Online Supplement:

Nestedness Promotes Stability in Maximum-Entropy Bipartite Food Webs

S1 Derivation of the MaxEnt Energy Flow Solution for General Topology

The Lagrangian is given by

$$\mathcal{L}\left(F^{*},\boldsymbol{\kappa},\boldsymbol{\lambda}\right)=-\sum_{i,j}f_{ij}\ln f_{ij}-\sum_{i}\kappa_{i}\left(\sum_{j}f_{ij}-d_{i}\right)-\sum_{j}\lambda_{j}\left(\sum_{i}f_{ij}-e_{j}\right),\tag{S1}$$

where we didn’t include the normalization condition eq. (1) as a constraint as it is already implied by the constraints eq. (3a) and eq. (3b).

We now need to solve for

$$\nabla\mathcal{L}=0.\tag{S2}$$

Note that $\frac{\partial\mathcal{L}}{\partial\kappa_i}=0$ and $\frac{\partial\mathcal{L}}{\partial\lambda_j}=0$ recover the constraint equations eq. (3a) and eq. (3b). On the other hand, $\frac{\partial\mathcal{L}}{\partial f_{ij}}=0$ gives

$$\begin{aligned}-\ln f_{ij}-1-\kappa_i-\lambda_j=0\\ f_{ij}=e^{-\kappa_i-\lambda_j-1}\\ f_{ij}=a_ib_j\end{aligned}\tag{S3}$$

where $a_i\equiv e^{-\kappa_i-\frac{1}{2}}>0$ and $b_j\equiv e^{-\lambda_j-\frac{1}{2}}>0$.

S2 Proof That the Linkage Existence Boundary Lies Below the Antidiagonal

Suppose there's a point on the linkage existence boundary that is above the antidiagonal. Let that point be the intersection of horizontal line l_h and vertical line l_v. Then the total flow below l_h (equal to its distance to the bottom side of the square) is larger than the total flow to the left of l_v (equal to its distance to the left side of the square), which is impossible since the former is a strict subset of the latter.

S3 Derivation of the MaxEnt Energy Flow Solution for a Perfectly Nested Topology

Let's verify that the MaxEnt energy flow formula given in the main text is correct.

Since x_{min} only depends on i and y_{min} only depends on j, we can factorize $\hat{f}_{ij} = a_i b_j$ where

$$a_i \propto d_i \exp\left(\int_{x_{\text{min}}}^{1} \frac{\mathrm{d}x}{h(x)}\right) \quad \text{and} \quad b_j \propto e_j \exp\left(\int_{y_{\text{min}}}^{1} \frac{\mathrm{d}y}{w(y)}\right).$$

It now suffices to verify

$$\sum_{i:(i,j)\in E} \hat{f}_{ij} = e_j \quad \text{and} \quad \sum_{j:(i,j)\in E} \hat{f}_{ij} = d_i.$$

Due to the symmetry between the indices i and j, we only need to show the former, and the proof for the latter would be nearly identical.

Substituting the expression for $\hat{f}_{ij}$ gives the following statement that we have to verify:

$$\sum_{i:(i,j)\in E} d_i \exp\left(\int_{y_{\min}}^{1} \frac{\mathrm{d}y}{w(y)} + \int_{x_{\min}}^{1} \frac{\mathrm{d}x}{h(x)}\right) = 1/C \tag{S4}$$

is constant in i.

Let's define $x_m(y)$ to be the maximum x such that (x, y) is in the white region (including the boundary) of the square, and $y_m(x)$ is similarly defined as the maximum y such that (x, y) is in the white region (including the boundary) of the square. Furthermore, for any x in column j (excluding its left and right edges), $h(x) = y_m(x) + x - 1$; and for any y in row i (excluding its top and bottom edges), $w(y) = x_m(y) + y - 1$.

For any x in column j (excluding its left and right edges), the left side of eq. (S4) can be rewritten as

$$\int_0^{y_m(x)} \exp\left(\int_{y_m(x)}^{1} \frac{\mathrm{d}y'}{x_m(y') + y' - 1} + \int_{x_m(y)}^{1} \frac{\mathrm{d}x'}{y_m(x') + x' - 1}\right) \mathrm{d}y$$

or

$$\exp\left(\int_{y_m(x)}^{1} \frac{\mathrm{d}y}{x_m(y) + y - 1}\right) \int_0^{y_m(x)} \exp\left(\int_{x_m(y)}^{1} \frac{\mathrm{d}x'}{y_m(x') + x' - 1}\right) \mathrm{d}y, \tag{S5}$$

which we wish to show to be independent of x over all of $[0, 1]$ except for those points at which $y_m(x)$ is discontinuous.

To avoid issues that may arise due to these discontinuities, let's write the linkage existence boundary as a parameterized curve $(x(t), y(t))$ for $0 \leq t \leq 1$, which goes from $(1, 0)$ in the top-right corner at $t = 0$ to $(0, 1)$ in the bottom-left corner at $t = 1$. The parameterization is such that the curve is continuously differentiable everywhere except at the corners, where it is only continuous.

Consider applying integration by parts to the following integral:

$$\begin{aligned}
&\int_0^{t_0} (\dot{x}(t) + \dot{y}(t)) \exp\left(\int_t^0 \frac{\dot{x}(t')}{x(t') + y(t') - 1}\,\mathrm{d}t'\right) \mathrm{d}t \\
&= (x(t_0) + y(t_0) - 1) \exp\left(\int_{t_0}^0 \frac{\dot{x}(t')}{x(t') + y(t') - 1}\,\mathrm{d}t'\right) \\
&\quad - \int_0^{t_0} (x(t) + y(t) - 1) \exp\left(\int_t^0 \frac{\dot{x}(t')}{x(t') + y(t') - 1}\,\mathrm{d}t'\right) \frac{-\dot{x}(t)}{x(t) + y(t) - 1}\,\mathrm{d}t \\
&= (x(t_0) + y(t_0) - 1) \exp\left(\int_{t_0}^0 \frac{\dot{x}(t)}{x(t) + y(t) - 1}\,\mathrm{d}t\right) + \int_0^{t_0} \dot{x}(t) \exp\left(\int_t^0 \frac{\dot{x}(t')}{x(t') + y(t') - 1}\,\mathrm{d}t'\right) \mathrm{d}t.
\end{aligned}$$

Moving the integral on the right-hand side to the left side cancels out the $\dot{x}(t)$ in the integrand on the left-hand side, giving

$$\int_0^{t_0} \dot{y}(t) \exp\left(\int_t^0 \frac{\dot{x}(t')}{x(t') + y(t') - 1}\,\mathrm{d}t\right) \mathrm{d}t = (x(t_0) + y(t_0) - 1) \exp\left(\int_{t_0}^0 \frac{\dot{x}(t)}{x(t) + y(t) - 1}\,\mathrm{d}t\right).$$

Therefore, we obtain

$$\begin{aligned}
S(t_0) &:= \exp\left(\int_{t_0}^1 \frac{\dot{y}(t)}{x(t) + y(t) - 1}\,\mathrm{d}t\right) \int_0^{t_0} \dot{y}(t) \exp\left(\int_t^0 \frac{\dot{x}(t')}{x(t') + y(t') - 1}\,\mathrm{d}t'\right) \mathrm{d}t \\
&= (x(t_0) + y(t_0) - 1) \exp\left(\int_{t_0}^0 \frac{\dot{x}(t)}{x(t) + y(t) - 1}\,\mathrm{d}t + \int_{t_0}^1 \frac{\dot{y}(t)}{x(t) + y(t) - 1}\,\mathrm{d}t\right). \qquad \text{(S6)}
\end{aligned}$$

Note that $S(t_0)$ becomes eq. (S5) when we set t_0 to the number such that $x(t_0) = x$.

It now suffices to show that $S(t_0)/S(t_1) = 1$ for any $0 < t_0, t_1 < 1$. Indeed,

$$\begin{aligned}
\frac{S(t_0)}{S(t_1)} &= \frac{x(t_0)+y(t_0)-1}{x(t_1)+y(t_1)-1} \exp\left(\int_{t_0}^{t_1} \frac{\dot{x}(t)}{x(t)+y(t)-1}\,\mathrm{d}t + \int_{t_0}^{t_1} \frac{\dot{y}(t)}{x(t)+y(t)-1}\,\mathrm{d}t\right) \\
&= \frac{x(t_0)+y(t_0)-1}{x(t_1)+y(t_1)-1} \exp\left(\int_{t_0}^{t_1} \frac{\dot{x}(t)+\dot{y}(t)}{x(t)+y(t)-1}\,\mathrm{d}t\right) \\
&= \frac{x(t_0)+y(t_0)-1}{x(t_1)+y(t_1)-1} \exp\left(\ln(x(t)+y(t)-1)\Big|_{t_0}^{t_1}\right) \\
&= 1,
\end{aligned}$$

hence completing the proof.

S4 Details of the Stability Argument

S4.1 The spectrum and perturbation theory of $\tilde{M}$

S4.1.1 Spectrum of $\tilde{M}$

Let's write $\tilde{M} = \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix}$ where $A = -\alpha F$ and $B = (\eta \odot F)^\top$.

Let's first look at the relationship between the eigenvalues/eigenvectors of AB and those of BA. For any $\mu \in \mathbb{C}$, let T_μ be the μ-eigenspace of AB, where $T_\mu = \{0\}$ if μ is not an eigenvalue of AB. Similarly, let S_μ be the μ-eigenspace of BA. We notice the following:

- If $\mu \neq 0$, then $ABT_\mu = T_\mu$ and $BAS_\mu = S_\mu$. Thus, T_μ and S_μ are related by $T_\mu = AS_\mu$ and $S_\mu = BT_\mu$, and $\dim(T_\mu) = \dim(S_\mu)$. As a result, μ is an eigenvalue of AB iff it is an eigenvalue of BA.

- If $\mu = 0$, then T_0 is just the kernel of AB and S_0 is the kernel of BA.

Now, $\tilde{M}^2 = \begin{bmatrix} AB & 0 \\ 0 & BA \end{bmatrix}$, whose μ-eigenspace is just $T_\mu \oplus S_\mu$.

- For a nonzero eigenvalue μ of $\tilde{M}^2$, the possible corresponding eigenvalues of $\tilde{M}$ are $\lambda = \pm\sqrt{\mu}$ and the corresponding eigenvectors take the form $v \oplus u$ where $v \in T_\mu$ and $u \in S_\mu$. Now, writing $\tilde{M}(v \oplus u) = \lambda(v \oplus u)$ gives $Au = \lambda v$ and $Bv = \lambda u$, so the eigenspace of $\tilde{M}$ corresponding to an eigenvalue $\lambda = \pm\sqrt{\mu} \neq 0$ is $\{(\lambda v) \oplus (Bv) : v \in T_\mu\}$.

- For an eigenvalue $\mu = 0$ of $\tilde{M}^2$ (if one exists), the corresponding eigenvalue of $\tilde{M}$ is $\lambda = \pm\sqrt{\mu} = 0$ and the corresponding eigenvectors take the form $v \oplus u$ where v is in the kernel of B and u is in the kernel of A.

S4.1.2 Perturbation theory of $\tilde{M}$

(Note: Every square matrix in this section is assumed to be diagonalizable.)

Consider adding to $\tilde{M}$ a random perturbation X whose entries are i.i.d. normal with mean 0 and variance σ^2 that is small. This perturbs the eigenvalues $\tilde{\lambda}_i$ of $\tilde{M}$ such that they become $\lambda'_i = \tilde{\lambda}_i + \delta\tilde{\lambda}_i$. Let's derive a formula for $\delta\tilde{\lambda}_i$ to first order in X to show that the mean is 0 and, when the entries of η are constant, the joint distribution is independent of F when all its singular values are distinct and positive.

For an eigenvalue $\tilde{\lambda}$ of $\tilde{M}$ with multiplicity $m \geq 1$, let $\tilde{v}_1, \tilde{v}_2, \ldots, \tilde{v}_m$ be a basis for its eigenspace $D_{\tilde{\lambda}}$. Since $\tilde{M}$ and $\tilde{M}^\top$ have the same characteristic polynomial, $\tilde{M}$ also has a left eigenspace $E^*_{\tilde{\lambda}}$ spanned by $\tilde{w}_1^\dagger, \tilde{w}_2^\dagger, \ldots, \tilde{w}_m^\dagger$, corresponding to eigenvalue $\tilde{\lambda}$.

Under the perturbation X, the perturbed eigenvectors can be decomposed as $\tilde{v}_i' + \delta\tilde{v}_i$, where $\tilde{v}_i' \in D_{\tilde{\lambda}}$ and $\delta\tilde{v}_i \perp E_{\tilde{\lambda}}$.[1] The deviations $\delta\tilde{v}_i$ from the eigenspace can be considered to be small, so

$$\begin{aligned}(\tilde{M} + X)(\tilde{v}_i' + \delta\tilde{v}_i) &= (\tilde{\lambda}_i + \delta\tilde{\lambda}_i)(\tilde{v}_i' + \delta\tilde{v}_i) \\ \tilde{M}\,\delta\tilde{v}_i + X\tilde{v}_i' &= \tilde{\lambda}_i\,\delta\tilde{v}_i + (\delta\tilde{\lambda}_i)\tilde{v}_i', \end{aligned} \tag{S7}$$

where second-order terms have been neglected. Multiplying both sides of eq. (S7) by $\tilde{w}_j^\dagger$ from the left gives

$$\begin{aligned}\tilde{\lambda}\tilde{w}_j^\dagger\delta\tilde{v}_i + \tilde{w}_j^\dagger X\tilde{v}_i' &= \tilde{\lambda}\tilde{w}_j^\dagger\delta v_i + \tilde{w}_j^\dagger(\delta\tilde{\lambda})\tilde{v}_i' \\ \tilde{w}_j^\dagger X\tilde{v}_i' &= \tilde{w}_j^\dagger\tilde{v}_i'(\delta\tilde{\lambda}). \end{aligned} \tag{S8}$$

Let's define the matrices

$$W = \begin{bmatrix} w_1 & w_2 & \ldots & w_m \end{bmatrix} \qquad V = \begin{bmatrix} v_1 & v_2 & \ldots & v_m \end{bmatrix} \qquad VV' = \begin{bmatrix} v_1' & v_2' & \ldots & v_m' \end{bmatrix},$$

where column i of V' is just the coefficients of writing v_i' as a linear combination of the v_j. Also,

[1] Here, we have assumed that $D_{\tilde{\lambda}} \oplus E^\perp_{\tilde{\lambda}} = V$, where $V = \mathbb{C}^{S_R+S_C}$ is the entire vector space. This assumption should generally hold since $\dim D_{\tilde{\lambda}} = m$ and $\dim E^\perp_{\tilde{\lambda}} = S_R + S_C - \dim E_{\tilde{\lambda}} = S_R + S_C - m$. Also, we know that it holds for Hermitian matrices where $D_{\tilde{\lambda}} = E_{\tilde{\lambda}}$.

define the diagonal matrix $\delta\Lambda = \mathrm{diag}(\delta\tilde{\lambda}_1, \delta\tilde{\lambda}_2, \ldots, \delta\tilde{\lambda}_m)$. Then eq. (S8) can be rewritten as

$$W^\dagger X V V' = W^\dagger V V' \delta\Lambda$$
$$(W^\dagger V)^{-1} W^\dagger X V = V' \delta\Lambda V'^{-1} \tag{S9}$$

Thus, the problem becomes diagonalizing the matrix on the left-hand side of eq. (S9), and the resultant eigenvalues will be the perturbations to the original degenerate eigenvalue $\tilde{\lambda}$.

When $\tilde{M}$ is Hermitian, we can choose W and V to be equal unitary matrices and eq. (S9) becomes the first-order perturbation theory result from quantum mechanics.

Let's now study the distribution of the eigenvalue perturbations $\delta\tilde{\lambda}_i$. If we uniformly choose from $\{\delta\tilde{\lambda}_1, \delta\tilde{\lambda}_2, \ldots, \delta\tilde{\lambda}_m\}$ at random,[2] then the expected value is $\frac{1}{m}\mathbb{E}[\mathrm{tr}(\delta\Lambda)] = \frac{1}{m}\mathbb{E}[\mathrm{tr}((W^\dagger V)^{-1} W^\dagger X V)] = 0$. The last step used the fact that every entry of $(W^\dagger V)^{-1} W^\dagger X V$ is a linear combination of entries of X, all of which have zero mean.

Now, consider the case where all singular values of F are distinct and positive. This is true almost surely when F is chosen to be the MaxEnt solution with random perturbations added to its entries to model deviations from the MaxEnt solution. Assume also that the entries of η are constant. WLOG, let them all be 1. Then $\tilde{M} = \begin{bmatrix} 0 & -\alpha F \\ \beta F^\top & 0 \end{bmatrix}$. According to section S4.1.1, the nonzero eigenvalues are the $\pm$ square roots of the nonzero eigenvalues of $-\alpha\beta F F^\top$ (or, equivalently, $-\alpha\beta F^\top F$). Therefore, there are $2\min(S_R, S_C)$ nonzero eigenvalues of the form $\pm i\sqrt{\alpha\beta}\sigma_k$ where σ_k are the $\min(S_R, S_C)$ singular values of F, and the remaining $|S_R - S_C|$ eigenvalues are

[2] This is the most reasonable thing to do when asking the question "What's the perturbation to eigenvalue $\tilde{\lambda}$?" when the eigenvalue is degenerate. Any other interpretation such as asking about the smallest value among the eigenvalue perturbations would break the symmetry of the degeneracy.

zero.

WLOG, suppose $S_R \geq S_C$. Write the SVD of F as $F = \sum_{k=1}^{S_C} \sigma_k v_k u_k^\top$. By section S4.1.1, the eigenvectors of $\tilde{M}$ corresponding to singular value σ_k of F are $(\pm i\sqrt{\alpha\beta}\sigma_k v_k) \oplus \beta F^\top v_k = (\pm i\sqrt{\alpha\beta}\sigma_k v_k) \oplus (\beta\sigma_k u_k)$. Let's scale it to $(\pm i\sqrt{\alpha} v_k) \oplus \left(\sqrt{\beta} u_k\right)$ for simplcity. Similarly, we can obtain the corresponding left eigenvectors: $\left(\mp i\sqrt{\beta} v_k^\top\right) \oplus \left(\sqrt{\alpha} u_k^\top\right)$. Finally, if $v_{S_C+1}, \ldots, v_{S_R}$ is a basis of the nullspace of $F^\top$, then $v_k \oplus 0$ ($S_C + 1 \leq k \leq S_R$) form a basis of the nullspace of $\tilde{M}$.

Consider the orthonormal basis consisting of the vectors $v_k \oplus 0$ ($1 \leq k \leq S_R$) and $0 \oplus u_k$ ($1 \leq k \leq S_C$). In this basis, the eigenspaces of $\tilde{M}$ are solely a function of α and β: the dependence on the entries of F has been absorbed into the change of basis. Therefore, if we always work in this basis constructed from the left and right singular vectors of F, then the set of V and W matrices for the eigenspaces and left eigenspaces of $\tilde{M}$ will be independent of F. Thus, the distribution of eigenvalue perturbations given by eq. (S9) will be independent of F as long as the distribution of the perturbation matrix X is invariant under the change of basis. Indeed, since the new basis is orthonormal, the change-of-basis matrix B is orthogonal. Then $(B^\top X B)_{i,j} = \sum_{k,l} B_{ki} B_{lj} X_{kl}$ is distributed normally with mean 0 and

$$\begin{aligned}
\mathrm{Cov}((B^\top X B)_{i,j}, (B^\top X B)_{i',j'}) &= \sum_{k,l,k',l'} B_{ki} B_{lj} B_{k'i'} B_{l'j'} \,\mathrm{Cov}(X_{kl}, X_{k'l'}) \\
&= \sum_{k,l} B_{ki} B_{lj} B_{ki'} B_{lj'} \sigma^2 \\
&= \delta_{ii'} \delta_{jj'} \sigma^2.
\end{aligned}$$

So $B^\top X B$ is a random matrix with i.i.d. normal random variables with mean 0 and variance σ^2, just like X. This completes the proof that the first-order perturbations to the eigenvalues of $\tilde{M}$ are independent of F when the entries of η are equal and the singular values of F are distinct and positive.

S4.2 Perfectly-nested non-singular F: $F^\top F$ has larger eigenvalues that are less spread out when N_p is high

Let's treat off-diagonal elements of $F^\top F$ as a perturbation δA to the diagonal part A of $F^\top F$. Note that both A and δA are symmetric matrices, so that we can directly cite perturbation theory results from quantum mechanics, which apply to Hermitian matrices.

The eigenvalues of A are just the elements on the diagonal, $a_{ii} = \vec{f_i} \cdot \vec{f_i}$. When N_p is high, the variance of the components of f_i is high, i.e., the mean of the squares of the components of f_i is high (the square of the mean of the components is a constant $1/S_R^2$). Thus, $\vec{f_i} \cdot \vec{f_i}$ is high, thus showing that the eigenvalues of A are larger for larger N_p.

Let's now look how the perturbation δA changes the eigenvalues.

For a non-degenerate eigenvalue a_{ii} corresponding to the unit eigenvector e_i, the first-order perturbation to the eigenvalue is δA_{ii} and the second-order perturbation is

$$\sum_{j \neq i} \frac{\delta A_{ij}^2}{a_{ii} - a_{jj}}. \tag{S10}$$

Each individual term in the sum is positive when $a_{ii} > a_{jj}$ and negative otherwise, meaning that eigenvalues smaller than a_{ii} push it upwards and eigenvalues larger than a_{ii} push it downwards, a phenomenon known as *level repulsion* in quantum mechanics. The effect of level repulsion spreads the eigenvalues out, with the effect greater when the perturbation is larger, i.e., when N_p is smaller.

For a degenerate or near-degenerate eigenspace spanned by unit eigenvectors e_i for $i \in I$, computing the second-order perturbation according to the formula above gives values that explode. This is interpreted as a nonzero first-order perturbation, which can be calculated using degenerate perturbation theory. The result is that the first order perturbations to a_{ii} for $i \in I$ are given by the eigenvalues of the symmetric matrix $\widetilde{\delta A} = (\delta A_{ij})_{i,j \in I}$. The sum of these eigenvalues is $\mathrm{tr}(\widetilde{\delta A}) = 0$, whereas the sum of the squares of the eigenvalues is $\mathrm{tr}(\widetilde{\delta A}^2) = \sum_{i,j \in I} \delta A_{ij}^2$. Thus, the original degenerate eigenvalues a_{ii} ($i \in I$) of A are now split (known as *level splitting* in quantum mechanics), and the spread of the resulting eigenvalues is greater when the entries of $\widetilde{\delta A}$ are greater, i.e., when N_p is smaller.

S4.3 Case where some consumers share the same set of resources

Since the eigenvalues of $F^\top F$ are the squares of the singular values of F, let's study those instead. Treating the random deviation from the MaxEnt matrix as a perturbation, i.e., $F = \overline{F} + \Delta F$, the zeros among the singular values of $\overline{F}$ become the singular values of $\widetilde{\Delta F}$, defined as ΔF with its domain restricted to the nullspace of $\overline{F}$ and image projected onto the left nullspace of $\overline{F}$ (see below). This is an $(S_R - S) \times (S_C - S)$ random matrix where S is the rank of F. If the entries of this random matrix are i.i.d. normal with variance σ^2, then according to the Marchenko-Pastur distribution, the smallest singular value is approximately

$$\sigma\left|\sqrt{S_R - S} - \sqrt{S_C - S}\right| = \sigma\frac{|S_R - S_C|}{\sqrt{S_R - S} + \sqrt{S_C - S}} \tag{S11}$$

for large $S_R - S, S_C - S$. The smallest singular value of $\widetilde{\Delta F}$ is larger when S is larger, i.e., when the topology is more nested according to N_p. In this case, the smallest singular value of F is also larger, translating to the smallest eigenvalues of $F^\top F$ being larger, hence increasing stability.

Derivation of the perturbation to the zero singular values of a matrix. Write the SVD of $\overline{F}$ as $\overline{F} = \sum_k \sigma_k U_k V_k^*$ where σ_k are the *distinct* singular values with multiplicities d_k, $U = \begin{bmatrix} U_1 & U_2 & \ldots U_k \end{bmatrix}$ is unitary, and $V = \begin{bmatrix} V_1 & V_2 & \ldots V_k \end{bmatrix}$ is unitary. Write the SVD of the perturbed matrix F as

$$\overline{F} + \Delta F = \sum_k (U_k + \Delta U_k)\tilde{U}_k(\sigma_k I_{d_k} + \Delta\Sigma_k)\tilde{V}_k^*(V_k + \Delta V_k)^*, \tag{S12}$$

where $\Delta\Sigma_k$ is diagonal, $\tilde{U}_k, \tilde{V}_k \in \mathrm{U}(d_k)$, $\tilde{U}_k\tilde{V}_k^* \approx I_{d_k}$ to first order (unless $\sigma_k = 0$), $U_k^*\Delta U_k \approx 0$ to first order and $V_k^*\Delta V_k \approx 0$ to first order. To first order, eq. (S12) can be rewritten as

$$\Delta F \approx \sum_k \left(\sigma_k(\Delta U_k)V_k^* + \sigma_k U_k(\Delta V_k)^* + \sigma_k U_k(\tilde{U}_k\tilde{V}_k^* - I_{d_k})V_k^* + U_k\tilde{U}_k(\Delta\Sigma_k)\tilde{V}_k^* V_k^*\right). \tag{S13}$$

To obtain the singular value perturbations $\Delta\Sigma_{k_0}$, we multiply from the left by $U_{k_0}^*$ and from the right by V_{k_0} to obtain

$$U_{k_0}^*(\Delta F)V_{k_0} \approx \sigma_{k_0}(\tilde{U}_{k_0}\tilde{V}_{k_0}^* - I_{d_{k_0}}) + \tilde{U}_{k_0}(\Delta\Sigma_k)\tilde{V}_{k_0}^*. \tag{S14}$$

If $\sigma_{k_0} = 0$, then this becomes

$$U^*_{k_0}(\Delta F)V_{k_0} \approx \tilde{U}_{k_0}(\Delta\Sigma_k)\tilde{V}^*_{k_0}, \tag{S15}$$

as desired.

S5 Comparison with Existing Nestedness Metrics

Here, we study the correlation between nestedness and instability using the following 4 metrics of nestedness and 2 metrics of instability.

4 metrics of nestedness:

- The *matrix dipole moment* N_p, as defined in the main text.
- *Nestedness by overlap and decreasing fill* NODF[1].
- The nestedness metric $N_T = 1 - T/100$, where T is the *matrix temperature*. T was defined first in[2], but the definition was ambiguous, so we implement T based on BINMATNEST[3].
- The nestedness metric $N_d = 1 - 4d/S_R S_C$, where d is the *nestedness discrepancy*[4]. Equivalently, this is $N_d = 1 - d_0$ where d_0 is a normalized version of discrepancy introduced by[5].

To reduce computational costs, the algorithm for finding the maximally packed configuration is by simply ordering the species in decreasing degree.

2 metrics of instability:

- $\overline{m}$, as defined in the main text.
- $p_{m>0}$, the proportion of m values that are positive among the $\tilde{M}$ matrices samples for a given topology.

For each pair S_R, S_C with $3 \leq S_R, S_C \leq 50$, we sampled 250 topologies with a roughly equal number of topologies for each s from 0 up to $0.4 S_R S_C$. For each topology, 250 $\tilde{M}$ matrices were sampled, from which the instability metrics $\overline{m}$ and $p_{m>0}$ were calculated. A linear regression was performed between each of the two instability metrics and each of the four nestedness metrics, and the resultant r-values are plotted on a heat map for each (S_R, S_C) pair (fig. S1).

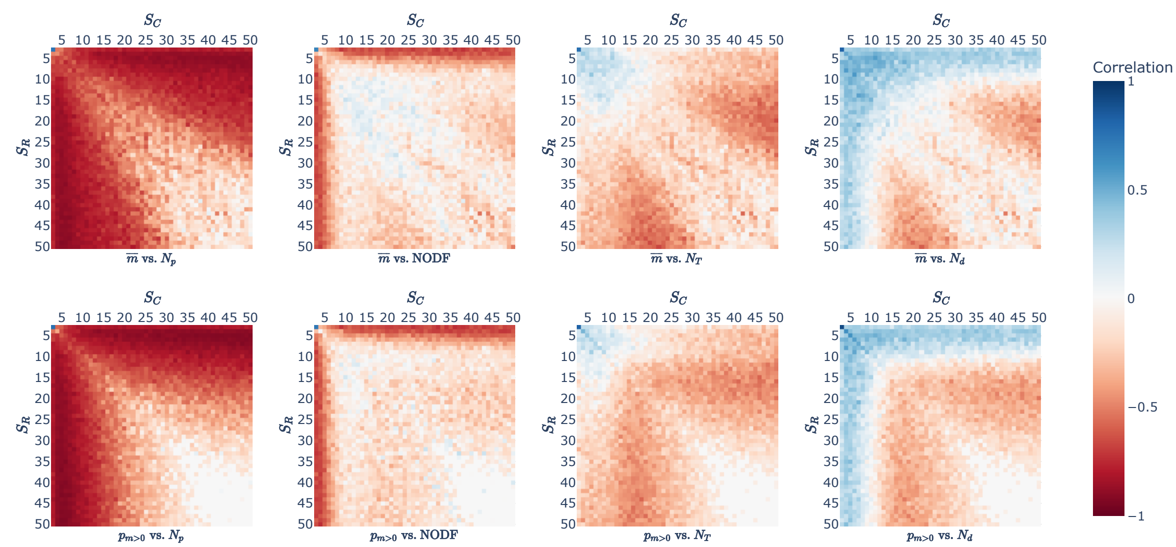

Figure S1: Correlation betwen instability and nestedness for each (S_R, S_C) pair with $3 \leq S_R, S_C \leq 50$ for each instability metric ($\overline{m}, p_{m>0}$) for each nestedness metric (N_p, NODF, N_T, N_d).

We notice significant negative correlation regardless of the instability metric or nestedness metric used, as long as S_R and S_C are not too close to each other and are sufficiently large. In addition, we notice that the negative correlation is present even when S_R or S_C are small if we use the N_p or NODF metrics, as long as $(S_R, S_C) \neq (3,3)$. As expected, the negative correlations are the strongest for N_p.

References

[1] Almeida-Neto, M., Guimaraes, P., Guimaraes Jr, P. R., Loyola, R. D. & Ulrich, W. A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos* **117**, 1227–1239 (2008).

[2] Atmar, W. & Patterson, B. D. The measure of order and disorder in the distribution of species in fragmented habitat. *Oecologia* **96**, 373–382 (1993).

[3] Rodríguez-Gironés, M. A. & Santamaría, L. A new algorithm to calculate the nestedness temperature of presence-absence matrices. *Journal of biogeography* **33**, 924–935 (2006).

[4] Brualdi, R. A. & Sanderson, J. G. Nested species subsets, gaps, and discrepancy. *Oecologia* **119**, 256–264 (1999).

[5] Greve, M. & L. Chown, S. Endemicity biases nestedness metrics: a demonstration, explanation and solution. *Ecography* **29**, 347–356 (2006).

[6] Cross, W. F. *et al.* Ecosystem ecology meets adaptive management: Food web response to a controlled flood on the Colorado River, Glen Canyon. *Ecological Applications* **21**, 2016–2033 (2011).

[7] Bumpers, P. M., Rosemond, A. D., Maerz, J. C. & Benstead, J. P. Experimental nutrient enrichment of forest streams increases energy flow to predators along greener food-web pathways. *Freshwater Biology* **62**, 1794–1805 (2017).

[8] de la Morinière, E. C. *et al.* Ontogenetic dietary changes of coral reef fishes in the mangrove-seagrass-reef continuum: stable isotopes and gut-content analysis. *Marine Ecology Progress Series* **246**, 279–289 (2003).

[9] Rudnick, D. & Resh, V. Stable isotopes, mesocosms and gut content analysis demonstrate trophic differences in two invasive decapod crustacea. *Freshwater Biology* **50**, 1323–1336 (2005).

[10] Judas, M. Gut content analysis of earthworms (Lumbricidae) in a beechwood. *Soil Biology and Biochemistry* **24**, 1413–1417 (1992).

[11] Baird, D. & Ulanowicz, R. E. The seasonal dynamics of the Chesapeake Bay ecosystem. *Ecological Monographs* **59**, 329–364 (1989).

[12] Goldwasser, L. & Roughgarden, J. Construction and analysis of a large Caribbean food web. *Ecology* **74**, 1216–1233 (1993).

[13] Christian, R. R. & Luczkovich, J. J. Organizing and understanding a winter's seagrass food-web network through effective trophic levels. *Ecological Modelling* **117**, 99–124 (1999).

Supplementary Tables

Table S1: Information regarding the food web networks involved in the empirical data analysis. For the “Resources” and “Consumers” columns, the format for each cell is [# of species]: [list of species separated by commas].

Reference	Ecosystem type	Resources	Consumers	Total # linkages	Topologically nested?
6	Riparian	6: Amorphous detritus, Diatoms, Leaves, Filamentous algae, Macrophytes, Fungi	8: Potamopyrgus antipodarum, Gammarus lacustris, Tubificida (a), Physidae, Lumbricidae, Chironomidae, Simuliidae, Other *	40	Yes
7	Forest streams	14: Elmidae, Copepoda, Ceratopogonidae, Hexatoma, Diptera (non-Tanypodinae), Tanypodinae, Maccaffertium, Paraleptophlebia, Serratella, Amphinemura, Isoperla, Leuctra, Tallaperla, Wormaldia	2: Desmognathus quadramaculatus, Eurycea wilderae	28	Yes
8	Mangrove; seagrass bed	12: Tanaidacea, Copepoda, Isopoda, Amphipoda, Mysidecea, Bivalvia, Gastropoda, Decapoda, Fish, Sediment, Other †, Unidentified	5: Haemulon flavolineatum, Haemulon sciurus, Lutjanus apodus, Lutjanus griseus, Ocyurus chrysurus	41	No
	Coral reef	4: Decapoda, Fish, Sediment, Unidentified		13	No
9	Freshwater tributary	4: Detritus, Algae, Invertebrates, Inorganic matter	2: Chinese mitten crab, Red swamp crayfish	8	Yes
10	Beech wood	2: Inorganic matter, Non-particulate organic matter	6: L. terrestris, L. castaneus, Aporrectodea caliginosa, Aporrectodea rosea, O. lacteum, O. cyaneum	12	Yes
11	Estuary (mesohaline)	2: Bay anchovy, Menhaden	3: Blue fish, Summer flounder, Striped bass	6	Yes
12	Tropical island	2: Coleoptera adult, Adult spider	2: Hummingbirds, Elaenia	4	Yes
		3: Coleoptera larva, Diptera larva, Ants	4: Bullfinch, Coleoptera adult, Centipede, Orthoptera	11	Yes
		2: Collembola, Mites	4: Coleoptera larva, Diptera larva, Ants, Thysanoptera	8	Yes
		2: Detritus, Fungi	2: Collembola, Mites	4	Yes
		2: Wood, Detritus	2: Coleoptera larva, Isoptera	4	Yes
13	Halodule wrightii community	3: Meiofauna, Benthic algae, Detritus	4: Brittle stars, Deposit-feeding peracaridan crustaceans, Deposit-feeding gastropods, Deposit-feeding polychaetes	11	Yes
		6: Meiofauna, Epiphyte-grazing amphipods, Deposit-feeding peracaridan crustaceans, Zooplankton, Macroepiphytes, Benthic algae, Detritus	4: Sheepshead minnow, Killifishes, Gobies and blennies, Pipefish and seahorses	20	No

* Ostracoda, Nematoda, Sphaeridae, Cladocera, Copepoda, Tubificida (b)

† Oligochaeta, Polychaeta, Echinoidea, Ostracoda, Seagrass, Foraminifera, Filamentous algae, Calcareous algae, Macroalgae