SUPPLEMENTARY INFORMATION

doi:10.1038/nature12438

Contents

1 Mathematical Definition of Nestedness

In the main text, we refer to nestedness as the classical NODF measure (nestedness metric based on overlap and decreasing fill) [1]. Alternative measures of nestedness are possible. In particular, we have also tested our results with the definition of nestedness η introduced by Bastolla and collaborators [2]. In this section we briefly report and summarize the characteristics of these two nestedness measures. Let a_{ij}^{PA} be the adjacency matrix of the bipartite graph, where the index i (j) runs over the plant (insect) indices: it is 1 if the i-th plant is pollinated by the j-th insect and zero otherwise. The analogous adjacency matrix for insects is $a_{ji}^{AP} = a_{ij}^{PA}$. Henceforth the set of plant (insect) indices will be denoted P (A), whereas the number of plant (insect) species will be denoted by n_P (n_A). The number of links (degree) of the i-th plant (insect) is $k_i^P = \sum_j a_{ij}^{PA}$ ($k_i^A = \sum_j a_{ij}^{AP}$) and it gives the number of insects (plants) that the i-th plant (insect) interacts with. The number of common links (i.e. the number of common pollinators) the i-th and the j-th plant have (also called the unweighted overlap matrix) is given by

$$o_{ij}^P \equiv \sum_k a_{ik}^{PA} a_{jk}^{PA}. \tag{S1}$$

In an analogous way, one defines the overlap matrix o^A with respect to insects.

$NODF$ measure is defined as [1]

$$NODF = \frac{\sum_{i<j:\ i,j\in P} T_{ij}^P + \sum_{i<j:\ i,j\in A} T_{ij}^A}{\left[\frac{P(P-1)}{2}\right] + \left[\frac{A(A-1)}{2}\right]}, \tag{S2}$$

where $T_{ij}^X = 0$ if $k_i^X = k_j^X$, and $T_{ij}^X = o_{ij}^X / \min(k_i^X, k_j^X)$, when both i and j belong to the same set $X = A,\ P$.

The alternative definition of nestedness η is [2]:

$$\eta = \frac{\eta^P + \eta^A}{2} \qquad \text{where} \qquad \eta^X \equiv \frac{\sum_{i<j} o_{ij}^X}{\sum_{i<j} \min\left(k_i^X, k_j^X\right)}, \tag{S3}$$

All the results presented in this work qualitatively hold for both these measures of nestedness.

Both nestedness take values in the interval $[0, 1]$, where 1 designates a perfectly nested network and 0 indicates a network with no nestedness. They are highly and significantly correlated [3]. From Eq. (S3) a perfect nestedness is achieved, when the following hierarchy among plants and insects is found:

$$\mathcal{I}_1 \subseteq \mathcal{I}_2 \subseteq ... \subseteq \mathcal{I}_{n_X}, \text{for } X = P, A \tag{S4}$$

where $\mathcal{I}_j$ is the subset of species with which species j interacts, and the sub-indices are ordered from the most specialist plant (insect), i.e. the species that has the minimum

number of inter-species interactions, to the most generalist, i.e. one that has the maximum number of inter-species interactions. The main difference between Eq. (S3) and (S2) is that, in the latter case, in order to have perfect nestedness, the strict inequality in Eq. (S4) must hold.

2 Connectance of mutualistic networks

The connectance C_Γ is the percentage of mutualistic interactions over all possible ones ($n_A \cdot n_P$). All the results we have presented in the main text are obtained for connectance that scales with the number of species as $C_\Gamma \approx 4/S^{0.8}$, and where $S = n_A + n_P$. In fact, this is the best fit (Table S1) for $C_\Gamma(S) = aS^b$ using empirical ecological networks (see Figure S1). Data sets of empirical mutualistic networks were taken from the work of Rezende et al. [4].The community size is defined as the total number of plant and animal pollinator species. This result is in qualitative agreement with previous results on the scaling of the connectivity with the size of mutualistic ecological networks [5].

Table S1: Fit results and corresponding goodness of fit tests.

$C_\Gamma(S)$	**Estimate**	**Standard Error**	**t-Statistic**	**P-Value**
a	4.21414	1.71557	2.45641	0.0171028
b	0.776538	0.1107747	7.0101	$3.06089 \cdot 10^{-9}$

For simplicity and in the absence of data, we also set the connectance of the intra-species competition interaction matrix in the same way, i.e. $C_\Omega \approx 4/S^{0.8}$, but the conclusions are robust with respect to other choices.

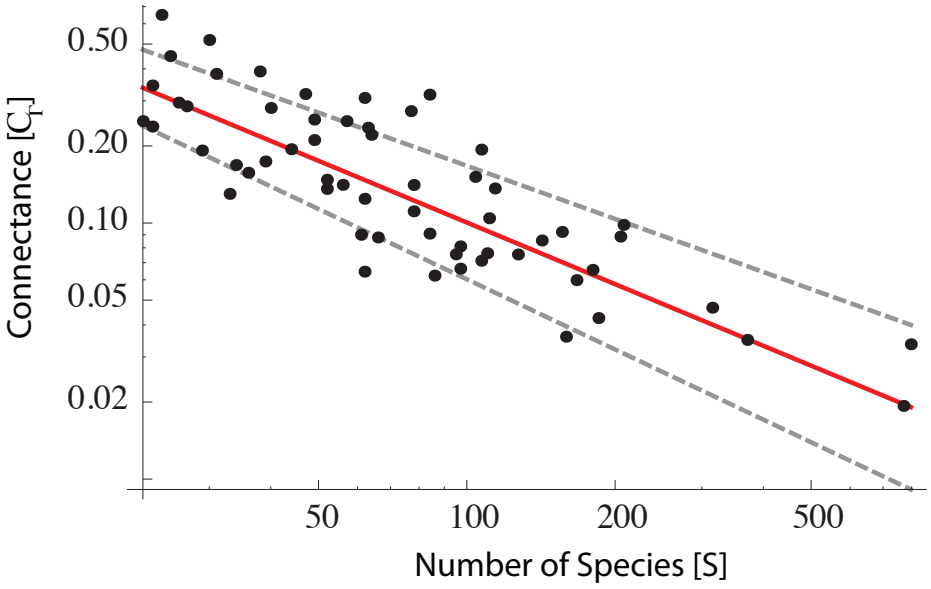

Figure S1: Best fit (red solid line) of the connectivity as a function of the number of species for 56 mutualistic communities. Dashed gray lines represent the region within the ± 1 standard deviation confidence interval for the exponent estimate. The plot is in log-log scale.

3 Mathematical Framework: Community Models

We present here the theoretical framework and mathematical details of the community dynamics model that incorporates mutualistic interactions between plants and pollinators or seed dispersers. We start with the general case of a community composed of S species comprising n_P plants and n_A animal insects or seed dispersers with $S = n_P + n_A$. Species in the same group are in competition with each other while interacting mutualistically with species in the other group.

x_i^P and x_j^A denote the abundances of the i-th plant species and the j-th animal pollinator species respectively and $\vec{x} = \{x_1^P, x_2^P, ..., x_{n_P}^P, x_{n_P+1}^A, ..., x_S^A\}$ is a S-component vector giving the abundance of each of the species in the community. x_i are variables. The intrinsic growth rates of the species in the absence of competition and mutualism is represented by the S-dimensional vector $\vec{\alpha}$. Species interact according to a given $S \times S$ interaction matrix M. The matrix is composed by four blocks, two describing the direct competitive interactions Ω_{PP} and Ω_{AA} among plants and insects respectively, and two assigning the mutualistic interactions Γ_{PA} and Γ_{AP} between insects and plants and vice-versa. Therefore the interaction matrix M has the following structure:

$$\begin{bmatrix} \Omega_{PP} & \Gamma_{PA} \\ \Gamma_{AP} & \Omega_{AA} \end{bmatrix} = \left(\begin{array}{cccc:cccc} d & \omega_{12} & \dots & \omega_{1n_P} & \gamma_{1n_P+1} & \dots & \dots & \gamma_{1S} \\ \omega_{21} & d & \dots & \dots & \gamma_{2,n_P+1} & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \omega_{n_P1} & \dots & \dots & d & \gamma_{n_Pn_P+1} & \dots & \dots & \gamma_{n_P,S} \\ \hdashline \gamma_{n_P+1,1} & \dots & \dots & \dots & d & \omega_{n_P+1n_P+2} & \dots & \omega_{n_P+1S} \\ \gamma_{n_P+2,1} & \dots & \dots & \dots & \dots & d & \dots & \omega_{n_P+2,S} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \gamma_{S1,1} & \dots & \dots & \dots & \omega_{Sn_P+1} & \dots & \dots & d \end{array} \right), \tag{S5}$$

where $M_{ii} = d > 0$ for $i = 1, ..., S$ represents the species self -limitation effect.

3.1 Holling Type I Functional Response

For the sake of mathematical simplicity, we start by representing both direct competitive and mutualistic interactions between species through a linear functional response of Holling Type I:

$$\frac{dx_i}{dt} = x_i \left(\alpha_i - \sum_j^S M_{ij} x_j \right) \equiv f(\vec{x}), \tag{S6}$$

The bipartite mutualistic interactions block matrices are built in the following way: for each pair γ_{ij}, γ_{ji} (with $i = 1, .., n_P$ and $j = n_P + 1, ...S$) we draw a random value p from a uniform distribution between zero and one ($U_{[0,1]}$). If $p < C_\Gamma$, then we set $\gamma_{ij} \sim \gamma_{ji} \sim -|\mathcal{N}(0, \sigma_\Gamma^2)|$, otherwise $\gamma_{ij} = \gamma_{ji} = 0$. $\mathcal{N}(\mu, \sigma^2)$ is the normal distribution with mean μ and variance σ^2, and determines the strength of the interactions among species. Analogously, the competition interactions ω_{ij} and ω_{ji} are drawn with probability C_Ω from

$|\mathcal{N}(0,\sigma_\Omega^2)|$. We note that because of this construction, the adjacency interaction matrix a (where $a_{ij}=1$ if i is connected to j and zero otherwise), is symmetric.

The local stability of the underlying population dynamics [6] is governed by the linearized system of differential equations $\dot{\overrightarrow{\delta x}} = \Phi\overrightarrow{\delta x}$ around the stationary point $\vec{x}^*$, where $\overrightarrow{\delta x} = \vec{x} - \vec{x}^*$ and the matrix Φ has elements

$$\phi_{ij} \equiv \left.\frac{\partial f_i(\vec{x}_i)}{\partial x_j}\right|_{\vec{x}^*} = -x_i^* M_{ij}, \tag{S7}$$

where $\vec{x}^* = M^{-1}\vec{\alpha}$ and $x_i^* > 0\ \forall i$. Therefore the main feature characterizing the local stability is the interaction matrix M.

Let us consider first the case where the stationary point of Eq. (S6) is $\vec{x}^* = (1,1,1...,1)$ so that $\Phi = -M$ and $\vec{\alpha} = M\cdot\vec{x}^*$. An approximate, but important indication of the stability of the stationary solution $\vec{x}^*$ for the community dynamics given Eq. (S6) can be obtained by extending methods recently used to study the stability of structured ecological networks [7] to our case (mixture of mutualistic and competition interactions). It can be shown that in the large S limit the maximum real part value of the eigenvalues (λ) of the linearized system Φ is

$$\begin{aligned}\max\Re(\lambda) &= -d + \max\left(\frac{(S-1)\left(C_\Gamma\sigma_\Gamma - C_\Omega\sigma_\Omega\right)}{\sqrt{2\pi}}, \frac{C_\Omega\sigma_\Omega - C_\Gamma\sigma_\Gamma}{\sqrt{2\pi}} + \right. \\ &+ \left.\frac{\sqrt{S}\left(4C_\Gamma\sigma_\Gamma C_\Omega\sigma_\Omega + \left(-2C_\Gamma+\pi+1\right)C_\Gamma\sigma_\Gamma^2 + \left(-2C_\Omega+\pi+1\right)C_\Omega\sigma_\Omega^2\right)}{\sqrt{2\pi}\sqrt{\left(2C_\Gamma\sigma_\Gamma C_\Omega\sigma_\Omega + \left(\pi-C_\Gamma\right)C_\Gamma\sigma_\Gamma^2 + \left(\pi-C_\Omega\right)C_\Omega\sigma_\Omega^2\right)}}\right).\end{aligned} \tag{S8}$$

By setting $\max\Re(\lambda) = 0$ and solving for σ_Γ we obtain the approximate critical strength for mutualistic interactions (for large S)

$$\sigma_c \approx \frac{\sqrt{2\pi}d}{S(1-\beta\nu)C_\Gamma} \quad \text{if } \beta\nu < 1 \tag{S9}$$

$$\begin{aligned}\sigma_c &\approx \frac{\sqrt{2\pi}d\sqrt{C_\Gamma\left(\pi\left(\beta^2+\nu^2\right)-(\beta-\nu)^2C_\Gamma\right)}}{\sqrt{S}C_\Gamma\left(\left(\beta^2+\nu^2\right)\left(-2C_\Gamma+\pi+1\right)+4\beta\nu C_\Gamma\right)} + \\ &+ \frac{\sqrt{2\pi}d(\beta\nu-1)\left((\beta-\nu)^2C_\Gamma - \pi\left(\beta^2+\nu^2\right)\right)}{S\left(\left(\beta^2+\nu^2\right)\left(-2C_\Gamma+\pi+1\right)+4\beta\nu C_\Gamma\right)^2} \quad \text{if } \beta\nu \geq 1\end{aligned} \tag{S10}$$

where we have set

$$C_\Omega = \nu C_\Gamma \tag{S11}$$

$$\sigma_\Omega = \beta\sigma_\Gamma. \tag{S12}$$

Setting $\nu = 1$, one has three different situations (see Table S2): if $\beta > 1$, then competition dominates over mutualism, if $\beta = 1$ competition and mutualism are comparable. Finally, if

$\beta < 1$, the mutualistic interactions dominate the competition. We remark that the results presented in Table S2 are only approximate. The extension to the Wigner semicircular law ([6]) for an asymmetric real matrix is exact only for the case when the correlation between matrix indices is solely between opposite diagonal elements, i.e.$\langle M_{ij} M_{kl} \rangle = 0$ if $k \neq j$ and/or $l \neq i$ [8]. For this reason, in order to ensure being within the stability region, it is enough to set $\sigma_\Gamma \ll \sigma_c$ (see Figure S2)

Table S2: *Different scenarios for the interactions strength and corresponding approximated stability regions for $S \gg 1$. We have imposed that the connectances in the direct competition and mutualistic matrices are the same ($\nu = 1$).*

Strengths of the interactions	**Scaling Stability Threshold**
weak Competition and strong Mutualism ($\beta < 1$)	$\sigma_c \sim 1/(C_\Gamma S)$
strong Competition and weak Mutualism ($\beta > 1$)	$\sigma_c \sim 1/(\sqrt{C_\Gamma S})$
Competition $\approx$ Mutualism ($\beta = 1$)	$\sigma_c \approx \pi/(\sqrt{C_\Gamma S}(1+\pi))$

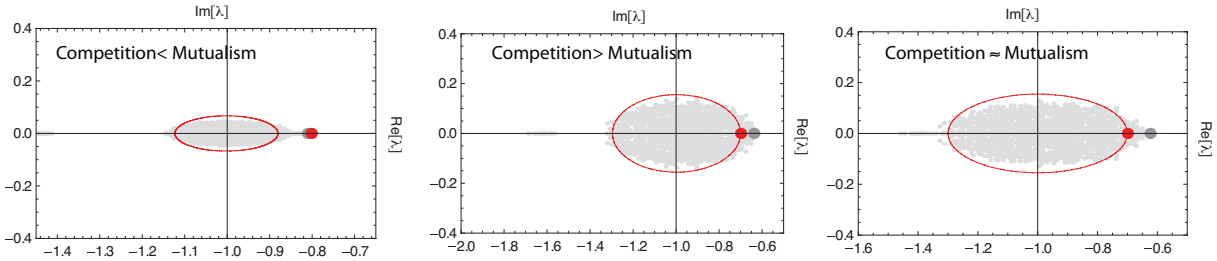

Figure S2: Eigenvalues spectrum of random structured ecological network (gray dots) with both mutualistic and intra-species competitions interactions as given by Eq. (S7). In red are shown the predictions of the extended Wigner semi-circular law [8, 7] of the distribution for the eigenvalues of the matrix Φ . We have analyzed three possible scenarios: $\beta = 3$ where competition is dominant with respect to mutualism, $\beta = 0.3$ (mutualism stronger than intra-species competition), and $\beta = 1$, i.e. the competition and mutualism are comparable. All networks are generated for S=100, $C_\Gamma = C_\Omega \sim 0.3$, $d = 1$ and $\sigma_\Gamma = 0.3\sigma_c$.

3.2 Holling Type II Functional Response

The theoretical framework can be extended to non linear functional responses. In particular, following Thebault and Fontaine [9], we model population community dynamics using a Holling type II functional response:

$$\frac{dx_i}{dt} = x_i\left(\alpha_i - dx_i - \frac{\sum_j M_{ij}x_j}{h_{ij} + \sum_k \theta_{ik}x_k}\right) \equiv g_i(\vec{x}), \tag{S13}$$

where $\theta_{ij} = 1$ if $i \in A(P)$ and $j \in P(A)$ and zero otherwise, while h_{ij} is a parameter having the role of a handling time for the interaction between i and j. In the following, for simplicity we set $h_{ij} = h$ for all i, j. This choice, although specific, leads to the same results one would get with random h_{ij} and it allows one to determine the stationary state in a much simpler way. We note that in this case we have isolated the self interaction term $(-dx_i)$ and thus the diagonal elements of M are zero, i.e. $M_{ii} = 0$ $(\forall i)$. Setting

$$b_{ij} = \frac{M_{ij}}{h + \sum_k \theta_{ik}x_k} = b_{ij}(x_P, x_A), \tag{S14}$$

where x_P (x_A) is the total plant (animal pollinator) population abundance, the stationary solution of Eq. (S13) is given by the self consistent equations

$$x_i^* = \sum_j [(d \cdot \mathbb{1} + b(x_P, x_A))^{-1}]_{ij} \cdot \alpha_j \tag{S15}$$

$$x_P = \sum_{j\in P} x_j^* \tag{S16}$$

$$x_A = \sum_{j\in A} x_j^*, \tag{S17}$$

that can be solved numerically.

The local stability of the dynamics governed by Eq. (S13) is given by the community matrix Φ $(\dot{\overrightarrow{\delta x}} = \Phi\overrightarrow{\delta x})$ of elements:

$$\phi_{ij} \equiv \frac{\partial g_i(\vec{x}_i)}{\partial x_j}_{|_{x^*}} = -x_i^*\left[\delta_{ij}d + \frac{M_{ij} + (d\,x_i^* - \alpha_i)\theta_{ij}}{h + \sum_j \theta_{ij}x_j^*}\right] \tag{S18}$$

where $\vec{x}^*$ is given by Eqs. (S15)-(S17) and $x_i^* > 0$ $\forall i$.

4 Algorithm for the Maximization of Species Abundance

4.1 Community-level Optimization

In order to implement the optimization principle, the most straightforward dynamics is the following. We start with $x_i = 1$ for $i = 1, ..., S$ and a randomly generated matrix M as explained above. $\vec{\alpha}$ is determined as $\vec{\alpha} = M\vec{x}^*$ ($\vec{\alpha} = (d\mathbb{1} + b)\vec{x}^*$ if Holling type II dynamics is considered). During the optimization dynamics $\vec{\alpha}$ is kept fixed and the matrix entries of Γ are swapped so to keep fixed the initial entry distribution. At each time step, we randomly pick a pair of interacting species and another pair which are not interacting (i.e. two pairs (i, j), (k, l) with $(i, j) \neq (k, l)$ and such that $\gamma_{ij} \neq 0$ and $\gamma_{kl} = 0$), and we switch the interaction between these two elements (i.e.$\gamma_{ij}^{n+1} = \gamma_{kl}^{n}$, $\gamma_{kl}^{n+1} = \gamma_{ij}^{n}$). We then calculate the new equilibrium point: $\vec{x}^*(n+1) = M^{-1}(n+1) \cdot \vec{\alpha}$. If $\sum_{j=1}^{S} x_j^*(n+1) \geq \sum_{j=1}^{S} x_j^*(n)$, then we accept the swap and $\Gamma = \Gamma(n+1)$, otherwise $\Gamma = \Gamma(n)$. We repeat this procedure for T steps. We checked that variants of the algorithm yields the same results. For instance, if we also allow for non-zero elements to switch, the resulting optimal interaction matrix does not change. For $\sigma < \sigma_c$ we also performe a simulated annealing [10] stochastic algorithm and we did not find relevant differences in the final results. Thus the final architecture of M_{opt} is largely independent of the details of the dynamics chosen to perform the maximization of the total population of the community. Our results do not change if the starting x_i^* are drawn at random uniformly in the range $[0.5, 1.5]$.

The total number of time steps T in the optimization procedure is chosen in a manner to ensure that in the final steps $\approx 90 - 95\%$ of the attempted interaction switches are rejected (see Figures S3-S5). The precise number of time steps depends on the particular setting for the optimization (S, σ_Γ, σ_Ω), but for simplicity we have set it to $T = 100S$.

4.2 Species-level Optimization

The above optimization principle may sound as if it is based on some sort of group selection mechanism. There is an interesting heated discussion on the role of group/kin selection in evolution [11, 12, 13]. Its not our purpose here to enter in this scientific discussion and our results do not have to be interpreted as evidence for a particular evolution mechanism.

In order to not rely on a interpretation of our algorithm based on group selection, we performed a variation of the optimization principle: The interactions of the community are rearranged in order to increase the abundance of the pollinator or plant species involved in the interaction rearrangement. In this way, the switching an interaction improves species' fitness (which, in the optimization, is approximated by abundance) and we have a more plausible mechanism behind the shift from a random to nested community.

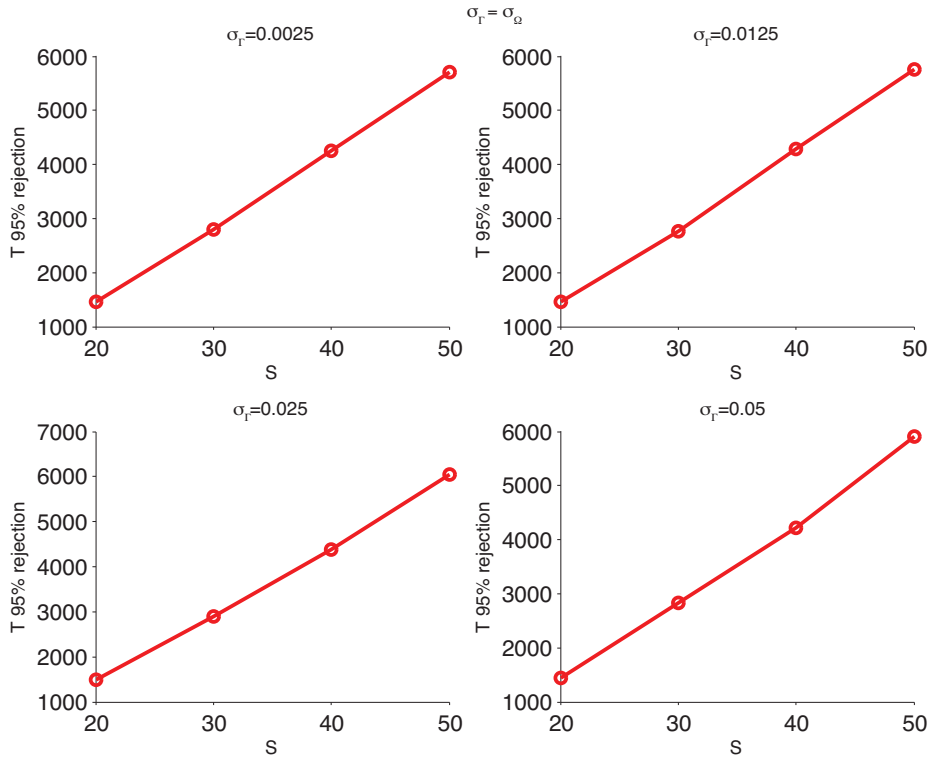

Figure S3: Number of time steps in order to reach the threshold corresponding to rejection of 95% of the attempted interaction switches during the optimization of the community population abundance for the case competition≈mutualism ($\beta = 1$) and with $C_\Gamma = C_\Omega = 4/S^{0.8}$.

The optimization algorithm is similar to the one presented in the previous section: We start with $x_i = 1$ for $i = 1, ..., S$ and a randomly generated matrix M; $\vec{\alpha}$ is also determined as before. During the optimization dynamics $\vec{\alpha}$ is kept fixed and the matrix entries of Γ are swapped so to keep fixed the initial entry distribution. At each time step, we select a species, j, randomly and an existing link to one of its partner species k; attempt a rewire between the $j-k$ and the $j-m$ links where m is a putative mutualistic partner species, that is γ_{jk} is interchanged with γ_{jm}. If the $j-m$ link already exists, i.e. γ_{jm} is different from zero, the switch leads to an interchange of interaction strengths, otherwise the swap corresponds to moving the $j-k$ link to $j-m$. The switch is accepted if and only if it does not lead to a decrease of the population abundance of species j in the steady state corresponding to the new network configuration. We repeat this procedure for T steps (see Figure S6). In order to be sure to reach a local maxima, for this case we set the number of steps scale super-linearly with the size $T = 100S^{3/2}$ because the population dynamics is slower than the previous algorithm. We find that also in this case a nested

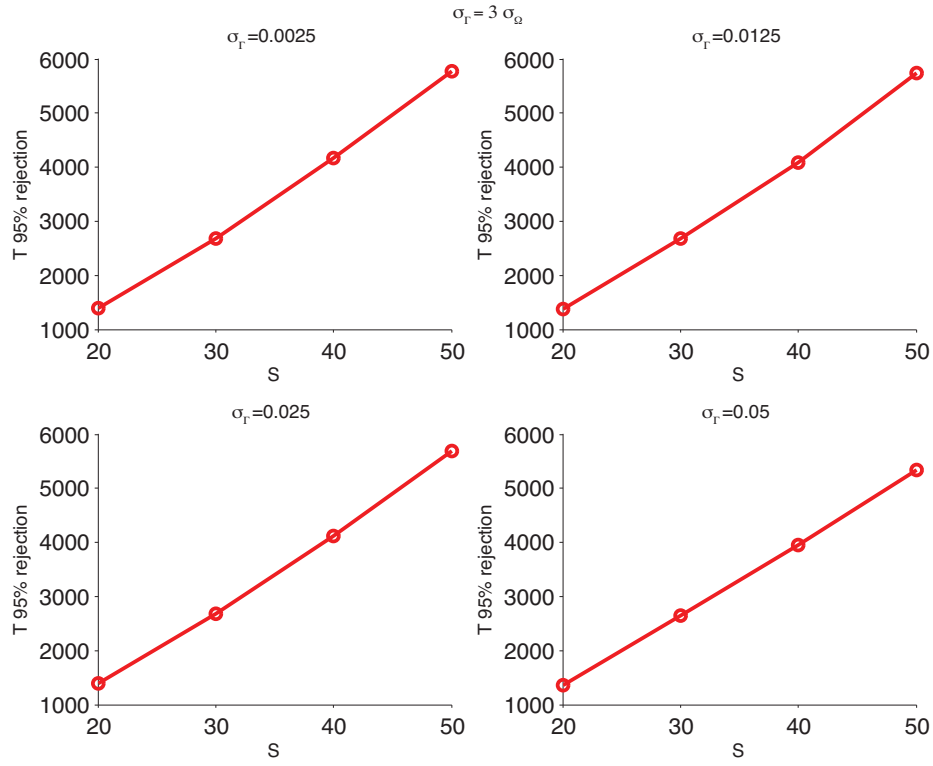

Figure S4: Number of time steps in order to reach the threshold corresponding to the rejection of 95% of the attempted interaction switches during the optimization of the community population abundance for the case competition>mutualism ($\beta = 3$) and with $C_\Gamma = C_\Omega = 4/S^{0.8}$.

community architecture emerges for the mutualistic community and all the related results still hold.

4.3 Relation Between Species-level Optimization and Community-level Optimization

The compatibility of individual species fitness and group selection is a convenient feature of mutualistic networks. This non-trivial collective behavior can be demonstrated in the following way: the stationary community abundance is given by $x_i^* = \sum_j M_{ij}^{-1}\alpha_j^*$. If we impose small variations to the structure of the interactions, i.e. $M + \delta b$, their effect propagates to the community populations in a highly non linear way: $x_i^* \rightarrow x_i^* - \sum_{j,k,l} M_{ik}^{-1}\delta M_{kl} M_{lj}^{-1}\alpha_j$. Thus, the local interaction switch has an impact on the populations of all the species in the communities, generating a non trivial "collective behavior" (Figure S7), where the benefit to a single species in the long run contributes to

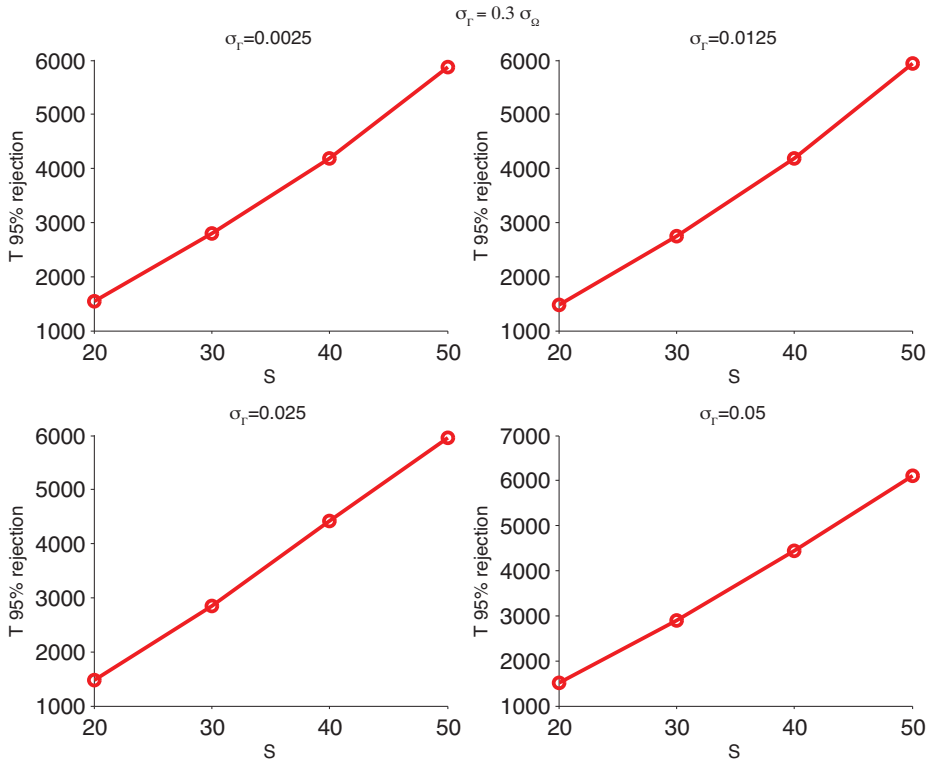

Figure S5: Number of time steps in order to reach the threshold corresponding to the rejection of 95% of the attempted interaction switches during the optimization of the community population abundance for the case competition<mutualism ($\beta = 0.3$) and with $C_\Gamma = C_\Omega = 4/S^{0.8}$.

increase the group fitness (i.e. increase the community population - see Figure S6).

As we will show in section 5, the relation between species-level optimization and the total number of individuals in the community can be also studied analytically using a mean field approximation. Indeed we show that an increase in the population of an individual species leads to a net increase of the population abundance of the whole community.

4.4 Optimization while Assembling Communities

We also tested a more realistic scenario where the interaction network is progressively "assembled" over the course of time [14, 15, 16] and meanwhile it is "reorganized" toward a more optimal state. We start with four mutualistic communities with five plant and five pollinators that randomly interact with probability C. In each one of the communities, we maximize the abundance of the individuals in the community as described in section 4.1. After 250 attempted swaps we merge communities 1 with 2, and 3 with 4 so to have two

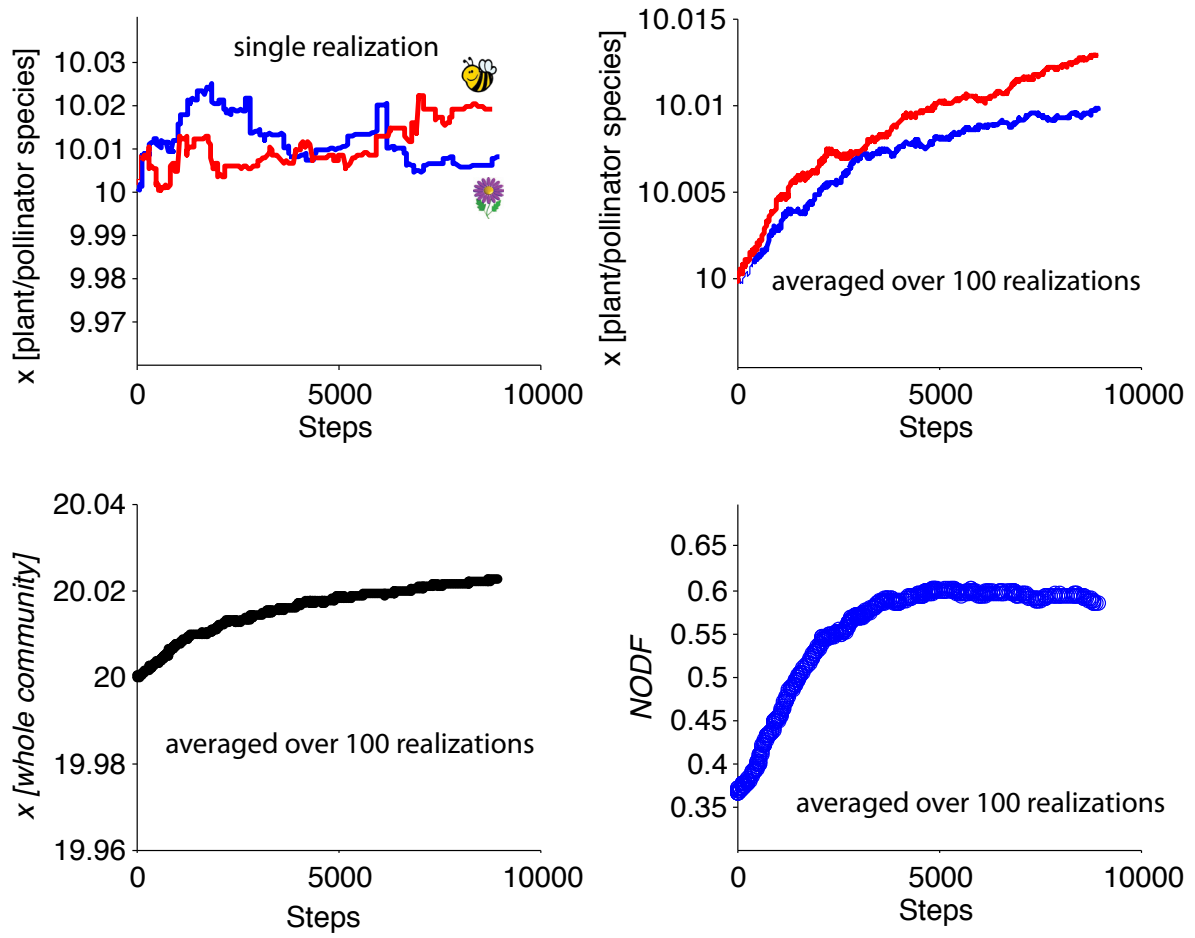

Figure S6: Adaptive framework algorithm: at each time step, we randomly pick a species i and one of its interacting partners j. We then pick another species k which may or may not interact with i. We rewire the interaction between these two elements. If the abundance of the initially picked species i does not decrease after the interaction switch, then we accept the swap. The parameters used in this simulations are $S = 20$ and $C = 4/S^{0.8}$. Holling Type II dynamics has been used.

communities with 10 plants and 10 pollinator species. We again turn on the optimization for $T = 500$ time steps. At this point we merge the two communities so to form a bigger mutualistic community with $S = 40$ species and we continue the optimization for $T = 2000$. As shown by Figure S8, the final architecture is again nested.

Indeed, our results thus prove to be very robust and largely independent of the details of the optimization algorithm. In fact the spontaneous emergence of generalist and specialist species from the optimization of the population abundances is independent of the

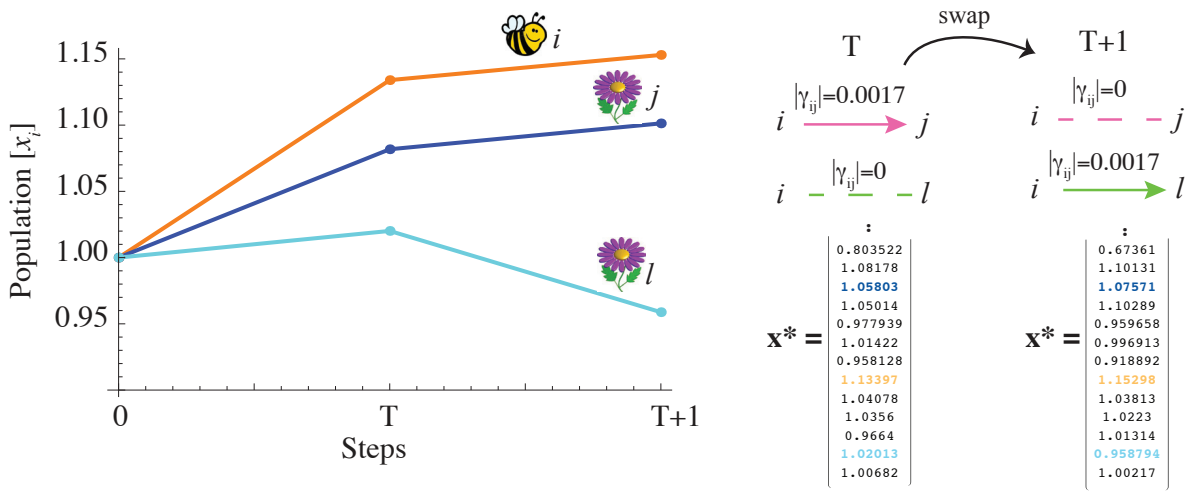

Figure S7: Example of the non trivial collective behavior emergent from the interaction switch of a species. The local rewiring between partners $i-j$ and $i-l$ impact the new stationary abundances $\vec{x}$ of all the species in the community. In a single interaction swap, some species may be disfavored but in the long run there is an average increase of the total population size of both pollinator and plant species (see Figure S6).

specific dynamical rules of the proposed variational principle as long as stationary conditions are reached. The dynamical rules are not meant to mimic the actual dynamics of a community. Rather, they are efficient means of probing the stationary state characterized by the maximal fitness (individual and collective).

4.5 Swapping algorithm: visualization of the architecture

In order to provide a visual indicator of the presence of a nested structure in the interaction matrix, we have applied the following procedure:

1. We considered the two off-diagonal blocks of the interaction matrix, containing the Plant-Animal Pollinator, Γ_{PA} $(n_A \times n_P)$, and Animal Pollinator-Plant, Γ_{AP} $(n_P \times n_A)$, interactions respectively. We applied the next steps to each block separately.

2. We set all non-zero interactions to one, i.e we consider only unweighted adjacency matrices $\gamma_{ij}^{PA(AP)} \to a_{ij}^{PA(AP)} = 1$ if $\gamma_{ij}^{PA(AP)} \neq 0$.

3. Using Γ_{PA} we calculate the number of Plant species that interact with each Animal Pollinator species, and vice-versa using Γ_{AP}. For example, for Animal Pollinator species i we compute $k_i^A = \sum_{j=1}^{n_P} a_{ij}^{PA}$, and for Plant species i we compute $k_i^P = \sum_{j=1}^{n_A} a_{ij}^{AP}$.

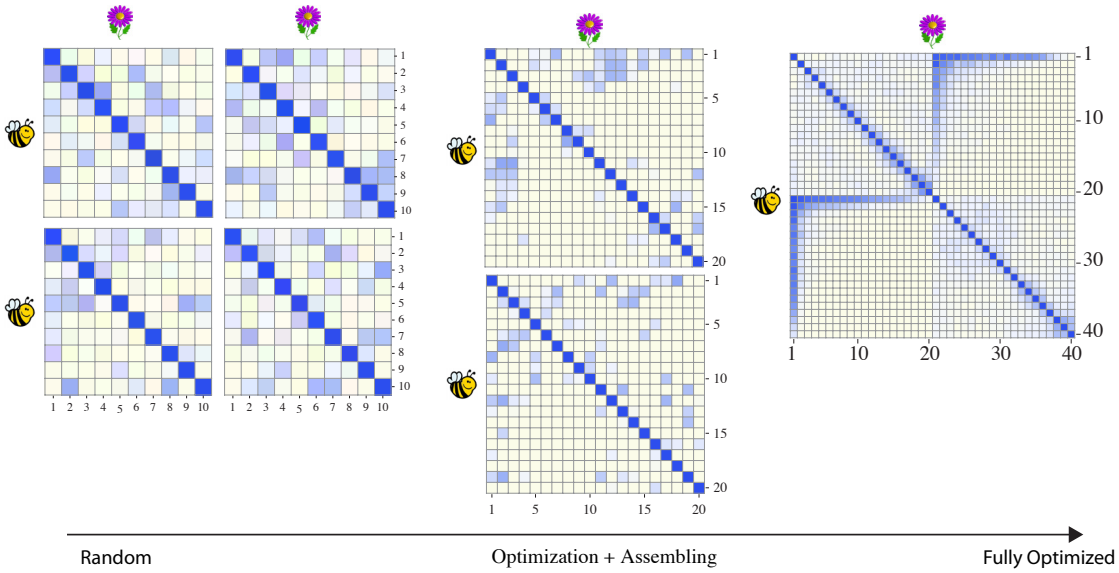

Figure S8: Maximization of the total community population, while assembling communities. Parameters used for this simulation are $C = 0.35$ and $\sigma_\Gamma = \sigma_\Omega = 0.05$. If the abundance of single populations is optimized, the same results are obtained.

4. We sort the species according to their k_i^A (k_i^P) values, in decreasing order: $(i', j', k', ..., l')$ with $k_{i'}^A \geq k_{j'}^A \geq k_{k'}^A \geq ... \geq k_{l'}^A$.

5. We relabel each species following the permutation $(i, j, k, ...) \rightarrow (i', j', k', ...)$, and we swap the rows and columns of the original interaction matrix accordingly.

In order to compare the average structure of many numerical realizations of random and optimized interaction matrices with the same S, C_Γ, n_A, n_P, σ_Γ, we applied the procedure described above to each of the n realizations, $\Gamma^{(k)}$, where $k = 1, ..., n$. Then we calculated the average interaction matrix $\bar{\Gamma}$ by averaging each entry over the n realizations: $\bar{\Gamma}_{ij} = \frac{1}{n} \sum_{k=1}^{n} M_{ij}^{(k)}$.

It is important to observe that this method is only one of many possible ways of relabeling species, and it may not yield the best visualization of the nestedness of the matrix. In order to have a fair comparison, independent of the relabeling procedure chosen, what matters is to be consistent and always apply the same procedure.

5 Analytical Results: Mean Field Approximation

In the following section, we present a mean field approximation that shows how: a) An increase in the population size of one individual species leads to a net increase of

the population abundance of the whole community; b) The communities with larger total population have an interaction matrix with higher nestedness and vice-versa. The approximation consists in considering the $2S \times 2S$ interaction matrix M with constant mutualistic/competitive interactions strength:

$$M = M_0 + V = \begin{bmatrix} \mathbb{1} + \Omega & \mathbb{0} \\ \mathbb{0} & \mathbb{1} + \Omega \end{bmatrix} + \begin{bmatrix} \mathbb{0} & \Gamma \\ \Gamma^T & \mathbb{0} \end{bmatrix}, \tag{S19}$$

where the T superscript denotes the transpose operation and with

$$\Omega_{ij} = C\omega(1 - \delta_{ij}) \qquad \omega > 0 \tag{S20}$$

$$\Gamma_{ij} = \begin{cases} -\gamma & \text{with probability } C \ (\gamma > 0) \\ 0 & \text{with probability } 1 - C \end{cases}$$

where $C = C_\Gamma = C_\Omega$ is the connectance and δ_{ij} the Kronecker $\delta(\delta_{ij} = 1$ if $i = j$, 0 otherwise). Therefore, in the mean-field approximation, we have set $C \cdot S^2$ non-zero elements of the same magnitude γ in the mutualistic interaction matrix, while all S^2 elements are set to $C\omega$ in the intra-species competition matrix.

5.1 Relation between species-level optimization and community-level optimization

Having in mind the population dynamics given by Eq. (S6), the species-level optimization algorithm and assuming that $\omega \sim \gamma \ll 1$, we now show using perturbative expansions that an increase in the population of one individual species leads to a net increase of the population abundance of the whole community. We note that the following results still hold if γ and ω are random variables drawn from a given distribution (and $\omega \sim \gamma \ll 1$).

Consider the interaction rewiring between species k and species m and n as given by Figure 9. Then the interaction network changes as $M \to \delta M$, with $\sum_l \delta M_{lk} = 0$ (the overall total interaction strength is kept fixed). In particular we have:

$$\delta M_{ki} = \begin{cases} M_{kn} - M_{km} & \text{if } i = m \\ M_{km} - M_{kn} & \text{if } i = n \\ 0 & \text{if } i \neq m, n. \end{cases} \tag{S21}$$

Similarly, we also have

$$\delta M_{ik} = \begin{cases} M_{nk} - M_{mk} & \text{if } i = m \\ M_{mk} - M_{nk} & \text{if } i = n \\ 0 & \text{if } i \neq m, n. \end{cases} \tag{S22}$$

These equations imply that $\delta M_{km} = -\delta M_{kn}$ and $\delta M_{mk} = -\delta M_{nk}$. From Eq. (S6) the stationary populations abundances are given by $\vec{x}^* = M^{-1}\vec{\alpha}$. After one optimization step we thus have

$$\vec{x}^* + \delta\vec{x}^* = (M + \delta M)^{-1}\vec{\alpha} = (1 + M^{-1}\delta M)^{-1}\vec{x}^*. \tag{S23}$$

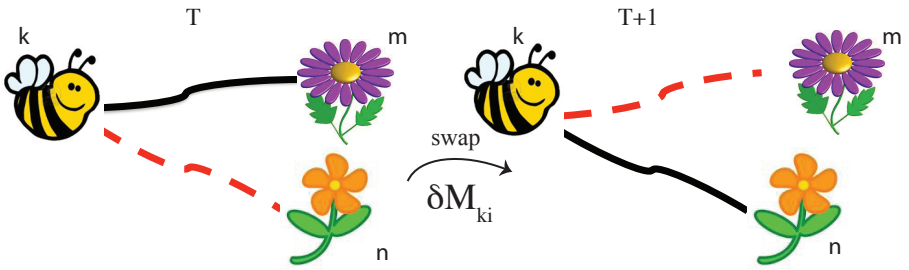

Figure S9: Example of swap during the optimization of the population abundance of the individual species, presented in section 4.2

We have assumed that $\omega \sim \gamma \ll 1$, i.e. $M = \mathbb{1} + \epsilon$ ($\| \epsilon \| \ll 1$) and thus $\delta M \sim \epsilon$. A perturbative expansion of Eq. (S23) to lowest order, yields the following change in the population abundances

$$\delta \vec{x}^* = -\delta M \cdot \vec{x}^*. \tag{S24}$$

From Eqs. (S21)-(S22), we have

$$\delta x_k^* = -\sum_i \delta M_{ki} x_i^* = \delta M_{km}(x_n^* - x_m^*) \tag{S25}$$

$$\delta x_n^* = -\sum_i \delta M_{ni} x_i^* = -\delta M_{nk} x_k^* \tag{S26}$$

$$\delta x_m^* = -\sum_i \delta M_{mi} x_i^* = -\delta M_{mk} x_k^* = -\delta x_n^* \tag{S27}$$

$$\delta x_l^* = 0 \ (l \neq n, m, k) \tag{S28}$$

and therefore

$$\delta x_{tot}^* = \sum_l \delta x^*{}_l = \delta x_k^*. \tag{S29}$$

We have thus demonstrated that if $\delta x_k^* > 0$ then $\delta x_{tot}^* > 0$, where δx_{tot}^* is the variation in the total number of individuals in the community. The key result is that the optimization of individual fitness leads to an increase of the community fitness (as measured by population abundances).

5.2 Relation between the total abundance of individuals in the community and nestedness

Given Eqs. (S19)-(S21), the inverse of the matrix M_0 can be calculated exactly [17]:

$$M_0^{-1} = \begin{bmatrix} A & \mathbb{0} \\ \mathbb{0} & A \end{bmatrix} \tag{S30}$$

with

$$A_{ij} = \frac{1}{(1-C\omega)[1+C\omega(S-1)]} \cdot \begin{cases} 1+C\omega(S-2) & \text{if } i=j \\ -C\omega & \text{if } i\neq j \end{cases} = \begin{cases} a \\ -b \end{cases} \tag{S31}$$

If we assume that the steady state of Eq. (S6) is $\vec{x}^* = (1, 1, ..., 1)$, then $\vec{\alpha}$ is fixed to $M\vec{x}^*$. Without loss of generality, we can select an initial matrix M such that the vector $\vec{\alpha}$ is constant, i.e. $\alpha_i = \alpha$, $\forall i$.

We define the following dynamics for the interaction matrix M: we select two elements from Γ, one interacting ($\gamma_{ij} = \gamma$) and one non-interacting ($\gamma_{kl} = 0$) pair of species, and we swap their values, i.e. $\gamma_{ij} = 0$ and $\gamma_{kl} = \gamma$. Using the new matrix M we recalculate the populations at the new steady state from the equation $\vec{x}^* = M^{-1}\vec{\alpha}$. The total population abundance, x^{tot}, is then proportional to the sum of all the entries of matrix M^{-1}: $x^{\text{tot}} = \alpha \sum_{i,j} M^{-1}_{ij}$. We accept the change if the total population relative to the new M is greater than the one calculated with the old M. This dynamics will lead to a matrix M that maximizes $\sum_{i,j} M^{-1}_{ij}$.

Let us therefore calculate the inverse of an interaction matrix of the form specified in Eq. (S19). By observing that $M = M_0 + V = (\mathbb{1} + VM_0^{-1})M_0$, then we have

$$M^{-1} = M_0^{-1}(\mathbb{1} + VM_0^{-1})^{-1} = M_0^{-1} - M_0^{-1}VM_0^{-1} + M_0^{-1}VM_0^{-1}VM_0^{-1} + ... \tag{S32}$$

The sum of all entries of M^{-1}, at the second order approximation ($o(\gamma^2)$), is equal to:

$$x^{\text{tot}} = \sum_{ij} M^{-1}_{ij} = \sum_{ij} (M_0^{-1})_{ij} - \sum_{ij} (M_0^{-1}VM_0^{-1})_{ij} + \sum_{ij} (M_0^{-1}VM_0^{-1}VM_0^{-1})_{ij} \tag{S33}$$

Inserting Eqs. (S19-S31) in (S33) we have

$$\sum_{ij} (M_0^{-1})_{ij} = 2\sum_{i,j=1}^{S} A_{ij} = 2S[a-b(S-1)]$$

$$\sum_{ij} (M_0^{-1}VM_0^{-1})_{ij} = \sum_{ij} \begin{bmatrix} \mathbb{0} & A\Gamma A \\ A\Gamma^T A & \mathbb{0} \end{bmatrix}_{ij} = 2\sum_{i,j=1}^{S} (A\Gamma A)_{ij} = 2[a-b(S-1)]^2 CS^2(-\gamma)$$

where we have used

$$\sum_{kl} \Gamma_{kl} = -\gamma CS^2 \tag{S34}$$

$$\sum_{ij} (M_0^{-1}VM_0^{-1}VM_0^{-1})_{ij} = \sum_{ij} \begin{bmatrix} A\Gamma A\Gamma^T A & \mathbb{0} \\ \mathbb{0} & A\Gamma^T A\Gamma A \end{bmatrix}_{ij}$$

$$= [a-b(S-1)]^2 \sum_{kl} \left[(\Gamma A\Gamma^T)_{kl} + (\Gamma^T A\Gamma)_{kl} \right] = \tag{S35}$$

$$= [a-b(S-1)]^2[(a+b)o\gamma^2 - 2b(\gamma CS^2)^2], \tag{S36}$$

where

$$o = \frac{1}{\gamma^2} \sum_{ij} \sum_{k} \left(\Gamma_{ki} \Gamma_{kj} + \Gamma_{ik} \Gamma_{jk} \right) \tag{S37}$$

is the overlap of the mutualistic interaction matrix.

Thus to second order in γ we have found

$$x^{tot} = K + \mathcal{C} o \;\Rightarrow\; o \propto \mathcal{C}^{-1} x^{tot} + constant, \tag{S38}$$

where

$$K = 2S[a - b(S-1)] + 2\gamma C S^2 [a - b(S-1)]^2 (1 - b\gamma C S^2) \tag{S39}$$

and

$$\mathcal{C} = (a+b)[a - b(S-1)]^2 \gamma^2 = \frac{\gamma^2}{(1 - C\omega)[(1 + C\omega)(S-1)]^2}. \tag{S40}$$

We note that the same conclusions hold for a random competitive interaction matrix Ω for which the off-diagonals elements are equal to ω (0) with probability C $(1 - C)$

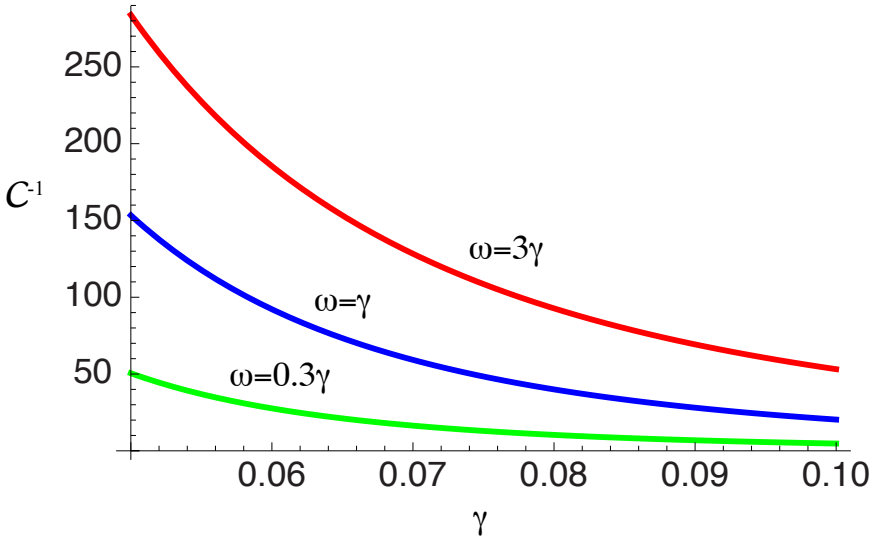

Figure S10: $\mathcal{C}^{-1}$ is the coefficient which relates community abundance $\sum_i x_i^*$ to the overlap of the mutualistic interaction network (o). For increasing population, we always obtain increased overlap, but the correlation decreases as mutualistic (competitive) interactions increases (decreases). Here $S = 50$ and $C = C_\Gamma = C_\Omega = 4/S^{0.8}$. The value of the mutualistic interaction strength is denoted by γ, whereas the competition strength is ω (mean-field case).

We have therefore found a direct relationship between the total population of a community and the overlap of the matrix representing the species' interactions:

Increasing the total population of a community is equivalent to increase the overlap of the species' interaction matrix.
Finally, observing that every definition of nestedness is an increasing function of the plant and animal pollinator overlaps, we conclude that:
Optimizing (increasing) the total abundance of individuals in the community leads to an interaction matrix with greater nestedness.

6 Numerical Results

6.1 Correlation between the total abundance of individuals in the community and nestedness

The correlation between community abundance and the architecture of mutualistic interactions revealed by our analytical calculations is also confirmed by extensive numerical simulations. Monte Carlo simulations presented in the main text (see Figure 2B) show the positive correlation between the total abundance of individuals in the community and random matrices characterized by different values of nestedness. In the numerical analysis we start with a random interaction M characterized by low NODF. We then swap interactions in order to increase the NODF. In this way we have a set of interaction matrices characterized by different values of NODF. For each one of them, given a fixed α, we calculate the corresponding community population abundance using Eq. (S6) or Eq. (S13). We repeat this procedure R times, so to achieve sufficient statistics. Another strong piece of evidence for the correlation between community population abundance and nestedness is given by the fact that if we minimize the nestedness of the mutualistic interaction network (we accept the swap only if NODF of the mutualistic network decreases), then a corresponding decrease in the total population is observed (see Figure S12).

These results demonstrate that population abundance of a mutualistic community and the structure of the interactions networks are correlated, independent of the particular optimization rule one chooses.

6.2 Architecture of Optimal Mutualistic Ecological Networks

As result of the maximization of the community abundance we obtain a new structure of the interaction matrix Γ, that we call the "optimal" interaction matrix Γ_{opt}. In this section, we present the resulting final architecture of Γ_{opt} for different regimes of interaction strengths presented in Table 1. The final optimal shape of the mutualistic interaction matrix Γ depends on the direct competition matrix Ω. If Ω is turned off ($\omega_{ij} = 0\ \forall i, j$) then, the maximization of the community abundance tends to generate rectangular blocks

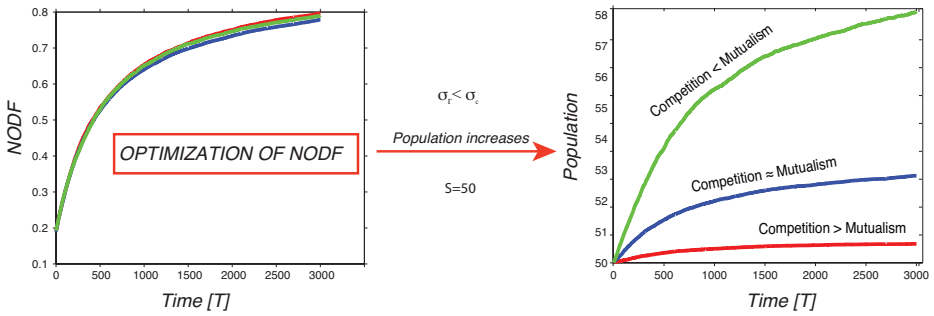

Figure S11: Maximization of nestedness of the mutualistic interaction network. We find that the corresponding total abundance increases.

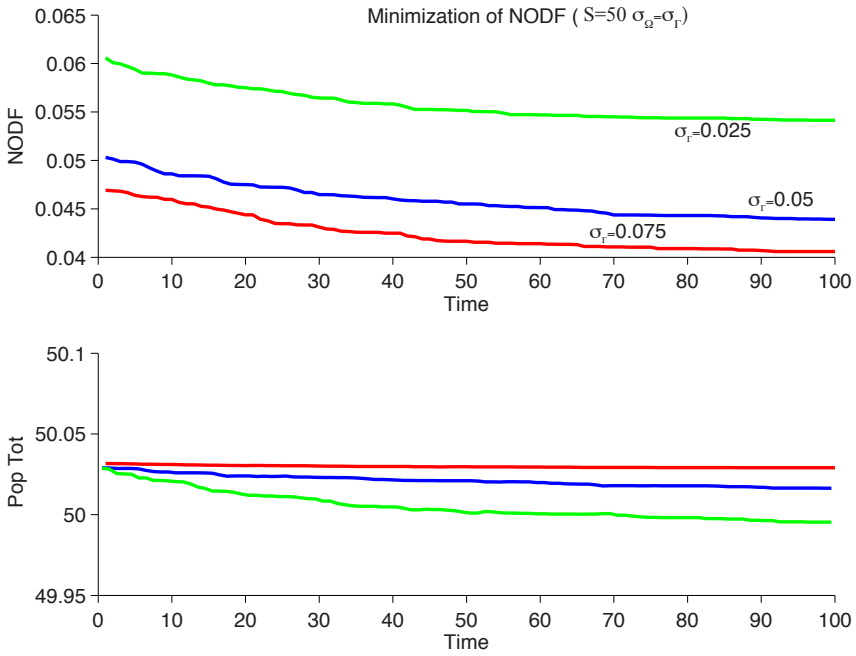

Figure S12: Minimization of nestedness of the mutualistic interaction network. We find that the corresponding total abundance decreases.

in the interaction matrix, i.e. there is a subgroup of plants and pollinators in which each insect species interacts with all plants and vice-versa (see Figure S14). All nestedness results are presented using the NODF measure, but as claimed in section 1 the same results hold if the alternative definition (η) is used, as shown in Figure S13

When direct competition is turned on ($\beta > 0$), the optimal mutualistic interaction

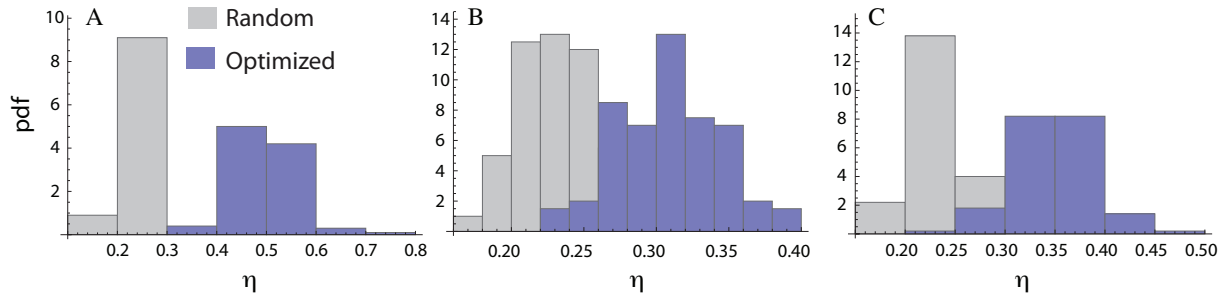

Figure S13: Comparison between the statistics of nestedness η - given by Eq. (S3) - for random (gray) and optimized (violet) mutualistic networks. We perform: (A) Maximization of community population using Holling Type I dynamics for the case of strong competition $\sigma_\Omega = 3\sigma_\Gamma$; (B) Optimization of the abundance of the species involved in the interaction rewiring for Holling Type I dynamics and weak competition ($\sigma_\Omega = \sigma_\Gamma/3$); (C). Optimization of the abundance of the species involved in the interaction rewiring for Holling Type I dynamics and $\sigma_\Omega = \sigma_\Gamma$. All plots are averages over 100 realizations of the optimization algorithm with $C_\Omega = C_\Gamma = 4S^{-0.8}$ and $\sigma_\Gamma < \sigma_c$.

matrix Γ_{opt} has an "upper diagonal" like architecture. The particular nested shape of the optimal mutualistic interaction matrix depends also on the strength of the mutualistic interaction. In fact, as noted in the main text and in SI sec. 5.2, the nestedness of the optimized system increases as σ_Γ decreases. Ecologically, this suggests that if mutualistic interactions strengths are weak, then it becomes crucial for the community to organize into a highly nested architecture in order to maximize its population and thus its persistence.

Figures S15-S19 show the predicted NODF from our optimization principle for the three typical cases of competitive and mutualistic interactions strengths. The simulations have been performed using the Holling Type I saturating functions and for the community-level optimization, but similar results are obtained with the Holling type II saturating function or species-level optimization.

6.3 Null Model Randomizations

Given that interactions are heterogeneously distributed across nodes, the question then arises whether they are distributed randomly or not. Note that it is not meaningful to compare absolute values of nestedness across networks of different sizes, connectance, or with different degree distributions[18, 19]. In order to address this issue, we will follow previous works [5, 9, 18] and define two null models to benchmark the degree of nestedness (see Figure S15).

In **Null Model 0** (also indicated as "random"), we keep fixed the size and the connectivity of the optimized networks, while the edges are placed randomly within the matrix

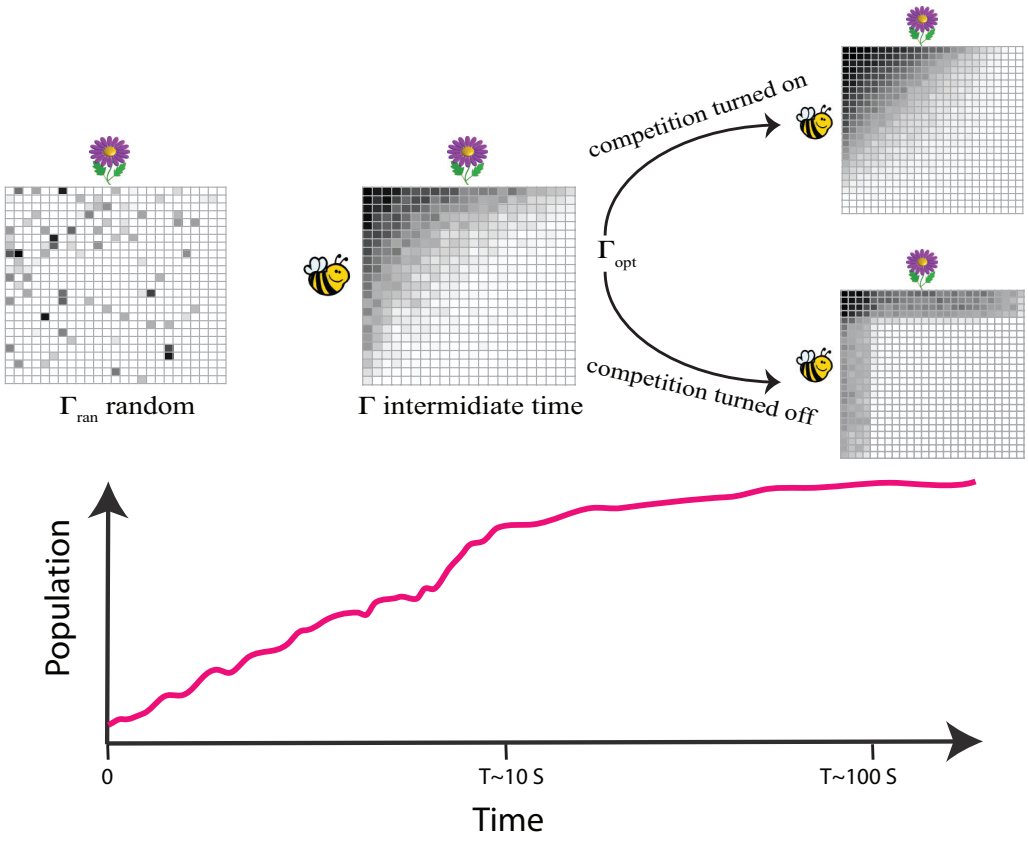

Figure S14: Architecture of a random mutualistic network Γ_{ran} and its evolution driven by the maximization of population abundance principle. Γ at intermediate time is independent of the presence or absence of direct competition Ω, while the final optimal network structure does depend on Ω. The second and the third interaction matrices from the left are averages performed over several runs of the iteration procedure and the scale of grey is a measure of the probability that a given entry is present. Time T is measured in terms of the number of iterations. All networks are generated for S=50, $C \sim S^{0.8}$ and $\sigma \ll \sigma_c$.

[5]. Moreover, since nestedness changes along with changes in the degree distribution [18, 19], we also perform a more stringent randomization (**Null Model 1**), where we keep fixed the average number of connections for each plant and insect, i.e., each element $A_{ij}^{(r)}$ of the bipartite randomized adjacency matrix is one with probability $k_i^P \cdot k_j^A/L$ and 0 otherwise (k_i^P, k_j^A are the degree of plant i and insect j, respectively and L is the total number of mutualistic links in the network) [5]. For each optimized network we compute 100 realizations of null model 1 and then calculate the corresponding average nestedness $\overline{NODF}_{nm1}$. In this way we can also define the relative nestedness $NODF^*$ [9], i.e., the

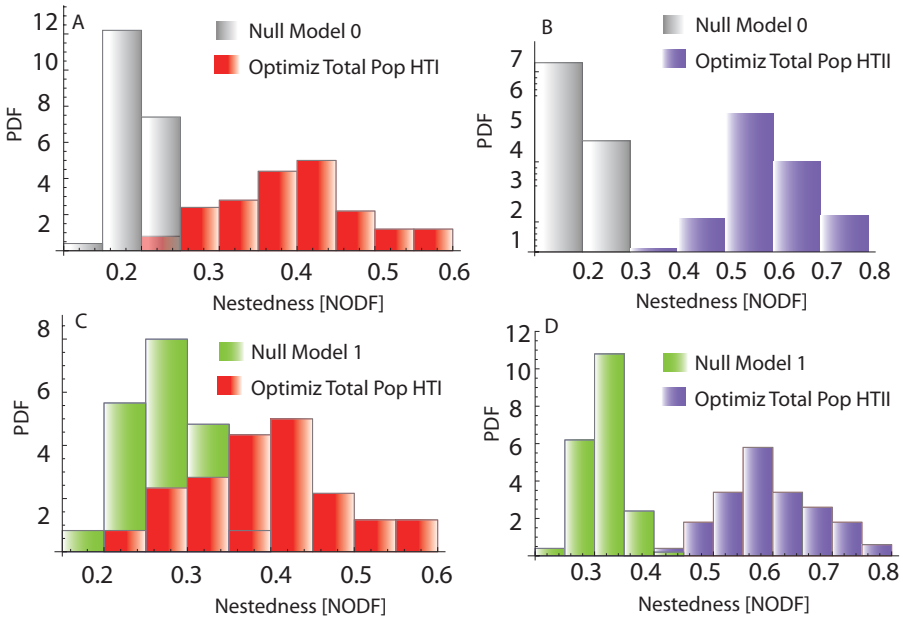

Figure S15: Histograms of the nestedness probability density (PDF) for randomizations under null model 0 (gray), under null model 1 (in green) and mutualistic networks, where community population has been maximized using Holling Type I (in red) and Type II (in violet) population dynamics. All plots are averages over 100 realizations of the abundance maximization algorithm for $\sigma_\Omega = \sigma_\Gamma$ and $C_\Omega = C_\Gamma = 4S^{-0.8}$.

nestedness normalized to null model 1 ($nm1$):

$$NODF^* = \frac{NODF - \overline{NODF}_{nm1}}{\overline{NODF}_{nm1}} \tag{S41}$$

As we can see from Figure S16 and Figure 3 in the main text, the relative nestedness increases along with the optimization of the interaction network.

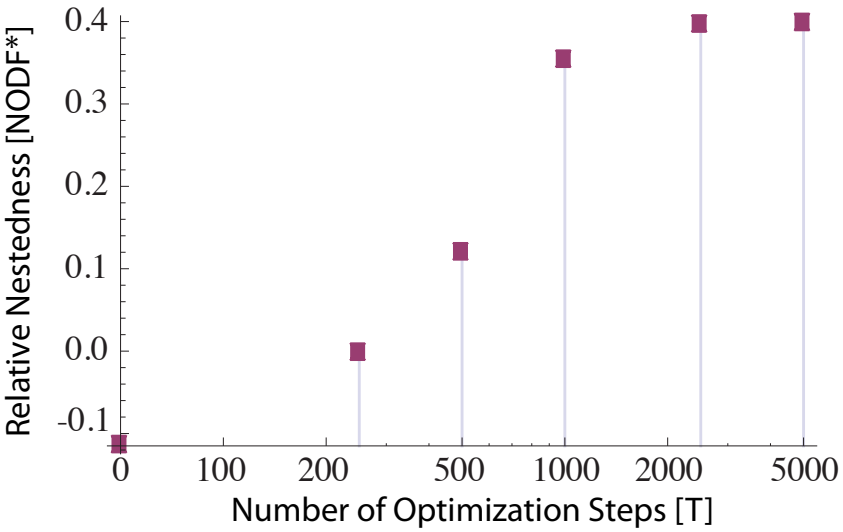

Figure S16: Relative nestedness as a function of the number of optimization steps T for the case of community-level optimization and using Holling Type II saturating function. We note that the results depend on the optimization dynamics. Parameters used in this simulation are S=50, $C_{Gamma} = 4S^{-0.8}$, $C_{\Omega} = 0$ and $\sigma_{\Gamma} = 0.15$.

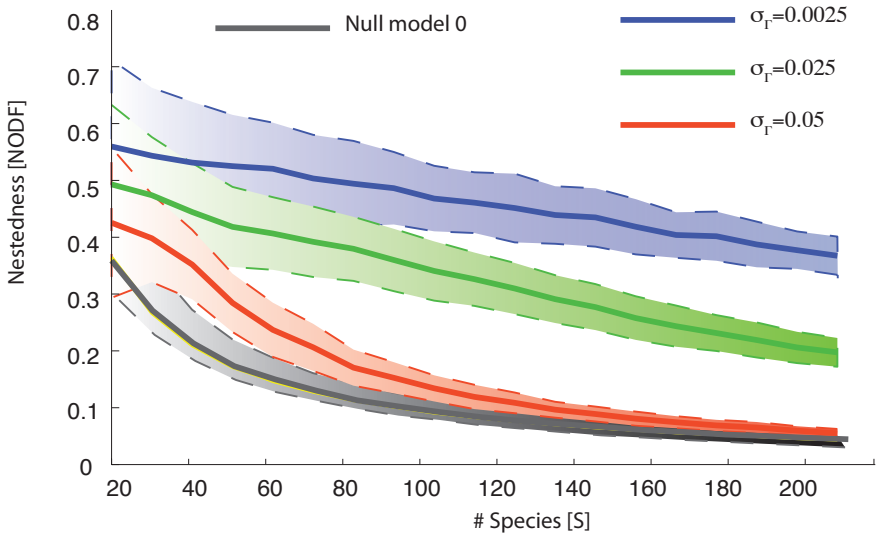

Figure S17: Nested architecture measured by NODF versus the number of species of the optimized mutualistic interaction network for three different values of interaction strengths for the case competition= mutualism ($\beta = 1$). Community population has been maximized using HTI saturating function. The dashed lines give the 1 standard deviation confidence interval.

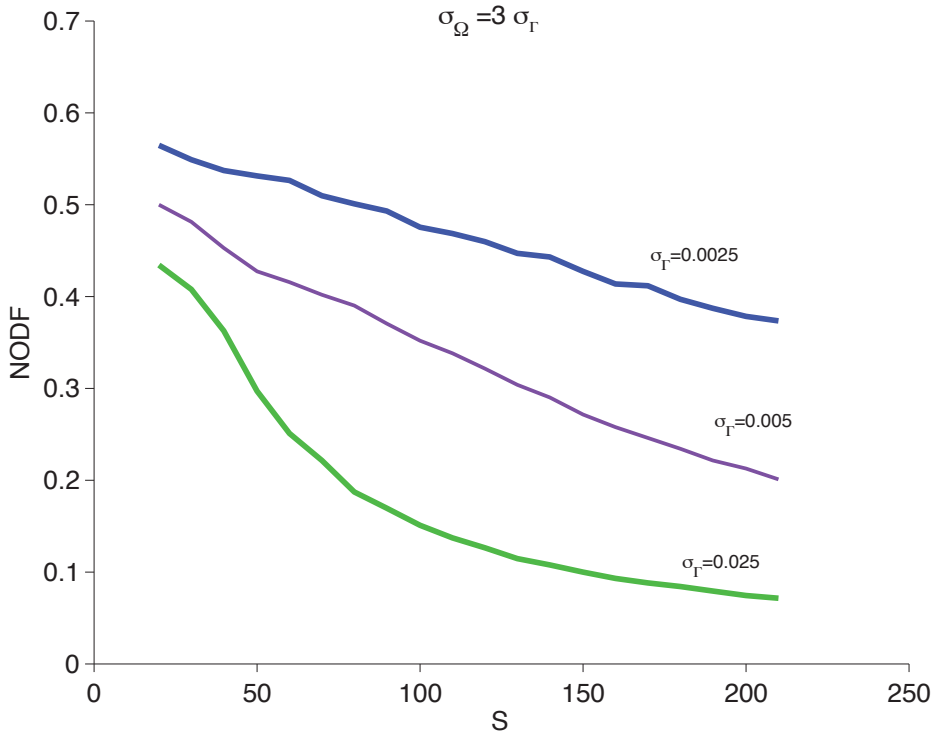

Figure S18: Nested architecture measured by NODF versus the number of species of the optimized mutualistic interaction network for three different values of interaction strengths for the case competition> mutualism ($\beta = 3$). Community population has been maximized using HTI saturating function. The standard deviation confidence interval behaves qualitatlvly as in Figure S17.

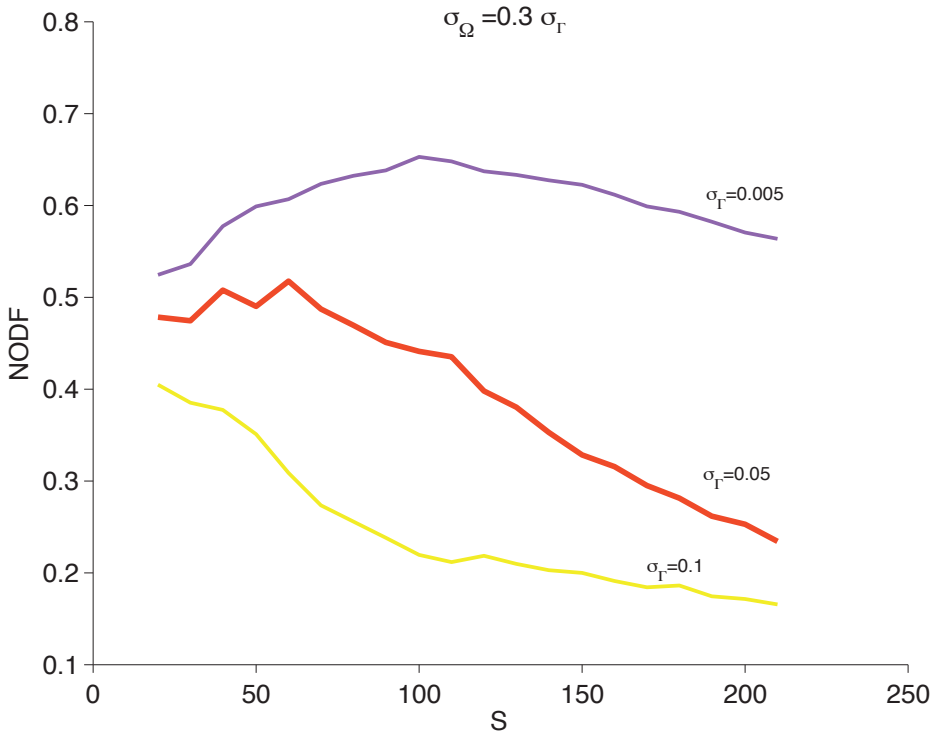

Figure S19: Nested architecture measured by NODF versus the number of species of the optimized mutualistic interaction network for three different values of interaction strengths for the case competition< mutualism ($\beta = 0.3$). Community population has been maximized using HTI saturating function.The standard deviation confidence interval behaves qualitatlvly as in Figure S17.

7 Stability Analysis of the Optimized Networks.

Community persistence and stability are known as important dynamical properties characterizing ecological networks, but the way in which the two are related in real systems is far from trivial. Community persistence has been considered as a possible driving force sculpting network architectures, i.e. the observed life web structures are the ones leading to high community persistence. Persistence can be defined as the fraction of initial species remaining once the dynamics has driven the system to a stable state. Consequently it is relevant only for unstable systems, and moreover, it is sensitive to initial conditions and transients, and to the distance of the system from stationarity. It is hard to disentangle the role of distinct ecological variables on the networks topology in a dynamically evolving system, since they all typically vary simultaneously. For instance, the reduction of the effective ecosystem diversity due to species extinction that occurs while reaching stationarity, necessarily increases the nestedness of the corresponding interaction matrix because, on average, even randomly assigned matrices of smaller size are more nested [20]. Therefore here we focus on the study of community stability for the optimized stationary networks.

7.1 Analytical Results: relation between Rarest Species and Community Resilience

Using perturbative expansion techniques, we find that rare species control community stability. In particular, we find that the dominant real part of the eigenvalues of Φ is related to the population of the rarest species through

$$Max[Re(\lambda)] \cong -Min[\vec{x}], \tag{S42}$$

when $\sigma_\Gamma \ll 1$.

7.1.1 Holling Type I

From Eq. (S7), we know that the stationary populations of the species $x_i^* > 0$ are related to the linearized matrix Φ of the Holling Type I dynamics $\dot{x}_i \equiv f_i(\vec{x})$ through $\phi_{ij} = -x_i^* M_{ij}$. The interaction matrix M is built so that

$$M_{ij} = \begin{cases} d & \text{if } i = j \\ m_X \sim C_X \mathcal{N}(0, \sigma_X) + (1 - C_X)\delta(m_X) & \text{if } i \neq j, \end{cases}$$

where the index $X = \Gamma, \Omega$ depending on whether we are considering the mutualistic or competitive block and δ is the Dirac $\delta-$function. Thus Eq. (S7) can be rewritten as

$$\phi_{ij} = -x_i^* M_{ij} = -x_i^* \delta_{ij} - (1 - \delta_{ij}) M_{ij} x_i^* = \phi_{ij}^0 + V_{ij}, \tag{S43}$$

where $\phi^0_{ij} = -x^*_i\delta_{ij}$ and $V_{ij} = -(1-\delta_{ij})M_{ij}x^*_i$. The eigenvalues of ϕ^0 are $-x^*_i$, while its eigenvectors are $u^{(i)}(0,0,...,0,1,0,...0)$ where the 1 is the $i-$th component of $u^{(i)}$. We can thus calculate the eigenvalues of Φ using perturbation theory. If λ_i is the $i-$th eigenvalue of Φ, then we have ($\sigma = max_x\sigma_x$)

$$\begin{aligned}\lambda_i &= -x^*_i + u^{(i)}\cdot Vu^{(i)} + \sum_{j\neq i}\frac{(u^{(i)}\cdot Vu^{(j)})(u^{(j)}\cdot Vu^{(i)})}{x^*_i - x^*_j} + o(\sigma^3) \qquad \text{(S44)}\\ &= -x^*_i + 0 + \sum_{j\neq i}\overbrace{M_{ij}M_{ji}}^{o(\sigma^2)}\frac{x^*_i x^*_j}{x^*_i - x^*_j} + o(\sigma^3) = -x^*_i + o(\sigma^2)\end{aligned}$$

from which Eq. (S42) is achieved. We have numerically tested the prediction given by Eq. (S44) and the result is shown in Fig. S20.

7.1.2 Holling Type II

This result is also valid for a community dynamics with Holling Type II saturating function. In fact, from Eq. (S18) we have that

$$\phi_{ij} \equiv -x^*_i\big[\delta_{ij}d + \frac{M_{ij} + (d\,x^*_i - \alpha_i)\theta_{ij}}{h + \sum_j \theta_{ij}x^*_j}\big] = \phi^0_{ij} + V_{ij}, \qquad \text{(S45)}$$

where this time $\phi^0_{ij} = -dx^*_i\delta_{ij}$ and $V_{ij} = -x^*_i\frac{M_{ij}+(d\,x^*_i-\alpha_i)\theta_{ij}}{h+\sum_j\theta_{ij}x^*_j}$.

The eigenvalues of ϕ^0 are $-dx^*_i$, while its eigenvectors are $u^{(i)}(0,0,...,0,1,0,...0)$ where the 1 is the $i-$th component of $u^{(i)}$. Therefore also in this case we have that the first term in the perturbative expansion goes to zero, i.e. $u^{(i)}\cdot Vu^{(i)} = 0$ (remember that $M_{ii} = 0$ and $\theta_{ii} = 0\ \forall i$). Consequently, given λ_i as the $i-$th eigenvalue of Φ, we have ($\sigma = max_x\sigma_x$)

$$\lambda_i = -dx^*_i + o(\sigma^2) \qquad \text{(S46)}$$

7.2 Numerical Results: Spectrum of the Optimized Networks

We performed an analysis of the eigenvalues λ of the matrix Φ associated with the optimal interaction matrix M_{opt} corresponding to the maximal population x_{opt} (see Eq. (S7)) in order to study the stability of the stationary states of the optimized networks.

Figures S21-S23 show the value of the maximum real part of the eigenvalues of Φ after the optimization procedure for the three standard cases when Eq. (S6) is used for the population dynamics. We find that, after the optimization, the system remains stable with high probability if the starting random interaction matrix is stable ($\sigma_\Gamma < \sigma_c$), while if the initial system is critical, i.e. at the edge of stability ($\sigma_\Gamma \approx \sigma_c$) or unstable ($\sigma > \sigma_c$), then also the optimized community is critical or unstable, respectively. At the same time we observe that there is a shift of the optimized system towards the edge of stability, i.e. Max($Re(\lambda)$) gets closer to zero.

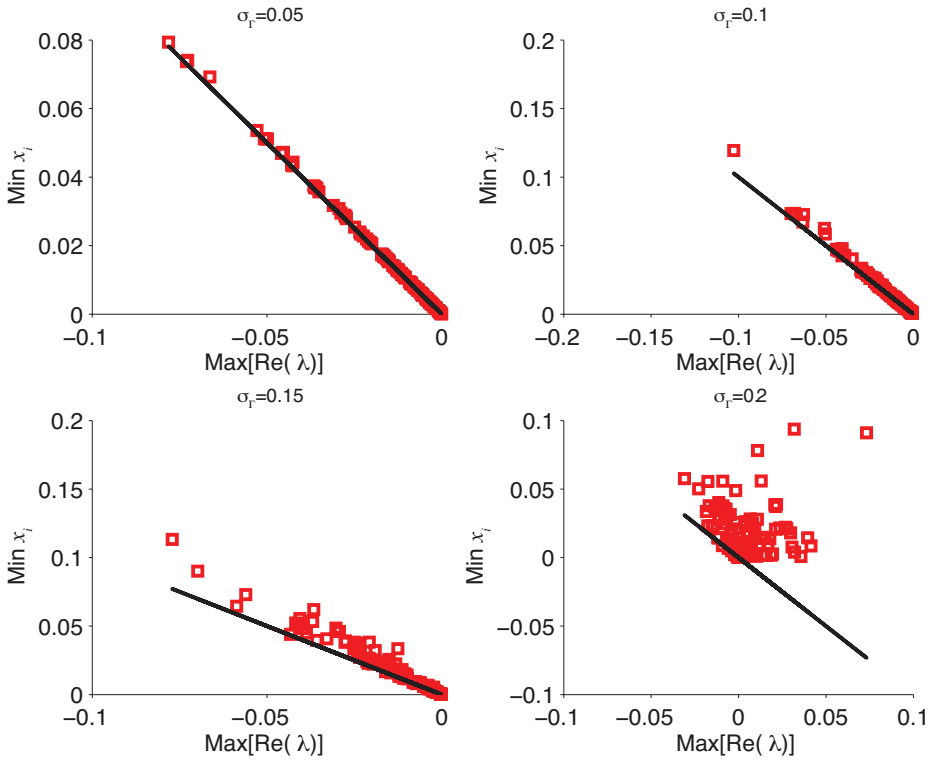

Figure S20: Numerical analysis of the relationship between the community resilience (given by the maximum real part eigenvalue of the interaction matrix Φ) and the population of the rarest species for different values of σ_Γ and for the case competition= mutualism ($\beta = 1$). Red squares represent the real part of the dominant eigenvalue of Φ for different realizations of the optimization starting from a random interaction matrix. The linear correlation (black line) predicted by our analytical calculation based on perturbative techniques holds within the stability region $\sigma_\Gamma < \sigma_c < 0.2$.

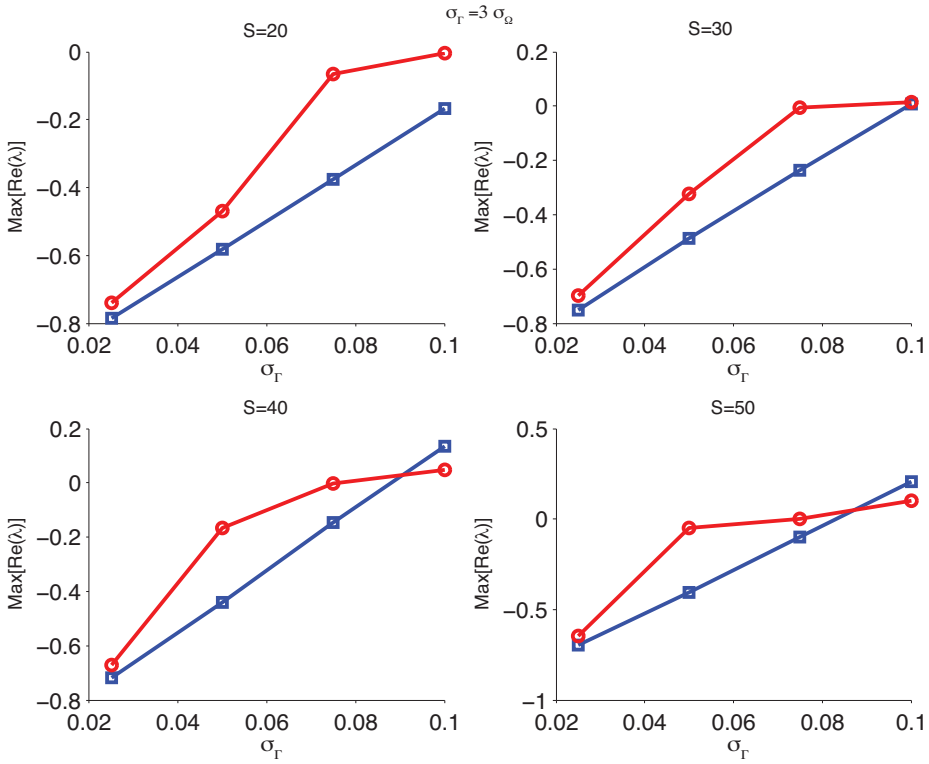

Figure S21: Maximum real part of the dominant eigenvalue of Φ before (blue squares) and after (red circles) the optimization procedure for the case competition> mutualism ($\beta = 3$).

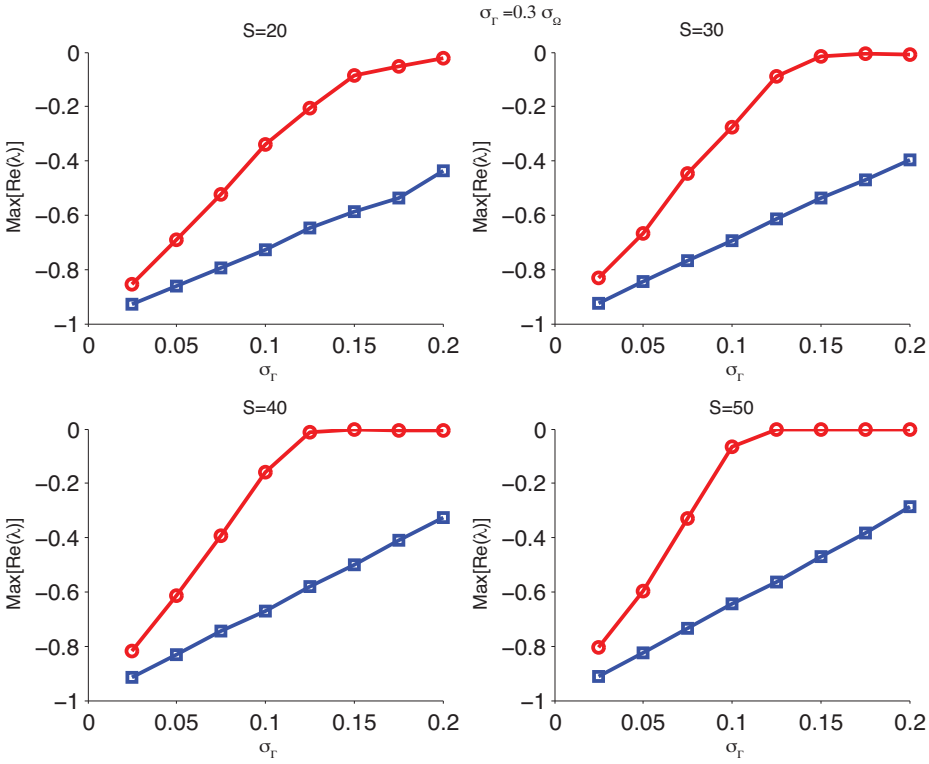

Figure S22: Maximum real part of the dominant eigenvalue of Φ before (blue squares) and after (red circles) the optimization procedure for the case competition< mutualism ($\beta = 0.3$).

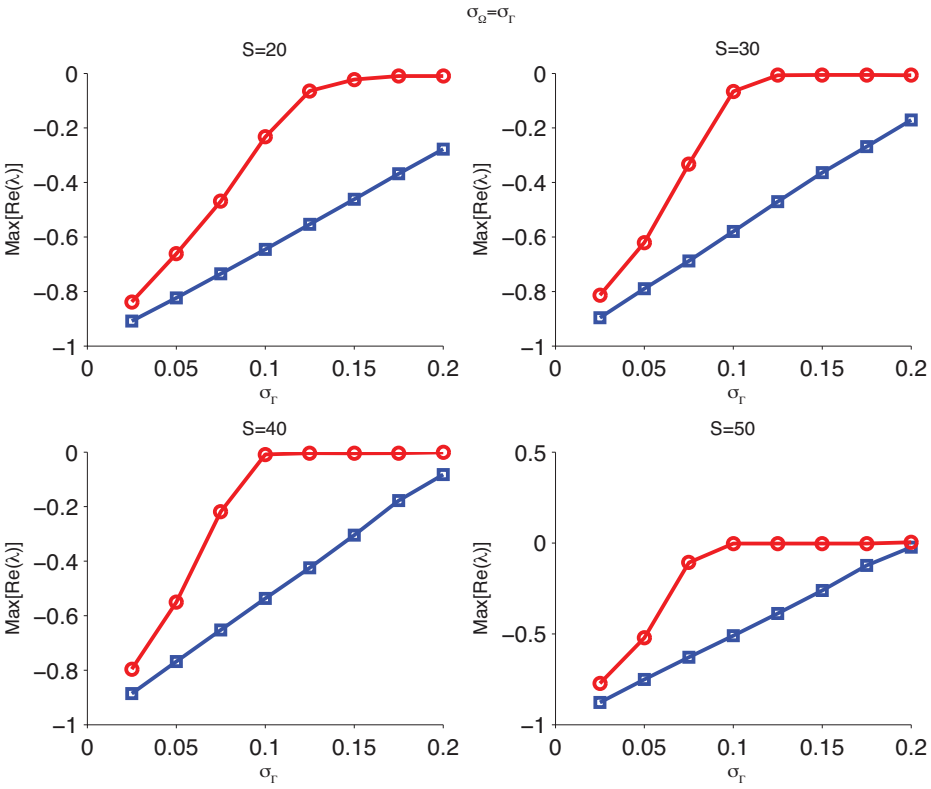

Figure S23: Maximum real part of the dominant eigenvalue of Φ before (blue squares) and after (red circles) the optimization procedure for the case competition$\approx$ mutualism ($\beta = 1$).

Figure S24 shows how also for Holling Type II dynamics, the optimized mutualistic community is less stable than the random starting one. In fact, another way to check the stability of the optimized community matrix is to study the population of the rarest species in the mutualistic community, as shown in previous section.

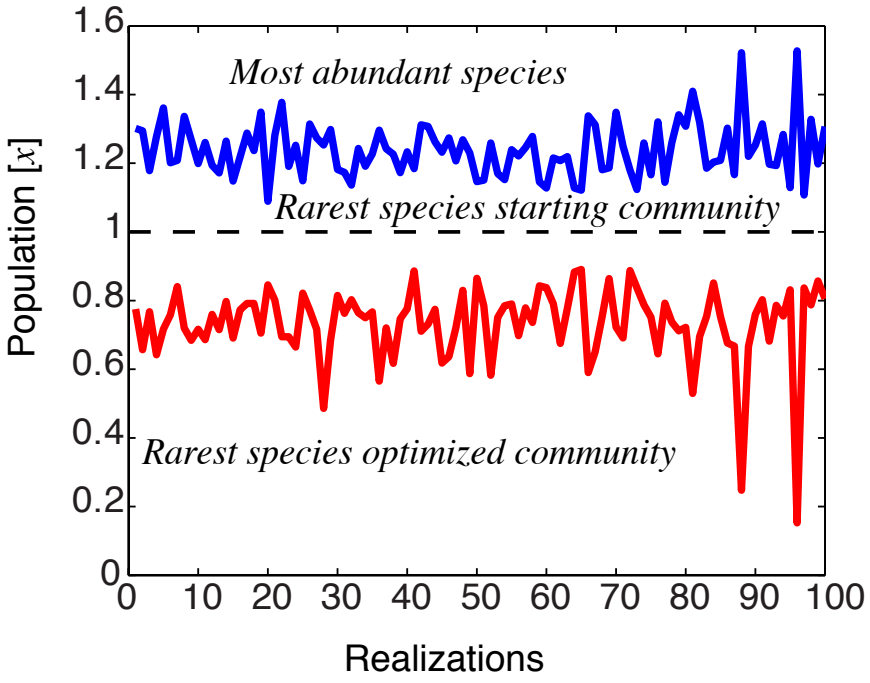

Figure S24: Population size of the rarest species before (dashed gray line) and after (red solid line) the single species optimization using Holling Type II saturating function for the population dynamics. The x axis represent different realizations of the optimization starting from the same initial condition with $\vec{x}_0^* = (1, 1, 1..., 1)$. The blue solid line likewise shows the population of the most abundant species in the optimized community in different realizations. This spontaneous symmetry breaking between species is a hallmark of nested architecture. Parameters of the simulations are $S = 20$, $C = 4S^{-0.8}$, $\sigma_\Gamma = \sigma_\Omega = 0.1$ and $h = 0.1$.

References

[1] M. Almeida-Neto, Paulo Guimara, Paulo R Guimara, and Werner Ulrich. A consistent metric for nestedness analysis in ecological systems : reconciling concept and measurement. *Oikos*, 117(March):1227–1239, 2008.

[2] Ugo Bastolla, Miguel a Fortuna, Alberto Pascual-García, Antonio Ferrera, Bartolo Luque, and Jordi Bascompte. The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature*, 458(7241):1018–20, April 2009.

[3] Serguei Saavedra, Daniel B Stouffer, Brian Uzzi, and Jordi Bascompte. Strong contributors to network persistence are the most vulnerable to extinction. *Nature*, 478(7368):233–235, 2011.

[4] Enrico L Rezende, Jessica E Lavabre, Paulo R Guimara, Pedro Jordano, and Jordi Bascompte. Non-random coextinctions in phylogenetically structured mutualistic networks. *Nature*, 448(August), 2007.

[5] Jordi Bascompte, Pedro Jordano, Carlos J Melián, and Jens M Olesen. The nested assembly of plant-animal mutualistic networks. *Proceedings of the National Academy of Sciences of the United States of America*, 100(16):9383–7, August 2003.

[6] Robert M May. Nature. *Nature*, 238:413, 1972.

[7] Stefano Allesina and Si Tang. Stability criteria for complex ecosystems. *Nature*, 483(7388):205–8, March 2012.

[8] H. Sommers, A. Crisanti, H. Sompolinsky, and Y. Stein. Spectrum of Large Random Asymmetric Matrices. *Physical Review Letters*, 60(19):1895–1898, May 1988.

[9] Elisa Thébault and Colin Fontaine. Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science (New York, N.Y.)*, 329(5993):853–6, August 2010.

[10] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by Simulated Annealing. *Science*, 220:671–680, 1983.

[11] Alan Grafen. Detecting kin selection at work using inclusive fitness. *Proceedings of the Royal Society B: Biological Sciences*, 274(1610):713–719, 2007.

[12] David Sloan Wilson and Edward O. Wilson. Evolution "for the good of the group". *American Scientist*, 96(5):380–389, 2008.

[13] Martin A Nowak, Corina E Tarnita, and Edward O Wilson. The evolution of eusociality. *Nature*, 466(7310):1057–1062, 2010.

[14] Paulo R Guimarães, Victor Rico-Gray, Paulo S Oliveira, Thiago J Izzo, Sérgio F dos Reis, and John N Thompson. Interaction intimacy affects structure and coevolutionary dynamics in mutualistic networks. *Current Biology*, 17(20):1797–1803, 2007.

[15] Nicolas Loeuille. Influence of evolution on the stability of ecological communities. *Ecology letters*, 13(12):1536–1545, 2010.

[16] Paulo R Guimarães Jr, Pedro Jordano, and John N Thompson. Evolution and coevolution in mutualistic networks. *Ecology letters*, 14(9):877–885, 2011.

[17] D. Brown and V. I. Smirnov. *Course in Higher Mathematics Vol. 3, Pt. 2 : Complex Variables Special Functions*, volume 60. The Harvard Classics. New York, 1964.

[18] Miguel A Fortuna, Daniel B Stouffer, Jens M Olesen, Pedro Jordano, David Mouillot, Boris R Krasnov, Robert Poulin, and Jordi Bascompte. Nestedness versus modularity in ecological networks: two sides of the same coin? *Journal of Animal Ecology*, 79(4):811–817, 2010.

[19] Nicholas J Gotelli and Werner Ulrich. Statistical challenges in null model analysis. *Oikos*, 121(2):171–180, 2012.

[20] Alex James, Jonathan W Pitchford, and Michael J Plank. Disentangling nestedness from models of ecological complexity. *Nature*, 000(1):10–13, June 2012.