

Supporting Information

Volkov et al. 10.1073/pnas.0903244106

SI Text

Classical approaches for estimating species interactions in ecological communities were largely built on the theoretical construct of Lotka–Volterra equations (1–3), or their resource-based extensions (4, 5). Most commonly, consumer–resource interaction strengths have been determined from food webs. In predator–prey systems, the interaction strengths have been quantified by obtaining measures of the prey abundances in the absence or presence of predators (6–9) or by deleting species. It has been shown that the interaction strength may be captured as entries in the community matrix, the Jacobian matrix containing the partial derivatives of the system responding to a tiny perturbation of a single species from equilibrium, the inverted Jacobian matrix, or the removal matrix. Interaction strengths can also be determined by measures of matter and energy flow through ecosystems (10) or through allometric approaches. As pointed out by Wootton and Emmerson (11) in a comprehensive review, “measuring interaction strength is difficult because of the large number of interactions in any natural system, long-term feedback, multiple pathways of effects between species pairs, and possible nonlinearities in interaction-strength functions.” Testing this body of theory in the field, particularly in species-rich ecological communities, has been very challenging because of the large number of species-specific parameters (12), intraspecific variability (13), and the complexity introduced by environmental heterogeneity.

Maximum Entropy Framework. Entropy is conceptually linked to the amount of disorder or uncertainty in a system: The higher a system’s entropy, the less certain one can be about its exact state. This connection is formalized in Shannon’s theory of information (14). The amount of information about a system that is provided by the outcome of a measurement of its state is equal to $-\ln(\rho(s))$, where $\rho(s)$ is the probability that the system is in the measured state, s . The expectation value of the information that is contained in an arbitrary measurement of a system’s state is just the weighted average of the information provided by all possible outcomes of the measurement, which is defined as the entropy, $\mathcal{H} = -\sum_i \rho(i) \ln \rho(i)$. For a discrete system with n possible states, entropy is maximized when all states occur with equal probability, $\rho(i) = \frac{1}{n} \forall i$. Conversely, the system about which a set of n outcomes of a measurement provides the most information is that system for which each outcome is equally likely. Thus, if we are given n observed states, the system that most likely generated this set of states is one that maximizes the entropy.

We ask what is the probability distribution function, $\rho(\vec{x})$, which is consistent with the partial knowledge contained in the tropical forest data. As argued by Jaynes (15, 16), the best guess is one that maximizes the entropy but, at the same time, is constrained by available information, such as the mean values of species abundances and 2-species covariances. The system entropy is defined as

$$\mathcal{H} = -\sum_{\vec{x}} \rho(\vec{x}) \ln \rho(\vec{x}) \quad [\text{S1}]$$

and the constraints are provided by the equations

$$\sum_{\vec{x}} \rho(\vec{x}) = 1 \quad [\text{S2}]$$

$$\langle x_i \rangle = \sum_{\vec{x}} \rho(\vec{x}) x_i = \frac{1}{P} \sum_{k=1}^P x_i^k \quad [\text{S3}]$$

and

$$\langle x_i x_j \rangle = \sum_{\vec{x}} \rho(\vec{x}) x_i x_j = \frac{1}{P} \sum_{k=1}^P x_i^k x_j^k. \quad [\text{S4}]$$

Here, x_i^k corresponds to the abundance of the i th species in the k th quadrat. Eq. S2 provides the normalization condition that the probabilities of all observable states sum to 1. Eqs. S3 and S4 ensure that the distribution $\rho(\vec{x})$ preserves the mean abundance of each species as well as the correlations between species. It should be noted that our implicit assumption of using entropy rather than relative entropy corresponds to a uniform distribution for the maximally uninformative prior. Such an assumption is valid when the states (in our case, the 20 most abundant species) are discrete, mutually exclusive, fixed in number, and do not have any preferential order.

It should be noted that, in principle, one could incorporate additional conditions on the entropy that relate to higher-order, e.g. triple-point, correlation measures. However, the associated mathematics becomes much more complicated than due to the resulting non-Gaussian nature of the probability distribution function. Instead, we truncate the list of conditions at the 2-point level to get the leading form of $\rho(\vec{x})$. The constraints in Eqs. S3 and S4 can be used to define the covariance matrix, Σ , as

$$\Sigma_{ij} = \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle. \quad [\text{S5}]$$

The entropy maximization is carried out by introducing the Lagrange multipliers ν , μ_i and Ω_{ij} and maximizing

$$\begin{aligned} \zeta &= \mathcal{H} - \nu \sum_{\vec{x}} \rho(\vec{x}) - \sum_{i=1}^S \mu_i \sum_{\vec{x}} \rho(\vec{x}) x_i - \frac{1}{2} \sum_{i,j=1}^S \Omega_{ij} \sum_{\vec{x}} \rho(\vec{x}) x_i x_j \\ &= -\sum_{\vec{x}} \rho(\vec{x}) \left[\ln \rho(\vec{x}) + \nu + \vec{\mu} \cdot \vec{x} + \frac{1}{2} \vec{x} \Omega \vec{x} \right]. \end{aligned} \quad [\text{S6}]$$

We set $\partial \zeta / \partial \rho(\vec{x}) = 0$ and get

$$\rho(\vec{x}) = e^{-1-\nu-\vec{\mu} \cdot \vec{x} - \frac{1}{2} \vec{x} \Omega \vec{x}} = \frac{1}{Z} e^{-\frac{1}{2} \vec{y} \Omega \vec{y}}, \quad [\text{S7}]$$

where $\vec{y} = \vec{x} + \Omega^{-1} \vec{\mu}$ is the species abundance vector that is shifted by a constant, and the quantity $Z = e^{-\frac{1}{2} \vec{\mu} \Omega^{-1} \vec{\mu}} e^{1+\nu}$ serves to normalize the probability. We use a simplified notation in which a vector written to the left of a matrix is meant to be in a transposed form.

The matrix element $-\Omega_{ij}$ has the natural interpretation of the interaction between species i and j . In order to calculate the elements of Ω , we note that the states \vec{x} span a continuous space, and we replace the summations in the constraints (Eqs. S2–S4) with integrals. Thus, the normalization condition of Eq. S2 becomes

$$1 = \int d^S x \rho(\vec{x}) = \frac{1}{Z} \int d^S y e^{-\frac{1}{2} \vec{y} \Omega \vec{y}}, \quad [\text{S8}]$$

and Eq. S3 can similarly be rewritten as

$$\langle x_i \rangle = \int d^S x \rho(\vec{x}) x_i \quad [\text{S9}]$$

$$= \frac{1}{Z} \int d^S y e^{-\frac{1}{2} \vec{y} \Omega \vec{y}} \left(y_i - \sum_j \Omega_{ij}^{-1} \mu_j \right) = -\sum_j \Omega_{ij}^{-1} \mu_j. \quad [\text{S10}]$$

Note that Eq. S9 and the previous definition $\bar{y} = \bar{x} + \Omega^{-1}\bar{\mu}$ imply that $\bar{y} = \bar{x} - \langle \bar{x} \rangle$. Thus, when the data are shifted so that $\langle x_i \rangle = 0 \forall i$, the elements of Ω are the coupling constants that link the species directly.

Applying the continuum treatment to Eq. S4, we find

$$\langle x_i x_j \rangle = \int d^S x \rho(\bar{x}) x_i x_j \quad [\text{S11}]$$

$$= \frac{1}{Z} \int d^S y e^{-\frac{1}{2} \bar{y} \Omega \bar{y}} [\langle x_i \rangle \langle x_j \rangle + y_i \langle x_j \rangle + y_j \langle x_i \rangle + y_i y_j] \quad [\text{S12}]$$

$$= \langle x_i \rangle \langle x_j \rangle + \frac{1}{Z} \int d^S y e^{\frac{1}{2} \bar{y} \Omega \bar{y}} y_i y_j. \quad [\text{S13}]$$

The integral in Eq. S11 can be solved by defining the generating function

$$\Upsilon(\vec{J}) = \int d^S y e^{-\frac{1}{2} \bar{y} \Omega \bar{y} + \vec{J} \cdot \bar{y}}. \quad [\text{S14}]$$

The source field \vec{J} in Eq. S14 allows us to write the integral in Eq. S11 in terms of derivatives of $\Upsilon(\vec{J})$ taken with respect to \vec{J} and then setting $\vec{J} = 0$:

$$\left. \frac{\partial^2 \Upsilon(\vec{J})}{\partial J_i \partial J_j} \right|_{\vec{J}=0} = \int d^S y e^{-\frac{1}{2} \bar{y} \Omega \bar{y}} y_i y_j. \quad [\text{S15}]$$

By making the substitution $\bar{z} = \bar{y} - \Omega^{-1} \vec{J}$, Eq. S14 can be rewritten as

$$\Upsilon(\vec{J}) = e^{\frac{1}{2} \vec{J} \Omega^{-1} \vec{J}} \int d^S z e^{-\frac{1}{2} \bar{z} \Omega \bar{z}} = e^{\frac{1}{2} \vec{J} \Omega^{-1} \vec{J}} \Upsilon(\vec{J})|_{\vec{J}=0}, \quad [\text{S16}]$$

yielding

$$\frac{1}{Z} \int d^S y e^{\frac{1}{2} \bar{y} \Omega \bar{y}} y_i y_j = \frac{1}{\Upsilon(\vec{J})} \left. \frac{\partial^2}{\partial J_i \partial J_j} \Upsilon(\vec{J}) \right|_{\vec{J}=0} = \Omega_{ij}^{-1}. \quad [\text{S17}]$$

Substituting back into Eq. S11, we find that $\Omega^{-1} = \Sigma$, and the interaction matrix $M = -\Omega$ is simply the negative inverse of the covariance matrix.

It should be noted that the form of M is independent of the mean species abundances $\langle x_i \rangle$. Because the covariance matrix Σ is not affected by changes in the mean abundances, neither is its inverse. Furthermore, this is a robust result for linear systems and can be derived in several ways. An alternative way of arriving at this result without invoking the maximization of entropy follows from the assumptions that $\ln \rho(\bar{x})$ peaks at $\bar{x}^{(0)}$, is normalizable and is a smooth function that can be expressed in a Taylor expansion up to quadratic order:

$$\ln \rho(\bar{x}) = \ln \rho(\bar{x}^{(0)}) + \frac{1}{2} \sum_{ij} (x_i - x_i^{(0)}) M_{ij} (x_j - x_j^{(0)}) + \dots, \quad [\text{S18}]$$

where the neglected terms are of cubic order in $(x_i - x_i^{(0)})$ and M , the matrix of the second derivative of $\ln \rho(\bar{x})$ with respect to \bar{x} , is negative definite. Note that $\bar{x}^{(0)} = \langle \bar{x} \rangle$. Within this Gaussian approximation, one again obtains the result that $-M$ is the inverse of Σ .

The elements of the matrix M are, by definition, the effective pairwise species interactions that reproduce the species covariances exactly while maximizing the entropy of the system. The strength and the sign of the interaction represent the mutual influence on each other of a pair of species. The magnitude of the element M_{ij} is a measure of the strength of the net interaction between species i and j . The sign of the interaction indicates the nature of the coupling: A positive coupling between species

indicates that a change in abundance of either species is accompanied by a similar change in the abundance of the other species. Conversely, a negative coupling indicates that a change in one is accompanied by an opposite change in the other.

To guarantee self-consistency, the method must allow one to integrate an existing interaction matrix over individual species, thereby calculating new effective interactions without explicit consideration of the integrated species. For example, one might start with the probability distribution of Eq. S7. Assuming that some of the species considered in M are not of interest, and one wishes instead to determine the form of the interactions when these species are integrated over. In other words, we wish to find the effective interactions considering the mediating influence of the species we are integrating over. We aim to find the form of the interaction matrix R given by

$$\frac{1}{Z} \int \prod_{k \notin \{q\}} dx_k e^{-\frac{1}{2} \bar{x} \Omega \bar{x}} = \frac{1}{Z'} \exp \left\{ -\frac{1}{2} \sum_{ij \in \{q\}} x_i R_{ij} x_j \right\}, \quad [\text{S19}]$$

where Z' is some normalization factor. This can be readily solved by introducing the matrix $\bar{\Omega}$ with elements $\bar{\Omega}_{ij} = \Omega_{ij}$ if neither i nor j are in $\{q\}$ and $\bar{\Omega}_{ij} = 0$ otherwise, as well as the vector \vec{J} with elements $J_i = \sum_{j \in \{q\}} x_j \Omega_{ji}$. Then the integral in Eq. S19 can be written as

$$\int \prod_{k \notin \{q\}} dx_k e^{-\frac{1}{2} \bar{x} \Omega \bar{x}} = \exp \left\{ -\frac{1}{2} \sum_{ij \in \{q\}} x_i \Omega_{ij} x_j \right\} \int \prod_{k \notin \{q\}} dx_k e^{-\frac{1}{2} \bar{x} \bar{\Omega} \bar{x} - \vec{J} \cdot \bar{x}}. \quad [\text{S20}]$$

Solving and equating to the right-hand-side of Eq. S19, we find that the ‘‘dressed’’ interactions between the species in the set $\{q\}$ are given by

$$R_{ij} = \Omega_{ij} - \sum_{\alpha\beta} \Omega_{i\alpha} \bar{\Omega}_{\alpha\beta}^{-1} \Omega_{\beta j}. \quad [\text{S21}]$$

Higher-Order Interactions. The mathematics of our method is greatly simplified by assuming that the interactions between species are primarily pairwise and the simple result is that the effective couplings between species can be derived from the matrix of their covariances. Nevertheless, the network need not be strictly Gaussian, and higher order interactions are measures of non-Gaussian effects. Our approach can be generalized to systematically handle 3-body and higher-order interactions in a perturbative fashion because the nonlinear case with the 3-species interaction is not amenable to exact analytic calculation. We use standard perturbation theory, which assumes that the characteristic 3-species correlation function is much smaller than the pairwise correlation. We begin by considering a system for which the data have been adjusted such that $\langle x_i \rangle = 0 \forall i$ and defining the unperturbed probability distribution function

$$\rho_0(\bar{x}) = \frac{1}{Z_0(\vec{h})} \exp \left[-\frac{1}{2} \bar{x} \Omega \bar{x} + \vec{h} \cdot \bar{x} \right], \quad [\text{S22}]$$

where $\Omega^{-1} \equiv \Sigma$ and the unperturbed partition function is given by

$$Z_0(\vec{h}) = \int d^S x \exp \left[-\frac{1}{2} \bar{x} \Omega \bar{x} + \vec{h} \cdot \bar{x} \right] = Z_0(0) e^{\frac{1}{2} \vec{h} \Omega^{-1} \vec{h}}. \quad [\text{S23}]$$

As before, we will express quantities of interest in terms of derivatives with respect to the components of the field term, \vec{h} , and then set $\vec{h} = 0$ at the end of the calculations.

When the 3-species term is included, we find the perturbed probability distribution function

$$\rho(\vec{x}) = \frac{1}{Z} \exp \left[-\frac{1}{2} \sum_{ij} x_i \Omega_{ij} x_j + \sum_i h_i x_i + \frac{1}{6} \sum_{ijk} \Gamma_{ijk} x_i x_j x_k \right], \quad [\text{S24}]$$

where Z is the perturbed partition function and can be expressed in terms of the unperturbed distribution. Defining the unperturbed expectation value of an arbitrary observable \cdot by $\langle \cdot \rangle_0 = \int d^S x \rho_0(\vec{x}) \cdot$, we have

$$Z = \int d^S x \exp \left[-\frac{1}{2} \sum_{ij} x_i \Omega_{ij} x_j + \sum_i h_i x_i + \frac{1}{6} \sum_{ijk} \Gamma_{ijk} x_i x_j x_k \right] \quad [\text{S25}]$$

$$= Z_0 \left\langle \exp \left[\frac{1}{6} \sum_{ijk} \Gamma_{ijk} x_i x_j x_k \right] \right\rangle_0 \quad [\text{S26}]$$

$$= Z_0 \left(1 + \frac{1}{6} \sum_{ijk} \Gamma_{ijk} \langle x_i x_j x_k \rangle_0 + \mathcal{O}(\Gamma^2) \right), \quad [\text{S27}]$$

where the exponent in Eq. S26 has been expanded, and terms beyond the leading order dropped. We can find the form of the 3-species interactions through the 3-body correlation,

$$\langle x_\alpha x_\beta x_\gamma \rangle = \frac{\partial^3 \ln Z}{\partial h_\alpha \partial h_\beta \partial h_\gamma} + \langle x_\alpha \rangle \langle x_\beta x_\gamma \rangle + \langle x_\beta \rangle \langle x_\alpha x_\gamma \rangle + \langle x_\gamma \rangle \langle x_\alpha x_\beta \rangle - 2 \langle x_\alpha \rangle \langle x_\beta \rangle \langle x_\gamma \rangle. \quad [\text{S28}]$$

Because the data all have zero mean, only the first term is nonzero, leaving

$$\langle x_\alpha x_\beta x_\gamma \rangle = \frac{\partial^3 \ln Z}{\partial h_\alpha \partial h_\beta \partial h_\gamma} \quad [\text{S29}]$$

$$= \frac{\partial^3 \ln Z_0}{\partial h_\alpha \partial h_\beta \partial h_\gamma} + \frac{\partial^3}{\partial h_\alpha \partial h_\beta \partial h_\gamma} \frac{1}{6} \sum_{ijk} \Gamma_{ijk} \langle x_i x_j x_k \rangle_0 \quad [\text{S30}]$$

$$= \frac{1}{6} \sum_{ijk} \Gamma_{ijk} \frac{\partial^3 \langle x_i x_j x_k \rangle_0}{\partial h_\alpha \partial h_\beta \partial h_\gamma}. \quad [\text{S31}]$$

Differentiating and dropping terms containing $\langle x \rangle_0$ or $\langle x^3 \rangle_0$, we find:

$$\frac{\partial^3 \langle x_i x_j x_k \rangle_0}{\partial h_\alpha \partial h_\beta \partial h_\gamma} = \langle x_i x_j x_k x_\alpha x_\beta x_\gamma \rangle_0 - \langle x_\alpha x_\beta \rangle_0 \langle x_i x_j x_k x_\gamma \rangle_0 - \langle x_\alpha x_\gamma \rangle_0 \langle x_i x_j x_k x_\beta \rangle_0 - \langle x_\beta x_\gamma \rangle_0 \langle x_i x_j x_k x_\alpha \rangle_0. \quad [\text{S32}]$$

This can be simplified using Wick's theorem (17), which states that the many-body correlation averaged over Gaussian distributions can be written as the sum of all distinct products of pair correlations. A 4-species average $\langle x_1 x_2 x_3 x_4 \rangle_0$ can then be written as

- Levins R (1968) *Evolution in Changing Environments* (Princeton Univ Press, Princeton).
- Vandermeer JH (1969) The competitive structure of communities: An experimental approach using protozoa. *Ecology* 50:362–371.
- Neill WE (1974) The community matrix and the interdependence of competition coefficients. *Am Nat* 108:339–408.
- Hsu SB, Hubbell SP, Waltman P (1977) A mathematical theory for single-nutrient competition in continuous cultures of microorganisms. *J Appl Math* 32:266–383.
- Tilman D (1982) *Resource Competition and Community Structure* (Princeton Univ Press, Princeton).
- Paine RT (1992) Food-web analysis through field measurement of per capita interaction strength. *Nature* 355:73–75.
- Fagan WF, Hurd LE (1994) Hatch density variation of a generalist arthropod predator: Population consequences and community impact. *Ecology* 75:2022–2032.
- Raffaelli DG, Hall SJ (1995) Assessing the relative importance of trophic links in food webs. *Food Webs: Integration of Patterns and Dynamics*, eds Polis G, Winemiller K (Chapman and Hall, New York), pp 185–191.
- Berlow EL (1999) Strong effects of weak interactions in ecological communities. *Nature* 398:330–334.

$$\langle x_1 x_2 x_3 x_4 \rangle_0 |_{\vec{h}=0} = \langle x_1 x_2 \rangle_0 \langle x_3 x_4 \rangle_0 |_{\vec{h}=0} + \langle x_1 x_3 \rangle_0 \langle x_2 x_4 \rangle_0 |_{\vec{h}=0} + \langle x_1 x_4 \rangle_0 \langle x_2 x_3 \rangle_0 |_{\vec{h}=0}. \quad [\text{S33}]$$

Similarly, the 6-species correlation $\langle x_1 x_2 x_3 x_4 x_5 x_6 \rangle_0$ can be expanded into a sum of 15 sets of products of 3 pair correlations. Expanding Eq. S32 in this way and using the identity $\langle x_i x_j \rangle_0 |_{\vec{h}=0} = \Omega_{ij}^{-1}$, we find

$$\frac{\partial^3 \langle x_i x_j x_k \rangle_0}{\partial h_\alpha \partial h_\beta \partial h_\gamma} \Big|_{\vec{h}=0} = \Omega_{\alpha i}^{-1} \Omega_{\beta j}^{-1} \Omega_{\gamma k}^{-1} + \Omega_{\alpha i}^{-1} \Omega_{\beta k}^{-1} \Omega_{\gamma j}^{-1} + \Omega_{\alpha j}^{-1} \Omega_{\beta i}^{-1} \Omega_{\gamma k}^{-1} + \Omega_{\alpha j}^{-1} \Omega_{\beta k}^{-1} \Omega_{\gamma i}^{-1} + \Omega_{\alpha k}^{-1} \Omega_{\beta i}^{-1} \Omega_{\gamma j}^{-1} + \Omega_{\alpha k}^{-1} \Omega_{\beta j}^{-1} \Omega_{\gamma i}^{-1}, \quad [\text{S34}]$$

and the 3-species interaction (Eq. S29) is

$$\Gamma_{ijk} = \sum_{\alpha\beta\gamma=1}^S \langle x_\alpha x_\beta x_\gamma \rangle_0 \Omega_{i\alpha} \Omega_{j\beta} \Omega_{k\gamma}. \quad [\text{S35}]$$

The distribution of the strengths of the perturbatively obtained 3-species interactions is shown here in Fig. S5. The majority of the 3-species interactions are significantly smaller than the 2-body ones. However, when 2-body interactions between species are very small, the 3-body effects may in fact play an important role in the dynamics of the system.

Our discussion here closely parallels an earlier study by Lezon et al. (18) on the inference of gene interactions from microarray data. There is a key difference between the 2 approaches. In the analysis of gene interactions, the variance of gene expressions was normalized to be equal to 1 for all genes, whereas here, we leave the data unnormalized. Thus instead of treating the correlation matrix of the gene data, here, we have the covariance matrix of the unnormalized species abundance data.

Derivation of the Results Presented in Fig. 3 in the Main Text. In the absence of interspecific interactions, consider the following expressions for the birth and death rates of a species as a function of its abundance n : $b_n = bn + cn^2$ ($c < 0$), $d_n = n$, $d_1 = b_0 = 0$ (reflecting boundary conditions).

The master equation for $P_n(t)$, the probability of having n individuals at time t , is given by

$$\frac{\partial P_n(t)}{\partial t} = P_{n-1}(t)b_{n-1} + P_{n+1}(t)d_{n+1} - P_n(t)(b_n + d_n). \quad [\text{S36}]$$

Assuming that the system is in steady state and setting $\frac{\partial P_n(t)}{\partial t} = 0$, one obtains the following expression for the probability:

$$P_n = P_1 \prod_{i=1}^{n-1} \frac{b_i}{d_{i+1}} = P_1 \frac{(-c)^{n-1}}{n} \frac{\Gamma(\alpha)}{\Gamma(\alpha - n + 1)}, \quad [\text{S37}]$$

where $\alpha \equiv -b/c$. The average abundance of the species can then be deduced to be

$$\langle n \rangle = \sum_{n=1}^{\alpha} n P_n = \frac{(-c)^\alpha e^{-1/c} \Gamma(1 + \alpha, -1/c) - 1}{\sum_{m=1}^{\alpha} \frac{(-c)^m}{m} \frac{\Gamma(\alpha+1)}{\Gamma(\alpha-m+1)}}. \quad [\text{S38}]$$

- Hall RO, Jr, Wallace BJ, Eggert SL (2000) Organic matter flow in stream food webs with reduced detrital resource base. *Ecology* 81:3445–3463.
- Wootton JT, Emmerson M (2005) Measurement of interaction strength in nature. *Annu Rev Ecol, Evol Systematics* 36:419–444.
- Hubbell SP (2006) Neutral theory in ecology and the evolution of ecological equivalence. *Ecology* 87:1387–1398.
- Hubbell SP (2005) Neutral theory in ecology and the hypothesis of functional equivalence. *Funct Ecol* 19:166–177.
- Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27:379–423, 623–656.
- Jaynes ET (1957) Information theory and statistical mechanics. *Phys Rev* 106:620–630.
- Jaynes ET (1957) Information theory and statistical mechanics II. *Phys Rev* 108:171–190.
- Zinn-Justin J (2007) *Phase Transitions and Renormalisation Group*. (Oxford Univ Press, New York).
- Lezon T, Banavar JR, Cieplak M, Maritan A, Fedoroff NV (2006) Using entropy maximization to infer genetic interaction networks from gene expression patterns. *Proc Natl Acad Sci USA* 103:19033–19038.

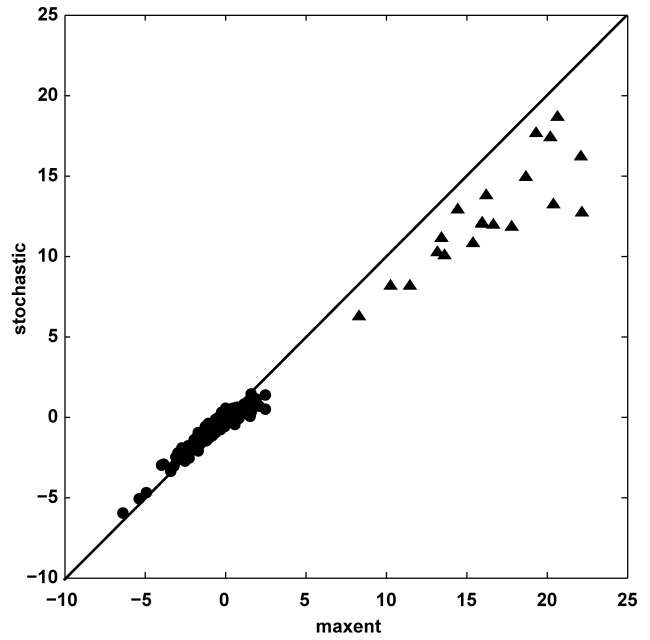
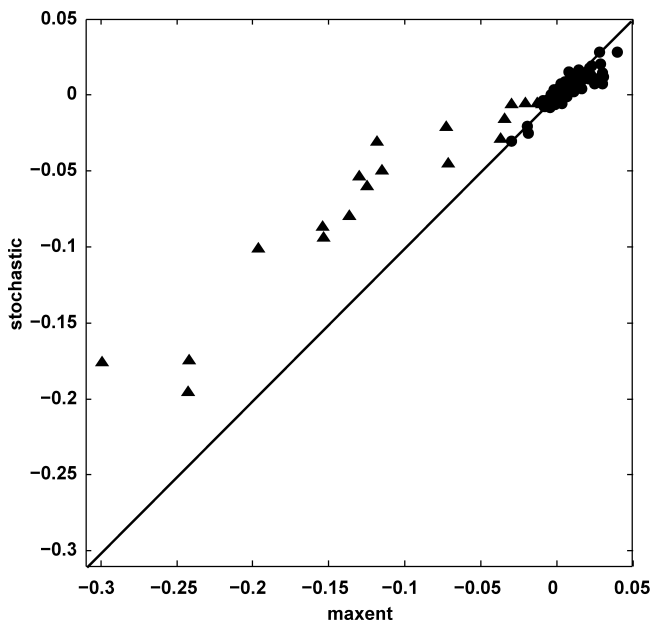


Fig. S1. Comparison of the interaction parameters deduced by using the maximum entropy and the stochastic dynamics frameworks. (Left) Scatter plot of the interaction coefficients obtained using the 2 methods. The correlation is $\approx 97\%$. Note, however, the systematic deviation from the $y = x$ line for the diagonal (intraspecific) interactions (filled triangles). This deviation arises from the fact that intraspecific interactions in the maxent case is of the form $c_{ii}n_i^2$, whereas in the stochastic dynamics framework, there is an additional birth/death rate contributions that depend linearly on n_i . (Right) Scatter plot of the deviations of the interaction parameters from the randomized plots measured in units of the respective standard deviations obtained using the 2 methods. The correlation is $\approx 99\%$. These results underscore the consistency in the inferences using these 2 methods. Interestingly, studies of 1,000 configurations of randomly shuffled trees in the forest yielded a correlation $>99\%$ between the inferences of the 2 methods in every configuration underscoring the virtually perfect accord between the 2 methods.

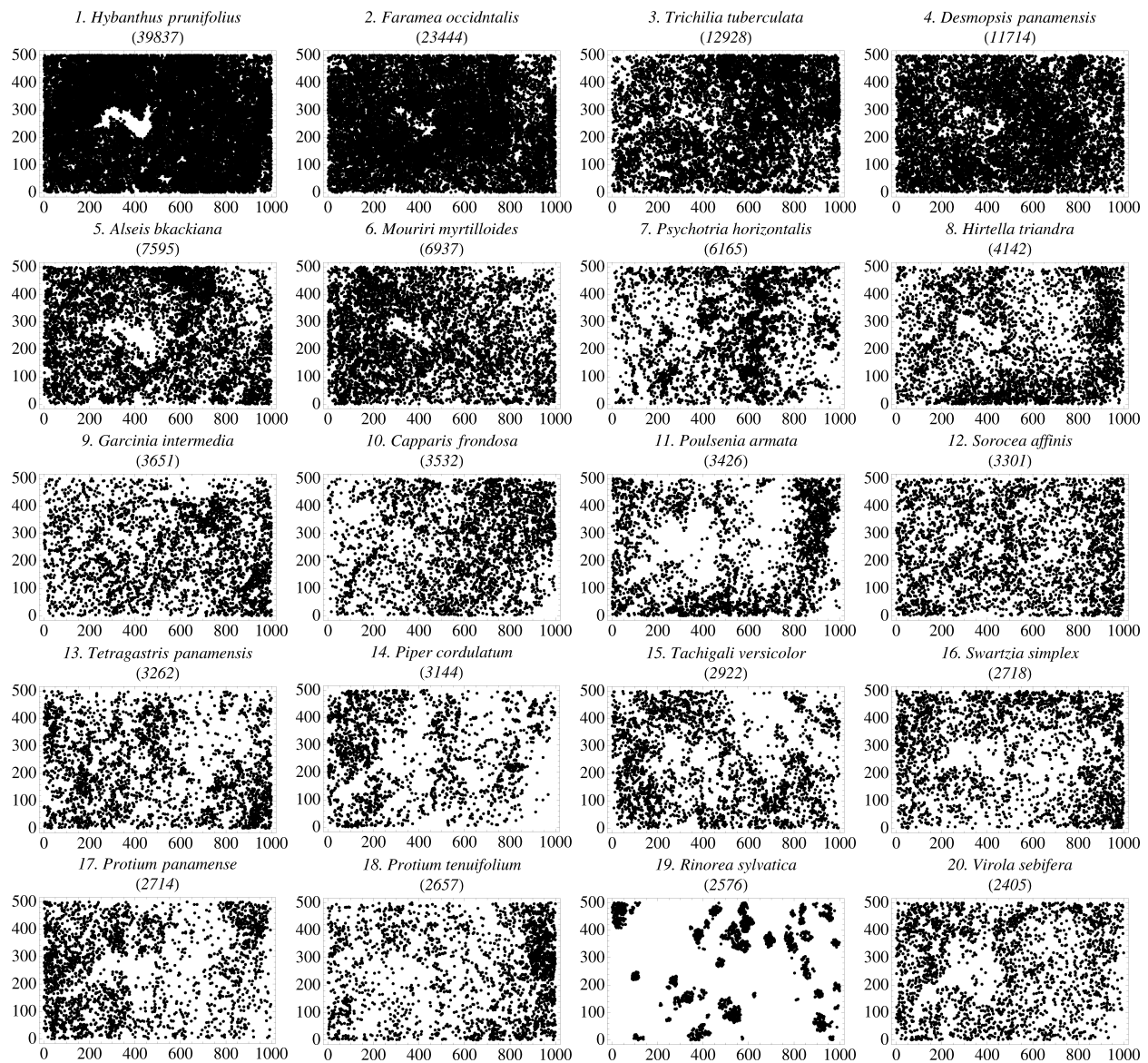


Fig. S2. Locations of the trees of the 20 most abundant species in the 1,000 × 500 m area in the BCI tropical forest. Note the anomalous distribution of species 19.

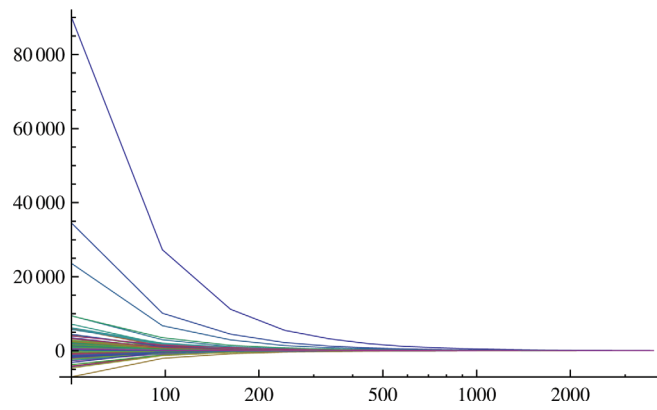


Fig. S3. Sensitivity of the results on the number of quadrats. Shown is a plot of the elements of the covariance matrix as a function of the number of quadrats that the plot has been divided into. The data become substantially independent of the number of quadrats for a large enough number of quadrats. Our choice of 1,250 quadrats is dictated by the requirement that we are in this flat regime but with as few quadrats as possible so that the mean occupancy of each quadrat for even the 20th ranked species is still ≈ 2 .

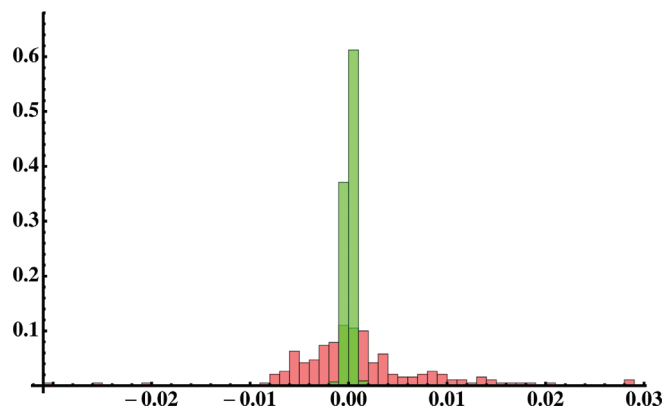


Fig. S5. Distribution of the coefficients for 2- and 3-species interactions. Green histogram: $\Gamma_{ijk}, i \neq j \neq k$; red histogram: $M_{ij}, i \neq j$.

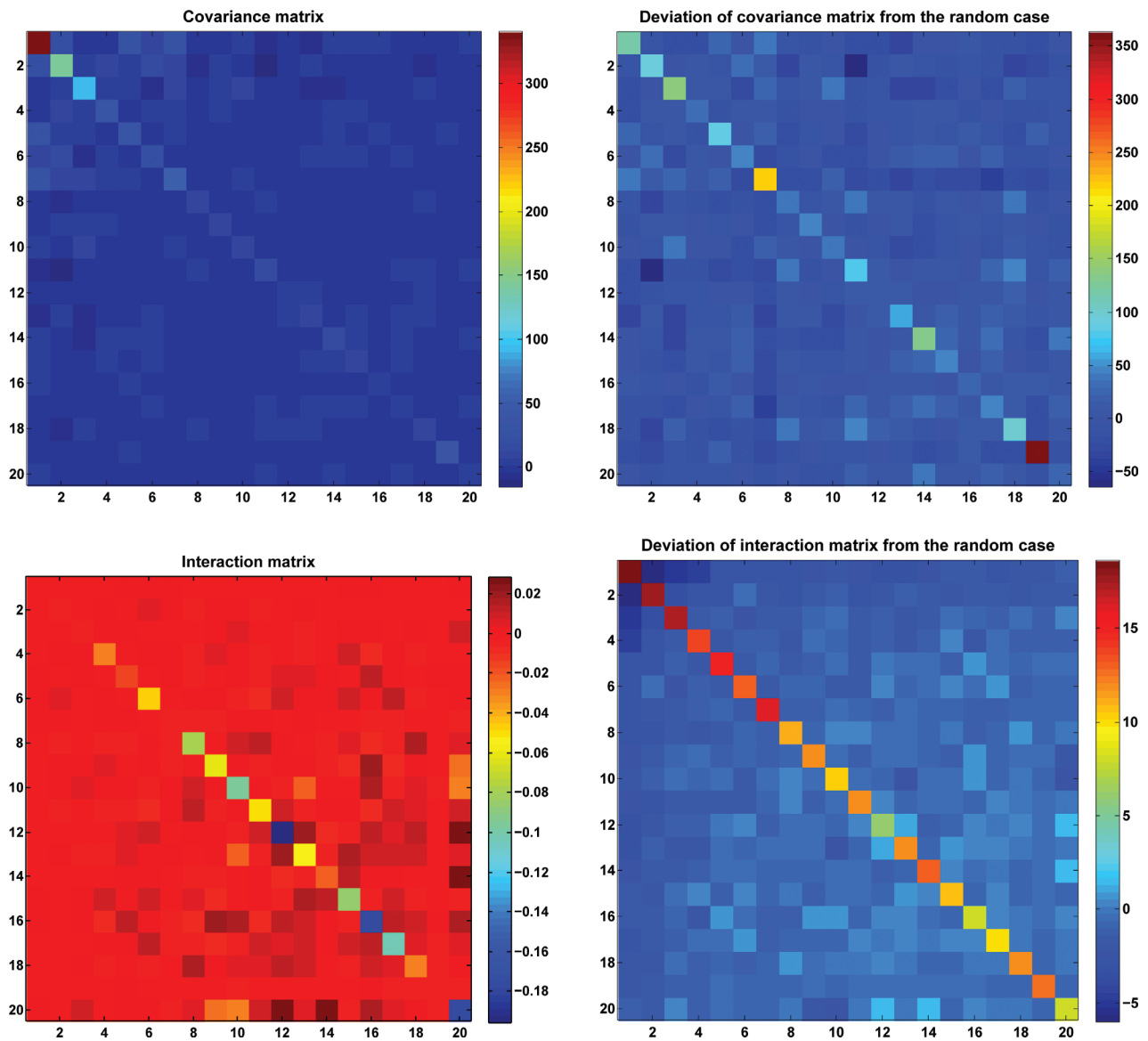


Fig. S6. Plots of covariance and interaction matrices (see Fig. 1 in the main text) with included diagonal values.