SCIENTIFIC REPORTS

OPEN

Reconciling cooperation, biodiversity and stability in complex ecological communities

Chengyi Tu[1,2], Samir Suweis[3], Jacopo Grilli[4,5], Marco Formentin[6] & Amos Maritan[3]

Empirical evidences show that ecosystems with high biodiversity can persist in time even in the presence of few types of resources and are more stable than low biodiverse communities. This evidence is contrasted by the conventional mathematical modeling, which predicts that the presence of many species and/or cooperative interactions are detrimental for ecological stability and persistence. Here we propose a modelling framework for population dynamics, which also include indirect cooperative interactions mediated by other species (e.g. habitat modification). We show that in the large system size limit, any number of species can coexist and stability increases as the number of species grows, if mediated cooperation is present, even in presence of exploitative or harmful interactions (e.g. antibiotics). Our theoretical approach thus shows that appropriate models of mediated cooperation naturally lead to a solution of the long-standing question about complexity-stability paradox and on how highly biodiverse communities can coexist.

Research in population dynamics has a long history dating back to almost one thousand years ago with Fibonacci's modeling of the rabbit population. Nevertheless, it is still under debate which are the mechanisms allowing the coexistence of many interacting species in the same environment[1–3]. The current loss of Earth's biodiversity[4] makes this open question of great relevance today more than ever, and this challenge calls for interdisciplinary approaches. Historically, the Lotka and Volterra (LV) equations have provided much theoretical guidance and several microscopic derivations of these equations have been proposed[1,5–7]. Furthermore, these equations are the core of most of the multi-species deterministic population dynamics models based on the ecological concept of niche partitioning: competing species in order to coexist need to interact with the environment differently and to rely on not-overlapping resources. Recent works[8–10] have studied the conditions for persistence of many species in random Lotka-Volterra equations, obtaining a trade-off between diversity and strength of interactions.

While consumer-resource dynamics and related models of population dynamics with prey-predator and competitive interactions have been extensively studied, cooperative interactions (mutualism and/or commensalism), which are beneficial to one or both the involved species, have historically received less attention[11]. The current approach to mutualistic population dynamics is a simple extension of the LV types of models, which does not change the functional form of the two-species interaction in the phenomenological equations. Indeed beneficial ($+$$+$) instead of predator-prey ($+$$-$, here called exploitative) interactions are utilized[1,12]. In particular, a microscopic derivation of the phenomenological equations specific for the population dynamics in mutualistic communities is still missing. Moreover, a generalization of the stability-complexity theorem[1,13] has revealed that mutualism is even more detrimental to stability as the product $S \times c$ increases[3,14–16], where S is the number of species and c the connectivity, i.e. the fraction of non-zero pairwise interactions between species. This prediction clashes with the observation of widespread mutualistic interactions (or other cooperative interactions) in many natural and laboratory communities and where the biodiversity is very high[17–20], although other cases have been also reported[21].

[1]School of Ecology and Environmental Science, Yunnan University, 650091, Kunming, China. [2]Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA, 94720-3114, USA. [3]Department of Physics and Astronomy "Galileo Galilei", University of Padova, Via Marzolo 8, 35131, Padova, Italy. [4]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM, 87501, USA. [5]The Abdus Salam International Centre for Theoretical Physics (ICTP), Strada Costiera 11, 34014, Trieste, Italy. [6]Department of Mathematics "Tullio Levi-Civita", University of Padova, Via Trieste 63, 35129, Padova, Italy. Chengyi Tu and Samir Suweis contributed equally. Correspondence and requests for materials should be addressed to S.S. (email: suweis@pd.infn.it) or M.F. (email: marco.formentin@unipd.it) or A.M. (email: maritan@pd.infn.it)

1
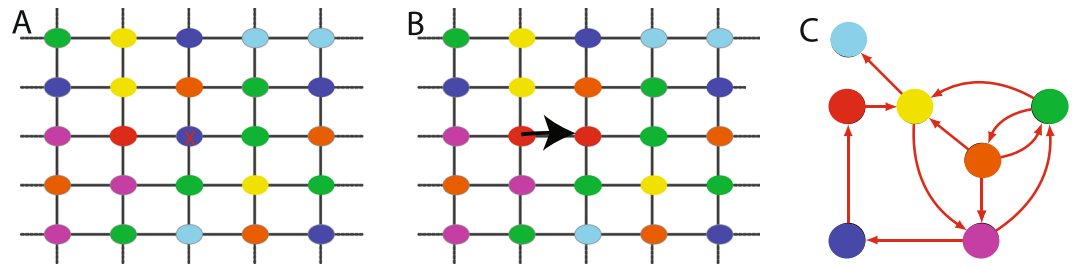
Figure 1. Graphical representation of our model. (**A,B**) The population dynamics is stochastic: at each time step a randomly chosen individual of a given species (denoted by the color) die and it is replaced by another species that is also picked at random and give birth to an offspring. The species birth rate depends on the species population abundance and on the species interaction matrix (**C**) - as given by Eq. (1).

An alternative theoretical approach to niche-based multi-species deterministic modeling is the Neutral Theory (NT) of Biodiversity[2,22,23]. In NT organisms of a community have identical per-capita probabilities of giving birth, dying, migrating, and speciating, regardless of the species they belong to. In this sense, NT is symmetric and aims to model only species on the same trophic level, therefore competing for the same pool of resources. An important example of the neutral model is the voter model (VM)[24]. VM was initially introduced to model the dynamics of two or more competing opinions of a group of people and its evolution in a discrete time setting can be described as follow: at each time step a randomly chosen individual changes his/her opinion to agree with one of her/his neighbors picked at random (see Fig. 1). In the ecological context one deals with a community of N individuals belonging to S different species. In its simplest version, a randomly selected individual dies and the corresponding resources are freed up for colonization by other species at the rate proportional to their abundances. An important limitation of this modeling is that, apart from competition, it does not explicitly consider species interactions (e.g. mutualism/commensalism). Although it has been already shown that niche and neutral approaches are only apparently contrasting[25,26], and few stochastic models have studied the impact of cooperative interactions on species coexistence and biodiversity patterns[19,27], two crucial issues in the current literature are: (i) the lack of a general framework specifically developed to model indirect (e.g. habitat modification or cross-feeding) cooperative interactions; (ii) the role of cooperative interactions in determining species coexistence and how they impact on patterns such as ecosystem stability[3,14,28].

In this work, we thus present a theoretical framework where, starting from a VM-like microscopic stochastic modelling, we add interactions among species and we properly account for the effect of the different type of ecological interactions. These interactions affect neutrality and lead, in their mean field formulation, to an emergent multi-species (implicit resource) model where cooperation is specifically modelled to take in account niche construction (by habitat modification through, for example, of release of secreted compounds). Reconciling apparently contrasting observations and previous results[1,3,13–15], we show that, in our model ecosystem, cooperation promotes biodiversity and diversity increases stability.

Results

Voter model with ecological interactions: breaking neutrality. In details, let be η_z the species label at spatial position z, where $\eta_z \in \{1, \ldots, S\}$ and $z = 1, \ldots, N$. The state at time t of the system is given by $\eta(t) = (\eta_1(t), \eta_2(t), \ldots, \eta_N(t)) \in \{1, \ldots, S\}^N$. We also set $\bar{\eta}^k$ to be the fraction of individuals of the k-species present in the system. Both intra- and inter-species competition is indirectly accounted by fixing an hard constraint on the system carrying capacity[2,23], i.e. an organism comes into the system always at the expenses of another organism. We now consider two main types of explicit ecological relationships: predator-prey or exploitative interactions $(+ -)$ and cooperation $(+ +, + 0)$. In the following we will also study the effect of harmful relationships $(- 0)$ that can be mediated by secreted compounds (e.g. fosfomycin in bacteria communities).

We thus introduce a directed graph on the set $\{1, \ldots, S\}$, where the nodes correspond to species and directed links represent the network of species interactions. Such a network is defined through cooperation matrix M_{ij} and exploitation matrix L_{ij} satisfying the following conditions: (i) For all $i, j = 1, \ldots, S$, $M_{ij} \geq 0$; (ii) For all $i, j = 1, \ldots, S$, it must be $L_{ij}L_{ji} < 0$ or $L_{ij} = L_{ji} = 0$; (iii) For all $i, j = 1, \ldots, S$, we have $L_{ij}M_{ij} = 0$, i.e. species i and j cannot simultaneously have both mutualistic and exploitative interactions.

In ecological terms, given two species i and j, a directed link of strength M_{ij} from i to j means that the j-th species receives a beneficial effect from the interaction with the i-th species, while $L_{kl} > 0 (<0)$ and $L_{lk} < 0 (>0)$ denotes that the l-th species exploits (is exploited by) the k-th species. For instance, in the former case we can think a microbial community where the presence of a certain species creates an environment by secreting metabolites, which modifies the niches and favors the growth of other bacteria[20]; in the latter one, we may think to host-parasite symbiosis. Typically, it is very difficult to measure the strength of the interactions among two species, so we adopt the standard approach of drawing the matrix entries from a given bivariate probability distribution (e.g. Gaussian or Uniform) in the same spirit as traditionally done[1,14]. If Q is a given matrix, we define c_Q its connectivity as well as μ_Q, σ_Q and ρ_Q the mean, standard deviation and correlation of the ij and ji entries, respectively.

The dynamics is described by a continuum time stochastic Markov process: a randomly chosen individual is removed and substituted by an individual of the j-th species at a rate

$$\omega(j, \eta, M, L) = \overline{\eta}^j + \epsilon_1\sum_{k=1}^{S}\overline{\eta}^k M_{kj}\theta(\overline{\eta}^j) + \epsilon_2\sum_{k=1}^{S}\overline{\eta}^k L_{kj}\overline{\eta}^j \tag{1}$$

where $\epsilon_1 > 0$ and $\epsilon_2 > 0$ give the cooperation and exploitation intensity, and $\theta(\cdot)$ is the Heaviside step function, i.e. $\theta(x) > 0$ when $x > 0$ and 0 otherwise. The presence of the θ-function in the mutualistic contribution, guarantees that the transition rate is zero if the j-th species is extinct. For $\epsilon_1 = \epsilon_2 = 0$ we recover the standard VM. When $\epsilon_1 > 0$ the species j is favored by the presence of the other species (k in the summation) to which it is connected and by their population. On the other hand, $\epsilon_2 > 0$ allows the possibility that a species exploits (or is exploited by) one or more other species.

It is important to highlight the differences of the contribution on Eq. (1) between exploitative and cooperative interactions. In the first case, the interaction term is quadratic in $\overline{\eta}$ (i.e. $\overline{\eta}^k L_{kj}\overline{\eta}^j$), as exploitative interactions can be derived using the law of mass-action used to describe chemical reactions[5,6]: a contact must occur between species and the chance of this interaction is, in the simplest hypothesis, proportional to both species concentrations. On the other hand, in cooperative interactions used to model effect of habitat modification or resource exchanges have a linear contribution to the birth rate in $\overline{\eta}$ (i.e. $\overline{\eta}^k M_{kj}$). In fact, if some species can generate a favorable environment for the growth of other species (like for example anaerobic conditions by aerobes, or the widespread phenomenon of growth facilitation through released metabolites driving non-limiting species co-occurrence[20]), then the linear term present in our equations just takes into account that the growth promoting effects are proportional to the population density of the species creating these conditions, but not to the population density of the species profiting of them. A further mechanism possibly behind the linear interaction term is indeed when interspecific cross-feeding (e.g. pollen, faecal pellets, metabolic waste) is present. Indeed, if a species is favored by a secreted compound released by another species, then what really contributes to the birth rate of the former is the amount of proper resources in the environment, which is proportional to the abundance of the latter (we assume that these resources are always fully utilized in the community and do not consider the limiting case of very low densities). If a species generates nourishment for more than one species, say k species, the hypothesis under which this leads to the above mentioned linear interaction term is when the secreted compound is used by each of the k species at a different time (see Supplementary Information, Section S1). This hypothesis has been used in many studies in cases where microbes tend to utilize nutrients in a specific sequential order[29], as observed in Bacteroide species[30] and in other studies[29]. Additionally, we also provide a mathematical derivation of the linear contribution to the birth rate based on the above mechanisms (see Supplementary Information, Section S1).

The microscopic dynamics given by rates Eq. (1) induces a Markovian evolution on the relative abundance $\overline{\eta}^s$ of each species. Standard techniques[31] can be used to prove that as $N \to \infty$, the process $(\overline{\eta}^1(t), \ldots, \overline{\eta}^S(t))_{t\geq 0}$ weakly converges to the solution of the system of ordinary differential (mean field) equation:

$$\frac{d}{dt}\overline{\eta}^s(t) = \epsilon_1\sum_{k=1}^{S}\overline{\eta}^k(t)M_{ks}\theta(\overline{\eta}^s(t)) + \epsilon_2\sum_{k=1}^{S}\overline{\eta}^k(t)L_{ks}\overline{\eta}^s(t)$$
$$- \overline{\eta}^s(t)(\epsilon_1\overline{\eta}^k(t)M_{ki}\theta(\overline{\eta}^i(t)) + \epsilon_2\overline{\eta}^k(t)L_{ki}\overline{\eta}^i(t)) \tag{2}$$

for $s = 1, \ldots, S$ where $\sum_{j=1}^{S}\overline{\eta}^j(t) = 1$ and is conserved by the dynamics.

Emergent ecological patterns.   Through Eq. (2) we can study many ecosystem properties of interest. One of the most important and studied emergent pattern in ecology, which we can determine within our model, is the relative species abundance (RSA)[2,22,23]. It describes commonness and rarity of species, thus characterizing the biodiversity of an ecological community. In our model, the RSA is given by the mean field stationary solution $(m_1, \ldots, m_S)$, which in turn depends on the species interaction matrices M and L.

The cumulative RSA is thus defined as the fraction of species with population greater that a certain value n, $P_>[n] = \frac{1}{S}\sum_{k=1}^{S}\theta(n - Nm_k)$ where we have fixed $N = 1/\min(m_1, \ldots, m_S)$ when all species coexist, i.e. we have made the choice that the rarest species has population equal to 1. We numerically find that the stationary RSA displays a log-normal shape, as the one found in many real ecosystems[2], and weakly depends on the specific distribution of the matrix elements M_{ij} and L_{ij} (see Fig. 2). Indeed, it is mainly determined only on its coefficient of variation, CV, i.e. the standard deviations of interaction strengths M and L relative to its mean (see Supplementary Information, Section S2). This allows to constrain the model parameters: in order to parametrize species interactions strengths, that are typically unknown, we use the random matrix approach[1,14,15], where we fix the mean and the standard deviation according to the desired RSA one needs to fit. Our deterministic approximation for the RSA neglects the fluctuations due to demographic stochasticity, which on the other hand decreases as $1/\sqrt{N}$ and thus when N is large enough, its effect becomes negligible.

We now consider species abundance fluctuations, defined as $x_N^i(t) = \sqrt{N}(\overline{\eta}^i(t) - m_i)$ for $i = 1, \ldots, S$. Another relevant quantity characterizing the ecosystem biodiversity is the covariance matrix V, $V_{ij} = \langle x^i(t)x^j(t)\rangle - \langle x^i(t)\rangle\langle x^j(t)\rangle$, describing the correlations in the population abundance fluctuations between pairs of species population abundances[32]. In our setting, we can compute analytically this quantity in the limit of normal fluctuations. The stochastic process $(x_N^1(t), \ldots, x_N^S(t))$ converges in distribution to a Gaussian Markov process $X := (X^1(t), \ldots, X^S(t))$, which solves the stochastic differential equation $dX = AX\,dt + \Phi dB_t$, where B_t is a S-dimensional Brownian motion, which corresponds to a S-dimensional Ornstein-Uhlenbeck process[31,33]. The analytical expressions for the matrices A and Φ in terms of the interaction matrices M and L, and of the equilibria, $(m_1, \ldots, m_S)$, of Eq. (2), are given in the Supplementary Information, Section S2. The covariance matrix, V, can be obtained by solving the following Lyapunov matrix equation $AV + VA^T + \Phi\Phi^T = 0$.
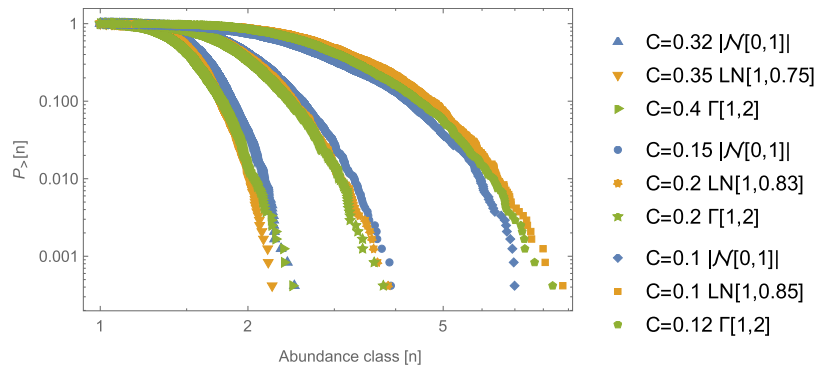
Figure 2. Cumulative RSA for a network of 300 species. The matrix elements of both M_{ij} and L_{ij} have been drawn from three different probability distributions ($z_h \sim p_h(z)$, $M_{ij} \sim z_h$, $L_{ij} \sim z_h$, $L_{ji} \sim -z_h$, $h = 1, 2, 3$): the modulus of a Normal distribution $z_1 \sim |\mathcal{N}(\alpha, \beta)|$ (blue lines), Gamma distribution $z_2 \sim \Gamma(\alpha, \beta)$ (green lines) and LogNormal distribution $z_3 \sim LN(\alpha, \beta)$ (orange lines). Connectivity for mutualistic interaction, M, is denoted by $C_M = C$, while for exploitative interactions, L, is denoted by $C_L = 0.1C$ (in all the studied cases $C_M + C_L \geq 1$). The cooperative and exploitative intensities are $\epsilon_1 = \epsilon_2 = 1$. We set the distribution parameters α, β (see legend) so that in each case we build interaction matrices with three different values of coefficient of variation CV $\approx 2, 3, 4$. Log-log plots display the relative species abundances (RSA) from the stationary solution of Eq. (1) averages over 100 realizations and normalized so that the smallest population is 1. As we can see, the cumulative RSA is not very sensible to the distribution from which the matrix elements of both M_{ij} and L_{ij} are drawn, but only on the CV.

This quantity is typically measured from species population time series, through the Pearson (or other type of) correlations[34]. Moreover, in many studies once the threshold is set opportunely, it is used as an empirical proxy of the species interactions matrix[34,35]. In other words, many works assume that $M + L$ can be approximated through V. Other works, applying maximum entropy approach, use V^{-1} as the quantity to describe the species interactions network[32]. However, we find that both V and V^{-1} are not good proxies of the species interactions matrix $M + L$ (see Supplementary Information, Section S2). This result highlights the importance to properly infer interaction networks from data[34] by considering a suitable model, which explicitly takes into account species interactions.

Our shift in the assumptions behind cooperative interactions enlightens some theoretical problematic aspect in order to explain species coexistence and ecosystem stability as observed in many real complex ecological communities. In fact, we now show that cooperation promotes species biodiversity and strongly stabilizes the ecosystem dynamics.

The importance of mediated cooperation for preserving biodiversity. If we focus on the purely cooperative voter model ($\epsilon_1 = \epsilon$ and $\epsilon_2 = 0$), we are able to analytically relate key dynamical features of Eq. (2) to the topology of the interaction matrix M and prove various results of ecological importance.

First, we show that the presence of non-supported species–the i-th species is non-supported if $\sum_j M_{ji} = 0$ – inhibits coexistence equilibria of the whole ecological community. More precisely, if species i is non-supported by other species then at stationarity Eq. (2) implies that $m_i = 0$. The extinction of the i-th may create new unsupported species that goes to zero in the large time limit. Such a cascade of extinctions may eventually end only when $\sum_j M_{ji} > 0$ for all nodes/species i of the network (see Supplementary Information, Section S3). The elimination of nodes of the interaction network corresponding to all non-supported species will be called pruning in the following.

Furthermore, we have found sufficient conditions on the topology of the cooperative interaction matrix M for the existence of stable stationary states of Eq. (2). In fact, if M is irreducible, i.e. if for any node i we can reach any other node j through a path of oriented links (k, l) such that $M_{kl} > 0$, then the Perron-Frobenius (PF) theorem holds[36] and there is a unique non-trivial stationary state $(m_1, \ldots, m_S)$ with only positive entries. This solution is proportional to the left eigenvector, v, of M corresponding to the eigenvalue of M with the largest modulus, which turns out to be non-degenerate, real and positive[36], denoted by α in the following (and that for brevity we will refer to it as PF eigenvalue). In other words, if M satisfies the PF theorem, then α tells us how the stationary species abundances m are distributed (see Methods). The corresponding right eigenvector will be denoted by w, and it gives information on how press perturbations spread throughout the network[15]. All components of both v and w are strictly positive and $m_i = v_i / \sum_k v_k$. An example of irreducible matrix M occurs when $M_{ij} > 0$ implies $M_{ji} > 0$ and the network has a single connected component. Many network architectures that have been observed in natural ecological communities satisfy this condition (e.g. hierarchical modular structure in mutualistic networks)[18].

Therefore, within our framework, we can analytically study the impact of the species interaction network architecture on system stability and species extinction. Two simple examples are shown corresponding to an ecosystem with no extinction (see Fig. 3A,B) and with extinction (see Fig. 3C,D). Additionally, a real-world example is also shown (see Fig. 3E,F).
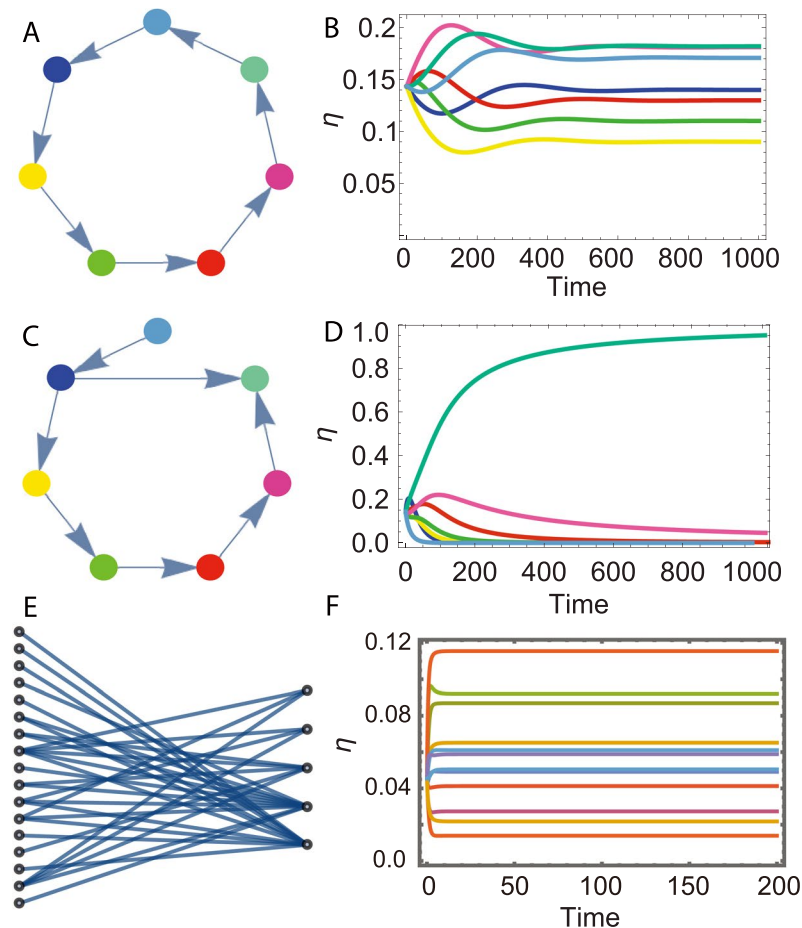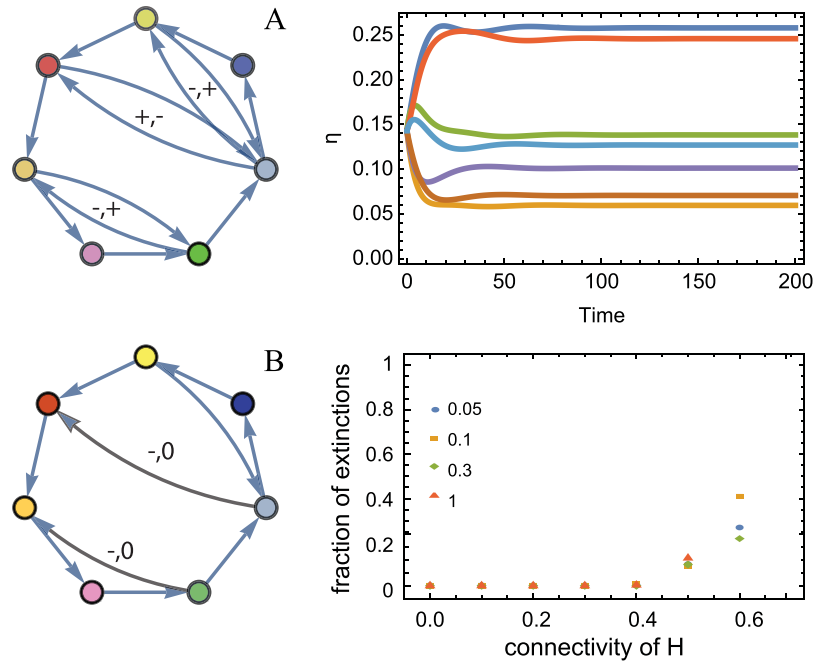
Figure 3. The results of the mean field predictions. (**A**) Species interaction network for 7 species where each species i has one mutualistic partner j, i.e. $M_{ij} = 1, \epsilon_1 = 1, \epsilon_2 = 1$. (**B**) Time evolution of the populations of the 7 species as predicted by the mean field dynamics Eq. (2). (**C**) Species interaction network for 7 species where one species is not helped by any species and the iterative pruning process, as described in the main text, leads to a cascade of extinctions. (**D**) As the time evolution of the mean field Eq. (2), only one species dominates the community. (**E**) Nested structure for fruit eating birds community in Mexico[39]. (**F**) All species coexist, as predicted by our theoretical framework. In the ordinate axis use the notation $\overline{\eta}$ and not η.

We then turn on also exploitative interactions ($\epsilon_2 > 0$) and we numerically find that, even in presence of relatively large concentrations of exploitative interactions (c_L), at stationarity the system still admits an high biodiversity and full coexistence is observed, as long as a mutualistic network of interactions is present, corresponding to an irreducible matrix, M (see Fig. 4A and Methods). Further discussion can be found in the Supplementary Information, Section S4.

Effect of harmful interactions.   The current model can also take into account the harmful interactions such as impact of noxious secreted compounds (e.g. antibiotics and fosfomycin molecule) on species biodiversity. We can thus introduce a further directed network H on the set $\{1, …, S\}$, where H_{ij} representing indirect harmful relationships (-0), i.e. a directed link of strength H_{ij} from i to j means that the j-th species receives a harmful effect from the i-th species. The overall species interaction network is defined through three matrices M, L and H where M_{ij} representing cooperation interactions $(++, +0)$, L_{ij} representing exploitation interactions $(+-)$ and H_{ij} representing harmful interactions (-0). These matrices satisfy the following conditions for all $i, j = 1, …, S$: (i) $M_{ij} \geq 0$. A directed link of strength M_{ij} from i to j means that the j-th species receives a beneficial effect from the interaction with the i-th species; (ii) $L_{ij}L_{ji} < 0$ or $L_{ij} = L_{ji} = 0$. $L_{ij} > (<0)$ and $L_{ji} < 0(>0)$ denotes that the j-th species exploits (is exploited by) the i-th species; (iii) $H_{ij} < 0, H_{ji} = 0$ or $H_{ij} = 0, H_{ji} < 0$ or $H_{ij} = H_{ji} = 0$. A directed link of strength H_{ij} from i to j means that the j-th species receives a harmful effect from the interaction with the i-th species; (iv) $L_{ij}M_{ij}H_{ij} = 0$ that is each pair of species, i and j, can have at most one of the above interactions.

The term of harmful interactions is linear in $\overline{\eta}$ (i.e. $\overline{\eta}^k H_{kj}$), so the dynamics as induced by the Markovian evolution on the relative abundance $\overline{\eta}^s$ of each species lead to the following mean field equation:

Figure 4. The fraction of extinctions of network with exploitative or harmful interactions as well as cooperative interactions. (**A**) Species interaction network for 7 species where each species i has one mutualistic partner j, i.e. $M_{ij} = 1$, some species have exploitative interactions ($+ -$) and cooperative and exploitative intensities are $\epsilon_1 = \epsilon_2 = 1$. The corresponding time evolution of the populations of the 7 species, as predicted by the mean field dynamics Eq. (2), are also shown. During the time evolution, the rates given by Eq. (1) remain positive and extinctions are not observed. (**B**) Example of a network with harmful interactions and plot of the fraction of extinctions (defined as the fraction of species extinct in the community) as a function of the connectivity of harmful matrix H (c_H) and with connectivity of the cooperative matrix M ($c_M = 1 - c_H$) for different average interaction strengths (colored points) $\mu_M = -\mu_H = 0.05, 0.1, 0.3, 1$ (see legend) and cooperative and harmful intensities are $\epsilon_1 = \epsilon_H = 1$. The off-diagonal elements of matrices M and H are drawn from normal distribution, i.e. $M_{ij} \sim z$, H_{ij}, $H_{ji} \sim 0, -z$, $z \sim N(\mu_M, \mu_M/3)$. The network size considered in these simulations is $S = 50$ and each point is the mean of 10 realizations. Similar results are found also for $S = 20$, $S = 100$ (not shown here).

$$\frac{d}{dt}\overline{\eta}^s(t) = \epsilon_1 \sum_{k=1}^{S} \overline{\eta}^k(t) M_{ks} \theta(\overline{\eta}^s(t)) + \epsilon_2 \sum_{k=1}^{S} \overline{\eta}^k(t) L_{ks} \overline{\eta}^s(t) + \epsilon_3 \sum_{k=1}^{S} \overline{\eta}^k(t) H_{ks} \theta(\overline{\eta}^s(t))$$
$$- \overline{\eta}^s(t) \sum_{i,k=1}^{S} (\epsilon_1 \overline{\eta}^k(t) M_{ki} \theta(\overline{\eta}^i(t)) + \epsilon_2 \overline{\eta}^k(t) L_{ki} \overline{\eta}^i(t) + \epsilon_3 \overline{\eta}^k(t) H_{ki} \theta(\overline{\eta}^i(t)))$$

(3)

for $s = 1, \ldots, S$ where $\sum_{j=1}^{S} \overline{\eta}^j(t) = 1$ and is conserved by the dynamics.

For simplicity but without loss of generality, we turn off the exploitations (given by the L matrix) and thus consider only the interaction matrices M and H, i.e. we focus on cooperation and on the effects harmful interactions mediated by secreted compounds. We discuss the results for a fully connected network, $c_M + c_H < 1$, since the case $c_M + c_H < 1$ does not alter our main conclusions. Of course, we only consider feasible and stable solutions. To do that, we first find the stationary solutions and then calculate the Jacobian matrix evaluated at that solution to check the stability. Figure 4B shows the fraction of extinction as a function of c_H. We consider an extinction to occur if the species abundance goes below an extinction threshold $1/S^2$, where S is the number of species (notice that the average population density in the uniform case is $1/S$). When $c_H > 0$ the PF theorem cannot be applied. However, we start to detect extinctions only after a critical threshold of harmful interactions in the system and the extinction rate is an increasing function of c_H. For c_H larger than 0.6, the only feasible solutions are not stable and thus the extinction rate is not shown. Therefore, as expected, the existence of such molecules challenges the unlimited coexistence seen in the purely cooperative model.

Cooperation determines ecosystem stability. We now study analytically the stability of the equilibria for an ecosystem with cooperative and exploitative interactions, by analyzing the eigenvalues of the linearization of Eq. (2), i.e. the Jacobian matrix J, around the equilibria, m_i, as a function of ecological complexity ($c \times S$). We set the diagonal elements of M to zero and the off-diagonal pair (M_{ij}, M_{ji}) is equal to $(0, 0)$ with probability $1 - c_M$ and with probability c_M drawn from a bivariate Gaussian distribution of means $(\mu, \mu)^T$ and covariance matrix $\Sigma = (\sigma^2, \rho\sigma^2; \rho\sigma^2, \sigma^2)$. This guarantees that, for a connected cluster, coexistence of all species occurs. The mean, variance and correlation of the non-diagonal elements of matrix M are $\mu_L = c_L\mu$, $\sigma_M^2 = c_M\sigma^2 + c_M(1 - c_M)\mu^2$

6
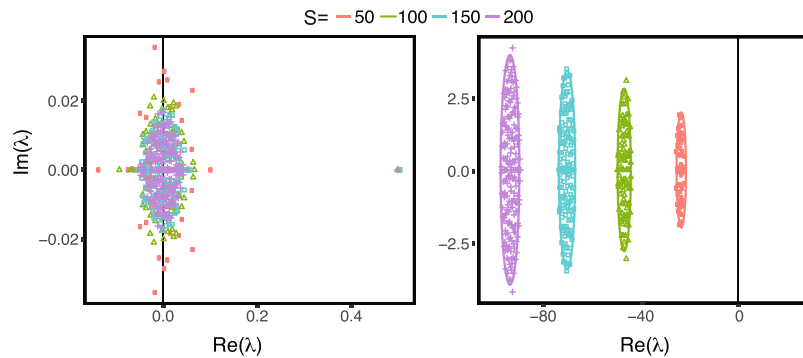
Figure 5. Eigenvalues spectrum of the Jacobian matrix J. (**A**) Pure exploitative interactions where $C_L = 1$, $\epsilon_2 = 1$ and $\epsilon_1 = 0$. (**B**) Exploitative and mutualistic interactions where $C_L = C_M = 0.5$, $\epsilon_1 = 0.75$ and $\epsilon_2 = 1$. The off-diagonal elements of matrices M and L are drawn uniformly between 0 and 1, i.e. $M_{ij} \sim z$, L_{ij}, $L_{ji} \sim \pm z$, $z \sim U(0, 1)$. The points are the eigenvalues of one Jacobian matrix obtained sampling at random the matrices M and L, while the lines indicate the analytical prediction for the support of the J eigenvalues in the corresponding cases (see Eq. (4)). The black vertical line indicates the instability threshold.

and $\rho_M = \frac{\rho \sigma^2 + (1 - c_M)\mu^2}{\sigma^2 + (1 - c_M)\mu^2}$. The case in which each element of M_{ij} is assigned independently of M_{ji} simply corresponds to the case $\rho = 0$ (notice that even if $\rho = 0$ we can have $\rho \neq 0$). The exploitative matrix L is considered similarly. If $\mu_M \geq \sigma_M \sqrt{(1 + \rho_M)/S}$, the leading eigenvalue $\lambda_M = S\mu_M = Sc_M\mu$ and the corresponding eigenvector has positive components[14]. Moreover, the components of the leading eigenvector are approximately constant, i.e. the equilibria of system given by Eq. (2) can be written as $m_i = (1 + \xi_i)/S$ for $i = 1, \ldots, S$ with $\sum_i \xi_i = 0$. Using the fact that $1 = Sm_i - \xi_i$, $\lambda_M = S\mu_M$ and taking into account that all the terms involving ξ_j are sub-leading in S, we obtain that the leading term of the system Jacobian does not depend on L (see Methods) and it is equal to:

$$J_{ij} = -\delta_{ij}S\mu_M + \left(M_{ij} - \mu_M\right) = -\delta_{ij}S\mu_M + M'_{ij} \tag{4}$$

where $M'_{ij} := M_{ij} - \mu_M$ is a random matrix with zero mean, variance σ_M^2 and correlation ρ_M. This implies that the eigenvalues are uniformly distributed in an ellipse centered around $-S\mu_M$ with semi-axis $\sqrt{S}\sigma_M(1 + \rho_M)$ and $\sqrt{S}\sigma_M(1 - \rho_M)$[14]. The largest eigenvalue of the Jacobian is therefore given by $-S\mu_M + \sqrt{S}\sigma_M(1 + \rho_M)$. Thus, for fixed connectivity, c, in the presence of cooperation the system stability increases with S, where as if only predator-prey interactions are present, then the system is always unstable for large enough S, as the May theorem would predict (see Fig. 5 and Supplementary Information, Section S5).

Discussion

In this work we have proposed how to incorporate the effect of indirect cooperation induced by habitat modification or cross-feeding[19,27,37] in order to have effective equations where resources are not explicitly modelled (as in Lotka Volterra equations). This framework is a possible choice to model species dynamics with implicit resources and cooperative interactions. The introduction of indirect mutualism changes the whole stability profile of the corresponding system's dynamics. This result highlights the importance of having a microscopic derivation of the phenomenological models and the birth rate should be accurately chosen depending on the specific ecological system one wants to model. Our results can be applied to study the effect of the interaction network topology on species coexistence in real ecological communities. In particular, we found that nested architecture, observed in plant-pollinators ecological communities[15,18], where specialist species, with only few mutualistic links, tend to interact with a proper subset of the many mutualistic partners of any of the generalist species, (see Fig. 3E) satisfies the hypothesis of the PF theorem and thus favor species coexistence (see Fig. 3F). We have then shown that by properly deriving the contribution of cooperation in the species population dynamics, we solve two long-standing problems in theoretical ecology: how a large number of species can coexist together and the complexity-stability paradox. In fact, we found that cooperation promotes ecosystem biodiversity, that in turn increases its stability without any fine-tuning of the species interaction strengths or of the self-interactions. Even moderate cooperative interactions can stabilize the dynamics and the stability increases with the ecosystem complexity (see Fig. 5). In our framework we can also model and studies the effect of harmful secreted compound, such as antibiotics. We observe that these harmful interactions may lead to the extinctions of targeted species, but also in this case our conclusions still hold: cooperation promotes ecosystem biodiversity and stability. Of course, if we add direct cooperation beside the mediated mutualism, such as classic Holling quadratic interactions term, as long as indirect linear mutualism is present, we observe violation of the classic complexity-stability paradox (see Supplementary Information, Section S6).

Our results can be compared to the ones obtained recently by Servan et al.[9], where it was shown that in random Lotka-Volterra systems many species were coexisting. It is important to underline, that that result was obtained assuming the stability of the interaction matrix, as the focus was to study the effect of feasibility. If the interaction matrix is not stable, in fact, one should expect more extinctions[10]. Conversely, in this work, a high

level of coexistence is obtained without assuming the stability of the interaction matrix or, equivalently, small interaction strengths. All presented results consider effective competition by fixing the system carrying capacity, i.e. an organism comes into the system always at the expenses of another organism. The results do not change when the hard constraint of total fixed population size is relaxed by introducing the possibility for a site to become empty (see Supplementary Information, Section S7).

Finally, we may also consider a constant effort hypothesis[38] in regulating species interactions. This would translate in normalizing the strengths of each interaction matrix by the by corresponding species in-degree (see Supplementary Information, Section S8). Nevertheless, in all these different scenarios our conclusions still hold: we find that a shift in the assumptions behind cooperative interactions resolve long-standing open theoretical question on the relation between stability and complexity and provides a unifying modeling approach to describe emergent patterns in ecology and interacting large ecological systems, as observed recently in real microbial communities[27,37].

Methods

Application of the Perron-Frobenius theorem to the model equations.
Let us consider the dynamics given by Eq. (2) for $\epsilon_2 = 0$. If M is irreducible, then the PF theorem holds[36] and given the initial condition $\bar{\eta}^s(0) > 0$ where $s = 1, \ldots, S$, the time dependent solution for the species fractions is

$$\bar{\eta}(t) = \frac{\bar{\eta}(0)^T e^{\epsilon M t}}{\sum_s (\bar{\eta}(0)^T e^{\epsilon M t})_s} \tag{5}$$

Since for any eigenvalue, $\beta \neq \alpha$, of M we have $\Re(\beta) < \alpha$, the dominant term in both numerator and denominator in Eq. (5) is $v\, e^{\alpha t}(\bar{\eta}(0) \cdot w)$ leading to $\lim_{t \to \infty} \bar{\eta}(t) = \frac{v}{\sum_s v_s} = m$. This is an easy computation when M has a basis of eigenvectors and in general can be derived using the Jordan decomposition. As a corollary of the derivation above, we have also that the stationary solution is globally stable in the region $\bar{\eta}^s(0) > 0$ for all $s = 1, \ldots, S$.

Analytical justification of the coexistence condition in the presence of exploitative interactions.
As explained in the main text, if the matrix M is irreducible and the transition rates given by Eq. (1) are positive during the time evolution (a necessary condition in order that the derivation of the mean field is justified), then we find numerically that, even in presence of a large concentrations of exploitative interactions, at stationarity the system still admits an high biodiversity and full coexistence is observed. Here we want to heuristically justify what we have observed numerically. Adding exploitative interactions does not lead to extinction, as long as the mutualistic network of interactions is present, corresponding to an irreducible matrix, M. We argue that, under this hypothesis, when $\bar{\eta}^s$ is positive but close to zero the complete mean field equations - where both $\grave{o}_1$ and $\grave{o}_2$ are positive - are perturbation of the mean field equation where only mutualistic interaction are present, since we have proved that a pure mutualistic system has no extinction as long as the matrix M is irreducible. Following the notation in Results, our continuous time Markov process is defined by the rule: a randomly chosen individual is removed and substituted by an individual of the j-th species at a rate

$$\omega(j, \eta, M, L) = \underbrace{\bar{\eta}^j + \epsilon_1 \sum_{k=1}^S \bar{\eta}^k M_{kj}\theta(\bar{\eta}^j)}_{:=\omega_j^M} + \underbrace{\epsilon_2 \sum_{k=1}^S \bar{\eta}^k L_{kj}\bar{\eta}^j}_{:=\omega_j^L} \tag{6}$$

where $\epsilon_1 > 0$ and $\epsilon_2 > 0$ give the cooperation and exploitation intensity, and $\theta(\cdot)$ is the Heaviside step function, i.e., $\theta(x) > 0$ when $x > 0$ and 0 otherwise. As $N \to \infty$ the relative abundance $\bar{\eta}^s$ converges to the solution of the system of ordinary differential equation for $s = 1, \ldots, S$. Equation for $\bar{\eta}^s$, when $\bar{\eta}^s$ is positive but close to zero, can be written in the following form

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{\eta}^s(t) = \underbrace{\omega_s^M - \bar{\eta}^s(t)\sum_{i=1}^S \omega_i^M}_{\simeq \delta > 0} + \mathcal{O}(\bar{\eta}^s) \tag{7}$$

The first two terms in Eq. (7) are the vector fields corresponding to mean field equation for M irreducible and no exploitation (i.e. $\epsilon_2 = 0$). We know that such a system has no extinction and its vector field is typically greater than $\delta > 0$ out of equilibrium when $\bar{\eta}^s \simeq 0$. The last term in Eq. (7) contains terms which are linear dependent of ω_j^L which is $\mathcal{O}(\bar{\eta}_s)$. Thus $\frac{\mathrm{d}}{\mathrm{d}t}\bar{\eta}^s(t)$ is positive for $\bar{\eta}^s$ close to zero. The requested transition rates never become negative during the time evolution of the mean field equation. This is a necessary condition otherwise the derivation itself of the mean filed equation would be meaningless.

Stability of the equilibria.
In the case of $\epsilon_2 \neq 0$, the entries of the Jacobian read

$$J_{ij} = \epsilon_1 \left(M_{ij}^T - \delta_{ij}\sum_{h,k=1}^S m_h M_{hk} - m_i \sum_{k=1}^S M_{jk} \right) + \epsilon_2 \left(L_{ji}m_i + \delta_{ij}\sum_{h=1}^S m_h L_{hi} - \delta_{ij}\sum_{h,k=1}^S m_h L_{hk}m_k \right.$$
$$\left. - m_i \sum_{k=1}^S L_{jk}m_k - m_i \sum_{k=1}^S L_{kj}m_k \right) \tag{8}$$

The diagonal entries of the Jacobian are

$$J_{ii} = -\epsilon_1 \sum_{h,k=1}^{S} m_h M_{hk} + \epsilon_2 \sum_{h=1}^{S} m_h L_{hi} - \epsilon_2 \sum_{h,k=1}^{S} m_h L_{hk} m_k \tag{9}$$

Since $m_i \sim 1/S$, it is simple to observe that the term proportional to ϵ_1 is of order S (plus sub-leading fluctuations). On the other hand, the leading order of the terms proportional to ϵ_2, is of order 1 and therefore always sub-leading if $\epsilon_1 > 0$. A similar argument applies to the off-diagonal elements. In that case, the terms proportional to ϵ_2 are of order 1, while the ones proportional to ϵ_2 are of order $1/S$.

Similarly to what found in the case $\epsilon_2 = 0$, we have that the following relations hold: $\mu_L = C_L \mu$, $\sigma_L = \sqrt{c_L(\sigma^2 + (1 - c_L)\mu^2)}$ and $\rho_L = \frac{\rho\sigma^2 + (1 - c_L)\mu^2}{\sigma^2 + (1 - c_L)\mu^2}$ where μ and σ are the mean and the standard deviation of the distribution from which we draw the value for the exploitative interaction strengths. These expressions have been used together with μ_M, σ_M and ρ_M, when calculating the coefficient of variation. The above considerations indicate that the distribution of the eigenvalues of the Jacobian, Eq. (9), is the same as the $\epsilon_2 = 0$ case. Figures visualizing these results are presented in the Supplementary Information, Section S5.

References

1. May, R. M. Will a large complex system be stable? *Nature* **238**, 413 (1972).
2. Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton University Press, 2001).
3. Coyte, K. Z., Schluter, J. & Foster, K. R. The ecology of the microbiome: networks, competition, and stability. *Science* **350**, 663–666 (2015).
4. Hull, P. M., Darroch, S. A. F. & Erwin, D. H. Rarity in mass extinctions and the future of ecosystems. *Nature* **528**, 345 (2015).
5. Matsuda, H., Ogita, N., Sasaki, A. & Satō, K. Statistical mechanics of population: the lattice Lotka-Volterra model. *Progress of theoretical Physics* **88**, 1035–1049 (1992).
6. Constable, G. W. A. & McKane, A. J. Models of genetic drift as limiting forms of the Lotka-Volterra competition model. *Physical review letters* **114**, 038101 (2015).
7. Chesson, P. Mechanisms of maintenance of species diversity. *Annual review of Ecology and Systematics* **31**, 343–366 (2000).
8. Grilli, J. *et al.* Feasibility and coexistence of large ecological communities. *Nature communications* **8**, 0 (2017).
9. Serván, C. A. *et al.* Coexistence of many species in random ecosystems. **2**, 1237 (2018).
10. Bunin, G. J. P. R. E. Ecological communities with Lotka-Volterra dynamics. **95**, 042414 (2017).
11. Kéfi, S., Holmgren, M. & Scheffer, M. When can positive interactions cause alternative stable states in ecosystems? *Functional Ecology* **30**, 88–97 (2016).
12. Holling, C. S. Some characteristics of simple types of predation and parasitism. *The Canadian Entomologist* **91**, 385–398 (1959).
13. McCann, K. S. The diversity–stability debate. *Nature* **405**, 228 (2000).
14. Allesina, S. & Tang, S. Stability criteria for complex ecosystems. *Nature* **483**, 205 (2012).
15. Suweis, S., Simini, F., Banavar, J. R. & Maritan, A. Emergence of structural and dynamical properties of ecological mutualistic networks. *Nature* **500**, 449 (2013).
16. Suweis, S., Grilli, J., Banavar, J. R., Allesina, S. & Maritan, A. Effect of localization on the stability of mutualistic ecological networks. *Nature communications* **6**, 10179 (2015).
17. Posfai, A., Taillefumier, T. & Wingreen, N. S. Metabolic trade-offs promote diversity in a model ecosystem. *Physical review letters* **118**, 028103 (2017).
18. Bascompte, J. & Jordano, P. Plant-animal mutualistic networks: the architecture of biodiversity. *Annu. Rev. Ecol. Evol. Syst.* **38**, 567–593 (2007).
19. Morris, B. E. L., Henneberger, R., Huber, H. & Moissl-Eichinger, C. Microbial syntrophy: interaction for the common good. *FEMS microbiology reviews* **37**, 384–406 (2013).
20. Zelezniak, A. *et al.* Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proceedings of the National Academy of Sciences*, 201421834 (2015).
21. Foster, K. R. & Bell, T. Competition, not cooperation, dominates interactions among culturable microbial species. *Current biology* **22**, 1845–1850 (2012).
22. Volkov, I., Banavar, J. R., Hubbell, S. P. & Maritan, A. Neutral theory and relative species abundance in ecology. *Nature* **424**, 1035 (2003).
23. Azaele, S. *et al.* Statistical mechanics of ecological systems: Neutral theory and beyond. *Reviews of Modern Physics* **88**, 035003 (2016).
24. Liggett, T. M. *Stochastic interacting systems: contact, voter and exclusion processes*. Vol. 324 (springer science & Business Media, 2013).
25. Segura, A. M. *et al.* Emergent neutrality drives phytoplankton species coexistence. *Proceedings of the Royal Society of London B: Biological Sciences* **278**, 2355–2361 (2011).
26. Fisher, C. K. & Mehta, P. The transition between the niche and neutral regimes in ecology. *Proceedings of the National Academy of Sciences* **111**, 13111–13116 (2014).
27. Goldford, J. E. *et al.* Emergent simplicity in microbial community assembly. *Science* **361**, 469–474 (2018).
28. Butler, S. & O'Dwyer, J. P. Stability criteria for complex microbial communities. *Nature communications* **9**, 2970, https://doi.org/10.1038/s41467-018-05308-z (2018).
29. Monod, J. The growth of bacterial cultures. *Annual Reviews in Microbiology* **3**, 371–394 (1949).
30. Tuncil, Y. E. *et al.* Reciprocal prioritization to dietary glycans by gut bacteria in a competitive environment promotes stable coexistence. *mBio* **8**, e01068–01017 (2017).
31. Ethier, S. N. & Kurtz, T. G. *Markov processes: characterization and convergence*. Vol. 282 (John Wiley & Sons, 2009).
32. Volkov, I., Banavar, J. R., Hubbell, S. P. & Maritan, A. Inferring species interactions in tropical forests. *Proceedings of the National Academy of Sciences* **106**, 13854–13859 (2009).
33. Gardiner, C. W. *Handbook of stochastic methods for physics, chemistry and the natural sciences*. Vol. 25 (1986).
34. Faust, K. & Raes, J. Microbial interactions: from networks to models. *Nature Reviews Microbiology* **10**, 538 (2012).
35. Lima-Mendez, G. *et al.* Determinants of community structure in the global plankton interactome. *Science* **348**, 1262073 (2015).
36. Seneta, E. *Non-negative matrices and Markov chains.* (Springer Science & Business Media, 2006).
37. Pacheco, A. R., Moel, M. & Segre, D. Costless metabolic secretions as drivers of interspecies interactions in microbial ecosystems. *bioRxiv*, 300046 (2018).
38. Suweis, S., Grilli, J. & Maritan, A. Disentangling the effect of hybrid interactions and of the constant effort hypothesis on ecological community stability. *Oikos* **123**, 525–532 (2014).
39. Kantak, G. E. Observations on some fruit-eating birds in Mexico. *The Auk* **96**, 183–186 (1979).

Acknowledgements

Author Contributions

S.S., M.F. and A.M. designed the study, C.T., J.G. and S.S. performed numerical simulations, A.M., M.F. and J.G. performed analytical calculations. All authors contributed in writing the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-41614-2.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.