815–2824 (2005).
[34] E. Bairey, E. D. Kelsic, and R. Kishony, "High-order species interactions shape ecosystem diversity," Nat. Commun. 7, 1–7 (2016).
[35] J. Grilli, G. Barabas, M. J. Michalska-Smith, and S. Allesina, "Higher-order interactions stabilize dynamics in competitive network models," Nature 548(7666), 210–213 (2017).
[36] C. Liao, T. Wang, S. Maslov, and J. B. Xavier, "Modeling microbial cross-feeding at intermediate scale portrays community dynamics and species coexistence," PLoS Comput. Biol. 16(8), e1008135 (2020).
[37] J. Fernandez-de-Cossio-Diaz and R. Mulet, "Statistical mechanics of interacting metabolic networks," Phys. Rev. E 101(4), 042401 (2020).