# Size Distribution of Ecological Niches

Jason M. Lewis,[1, *] Charo I. Del Genio,[1, †] and Thilo Gross[2, ‡]

[1]*Centre for Complexity Science, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK*
[2]*Department of Engineering Mathematics, University of Bristol, University Walk, Bristol BS8 1TR, UK*

(Dated: July 1, 2014)

We introduce a model of ecological systems based on random networks generated using the directed Erdős-Rényi model. Species are represented as nodes with Boolean state denoting their viability within the ecosystem, and interspecies dependence is encoded as one of two edge types between nodes; either necessity or incompatibility. We study the size distribution of the remaining viable subset of species under variation of both increased connectivity and dependence type. We find two exponential relationships describing the variation of viable species subset size with the parameters of the model. One of these is directly connected to the giant component phase transition in Erdős-Rényi random graphs and the connectance threshold. Ecological interpretations of these relationships are explored and possibilities for direct extension of the model are discussed.

## I. INTRODUCTION

A biological ecosystem consists of a collection of species interacting with one another in conjunction with their environment [1, 2]. In this sense one can define an ecosystem as the network of interactions between all constituent species and their environment, hence dependency between species is a key factor for the structure of an ecosystem - nothing exists in isolation [2–4]. An interaction between species is often classified by what benefit or detriment (or neutrality) has been caused to each interacting member. For example, predation is of benefit to the antagonistic organism and detriment to the prey, or mutualism, such as pollination, which is beneficial to both species [5–8].

Modelling approaches to understanding ecosystems have been used for quite some time. A famous example is of Lotka [9] and Volterra [10] from the mid-1920s who used a pair of coupled ordinary differential equations to model fluctuations in shark and fish populations in the Adriatic Sea, but more generally modelling predator-prey interactions (and later a ratio-based model of this type by Arditi and Ginsberg in 1989 [11]). In recent years, biological systems are typically modelled with respect to the specific interactions between specific constituents, and attention is drawn to the dynamics of such systems as a whole, or under introduction of external factors. This is often done using an individual-based approach where populations, communities, and ecosystems are modelled by looking at the behaviour of the individuals to determine the behaviour of the whole system. This approach has been successful not only in the field of ecology, but in epidemiology, social science, and far beyond [12–14].

We present a simple model (overview in Section III which pays no heed to the dynamics or specifics of a certain ecosystem, but instead looks to see how a general population of species can persist under a set of simple rules describing their connections. The aim is to determine what fraction of an initial population of species will be viable based on their interdependencies. The distinction between looking at a general model of species viability, and not a specific ecosystem, is useful in looking at this sort of question as we look at *typical* behaviour of ecosystems rather than being interested in the population fluctuation and dynamics of individual species. It also allows us to deal with the problem of complexity: ecosystems, food webs, and biological communities in general are highly interconnected and multi-scaled, and are a good example of a complex system [15–18]. It can be difficult to classify and describe the interactions between constituents and simplifications in the understanding of these connections are required to model them, so it can be hard to look at large systems in this way. As we are interested in how behaviour scales with system size, and not the dynamics of a specific system, a simplified approach based on a specified type of interconnection and ensemble averaging allows us to make this type of query. This level of generality also speaks of the extendibility of the model, both via modification to the rule-set and for application to different fields in terms of viability modelling; possibly in finance where one has a landscape of interdependent entities with various motives of financial exploitation and symbiosis.

## II. BACKGROUND

In this section we describe the model with context via description of the relevant background material. In the section after this, we provide an outline of the model for clarity, as some readers may or may not be familiar with the concepts discussed below.

We aim to represent the ecosystem as a random network with species represented as nodes and links corresponding to interdependencies between them. The via-

---

\* Electronic address: J.M.Lewis@warwick.ac.uk
† Electronic address: C.I.del-Genio@warwick.ac.uk
‡ Electronic address: Thilo@biond.org

bility of each species is encoded as the state of its node: if viable, 1, and if not, 0. We begin with all species as viable and iterate through the set of dependencies, updating each node as appropriate, until all rules are satisfied - that is no states are changed. At the end of this process a subset of species will remain, and we are interested how the fraction of this subset from the initial set scales with system size, and this is investigated for different types of dependency between species.

### A. Erdős-Rényi Model

The random networks are generated using the Erdős-Rényi model for random graphs so we will first cover some background on this subject. A graph is a collection of vertices $V$ and edges $E$ between them (and we denote the number of nodes $N = |V|$) [18, 19]. To represent these it is useful to construct an adjacency matrix $A$ of dimension $N \times N$ whose entries correspond to the number of edges between nodes; i.e. $A_{ij} = 1$ denotes one edge from node $i$ to node $j$. A directed Erdős-Rényi random graph [20] is generated by building adjacency matrices with each entry as a Bernoulli random variable, that is a value 1 with probability $p$ and 0 otherwise [21], with zero diagonal entries (i.e. no self edges):

$$A_{ij} = \begin{cases} 0 & \text{if } i = j \\ \text{Bern}(p) & \text{otherwise} \end{cases} \quad (1)$$

Therefore the existence of an edge is independent of all other edges in this model [22]. Note that this network has no multiple edges ($A_{ij}$ can only be zero or one). Using a probabilistic model like this (parameterised by the link probability $p$) means that we will average over an ensemble, which will be a sample of the set of all graphs with a given parameter, therefore this approach is an approximation which improves as the ensemble size increases.

### B. Encoding Interdependence

In order to impose dependency relations on this network we label the edges according to dependence type. In this model we make a simplification for the types of interdependence between species, namely that there are two types:

i) *Necessity:* species A requires the presence of species B to be viable.

ii) *Incompatibility:* species A is only viable if species B is not viable.

This presents the strongest type of dependence, where one species is directly and absolutely determining whether another is viable or not, within a specific ecosystem. One can interpret these sorts of dependency

in a number of ways, for example incompatibility may be the existence of a predator completely causing the prey to go extinct, but also may describe a situation of one species out-competing another to deny its viability in the ecosystem. If we recall the framework for dependence as classification by benefit/detriment/neutrality [5–8] from Section I, then this type of network can encode this interdependence (neutrality via the lack of an edge between nodes), albeit in the strongest sense.

We label edges similarly to how we build network links; If a link exists then we assign it of type necessity with probability $q$, and of type incompatibility otherwise. Therefore it is once again determined via in independent Bernoulli trial with parameter $q$. It is also clear that for $N$ species one would expect there to be $Np$ links, and hence $Npq$ links of type necessity, as expected values (and other key properties such as variance and so on) are well known for this distribution.

The model as described in this way can also be thought of as a type of Boolean satisfiability problem (SAT). We have a set of species with general dependencies between them, which we can interpret as rules/constraints or Boolean formulae, and we are sequentially searching to see the outcome which satisfies these rules. SAT is a NP-complete problem [23, 24], therefore the solution cannot be quickly determined, hence our approximation-based approach is appropriate. Also algorithmically it restricts our approach to iterating through the rule-set and checking if updates need to be made to the state of any species exhaustively, until no changes need to be made.

### C. Random Boolean Network

The model as described can also be considered as a random Boolean network (RBN), already of biological fame in the area of genetic regulatory networks [25–27]. Each of the nodes may take one of two values (0 or 1) and the state of each node is determined by a Boolean function. Iteration is done in discrete steps, such that the state of the system at $t + 1$ is entirely determined by the state at $t$. These functions, with the dependencies written in this way, are shown in Table I.

| Type | $t$ | | $t+1$ | |
|---|---|---|---|---|
| | **A** | **B** | **A** | **B** |
| Necessity | 1 | 0 | 0 | 0 |
| Incompatibility | 1 | 1 | 0 | 1 |

TABLE I: Representation of species dependency rules as Boolean functions.

Of course, due to the biological significance of the node states, this is a Boolean network with a twist: if a species is non-viable, it remains non-viable and cannot resurrect. We also consider the fact that our initial condition (naturally) is all species viable, which implies there are only

two functions which can alter the state of a node. The second difference is that this system has a guaranteed fixed-point attractor: we are guaranteed to converge to some subset of species in a fixed number of iterations. We will either have no species left at some stage, or there will be a point where no updates have been performed to the state of the nodes. This is again a consequence of the one-directional nature imposed by non-viability, and we cannot see the cellular-automata-style dynamics typical of RBNs since there are no limit cycles that can exist [28, 29]. We update the state of each node at the same time, hence this RBN would be termed synchronous or classical [28].

Finally, there is an additional way of looking at this sort of model. One could consider the network of necessity and incompatibility links to be separate from one another and treat this system as a multiplex network [30–33] with two layers. This could allow for some interesting possibilities of extending the model via introduction of new interaction types as new layers however does have the downfall of possibly unnecessarily obfuscating the simplicity of the model. Also if one were to extend this model it would be more natural to allow the current strong rules to be weakened in a tunable way, governed by a single parameter, rather than introducing large number of distinct rules (more in Section V.

## III.   OUTLINE OF MODEL/METHODOLOGY

The model we propose describes ecosystems as random networks, where species are nodes and the dependency relations between them are links. Networks are built as directed Erdős-Rényi random graphs and links are labelled as either necessity or incompatibility type. The viability of each species is held as a Boolean variable and sequentially updated based on these links until no change can occur. For a given system, the model is parametrised by two variables: $p$, the probability of a link between two species, and $q$, the probability each link is of type necessity. These parameters vary continuously between 0 and 1. The variable of interest is system size $N$, that is the number of species in the ecosystem, and system properties can be investigated via ensemble averaging.

From a technical perspective, the model was coded in C, as we are interested in performing a large number of simple operations quickly. We are generating ensembles of matrices, with $N(N-1)$ elements defined by a random number. A certain number of these (edge existing) require a random number to query the type of link to assign. Therefore we require a good pseudo-random number generator as we could run into trouble with low periodicity generators. We use the Mersenne twister algorithm [34] with a fixed seed for reproducibility. This is used in a number of scientific libraries as it has a very long period and passes numerous tests for randomness, so should be suitable here. We pull a random number from the uniform distribution on the interval [0,1] and make
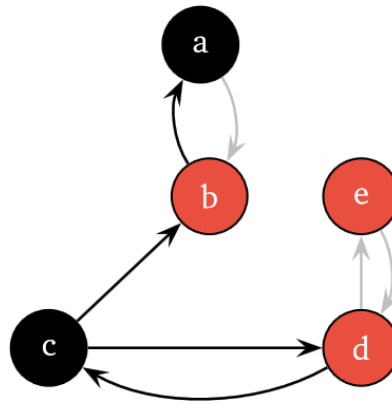


FIG. 1: Example random network with $N = 5$ during update. Nodes (a) and (c) are still viable (black), but the others are not (red). Directed edges (grey/black) encode two types of species dependence: necessity and incompatibility.

the logical query $< p$? to generate elements as 1 with probability $p$ and 0 with probability $1 - p$. For those as 1 we pull another random number from this distribution and make the logical query $< q$?, to tag the link as necessity with probability $q$ and incompatible with probability $1 - q$. This process generates adjacency matrices whose diagonal elements are zero and non-diagonal elements are like random variables distributed as $A_{ij} \sim Bern(p)$ iid. In order to ensure this process was appropriate we created plots of the resulting distribution from these values to graphically compare. This is not included (as it is entirely uninteresting and purely for verification purposes), but showed that within $1-2\%$ the generated matrices elements are distributed as required. Also, naturally matrices were calculated with freshly generated numbers each time to prevent bias.

## IV.   RESULTS AND DISCUSSION

The key aim of this model is to understand how the fraction of species remaining varies with increasing number of initial species $N$ in the ecosystem. As we have mentioned, for a given system size there are two parameters which determine the interdependency of species. We wish to understand how the size of the remaining subset of species varies for the full parameter space in order to determine the behaviour of the model.

Figure 2 shows how the remaining fraction varies as the connectivity changes (via $p$) and link type changes (via $q$), for a selection of system sizes. One can see how more of the parameter space results in extinction as we increase the number of species, and how the transition curve (from all to zero species remaining) becomes sharper at larger $N$. Note that simulations have been carried out well above $N = 25$, but they are less important to include here, due to the vanishing non-trivial region
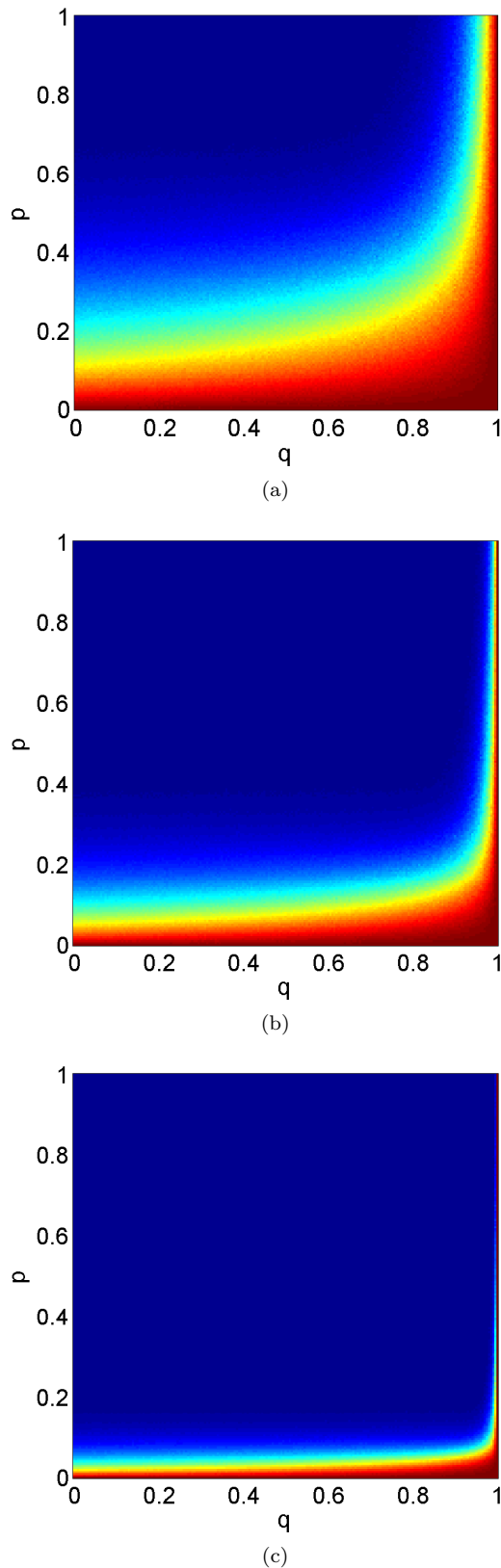
FIG. 2: Full parameter space for various system sizes, $N$: 5 (a), 10 (b), and 25 (c). Red corresponds to all species viable, and the blue to all species non-viable. Resolution ($p$,$q$ step size) 0.005, ensemble size 1000.

with increasing $N$. It is also difficult to get high resolution plots with appreciable ensemble sizes as is computationally intensive as was mentioned in Section II B. Later we see more specific regions of this parameter space probed with much larger system sizes, in order to really separate underlying relations from finite size effects.

This overall view is very good at identifying the general picture of how size distribution varies with ecosystem size, however we are interested in making quantitative statements about the relationships governing this picture. To do this we take a reductionist approach by taking parameters to their extreme cases and intensifying our study of these regions, in the hope that overall behaviour is a combination of some underlying set of relationships.

There are two trivial cases: first $p = 1$ and we are looking at a completely disconnected graph (zero edges) where no species may interact with another (thus link type is irrelevant). This is quite non-physical as we stated in Section I that organisms do not exist in isolation, but from the perspective of this model, we see no interactions amongst species, therefore all species are viable (recalling our initial condition). Note how we seek to query viability of a species within an ecosystem, so we say nothing of the species ability to exist without external factors, merely that its niche within this ecosystem exists. The second trivial case is $q = 1$. If we recall how the parameter $q$ was defined, as the probability of a link being of type necessity, then when $q = 1$ we have all links of type necessity. Once again recalling our initial condition of all species viable, it is understandable that for any number of links placed no species will become non-viable, as each link only imposes a rule that both are viable and there is a complete absence of any rule which can cause that to not be the case.

We are therefore much more interested in the two non-trivial extremes, $q = 0$ and $p = 1$ which, as one can see in Figure 2, have values which appear to scan the complete range for each $N$. First we look at the $q = 0$ case, that is every link that is placed is of type incompatibility. Figure 3 shows how the logarithm of fraction of species remaining varies with $p$ for various values of $N$. We observe a linear fit in these two variables for large values of $N$. In order to characterise this fit we did this for increasing values of $N$, between 5 and 1000, and looked at how the fit changed with system size. This can be seen in Table II. We observe that the gradient of this straight line $m$ tends to $N$ for large $N$, and the intercept $c$ tends to zero. A linear relationship between these two variables therefore implies an equation which describes this below:

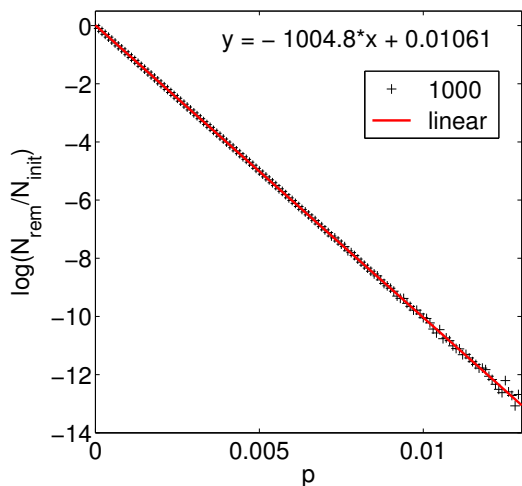$$\frac{N_f}{N} = e^{-Np} \qquad (2)$$

Where $N$ is the initial number of species (nodes), $N_f$ is the number of species remaining, and $p$ is the link probability. Our attention is immediately drawn to the term $Np$ as the exponent: this is defined as the mean connectivity, i.e. how many links each node is expected to
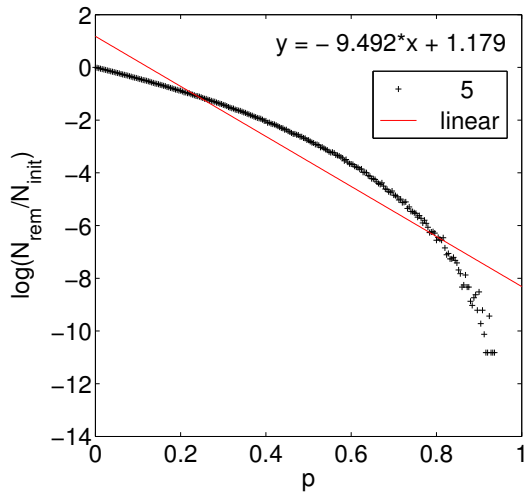
(a)
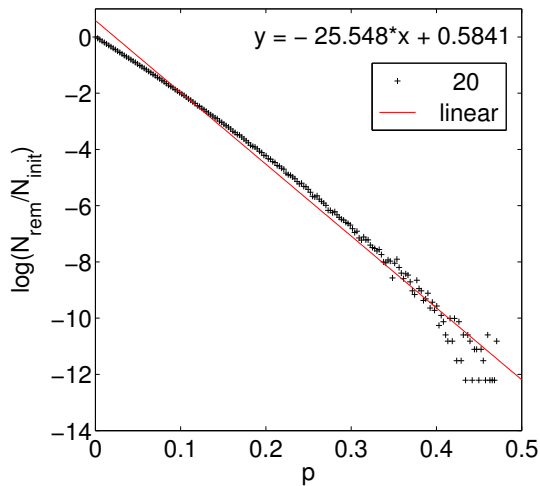


(b)



(c)

FIG. 3: For fixed $q = 0$, logarithm of normalised number of species for varying $p$, and linear fit, for various values of $N = 5$ (a), 20 (b), 1000 (c); Ensemble size 10000.

| N | m | c | p |
|---|---|---|---|
| 5 | $-9.49$ | 1.18 | $0.0 - 0.936$ |
| 10 | $-14.9$ | 0.847 | $0.0 - 0.708$ |
| 20 | $-25.5$ | 0.584 | $0.0 - 0.471$ |
| 40 | $-45.5$ | 0.315 | $0.0 - 0.260$ |
| 80 | $-85.4$ | 0.161 | $0.0 - 0.146$ |
| 160 | $-167.8$ | 0.139 | $0.0 - 0.0834$ |
| 1000 | $-1004.8$ | 0.011 | $0.0 - 0.013$ |

TABLE II: Parameters of linear fit $y = mx + c$ for increasing system size for $q = 0$. We restrict ourselves to values of $p$ which do not result in no species remaining.

have to other nodes. We immediately turn our attention toward the graph structure of our network, Erdős-Rényi, and recall a key result: the existence of a giant component [18, 20, 21].

As links are placed from the successful outcome of an independent Bernoulli trial the vertex degree distribution will be Binomial [19] (counting successes/failures of Bernoulli trials). Taking the limit of system size to infinity, whilst keeping the product $Np$ constant, gives us a Poisson degree distribution. It is then very intuitive to consider a Poisson branching process from an arbitrary node: if there is on average more than one link per node, we obtain a sustained branching process, which translates to a graph component that spans the system (the so-called infinite spanning cluster, or giant component). This was formally investigated by Erdős and Rényi who found that when $Np$ exceeds the threshold value of one, a unique giant component almost surely exists [21]. This transition is related to the continuous phase transition in two-dimensional percolation problems, and can intuitively be understood in that formalism [35, 36].

We observe finite size effects when trying to fit this relationship to the data: for low values of $N$ the fit is poor, but it becomes extremely good for large $N$ (less than 1% deviation). Figure 4 shows this scaling very clearly; it shows the results for multiple system sizes collapsed onto the same domain. One can clearly see the data tend toward a straight line as system size increases and finite size effects become negligible, as the giant component only exists for infinite system size (as discussed above). Therefore we conclude that this relationship is connected to the giant component phase transition in the network structure.

We must now ask: what ecological interpretation does this relationship have? We are observing a sort of dying giant component of species. When there are few links, adding one essentially causes a species to become nonviable. As we add more, there is a greater chance that the new link will remove a species that is already nonviable, so no removal, which explains the exponential decay in remaining species with increasing $p$. After a certain number of links have been placed - related to the giant component phase transition - no other component
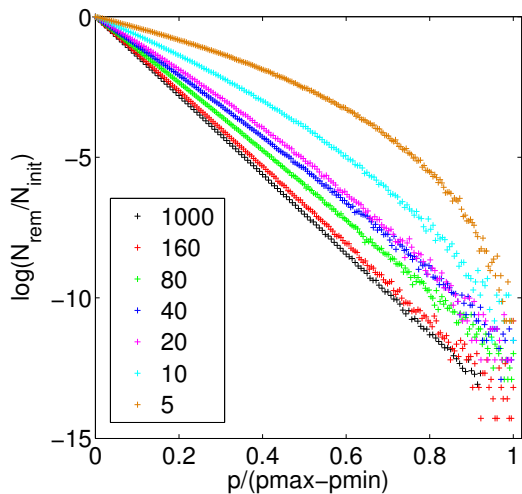
FIG. 4: For fixed q=0. Normalisation allows collapse of data from varying $N$ onto one another. Finite size effects observed as trend becomes linear for large $N$.



FIG. 5: For fixed p=1. Normalisation allows collapse of data from varying $N$ onto one another. Trend can be seen as linear for all $N$.

will contain more than $O(log(N))$ vertices [21], and the majority of species will be non-viable. Therefore ecologically speaking, this scenario requires species to form clusters in order to be viable. In fact, this is what one observes in nature: large ecosystems tend to form so-called communities or an ecological niche which is stable and more robust to global changes (such as extinction of other communities).

We can also see this more clearly by rewriting the relationship as:

$$p = \frac{ln(N)}{N}\left(1 - \frac{ln(N_f)}{ln(N)}\right) \qquad (3)$$

This looks a little more complicated but allows us to recall one of the results of Erdős and Rényi, who note that if $p < (1-\epsilon)\frac{ln(N)}{N}$ then a graph will almost surely contain isolated vertices and be disconnected. If however $p > (1+\epsilon)\frac{ln(N)}{N}$ then it will almost surely be connected, implying a threshold of connectedness for the graph of $\frac{ln(N)}{N}$ [20, 21]. We can identify the term $\frac{ln(N_f)}{ln(N)}$ above as $\epsilon$, and as both $N$ and $N_f$ are positive constants, and greater than zero (as we restricted ourselves to values of $p$ with non-zero species remaining; see Table II), then this term must be positive. This means we are in the region of disconnectedness when there are non-zero species remaining.

The second non-trivial extreme case is $p = 1$. This corresponds to the complete graph: all nodes are connected to all other nodes (self-excluding), and we vary link type via the $q$ parameter. We perform the same procedure as for the $q = 0$ case: obtaining linear fits for the variables $1 - q$ and the logarithm of remaining fraction. The results are outlined in Table III, and the normalised data
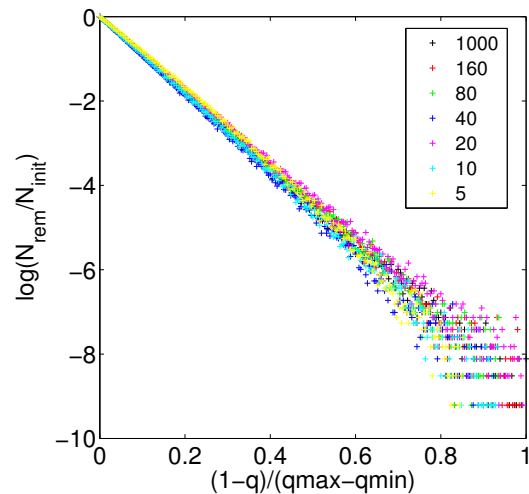
is presented in Figure 5. In this case we do not observe finite size effects, instead obtaining a good linear fit for a wide range of system sizes. Once again the intercept $c$ tends to zero, but this time very quickly, without much early deviation. The gradient appears to scale as $N^2$. Actually we believe it to scale as $N(N-1)$ (obviously not distinguishable, especially for large $N$), which lends itself to a very simple and intuitive explanation of the relationship shown below:

$$\frac{N_f}{N} = e^{-N(N-1)(1-q)} \qquad (4)$$

If we draw parallels between the exponent terms in each of these relationships, it makes sense to query the threshold $N(N-1)(1-q) = 1$. The term $1-q$ is the probability a link is of type incompatible and the factor $N(N-1)$ is the number of edges in the graph. The product of these is therefore the expected number of links of type incompatibility in the graph, as mentioned in Section II B. The threshold underlying this is when there is just one link of type incompatibility in the graph; when there is less than one, some species can remain, but greater than one and all species become non-viable. This is understandable as $p = 1$ corresponds to the complete graph, and starting with all necessity links and adding just one incompatibility type link causes a sort of propagation of non-viability throughout the system, until all species are non-viable; since when one is removed all others depend on that, hence get removed, until all become non-viable. This relationship therefore describes the chance of destruction of the complete graph by a destabilising organism (such as a keystone species [37]). This is a direct consequence of the strength of the link types and is also why this is not a system-size-based effect. With this, all extreme cases have been classified

and relationships between relevant variables found. We can use these to describe the behaviour of non-extreme cases, where a combination of these effects will come into play.

| N | m | c | q |
|---|---|---|---|
| 5 | $-24.6$ | 0.295 | $0.659 - 1.0$ |
| 10 | $-95.5$ | 0.100 | $0.907 - 1.0$ |
| 20 | $-382.0$ | $-0.006$ | $0.978 - 1.0$ |
| 40 | $-1568.6$ | 0.003 | $0.995 - 1.0$ |
| 80 | $-6497.2$ | 0.056 | $0.9987 - 1.0$ |
| 160 | $-25860$ | 0.037 | $0.9997 - 1.0$ |
| 1000 | $-994019$ | $-0.017$ | $0.999992 - 1.0$ |

TABLE III: Parameters of linear fit $y = mx + c$ for increasing system size for $p = 1$. We restrict ourselves to values of $q$ which do not result in zero species remaining.

## V.  CONCLUSIONS & FUTURE WORK

We have presented a model which represents an ecosystem as a random network, with species as nodes, and the dependencies between species encoded in the links. Species are denoted as viable or not via a Boolean state variable which is synchronously updated given the rule-set defined by the network edges. We used this model to investigate the size variation of ecosystems based on different levels and types of connectivity. We found that increasing system size causes a great deal of the parameter space to result in non-viability. By looking at extremal cross-sections of this space we determined relationships describing how the subset of viable species varies with increased system size. Most prominently we found that in this model is it necessary for a large ecosystem to stay disconnected in order to allow for any species to remain. This is due to the strength of the link types, deterministically removing incompatible species, so highly con-

nected systems can be completely non-viable due to this strength.

A natural extension to this study would be to look at altering the rule-set. As previously mentioned this can be done in a number of frameworks, such as SAT or multiplex networks. One interesting way to do this however would be to, instead of accumulating a large number of rules, introduce a tunability parameter to the existing model. For example, a link could be of type incompatibility, but it could have a strength parameter which allows it to limit other species viability in a probabilistic way, rather than deterministic. Often in nature we have a great deal of variation in the level of interaction between species, so this could capture some of that behaviour. Indeed it is often suggested that a great deal of weak connectance is the key to stability [15, 38, 39], and this could incorporate that without making the model limited by its specificity.

It would also be useful to further investigate the ecological interpretation of this model to see in what other ways it might match up to nature, and indeed to similar systems in other fields. We have also not specifically characterised the curve in parameter space (Figure 2), and it appears non-trivial to do so. It would be good to investigate this further to determine whether an analytic expression might be found, and if it is indeed the combination of the described relationships as it appears to be. Finally it would be interesting to look at the evolution of local graph properties during the dynamics of determining the viable subset, as we have obtained a good global view. For example, to look at the formation of clusters and their size distributions.

## VI.  ACKNOWLEDGEMENTS

[1] S. Levin, S. Carpenter, H. Godfray, A. Kinzig, M. Loreau, J. Losos, B. Walker, and D. Wilcove, *The Princeton Guide to Ecology* (Princeton University Press, 2009).

[2] E. Schulze, E. Beck, and K. Müller-Hohenstein, *Plant Ecology* (Springer-Verlag, 2005).

[3] M. Molles, *Ecology: Concepts and Applications* (WCB/McGraw-Hill, 1999).

[4] T. Smith and R. Smith, *Elements of Ecology* (Pearson Benjamin Cummings, 2012).

[5] S. Pimm, *The Balance of Nature?: Ecological Issues in the Conservation of Species and Communities* (University of Chicago Press, 1991).

[6] W. Z. Lidicker, BioScience **29**, 475 (1979).

[7] D. H. Boucher, S. James, and K. H. Keeler, Annual Review of Ecology and Systematics , 315 (1982).

[8] M. Begon, C. Townsend, and J. Harper, *Ecology: From Individuals to Ecosystems* (Wiley, 2009).

[9] A. J. Lotka, Williams and Wilkins Co (1925).

[10] V. Volterra, Nature **118**, 558 (1926).

[11] R. Arditi and L. R. Ginzburg, Journal of Theoretical Biology **139**, 311 (1989).

[12] V. Grimm and S. F. Railsback, *Individual-based Modeling and Ecology* (Princeton University Press, 2013).

[13] D. DeAngelis and L. Gross, *Individual-based Models and Approaches in Ecology: Populations, Communities, and Ecosystems* (Chapman & Hall, 1992).

[14] V. Grimm, U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S. K. Heinz, G. Huse, *et al.*, Ecological Modelling **198**, 115 (2006).

[15] G. A. Polis and D. R. Strong, American Naturalist , 813 (1996).

[16] P. Bak, C. Tang, and K. Wiesenfeld, Physical Review A **38**, 364 (1988).

[17] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, Science **298**, 824 (2002).

[18] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Physics Reports **424**, 175 (2006).

[19] M. E. Newman, S. H. Strogatz, and D. J. Watts, Physical Review E **64**, 026118 (2001).

[20] P. Erdős and A. Rényi, Publ. Math. Debrecen **6**, 290 (1959).

[21] P. Erdős and A. Rényi, Publ. Math. Inst. Hungar. Acad. Sci **5**, 17 (1960).

[22] B. Bollobás, *Random Graphs*, Cambridge Studies in Advanced Mathematics (Cambridge University Press, 2001).

[23] S. A. Cook, in *Proceedings of the third annual ACM symposium on theory of computing* (ACM, 1971) pp. 151–158.

[24] L. A. Levin, Problemy Peredachi Informatsii **9**, 115 (1973).

[25] K. Bassler, "Symmetry in critical random boolean networks dynamics," (May 2014), talk at Centre for Complexity Science, University of Warwick.

[26] S. A. Kauffman, Journal of Theoretical Biology **22**, 437 (1969).

[27] S. A. Kauffman, *The Origins of Order: Self-organization and Selection in Evolution* (Oxford University Press, 1993).

[28] C. Gershenson, Workshop and Tutorial, Proc. of the 9th Int. Conf. on Simulation and Synthesis of Living Systems (ALife IX) , 160 (2004), nlin/0408006.

[29] C. Gershenson, in *Artificial Life 8* (MIT Press, 2003) pp. 1–8.

[30] P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, and J.-P. Onnela, Science **328**, 876 (2010).

[31] S. Gomez, A. Diaz-Guilera, J. Gomez-Gardeñes, C. J. Perez-Vicente, Y. Moreno, and A. Arenas, Physical Review Letters **110**, 028701 (2013).

[32] J. Gómez-Gardeñes, I. Reinares, A. Arenas, and L. M. Floría, Scientific Reports **2** (2012).

[33] A. Arenas, "Multiplex networks, structure and dynamics," (February 2014), talk at Centre for Complexity Science, University of Warwick.

[34] M. Matsumoto and T. Nishimura, ACM Transactions on Modeling and Computer Simulation (TOMACS) **8**, 3 (1998).

[35] A. Aharony and D. Stauffer, *Introduction To Percolation Theory* (Taylor & Francis, 2003).

[36] G. Grimmett, *Percolation*, Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen (Springer, 1999).

[37] R. Paine, Conservation Biology **9**, 962 (1995).

[38] T. Gross, L. Rudolf, S. A. Levin, and U. Dieckmann, Science **325**, 747 (2009).

[39] N. M. Haddad, G. M. Crutsinger, K. Gross, J. Haarstad, and D. Tilman, Ecology Letters **14**, 42 (2011).