

Constructing random matrices to represent real ecosystems

Alex James¹

Michael J. Plank²

Axel G. Rossberg³

Jonathan Beecham⁴

Mark Emmerson⁵

Jonathan W. Pitchford⁶

1. Corresponding author. alex.james@canterbury.ac.nz.

Biomathematics Research Centre, University of Canterbury, New Zealand.

2. michael.plank@canterbury.ac.nz.

Biomathematics Research Centre, University of Canterbury, New Zealand.

3. Axel@rossberg.net.

Centre for Environment, Fisheries and Aquaculture Science (Cefas), Lowestoft, UK

School of Biological Sciences, Queens University Belfast, UK.

4. jonathan.beecham@cefas.co.uk.

Centre for Environment, Fisheries and Aquaculture Science (Cefas), Lowestoft, UK.

5. m.emmerson@qub.ac.uk.

School of Biological Sciences, Queens University Belfast, UK.

6. jon.pitchford@york.ac.uk.

Departments of Biology and Mathematics, University of York, UK.

Keywords: community matrix; complexity; food web; interaction strength; stability;

predator–prey interaction.

Online version also includes Appendices.

Article type: Article

Abstract

22 Models of complex systems with n components typically require knowledge of
the order of n^2 parameters, because each component can, in principle, interact
24 with every other. When it is impractical to measure these parameters one may
choose random parameter values and study the emergent statistical properties at
26 the system level. Many influential results in theoretical ecology have been derived
from two key assumptions: that species interact with random partners at random
28 intensities, and that intraspecific competition is comparable between species. Under
these assumptions, community dynamics can be described by a community matrix
30 which is often amenable to mathematical analysis. We combine empirical data with
mathematical theory to show that both these assumptions lead to results that must
32 be interpreted with caution.

We examine 21 empirically derived community matrices constructed using three
34 established independent methods. The empirically derived systems are more stable
by some orders of magnitude than the results from random matrices. This surpris-
36 ing and consistent disparity is not explained by existing results on predator-prey
interactions. We ask the important question: what are the key properties of em-
38 pirical community matrices which distinguish them from random matrices? We
show that network topology typically has less effect than **the relationship be-**
40 **tween a species' trophic position within the food web and its interaction**
strengths. We identify key features of empirical networks that must be preserved
42 if models based on random matrices are to capture the features of real ecosystems.

1 Introduction

44 Interactions between species are central to the concept of an ecosystem. They are, how-
ever, both expensive and technically challenging to measure empirically. It is natural,
46 therefore, that ecologists have sought to understand to what extent these interactions
can be thought of as random, and furthermore to understand and quantify the possible
48 relationships between these interactions which best confer the features of ecosystem sta-
bility, resilience and dynamics observed in nature. For these reasons, theories involving
50 randomly generated interactions between species have underpinned many influential ideas
concerning the stability of complex ecological networks (May, 1972; Pimm and Lawton,
52 1978; Yodzis, 1981; Bastolla et al., 2009; Allesina and Tang, 2012).

May (1972) famously showed that, under certain assumptions, there is a limit on how
54 “complex” a network ecosystem with a stable equilibrium can be. He then hypothesized
that this was relevant to the stability of ecological networks. Stability, in this context,
56 means there is an equilibrium in which all species in the network survive at some pos-
itive density, and that this equilibrium is robust to sufficiently small perturbations in
58 the species densities. Under this definition, stability of the equilibrium is quantified by
the leading eigenvalue of the Jacobian matrix evaluated at the equilibrium point, which
60 describes the behaviour of the system close to equilibrium. Whether complex ecosystems
operate close to equilibrium is matter of some debate, and local stability is not the only
62 way of quantifying the resilience of an ecosystem to change (Grimm and Wissel, 1997; Mc-
Cann, 2000). Other measures include, for example, permanence (Jansen, 1987; Law and
64 Blackford, 1992), persistence (Bastolla et al., 2009; Thébault and Fontaine, 2010; Gravel
et al., 2011) and species-deletion stability (Pimm, 1980). Nevertheless, local stability is a
66 necessary condition for a persistent equilibrium and is a widely used measure of ecosystem
robustness (Thébault and Fontaine, 2010; Allesina and Tang, 2012; Staniczenko et al.,
68 2013) and we focus on local stability in this paper.

The conclusions of May (1972) contradicted earlier ideas that ecosystems with more
70 species, and more interactions among those species, are more likely to be stable (MacArthur,
1955; Elton, 1958). In reality, highly complex ecological networks, i.e. networks with
72 many species, many interactions and strong interactions, do exist (O’Gorman and Em-
merson, 2010; Twomey et al., 2012) and empirical evidence frequently points to a posi-
74 tive relationship between complexity and stability (Yodzis, 1981; De Ruiter et al., 1995;
McCann, 2000; Gravel et al., 2011). Various theoretical explanations for this apparent
76 “stability despite complexity” have been proposed, including a skew in the interaction
strength distribution towards weak links (McCann et al., 1998; Emmerson and Yearsley,
78 2004), weak links in long loops (Neutel et al., 2002), a pyramidal distribution of biomass
across trophic levels (Neutel et al., 2007), the stabilising effect of predator–prey interac-
80 tions (De Angelis, 1975; Allesina and Pascual, 2008) and spatial colonization–extinction
dynamics (Gravel et al., 2011). Interaction strengths are widely accepted to be very
82 important, but the consequences for theoretical structure–stability relationships are not
well understood.

84 Our aim is to explore the relative importance of some of these hypotheses for stability
despite complexity, and to investigate the consequences of two key assumptions we iden-
86 tify in random matrix models. To do this, we use 21 networks constructed from empirical
data via three independent methods. A combination of numerical simulations, statistical
88 analysis and algebraic calculations is used to assess the important features of these em-
pirically derived networks within the wider class of random models which aim to describe
90 them.

The first assumption of May (1972) concerns the interactions between species. Species
92 were assumed to interact with one another randomly, so that the architecture of the net-
work, i.e. which elements of the community matrix are non-zero, was described by an
94 Erdős–Rényi random graph (Erdős and Rényi, 1960). The strengths assigned to species
interactions, i.e. the weights of the non-zero network links, were normally distributed.

96 To test the assumption of randomly generated interactions and intensities, we quantify
the stability of the empirically derived food webs and compare this to the stability pre-
98 dicted by a random matrix. We use a suite of randomization algorithms with different
assumptions about food web structure and interaction strength distributions. In almost
100 all cases, we find that each empirically derived food web is substantially more stable than
the corresponding randomized food webs. This shows that the assumptions built into the
102 random matrices bias the results towards instability.

The second assumption concerns the intraspecific competition of each species. May (1972)
104 assumed that all species had the same timescale for self-regulation at equilibrium, result-
ing in equal elements on the diagonal of the community matrix. We use simple dynamical
106 systems analysis to show that normalising the self-regulation terms in this way leads to
a model **that is not representative of real communities**. In ecological terms, this
108 requires a highly restrictive and untested assumption about trade-offs between species'
intrinsic growth rates and their interactions with other species. We investigate the conse-
110 quences of this assumption by exploring the effect on matrix stability of different meth-
ods for estimating the self-regulation terms. In contrast to the “random interactions”
112 assumption above, the consequence of assuming equal self-regulation is a strong bias
towards stability.

114 2 Random matrix models

We assume that a given ecosystem comprises n species, and that the population of each
116 species is represented by its biomass density x_i , where $i = 1, \dots, n$. The dynamics of the
system are then represented by a general system of differential equations,

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}),$$

118 where $\mathbf{x} = (x_1, \dots, x_n)$ and \mathbf{f} is a function which depends on \mathbf{x} . We define the Jacobian
matrix in its mathematical sense to mean the matrix describing the linearised dynamics
120 at any given location (Wiggins, 2003) i.e.

$$J_{ij} = \frac{\partial f_i}{\partial x_j}.$$

When this matrix is evaluated at an equilibrium point of the system, i.e. at a point \mathbf{x}^*
122 such that $\mathbf{f}(\mathbf{x}^*) = \mathbf{0}$, it is referred to as a community matrix. This is closely related to, but
distinct from, the matrix of coefficients of species interaction rates used to parameterise,
124 for example, a Lotka–Volterra model (see Sec. 4). Local stability of an equilibrium is
determined by the real part of the leading eigenvalue (i.e. eigenvalue with largest real
126 part) of its associated community matrix: if that real part is negative then the equilibrium
is locally stable. Although stability is a binary on/off property, for stable equilibria we use
128 the term “less stable” to mean “further from stability” (i.e. having a leading eigenvalue
with larger real part), and conversely for “more stable”.

130 May (1972) modelled ecological networks of S species using random $S \times S$ community
matrices, A . Each off-diagonal element a_{ij} of A is set to zero with probability $1 - C$
132 and drawn from a distribution with mean zero and variance σ^2 with probability C . The
element a_{ij} represents the effect of a unit of species j on the rate of increase of species i
134 at equilibrium: if a_{ij} is zero, species j has no direct effect on species i . The parameter
 C is referred to as the connectance. May’s critical insight was to use Wigner’s semicircle
136 theorem and the circular law of Metha (1967) to show that the real parts of the eigenvalues
of this random matrix must all be less than

$$d_0 = \sigma\sqrt{SC}. \tag{1}$$

138 **The variable d_0 is commonly referred to as “complexity” and represents the
relative strength of intraspecific competition needed to stabilise a given food**

140 **web.** If the diagonal elements a_{ii} , representing the effects of intraspecific competition,
are all set to $-d_0$, then all the eigenvalues of the community matrix will have negative
142 real parts and the equilibrium will be stable (see Sec. 4). A matrix with a high value of
 d_0 requires strong intraspecific competition to stabilise it. May (1972) used Eq. (1) to
144 conclude that high connectance (large C), a large number of species (large S) or strong
interactions (high σ) in food webs leads to instability.

146 These results were generalised by Allesina and Tang (2012) to community matrices with a
more specific structure in the interaction terms a_{ij} . This included predator–prey systems
148 in which interactions are beneficial to one species and detrimental to the other (a_{ij} and
 a_{ji} have opposite signs). Tang et al. (2014) extended the stability condition to account
150 for pairwise correlation between a_{ij} and a_{ji} . They showed that, with high probability,
the leading eigenvalue of the community matrix will have real part less than

$$d_0 = \sqrt{SV}(1 + \rho) - E, \quad (2)$$

152 where ρ is the correlation between a_{ij} and a_{ji} and E and V are the mean and variance
of the off-diagonal elements (including zeros). This result, with its revised definition
154 of complexity, indicates that (under the assumptions of random network topology) the
negative pairwise correlation ($\rho < 0$) one might expect to find in a predator–prey system
156 should be a stabilising factor.

3 The relationship between random matrices and real 158 food webs

There are numerous methods of constructing a community matrix from data, many of
160 which rely on body-size data, allometric scaling relationships, bio-energetic models and/or
interaction strength data. Brose et al. (2006), Otto et al. (2007) and Woodward et al.

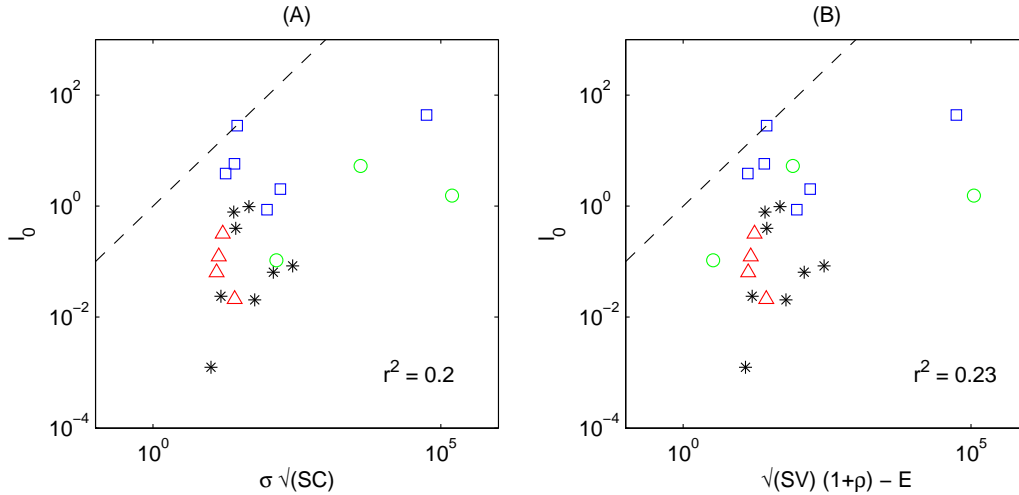


Figure 1: **The stability–complexity relationship in 21 empirically derived food webs.** The real part l_0 of the leading eigenvalue against complexity d_0 defined by: (A) Eq. (1); (B) Eq. (2). S is the number of species, C is the connectance, σ^2 is the variance of non-zero off-diagonal elements, E and V are the mean and variance of all off-diagonal elements and ρ is the correlation coefficient between a_{ij} and a_{ji} pairs. Diagonal elements are set to zero. The dashed lines show the line $l_0 = d_0$. Symbols indicate the method used to construct the community matrix: biomass flux = black stars (successional webs) or red triangles (soil webs); EwE = blue squares; PPMR = green circles. Although complexity is correlated with stability, the relationship is weak and of little practical relevance. Many of the empirically derived networks are many orders of magnitude more stable than predicted by the criteria in Eq. (1) and (2).

162 (2005a) explored the consequences of some of these assumptions and highlighted their
 importance in determining stability. We use three independent, established approaches,
 164 each relying on different sets of assumptions, to construct community matrices from
 empirical data from 21 food webs with distinct topologies. **These consist of: (i)**
 166 **eight successional food webs and four soil food webs constructed using the**
“biomass flux” method of Moore et al. (1996); Neutel et al. (2002); (ii) six
 168 **marine ecosystems modelled using Ecopath with Ecosim (EwE) (Christensen**
and Pauly, 1992); (iii) three freshwater/estuarine ecosystems constructed us-
 170 **ing the predator–prey mass ratio (PPMR) model of Emmerson and Raffaelli**
(2004). See Appendix A for details. The soil and freshwater systems are all food webs
 172 consisting only of predator–prey interactions; the marine systems contain a mixture of
 interaction types.

174 We set the diagonal elements of each empirically derived matrix to zero and calculate
the real part l_0 of the leading eigenvalue as a measure of how far that matrix is from
176 stability (May, 1972; Neutel et al., 2002; Allesina and Tang, 2012). Fig. 1A plots l_0
against complexity as defined by Eq. (1). **Complexity is inversely correlated with**
178 **stability, but only weakly**, and predicts stability within two orders of magnitude at
best. Fig. 1B repeats the plot using the modified stability condition incorporating
180 pairwise correlation in Eq. (2) (Tang et al., 2014). The only matrices for which Eq. (2)
provides a substantially better prediction are the food webs constructed with the PPMR
182 model (green circles), which result in a strong negative pairwise correlation (Emmerson
and Raffaelli, 2004). The construction methods used in the other networks (see Appendix
184 A) result in much weaker pairwise correlation and accounting for it using Eq. (2) has
almost no effect.

186 Fig. 1 raises the important question: given that complexity alone cannot usefully predict
stability in these empirically derived networks, which other network properties can? We
188 test the hypothesis that randomly generated community matrices capture the essential
properties of a real food web by comparing suites of randomly generated matrices to the
190 empirically derived matrices, where each suite is defined by a set of ecologically motivated
rules. We seek algorithms capable of emulating key properties of the empirical networks.

192 For each empirically derived matrix, we generate suites of 200 random matrices for each
of 12 randomization algorithms and calculate the value of l_0 for each randomization. If a
194 particular network property is fundamental to stability, then alterations to the network
that preserve this property should only minimally affect stability. Fig. 2 shows the results
196 of this process for one empirically derived community matrix. Every randomized matrix
has the same size and connectance as the empirically derived matrix. The red line shows
198 the real part of the leading eigenvalue of the empirically derived matrix, $l_0 = 0.1$. The
histograms show the distribution of l_0 generated by 200 realisations of each randomization
200 algorithm.

Fig. 2A shows the distribution for the random matrix proposed by May (1972): the network topology is Erdős–Rényi; the non-zero entries are drawn from a normal distribution with zero mean and the same variance as in the original matrix. The difference is striking: in the random matrices, the strength of intraspecific competition needed to stabilise the food web, l_0 , is between 100 and 10,000 times greater than in the original matrix. Fig. 2B shows the distribution of l_0 for a network as in A, **but with the same pairwise sign structure (i.e. the same number of predator–prey and competitive interactions) as the original matrix.** These interaction pairs are randomly placed in the matrix and the positive and negative elements are randomly and independently generated from half-normal distributions with the same mean as in the original matrix. This approach is similar to that of Pimm and Lawton (1978) and Allesina and Tang (2012). This randomization still gives values of l_0 approximately 100 times larger than in the original matrix. Fig. 2C uses the same algorithm as B to generate the interaction strengths, but preserves the empirical network topology rather than generating a topology at random. This is similar to an algorithm used by Jacquet et al. (2013). The difference in intraspecific interaction strength needed to stabilise the webs, l_0 , between the randomizations and the empirically derived matrix is still large.

The preceding randomization algorithms all randomly generate non-zero elements of the community matrix from normal distributions and give values of l_0 that are consistently more than 100 times greater than in the original matrix. Given this failure of randomly generated elements to capture the properties of the empirically derived matrices, we test a second category of algorithms that permute the actual matrix elements, rather than randomly generating new elements. Fig. 2E shows the distribution of l_0 from an algorithm which moves existing pairs of interactions, i.e. which destroys the topology of the original network, but holds (a_{ij}, a_{ji}) pairs together and preserves pairwise correlation ρ . This algorithm gives a marked improvement over Fig. 2A–C in retaining the original matrix properties: for this network, the value of l_0 for the empirically derived matrix is now within the interquartile range of the distribution of l_0 under the randomization. In

Fig. 2F, pairs of elements (a_{ij}, a_{ji}) are swapped with other existing pairs of elements, thus preserving the topology of the network. This additional constraint does very little to bring the stability of randomized matrices closer to that of the empirically derived matrices on which they are based. Finally, Fig. 2G applies a permutation of the positive elements and an independent permutation of the negative elements. This preserves the original topology and sign structure, but destroys any pairwise correlation. Again, this modification gives very little change relative to algorithms E and F. Algorithms F and G are similar to those used by Yodzis (1981); De Ruiter et al. (1995); Neutel et al. (2002).

Algorithms E–G, which use the original matrix elements rather than replacing them with random numbers, show a marked improvement in preserving the stability of the original matrix. **This shows that the methods used to estimate the community matrices produce off-diagonal elements that yield relatively stable communities, even if these elements are permuted and the topological structure of the original empirical networks is destroyed.** The observation that there is little difference between E, F and G shows that network topology (preserved in F and G) and pairwise correlation (preserved in E and F) are less important, in this particular community, than the distribution of matrix elements.

We repeated the above analysis using all 21 empirically derived community matrices (see Appendix A). We define the error for a random matrix to be $\log_{10}(l_{\text{rand}}/l_{\text{emp}})$, where l_{rand} and l_{emp} are the values of l_0 in the randomized and empirically derived matrix respectively. Fig. 3 shows the mean and the 5th to 95th percentile range of the error for each randomization algorithm, applied to each of the 21 food webs. When these intervals exclude zero, the value of l_0 for the empirically derived matrix lies in the tail of the distribution of l_0 under the randomization scheme, showing that the scheme does not capture the relevant properties of the original matrix. The results in Fig. 3 confirm that the patterns seen in Fig. 2 for a single successional soil food web extend across a range of food webs. For the randomization algorithms that use normally distributed elements (A–

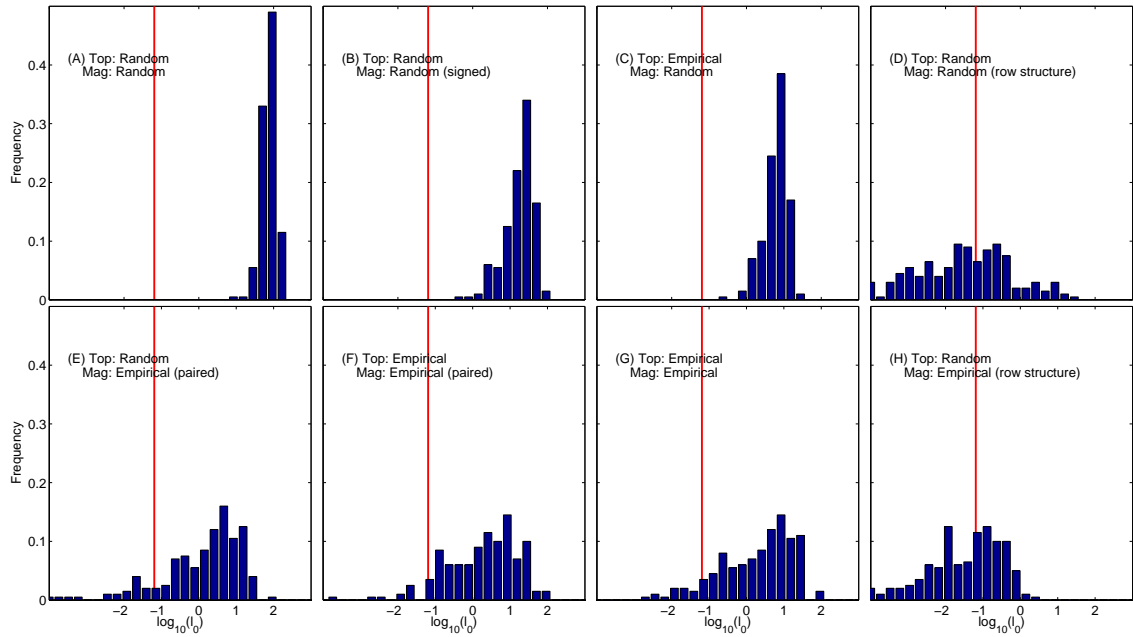


Figure 2: **Randomized matrices are statistically many orders of magnitude less stable than the empirically derived matrix on which they are based.** Graphs show the distribution of the strength of intraspecific competition needed to stabilise the web, l_0 (the real part of the leading eigenvalue), for six different randomization algorithms, changing either or both of network topology (Top) and magnitude of interaction strengths (Mag), applied to an empirically derived community matrix (a successional soil food web). The red line shows the value of l_0 for the empirically derived matrix. Randomizations that use a normal distribution to generate the non-zero elements (A–C) are orders of magnitude less stable than the empirically derived matrices; randomizations that move or swap the actual matrix elements (E–G) are closer, but still show a consistent bias towards instability; randomizations that preserve row structure (D and H) better reflect the stability of the empirically derived matrices. See Appendix B for details of the randomization algorithms.

256 C), the predictions are particularly biased, showing that the random matrices are much
further from stability than the empirically derived matrices. The distributions of l_0 for
258 these algorithms are relatively narrow, showing the similarity of all matrices generated by
these algorithms. Importantly, all the randomization algorithms tested are consistently
260 biased towards instability (error > 0) for the majority of matrices.

As seen in Fig 2, using the same distribution of elements (Fig. 3E–G) as in the empirically
262 derived matrix helps to stabilise the randomized networks. Preserving either the empirical
network topology (Fig. 3F, G) or pairwise correlation (Fig. 3E, F) does not further
264 improve the results for most networks. The notable exceptions to this are the food webs
constructed with the PPMR model (green circles), which have strong pairwise correlation
266 and therefore respond well to algorithms that preserve this feature. This is consistent
with the results of Tang et al. (2014), which also used PPMR in the construction methods.

268 The structure of predator–prey networks has been found to play an important role in
stability (Levins, 1979; Dambacher et al., 2003; Allesina and Tang, 2012). A pyramidal
270 biomass pattern leads to strong row structure in community matrices: the variance of
elements within a row is much less than the variance of elements between rows (Jacquet
272 et al., 2013). This is borne out by the definition of an element of the community matrix:
 a_{ij} is the effect of a unit of species j on the rate of increase of species i at equilibrium.
274 The rate of increase of species i is typically proportional to its abundance (see Sec. 4),
meaning that the magnitude of elements in row i of the community matrix should be
276 strongly correlated with the equilibrium abundance of species i .

To test the role of row structure, we designed an algorithm that preserves the average
278 magnitude of the elements in each row of the matrix. The positive/negative elements in
each row are generated from normal distributions with the same means as in the original
280 matrix (Figs. 2D, 3D). **To investigate the role of interaction strength distribution
in tandem with row structure, we designed a second algorithm that moves el-
282 ements of the original matrix within rows in the lower triangle and moves the**

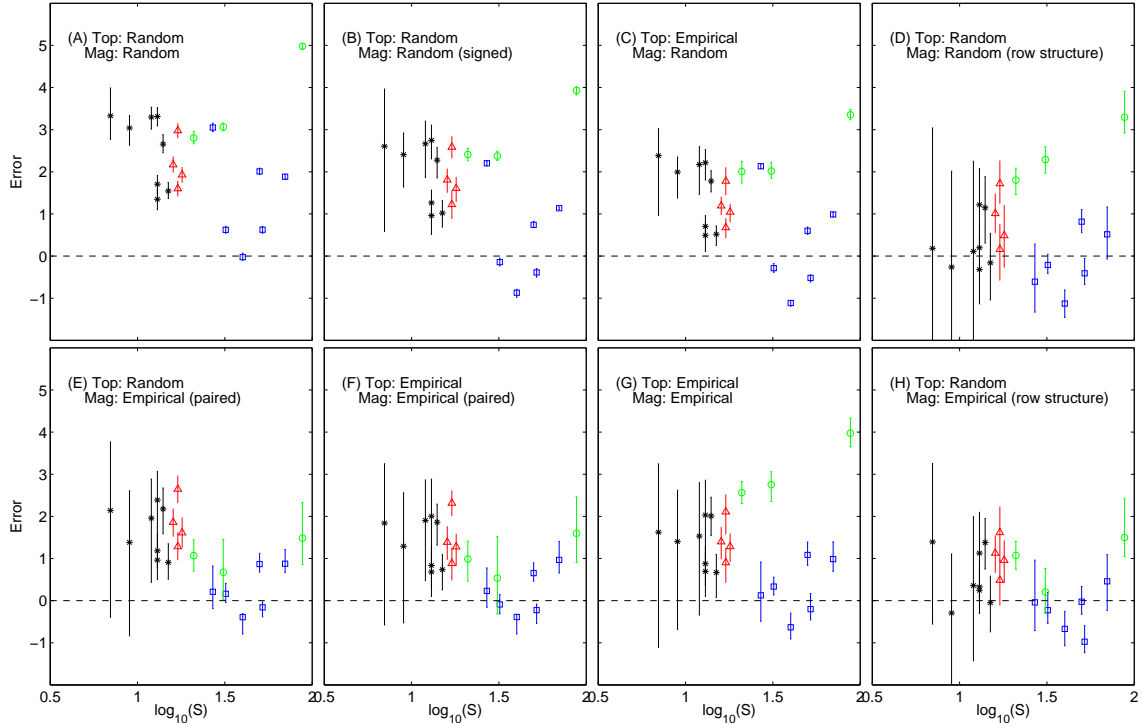


Figure 3: **Differences in stability between empirically derived and randomized matrices across 21 communities and three construction methods.** The distribution of error = $\log_{10}(l_{\text{rand}}/l_{\text{emp}})$ for the same six randomization algorithms (A–H) as in Fig. 2, changing either or both of network topology (Top) and magnitude of interaction strengths (Mag). The food webs are plotted against the number of species S in the web on the horizontal axis. See Appendix B for details of the randomization algorithms. Symbols indicate the method used to construct the community matrix: biomass flux = black stars (successional webs) or red triangles (soil webs); EwE = blue squares; PPMR = green circles.

corresponding elements within columns in the upper triangle, holding pairs
284 of elements (a_{ij}, a_{ji}) together (Figs. 2H, 3H) (see Appendix B for details). Both
these algorithms destroy the original network topology, but preserve stability remark-
286 ably well in most food webs. Using randomly generated elements (Fig. 3D)
destabilises community matrices constructed using the PPMR model (green
288 circles), because it destroys the strong pairwise correlation in these matrices;
the algorithm that preserves pairwise correlation (Fig. 3H) preserves stabil-
290 ity much better for these matrices. For the other food webs, there is little
difference between Figs. 3D and H, indicating that the row structure itself is
292 more important than the precise distribution of interaction strengths.

Cycles of length three (interaction chains from species i to species j to species
294 k and back to species i), which typically arise from omnivorous interactions
in food webs, can also affect stability. Neutel et al. (2002) showed that strong
296 top-down effects in omnivorous relations tend to be spread across different
cycles, meaning that the maximum cycle weight (where cycle weight is defined
298 as the geometric mean of the strengths of the interactions in the cycle) tends
to be lower in a real food web than in a randomized community matrix. Since
300 cycle weights are closely linked to eigenvalues (Hofbauer and Sigmund, 1998),
they proposed that this helps make food webs stable. To test the role of cycles of
302 in our empirically derived matrices, we designed simple algorithms that only change links
that are not in a cycle of length three, either destroying or preserving the original network
304 topology. Both algorithms preserve the number of cycles of length three, their weights
and the pairwise correlation of elements (see Appendix B for details). These are severe
306 constraints: in most networks, only 25% to 35% of links are changed by these algorithms.
To enable a fair comparison with randomization algorithms that do not preserve cycles of
308 length three, we devised corresponding algorithms that change the same number of links
as the cycle-preserving algorithms, but choose the links at random rather than because
310 of their involvement in a cycle of length three. **For the topology-changing algorithm**

(Fig. 4A), there is little difference in the mean error between the algorithm
312 that preserves cycles of length three and the algorithm that does not. For the
topology-preserving algorithm (Fig. 4B), preserving cycles of length three
314 reduces error (see Supplementary Fig. S2 for the error distributions for each
algorithm). This shows that stability is promoted by the way in which cycles
316 are positioned within the overall network topology (which is only preserved
in Fig. 4B), rather than simply by their weights (which are preserved in both
318 Fig. 4A and B).

Of the randomization algorithms investigated, the cycle-preserving, topology-preserving
320 algorithm (Fig. 4B) is least biased, but is highly restrictive, requiring information about
the network topology, cycle structure and distribution of matrix elements. In contrast, the
322 algorithm shown in Fig. 3D only requires the means of the community matrix elements
for each species and has purely random topology. This corresponds to May-type assump-
324 tions combined with realistic sign structure (Allesina and Tang, 2012) and equilibrium
biomasses for each species.

326 4 Random matrices and dynamic models

The results in Sec. 3 use the properties of general community matrices to indicate how
328 stability may arise in complex networks. However, to truly understand the ecological
implications of the results for random matrix models, it is helpful to consider a specific
330 dynamic model rather than just a community matrix linearised around a hypothetical
equilibrium point.

332 Most random matrix models, including those considered in Sec. 3, assume that all di-
agonal elements of the community matrix are the same, i.e. $a_{ii} = -d$ for some $d > 0$.
334 This was justified by May (1972) with the statement “to set a time-scale, these damp-
ing times are all chosen to be unity” (i.e. $d = 1$). However, introducing the variability

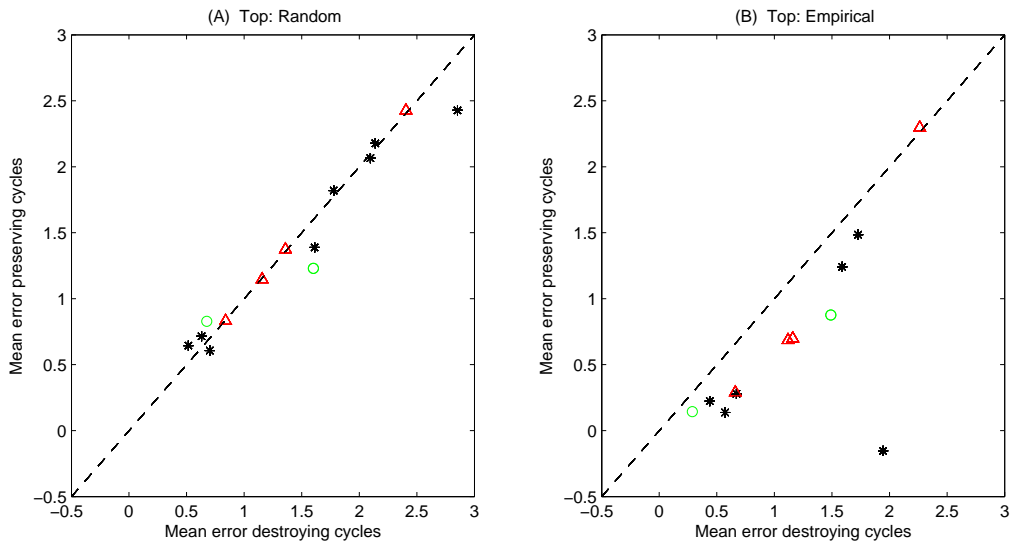


Figure 4: **Stability is enhanced by the position of cycles within the overall network topology.** The mean error = $\log_{10}(l_{\text{rand}}/l_{\text{emp}})$ of randomization algorithms that preserve cycles of length three plotted against the mean error of corresponding algorithms that change the same number of network links but do not preserve cycles. Randomizations in A and B destroy and preserve the original network topology respectively. See Appendix B for details of the randomization algorithms. Symbols indicate the method used to construct the community matrix: biomass flux = black stars (successional webs) or red triangles (soil webs); PPMR = green circles. Community matrices for which $< 90\%$ of randomizations produce distinct matrices are not shown; this includes all the EwE webs.

336 into the diagonal elements without changing their mean tends to increase the leading
 eigenvalue and this effect cannot be removed by a rescaling of variables (Haydon, 1994).
 338 Furthermore, the correspondence of diagonal elements with intraspecific effects is par-
 ticular to simple generalised Lotka–Volterra models. More general models, for example
 340 with type-II functional responses or other nonlinear interaction terms, do not have this
 property and generate diagonal elements that depend on interspecific interactions (De
 342 Angelis, 1975; Haydon, 1994). Therefore, some variation in the diagonal elements of the
 community matrix should be expected. The assumption of identical diagonal elements
 344 is crucial yet ecologically unjustified. In this section, we show that a dynamic model
 that is constrained to have equal diagonal elements in the community matrix behaves
 346 fundamentally differently from an unconstrained model and cannot exhibit one of the
 most common mechanisms for species loss.

348 Although the results of May (1972) did not rely on a specific dynamic model, the simplest
 model that generates a community matrix of the type considered by May (1972) and
 350 Allesina and Tang (2012) is a generalised Lotka–Volterra model

$$\frac{dx_i}{dt} = x_i \left(r_i + \sum_{j=1}^S q_{ij} x_j \right), \quad (3)$$

where x_i is the biomass density of species i , r_i is its intrinsic growth rate and the q_{ij}
 352 are referred to as Lotka–Volterra coefficients. This model assumes species interactions
 can be described by “mass action” (type-I) terms. The coefficients q_{ij} correspond to the
 354 dynamic index measure of interaction strength *sensu* Berlow et al. (1999) and can be
 estimated empirically from predator–prey mesocosms (Emmerson and Raffaelli, 2004).
 356 The diagonal element q_{ii} represents the strength of intraspecific competition, or self-
 limitation, for species i and must be non-positive to prevent boundless growth of species
 358 i in isolation. Although Eq. (3) is the simplest model of an ecological network, the
 conclusions of this section extend to more general models (see Appendix C).

360 The community matrix for Eq. (3) has elements $a_{ij} = q_{ij}x_i^*$. This helps to explain the
row weight patterns seen in the empirically derived matrices in Sec. 3: the elements in
362 row i of the community matrix are proportional to the equilibrium biomass density x_i^* of
species i . This emphasises that the community matrix element a_{ij} does *not* represent the
364 direct effect of species j on species i (which is q_{ij}) and a_{ii} does *not* represent the strength
of intraspecific competition (which is q_{ii}).

366 The system in Eq. (3) undergoes a transcritical bifurcation whenever one of the equilib-
rium abundances x_i^* becomes negative. During this transition, the abundance of species i
368 gradually declines to zero and the community moves smoothly to a different equilibrium,
in which species i is absent. Instability and species loss are thus associated with the
370 gradual decline of one species to zero density.

At a transcritical bifurcation, the equilibrium abundance x_i^* for species i is zero, and
372 so the diagonal element of the community matrix for species i , $a_{ii} = q_{ii}x_i^*$, is also zero.
Requiring all the diagonal elements of the community matrix to equal $-d < 0$ prevents
374 this type of transition from occurring. It also imposes a constraint on the intrinsic growth
rates r_i (see Appendix C):

$$r_i = d \sum_{j=1}^S \frac{q_{ij}}{q_{jj}}. \quad (4)$$

376 This constraint means that the equilibrium biomasses are $x_i^* = -d/q_{ii}$, which are always
all positive since $d > 0$ and $q_{ii} < 0$. Therefore, the only way in which a species can go
378 extinct is via a degenerate bifurcation that makes the positive equilibrium become unsta-
ble, causing the system to move suddenly to a different equilibrium point (see Appendix
380 C). Although sudden changes in community composition, such as regime shifts (Moell-
mann and Diekmann, 2012), are certainly possible, a model that precludes species loss via
382 gradual decline is unable to capture one of the simplest and most common mechanisms
for change in ecological communities (Rossberg, 2013). Hopf bifurcations, which lead to
384 oscillatory dynamics, are also possible, but are not directly associated with species loss.

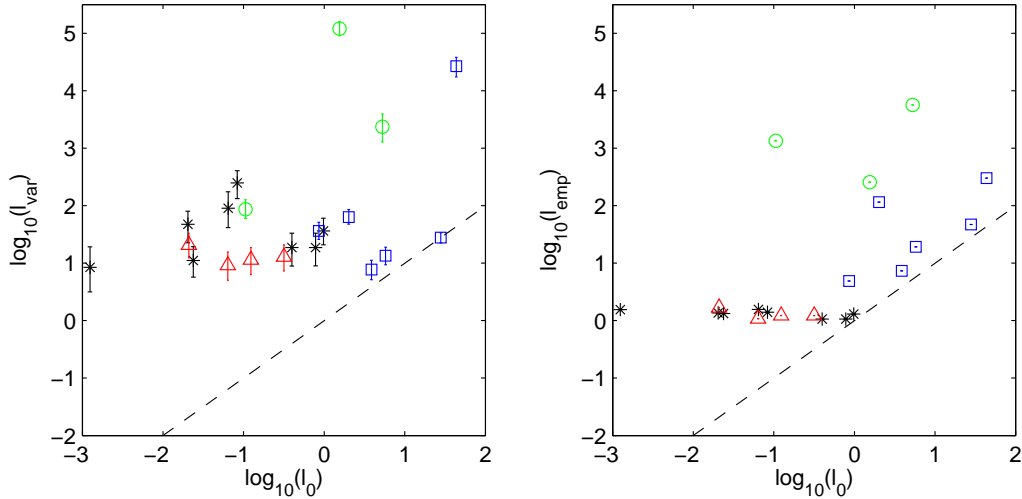


Figure 5: **Adding variability to the diagonals of empirically derived community matrices decreases stability by orders of magnitude.** The real part of the leading eigenvalue, l_0 , of the matrix with diagonal elements set to zero against: (a) l_{var} the real part of the leading eigenvalue of the matrix with randomly generated diagonal elements with the same variance as the off-diagonals; (b) l_{emp} the real part of the leading eigenvalue of the matrix with empirically derived estimates for the diagonal elements. For each matrix with variable diagonal elements, a constant was added to all the diagonal elements so that their sum was zero. Symbols indicate the method used to construct the community matrix: biomass flux = black stars (successional webs) or red triangles (soil webs); EwE = blue squares; PPMR = green circles.

Eq. (4) can be interpreted as representing an ecological trade-off. A species that receives
 386 a net benefit from interactions with other species must compensate by having a negative
 intrinsic growth rate. For example, a top predator benefits from consuming prey
 388 ($q_{ij} \geq 0$), but will die out in the absence of prey ($r_i < 0$). Conversely, a species that is
 negatively impacted from interactions with other species (for example a basal resource,
 390 or a species that competes with other species) must compensate by having a positive
 intrinsic growth rate. Although the idea that such trade-offs may operate in specific
 392 ecological circumstances is easy to argue, it is implausible that the intrinsic growth rates
 will be precisely tuned to satisfy Eq. (4). Any slight deviation from Eq. (4) results in a
 394 shift in model behaviour, meaning that the assumption of equal diagonal elements makes
 the model structurally unstable.

396 In the limit as $S \rightarrow \infty$, variation among diagonal elements has a negligible effect on

the leading eigenvalue compared to the effect of the off-diagonal elements (Allesina and
398 Tang, 2012). However, in moderately sized networks, variation in diagonal elements can
affect stability. To quantify the effect of variability in the diagonal elements of the 21
400 empirically derived community matrices, we added variability in one of two ways: firstly
using randomly generated diagonal elements with the same variance as the off-diagonal
402 elements, and secondly using the diagonal elements estimated from the empirical food
web data. To enable a meaningful quantitative comparison with the zero-diagonal case,
404 the diagonal elements were shifted so that their mean was zero.

We compared the real part of the leading eigenvalue of the matrix with diagonal variabil-
406 ity to that of the matrix with diagonal elements set to zero (l_0) (Fig. 5). For all 21 food
webs, using randomly generated diagonal elements (Fig. 5A) gives a leading eigenvalue
408 with substantially larger real part than using equal diagonal elements. Using the empiri-
cally derived diagonal elements also gives a leading eigenvalue with larger real part in the
410 majority of food webs (Fig. 5B). For one of the three community matrix construction
methods (that used for the soil food webs; black/red), the empirically derived diago-
412 nal elements dominate the matrix so strongly that all the matrices have approximately
the same leading eigenvalue when they are included. This shows that the standard as-
414 sumption of equal diagonal elements consistently gives predictions that overestimate the
stability of the equilibrium.

416 5 Discussion

In the absence of detailed ecological data on species interactions, the use of random
418 matrix models to make predictions about the relationship between ecosystem complexity
and stability is widespread (Thébault and Fontaine, 2010; Gravel et al., 2011; Allesina
420 and Tang, 2012). We have shown that using randomly generated network topology and
interaction strengths can lead to predictions about local stability that differ by orders of

422 magnitude from those of empirically derived models.

Various characteristics of real food webs, for example the sign structure associated with
424 predator–prey interactions (Allesina and Pascual, 2008; Jacquet et al., 2013), the network
topology (Haydon, 2000; Tylianakis et al., 2010), the distribution (McCann et al., 1998;
426 Emmerson and Raffaelli, 2004; Jacquet et al., 2013) and the relative positioning (Yodzis,
1981; De Ruiter et al., 1995; Neutel et al., 2002) of interaction strengths, have been sug-
428 gested to have stabilising influences. Our results show that including such characteristics
in models can increase stability but, even for larger networks where one would expect
430 predictions based on random matrices to give the best results, this is not always the case.

**Measuring species interaction coefficients directly in complex food webs is
432 impossible. The models themselves are necessarily simplifications of the real
ecology, ignoring, for example, details of age- and size-structure. They may
434 also amalgamate species into perceived functional groups, which reduces the
apparent species richness. Further assumptions and simplifications are nec-
436 essary to convert empirical data into estimated community matrices. One of
the strengths of our study is that the empirical matrices originate from three
438 independently established sets of modelling assumptions: biomass flux model
(Moore et al., 1996), Ecopath with Ecosim (Christensen and Pauly, 1992) and
440 the predator–prey mass ratio model (Emmerson and Raffaelli, 2004). None
of these approaches is “correct” and each has arisen from ecological and data-
442 driven constraints specific to the systems under study. Nevertheless, the re-
sulting community matrices represent our best estimates of ecological reality.
444 Where consistent patterns emerge from our analysis of empirical networks de-
rived in different ways, this provides evidence that these ecosystems contain
446 structure that is not captured by simple random matrix models.**

Our study shows that network topology may be less important for community stability
448 than widely thought. For example, Tylianakis et al. (2010) summarised the attributes

of network topology thought to confer stability. Our results in Fig. 3C, F, G show that
450 it is possible to have two networks with identical topology but with key properties that
vary by up to four orders of magnitude. Conversely, our results in Fig. 3D show that
452 a random network topology, with a particular organisation of interaction strengths, can
come closer to empirical data. **These results do not imply that topology has no
454 effect, but they do show that it is less important than the distribution of
interaction strengths.**

456 We designed new randomization methods to quantify the role of two important features
of real ecological networks, namely cycles of length three and row structure in the commu-
458 nity matrix, in determining stability. We have shown that randomizations that preserve
either of these properties better reflect the stability of the empirically derived network,
460 although for cycles of length three the results are less marked than for row structure.
That each of these properties involves a combination of network topology and the sizes
462 of the community matrix elements emphasises our finding that results based on random-
izations which ignore, or make unjustified assumptions about, either of these features do
464 not usefully reflect ecological reality.

Previous studies have focused on different properties of community matrices. Neutel
466 et al. (2007) saw a correlation between cycle weights and stability in data from soil
food webs; Jacquet et al. (2013) found that removing trophic structure in community
468 matrices from EwE models adversely affected stability; Tang et al. (2014) highlighted the
role of pairwise correlation in community matrices constructed using allometric scaling
470 laws based on body mass (Brose et al., 2006; Reuman et al., 2009; Pawar et al., 2012).
Our study unifies these seemingly disparate findings by including community matrices
472 constructed using each of these three methods and showing that different factors are
more important for stability in these different types of community. For instance, pairwise
474 correlation is important in communities where it is strong, but other factors, most notably
row structure, dominate in communities where it is weaker.

476 Based on empirical and theoretical considerations, Rossberg (2013) argued that transcrit-
478 ical bifurcations, involving gradual decline of a species to zero density, are the dominant
480 form of loss of local stability in ecological communities. Community matrices where the
482 diagonal elements are assumed to be equal prevent this type of change from happening
484 and cannot, therefore, characterize the main type of instability that leads to change in
community structure. The assumption of May (1972), that variations in the strength
of intraspecific competition are unimportant, is one that has been largely neglected (see
however De Angelis, 1975; Haydon, 1994; De Ruiter et al., 1995; Haydon, 2000; Neutel
et al., 2002). The effect of this assumption is at least as prominent as the assumptions
concerning off-diagonal community elements.

486 Technical detail and exhaustive testing are unavoidable ingredients in the preceding anal-
488 ysis, but a clear and practical biological message emerges. Although there is a relationship
490 between “stability” and “complexity” in ecological communities, the predictive power of
492 this relationship is weak. Honest and seemingly pragmatic attempts to replace ignorance
494 of ecological detail with random numbers, whether these relate to network structure or to
interactions between species, must be treated with extreme caution. Despite the limita-
tions explained here in the context of local stability, random matrix models may be useful
in the context of more general co-existence conditions, where they can lead to quantita-
tive predictions of community structure in good agreement with observations (Meszena
et al., 2006; Rossberg, 2013).

496 On a further constructive note, we argue that using explicit dynamic models to describe
ecological networks is preferable to directly assigning elements to a community matrix
498 without reference to the underlying population dynamics. Parameterising a dynamic
model requires additional data, for example species’ intrinsic growth rates or equilibrium
500 biomass densities, as well as estimates of interaction strengths. Collecting these data is
not always practical. Nevertheless, estimating them via established models (De Ruiter
502 et al., 1995; Emmerson and Raffaelli, 2004; Datta et al., 2010), or investigating the effects

of introducing variability in them, is preferable to simply assuming that they satisfy an
504 arbitrary set of mathematical constraints.

Acknowledgements

506 We thank Chris Hughes and Leonid Pastur for mathematical context and precision, Sarah
Collings, Richard Law and Anje-Margriet Neutel for discussions and comments, Stefano
508 Allesina for critical input and for sharing pre-prints of his publication with us, and Sebas-
tian Diehl for insightful reviews. A.G.R. and J.B. were supported by the U.K. Depart-
510 ment of Environment, Food and Rural Affairs (M1228), and the European Commission
(agreement no. 308392, DEVOTES).

512 Online Appendices

A Empirically derived community matrices

514 The 21 empirically derived community matrices are available in electronic form. This section describes the methods used to obtain these data for each class of food web studied.

516 Biomass flux method

Eight successional soil food webs and four soil food webs were supplied by Anje-Margriet Neutel from the data published in De Ruiter et al. (1993); Moore et al. (1993); De Ruiter et al. (1995); Moore et al. (1996); Neutel et al. (2002). **The elements of the community matrix were derived from a generalised Lotka–Volterra type model of the same form as Eq. (3) at equilibrium. For primary producers, $r_i > 0$ was the intrinsic rate of increase per year. For consumers, $r_i < 0$ was the non-predatory death rate per year. For a species i that is consumed by species j , the Lotka–Volterra coefficient q_{ij} was set equal to the negative of the consumption coefficient (in units of $\text{g}^{-1} \text{m}^2 \text{yr}^{-1}$) c_{ij} ; the Lotka–Volterra coefficient q_{ji} was set equal to the assimilation efficiency of species j times the production efficiency of species j times the consumption coefficient c_{ij} . Species were aggregated into functional groups of species with similar food sources. Intrinsic birth and death rates and efficiencies were estimated from microcosm studies. Trophic interactions among taxa were established by direct observation or gut content analysis (Moore et al., 1996). Consumption coefficients were estimated from the measured biomass flux from prey to predator (Moore et al., 1993). The original data included a row and a column corresponding to detritus; these were removed for this analysis.**

In the eight successional food webs, the average number of species in the webs was 12
 536 (range 7–15), mean connectance was 0.29 (range 0.28 to 0.34) and average pairwise cor-
 relation ρ (**i.e. correlation between a_{ij} and a_{ji}**) was -0.062 (range -0.02 to -0.13).
 538 In the four soil food webs, the average number of species in these webs was 17 (range
 16–18), mean connectance was 0.27 (range 0.23 to 0.31) and average pairwise correlation
 540 ρ was -0.055 (range -0.047 to -0.059). All the networks were strictly predator–prey
 (i.e. a_{ij} and a_{ji} either had opposite signs or were both zero).

542 **Predator–prey mass ratio method**

Data were obtained for the average adult body mass w_i of species i and which species
 544 predated on which other species for the well documented Ythan Estuary (Emmerson and
 Raffaelli, 2004), Broadstone Stream (Woodward et al., 2005b) and Tuesday Lake 1984
 546 (Jonsson et al., 2005) food webs. In the case of Tuesday Lake, 56 species were aggregated
 into 21 functional groups with identical links (trophic species). We used the average
 548 body mass of all species in a functional group as the body mass for that trophic species.
 Species with no observed predator or prey links were discarded.

550 The model of Emmerson and Raffaelli (2004) was used to estimate elements of the com-
 munity matrix for both these food webs. For each predator–prey pair, the Lotka–Volterra
 552 coefficient q_{ij} was assumed to be a power-law function of the ratio of predator body size
 w_j to prey body size w_i :

$$q_{ij} = -q_0 \left(\frac{w_j}{w_i} \right)^{q_1} .$$

554 The corresponding element q_{ji} is opposite in sign and reduced in magnitude by a factor
 ϵ_j representing the predator’s feeding efficiency: $q_{ji} = -\epsilon_j q_{ij}$. The equilibrium biomass
 556 x_i^* of species i was estimated to be

$$x_i^* = x_0 w_i^{q_2} ,$$

The community matrix elements were then calculated via $a_{ij} = q_{ij}x_i^*$. Parameter values:
558 $q_0 = 7 \times 10^{-4}$, $x_0 = 95.92$, $q_1 = 0.66$, $q_2 = -0.1836$. A fixed efficiency of $\epsilon = 0.1$ was used
for all species.

560 The average number of species in these webs was 47 (range 21–88), mean connectance
was 0.24 (range 0.11 to 0.32) and average pairwise correlation ρ was -0.75 (range -0.28
562 to -0.98).

Ecopath with Ecosim

564 The Ecosim community matrices were constructed in two steps. First, an Ecopath (Chris-
tensen and Pauly, 1992) model was set up, representing a balanced account of biomass
566 flows through the ecosystem, constrained by empirical data on abundances, metabolic
rates and feeding preferences. From such a static Ecopath model, a dynamic Ecosim
568 model was derived by modelling energy flows as outcomes of population-dynamic pro-
cesses (feeding, respiration, mortality). This led to expressions for $\mathbf{f}(\mathbf{x})$, from which the
570 community matrix can be computed. The time derivatives $\mathbf{f}(\mathbf{x})$ were given by the derivt
function of Ecosim. Numerically differentiating the output of this function with respect
572 to the biomass x_j of species j gives the j^{th} column of the community matrix.

The six marine food webs were: (i) the Tampa Bay model, which is a subset consisting
574 of 52 groups of the Gulf of Mexico model (Walters et al., 2006), (ii) the Georgia Strait
(British Columbia) model (Dalsgaard et al., 1998), a model consisting of 27 functional
576 groups, (iii) the Caribbean reef model (50 groups Opitz, 1996), (iv) the North East Pacific
model (40 groups Gu enette and Christensen, 2005), (v) the Great Barrier Reef model
578 (32 groups Gribble, 2005) and (vi) the Cefas North Sea model (70 groups Mackinson
and Daskalov, 2007). The Ecopath with Ecosim (EwE) models for datasets (i)–(v) can
580 be downloaded from the University of British Columbia website (sources.ecopath.org –
password required for (i) and (ii) and may be found at www.ecopath.org/models for

582 (iii)–(v)). The model for dataset (vi) is available by contacting one of the Cefas authors.

All of these food webs contain a substantial number of non-predator–prey interactions,
584 where either a_{ij} and a_{ji} have the same sign, or one of a_{ij} and a_{ji} is zero and the other
is non-zero. The Tampa Bay and North Sea models contain multi-stanza (i.e. age-
586 structured) groups for the same species. **Each stanza is a separate node and a part
of the population of stanza m of species n flows into stanza $m + 1$ of species n ,
588 creating a $(+, 0)$ type link. Thus, it is not possible to order the species so that
all elements in the lower matrix triangle are positive and all elements in the
590 upper matrix triangle are negative. We therefore designed the randomization
algorithms (see below) so that they can be applied to any community matrix,
592 regardless of sign structure.**

The average number of species in these webs was 45 (range 27–70), mean connectance was
594 0.56 (range 0.34 to 0.70) and average pairwise correlation ρ was -0.057 (range -0.0097
to -0.28).

596 Stability criteria

The values of d_0 in the stability criteria of Eq. (1) and Eq. (2) were calculated for each
598 empirically derived community matrix. In these equations: S is the number of species in
the matrix; C is the connectance (i.e. the number of links divided by $S(S - 1)$); σ is the
600 standard deviation of the non-zero off-diagonal matrix elements; E and V are the mean
and variance respectively of all off-diagonal elements; ρ is the pairwise correlation (Tang
602 et al., 2014)

$$\rho = \frac{E(a_{ij}a_{ji}) - E^2}{V}.$$

The stability criterion in Eq. (1) is only valid for matrices where $\sqrt{SV}(1 + \rho) > SE$
604 (Tang et al., 2014). This was checked and found to be true for all our empirically derived

community matrices.

B Randomization algorithms

These algorithms randomize an $S \times S$ matrix $A = [a_{ij}]$ to create an $S \times S$ matrix $B = [b_{ij}]$. N_l is the number of non-zero off-diagonal elements and $C = N_l/(S(S - 1))$ is the fraction of off-diagonal elements that are non-zero. The mean and standard deviation of the non-zero elements are μ and σ respectively. The mean of the positive and negative elements are μ_+, μ_- . The number of elements involved in a cycle of length three is N_c (in cases where links are not bidirectional, i.e. only one of a_{ij}, a_{ji} is non-zero, an element is considered to be in a cycle regardless of the signs of the element); the number of non-zero elements not involved in a cycle of length three is $N_0 = N_l - N_c$.

Off-diagonal algorithms

All the off-diagonal algorithms hold the diagonal elements zero. Any randomization that resulted in a species having no interactions (i.e. every entry in a row or column was zero) or contained no cycles was rejected as these randomizations have a zero leading eigenvalue. Algorithms used in Fig. 2 and 3 are:

A: Random topology; randomly generated entries. Each randomization has N_l entries sampled from a normal distribution $N(0, \sigma^2)$. Entries are positioned randomly.

B: Random topology with sign structure; randomly generated entries. Each randomization has N_l randomly positioned entries. Positive (negative) entries are sampled from the half-normal distribution with mean μ_+ (μ_-). For every element pair (a_{ij}, a_{ji}) (with $i < j$) in the empirically derived matrix with a particular sign structure, i.e. $(+, +)$, $(+, -)$, $(-, 0)$, etc, there is a random pair (b_{kl}, b_{lk}) (with $k < l$) with the same sign structure.

- 628 **C:** Empirical topology; randomly generated entries. An element b_{ij} is non-zero if and
only if the corresponding element in the empirically derived matrix, a_{ij} is non-zero.
630 Furthermore $\text{sign}(a_{ij}) = \text{sign}(b_{ij})$. Positive (negative) entries are sampled from the
half-normal distribution with mean μ_+ (μ_-).
- 632 **D:** Random topology; randomly generated entries preserving row structure. Each ran-
domization has N_l entries placed at random. The randomized matrix contains the
634 same proportion of each pair type $[(+, +), (-, 0), (+, -), \text{etc.}]$ as the original ma-
trix. The mean of the positive [negative] entries of row i of the empirically derived
636 matrix is μ_{i+} [μ_{i-}]. Where there are no entries of that sign in a row, the value of
 $\mu_{i\pm}$ from a populated row (chosen at random) is used instead. Positive [negative]
638 entries in row i of the randomized matrix are sampled from a normal distribution
with mean μ_{i+} [μ_{i-}] and CV 0.2.
- 640 **E:** Random topology; empirical entries (paired). Each element a_{ij} , where $i < j$, is moved
to element b_{kl} , where $k < l$. For every move of a_{ij} to b_{kl} there is a corresponding
642 move of a_{ji} to b_{lk} preserving the pair structure of the empirically derived matrix.
- F:** Empirical topology; empirical entries (paired). Each element pair (a_{ij}, a_{ji}) , where
644 $i < j$, is swapped with an element pair (a_{kl}, a_{lk}) , where $k < l$, which has the same
sign structure, i.e. $(+, +)$, $(-, +)$, $(0, -)$, etc.
- 646 **G:** Empirical topology; empirical entries (not paired). Every positive element is swapped
with another positive element. Every negative element is swapped with another
648 negative element.
- H:** Random topology; empirical entries (constrained within rows/columns). **This algo-**
650 **rithm permutes elements of the empirically derived matrix within rows,**
while preserving (a_{ij}, a_{ji}) pairs. Most of the empirically derived commu-
652 **nity matrices are organised with top predators in the upper rows and**
basal resources in the lower rows. This tends to lead a triangular struc-

654 ture in the matrix, where the lower-left triangle contains predominantly
negative elements and upper-right triangle contains predominantly pos-
656 itive elements. This means that the lower triangle typically contains
elements that have, on average, larger magnitude than the elements in
658 the upper triangle. The row structure of the empirically derived matrices
therefore tends to be stronger in the lower triangle, so we designed
660 the algorithm described above to preserve row structure in the lower
triangle (and therefore column structure in the upper triangle).

662 Each non-zero lower triangle element a_{ij} (where $i > j$) is moved within
the same row to b_{ik} (where $i > k$). To preserve the pair structure of the
664 original matrix, this move of a_{ij} to a_{ik} is accompanied by a corresponding
move of a_{ji} to b_{ki} .

666 We also tested a similar algorithm that preserved row structure in the
upper triangle and column structure in lower triangle; this produced ma-
668 trices with leading eigenvalues further away from the leading eigenvalue
of the empirically derived matrix.

670 The randomization algorithms for Fig. 4 are:

A (vertical axis): Random topology holding cycles of length three; empirical entries
672 (paired). This is a constrained version of randomization D from Fig. 2 and 3. The
 N_c elements that are part of a cycle of length three are fixed (i.e. their position was
674 not changed). The remaining N_0 elements are moved as described in D above.

A (horizontal axis): Random topology holding some entries; empirical entries (paired).
676 This is a constrained version of randomization D from Fig. 2 and 3. N_c elements
are chosen at random. These elements are fixed. The remaining N_0 elements are
678 moved as described in D above.

B (vertical axis): Empirical topology holding cycles of length three; empirical entries
 680 (paired). This is a constrained version of randomization E from Fig. 2 and 3. The
 682 N_c elements that are part of a cycles of length three are fixed. The remaining N_0
 elements are swapped as described in E above.

B (horizontal axis): Empirical topology holding some entries; empirical entries (paired).
 684 This is a constrained version of randomization E from Fig. 2 and 3. N_c elements
 are chosen at random. These elements are fixed. The remaining N_0 elements are
 686 swapped as described in E above.

Supplementary Table S1 gives a summary of the properties preserved by each algorithm
 688 and measures of the bias in d .

Diagonal algorithm

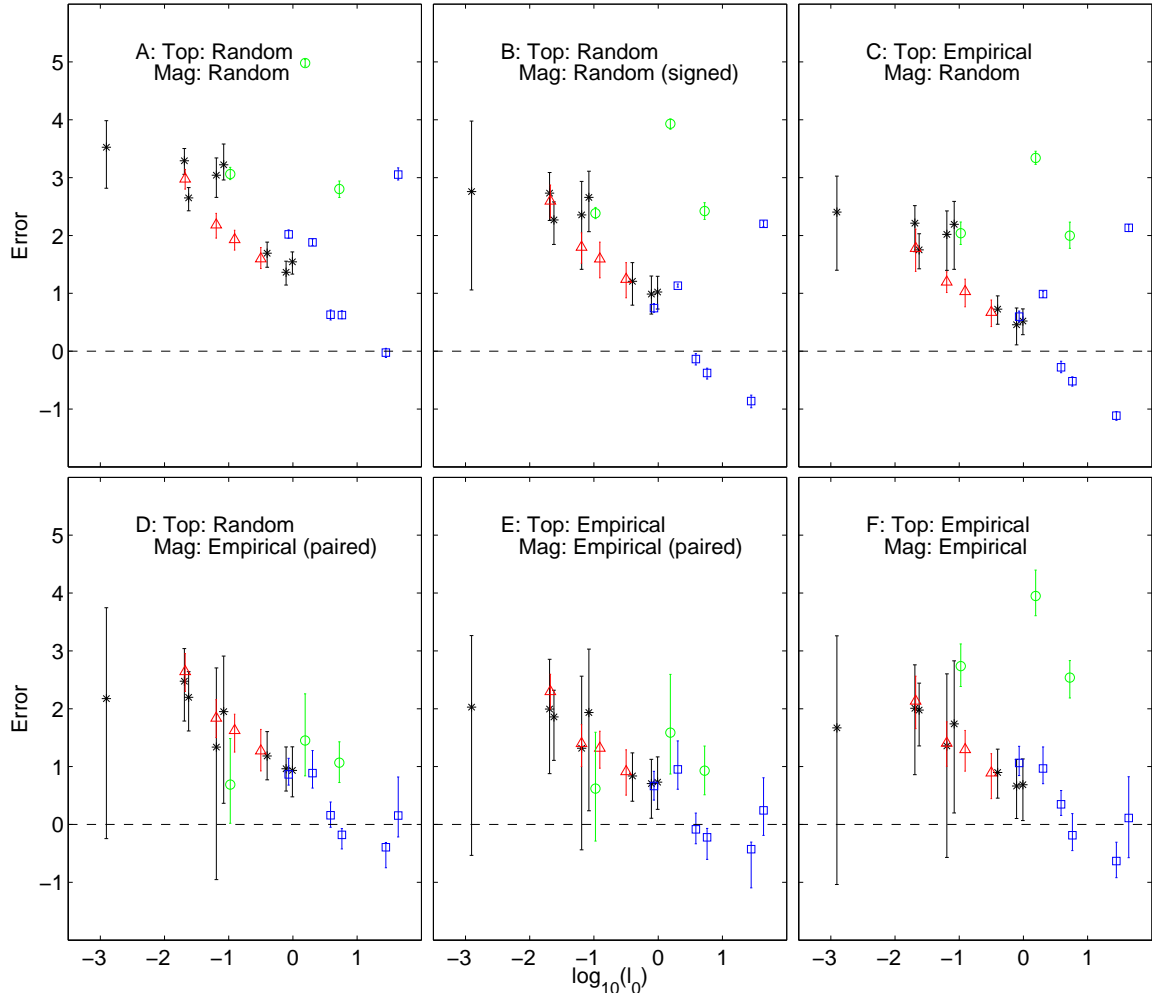
690 In Fig. 5A, the diagonal elements were replaced with normally distributed random num-
 bers with mean zero and variance σ^2 , i.e. the variance of the non-zero off-diagonal
 692 elements.

C Dynamic model bifurcation analysis

694 Consider the generalised Lotka–Volterra model

$$\frac{dx_i}{dt} = x_i \left(r_i + \sum_{j=1}^S q_{ij} x_j \right), \quad i = 1, \dots, S, \quad (\text{C.1})$$

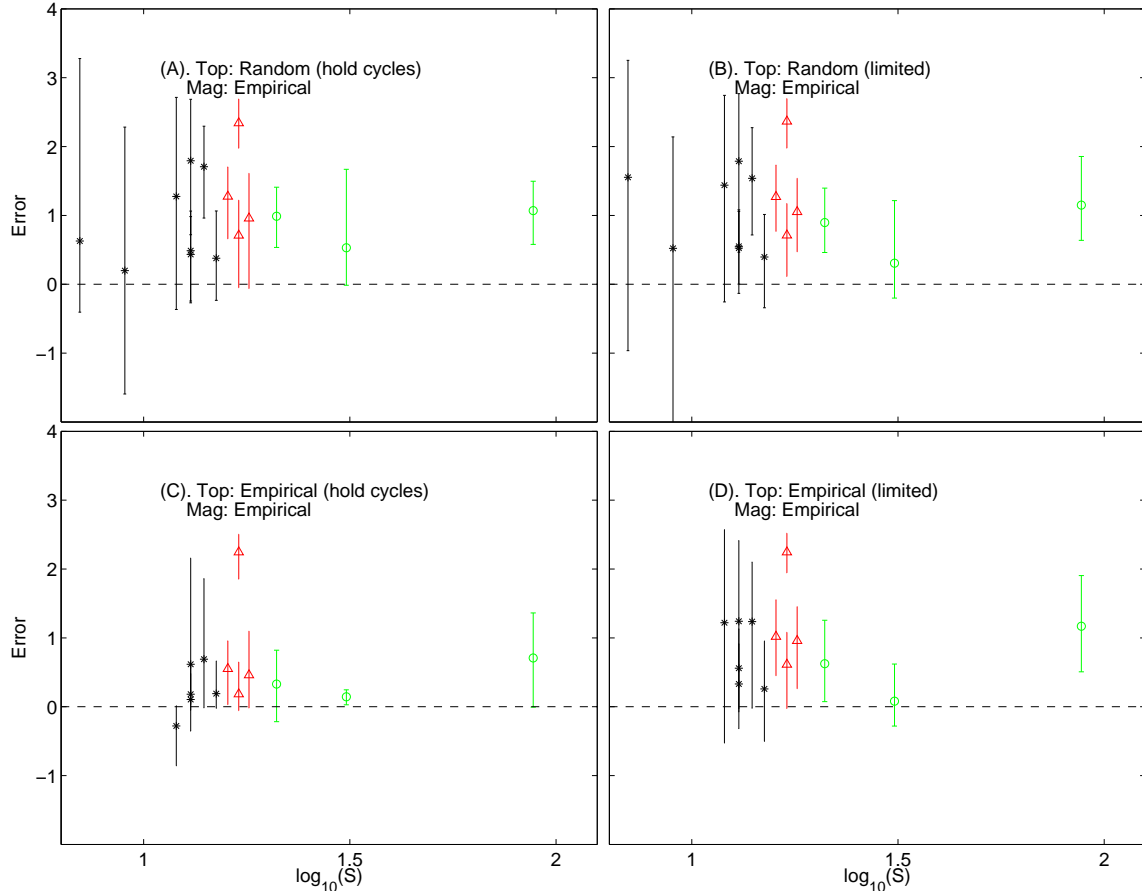
where x_i is the biomass density of species i , r_i is the intrinsic growth rate of species
 696 i and the q_{ij} are Lotka–Volterra coefficients. The elements q_{ii} must be non-positive to
 prevent boundless growth of species i in isolation. Eq. (C.1) has an equilibrium point \mathbf{x}^*
 698 satisfying $Q\mathbf{x}^* = -\mathbf{r}$. At this equilibrium, the diagonal elements of the community matrix



Supplementary Figure S1. Figure 3 redrawn with the matrices ordered by l_0 , the real part of the leading eigenvalue of the empirically derived matrix, for the same six randomization algorithms (A–F) as in Fig. 3, changing either or both of network topology (Top) and magnitude of interaction strengths (Mag). Symbols indicate the method used to construct the community matrix: biomass flux = black stars (successional webs) or red triangles (soil webs); EwE = blue squares; PPMR = green circles.

Figure	Pairs	Elements	Topology	Cycles	Rows	Limited changes	E(error)	P(under)
2A	×	×	×	×	×	×	2.29	0.03
2B	✓	×	×	×	×	×	1.7	0.14
2C	✓	×	✓	×	×	×	1.24	0.14
2D	×	×	×	×	✓	×	0.24*	0.42*
2E	✓	✓	×	×	×	×	1.21	0.13
2F	✓	✓	✓	×	×	×	1.01	0.16
2G	✓	✓	×	×	×	×	1.31	0.12
2H	✓	✓	×	×	✓	×	0.49	0.29
4A (<i>y</i>)	✓	✓	×	✓	×	✓	0.70	0.09
4A (<i>x</i>)	✓	✓	×	×	×	✓	0.77	0.05
4B (<i>y</i>)	✓	✓	✓	✓	×	✓	0.24	0.21
4B (<i>x</i>)	✓	✓	✓	×	×	✓	0.67	0.08

Supplementary Table S1. The properties preserved by each off-diagonal randomization algorithm. *Pairs*: $b_{ij} \neq 0$ if and only if $b_{ji} \neq 0$. *Elements*: the elements of the original matrix were permuted rather than random deviates from a probability distribution. *Topology*: the original network topology was preserved. *Cycles*: the positions and weights of the cycles of length three was preserved. *Rows*: row structure was preserved. *Limited changes*: The number of elements of the original matrix that were moved was limited to be the same as in the corresponding cycle-preserving algorithm. E(error) is the mean value of error = $\log_{10}(l_{\text{rand}}/l_{\text{emp}})$ of 200 realisations of the randomization scheme across the 21 empirically derived matrices. P(under) is the proportion of realisations for which the error is negative. An unbiased randomization scheme would have E(error) = 0 and P(under) = 0.5. Entries marked * do not include the freshwater webs as this algorithm does not preserve pairwise structure and as expected gives very poor results for these matrices.



Supplementary Figure S2. The distribution of error = $\log_{10}(l_{\text{rand}}/l_{\text{emp}})$ for the four randomization algorithms used in Fig. 4. Fig. 4A plots algorithm A against algorithm B; Fig. 4B plots algorithm C against algorithm D. The food webs are plotted against the number of species S in the web on the horizontal axis. Symbols indicate the method used to construct the community matrix: biomass flux = black stars (successional webs) or red triangles (soil webs); PPMR = green circles. Community matrices for which $< 90\%$ of randomizations produce distinct matrices are not shown; this includes all the EwE webs.

are $a_{ii} = q_{ii}x_i^*$. If $x_i^* > 0$ for all i , this equilibrium corresponds to the coexistence of all
700 species. If, as the model parameters are varied, x_j^* becomes negative for some j , there
is a transcritical bifurcation and a different equilibrium, with $x_j^* = 0$, becomes stable.
702 Hence, species j goes extinct via a gradual decline of x_j^* to zero and at the transcritical
bifurcation, $a_{jj} = 0$.

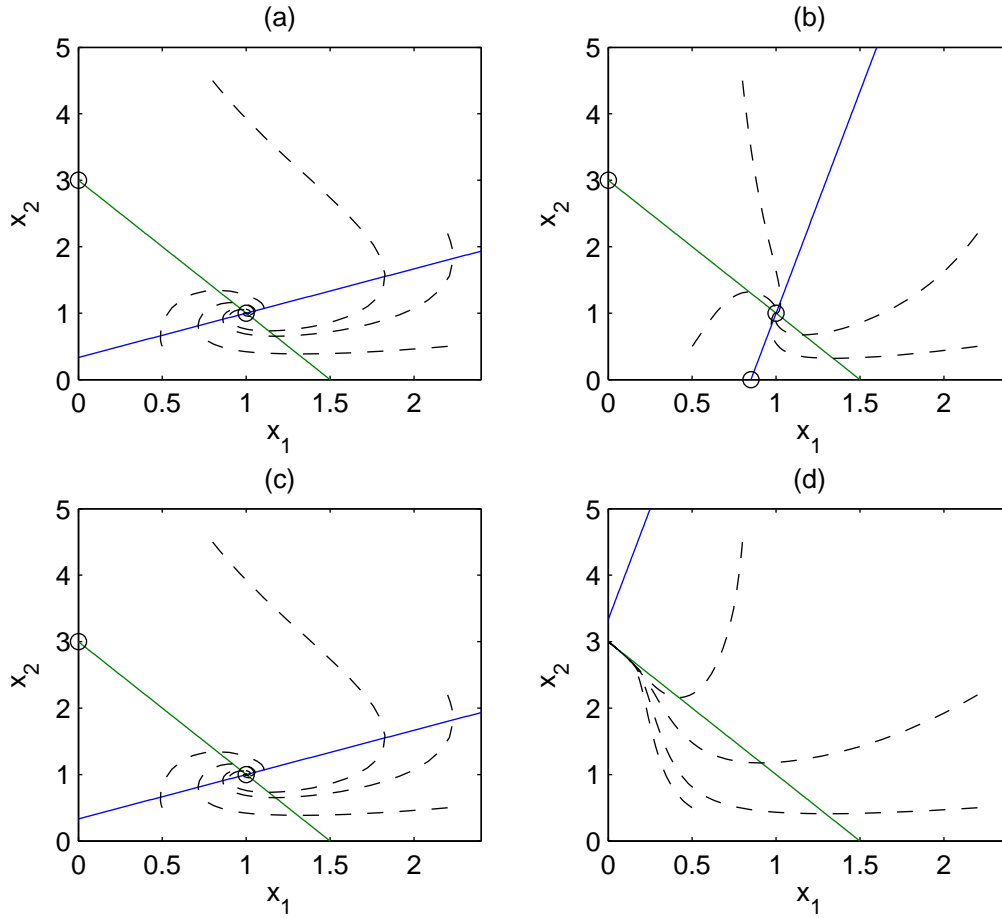
704 Now consider the effect of requiring the diagonal elements of the community matrix to
be equal, $a_{ii} = -d < 0$. Combining this with the equilibrium condition $Q\mathbf{x}^* = -\mathbf{r}$ shows
706 that the intrinsic growth rates r_i must satisfy

$$r_i = d \sum_{j=1}^S \frac{q_{ij}}{q_{jj}}, \quad i = 1, \dots, S, \quad (\text{C.2})$$

which is Eq. (4) in the main text. **This means that basal species i , which are**
708 **negatively influenced by their interactions with other species and so have**
 $q_{ij} < 0$, **will always have positive r_i (recall that $q_{jj} < 0$). Top predators, which**
710 **are positively influenced by their interactions and so have $q_{ij} > 0$ ($i \neq j$) will**
typically have negative r_i .

712 Under the constraint in Eq. (C.2), the equilibrium equation $Q\mathbf{x}^* + \mathbf{r} = \mathbf{0}$ becomes
 $Q(\mathbf{x}^* - \mathbf{v}) = \mathbf{0}$, where $v_i = -d/q_{ii}$, which is always positive since $d > 0$ and $q_{ii} < 0$.
714 Hence, there is always a positive equilibrium with $x_i^* = v_i$. This equilibrium cannot
undergo a transcritical bifurcation with another equilibrium for which $v_j^* = 0$ for some
716 j . The only way a species can go extinct is if the matrix Q becomes rank-deficient,
which leads to a line of non-isolated equilibria through the point $\mathbf{x}^* = \mathbf{v}$. The positive
718 equilibrium at $\mathbf{x}^* = \mathbf{v}$ suddenly becomes unstable and an equilibrium in which one or
more species is zero suddenly becomes stable.

720 Supplementary Fig. S3 illustrates the bifurcation structure of the model, with and with-
out the constraint imposed by Eq. (C.2), for the simple two-species case. In the con-
722 strained model (Fig. S3(a,b)), the two nullclines are forced to intersect at the point



Supplementary Figure S3. Phase planes for a two-species generalised Lotka–Volterra model (C.1). The blue line is the x_1 nullcline, the green line is the x_2 nullcline, the open circles are equilibria and the dashed curves are example orbits. **(a,b)** show the case where the parameters are constrained by Eq. (C.2): the nullclines are forced to cross at $(x_1, x_2) = (-d/q_{11}, -d/q_{22})$, which is always in the interior of the positive quadrant; at the bifurcation, the nullclines lie on top of another creating a line of non-isolated equilibria; after the bifurcation (b) one of the two species goes suddenly extinct (which species goes extinct is sensitive to initial conditions). **(c,d)** show the case where the parameters are not constrained: as the x_2 nullcline steepens, the equilibrium density of species 2 gradually declines gradually until the equilibrium collides with the x_1 axis, at which point a transcritical bifurcation occurs; after the bifurcation (d), species 2 is extinct. Parameter values in (a,c): $r_1 = -0.5$, $r_2 = 3$, $q_{11} = q_{22} = q_{21}$, $q_{12} = 1.5$, $q_{21} = -2$, $d = 1$. In (b), q_{12} is changed to 0.15 and r_1 is changed to 0.85 to satisfy Eq. (C.2). In (d), q_{21} is changed to 0.15 but r_2 is left as -0.5 so that Eq. (C.2) is not satisfied. This results in extinction of the predator (species 1).

$(x_1, x_2) = (-d/q_{11}, -d/q_{22})$. As the q_{ij} are varied, the equilibrium must always remain in
724 the interior of the positive quadrant, because q_{ii} must be non-positive and finite. There-
fore, no transcritical bifurcations are possible. **The only way in which a species can**
726 **go extinct is if the nullclines become parallel (in the two-species model, this**
requires q_{12} and q_{21} have the same sign). If this happens, the nullclines coincide
728 and there is a line of non-isolated equilibria passing through $(-d/q_{11}, -d/q_{22})$. After this
bifurcation, the equilibrium at $(-d/q_{11}, -d/q_{22})$ loses stability and one of the two species
730 goes extinct.

In the unconstrained model (Fig. S3(c,d)), the slopes and intercepts of the nullclines can
732 vary independently as model parameters are varied. The coexistence equilibrium point
can therefore move outside the positive quadrant and lose stability via a transcritical bi-
734 furcation. **For example, in the transition from Fig. S3(c) to S3(d), the predator**
(species 1) gradually declines to zero, at which point there is a transcritical
736 **bifurcation. The coexistence equilibrium moves outside the feasible region**
and the prey-only equilibrium $(x_1, x_2) = (0, -r_2/q_{22})$ becomes stable.

738 The result that constraining diagonal elements of the community matrix to be equal
precludes transcritical bifurcations is not particular to the generalised Lotka–Volterra
740 model in Eq. (C.1). Consider a more general model where the rate of change of species
i is

$$\frac{dx_i}{dt} = x_i (r_i + q_{ii}x_i + h_i(\mathbf{x})), \quad i = 1, \dots, S, \quad (\text{C.3})$$

742 for some function $h_i(\mathbf{x})$. If we require that $(x_j = 0 \ \forall j \neq i \Rightarrow h_i(\mathbf{x}) = 0)$ then, as for
the generalised Lotka–Volterra model, each species, in isolation, behaves according to a
744 logistic equation with intrinsic growth rate r_i (which can be positive or negative) and
carrying capacity $-r_i/q_{ii} > 0$. However, the generalised Lotka–Volterra model assumes
746 that $h_i(\mathbf{x})$ is a linear function of the x_i , whereas in Eq. (C.3) the interaction terms
contained in $h_i(\mathbf{x})$ can be nonlinear, for example representing a type-II response. The

748 diagonal elements of the Jacobian matrix are

$$a_{ii} = r_i + 2q_{ii}x_i + h_i(\mathbf{x}) + x_i \frac{\partial h_i}{\partial x_i}.$$

Using the equilibrium condition from setting Eq. (C.3) equal to 0, we obtain the diagonal
750 elements of the Jacobian matrix at equilibrium:

$$a_{ii} = x_i^* \left(q_{ii} + \frac{\partial h_i}{\partial x_i} \Big|_{\mathbf{x}=\mathbf{x}^*} \right).$$

This shows that, if species i goes extinct via a transcritical bifurcation ($x_i^* = 0$), then its
752 diagonal element in the community matrix (a_{ii}) must also be zero at the bifurcation point.
If instead the diagonal elements a_{ii} are all constrained to equal $-d$, then the equilibrium
754 satisfies

$$x_i^* = - \frac{d}{q_{ii} + \frac{\partial h_i}{\partial x_i} \Big|_{\mathbf{x}=\mathbf{x}^*}}. \quad (\text{C.4})$$

Unlike in the constrained generalised Lotka–Volterra model, there is not always a unique
756 positive equilibrium. There may be 0, 1 or more solutions to Eq. (C.4), some of which
may have $x_i^* < 0$ for some species. There may be saddle-node bifurcations that change
758 the number of equilibria and there may be Hopf bifurcations that change the stability of
equilibria. Nevertheless, Eq. (C.4) shows that it is impossible for x_i^* to pass smoothly
760 through 0 as a model parameter is varied. Hence, transcritical bifurcations are impossible.

In general, requiring the diagonal elements of the community matrix to be equal imposes
762 S constraints (e.g. Eq. (C.2)) on the model parameters. This is equivalent to taking a
codimension S slice through the full model parameter space. The full parameter space is
764 dominated by transcritical bifurcations, whereas the constrained codimension S parameter
space has no transcritical bifurcations. The constrained model is therefore a singular
766 case. It does not give a representative picture of the ways in which equilibria can gain or
lose stability.

768 References

- Allesina, S. and Pascual, M. (2008). Network structure, predator–prey modules, and
770 stability in large food webs. *Theoretical Ecology*, 1(1):55–64.
- Allesina, S. and Tang, S. (2012). Stability criteria for complex ecosystems. *Nature*,
772 483(7388):205–208.
- Bastolla, U., Fortuna, M. A., Pascual-García, A., Ferrera, A., Luque, B., and Bascompte,
774 J. (2009). The architecture of mutualistic networks minimizes competition and in-
creases biodiversity. *Nature*, 458(7241):1018–1020.
- 776 Berlow, E. L., Navarrete, S. A., Briggs, C. J., Power, M. E., and Menge, B. A. (1999).
Quantifying variation in the strengths of species interactions. *Ecology*, 80(7):2206–2224.
- 778 Brose, U., Williams, R. J., and Martinez, N. D. (2006). Allometric scaling enhances
stability in complex food webs. *Ecology Letters*, 9(11):1228–1236.
- 780 Christensen, V. and Pauly, D. (1992). Ecopath ii—a software for balancing steady-
state ecosystem models and calculating network characteristics. *Ecological modelling*,
782 61(3):169–185.
- Dalsgaard, J., Wallace, S. S., Salas, S., and Preikshot, D. (1998). Mass-balance model
784 reconstructions of the Strait of Georgia: the present, one hundred, and five hundred
years ago. In Pauly, D., Pitcher, T., Preikshot, D., and Hearne, J., editors, *Back to*
786 *the future: reconstructing the Strait of Georgia ecosystem*, volume 6 of *Fisheries Centre*
Research Reports, pages 72–91.
- 788 Dambacher, J. M., Luh, H.-K., Li, H. W., and Rossignol, P. A. (2003). Qualitative
stability and ambiguity in model ecosystems. *The American Naturalist*, 161(6):876–
790 888.
- Datta, S., Delius, G. W., and Law, R. (2010). A jump-growth model for predator–prey

- 792 dynamics: derivation and application to marine ecosystems. *Bulletin of mathematical*
biology, 72(6):1361–1382.
- 794 De Angelis, D. L. (1975). Stability and connectance in food web models. *Ecology*, 56:238–
243.
- 796 De Ruiter, P. C., Neutel, A.-M., and Moore, J. C. (1995). Energetics, patterns of inter-
action strengths, and stability in real ecosystems. *Science*, 269:1257–1257.
- 798 De Ruiter, P. C., Van Veen, J. A., Moore, J. C., Brussaard, L., and Hunt, H. W. (1993).
Calculation of nitrogen mineralization in soil food webs. *Plant and Soil*, 157(2):263–273.
- 800 Elton, C. S. (1958). *The ecology of invasions by animals and plants*. Methuen, London.
- Emmerson, M. C. and Raffaelli, D. (2004). Predator–prey body size, interaction strength
802 and the stability of a real food web. *Journal of Animal Ecology*, 73(3):399–409.
- Emmerson, M. C. and Yearsley, J. M. (2004). Weak interactions, omnivory and emergent
804 food-web properties. *Proceedings of the Royal Society of London. Series B: Biological*
Sciences, 271(1537):397–405.
- 806 Erdős, P. and Rényi, A. (1960). On the evolution of random graphs. *Publ. Math. Inst.*
Hungar. Acad. Sci, 5:17–61.
- 808 Gravel, D., Canard, E., Guichard, F., and Mouquet, N. (2011). Persistence increases with
diversity and connectance in trophic metacommunities. *PLoS One*, 6(5):e19374.
- 810 Gribble, N. A. (2005). Ecosystem modelling of the great barrier reef: A balanced trophic
biomass approach. In Zerger, A. and Argent, R. M., editors, *MODSIM 2005 Interna-*
812 *tional Congress on Modelling and Simulation*, pages 2561–2567.
- Grimm, V. and Wissel, C. (1997). Babel, or the ecological stability discussions: an
814 inventory and analysis of terminology and a guide for avoiding confusion. *Oecologia*,
109(3):323–334.

- 816 Guénette, S. and Christensen, V. (2005). Food web models and data for studying fisheries
and environmental impacts on eastern pacific ecosystems. Technical Report 13(1),
818 University of British Columbia.
- Haydon, D. (1994). Pivotal assumptions determining the relationship between stability
820 and complexity: an analytical synthesis of the stability-complexity debate. *American
Naturalist*, pages 14–29.
- 822 Haydon, D. T. (2000). Maximally stable model ecosystems can be highly connected.
Ecology, 81(9):2631–2636.
- 824 Hofbauer, J. and Sigmund, K. (1998). *Evolutionary games and population dynamics*.
Cambridge University Press, Cambridge.
- 826 Jacquet, C., Moritz, C., Morissette, L., Legagneux, P., Massol, F., Archambault, P., and
Gravel, D. (2013). No complexity-stability relationship in natural communities. *arXiv
828 preprint arXiv:1307.5364*.
- Jansen, W. (1987). A permanence theorem for replicator and lotka-volterra systems.
830 *Journal of Mathematical Biology*, 25(4):411–422.
- Jonsson, T., Cohen, J. E., and Carpenter, S. R. (2005). Food webs, body size, and species
832 abundance in ecological community description. *Advances in ecological research*, 36:1–
84.
- 834 Law, R. and Blackford, J. C. (1992). Self-assembling food webs: a global viewpoint of
coexistence of species in lotka-volterra communities. *Ecology*, pages 567–578.
- 836 Levins, R. (1979). Coexistence in a variable environment. *American Naturalist*, pages
765–783.
- 838 MacArthur, R. (1955). Fluctuations of animal populations and a measure of community
stability. *ecology*, 36(3):533–536.

- 840 Mackinson, S. and Daskalov, G. (2007). An ecosystem model of the North Sea to support
an ecosystem approach to fisheries management: description and parameterisation.
842 *Cefas Science Series Technical Report*, 142:195.
- May, R. M. (1972). Will a large complex system be stable. *Nature*, 238:413–414.
- 844 McCann, K., Hastings, A., and Huxel, G. R. (1998). Weak trophic interactions and the
balance of nature. *Nature*, 395(6704):794–798.
- 846 McCann, K. S. (2000). The diversity–stability debate. *Nature*, 405(6783):228–233.
- Meszéna, G., Gyllenberg, M., Pásztor, L., and Metz, J. A. J. (2006). Competitive exclu-
848 sion and limiting similarity: a unified theory. *Theoretical Population Biology*, 69(1):68–
87.
- 850 Metha, M. L. (1967). Random matrices and the statistical theory of energy levels. *Aca-
demic, New York*.
- 852 Moellmann, C. and Diekmann, R. (2012). Marine ecosystem regime shifts induced by
climate and overfishing: A review for the northern hemisphere. *Advances in Ecological
854 Research*, 47:303–347.
- Moore, J. C., De Ruiter, P. C., and Hunt, H. W. (1993). Influence of productivity on the
856 stability of real and model ecosystems. *Science*, 261:906–908.
- Moore, J. C., de Ruiter, P. C., Hunt, H. W., Coleman, D. C., and Freckman, D. W. (1996).
858 Microcosms and soil ecology: Critical linkages between fields studies and modelling food
webs. *Ecology*, 77(3):694–705.
- 860 Neutel, A.-M., Heesterbeek, J. A. P., and de Ruiter, P. C. (2002). Stability in real food
webs: weak links in long loops. *Science*, 296(5570):1120–1123.
- 862 Neutel, A.-M., Heesterbeek, J. A. P., van de Koppel, J., Hoenderboom, G., Vos, A.,
Kaldeway, C., Berendse, F., and de Ruiter, P. C. (2007). Reconciling complexity with
864 stability in naturally assembling food webs. *Nature*, 449(7162):599–602.

- O’Gorman, E. J. and Emmerson, M. C. (2010). Manipulating interaction strengths and
866 the consequences for trivariate patterns in a marine food web. *Advances in Ecological
Research*, 42:301–419.
- 868 Opitz, S. (1996). Trophic interactions in caribbean coral reefs. Technical Report 43,
International Center for Living Aquatic Resources Management.
- 870 Otto, S. B., Rall, B. C., and Brose, U. (2007). Allometric degree distributions facilitate
food-web stability. *Nature*, 450(7173):1226–1229.
- 872 Pawar, S., Dell, A. I., and Savage, V. M. (2012). Dimensionality of consumer search space
drives trophic interaction strengths. *Nature*, 486:485–489.
- 874 Pimm, S. L. (1980). Food web design and the effect of species deletion. *Oikos*, pages
139–149.
- 876 Pimm, S. L. and Lawton, J. H. (1978). On feeding on more than one trophic level. *Nature*,
275(5680):542–544.
- 878 Reuman, D. C., Mulder, C., Banašek-Richter, C., Cattin Blandenier, M.-F., Breure,
A. M., Hollander, H. D., Kneitel, J. M., Raffaelli, D., Woodward, G., and Cohen, J. E.
880 (2009). Allometry of body size and abundance in 166 food webs. *Advances in Ecological
Research*, 41:1–44.
- 882 Rossberg, A. G. (2013). *Food webs and biodiversity: foundations, models, data*. Wiley.
- Staniczenko, P. P. A., Kopp, J. C., and Allesina, S. (2013). The ghost of nestedness in
884 ecological networks. *Nature communications*, 4:1391.
- Tang, S., Pawar, S., and Allesina, S. (2014). Correlation between interaction strengths
886 drives stability in large ecological networks. *Ecology Letters*, 17(9):1094–1100.
- Thébault, E. and Fontaine, C. (2010). Stability of ecological communities and the archi-
888 tecture of mutualistic and trophic networks. *Science*, 329(5993):853–856.

- Twomey, M., Jacob, U., and Emmerson, M. C. (2012). Perturbing a marine food web:
890 Consequences for food web structure and trivariate patterns. *Advances in Ecological
Research*, 47:349–409.
- 892 Tylianakis, J. M., Laliberté, E., Nielsen, A., and Bascompte, J. (2010). Conservation of
species interaction networks. *Biological conservation*, 143(10):2270–2279.
- 894 Walters, C., Martell, S. J. D., and Mahmoudi, B. (2006). An Ecosim model for exploring
ecosystem management options for the Gulf of Mexico: implications of including multi-
896 stanza life history models for policy predictions. In *Presentation for Mote Symposium*,
volume 6, page 23.
- 898 Wiggins, S. (2003). *Introduction to applied nonlinear dynamical systems and chaos*, vol-
ume 2. Springer.
- 900 Woodward, G., Ebenman, B., Emmerson, M., Montoya, J. M., Olesen, J. M., Valido,
A., and Warren, P. H. (2005a). Body size in ecological networks. *Trends in ecology &
902 evolution*, 20(7):402–409.
- Woodward, G., Speirs, D. C., and Hildrew, A. G. (2005b). Quantification and resolution
904 of a complex, size-structured food web. *Advances in ecological research*, 36:85–135.
- Yodzis, P. (1981). The stability of real ecosystems. *Nature*, 289(5799):674–676.