



# Neutral theory of cooperative dynamics

Jordi Piñero<sup>a,1</sup> , Artemy Kolchinsky<sup>b,c,d,1</sup> , Sidney Redner<sup>e</sup> , and Ricard Solé<sup>b,c,e,f,2</sup>

Edited by Andrea Rinaldo, Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland; received June 12, 2025; accepted November 18, 2025

Mutualistic interactions are widespread in nature, from plant communities and microbiomes to human organizations. Along with competition for resources, cooperative interactions shape biodiversity and contribute to the robustness of complex ecosystems. We present a stochastic neutral theory of cooperator species. Our model shares with the classic neutral theory of biodiversity the assumption that all species are equivalent, but crucially differs in requiring cooperation between species for replication. With low migration, our model displays a bimodal species-abundance distribution, with a high-abundance mode associated with a core of cooperating species. This core is responsible for maintaining a diverse pool of long-lived species, which are present even at very small migration rates. We derive analytical expressions of the steady-state species abundance distribution, as well as scaling laws for diversity, number of species, and residence times. With high migration, our model recovers the results of classic neutral theory. We briefly discuss implications of our analysis for research on the microbiome, synthetic biology, and the origin of life.

neutral theory | cooperation | theoretical ecology | stochastic processes | mutualism

Cooperation is ubiquitous across scales in complex systems. In ecology, cooperative interactions shape ecosystem structures (1, 2). Within biology, mutualism (a reciprocal form of cooperation) is the engine of evolutionary transitions (3, 4) and constitutes an essential part of the architecture of biodiversity (5, 6). In microbiomes, cooperative interactions occur through extensive cross-feeding exchanges associated with shared diffusive metabolites (7–9), a phenomenon dubbed the “social network” of microorganisms (10). Moreover, cooperative interactions enhance community stability (11) and facilitate metabolic functions (12). Cooperation has also been studied in human organizations, for example, between companies engaged in jointly manufacturing a certain product (13).

Ecological theory has traditionally studied either *i*) the dynamics of large randomly assembled communities, leading to general stability–complexity principles (14, 15) or *ii*) the stationary properties of stochastic interactions between neutral species, inspired by the neutral theory of biodiversity (16, 17). Stability patterns in mutualistic communities under approach *i*) have been addressed using network models (18–20). The results of these models may depend strongly on the assumptions made about different species (in terms of size, lifestyle, or physiology) and their interactions, making it challenging to derive general theoretical lessons.

In contrast, approach *ii*) assumes that all species in the system are equivalent, shifting the focus to the effects of demographic noise as the main determinant of emerging ecological patterns (21–25). Despite its simplicity, neutral theory successfully accounts for many relevant statistical patterns, including abundance distributions, species–area relations, and diversity estimates in space and time (16, 22, 26, 27). However, a neutral theory that explicitly involves cooperative interactions is lacking, and little is known about the impact of mutualistic exchanges under a neutral picture. For example, microbiomes exhibit marked quantitative patterns that diverge significantly from those predicted by classic neutral theory, such as the presence of a persistent subset of species (12, 28, 29).

In this paper, we present a neutral theory for ecosystems of cooperators. Our model preserves the neutral hypothesis—that all species in the system follow the same replication rules, albeit possibly with frequency-dependent replication probabilities. However, it departs from the classic neutral theory of biodiversity by requiring cooperation between species for replication to occur. Our model contains only two independent parameters: the total number of individuals in the system  $N$ , typically assumed to be large, and the migration rate  $\mu$ , the probability that a new species enters the system at a given time step. One of the remarkable features of our model is the emergence of a bimodal species-abundance distribution, characterized by a “core” of cooperator species that remain at high abundances for much longer compared to species in the low-abundance mode. This core leads to the maintenance of species diversity even at very small migration rates,

## Significance

Cooperative interactions shape biodiversity across ecosystems. This study introduces a stochastic neutral theory of cooperator species, revealing how cooperation fosters diversity beyond traditional neutral models based on pure competition. A key finding is the emergence of a “cooperator core,” a set of species that sustain high abundance and longevity, even under low migration rates. These insights provide a theoretical framework for understanding diversity maintenance in complex ecosystems, with implications for the microbiome, synthetic biology, and the origin of life.

Author affiliations: <sup>a</sup>Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824; <sup>b</sup>Institució Catalana de Recerca i Estudis Avançats-Complex Systems Lab, Department of Medicine and Life Sciences, Universitat Pompeu Fabra, Barcelona 08003, Spain; <sup>c</sup>Barcelona Collaboratorium, Barcelona 08005, Spain; <sup>d</sup>Universal Biology Institute, University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan; <sup>e</sup>Santa Fe Institute, Santa Fe, NM 87501; and <sup>f</sup>Institut de Biologia Evolutiva (CSIC-UPF), Barcelona 08003, Spain

Author contributions: J.P., A.K., S.R. and R.S. designed research; J.P., A.K., S.R. and R.S. performed analyses; and J.P., A.K., S.R. and R.S. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2025 the Author(s). Published by PNAS. This article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](#).

<sup>1</sup>J.P. and A.K. contributed equally to this work.

<sup>2</sup>To whom correspondence may be addressed. Email: ricard.sole@upf.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2515423122/-DCSupplemental>.

Published December 19, 2025.

where, in the absence of cooperation, the system would rapidly fixate. In this sense, we show that cooperative interactions can provide a powerful stabilizing effect.

We solve our model analytically and confirm the results via extensive numerical simulations. Our analysis is split into two main parts: First, we focus on the statistical patterns displayed by the system's steady state; afterward, we consider the stochastic dynamics of abundance trajectories as species enter and leave the system.

## 1. Model Setup

Our model considers a well-mixed population of  $N$  individuals, each labeled by a species. The configuration of the system at a given time is represented by the abundance vector  $\mathbf{n} = (n_1, n_2, \dots)$ , where  $n_i$  is the abundance of species  $i$ . We use  $R(\mathbf{n})$  to indicate the number of nonzero elements of vector  $\mathbf{n}$ , i.e., the number of present species. The configuration evolves stochastically over a sequence of discrete steps involving either a cooperative replication or a migration event:

- 1) Replication with probability  $1 - \mu$ , Fig. 1, *Left*: Two individuals are randomly chosen from the population, individual  $A$  from species  $i$  and individual  $B$  from species  $j$ . If  $A$  and  $B$  belong to the same species ( $i = j$ ), nothing happens. Otherwise, an individual  $C$  of species  $k$  is randomly selected from the population and replaced by a new individual of species  $i$ . The abundance vector  $\mathbf{n}$  is updated as

$$(n_i, n_k) \rightarrow (n_i + 1, n_k - 1). \quad [1]$$

(If  $i = k$ , there is no change). Given our well-mixed population, the probability that this event occurs in a given step is

$$(1 - \mu) \frac{n_i}{N} \frac{N - n_i}{N} \frac{n_k}{N}. \quad [2]$$

In Eq. 2, the factors represent the probability that replication takes place,  $A$  belongs to species  $i$ ,  $B$  belongs to any species other than  $i$ , and  $C$  belongs to species  $k$ .

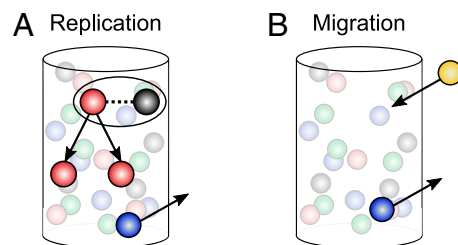
- 2) Migration with probability  $\mu$ , Fig. 1, *Right*: An individual  $A$  of species  $i$  is randomly chosen and replaced by an individual of a new species. The abundance vector  $\mathbf{n}$  is updated as

$$(n_i, n_{R(\mathbf{n})+1}) \rightarrow (n_i - 1, 1). \quad [3]$$

This event happens with probability  $\mu(n_i/N)$  per step. For notational convenience, if any species becomes extinct ( $n_i \rightarrow 0$ ), the species are reindexed so that the first  $R(\mathbf{n})$  entries of  $\mathbf{n}$  are strictly positive.

Total population size is conserved under the above rules ( $\sum_i n_i = N$ ), resulting in competition between species for limited space. Also, because every migration event introduces a new species, migration represents an inflow from an infinitely diverse external reservoir.

In the following, we consider the stationary properties of this stochastic process. We are particularly interested in the species abundance distribution  $P_n$  (the probability that a species has abundance  $n$  in steady state) and the system's overall diversity. We measure diversity using the Simpson index (30), defined for



**Fig. 1.** Neutral model of cooperators. The population evolves according to two rules: (A) during replication, individuals of different species (black and red balls) cooperate. The red individual replicates and replaces a randomly chosen individual (blue ball). (B) during migration with probability  $\mu$ , an individual from a new species (yellow ball) enters from an external pool and substitutes a randomly chosen individual (blue ball).

abundance vector  $\mathbf{n}$  as\*

$$\lambda(\mathbf{n}) := \sum_{i=1}^{R(\mathbf{n})} \left( \frac{n_i}{N} \right)^2. \quad [4]$$

The Simpson index is the probability that two randomly chosen individuals belong to the same species (30), and it is commonly used in ecology as an inverse measure of diversity (31). It is bounded as  $1/R(\mathbf{n}) \leq \lambda(\mathbf{n}) \leq 1$ , with the lower bound achieved when the population is evenly distributed among  $R(\mathbf{n})$  species (maximum diversity) and the upper bound achieved when the population is fixated on a single species (minimum diversity). The inverse Simpson index  $1/\lambda(\mathbf{n})$  may be interpreted as the “effective” number of species in the system.

As shown below, our model has two different steady-state regimes, depending on the migration rate. When migration is high, diversity is very large ( $\lambda \approx 0$ ), competition for space is strong, and the species abundance is well described by the Fisher’s Logseries distribution. This regime reduces to the classic neutral model of Hubbell, which does not involve cooperative interactions (32). When migration is low, the system forms a “core” of high-abundance species, and the species-abundance distribution acquires a characteristic bimodal shape, shown schematically in Fig. 2. In what follows, we will study the emergence and nature of this core.

## 2. Steady-State Distribution

To derive the steady-state species abundance distribution, we exploit the neutrality of our model, which allows us to treat all present species as statistically equivalent. Thus, without loss of generality, we consider species  $i = 1$  as the representative species.

The stochastic dynamics of the abundance of the representative species, written  $n \equiv n_1$ , follows a birth–death process. The birth probability ( $n \rightarrow n + 1$  for  $n \in \{1, \dots, N - 1\}$ ) during a single time step is

$$b_n = (1 - \mu) \frac{n}{N} \left( 1 - \frac{n}{N} \right)^2, \quad [5]$$

while the death probability ( $n \rightarrow n - 1$  for  $n \in \{1, \dots, N\}$ ) is

$$d_n = (1 - \mu) \frac{n}{N} \left( 1 - \frac{n}{N} \right) \left[ 1 - \left( 1 - \frac{n}{N} \right) \lambda_o(\mathbf{n}) \right] + \mu \frac{n}{N}. \quad [6]$$

\*The Simpson index  $\lambda(\mathbf{n})$  is sometimes alternatively defined as  $\sum_{i=1}^{R(\mathbf{n})} n_i(n_i - 1)/(N(N - 1))$ , i.e., as the probability of drawing two individuals of the same species when sampling without replacement. The definition [4] corresponds to the same probability when sampling with replacement. The two definitions differ by a small term of order  $1/(N - 1)$ , which is irrelevant for our analysis.

(For a detailed derivation of Eqs. 5 and 6 see *SI Appendix, section S1*.) The quantity  $\lambda_o(\mathbf{n})$  refers to the Simpson index of the other species (i.e., excluding the representative species  $i = 1$ ):

$$\lambda_o(\mathbf{n}) := \sum_{i=2}^N \left( \frac{n_i}{N-n} \right)^2. \quad [7]$$

Although different species interact during replication, the stochastic dynamics of each species depends on all others only through a single number,  $\lambda_o(\mathbf{n})$ . We remark that similar coupled birth–death processes have been considered in work on interacting particle systems and nonlinear chemical master equations (33).

The above birth–death process is absorbing into the extinction state  $n = 0$ , since any particular species will eventually go extinct. To study the abundance of nonextinct species, we may make the process ergodic by adding a positive birth probability  $b_0 > 0$  out of the extinction state  $n = 0$ . In the following, we focus on the steady-state distribution restricted to positive abundances  $n \in \{1, 2, \dots\}$ , which does not depend on the choice of  $b_0$ .

We now find the steady-state species abundance distribution  $P_n$  of the representative species. Since births and deaths must balance in steady state,  $P_n b_n = P_{n+1} d_{n+1}$  for  $n \in \{1, \dots, N-1\}$ , we have

$$P_n \propto \prod_{k=1}^{n-1} \frac{b_k}{d_{k+1}}. \quad [8]$$

This is not yet a closed equation because the death probability [6] depends on the fluctuating quantity  $\lambda_o(\mathbf{n})$ , i.e., the Simpson index of the nonrepresentative species. However, when the number of species is large,  $\lambda_o(\mathbf{n}) \approx \lambda(\mathbf{n})$ . Furthermore, due to self-averaging,  $\lambda(\mathbf{n})$  is tightly peaked around its expected steady-state value,

$$\lambda_o(\mathbf{n}) \approx \lambda^* := \langle \lambda(\mathbf{n}) \rangle, \quad [9]$$

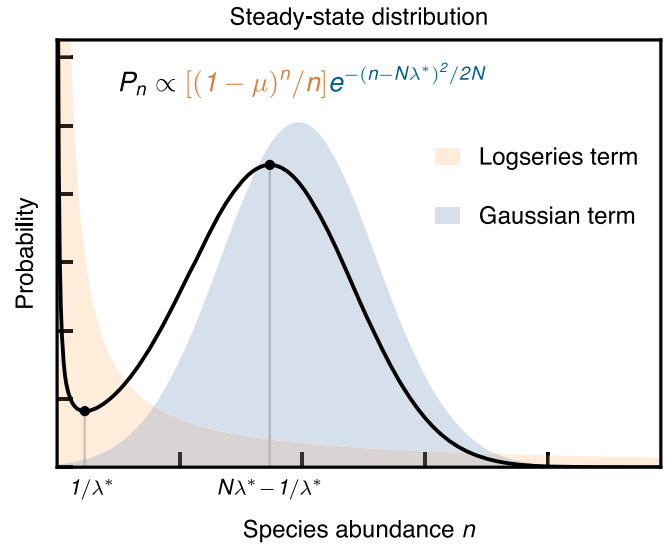
where  $\langle \cdot \rangle$  indicates steady-state expectations over the entire population. Plugging [9] into [6] and using [8], we find that under reasonable assumptions, the steady-state abundance distribution is approximated as

$$P_n \propto \frac{(1-\mu)^n}{n} e^{-(n-N\lambda^*)^2/2N}. \quad [10]$$

(See *SI Appendix, section S2* for a detailed derivation.) We also analyze the approximation Eq. 9 and show that it becomes exact in the limit of large  $N$ .

Eq. 10 is one of our main results, showing how the abundance distribution depends on the population size  $N$ , migration rate  $\mu$ , and the steady-state Simpson index  $\lambda^*$ . In reality,  $\lambda^*$  itself depends on  $N$  and  $\mu$ —the only two parameters that describe the system—although finding the explicit expression of  $\lambda^*$  in terms of  $N$  and  $\mu$  is not trivial (we will do so in the next section). For now, we note that the Simpson index  $\lambda^*$  and the migration probability  $\mu$  move in opposite directions. When  $\mu$  decreases to 0 (migration vanishes),  $\lambda^*$  increases to 1 (fixation is reached). Conversely, when  $\mu$  increases to 1 (only migration occurs),  $\lambda^*$  decreases to its minimum value of  $1/N$  (maximum diversity).

Observe that Eq. 10 expresses  $P_n$  as the product of two distributions, as illustrated in Fig. 2. The first is the Logseries distribution,  $(1-\mu)^n/n$ , the abundance distribution of Hubbell’s neutral

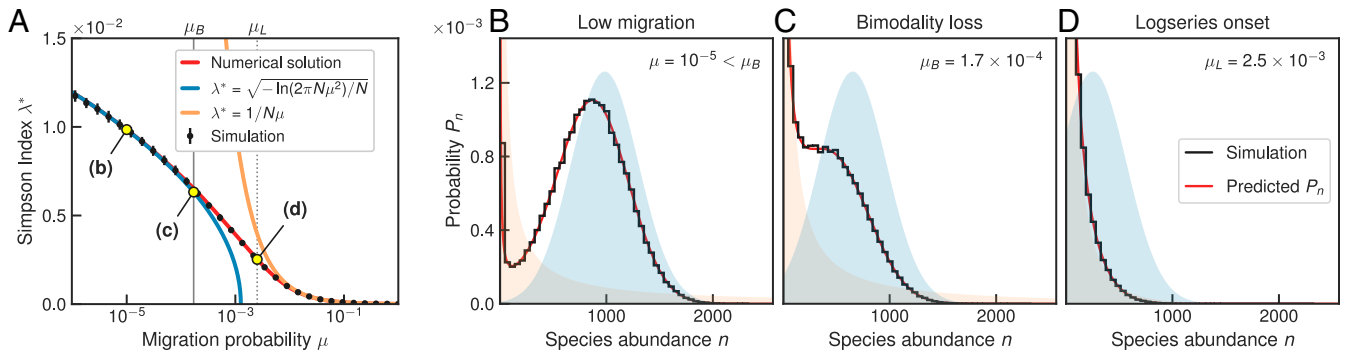


**Fig. 2.** Schematic illustration of the species abundance distribution  $P_n$  in steady state. This distribution depends on population size  $N$ , migration rate  $\mu$ , and the steady-state Simpson index  $\lambda^*$ . In Eq. 10, it is approximated as the product of Logseries distribution, arising from competition for space, and a Gaussian distribution, arising from cooperative interactions. At low migration, the combined distribution exhibits a bimodal shape, with local maxima at  $n = 1, n = N\lambda^* - 1/\lambda^*$  and a local minimum at  $n = 1/\lambda^*$ , see Eq. 15. Note that the two terms combine multiplicatively, not additively, so their areas do not add up to the area of the combined distribution.

model, which does not have cooperative interactions (34). This contribution represents migration-driven competition for space, and it dominates the system at high migration rates. The second is a Gaussian distribution,  $e^{-(n-N\lambda^*)^2/2N}$ , with mean abundance  $N\lambda^*$  and SD  $\sqrt{N}$ . This contribution represents a high-abundance mode that emerges due to cooperative interactions. We term this contribution the “cooperator core.”

The emergence of the cooperator core is a highly nontrivial phenomenon, and it implies various other interesting aspects of this model (maintenance of high diversity at low migration, long residence times, etc.). In simple terms, the cooperator core arises due to the frequency-dependent (i.e., abundance-dependent) replication rates. Individuals from low-abundance species are unlikely to be randomly paired with an individual from the same species, thus they have a high per-capita probability of replicating. Conversely, individuals from high-abundance species are more likely to be randomly paired with an individual from the same species, thus they have a lower per-capita replication probability. As a result, species feel an effective force toward an intermediate abundance at  $n \approx N\lambda^*$ , leading to a bimodal abundance distribution as in Fig. 2.

Importantly, the cooperator core only emerges at low migration rates. At high migration, when  $\mu$  is large and  $\lambda^*$  is low, most species exhibit low abundances near  $n \approx 1$ . In this regime, there is very little variation in species abundances, so the frequency dependence of the replication rates is negligible, and our system approaches Hubbell’s neutral model. We also remark that bimodality requires that the mean of the Gaussian contribution ( $N\lambda^*$ ) be many SDs ( $\sqrt{N}$ ) away from the origin or, in other words, that  $\lambda^* \gg 1/\sqrt{N}$ . Clearly, this cannot occur at very high migration probability, since  $\lambda^*$  decreases to its minimum value of  $1/N$  as  $\mu$  increases to 1. In the next section, we will derive the explicit migration probability  $\mu$  that allows for bimodality.



**Fig. 3.** Scaling of Simpson index and species abundance distributions. (A) Scaling of Simpson index versus migration probability  $\mu$  in a system with  $N = 10^5$  individuals. We compare mean and SDs from simulations of the full system (100 runs; black), numerical inversion of Eq. 12 (red), approximations in low migration [13] (blue) and high migration [18] (orange) regimes. The solid vertical line indicates  $\mu_B$  [16] where bimodality is lost; the dotted vertical line indicates  $\mu_L$  [17] where system transitions to Logseries regime. (B–D) Empirical histograms from simulations (across 1,000 runs) versus predicted steady-state distributions [10] for three migration probabilities. As in Fig. 2, shaded areas represent Logseries (orange) and Gaussian (blue) contributions. (B) Low-migration regime exhibits a bimodal distribution with a cooperator core. (C) Bimodality disappears once migration probability increases beyond  $\mu_B$  (Eq. 16). (D) At higher migration probabilities, including the Logseries transition point where  $\mu_L = \lambda^*$  [17], the distribution approaches the Logseries, as predicted by Hubbell's neutral theory.

### 3. Steady-State Properties

In this section, we derive several important properties of the steady-state species abundance distribution and the cooperator core. In particular, we derive expressions of the steady-state Simpson index, the characteristic migration probability  $\mu_B$  below which the formation of the cooperator core takes place, and the expected number of species in steady state. Most of our analyses will consider separately the case of low migration (when the core is present) and high migration. Fig. 3 illustrates our results, including the scaling of  $\lambda^*$  with the migration probability  $\mu$  and a plot of three steady-state distributions.

**3.1. Steady-State Simpson Index.** In this section, we find the Simpson index as a function of  $\mu$  and  $N$ . We begin by writing  $\lambda^*$  in terms of the species abundance distribution  $P_n$  as

$$\lambda^* \approx \frac{1}{N} \frac{\sum_{n=1}^N P_n n^2}{\sum_{n=1}^N P_n n}. \quad [11]$$

This expression (derived in *SI Appendix, section S3*) implies that the expected Simpson index is proportional to the ratio of the second and first moments of  $P_n$ . Since  $P_n$  itself depends on  $\lambda^*$ , we must solve Eq. 11 using self-consistency. We do this separately in the low-migration and the high-migration regimes. We also consider several important transition points.

**3.1.1. Low migration.** In the low-migration regime, we may approximate the two sums in Eq. 11 by integrals (*SI Appendix, section S3*). This leads to the equation

$$\frac{1}{\mu} = \sqrt{\frac{\pi N}{2}} \operatorname{erfcx} \left( \sqrt{\frac{N}{2}} (\mu - \lambda^*) \right), \quad [12]$$

where  $\operatorname{erfcx}(z) := e^{z^2} \operatorname{erfc}(z) = e^{z^2} \sqrt{4/\pi} \int_z^\infty e^{-t^2} dt$  is the scaled complementary error function (35). Although the function  $\operatorname{erfcx}(z)$  does not have a closed-form inverse, it has an efficient numerical implementation (36) that allows us to quickly compute  $\lambda^*$  using numerical root-finding algorithms. The scaling predicted by solving this equation for fixed  $N$  and varying  $\mu$  is shown as a red curve in Fig. 3A.

A closed-form approximation is possible for sufficiently small  $\mu$ . For  $\mu - \lambda^* \ll 0$ , we may use  $\operatorname{erfcx}(z) \approx 2e^{z^2}$  for  $z \rightarrow -\infty$ .

Plugging into Eq. 12 and solving gives

$$\lambda^* \approx \mu + \sqrt{\frac{-\ln(2\pi N \mu^2)}{N}} \approx \sqrt{\frac{-\ln(2\pi N \mu^2)}{N}}. \quad [13]$$

Note that this approximation is only defined for  $\mu \leq 1/\sqrt{2\pi N}$ , so that the argument of the square root is nonnegative.

Eq. 13 implies that, in the low-migration regime, the Simpson index scales as  $\sim 1/\sqrt{N}$  in population size and as  $\sim \sqrt{-\ln \mu}$  in the migration probability. The scaling predicted by Eq. 13 is shown as a blue curve in Fig. 3A. The predicted abundance distribution is compared against simulations in Fig. 3B.

Interestingly, an expression similar to Eq. 12 has previously appeared in a different ecological model, which considers a population embedded in a fluctuating spatiotemporal fitness landscape (37, Appendix C).

**3.1.2. Bimodality point.** We now find the migration probability below which the species abundance distribution becomes bimodal. To do so, we consider  $P_n$  [10] as a differentiable function of  $n$  and find the critical abundances  $\hat{n}$  where its derivative vanishes. With some algebra, these are found to be

$$\hat{n} = \frac{N}{2} \left[ \lambda^* + \ln(1 - \mu) \pm \sqrt{(\lambda^* + \ln(1 - \mu))^2 - 4/N} \right]. \quad [14]$$

We simplify by expanding around large  $N$  and small  $\mu$  to express the two critical abundances as

$$\hat{n}_{\min} \approx \frac{1}{\lambda^*} \quad \text{and} \quad \hat{n}_{\max} \approx N\lambda^* - \frac{1}{\lambda^*}, \quad [15]$$

shown as the local minimum and local maximum in Fig. 2. These two points become distinct when

$$\lambda^* \geq \frac{2}{\sqrt{N}} \quad \text{and} \quad \mu \leq \mu_B := \frac{e^{-2}}{\sqrt{2\pi N}}, \quad [16]$$

where  $\mu_B$  is found by plugging  $\lambda^* = 2/\sqrt{N}$  into Eq. 13 and solving. Thus,  $\mu_B$  is the migration probability below which we see the formation of a bimodal abundance distribution.



**3.1.3. Logseries onset.** Another interesting value for the migration probability separates the cooperation- from the migration-dominated regimes. As we increase  $\mu$  toward 1, the abundance distribution becomes dominated by the Logseries term, recovering the distribution predicted by Hubbell's neutral theory of biodiversity.

To identify this transition point, we consider the migration rate at which  $\mu$  becomes larger than  $\lambda^*$ . Plugging  $\mu - \lambda^* = 0$  into Eq. 12 and using  $\text{erfcx}(0) = 1$  specifies this point as

$$\mu_L := \sqrt{\frac{2}{N\pi}}. \quad [17]$$

**3.1.4. High migration.** As mentioned above, when  $\mu$  is large, the abundance distribution approaches that of Hubbell's neutral model. To estimate the steady-state Simpson index  $\lambda^*$  in this regime, observe that the term  $(1 - \mu)^n$  in Eq. 10 imposes an exponential cutoff on abundances  $n > 1/\mu$ . Considering the regime  $\mu > \mu_L$  from Eq. 17, we may restrict our attention to  $n < 1/\mu < \sqrt{N\pi/2}$  and  $\lambda^* < \mu$ . For these values, the Gaussian term can be approximated as a constant factor,  $e^{-(n-N\lambda^*)^2/2N} \approx e^{-N\lambda^{*2}/2}$ . We reevaluate  $\lambda^*$  [11] using the Logseries distribution  $P_n \propto (1 - \mu)^n/n$  and large  $N$ , giving

$$\lambda^* \approx \frac{1}{N\mu}. \quad [18]$$

Thus, with high migration, the Simpson index scales as  $\lambda^* \sim N^{-1}$  in population size and  $\lambda^* \sim \mu^{-1}$  in migration probability. This scaling is shown with the orange curve in Fig. 3A.

In fact, the high-migration expression [18] can be derived from Eq. 12 (as shown in *SI Appendix, section S3B*). This is surprising because Eq. 12 was derived under the assumption of  $\mu \ll \lambda^*$ . Empirically, we found that the value of  $\lambda^*$  given by the numerical solution to Eq. 12 provides an excellent match to  $\lambda^*$  calculated from simulations, even for large migration probabilities (see Fig. 3A, red curve).

**3.2. Steady-State Number of Species.** As another important measure of diversity, we consider the expected number of different species in steady state:

$$R^* := \langle R(\mathbf{n}) \rangle. \quad [19]$$

The expected steady-state abundance of the representative species obeys  $\sum_n P_n n = N/R^*$ , therefore the number of species can be found as

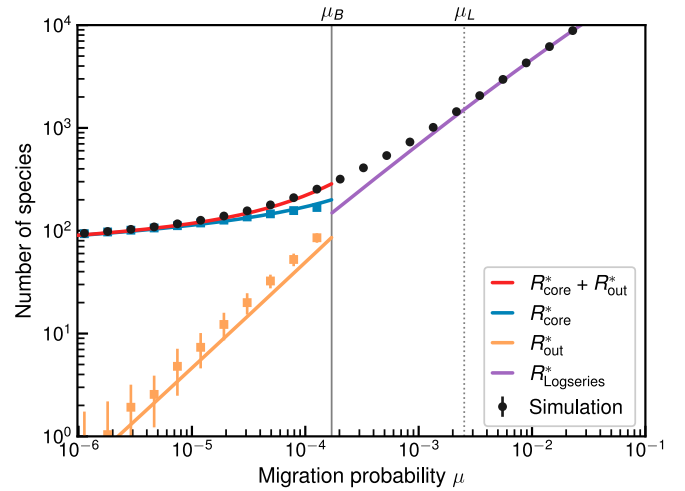
$$R^* = \frac{N}{\sum_n P_n n}. \quad [20]$$

With high migration,  $R^*$  can be estimated by evaluating the denominator of Eq. 20. We may ignore the Gaussian contribution in Eq. 10, since the most relevant contribution to Eq. 10 is the Logseries distribution. For large  $N$ , the normalization constant of the Logseries term is  $-\ln \mu$ , thus we compute

$$\sum_n P_n n \approx \frac{1}{-\ln \mu} \sum_n (1 - \mu)^n \approx \frac{1 - \mu}{-\mu \ln \mu}. \quad [21]$$

Substituting back into Eq. 20 gives

$$R^* \approx R_{\text{Logseries}}^* := -\frac{\mu N}{1 - \mu} \ln \mu, \quad [22]$$



**Fig. 4.** Scaling of number of species  $R^*$  versus migration probability  $\mu$  ( $N = 10^5$ ). We compare mean and SDs from simulations of the full system (100 runs; black); approximations in the low-migration regime of the number of total species (red), core species [24] (blue), and noncore species [25] (orange); approximations in the high-migration regime of the total number of species (purple). The solid vertical line indicates  $\mu_B$  [16] where bimodality is lost; the dotted vertical line indicates  $\mu_L$  [17] where system transitions to Logseries regime.

shown in purple in Fig. 4. As expected, in the limit of  $\mu \rightarrow 1$ ,  $R_{\text{Logseries}}^* \rightarrow N$ . This indicates that if only migration occurs (with no replication taking place), then the system acquires the maximum number of species with each of the  $N$  individuals belonging to a different species.

In the low-migration regime, below the point of bimodality [16], there are two relevant species counts. The first is  $R_{\text{core}}^*$ , the number of high-abundance species that belong to the cooperator core. We define these species as those with abundances larger than the local minimum,  $n > 1/\lambda^*$ , see Fig. 2. The second is  $R_{\text{out}}^*$ , the number of low-abundance species that remain outside of the cooperator core and have abundance  $n \leq 1/\lambda^*$ . The total number of species is given by

$$R^* \approx R_{\text{core}}^* + R_{\text{out}}^*. \quad [23]$$

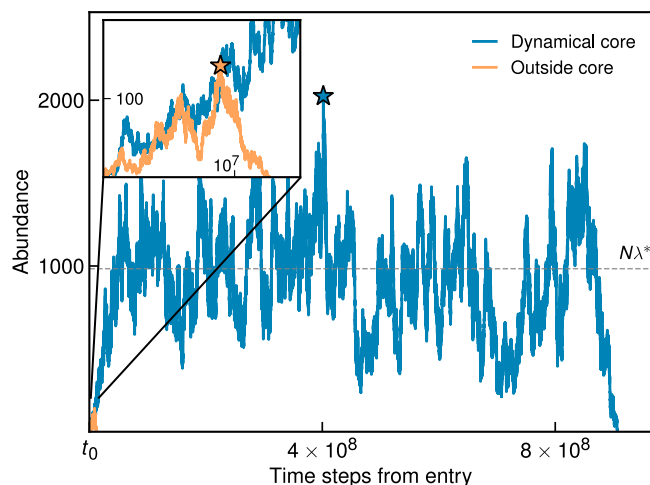
Let us introduce  $N_{\text{core}}$  and  $N_{\text{out}}$  as the total number of individuals inside and outside the core, respectively, where  $N_{\text{core}} + N_{\text{out}} = N$ . We also denote by  $\langle n \rangle_{\text{core}}$  and  $\langle n \rangle_{\text{out}}$  the expected abundances of a species conditioned to be inside and outside the core, respectively. To estimate  $R_{\text{core}}^*$ , we use the relation  $R_{\text{core}}^* = N_{\text{core}}/\langle n \rangle_{\text{core}}$ . To a first approximation, we may take  $N_{\text{core}} \approx N$ , i.e., in the low-migration regime, most individuals belong to the core (*SI Appendix, section S4*). Then, using  $\langle n \rangle_{\text{core}} \approx \hat{n}_{\text{max}}$  from [15], we obtain

$$R_{\text{core}}^* \approx \frac{1}{\lambda^* - (N\lambda^*)^{-1}}, \quad [24]$$

shown in blue in Fig. 4. Similarly, we estimate  $R_{\text{out}}^*$  by using  $R_{\text{out}}^* = N_{\text{out}}/\langle n \rangle_{\text{out}}$ . We may estimate this ratio as (*SI Appendix, section S4*):

$$R_{\text{out}}^* \approx -N\mu \ln \lambda^*, \quad [25]$$

shown in orange in Fig. 4. The red curve shows the total number of species  $R^* \approx R_{\text{core}}^* + R_{\text{out}}^*$ . Observe that at low migration rate  $\mu$ , almost all species belong to the core.



**Fig. 5.** Abundance trajectories from entry to extinction, illustrating that species that enter the dynamical core reside for much longer times. We sample two abundance trajectories from the stationary process ( $N = 10^5$ ,  $\mu = 10^{-5}$ ), one for a species that enters the dynamical core and one for a species that does not. The dashed gray line indicates the abundance position of the local minimum discussed in Section 4.1, stars indicate maximum abundances reached.

## 4. Dynamical Properties

We now study the stochastic dynamics of our model by considering trajectories of species as they enter and leave the population. We focus mainly on the low-migration regime, where cooperation plays an important role.

The dynamics of any given species is described by a trajectory of nonnegative abundance values  $n(t)$ , starting from  $n(t_0) = 1$  when that species enters the population at time  $t = t_0$  and ending on  $n(t_f) = 0$  when that species goes extinct at time  $t = t_f$ . We characterize each trajectory by two statistics: the maximum abundance reached,  $\max_{t: t_0 \leq t \leq t_f} n(t)$ , and the residence time before extinction,  $t_f - t_0$ .

For concreteness, we illustrate two typical abundance trajectories in Fig. 5. We see that the blue trajectory resides in the system for a very long time and reaches a high maximum abundance value, while the orange trajectory does not reach a high abundance and quickly goes extinct.

In fact, simulations show that all trajectories cluster into two well-defined classes: one with long residence times and high abundances, which we term the “dynamical core,” and one with short residence times and low abundances. To illustrate this, Fig. 6 shows a scatter plot of the maximum abundances and residence times of  $10^6$  randomly sampled species. Species that fall into the dynamical core (*Top Right* cluster) tend to have residence times orders of magnitude larger than those that do not enter the dynamical core. Moreover, the distribution of maximum abundances has a clear bimodal shape (*Inset*), allowing us to define a quantitative threshold for classifying species as belonging to the dynamical core.

In the following, we derive the distribution of maximum abundances, shown in red in Fig. 6. We then use this distribution to define the abundance threshold for the dynamical core. We also derive the mean residence times of species that belong and do not belong to the dynamical core.

**4.1. Maximum Abundances.** Let  $Q_m$  indicate the probability that a typical species trajectory reaches a maximum abundance  $m$ . To

find the form of  $Q_m$ , it will be helpful to introduce the quantity  $u_n(\ell)$  to represent the probability that a species with abundance  $n$  reaches abundance  $\ell$  or higher before going extinct.  $Q_m$  and  $u_n(\ell)$  are related by

$$Q_m = u_1(m) - u_1(m+1). \quad [26]$$

In other words,  $Q_m$  is equal to the probability that a species that enters the system at abundance  $n = 1$  eventually reaches abundance  $m$  but not abundance  $m+1$ .

Importantly, the birth–death process defined by Eqs. 5 and 6,  $u_n(\ell)$  obeys the following recurrence:

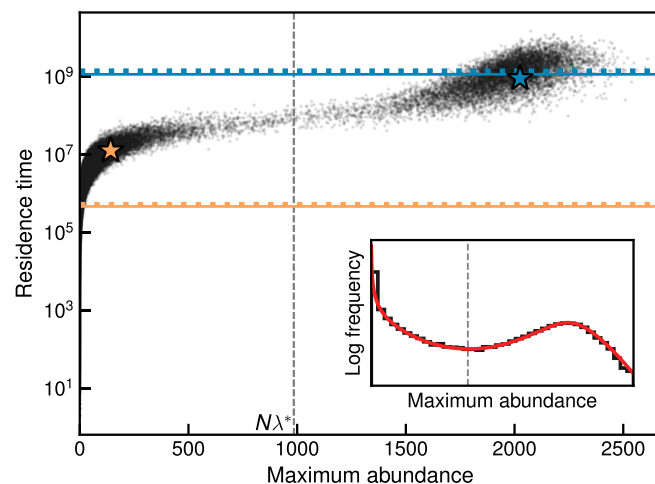
$$u_n(\ell) = b_n u_{n+1}(\ell) + d_n u_{n-1}(\ell) + (1 - b_n - d_n) u_n(\ell). \quad [27]$$

The boundary conditions are  $u_0(\ell) = 0$  (extinct species never come back into the system) and  $u_\ell(\ell) = 1$ . A general solution to the recurrence [27] can be derived (for details, see *SI Appendix, section S5*). Then, we may approximate  $u_1(\ell)$ , the quantity that enters into Eq. 26, as

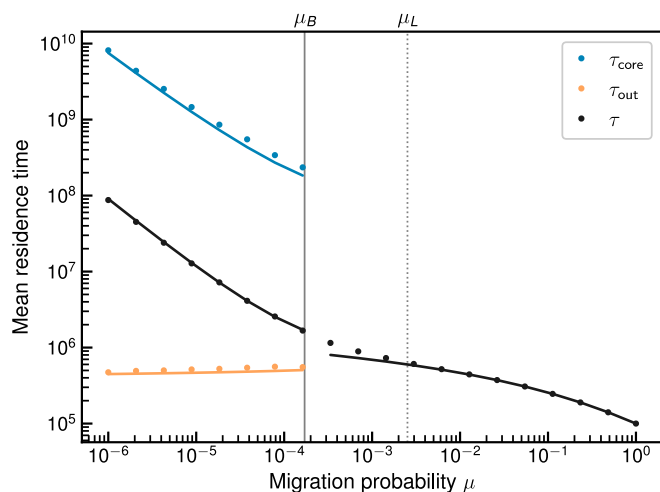
$$u_1(\ell) \approx \frac{1}{e^{\ell^2/2N - (\ell-1)\lambda^*} (\sqrt{2N} \mathcal{D}(x_\ell) + \frac{1}{2}) - \sqrt{2N} \mathcal{D}(x_1) + \frac{1}{2}}, \quad [28]$$

where we introduce the rescaled abundances  $x_\ell := (\ell - N\lambda^*)/\sqrt{2N}$  and the Dawson function  $\mathcal{D}(z) := e^{-z^2} \int_0^z e^{t^2} dt$ .

Plugging this result into Eq. 26 gives the probability  $Q_m$  that a species reaches maximum abundance  $m$ . The predicted and simulated distributions of maximum abundances are shown in Fig. 6, *Inset*. In addition, the distribution  $Q_m$  is bimodal, with a local minimum located at approximately  $N\lambda^*$  (*SI Appendix, section S5*). This bimodality allows us to classify species into two sets: those whose abundances reach  $N\lambda^*$  (or higher) and those



**Fig. 6.** Species residence times and maximum abundances for  $10^6$  species sampled from stationary dynamics ( $N = 10^5$ ,  $\mu = 10^{-5}$ ). The species fall into two clusters: one with long residence times and high maximum abundances, the dynamical core, and one with short residence times and low maximum abundances. The residence times and maximum abundances of the two trajectories from Fig. 5 are indicated with star symbols. Horizontal lines indicate predicted (solid) and simulated (dashed) mean residence times of species in the dynamical core (blue) and outside the dynamical core (orange). *Inset*: the simulated histogram (black) and prediction (red) from Eqs. 26 and 28 of maximum abundances distribution  $Q_m$ . The distribution has a bimodal shape, with the dividing point located near abundance  $N\lambda^*$  (dashed vertical line in main plot and *Inset*).



**Fig. 7.** Scaling of mean residence times with migration probability  $\mu$  ( $N = 10^5$ ). Simulations (dots) are compared to predictions (solid curves) for: core species,  $\tau_{\text{core}}$  [30]; noncore species,  $\tau_{\text{out}}$  [31]; and all species,  $\tau$  [32] and [33] for low and high values of  $\mu$ . The solid vertical line indicates  $\mu_B$  [16] where bimodality is lost; the dotted vertical line indicates  $\mu_L$  [17] where system transitions to Logseries regime.

that do not. We refer to the former as the dynamical core. For example, Fig. 5 shows a blue trajectory that reaches abundance values larger than  $N\lambda^*$  (dashed line) and thus belongs to the dynamical core, while the orange trajectory does not reach this threshold and thus remains outside the dynamical core.

Finally, we use the term “infiltration probability”  $\beta$  to refer to the probability that a new species will enter the dynamical core. Interestingly, this probability is approximately equal to the Simpson index (as shown in *SI Appendix, section S5*),

$$\beta \approx \lambda^*, \quad [29]$$

suggesting that new species are less likely to enter the dynamical core in more diverse populations.

**4.2. Residence Times.** We now derive the mean residence times by separately considering species that enter the dynamical core ( $\tau_{\text{core}}$ ), that do not enter the dynamical core ( $\tau_{\text{out}}$ ), and all species ( $\tau$ ). For simplicity, we will also use our calculations of the expected number of all present species  $R^*$  [22] and [23], core species  $R_{\text{core}}^*$  [24], and noncore species  $R_{\text{out}}^*$  [25].

Here, there is a subtle point to be raised. Our calculations of  $R_{\text{core}}^*$  and  $R_{\text{out}}^*$  were based on a “static” definition of the core, as the set of species that have abundance greater or smaller than  $1/\lambda^*$  at a single point in time. This is different from the “dynamic” definition of the core considered in this section, as the set of species whose trajectory reaches an abundance greater than  $N\lambda^*$  at any point in time. The mean residence times can be alternatively derived using a mean first-passage time (MFPT) calculation that does not explicitly invoke Eqs. 22–25 (see full derivation in *SI Appendix, section S6*). This alternative MFPT-based analysis, which we do not include here for simplicity, leads to the same quantitative results. This suggests that the two definitions of the core are essentially equivalent when restricted to the set of species present in steady state.

We now consider the low-migration regime, and derive the mean residence times of species that enter the core,  $\tau_{\text{core}}$ , and those that do not enter the core,  $\tau_{\text{out}}$ . The rate at which new species migrate into the system and eventually enter the dynamical core is given by  $\mu\beta$ . On the other hand, the rate at

which core species go extinct is  $R_{\text{core}}^*/\tau_{\text{core}}$ , where  $R_{\text{core}}^*$  is the expected number of species in the core at any one time. Since these rates must balance in steady state, we have  $\mu\beta = R_{\text{core}}^*/\tau_{\text{core}}$ . Using Eqs. 24 and 29, we arrive at

$$\tau_{\text{core}} \approx \frac{1}{\mu\lambda^*[\lambda^* - (N\lambda^*)^{-1}]}. \quad [30]$$

Similarly, for species that never reach the core, entry rates must be balanced against extinction rates as  $\mu(1 - \beta) = R_{\text{out}}^*/\tau_{\text{out}}$ . Using Eqs. 25 and 29, this gives

$$\tau_{\text{out}} \approx -\frac{N}{1 - \lambda^*} \ln \lambda^* \quad [31]$$

We estimate the overall mean residence time by combining these results with the infiltration probability as

$$\tau \approx \beta\tau_{\text{core}} + (1 - \beta)\tau_{\text{out}}. \quad [32]$$

These predictions are compared against simulated data in Fig. 6 (solid and dotted horizontal lines).

The scaling of mean residence times against migration probability  $\mu$  is shown in Fig. 7. For low migration probabilities, core species live orders of magnitude longer than species that do not enter the core.

Our analysis above has mostly focused on the low-migration regime, in which the distinction between core and noncore species is meaningful. In the high-migration regime, we consider the mean residence time of all species  $\tau$ . In steady state, entry and exit rates of all species into the system must be balanced,  $\mu = R^*/\tau$ . Using our estimate of  $R^*$  [22] gives

$$\tau \approx \frac{-N}{1 - \mu} \ln \mu. \quad [33]$$

## 5. Discussion

In this paper, we introduced a neutral model for cooperative ecosystems, deliberately choosing the simplest set of rules consistent with cooperative interactions. This minimalist design, inspired by the neutral theory of biodiversity (32), allows a full analytical treatment while capturing the essential features of cooperative dynamics. Our analysis addresses both steady-state behavior and the dynamical properties of the system, providing a baseline for future comparisons with more complex models. It also complements existing models of cooperation, such as mutualistic voter models (38), where agents adopt neighbor states with a bias toward cooperation, or mutualistic Lotka–Volterra systems (39), where species benefit each other’s growth.

In the first part, we derived an expression for the steady-state Simpson diversity index, which allowed us to classify the system into different regimes depending on the migration probability. At high migration rates, species abundances are very low and frequency-dependent effects become negligible. In this regime, we recover the predictions of Hubbell’s neutral theory of biodiversity, including a Logseries species abundance distribution in steady state. In contrast, at low migration rates, our model exhibits frequency-dependent reproduction rates that result from cooperative interactions. In this regime, our model predicts the emergence of a bimodal abundance distribution (Fig. 2), which cannot be derived from the classic neutral theory.

Bimodality allows us to define a core of cooperators, defined as the set of species that belong to the high-abundance component.

We derive scaling laws for the core, showing that the effective number of species scales roughly as  $N^{1/2}$  with population size and  $(-\ln \mu)^{-1/2}$  with migration rate. Due to the presence of the core, the system preserves diversity (maintains low  $\lambda^*$ ) even at exponentially small migration rates. Hence, cooperative interactions can dramatically increase ecosystem stability.

Several studies have reported bimodal abundance distributions in gut microbiomes; see, for example, refs. 40, figure 1 and 41, figure 2. Previous research has attributed bimodality to emergent niche partitioning (42), intrinsic bistability (43), or a combination of multiple processes (41). However, our theory suggests a neutral mechanism that generates bimodality under very minimal assumptions, offering a unique explanation of this phenomenon.

In the second part of our paper, we studied the stochastic dynamics of species as they enter and leave the system. In the low-migration regime, the distribution of maximum abundances again exhibits a bimodal shape, allowing us to classify species into two types. This classification captures the dynamical signature of the cooperator core, as discussed in the steady-state analysis in the first part of this paper. We showed that species that enter the core achieve much higher abundances and longer residence times than species that do not enter the core.

It is worth noting that other models involving higher-order interactions have been proposed and investigated in various contexts. For example, Hinrichsen (44) investigated the universality class of the pair contact process  $2A \rightarrow 3A$ ,  $2A \rightarrow 0$  with diffusion and suggested that this model does not belong to the directed percolation universality class. In ref. 45, the authors studied a quadratic contact process and determined a bound on the critical value of the infection rate, while in refs. 46 and 47, the authors introduced general higher-order social contagion models and showed that these higher-order interactions fundamentally change the nature of the phase transition to a contagion phase. Although these models are related to ours in their use of higher-order interactions, they do not capture key features of our system, such as the bimodal species abundance distribution and the emergence of a cooperator core. Moreover, other similarities can be found when comparing our work with a recent Generalized Lotka–Volterra model with annealed disorder, which shows how temporal fluctuations in interactions act as effective environmental noise, promoting diversity (48). Though based on different mechanisms, both highlight how nonequilibrium stochastic dynamics can generate high diversity and bimodal species abundance patterns.

In future work, it may be possible to generalize the method developed in this paper to other models of neutral cooperators, including three-way generalizations of our pairwise cooperation rule as well as other models considered in the literature. For a general neutral model, one may describe the abundance fluctuations of a “representative species” using a nonlinear birth–death master equation (33), with transition rates encoding mean-field effects of the rest of the population. Steady-state properties can then be identified by solving with self-consistency, in a similar way as done in this study.

Another promising direction to explore is the use of spatially extended models, which can offer additional insights into the effects of local interactions. This approach has also been extensively developed within the framework of neutral

theory (25, 32, 49), where it has enabled the study of the dynamics of metacommunities. In our cooperative context, spatial structure would introduce correlations between neighboring organisms, effectively reducing the reproduction rate of existing species. We anticipate that this rate would scale with the surface area that separates a given species from its neighbors. Consequently, less frequent species—having a higher surface-area-to-volume ratio—would gain a relative advantage over more common species, which have a lower ratio. As a result, we expect that the mean abundance of the dominant core species would decrease in spatially structured environments.

Experimental validation of our predictions could be achieved using engineered microbial consortia. For instance, a synthetic community could be constructed in which each strain lacks the ability to synthesize a single essential metabolite but can obtain it through cross-feeding with the others. Such syntrophic systems have been successfully engineered, ranging from two-species (50, 51) to multispecies communities (52). This approach would allow us to test our theoretical predictions under relaxed connectivity constraints, extending beyond the fully cooperative scenario assumed here. Minimal genetic modifications could be introduced into a single model organism (such as *E. coli*) to generate multiple auxotrophic variants that remain metabolically and physiologically equivalent, thus approximating a neutral interaction regime. By labeling some of these variants with a distinct fluorescent reporter, one could perform time-resolved sampling and quantify strain-specific abundance distributions using flow cytometry.

Further extensions of the theory should consider the heterogeneity of interactions, complex network topologies, and spatial effects. Another interesting direction would be to build a connection to the theory of autocatalytic chemical reaction networks, as studied in research on the origin of life. In particular, our approach is related to the “hypercycle” (53–56), a proposed model of collective chemical replication. In this model, a directed network of chemical species ( $X_i$ ) cross-catalyze the replication of other species ( $X_i + X_j \rightarrow X_i + X_i + X_j$  for neighbors  $i \rightarrow j$ ). This reaction scheme is equivalent to our replication rule, Fig. 1, therefore our model can be interpreted as a (stochastic) hypercycle on a fully connected network. In this context, our analysis could shed light on the emergence of cross-catalytic cores in replicating chemical systems.

**Data, Materials, and Software Availability.** There are no data underlying this work.

**ACKNOWLEDGMENTS.** J.P. and R.S. thank the hospitality of the Santa Fe Institute, where this work started. J.P. also thanks the hospitality of the Simon Levin Lab where some of this work was developed. R.S. acknowledges support from the Institució Catalana de Recerca i Estudis Avançats, the AGAUR 2021 SGR 0075 grant, a PID2023-152129NB-I00 grant funded by MICIU/AEI/10.13039/501100011033, and by the Santa Fe Institute. This project was partly supported by Grant No. 62417 from the John Templeton Foundation. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation. A.K. was partly supported by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 101068029. Finally, S.R. thanks NSF grant DMR-1910736 for financial support.

1. J. L. Bronstein, *Mutualism* (Oxford University Press, 2015).
2. E. G. Leigh Jr., The evolution of mutualism. *J. Evol. Biol.* **23**, 2507–2528 (2010).
3. F. Lutzoni, M. Pagel, Accelerated evolution as a consequence of transitions to mutualism. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11422–11427 (1997).

4. P. Schuster, How does complexity arise in evolution? *Evol. Complex Syst.* **22**, 147–159 (2001).
5. J. Bascompte, P. Jordano, Plant–animal mutualistic networks: The architecture of biodiversity. *Annu. Rev. Ecol. Evol. Syst.* **38**, 567–593 (2007).
6. S. Suweis, F. Simini, J. R. Banavar, A. Maritan, Emergence of structural and dynamical properties of ecological mutualistic networks. *Nature* **500**, 449–452 (2013).



7. M. J. Müller, B. I. Neugeboren, D. R. Nelson, A. W. Murray, Genetic drift opposes mutualism during spatial population expansion. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 1037–1042 (2014).
8. S. Germerodt *et al.*, Pervasive selection for cooperative cross-feeding in bacterial communities. *PLoS Comput. Biol.* **12**, e1004986 (2016).
9. E. J. Culp, A. L. Goodman, Cross-feeding in the gut microbiome: Ecology and mechanisms. *Cell Host Microbe* **31**, 485–499 (2023).
10. K. Zengler, L. S. Zaramela, The social network of microorganisms-how auxotrophies shape complex communities. *Nat. Rev. Microbiol.* **16**, 383–390 (2018).
11. K. Z. Coyte, J. Schluter, K. R. Foster, The ecology of the microbiome: Networks, competition, and stability. *Science* **350**, 663–666 (2015).
12. B. E. Morris, R. Henneberger, H. Huber, C. Moissl-Eichinger, Microbial syntrophy: Interaction for the common good. *FEMS Microbiol. Rev.* **37**, 384–406 (2013).
13. S. Saavedra, F. Reed-Tsochas, B. Uzzi, A simple model of bipartite cooperation for ecological and organizational networks. *Nature* **457**, 463–466 (2009).
14. R. M. May, *Stability and Complexity in Model Ecosystems* (Princeton University Press, 2019).
15. S. Allesina, S. Tang, Stability criteria for complex ecosystems. *Nature* **483**, 205–208 (2012).
16. S. P. Hubbell, *The Unified Neutral Theory of Biodiversity and Biogeography* (MPB-32) (Princeton University Press, 2011).
17. R. V. Solé, D. Alonso, A. McKane, Self-organized instability in complex ecosystems. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* **357**, 667–681 (2002).
18. U. Bastolla *et al.*, The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature* **458**, 1018–1020 (2009).
19. F. S. Valdivinos, Mutualistic networks: Moving closer to a predictive theory. *Ecol. Lett.* **22**, 1517–1534 (2019).
20. P. R. Guimarães Jr., The structure of ecological networks across levels of organization. *Annu. Rev. Ecol. Syst.* **51**, 433–460 (2020).
21. I. Volkov, J. R. Banavar, S. P. Hubbell, A. Maritan, Neutral theory and relative species abundance in ecology. *Nature* **424**, 1035–1037 (2003).
22. J. Harte, *Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics* (OUP Oxford, 2011).
23. S. Azaele *et al.*, Statistical mechanics of ecological systems: Neutral theory and beyond. *Rev. Mod. Phys.* **88**, 035003 (2016).
24. J. Grilli, Macroecological laws describe variation and diversity in microbial communities. *Nat. Commun.* **11**, 4743 (2020).
25. J. P. O'Dwyer, J. L. Green, Field theory for biogeography: A spatially explicit model for predicting patterns of biodiversity. *Ecol. Lett.* **13**, 87–95 (2010).
26. R. Muneerakul *et al.*, Neutral metacommunity models predict fish diversity patterns in Mississippi-Missouri basin. *Nature* **453**, 220–222 (2008).
27. J. Rosindell, S. P. Hubbell, R. S. Etienne, The unified neutral theory of biodiversity and biogeography at age ten. *Trends Ecol. Evol.* **26**, 340–348 (2011).
28. A. T. Neu, E. E. Allen, K. Roy, Defining and quantifying the core microbiome: Challenges and prospects. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2104429118 (2021).
29. G. Wu *et al.*, A core microbiome signature as an indicator of health. *Cell* **187**, 6550–6565 (2024).
30. E. Simpson, Measurement of diversity. *Nature* **163**, 688 (1949).
31. A. E. Magurran, *Measuring Biological Diversity* (John Wiley & Sons, 2003).
32. S. P. Hubbell, A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs. *Coral Reefs* **16**, S9–S21 (1997).
33. M. Malek-Mansour, G. Nocolis, A master equation description of local fluctuations. *J. Stat. Phys.* **13**, 197–217 (1975).
34. A. McKane, D. Alonso, R. V. Solé, Mean-field stochastic theory for species-rich assembled communities. *Phys. Rev. E* **62**, 8466 (2000).
35. M. R. Zaghloul, Efficient multiple-precision computation of the scaled complementary error function and the Dawson integral. *Numer. Algorithms* **95**, 1291–1308 (2024).
36. S. G. Johnson, Faddeeva W function implementation (2012). <http://ab-initio.mit.edu/Faddeeva>.
37. B. Ottino-Löffler, M. Kardar, Population extinction on a random fitness seascape. *Phys. Rev. E* **102**, 052106 (2020).
38. C. Tu, S. Suweis, J. Grilli, M. Formentin, A. Maritan, Reconciling cooperation, biodiversity and stability in complex ecological communities. *Sci. Rep.* **9**, 5580 (2019).
39. J. N. Holland, D. L. DeAngelis, J. L. Bronstein, Population dynamics and mutualism: Functional responses of benefits and costs. *Am. Nat.* **163**, 277–295 (2005).
40. M. Loftus, S. A. D. Hassounah, S. Yooseph, Bacterial associations in the healthy human gut microbiome across populations. *Sci. Rep.* **11**, 2828 (2021).
41. P. I. Costea *et al.*, Enterotypes in the landscape of gut microbial community composition. *Nat. Microbiol.* **3**, 8–16 (2018).
42. R. Vergnon, E. H. Van Nes, M. Scheffer, Emergent neutrality leads to multimodal species abundance distributions. *Nat. Commun.* **3**, 663 (2012).
43. D. Gonze, L. Lahti, J. Raes, K. Faust, Multi-stability and the origin of microbial community types. *ISME J.* **11**, 2159–2166 (2017).
44. H. Hinrichsen, Pair contact process with diffusion: A new type of nonequilibrium critical behavior? *Phys. Rev. E* **63**, 036102 (2001).
45. M. Bessonov, R. Durrett, Phase transitions for a planar quadratic contact process. *Adv. Appl. Math.* **87**, 82–107 (2017).
46. I. Iacopini, G. Petri, A. Barrat, V. Latora, Simplicial models of social contagion. *Nat. Commun.* **10**, 2485 (2019).
47. J. H. Kim, K. I. Goh, Higher-order components dictate higher-order contagion dynamics in hypergraphs. *Phys. Rev. Lett.* **132**, 087401 (2024).
48. S. Suweis, F. Ferraro, C. Grilletta, S. Azaele, A. Maritan, Generalized Lotka-Volterra systems with time correlated stochastic interactions. *Phys. Rev. Lett.* **133**, 167101 (2024).
49. R. S. Etienne, D. Alonso, Neutral community theory: How stochasticity and dispersal-limitation can explain species coexistence. *J. Stat. Phys.* **128**, 485–510 (2007).
50. W. Shou, S. Ram, J. M. Vilar, Synthetic cooperation in engineered yeast populations. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 1877–1882 (2007).
51. D. R. Amor, R. Montañez, S. Duran-Nebreda, R. Solé, Spatial dynamics of synthetic microbial mutualists and their parasites. *PLoS Comput. Biol.* **13**, e1005689 (2017).
52. M. T. Mee, J. J. Collins, G. M. Church, H. H. Wang, Syntrophic exchange in synthetic microbial communities. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E2149–E2156 (2014).
53. M. Eigen, P. Schuster, *The Hypercycle: A Principle of Natural Self-Organization* (Springer Science & Business Media, 2012).
54. E. Szathmáry, L. Demeter, Group selection of early replicators and the origin of life. *J. Theor. Biol.* **128**, 463–486 (1987).
55. P. Schuster, Some mechanistic requirements for major transitions. *Philos. Trans. R. Soc. B Biol. Sci.* **371**, 20150439 (2016).
56. R. Solé, M. De Domenico, Bifurcations and phase transitions in the origins of life. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **380**, 20240295 (2025).