A stochastic theory of community food webs
I. Models and aggregated data

BY J. E. COHEN[1] AND C. M. NEWMAN[2]

[1] *Rockefeller University*, 1230 *York Avenue, New York,
New York* 10021-6399, *U.S.A.*

[2] *Department of Mathematics, University of Arizona,
Tucson, Arizona* 85721, *U.S.A.*

Three recently discovered quantitative empirical generalizations describe
major features of the structure of community food webs. These general-
izations are: (i) a species scaling law: the mean proportions of basal,
intermediate and top species remain invariant at approximately 0.19,
0.53, and 0.29, respectively, over the range of variation in the number
of species in a web; (ii) a link scaling law: the mean proportions of trophic
links in the categories basal–intermediate, basal–top, intermediate–
intermediate, and intermediate–top remain invariant at approximately
0.27, 0.08, 0.30 and 0.35, respectively, over the range of variation in the
number of species in a web; and (iii) a link-species scaling law: the ratio
of mean trophic links to species remains invariant at approximately 1.86,
over the range of variation in the number of species in a web. This paper
presents a model, the only successful one among several attempts, in
which the first two of these empirical generalizations can be derived as
a consequence of the third. The model assumes that species are ordered
in a cascade or hierarchy such that a given species can prey on only those
species below it and can be preyed on by only those species above it in
the hierarchy.

1. INTRODUCTION

A *food web* is a set of kinds of organisms and a relation that shows which kinds
of organisms, if any, each kind of organism in the set eats. A *community food web*
is a food web obtained by picking, within a habitat or set of habitats, a set of kinds
of organisms on the basis of taxonomy, size, location, or other criteria, without
prior regard to the eating relations among the organisms (Cohen 1978, pp. 20–21).
In the past hundred years, ecologists have reported many community food webs.
Briand (1983, and unpublished data) collected and edited 62 of these, including
13 of those assembled by Cohen (1978). Several simple empirical generalizations
describe the major features of these community food webs, viewed as an ensemble
(Briand & Cohen 1984; Cohen & Briand 1984).

The purpose of this paper is to propose a simple explanation that accounts for
these empirical generalizations in an economical way. The proposed explanation
(the 'cascade' model of §6) is one of several attempted models. The unsuccessful
models will also be reviewed to show why models that are simpler than the one
we ultimately propose do not account for the major features of the data.

Section 2 introduces our terminology and summarizes the empirical generalizations that this work aims to explain. Sections 3–6 describe successively more restricted stochastic models, based on random directed graphs, and their failures and successes in accounting for the observed generalizations. Section 7 reviews the results obtained, relates them to prior results, and points out some of their limitations.

The companion to this paper (Cohen *et al.* 1985) tests further the most successful model proposed here, by using disaggregated data on individual community food webs.

2. TERMINOLOGY AND EMPIRICAL GENERALIZATIONS

We shall follow the terminology and restate the major conclusions of Briand & Cohen (1984) and Cohen & Briand (1984).

By a *species*, we mean a class of organisms that prey on the same kinds of organisms and are preyed on by the same kinds of organisms. A species in this sense may result from lumping together kinds of organisms that were identified as separate by a reporting ecologist but that were recorded as having the same sets of prey and the same sets of predators (Briand & Cohen 1984). A species in this sense bears no necessary relationship to a biological species.

By a *link*, we mean any reported feeding or trophic relation between two species in a community food web. Observers use various criteria to decide how much feeding justifies the reporting of a link and how much failure to observe feeding justifies reporting the absence of a link (Cohen & Briand 1984).

A community food web graph represents a community food web as a directed graph or digraph. (The use of digraphs to represent food webs was proposed, apparently independently, by Harary (1961) and Gallopin (1972).) The vertices of the digraph correspond to the set of species in the community food web, and there is an arrow or directed edge from vertex i to vertex j in the digraph if and only if species j feeds on species i, that is, food flows from species i to species j. In the description of the theory of digraphs by Robinson & Foulds (1980), the possibility that $i = j$, that is, cannibalism, is excluded. As will be explained below, cannibalism was excluded from our data, independently of the theory of digraphs. Consequently the data are consistent with the assumptions of Robinson & Foulds (1980). Henceforth we shall use the single word *web* to mean a digraph that represents a community food web. We shall sometimes use the words species and vertex interchangeably.

A *predator* is a species that eats at least one species in the web. A *prey* is a species that is eaten by at least one species in the web. A *top* species is a species not eaten by any species in the web. Such a species is represented in the web by a vertex that is called a sink (Robinson & Foulds 1980, p. 20). An *intermediate* species is a species that has both at least one predator and at least one prey. A *basal* species is a species that eats no species. Such a species is represented in the web by a vertex that is called a source (Robinson & Foulds 1980, p. 20).

A species that neither eats nor is eaten by any species (an *isolated* species) is, according to the definitions just given, both a top and a basal species. However,

either such species do not exist in reality or reports of webs, with rare exceptions, exclude them. In the whole collection of 62 webs that we shall analyse, only two or three isolated species in total were reported by the original sources, and these isolated species have been excluded in the editing of the data (F. Briand, personal communication).

We now distinguish special subsets of top and basal species. A *proper top* species is a top species that is also a predator, that is, a species that is eaten by none, but that eats at least one other species. A proper top species is represented by a vertex that is a proper sink in the terminology of Robinson & Foulds (1980, p. 20). A *proper basal* species is a basal species that is also a prey, that is, a species that eats none, but that is eaten by at least one other species. A proper basal species is represented by a vertex that is a proper source (Robinson & Foulds 1980, p. 20).

Because isolated species are absent from our data, all reported top species are proper top species and all reported basal species are proper basal species. In the absence of isolated species, we can partition all species in a web into the sets of proper top, intermediate, and proper basal species.

A *basal–intermediate link* is a link from a (necessarily proper) basal species to an intermediate species; similarly for a *basal–top link*, an *intermediate–intermediate link*, and an *intermediate–top link*.

For a given reported web, let S denote the total number of species (vertices), T the number of (proper) top species, I the number of intermediate species, B the number of (proper) basal species, L the total number of links, L_{BI} the number of basal–intermediate links, L_{BT} the number of basal–top links, L_{II} the number of intermediate–intermediate links, and L_{IT} the number of intermediate–top links.

The adjacency matrix A of a web (or of any digraph) is an $S \times S$ matrix in which the element a_{ij} in row i and column j equals 1 if species i is eaten by species j, and equals 0 if species i is not eaten by species j. Thus species j is a basal species if and only if column j of A is 0, because column j of A is 0 if and only if species j eats no species in the web. Species j is a proper basal species if and only if column j of A is 0 and row j is not 0. Similarly species i is a top species if and only if row i of A is 0. Species i is a proper top species if and only if row i is 0 but column i is not 0. Species i is isolated if both row i and column i are 0.

As is conventional, let $E(.)$ denote the expectation or average of the random variable enclosed in parentheses. Let a bar denote the sample mean of the random variable it covers. Thus $\bar{B}$ is the sample mean number of basal species, while $E(B)$ is the expected number of basal species according to some model.

The three major findings of Briand & Cohen (1984) and Cohen & Briand (1984) may be stated as 'scaling laws', that is, as summaries of how the variables just defined change, or scale, as the total number of species in a web increases. Each of these scaling laws has two parts: (i) a qualitative part that states the approximate form of a scaling relationship, and (ii) a quantitative part that estimates the numerical value of the parameter or parameters in the scaling law. The scaling laws are cross-sectional, not longitudinal: they describe a comparison of many webs at the moment of observation, not the development of a single web over time resulting from the sequential addition of species.

Species scaling (Briand & Cohen 1984)

(i) As S varies from 3 to 33 *lumped* species, $\bar{B}$, $\bar{I}$ and $\bar{T}$ are all approximately proportional to S. Equivalently, the proportions of species that are basal, intermediate and top show no pronounced trend, neither increasing nor decreasing, as S varies from 3 to 33.

(ii) Approximately, $\bar{B} = 0.19S$, $\bar{I} = 0.53S$, and $\bar{T} = 0.29S$ for all webs. (The sum $0.19 + 0.53 + 0.29$ exceeds 1 due to rounding. For more exact figures, see table 1.)

TABLE 1. SUMMARY STATISTICS OF THE NUMBERS OF SPECIES AND LINKS IN 62 COMMUNITY WEBS, BY TYPE OF WEB, TYPE OF SPECIES, AND CATEGORY OF LINK (FROM COHEN & BRIAND 1984)

	constant webs		fluctuating webs†		all webs	
	number	fraction	number	fraction	number	fraction
webs ...	19		43		62	
all species	351	1.000	683	1.000	1034	1.000
basal	66	0.188	130	0.190	196	0.190
intermediate	177	0.504	366	0.536	543	0.525
top	108	0.308	187	0.274	295	0.285
all links	811	1.000	1108	1.000	1919	1.000
basal–intermediate	198	0.244	327	0.295	525	0.274
basal–top	92	0.113	56	0.051	148	0.077
intermediate–intermediate	260	0.321	318	0.287	578	0.301
intermediate–top	261	0.322	407	0.367	668	0.348

† The environment of a web is considered to be 'fluctuating' if the original report indicates temporal variations of substantial magnitude in temperature, salinity, water availability or any other major physical parameter. Otherwise, the environment of the web is considered to be 'constant'.

It seems plausible (Pimm 1982) that ecologists have been more interested in species at the top of webs than in species at the bottom, and that the coefficient 0.19 for the observed fraction of basal species is lower than the true fraction of basal species. When Briand & Cohen (1984) 'lumped' trophic species, they found that the ratio of basal species to top species increased relative to the ratio observed by Cohen (1977, 1978), as expected from Pimm's suspicions. Supposing that the number of top species in table 1 were correctly observed, and that the number of basal species were increased to equal the number of top species, as predicted by all of our models, the fraction of all species that are top species would decline to 0.26 and the fraction of all species that are basal would increase to 0.26. This number seems a reasonable estimate of the fractions of top and basal species, corrected for the possible undercount of basal species.

Link scaling (Cohen & Briand 1984)

(i) As S varies from 3 to 33, $\bar{L}_{BI}$, $\bar{L}_{BT}$, $\bar{L}_{II}$ and $\bar{L}_{IT}$ are all approximately proportional to L. Equivalently, the proportions of links that are basal–

intermediate, basal–top, intermediate–intermediate and intermediate–top show no pronounced trend, neither increasing nor decreasing, as S varies from 3 to 33.

(ii) Approximately, $\bar{L}_{BI} = 0.27L$, $\bar{L}_{BT} = 0.08L$, $\bar{L}_{II} = 0.30L$, and $\bar{L}_{IT} = 0.35L$, for all webs.

Link-species scaling (*Cohen & Briand* 1984)

(i) As S varies from 3 to 33, $\bar{L}$ is approximately proportional to S. Equivalently, the ratio of total links to total species in a web shows no pronounced trend, neither increasing nor decreasing, as S varies from 3 to 33.

(ii) Approximately, $\bar{L} = 1.86S$, for all webs. (More precisely, the coefficient of proportionality is 1.8559 with a standard deviation of 0.0740.) It will be convenient later to have a notation for the empirically observed ratio of links to species; we denote this quantity by d, to suggest 'density of links per species'. Thus in our data $d = 1.86$ approximately.

In stating these empirical generalizations, we have repeatedly emphasized that the range of variation in the total number of lumped species S among the webs collected by Briand is from 3 to 33. We cannot know whether these generalizations will continue to hold in webs with substantially larger S. The theory to be developed predicts that the scaling laws will continue to hold for larger S.

The scaling laws just stated are all first-order laws that describe trends only. They neglect entirely variability with respect to the trends. We shall discuss variability briefly in connection with the cascade model of §6.

A fourth empirical generalization plays a major role in attempts to explain the first three. Gallopin (1972, p. 266) observed that 'directed food webs are in general acyclic, although exceptions are possible'. Cohen (1978, p. 57) found one case of cannibalism, but no larger cycles, in four webs. In the 62 webs of Briand, cannibalism was reported by very few of the original sources, and then only for one species in the web. Because cannibalism is widespread in nature, particularly among invertebrates, the original investigators must have largely, but not consistently, ignored cannibalism. Consequently, Briand chose to exclude all of the few reported cases of cannibalism (F. Briand, personal communication).

To be precise in describing trophic cycles other than cannibalism, we now define (Robinson & Foulds 1980, pp. 24–25, 70) a *walk* in a digraph to be a finite sequence, consisting of vertices and edges alternately, beginning and ending with vertices, in which each edge goes from the vertex written on its left to the vertex written on its right. For example a, (a, b), b, (b, c), c is a walk from vertex a to vertex c through the edges (a, b) and (b, c). A walk in which the first vertex is the same as the last vertex is a closed walk. The *length* of a walk is the number of edges it contains, each counted as often as it occurs. A *cycle* is a closed walk in which all vertices are distinct except the first and last. If two cycles pass through the same set of vertices in the same order, differing only in the vertex that is written down first, the two cycles are considered to be identical. A *k-cycle* is a cycle of length $k > 0$. (If cannibalism is excluded, then 1-cycles are impossible. However, it will sometimes be convenient later to consider the possibility of a directed edge from a vertex to itself, that is, an 1-cycle or loop.) For $k > 0$, a digraph is *k-acyclic* if

it contains no h-cycles, for $h = 1, \ldots, k$. A digraph is *acyclic* if it contains no k-cycles, for any $k > 0$.

If we use this language, all 62 of Briand's webs are acyclic except cases 21 and 30 (in the numbering of Briand (1983) where the unlumped matrices are given). Webs 21 and 30 each contain a single cycle of length 2, and no longer cycles (F. Briand, personal communication). We summarize the distribution of cycles in the webs assembled by Briand with a fourth empirical generalization: acyclicity.

Acyclicity (Gallopin 1972)

Nearly all webs are acyclic.

It seems nearly certain that decomposers feed on what appear as top species and are food for what appear as basal species. The absence of cycles of length greater than 2 implies that the reporters of webs ignore the decomposers. Therefore we cannot examine the ecological role of decomposers with these data.

3. MODEL 0: ANARCHY

We shall say that a random variable Y is Bernoulli with parameter p, and shall write $Y \sim B(p)$, for $0 \leqslant p \leqslant 1$, if $Y = 1$ with probability p and $Y = 0$ with probability $q = 1 - p$. (In our notation, $\sim$ means 'has the distribution of' or 'distributed as' rather than 'asymptotically or approximately equals'.) We shall say that a random matrix X is independently and identically distributed (i.i.d.) Bernoulli with parameter p, and shall write $X \sim$ i.i.d. $B(p)$, if every element X_{ij} of X is $\sim B(p)$, and all elements of X are i.i.d. (Since p is assumed constant for all elements of the matrix, the additional requirement that they be *identically* distributed is redundant, but is retained to accord with convention.)

Suppose A, the $S \times S$ adjacency matrix of a model web, were $\sim$ i.i.d. $B(p)$. Then $E(L) = pS^2$. However, according to the link-species scaling law, $\bar{L} = dS$. These two equations are simultaneously satisfied (with $E(L) = \bar{L}$) if $p = d/S$. To avoid confusing the empirical estimate of density d with a model parameter, we shall specify p in all of our models as c/S. The relation between the model parameter c and the sample statistic d will vary from model to model.

Suppose that each species in a web of S species has an identical and independent chance p of eating any species, including itself, in the web, where, as the number S of species increases, the probability p decreases according to c/S, that is, let $A \sim$ i.i.d. $B(c/S)$ for $S \geqslant c$.

We now analyse the properties of model 0 and compare them with the empirical generalizations above, using $d = 1.86$ as an estimator for c.

Species scaling

The probability that a species is a top species is q^S, and this is also the probability that a species is a basal species. Thus

$$E(T)/S = E(B)/S = q^S = (1 - c/S)^S \tag{3.1}$$

is the expected fraction of species that are top species in a web of S species, and also the expected fraction of species that are basal species. Thus model 0 predicts that the fractions of top and basal species should be equal.

Similarly, the probability that a species is a proper top species is $q^S(1 - q^{S-1})$,

and this is also the probability that a species is a proper basal species. Thus model 0 predicts that the fractions of proper top and proper basal species should be equal.

This prediction is only roughly consistent with the empirical observation that 0.19 of species are basal and 0.29 of species are top. However, like Pimm (1982), we believe that ecologists often are more interested in species at the top of food chains than in species at the bottom. If other predictions of the model turned out to be correct, we would be prepared to accept the model's prediction that in properly collected data, the expected fractions of top and of basal species are equal.

The right member of (3.1) increases monotonically and becomes close to the limiting value

$$\lim_{S \to \infty} E(T)/S = \lim_{S \to \infty} E(B)/S = e^{-c} \tag{3.2}$$

even for moderate values of S. For example, $(1-d/10)^{10} = 0.13$ and $(1-d/20)^{20} = 0.14$ while $e^{-d} = 0.16$. Thus model 0 predicts that the fractions of top, basal and intermediate species should be very nearly independent of the number S of species in the web. The same conclusion applies to the asymptotic fractions of proper top and proper basal species, which are both equal to $e^{-d} - e^{-2d}$. The predicted change in these proportions for S between 3 and 33 would probably be undetectably small, given the variation among webs in the observed proportions (Briand & Cohen 1984).

While model 0 explains the qualitative part of the species scaling law, its predicted asymptotic fractions of top and basal species seem too low to explain the quantitative part of the species scaling law. The predicted asymptotic fraction 0.16 is substantially lower than the fraction 0.26 estimated above. The predicted asymptotic ratio of the expected number of proper top or proper basal species to the expected number of non-isolated species, given by $[e^{-c} - e^{-2c}]/[1 - e^{-2c}]$, is 0.14, further still from the estimated fraction 0.26.

In going from the proportions of top or basal species to the proportions of proper top or proper basal species, the term involved in the corrections, $e^{-2c} = 0.024$, is small compared with the terms being corrected, given the observed ratio 1.86 of links to species, and appears in both numerator and denominator. When the proportions of top and basal species are corrected to the proportions of proper top and proper basal species, they decrease slightly. This slight decrease holds in the remaining models as well. For this reason, we shall not discuss proper top or proper basal species further until we come to model 3.

Link scaling

We skip the analysis of link scaling because model 0 will be evaluated on other grounds.

Species-link scaling

The assumed behaviour of p as a function of S is chosen to reproduce the observed species-link scaling.

Acyclicity

Although model 0 predicts that about 84 % (that is, a fraction $1 - e^{-d}$) of webs will display cannibalism, model 0 should not be rejected on this basis because cannibalism has been suppressed from the data. However, according to theorem 1 below, model 0 also predicts that about 82 % (that is, a fraction $1 - e^{-d^2/2}$) of webs will have one or more 2-cycles, which is grossly contrary to observation.

Effect of lumping

According to model 0, it could happen that, for some $i < j$, column i is identical to column j and row i is identical to row j. In this case, if the simulated matrix were to be treated in the same way as the real data were treated, species i and j should be lumped. Our analysis so far has ignored the possible need to lump species in the simulated webs. We now show that the probability of needing to lump two non-isolated simulated species according to model 0 is so small that it is perfectly reasonable to ignore lumping, given the observed ratio 1.86 of links to (lumped) species.

Choose $i < j$. Define P (lump i and j) to be the probability that, in a matrix A with entries a_{hk} generated by model 0, column i equals column j and row i equals row j. Similarly, define P (lump non-isolated i and j) to be the probability that, in a matrix A generated by model 0, column i equals column j, row i equals row j, and column i or row i or both are not all zero. The $4S-4$ entries in the two columns and rows consist of $2(S-2)$ pairs of entries and one quartet of entries $(a_{ii}, a_{ij}, a_{ji}, a_{jj})$. To lump species i and j, we require that the two entries of each pair be equal and the four entries in the quartet be equal. Hence P (lump i and j) $= (p^2+q^2)^{2S-4}(p^4+q^4)$, and P (lump non-isolated i and j) $= P$ (lump i and j) $- P$ (i and j are isolated) $= (p^2+q^2)^{2S-4}(p^4+q^4) - q^{4S-4}$. The expected fraction of species that are non-isolated but lost by lumping is then less than or equal to

$$(1/S) \sum_{j=2}^{S} \sum_{i=1}^{j} P \text{ (lump non-isolated } i \text{ and } j)$$

$$= (1/2)(S-1)[(1-2pq)^{2S-4}(p^4+q^4) - q^{4S-4}], \quad (3.3)$$

which, as S increases, approaches $c^2 e^{-4c} = 0.002$ when $c = 1.86$. Thus the expected fraction of non-isolated vertices of a random web generated according to model 0 that should be lumped is negligible, so we do not correct the previous calculations for lumping.

Effect of disconnected weak components

All reported webs are *weakly connected* in the sense that the set of species cannot be divided into two non-empty subsets with no link between the two subsets. (The adjective 'weak' allows for the possibility that the linkage might be in one direction only.) A *weak component* is a maximal set of vertices (species) that is weakly connected. Thus all reported webs have only a single weak component. We now show that the expected fraction of non-isolated species that belong to a single weak component according to model 0 is asymptotically so close to 1 that it is reasonable to ignore the effect of disconnected weak components, given $d = 1.86$.

According to Erdös & Rényi (1960, p. 56, theorem 9b), the fraction of all species (including isolated species) that belong to the largest weak component of a web is asymptotically

$$1 - (2c)^{-1} \sum_{k=1}^{\infty} k^{k-1} (2ce^{-2c})^k / k!. \quad (3.4)$$

Hence the fraction of all species that are not isolated and do not belong to the largest weak component is

$$(2c)^{-1} \sum_{k=1}^{\infty} k^{k-1} (2ce^{-2c})^k / k! - e^{-2c} = (2c)^{-1} \sum_{k=2}^{\infty} k^{k-1} (2ce^{-2c})^k / k! \qquad (3.5)$$

which is approximately 0.002 when $c = 1.86$. Thus 99.8% of the non-isolated species of a random web generated according to model 0 belong to a single weak component, so we do not correct the previous calculations for disconnected components.

In summary, model 0 can explain roughly the observed scale-invariance in the proportion of top, intermediate and basal species and the numerical similarity in the proportions of top and basal species. But it predicts fractions of top and basal species that are too low and fractions of food webs with cycles that are far too high.

4. Model 1: finitely acyclic democracy

The most straightforward way to eliminate the problem of too many cycles is by assumption. We start with the weakest assumption that is *a priori* plausible.

Suppose there is a finite positive integer k and a finite positive real number c such that, for $S \geq c$, the adjacency matrix A of a web with S species is $\sim$ i.i.d. $B(c/S)$, conditional on A being k-acyclic.

Biologically, this model assumes that any species can eat any species with equal probability c/S provided that, in the resulting feeding relations, it never happens that species X eats species X (no 1-cycles), nor that species X eats species Y and species Y eats species X (no 2-cycles), nor that species X eats species Y, species Y eats species Z and species Z eats species X (no 3-cycles), nor that there are any cycles of length up to and including k, which is fixed and independent of S.

One way to simulate this model would be to generate Bernoulli matrices according to model 0 and then throw away those matrices A in which the trace (sum of the diagonal elements) of $A + A^2 + \ldots + A^k$ exceeds 0.

Before considering general k, we consider the special case of 1-acyclic democracy.

1-acyclic democracy

To generate an $S \times S$ Bernoulli matrix A with parameter c/S, conditional on no cannibalism (no 1-cycles), set the diagonal elements of A equal to 0 with probability 1. The off-diagonal elements of A are to be filled with independent random variables $\sim B(c/S)$ as before. Then $E(L) = (c/S)S(S-1)$. Since the species-link scaling law gives $\bar{L} = dS$, we can estimate c by $c = dS/(S-1)$, which approaches d for large S but is larger than d for finite S.

The probability that a species is a top species is q^{S-1}, where $q = 1 - c/S$, and this is also the probability that a species is a basal species. Thus

$$E(T)/S = E(B)/S = q^{S-1} = (1 - d/[S-1])^{S-1} \qquad (4.1)$$

is the expected fraction of species that are top species in a web of S species, and also the fraction of species that are basal species. This model predicts that the

fractions of top and basal species should be equal. The asymptotic behaviour of $E(T)/S$ and $E(B)/S$ for large S is identical to that in (3.2) for model 0. The predicted asymptotic fractions of top and basal species are too low to accord well with observation.

A web will have a 2-cycle if there exist indices $i, j \neq i$ such that $a_{ij} = 1$ and $a_{ji} = 1$. For a given i and j, the probability that there is a 2-cycle through i and j is p^2, so the probability that there is no 2-cycle through i and j is $1 - p^2$. The probability that there is no 2-cycle in the entire web is

$$(1-p^2)^{S(S-1)/2} = (1-\{d/[S-1]\}^2)^{S(S-1)/2} \to e^{-d^2/2} = 0.18 \qquad (4.2)$$

so that about 82 % of such model webs would have at least one 2-cycle. This proportion is grossly too high and we are forced to abandon 1-acyclic democracy as unrealistic.

The calculated asymptotic fraction e^{-c} of webs under model 0 (anarchy) that have no 1-cycles may be multiplied by the calculated asymptotic fraction $e^{-c^2/2}$ of webs under 1-acyclic democracy that have no 2-cycles to give the predicted asymptotic fraction $e^{-c-c^2/2}$ of webs under model 0 that are 2-acyclic, that is, have neither 1-cycles nor 2-cycles, because under model 0 the diagonal elements of the adjacency matrix are independent of the off-diagonal elements.

k-Acyclic democracy: the general case

From the perhaps surprising finding that the predicted asymptotic fraction of top or basal species is e^{-c} under the anarchy model as under the model of 1-acyclic democracy, one might conjecture that the proportion is the same under the k-acyclic democracy model, for any finite $k > 0$. From the formula $e^{-c-c^2/2}$ for the asymptotic fraction of webs under model 0 that are 2-acyclic, and from the analogous formulas for undirected graphs of Erdös & Rényi (1960), one might conjecture that the asymptotic proportion of k-acyclic digraphs under model 0 is $\exp\left(-\sum_{h=1}^{k} c^h/h\right)$. The following theorem and corollary establish that both of these conjectures are correct.

THEOREM 1. *Suppose that for some $c \geqslant 0$ and for $S \geqslant c$, the adjacency matrix A of a web with S species is $\sim i.i.d.$ $B(c/S)$. (This is model 0.) Let $M_k(S)$ be the number of distinct k-cycles in the web, $k = 1, 2, \ldots, S$, and let $Y(S)$ be the number of prey species of species 1, that is, the sum of column 1 of A. Let $M(S) = \sum_{h=1}^{S} M_h(S)$ be the total number of distinct cycles in A. Then for any $k \geqslant 0$, the random vector $(Y(S), M_1(S), \ldots, M_k(S))$ (which is interpreted as the scalar $Y(S)$ if $k = 0$) converges in distribution as $S \to \infty$ to a random vector with independent Poisson-distributed components with mean $(c, c, c^2/2, c^3/3, \ldots, c^k/k)$, that is, for any non-negative integers $y, m_1, \ldots, m_k$,*

$$\lim_{S \to \infty} P(Y(S) = y, M_1(S) = m_1, \ldots, M_k(S) = m_k)$$

$$= e^{-c}[c^y/y!]\prod_{h=1}^{k} \{e^{-(c^h/h)}[(c^h/h)^{m_h}]/m_h!\}. \qquad (4.3)$$

For $0 \leqslant c < 1$, $(Y(S), M(S))$ converges in distribution as $S \to \infty$ to a bivariate random vector with independent Poisson-distributed components with mean $(c, -\ln(1-c))$.

COROLLARY. *Under the above assumptions, for any $c \geqslant 0$ and any $k \geqslant 1$, as $S \to \infty$*

$$P(Y(S) = y | M_1(S) = 0, \ldots, M_k(S) = 0) \to e^{-c}[c^y/y!]. \tag{4.4}$$

The left member of (4.4) *is the probability that species 1 has y prey in the model of k-acyclic democracy. For $0 \leqslant c < 1$, the asymptotic probability that species 1 has y prey in an acyclic web is also Poisson, that is,*

$$P(Y(S) = y | M(S) = 0) \to e^{-c}[c^y/y!]. \tag{4.5}$$

The corollary follows immediately from theorem 1 and the definition of conditional probability. The proof of theorem 1 is deferred to appendix 1.

The corollary (with $y = 0$) implies that, in the model of k-acyclic democracy, the fraction of species with no predators, and the fraction of species with no prey, both approach e^{-c} as $S \to \infty$. The mean number of species on which a given species preys, and the mean number of species that prey on a given species, both approach c.

In summary, for fixed finite k, the model of k-acyclic democracy predicts that the expected fractions of top and basal species are equal and, asymptotically for large numbers S of species in a web, independent of S. These predictions are roughly consistent with the data. The model also predicts that the numerical value of this asymptotic fraction should be lower than that observed. However, in concluding that this discrepancy exists, we are assuming that it is appropriate to use the ratio $d = 1.86$ of links to species, observed in the *finite* range of S from 3 to 33, to estimate the *asymptotic* effective density of links c.

5. MODEL 2: ACYCLIC DEMOCRACY

Excluding cycles up to any fixed finite order k, as in model 1, might be qualitatively different, in the limit of large S, from excluding cycles of all lengths in the limit of large S. To investigate this possibility, we have partly analysed the next model.

Suppose there is a finite positive real number c such that, for $S \geqslant c$, the adjacency matrix A of a web with S species is $\sim$ i.i.d. $B(c/S)$, conditional on A being acyclic.

Biologically, this model assumes that any species can eat any species with equal probability c/S, provided that, in the resulting feeding relations, it never happens that species X eats species X (no 1-cycles), nor that species X eats species Y and species Y eats species X (no 2-cycles), nor that species X eats species Y, species Y eats species Z and species Z eats species X (no 3-cycles), and so on, excluding all cycles of length up to and including S.

The theoretical results available to us so far require us to discuss separately two cases: $0 \leqslant c < 1$, and $1 \leqslant c$.

In the first case, (4.5) implies that the fractions of top and basal species are equal and, asymptotically for large S, independent of S. These predictions are roughly consistent with the data. However, since the expected value of the observed density d must be no larger than the model parameter c, and since $d > 1$, this first case is not of empirical interest, given our data.

In the second case, $1 \leqslant c$, we have so far no exact results concerning the

asymptotic proportions of top and of basal species. By symmetry these proportions must be equal. The results of our numerical investigations, which we will now describe, can be interpreted to be consistent with the conjecture that, for $S \gg c$, the fraction of top species and the fraction of basal species both approach e^{-d^*}, where d^* is the asymptotic (large S) effective density of links. We know that this is also the case when $c < 1$ since then $d^* = c$. However, when $c \geqslant 1$, we have no theory so far that permits us to compute c from d^* or vice versa.

To estimate the fractions of zero rows and of zero columns according to model 2, we have resorted to simulation, settling at last on the third of three approaches described in appendix 2. This approach to simulation, which is actually a slight modification of model 2, guarantees that the model parameter c equals the asymptotic ratio d^* of links to species. For $S = 10$ and $S = 20$, and for each value of $c = 0.5(0.5)4.0$ (an abbreviation for the sequence of numbers 0.5, 1.0, 1.5, ..., 4.0), table 2 compares the simulated mean fractions of zero rows and of zero columns in 100 acyclic matrices with the conjectured asymptotic fraction e^{-c}. For the lower values of c, the agreement between the sampled fractions of zero rows or columns and e^{-c} is excellent. For the larger values of c, e^{-c} falls more rapidly than the sampled fractions of zero rows or columns. For large c, the difference between e^{-c} and the sampled fraction of zero rows or columns is slightly smaller for $S = 20$ than for $S = 10$.

TABLE 2. THE SIMULATED MEAN FRACTIONS OF ZERO ROWS OR ZERO COLUMNS IN 100 ACYCLIC $S \times S$ MATRICES WITH EXACTLY Sc POSITIVE ELEMENTS, GENERATED BY THE THIRD APPROACH (APPENDIX 2) TO SIMULATING MODEL 2, AND THE FRACTIONS PREDICTED BY THE ASYMPTOTIC FUNCTION e^{-d^*} CONJECTURED IN (5.1) AND BY THE FUNCTION (6.2a) (WITH c REPLACED BY $2c$) DERIVED FOR MODEL 3, THE CASCADE MODEL

	$S = 10$		predictions		$S = 20$	
c	rows	columns	$\exp(-c)$	model 3	rows	columns
0.5	0.5854	0.5974	0.6065	0.6321	0.5959	0.5962
1.0	0.3500	0.3494	0.3679	0.4323	0.3644	0.3600
1.5	0.2313	0.2445	0.2231	0.3167	0.2341	0.2470
2.0	0.1669	0.1769	0.1353	0.2454	0.1585	0.1642
2.5	0.1620	0.1232	0.0821	0.1987	0.1522	0.1332
3.0	0.1318	0.1042	0.0498	0.1663	0.1121	0.1142
3.5	0.1147	0.1056	0.0302	0.1427	0.0959	0.1133
4.0	0.1032	0.1000	0.0183	0.1250	0.0859	0.0820

For $S = 10$, the standard deviation (computed from the numerical simulation) of the proportion of zero rows (in a single matrix, not in the mean proportion) at first increases with increasing c and then declines slowly from a maximum of approximately 0.09 when $c = 1$ to a minimum of approximately 0.02 when $c = 4$. Since 100 matrices were generated, the standard deviation of the simulated mean proportions given in table 2 is one-tenth as large, that is, not exceeding 0.01. The standard deviations when $S = 20$ are similar, and the same conclusion applies. Thus the difference in table 2 between the sampled proportion of zero columns or rows and $e^{-c} = e^{-d^*}$ for the larger values of c appears to be real.

If this difference approaches 0 as $S \to \infty$, then we may conjecture, pending further theoretical progress, that in model 2,

$$\lim_{S \to \infty} E(T)/S = \lim_{S \to \infty} E(B)/S = e^{-d^*}. \qquad (5.1)$$

If this is so, then, like models 0 and 1, model 2 can explain roughly the observed scale-invariance in the proportion of top, intermediate and basal species and the numerical similarity in the proportions of top and basal species. But it predicts fractions of top and basal species that are too low according to the conjecture (5.1), and that are too low (according to our simulations) even for $S = 10$ (in table 2, $c = 2.0$ gives a fraction of 0 rows near 0.17, lower than the estimate from data of 0.26).

6. MODEL 3: CASCADE

Many biologists might be reassured by the failure of the models considered so far because these models make the biologically implausible assumption that any species is capable, in principle, of eating any other species. These models assume that it is only a matter of chance that the grass does not eat the cow, nor the lamb the wolf. Yet it is not absurd to consider such models. It is a healthy discipline to require that they be rejected by quantitative data and not by 'intuitions' that are often wrong.

Now that the previous models have been rejected for their quantitative failures, we must abandon the assumption that each species could potentially eat any other, while imposing the least possible additional structure. We shall do so by noticing an important feature of acyclic matrices.

An $S \times S$ matrix A is called *strictly upper triangular* if $a_{ij} = 0$ whenever $i \geqslant j$. This means that the main diagonal and all matrix elements below the main diagonal are zero; the non-zero elements of A, if any, lie strictly above the main diagonal. For brevity, we shall henceforth call such a matrix *triangular*.

If the adjacency matrix of a web with S species is triangular, the species labelled 1 can potentially be eaten by any species other than itself, but can eat none. The second species can potentially be eaten by the species labelled 3 to S, but can eat only species 1. And so on: the species labelled S can potentially eat all the other species, but can be eaten by none of them. Thus a triangular adjacency matrix describes a strict trophic hierarchy or cascade.

A digraph is acyclic if and only if its vertices can be numbered in such a way that its adjacency matrix is triangular (for example, Robinson & Foulds 1980, p. 176). Thus the adjacency matrix A of a web is acyclic if and only if some permutation, applied to both rows and columns of A, changes the matrix to triangular form. Model 2 can be interpreted as saying that the luck of the draw determines which species eat which others, provided that, when all is done, the species can be arranged in a cascade. The order of species in the cascade is determined (non-uniquely) after the trophic links are chosen.

We now suppose that the order of species in the cascade is determined before the trophic links are chosen.

Suppose there is a finite positive real number c such that, for $S \geqslant c$, the elements

above the main diagonal of the adjacency matrix A are $\sim$ i.i.d. $B(c/S)$, while the elements on or below the main diagonal are fixed with probability 1 at 0.

THEOREM 2. *Suppose that for some $c \geqslant 0$ and for $S \geqslant c$, the adjacency matrix A of a web with S species is triangular, with the elements above the main diagonal $\sim$ i.i.d. $B(c/S)$. (This is model 3.) Let T be the number of zero rows (top species) and B be the number of zero columns (basal species) in A. Then, with $p = c/S$, $q = 1-p$,*

$$E(T) = E(B) = [1-q^S]/p, \tag{6.1a}$$

$$\mathrm{var}\,(T) = \mathrm{var}\,(B) = (1-q^S)/p - (1-q^{2S})/(1-q^2). \tag{6.1b}$$

Asymptotically,

$$\lim_{S\to\infty} E(T)/S = \lim_{S\to\infty} E(B)/S = (1/c)\,(1-\mathrm{e}^{-c}), \tag{6.2a}$$

$$\lim_{S\to\infty} \mathrm{var}\,(T/S) = \lim_{S\to\infty} \mathrm{var}\,(B/S) = 0. \tag{6.2b}$$

If T_P is the number of proper top species, B_P is the number of proper basal species, N is the number of not isolated species, and I is the number of intermediate species, then

$$E(T_P) = E(B_P) = S[(1-q^S)/c - q^{S-1}], \tag{6.3a}$$

$$E(I) = S[1 - 2(1-q^S)/c + q^{S-1}], \tag{6.3b}$$

$$E(N) = S(1-q^{S-1}), \tag{6.3c}$$

and asymptotically

$$\lim_{S\to\infty} E(T_P)/E(N) = \lim_{S\to\infty} E(B_P)/E(N) = \{[1-\mathrm{e}^{-c}]/c - \mathrm{e}^{-c}\}/[1-\mathrm{e}^{-c}], \tag{6.4a}$$

$$\lim_{S\to\infty} E(I)/E(N) = \{1 - (2/c)\,[1-\mathrm{e}^{-c}] + \mathrm{e}^{-c}\}/[1-\mathrm{e}^{-c}]. \tag{6.4b}$$

For large c, e^{-c} is nearly zero so the asymptotic fraction of top or proper top or basal or proper basal species approaches $1/c$. Also, the total number L of trophic links is binomially distributed with mean and variance

$$\left.\begin{array}{l} E(L) = pS(S-1)/2 = c(S-1)/2, \\ \mathrm{var}\,(L) = pqS(S-1)/2 = c(S-c)\,(S-1)/(2S), \end{array}\right\} \tag{6.5}$$

and the numbers of links of each kind have means

$$E(L_{BI}) = E(L_{IT}) = (S-1)\,(1+q^{S-1}) - (1+q)\,(1-q^{S-1})/p, \tag{6.6a}$$

$$E(L_{BT}) = (1-q^{S-1})/p - (S-1)\,q^{S-1}, \tag{6.6b}$$

$$E(L_{II}) = pS(S-1)/2 - (S-1)\,(2+q^{S-1}) + (1-q^{S-1})\,(1+2q)/p. \tag{6.6c}$$

Asymptotically, as $S \to \infty$,

$$E(L_{BI})/E(L), \quad E(L_{IT})/E(L) \to 2[c(1+\mathrm{e}^{-c}) - 2(1-\mathrm{e}^{-c})]/c^2, \tag{6.7a}$$

$$E(L_{BT})/E(L) \to 2[1-\mathrm{e}^{-c} - c\mathrm{e}^{-c}]/c^2, \tag{6.7b}$$

$$E(L_{II})/E(L) \to 1 - 2[c(2+\mathrm{e}^{-c}) - 3(1-\mathrm{e}^{-c})]/c^2. \tag{6.7c}$$

Proof. Only elementary calculations are required, noting that the probability that species i is basal is q^{i-1}, the probability that species i is top is q^{S-i}, the probability that species i is proper top is $q^{S-i} - q^{S-1}$, the probability that species i is

proper basal is $q^{i-1}-q^{S-1}$, the probability that species i is intermediate is $1-q^{i-1}-q^{S-i}+q^{S-1}$, and the probability that species i is not isolated is $1-q^{S-1}$. Also,

$$
\left.
\begin{aligned}
E(L_{BI}) &= p \sum_{j=2}^{S} \sum_{i=1}^{j-1} q^{i-1}(1-q^{S-j}), \\
E(L_{BT}) &= p \sum_{j=2}^{S} \sum_{i=1}^{j-1} q^{i-1} q^{S-j}, \\
E(L_{IT}) &= p \sum_{j=2}^{S} \sum_{i=1}^{j-1} (1-q^{i-1}) q^{S-j}, \\
E(L_{II}) &= p \sum_{j=2}^{S} \sum_{i=1}^{j-1} (1-q^{i-1})(1-q^{S-j}).
\end{aligned}
\right\} \tag{6.8}
$$
■

When (6.5) is solved for p and $E(L)$ is replaced by the observed number of links, it becomes apparent that p is what ecologists call the (lower) connectance (F. Briand, personal communication).

To compare the predictions of model 3 with observation requires an estimate of c. From (6.5),

$$c = 2E(L)/(S-1). \tag{6.9}$$

For a single finite S, replacing $E(L)$ by the total number of links, we estimate c as twice the total number of links divided by $S-1$. However, for a single value of c common to all webs, we use an asymptotic estimate. Asymptotically, as $S \to \infty$, the link scaling law indicates that $\bar{L}$ is dS, and $S/(S-1) \downarrow 1$ as $S \to \infty$, so that c is estimated as $2d = 3.72$. We now examine the macroscopic predictions of model 3, using this single estimate of $c = 3.72$. We shall review the scaling laws stated in section 2.

Species scaling

Figure 1 shows the predicted mean proportion of top species and a confidence interval of ± 2 standard deviations as a function of S, using (6.1) with a single value of $c = 3.72$, superimposed on the data of Briand & Cohen (1984). Figure 2 shows the same for basal species.

The predicted mean proportion of top or of basal species changes so slowly in the observed range of S as to defy discrimination from constancy. According to (6.1) with $c = 3.72$, model 3 predicts the mean and variance in the proportion of top species to be (with identical results for basal species) as shown in table 3. Thus model 3 reproduces qualitatively the species scaling law.

Quantitatively, model 3 predicts asymptotic proportions of basal, intermediate and top species equal to 0.26, 0.48, and 0.26. (By using the remark after (6.4), we can easily see why the predicted proportion of top species is near one quarter. Because $e^{-3.72} = 0.024$, the fraction of top species is predicted to be slightly greater than one quarter.) The observed proportions are 0.19, 0.53, and 0.29. As we suggested above, if observer bias has lowered the fraction of basal species, a plausible estimate of the proportion of top and of basal species is 0.26, exactly as predicted by model 3. Thus the quantitative agreement between the predicted asymptotic mean and the observed mean is good. The model predicts a decrease in the standard deviation that is suggested by the data on basal species but that is not observed in the data on top species.
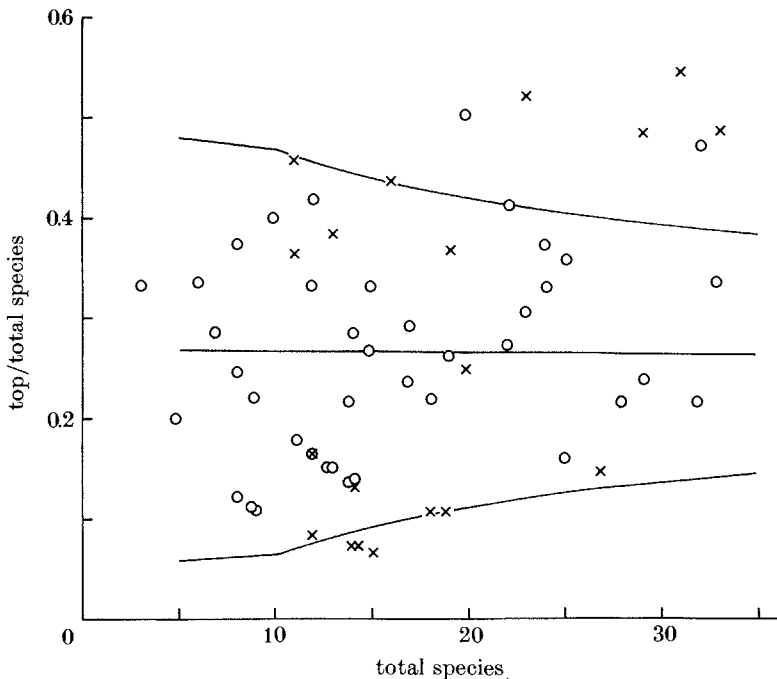
FIGURE 1. The predicted mean proportion of top species (middle line) and a confidence interval of ±2 standard deviations (upper and lower lines) as a function of total species S, according to the cascade model. In this figure and figure 2, × is constant environment, ○ is fluctuating environment. The symbols × and ○ have been perturbed from their exact locations by a small random amount to indicate when several food webs have exactly the same coordinate. The data are replotted from Briand & Cohen (1984).

In summary, model 3 predicts the form and the parameter value of the species scaling law. It is only partly successful in explaining the variation with respect to the species scaling law.

We now show that models 0 and 1, and perhaps 2 (if conjecture (5.1) is valid), predict asymptotic fractions of top or basal species that are lower than those predicted by model 3. From (3.2), (4.4) and (6.2a), we must establish that for any non-negative c (for example, $c = 1.86$), $e^{-c} \leqslant (1-e^{-2c})/(2c)$. We use $2c$ in place of c on the right of (6.2a) so that, asymptotically, models 0, 1 and 3 will all have the same effective density d^* of links. The inequality is equivalent to the inequality $c \leqslant (e^c - e^{-c})/2$, which is easily proved by noting that both sides approach 0 when $c \downarrow 0$ and by comparing derivatives of both sides with respect to c.

This inequality raises a question. In table 2, the simulated fractions of top and basal species exceeds e^{-c}. We have just shown that $(1-e^{-2c})/(2c)$ exceeds e^{-c}. Might not $(1-e^{-2c})/(2c)$, shown in table 2 under the column headed 'predictions, model 3', be a better description of the simulated fractions of top and basal species in model 2 than e^{-d^*}? Table 2 gives a weak hint that this may not be the case. Though, for $c = 4.0$, the simulated fractions of top and basal species are near those predicted by model 3, as S increases from 10 to 20 the simulated fractions move slightly away from $(1-e^{-2c})/(2c)$ and towards e^{-c}.
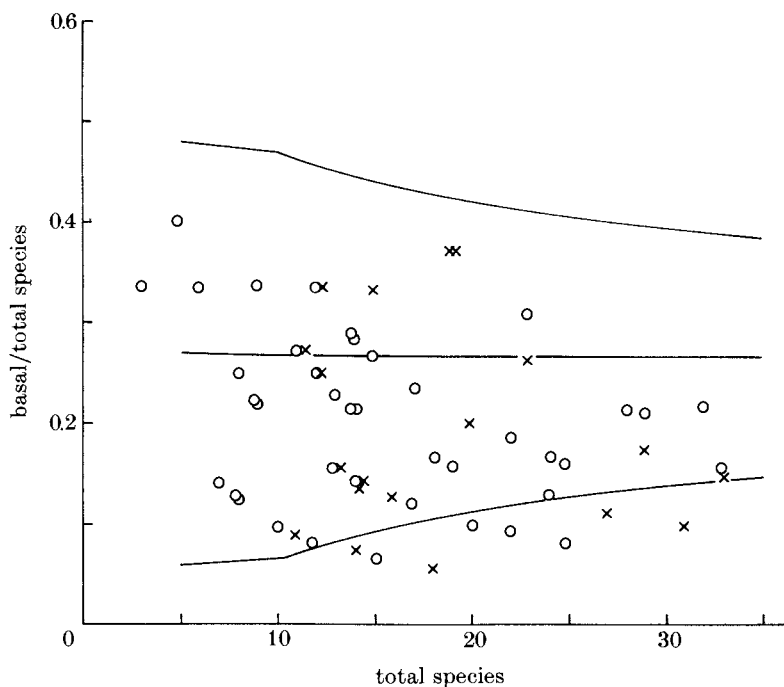
FIGURE 2. The predicted mean proportion of basal species (middle line) and a confidence interval of ± 2 standard deviations (upper and lower lines) as a function of total species S, according to the cascade model. The symbols and source of data are as in figure 1.

TABLE 3. PREDICTED MEAN AND VARIANCE IN THE PROPORTION OF TOP SPECIES, ACCORDING TO MODEL 3

S	$E(T)/S$	$[\mathrm{var}\,(T/S)]^{\frac{1}{2}}$
5	0.269	0.104
15	0.265	0.086
25	0.264	0.069
35	0.264	0.059
∞	0.262	0

Link scaling

Figure 3 shows the ratio of the expected number of links of each kind to the expected total number of links, based on (6.5) and (6.6) with $c = 3.72$, for S between 4 and 40. For $S > 10$, the ratios are effectively constant. For $S \leqslant 10$, the predicted curves for $E(L_{BI})/E(L)$ and $E(L_{IT})/E(L)$ reproduce the suggestion of a decline in the observed values of L_{BI}/L in figure 2a of Cohen & Briand (1984) and in the observed values of L_{IT}/L in figure 2d of Cohen & Briand (1984). The predicted increase in $E(L_{IT})/E(L)$ might even be reflected in the data of their figure 2c. However, few of the real webs had 10 or fewer species, so these suggestions from the data are very weak. Overall, the qualitative predictions of model 3 are consistent with the qualitative link scaling law.

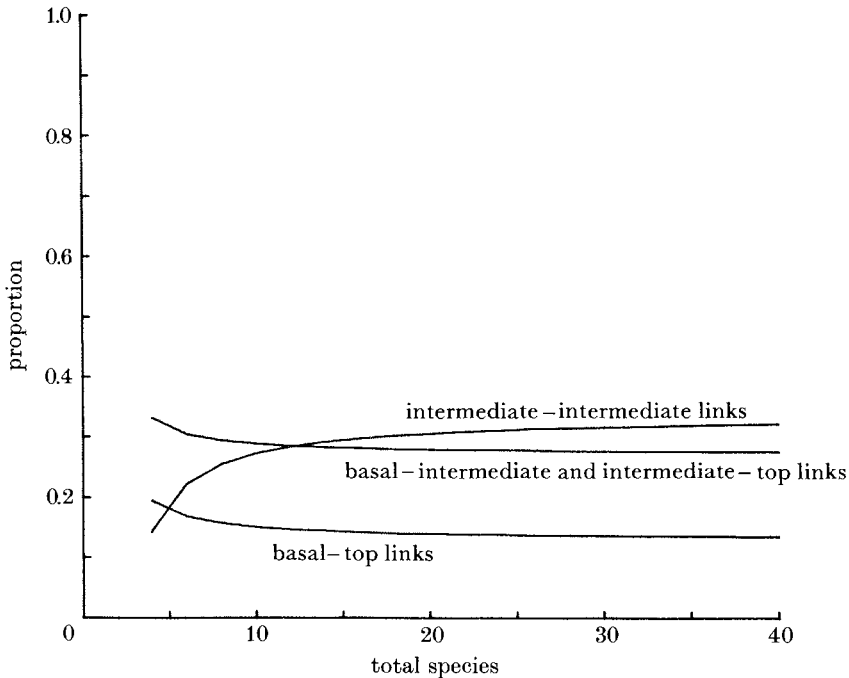Quantitatively, model 3 predicts the asymptotic proportions of each kind of

FIGURE 3. The predicted ratio of the expected number of links of each kind to the expected total number of links, according to the cascade model with $c = 3.72$, for total numbers of species $S = 4(2)40$. For $S > 10$, the ratios change little.

TABLE 4. OBSERVED PROPORTIONS OF EACH KIND OF LINK, AND ASYMPTOTIC PREDICTED PROPORTIONS ACCORDING TO MODEL 3

type of link	observed proportion	predicted proportion from (6.7) with $c = 3.72$
basal–intermediate	0.27	0.27
basal–top	0.08	0.13
intermediate–intermediate	0.30	0.33
intermediate–top	0.35	0.27

link shown in table 4. The principal discrepancy between the data and the model is that fewer basal–top links and more intermediate–top links are observed than predicted.

Link-species scaling

That model 3 correctly predicts the qualitative relation between total links and total species follows from (6.5). Quantitative agreement is guaranteed by the choice of $c = 3.72$.

Acyclicity

Acyclicity is guaranteed by making the adjacency matrices triangular.

In summary, model 3 correctly predicts the qualitative species scaling and link scaling laws in webs with more than a handful of species. Quantitatively, model 3 also predicts, to a first approximation, the observed proportions of basal, intermediate and top species and the observed proportions of each kind of link.

Sensitivity analysis

We are sceptical about the completeness of observation of trophic links, especially those that involve what are currently described as basal species. If moderately more trophic links were observed, would our quantitative predictions be radically altered? If so, the present quantitative estimates of model 3 are approximately right for the wrong reason, namely, that the effective density of links happened to be low. Thus it is important to know how the predicted asymptotic proportions of species and of links of each kind vary as c varies in the neighbourhood of its estimated value 3.72.

Figure 4 plots the predicted asymptotic proportions of basal, top, proper basal, proper top, and intermediate species among all non-isolated species, based on (6.2) and (6.4), as a function of $c = 0.5(0.5)10$. As c increases from 3.5 to 4.5, the predicted asymptotic proportions of proper basal or proper top species declines from 0.25 to 0.21 while the predicted asymptotic proportion of intermediate species among non-isolated species increases from 0.49 to 0.58. Neither range of variation seems incompatible with the data.

Figure 5 plots the predicted asymptotic proportions of links of each kind, based on (6.7), as a function of $c = 0.5(0.5)10$. As c increases from 3.5 to 4.5, the predicted asymptotic proportions of basal–intermediate or intermediate–top links declines from 0.27 to 0.25, the proportion of basal–top links declines from 0.14 to 0.09, and
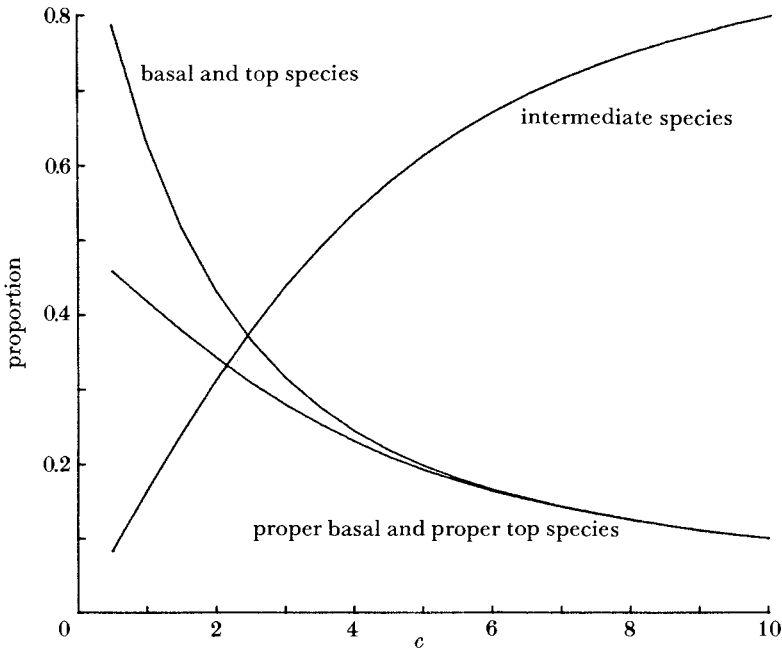

FIGURE 4. The predicted asymptotic proportions of basal, proper basal, top, proper top, and intermediate species, as a fraction of non-isolated species, according to the cascade model, for $c = 0.5(0.5)10$. Because basal and top species are plotted here as a fraction of non-isolated species, the sum of proportions (basal + intermediate + top) exceeds 1. The excess over 1 is small once $c > 3$. The sum of proportions (proper basal + intermediate + proper top) equals 1.
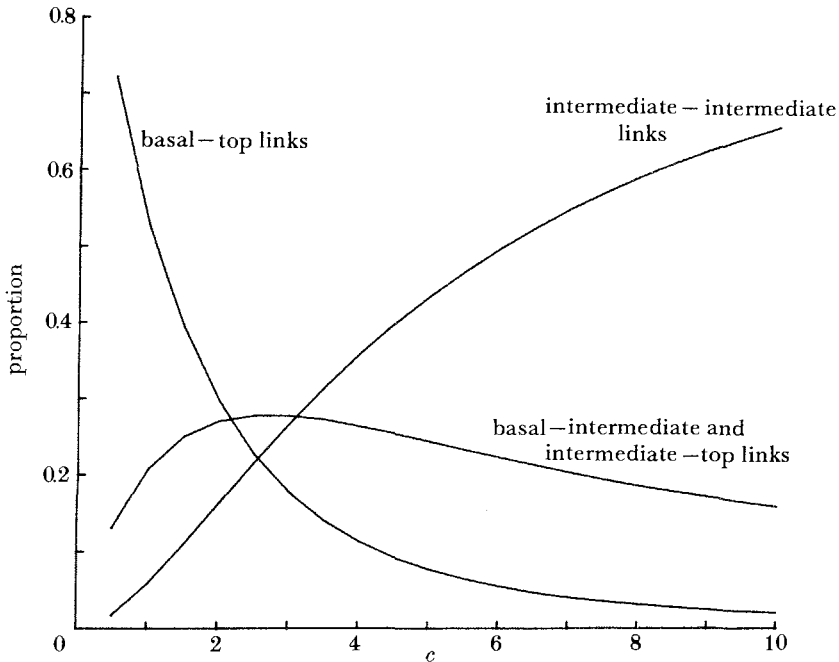
FIGURE 5. The predicted asymptotic proportions of links of each kind, according to the cascade model, for $c = 0.5(0.5)10$.

the proportion of intermediate–intermediate links increases from 0.31 to 0.40. Such changes improve the agreement between the observed and predicted proportions of basal-top links but worsen the agreement between the observed and the predicted proportions of the remaining classes of links. However, the changes in the predicted asymptotic proportions are not very radical in any case. In particular, the estimate of $c = 3.72$ happens to fall very near where the curve for basal–intermediate and intermediate–top links is flattest.

We conclude that the predicted asymptotic proportions of species and links of each kind are not so sensitive to the exact value of the observed ratio of links to species as to exclude the possibility of a somewhat greater effective density of links.

Lumping

Would lumping substantially alter the number of species and hence the proportions of interest in the cascade model? The same approach used to analyse lumping in model 0 shows that, for $i < j$, $P(\text{lump } i \text{ and } j) = (1-2pq)^{S+i-j-1}q^{2(j-i)-1}$ while $P(\text{lump non-isolated } i \text{ and } j) = (1-2pq)^{S+i-j-1}q^{2(j-i)-1} - q^{2(S-1)-1}$. The expected fraction of species that are not isolated and lost by lumping is then less than or equal to

$$(1/S) \sum_{j=2}^{S} \sum_{i=1}^{j} P(\text{lump non-isolated } i \text{ and } j)$$

$$= (q^{2S-3}/[Sr])[(S-1)(1+r)^{S-1} - \{(1+r)^{S-1}-1\}/r - S(S-1)r/2], \quad (6.10)$$

where $r = (p/q)^2$. As $S \to \infty$, (6.10) approaches $c^2 e^{-2c}/3 = 0.003$ when $c = 3.72$. In model 3 as in model 0, the effect of lumping non-isolated species is negligible.

Effect of disconnected components

The effect of weak components is essentially identical in models 0 and 3. The calculation based on (3.4) remains the same, with the parameter c of (3.4) still estimated by $d = 1.86$ rather than by $2d$. As in model 0, asymptotically all but a negligible fraction of species belong to the largest weak component.

So far, we have taken c as exogenously determined, for example, by the feeding apparatus or behavioural flexibility of species, and have attempted to predict other structural features of webs from that parameter. Why might c assume a value in the vicinity of 3.72? Figure 5 shows that $c = 3.72$ is in the range around 2.69 where the predicted asymptotic proportions of basal–intermediate and intermediate–top species are maximal. It is tempting to speculate, but without theoretical or additional empirical support at the moment, that the effective density of links is adjusted to maximize the proportion of links between basal and intermediate species, and between intermediate and top species.

7. CONCLUSIONS

In this section, we shall first summarize the conclusions we draw from the four models we have considered. We then relate our results to some earlier efforts to model webs. Finally, we mention two important limitations on our results.

Briand (1983), using 'unlumped' webs, first suggested, and Cohen & Briand (1984), using 'lumped' webs, demonstrated that the average total links of a web are nearly proportional to the total species of the web. Within the framework of the random digraph models considered here, this observation has the important implication that the probability of a given species eating or being eaten by another given species must vary as the reciprocal of the total number of species in the web. This has the further consequence that the number of predators or prey of a randomly chosen species is asymptotically independent of the total number of species in the web.

The exclusion of cycles of finite lengths or of all lengths as S increases is insufficient to reproduce quantitatively the species scaling law, although an open mathematical question remains in the analysis of model 2. That question is: when $c \geqslant 1$, what is the asymptotic mean fraction of zero columns or of zero rows in a random $S \times S$ matrix whose elements are independently and identically distributed Bernoulli random variables with mean c/S, conditional on the matrix being acyclic?

To explain the observed proportions of top and basal species, it appears to be necessary to suppose that there is an ordering, hierarchy, or cascade of species that constrains the possible predators and prey of each species. Under this assumption, it is possible to predict qualitatively, and to fair approximation quantitatively, the species scaling law and the link scaling law, by using a single parameter from the data, the ratio of total links to total species.

In evaluating the quantitative discrepancies between the observed and predicted proportions of each kind of species and each kind of link, it is important to recall that no fitting is involved in generating the predicted proportions. The only numerical parameter taken from the data is the observed ratio of the total number

of links to the total number of species. In addition to its qualitatively correct predictions, model 3 gives seven numbers for the price of one. (Of these seven, only five are independent: two of the three proportions of kinds of species, and three of the four proportions of kinds of links.)

The gross testing presented here demonstrates that the overall proportions of species or links are consistent with the predictions of model 3. The following paper (Cohen *et al.* 1985) examines how well model 3 describes individual webs.

Cohen (1978, pp. 58–61) considered six stochastic models of webs that are similar to those considered here. His model 6 models the adjacency matrix ('food web matrix') of a web with m prey and n predators by constructing an $m \times n$ matrix in which each element equals 1 with probability $L'/(mn)$ and equals 0 with probability $1 - L'/(mn)$, where L' is the observed number of links, independently for all elements. This model 6 is similar to model 0 here, but model 6 limits the number of prey to m and the number of predators to n. Model 0 here allows the adjacency matrix to be $S \times S$ so that the numbers of prey and predators are limited only by S. None of the models of Cohen (1978) rules out cycles (like our models 1 and 2) or imposes a cascade structure (like our model 3).

Lawlor (1978) observed that in randomly constructed matrix models of ecosystems, when the probability of a non-zero entry in the matrix is independent of the number of species, an overwhelming majority have 3-cycles if the number of species increases beyond 20 (contrary to his and others' informal observations that such cycles are rare in real webs). However, when the probability of a non-zero entry varies inversely as the number of species (as we suppose in this paper, on the basis of the link-species scaling law), Lawlor found (without giving the details of the calculations) that the proportion of random matrix models without 3-cycles increases with increasing numbers of species. He concluded that the usefulness of 'random' models of ecosystems depends critically on whether the models possess the specific structural patterns characteristic of real ecosystems. This conclusion we share.

We are aware of at least two major limitations of the scope of the models and data we have investigated here. First, we have dealt only with the combinatorial structure of webs, rather than with quantities of stocks and flows. Our approach is more like gross anatomy than like physiology. Second, we have dealt only with a static snapshot of webs, ignoring cyclical, successional, or other changes. That is, the gross anatomy is frozen, rather than in motion. In spite of these important limitations, we have provided, in the cascade model, a unifying perspective of simplicity and potential usefulness.

APPENDIX 1: PROOF OF THEOREM 1

In this proof, we shall omit the explicit dependence on S where possible; for example, we replace $Y(S)$ by Y, $M(S)$ by M. Let $C^{(k)}$ be the set of possible distinct k-cycles. For $s \in C^{(k)}$, let $B_s^{(k)} = 1$ if cycle s occurs in (the web specified by) the random adjacency matrix A, $B_s^{(k)} = 0$ if s does not occur. Then the number of k-cycles in A is $M_k = M_k(S) = \Sigma_{s \in C^{(k)}} B_s^{(k)}$.

Since the random variables $\{a_{i1}\}_{i=1}^S$ are independent with $E(a_{i1}) = c/S$, it is a standard fact that Y converges in distribution to a Poisson variable with mean C.

Let $\#(.)$ denote the cardinality (number of elements) of the set in parentheses. Then

$$\#(C^{(k)}) = S!/[(S-k)!k], \quad E(B_s^{(k)}) = (c/S)^k$$

so that

$$\#(C^{(k)}) E(B_s^{(k)}) \to c^k/k \quad \text{as} \quad S \to \infty.$$

The random variables $\{B_s^{(k)}\}_{s \in C^{(k)}}$ are non-decreasing functions of the independent elements $\{a_{ij}\}$ of A and hence are associated. (Recall that a finite family $\{X_1, \ldots, X_n\}$ of random variables is defined to be *associated* if cov $(f(X_1, \ldots, X_n), g(X_1, \ldots, X_n)) \geqslant 0$ for any real functions f and g that are coordinate-wise increasing.) A theorem independently discovered by Wood (1982) and Newman *et al.* (1984) (and stated as theorem 11 by Newman (1984)) then implies that M_k converges in distribution to a Poisson variable with mean c^k/k, for $c \geqslant 0$, $k \geqslant 1$, provided that

$$\lim_{S \to \infty} \Sigma \, \text{cov} \, (B_s^{(k)}, B_{s'}^{(k)}) = 0, \tag{A 1}$$

where the summation extends over pairs $s, s' \in C^{(k)}$ such that $s \neq s'$.

Similarly, according to theorem 10 of Newman (1984), which is taken from Newman (1980), (4.3) holds if, in addition,

$$\lim_{S \to \infty} \text{cov} \, (Y, M_h) = 0, \tag{A 2a}$$

$$\lim_{S \to \infty} \text{cov} \, (M_h, M_j) = 0, \tag{A 2b}$$

for all $c \geqslant 0$ and all h, j such that $1 \leqslant h \neq j \leqslant k$.

So we must prove (A 1) and (A 2).

For $k = 1$, as noted in the text, each $B_s^{(k)}$ is just an a_{ii} so that cov $(a_{ii}, a_{jj}) = 0$ for $i \neq j$ and (A 1) holds. Also, for $k = 1$, cov $(Y, M_1) = \text{cov} \, (a_{11}, a_{11}) = (c/S)(1-c/S)$ so that (A 2a) holds for $h = 1$. Similarly cov $(M_1, M_j) = 0$ for $j \neq 1$. We may henceforth assume $k, h, j \geqslant 2$.

Unless the two cycles, s and $s' \neq s$ share some directed edge, $B_s^{(k)}$ and $B_{s'}^{(k)}$ are independent. Similarly, a_{i1} and $B_s^{(k)}$ are independent unless the edge $(i, 1)$ is in s. Since, for $k > 1$, a_{11} and $B_s^{(k)}$ are independent,

$$\text{cov} \, (Y, M_k) = \sum_{i=1}^{S} \sum_{s \in C^{(k)}} \text{cov} \, (a_{i1}, B_s^{(k)})$$

$$= (S-1) \sideset{}{'}\sum_{s \in C^{(k)}} \text{cov} \, (a_{21}, B_s^{(k)}) + \sum_{s \in C^{(k)}} \text{cov} \, (a_{11}, B_s^{(k)})$$

$$= (S-1) \sideset{}{'}\sum_{s \in C^{(k)}} \text{cov} \, (a_{21}, B_s^{(k)}),$$

where Σ' is over those cycles s that include the edge $(2, 1)$. There are exactly $(S-2)!/(S-2-(k-2))!$ such k-cycles. If $i_1 = 2$, $i_2 = 1$, i_3, $\ldots$, i_k are the vertices of such a cycle, with $i_{k+1} = i_1 = 2$, then

$$\operatorname{cov}(a_{21}, B_s^{(k)}) = \operatorname{cov}(a_{21}, a_{21} \prod_{j=2}^{k} a_{i_j i_{j+1}}) = (c/S)^{k-1} \operatorname{var}(a_{21}) = (c/S)^k (1 - c/S).$$

$$\text{(A 3)}$$

Thus

$$\operatorname{cov}(Y, M_k) = [(S-1)!/(S-k)!] (c/S)^k (1 - c/S) \to 0 \quad \text{as} \quad S \to \infty,$$

which proves (A 2 a).

Suppose s is an h-cycle and s' is a j-cycle. Let $\beta = \beta(s, s')$ denote the number of edges shared in both s and s'. Analogously to (A 3), we have

$$\operatorname{cov}(B_s^{(h)}, B_{s'}^{(j)}) = (c/S)^{h-\beta} (c/S)^{j-\beta} \operatorname{var}(\prod_{g=1}^{\beta} a_{i_g i_{g+1}})$$

$$= (c/S)^{h+j-2\beta} [(c/S)^\beta - (c/S)^{2\beta}]$$

$$= (c/S)^{h+j-\beta} [1 - (c/S)^\beta] = O(S^{-h-j+\beta}).$$

$$\text{(A 4)}$$

Then, for some fixed $s_0 \in C^{(k)}$, and for $C_1 = \{(s, s') \mid s \neq s'\}$, $C_2 = \{s \in C^{(k)} \mid s \neq s_0\}$,

$$\sum_{C_1} \operatorname{cov}(B_s^{(k)}, B_{s'}^{(k)}) = \#(C^{(k)}) \sum_{C_2} \operatorname{cov}(B_{s_0}^{(k)}, B_s^{(k)})$$

$$= \#(C^{(k)}) (\Sigma^0 + \Sigma^1 + \ldots + \Sigma^{k-1})$$

$$= \#(C^{(k)}) (\Sigma^1 + \ldots + \Sigma^{k-2})$$

$$\text{(A 5)}$$

where Σ^β denotes $\sum \operatorname{cov}(B_{s_0}^{(k)}, B_s^{(k)})$ over those $s \in C^{(k)}$ such that $\beta(s_0, s) = \beta$. The last equality in (A 5) holds because there can be no $s \in C^{(k)}$ with $\beta(s_0, s) = k-1$ and because each term in the Σ^0 sum vanishes.

Now the removal of $k - \beta > 0$ edges from a k-cycle s_0 leaves some number $\eta \geqslant 1$ of disconnected walks. A cycle $s \in \Sigma^\beta$ must reconnect these walks into a cycle (in an order that may differ from the order in s_0). Thus s is specified by the order of the walks shared with s_0 and by the η new walks leading from an end point of one shared walk to a starting point of another shared walk. If these η new walks have lengths L_1, $\ldots$, L_η, with $L_1 + \ldots + L_\eta = k - \beta$, then for a given ordering of the shared walks, the number of such new walks is bounded above by $S^{L_1-1} S^{L_2-1} \ldots$ $S^{L_\eta-1} = S^{k-\beta-\eta}$ and thus by $S^{k-\beta-1}$. So for $\beta < k$ the number of terms in any Σ^β is bounded by $S^{k-\beta-1}$ times a combinatorial coefficient that depends only on k and β but not on S. By using (A 4) with $h = j = k$, we may bound (A 5) above by

$$\#(C^{(k)}) \sum_{\beta=1}^{k-2} O(S^{-2k+\beta} S^{k-\beta-1}) = O(S^k S^{-k-1}) = O(S^{-1}),$$

$$\text{(A 6)}$$

which proves (A 1).

We now prove (A 2 b). As in (A 5), for $h < j$,

$$\operatorname{cov}(M_h, M_j) = \#(C^{(h)}) \sum_{s \in C^{(j)}} \operatorname{cov}(B_{s_0}^{(h)}, B_s^{(j)})$$

$$= \#(C^{(h)}) (\Sigma^0 + \Sigma^1 + \ldots + \Sigma^h)$$

$$= \#(C^{(h)}) (\Sigma^1 + \ldots + \Sigma^{h-1}).$$

$$\text{(A 7)}$$

In the last equality, $\Sigma^h = 0$ because no j-cycle can share h edges with the h-cycle s_0 if $h < j$. As in the derivation of (A 6) as an upper bound for (A 5), we see that with $h - \beta > 0$, the number of terms in Σ^β is $O(S^{j-\beta-1})$. Then (A 4) implies that (A 7) is bounded by

$$\#(C^{(h)}) \sum_{\beta=1}^{h-1} O(S^{-h-j+\beta} S^{j-\beta-1}) = O(S^h S^{-h-1}) = O(S^{-1}),$$

which proves (A 2b).

The claimed limiting behaviour of $(Y(S), M(S))$ for $0 \leqslant c < 1$ now follows from (4.3) by approximating $M(S)$ by $M_k^*(S) = \Sigma_{h=1}^k M_h(S)$ for large fixed k. For fixed k, (4.3) implies that $(Y(S), M_k^*(S))$ converges in distribution as $S \to \infty$ to a 2-vector with independent Poisson components and mean $(c, \Sigma_{h=1}^k c^h/h)$. Moreover,

$$E \,|\, M(S) - M_k^*(S) \,|\, = E(M(S) - M_k^*(S)) = \sum_{h=k+1}^{\infty} \#(C^{(h)}) \, E(B_{s_0}^{(h)})$$

$$\leqslant \sum_{h=k+1}^{\infty} c^h/h$$

$$\to 0 \quad \text{as} \quad k \to \infty \quad \text{for} \quad c < 1.$$

Now for any real numbers r and t

$$|\, E(\exp\{i(r\,Y(S) + tM(S))\}) - E(\exp\{i(r\,Y(S) + tM_k^*(S))\}) \,|$$

$$\leqslant E(|\exp\{it(M(S) - M_k^*(s))\} - 1|)$$

$$\leqslant E|t(M(S) - M_k^*(S))| \to 0 \quad \text{as} \quad k \to \infty.$$

Therefore, the limiting distribution of $(Y(S), M(S))$ equals the limiting distribution, as $k \to \infty$, of the limiting distribution, for any fixed k, as $S \to \infty$, of $(Y(S), M_k^*(S))$. This proves the claimed results when $0 \leqslant c < 1$.

APPENDIX 2: NUMERICAL SIMULATION OF ACYCLIC RANDOM DIGRAPHS

We have programmed three numerical approaches to investigating the fraction of zero rows or columns in a matrix that is $\sim$ i.i.d. $B(c/S)$, conditional on being acyclic.

The first, and most naive, approach is to generate matrices that are $\sim$ i.i.d. $B(c/S)$ and reject those that have a cycle of any length. There are two difficulties with this approach. First, given a value of c, this approach generates acyclic webs very inefficiently. For example, with an arbitrarily chosen $c = 2.1$, the number of Bernoulli matrices that had to be generated to find 100 acyclic matrices of each size in a sample calculation was as shown in table A 1.

We lack theory for what the numbers on the right of the table should be, either for finite S or in the limit as $S \to \infty$. (These results show, incidentally, first that the fraction of acyclic matrices among Bernoulli matrices, for fixed c, need not be a monotone decreasing function of S, and second that the fraction of 10-acyclic Bernoulli matrices, asymptotically as $S \to \infty$, according to theorem 1, bears no close relation to the fraction of 10-acyclic 10×10 Bernoulli matrices. According to (4.3), the former fraction is $\exp(-\Sigma_{k=1}^{10} c^k/k)$, which is less than 10^{-157} when

TABLE A 1. MATRICES GENERATED ACCORDING TO A FIRST NAIVE APPROACH

size of matrix (S)	number of matrices generated to get 100 acyclic $S \times S$ matrices
3	34 972
5	16 113
10	28 726
15	62 825
20	279 401

$c = 2.1$, while according to the numerical results above the latter fraction is approximately $100/28\,726$.)

A second difficulty with this first approach is that, so far, we lack theory to guide the choice of c when we want to compare the computed fractions of zero rows or columns with data. By throwing away the matrices with cycles, we change the expected number of matrix elements that equal 1 from $pS^2 = cS$ to some (so far) unknown smaller function of c and S. For comparison with data, we want to choose c so that the 'effective density' of links, estimated as (average number of matrix elements equal to 1)/S, equals the observed $d = 1.86$. In the numerical simulations described above, with $c = 2.1$, the total number of elements equal to 1, summed over 100 acyclic matrices, and the average effective density per matrix, were as shown in table A 2.

TABLE A 2. NUMBER AND EFFECTIVE DENSITY OF LINKS IN NAIVELY GENERATED
ACYCLIC MATRICES

size of matrix (S)	number of 1s in 100 acyclic matrices	effective density
3	239	0.80
5	537	1.07
10	1293	1.29
15	2124	1.42
20	3094	1.55

Depending on the matrix size S, the effective density can be quite different from c in model 2. Again, we lack theory for what the numbers on the right should be, either for finite S or as $S \to \infty$.

A second approach, based on the ideas of Erdös & Rényi (1960) avoids both of these difficulties, but encounters a subtler third difficulty. In this approach, to obtain an effective density c, we construct a random acyclic matrix with the integer part of cS (denoted int (cS)) edges. This is impossible if $cS > S(S-1)/2$ (or more generally if cS exceeds the maximum number of links possible in an $S \times S$ acyclic matrix). Provided int (cS) is sufficiently small, we add one edge at a time. We choose a 0 element of the matrix, with probability equal to 1 divided by the number of 0 elements that could be changed to 1 without creating a cycle. To identify the 0 elements that are available to be changed to 1 without creating a cycle, we maintain in an auxiliary matrix the transitive closure of the adjacency matrix. We continue adding edges until int (c/S) edges have been added. If, because of the

sequence of edges chosen, the required number of 1's cannot be added to the matrix, then the partly completed matrix is abandoned and a fresh start is made. This generates a random acyclic matrix with effective density close to c.

The virtue of this second approach is that it guarantees $L/S = c = d^* = d$ approximately (recall that c is the model parameter with c/S being the probability of a random link, d^* is the asymptotic (large S) effective ratio of links to species, and d is the observed ratio L/S of links to species in real webs). A drawback, which we overlooked at first, is that this approach does not generate all random digraphs with S vertices and, say, E (always directed) edges with equal probability. In the probability distribution over digraphs assumed by model 2, any two digraphs with S vertices and E edges occur with equal probability. However, in the numerical approach just described, suppose $S = 6$ and we wish to choose randomly $E = 3$ edges. There are $6 \times 5 = 30$ ways to choose the first directed edge without creating a loop. Suppose, without loss of generality, that the edges are labelled so that the first edge is $(1, 2)$, that is, the edge goes from vertex 1 to vertex 2. There are then 28 ways to choose the second edge (edge $(1, 2)$ has already been chosen and edge $(2, 1)$, which would create a cycle, is forbidden). If the second edge is, for example $(3, 4)$, then there are 26 ways to choose the third edge. But if the second edge is $(2, 3)$, then there are only 25 ways to choose the third edge because two edges have already been chosen and *three* edges are forbidden ($(2, 1)$, $(3, 2)$ and $(3, 1)$).

Our third approach modifies the procedure just described to avoid this difficulty. As each randomly chosen edge is added to a digraph, the number of available edges that could have been chosen at that stage is noted. The product of all the numbers of available edges is assigned to the generated digraph as a weight. This weight is the inverse of the probability of choosing the edges in the particular random digraph *in the order in which the edges occurred*. The weight assigned to a given digraph may vary depending on the order in which the edges are chosen. All the statistics (such as the mean or variance of the fraction of species that are top or basal) computed from the random digraphs generated according to this third approach incorporate the weights, so that all digraphs with a given number of vertices and edges are represented with equal probability.

When the unweighted simulations based on the second approach are compared with the weighted simulations based on the third approach, the simulated mean fractions of 0 rows and columns were generally slightly larger when weighted, but usually not by more than 0.01 and never (for the range of parameters in table 2) by more than 0.04. A conjecture that for large S and for c small compared to S the two approaches give identical mean proportions of 0 rows and columns seems plausible.

The simulations based on the second and third approach are not identical to those based on the first, naive approach. There is no variation in the number of edges (links) per acyclic digraph generated according to the second or third approach, while there is variation in the number of edges per acyclic digraph generated naively by the first approach. As in the parallel case of undirected graphs considered by Erdös & Rényi (1960), we expect (but have not proved) that this difference in approach to simulating model 2 has no effect in the limit of large S.

REFERENCES

Briand, F. 1983 Environmental control of food web structure. *Ecology* **64**, 253–263.

Briand, F. & Cohen, J. E. 1984 Community food webs have scale-invariant structure. *Nature, Lond.* **307**, 264–266.

Cohen, J. E. 1977 Ratio of prey to predators in community food webs. *Nature, Lond.* **270**, 165–167.

Cohen, J. E. 1978 *Food webs and niche space.* Princeton: Princeton University Press.

Cohen, J. E. & Briand, F. 1984 Trophic links of community food webs. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4105–4109.

Cohen, J. E., Newman, C. M. & Briand, F. 1985 A stochastic theory of community food webs. II. Individual webs. *Proc. R. Soc. Lond.* B **224**, 449–461.

Erdös, P. & Rényi, A. 1960 On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kut. Int. Kozl.* **5**, 17–61. Reprinted in *The art of counting: selected writings* (ed. J. Spencer), pp. 574–617. Cambridge: M.I.T. Press (1973).

Gallopin, G. C. 1972 Structural properties of food webs. In *Systems analysis and simulation in ecology*, vol. 2 (ed. B. C. Patten), pp. 241–282. New York: Academic Press.

Harary, F. 1961 Who eats whom? *Gen. Systems* **6**, 41–44.

Lawlor, L. R. 1978 A comment on randomly constructed model ecosystems. *Am. Nat.* **112**, 445–447.

Newman, C. M. 1980 Normal fluctuations and the FKG inequalities. *Comm. Math. Phys.* **74**, 119–128.

Newman, C. M., Rinott, Y. & Tversky, A. 1984 Nearest neighbors and Voronoi regions in certain point processes. *Adv. appl. Probab.* **15**, 726–751.

Newman, C. M. 1984 Asymptotic independence and limit theorems for positively and negatively dependent random variables. In *Inequalities in statistics and probability* (ed. Y. L. Tong). Institute of Mathematical Statistics Lecture Notes – Monograph Series, vol. 5, pp. 127–140.

Pimm, S. 1982 *Food webs.* London: Chapman and Hall.

Robinson, D. F. & Foulds, L. R. 1980 *Digraphs: theory and techniques.* New York: Gordon and Breach.

Wood, T. E. 1982 Sequences of associated random variables. Ph.D. dissertation, University of Virginia.