How Tangled is Nature? A Model of Evolutionary Ecology

Paul Anderson and Henrik Jeldtoft Jensen

Department of Mathematics, Imperial College London, South Kensington campus, London SW7 2AZ, United Kingdom

Introduction

An important characteristic of an ecosystem is the set of all interactions between the various individuals. Organisms may influence each other in many ways and it is difficult to monitor and quantify most relationships except for the most obvious. Here, we look at the effect of different levels of connectivity between species within the framework of a simple model of ecosystem assembly and evolution: the *Tangled Nature model* [1, 2, 3]. All work presented in this poster appears in [4]. We compare the early and late time connectivity and cluster properties of ecosystems evolving in two differently connected spaces: genotypes influence either a small or a large number of other genotypes.
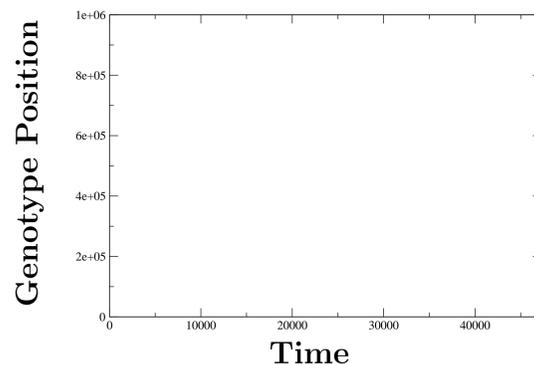
The Model and Methods

↪ An individual is represented by a vector $\mathbf{S}^\alpha = (S_1^\alpha, S_2^\alpha, ..., S_L^\alpha)$ in the genotype space $\mathcal{S}$, where the "genes" S_i^α may take the values ± 1, i.e. $\mathbf{S}^\alpha$ denotes a corner of the L-dimensional hypercube. We take $L = 20$. The evolutionary dynamics determines whether a genotype is occupied or not. The total number of occupied sites is called the *diversity*.

↪ For simplicity, an individual is removed from the system with a constant probability p_{kill} per time step.

↪ The probability that an individual reproduces, p_{off}, is controlled by a weight function $H(\mathbf{S}^\alpha, t)$ related to its interactions with other sites. Reproduction is *asexual* and mimics fission: two individuals are produced with the parent being killed.

↪ Each gene of the offspring has a fixed probability of mutating per time step, p_{mut}. 500 individuals are placed randomly on the network to start the simulation.

↪ A time step consists of *one* annihilation attempt followed by *one* reproduction attempt. One generation consists of $N(t)/p_{kill}$ time steps, which is the average time taken to kill all currently living individuals. Generation time is used throughout.

We are interested in the effect of changing the background connectivity, θ. This determines the probability that any two sites are interacting. If they are, then the strength of the interaction is given by $J^{ab} = J(\mathbf{S}^a, \mathbf{S}^b)$, a number between -1 and $+1$. All connections are calculated at $t = 0$. Thus the network properties at any given time depend on which sites are occupied. Interactions between other genotypes can be explored by mutations away from the current site.
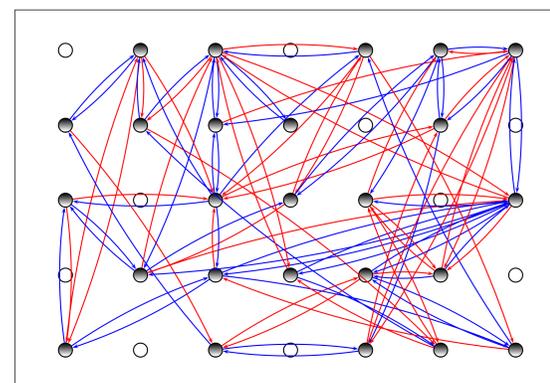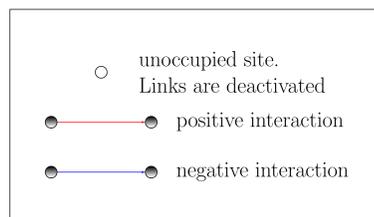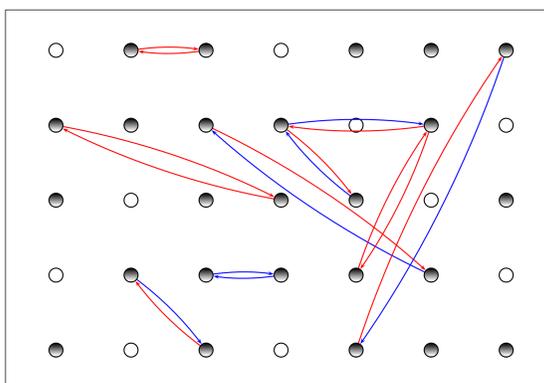
Our main results are explained in the figures. We consider two values for θ: $\frac{1}{200}$ (*low θ*) and $\frac{1}{4}$ (*high θ*), and three time values: $t = 500$ (*primal time*), $t = 5000$ (*early time*) and $t = 500000$ (*late time*). An ensemble of 500 runs for low and high θ were run on a cluster of undergraduate machines left running overnight and at weekends.

The degree and strength distribution plots below show results from the simulation and the *null hypothesis*. For this, the number of individuals at a given time was read in from the simulation and these were then thrown down at random on to the network with the constraint that the diversity was the same. This provides a check on whether any trends are real or just illusions created by an expanding diversity.

Other features of the Tangled Nature model include a punctuated dynamic as shown below — where the network spends long periods in a so-called quasi-Evolutionary Stable Strategy (q-ESS) terminated by hectic rearrangements of genotype space until a new q-ESS is found — and the appearance of quasi-species [5].

Conclusion

Our most important results are that non-trivial temporal evolution of the network properties of an ecosystem and a realistic form for the species abundance are only seen if the genotype space is well connected. This is interpreted here as meaning that an occupied genotype is likely to interact with many other (potentially occupied) genotypes. No evolution at the level of ecosystems can occur in a world where most genotypes have very little influence on other organisms. It is easy to overlook the importance of the entire network of interactions when dealing with small communities of organisms on a macroscopic scale, but easier to visualise with colonies of billions of bacteria.

From our results, it is tempting to speculate that the observed degree of diversity, complexity and adaptation of living matter may be directly related to a high level of interdependence between organisms. Hence, Darwin's entangled bank may be a particularly useful image to keep in mind when studying the evolution of large collections of individuals.

References

[1] Kim Christensen, Simone A. di Collobiano, Matt Hall, and Henrik J. Jensen. Tangled Nature: A Model of Evolutionary Ecology. *J. theor. Biol.*, 216:73–84, 2002.

[2] Matt Hall, Kim Christensen, Simone A. di Collobiano, and Henrik J. Jensen. Time-dependent extinction rate and species abundance in a tangled-nature model of biological evolution. *Phys. Rev. E*, 66(011904):1–10, 2002.

[3] Simone A. di Collobiano, Kim Christensen, and Henrik J. Jensen. The tangled nature model as an evolving quasi-species model. *J. Phys. A: Math. Gen.*, 36:883–891, 2003.

[4] Paul Anderson and Henrik Jensen. Network properties, species abundance and evolution in a model of evolutionary ecology. submitted to Proceedings of the Royal Society B. Pre-print at cond-mat/0307114.

[5] Manfred Eigen, John McCaskill, and Peter Schuster. Molecular Quasi-Species. *J. Phys. Chem.*, 92:6881–6891, 1988.

Acknowledgements

Further information

More information can be found in the papers cited above and at H. J. Jensen's webpage: *http://www.ma.imperial.ac.uk/~hjjens*. Please send e-mail to *h.jensen@imperial.ac.uk*.

Results

Not all sites are occupied. There are several isolated species, in the sense that they are not interacting with anyone. Most sites are in two-clusters. These act as building blocks for larger groups. They are usually plugged together by mutants. Large clusters do not persist and the mutually positive two-clusters are the only long-living structures. There is no tendency to form larger clusters at later times.

Not all sites are occupied. Notice how all nodes are connected in one giant cluster and there are no isolated species. With such a high background connectivity, all occupied sites belong to one cluster at all time steps, although an individual species may only be interacting with a few other genotypes. In the simulation, the nodes sit on the corners of a 2^{20} dimensional hypercube.
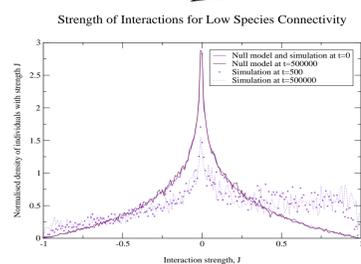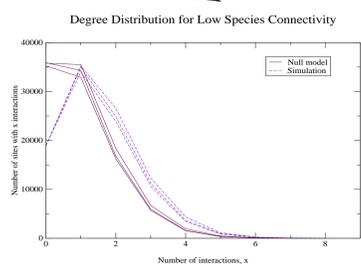

The degree distribution shifts out at later times due to an increased diversity but does *not* evolve away from the null model since isolated sites are over represented in the null case.

A change from the null model is seen, but this is not due to any fundamental change in cluster structure but rather the eventual dominance of mutually positive two-clusters.
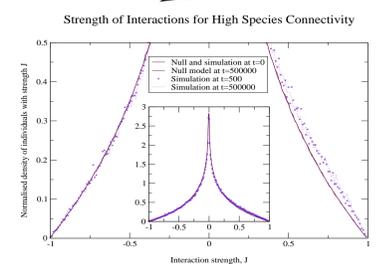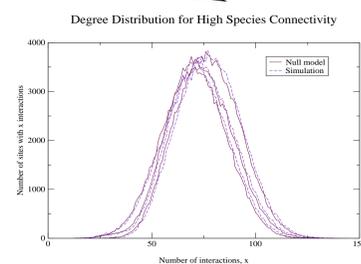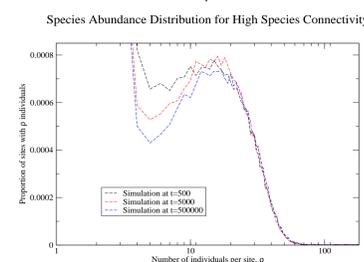
The degree distribution shifts out at later times due to an increased diversity but, as for the low connectivity case, does not evolve away from the null model.

A definite shift towards more positive interactions occurs. *This* is what drives the increasing diversity and is non-trivial since all sites are tangled together in one giant cluster.
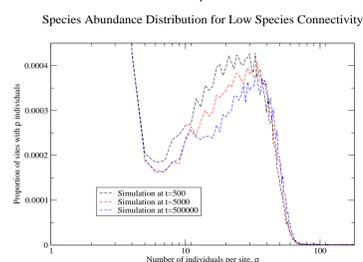
The species abundance distribution (SAD) is skewed to the right at later times as the heavily populated two-clusters flourish. Thus patches of clusters do not produce the log-normal form expected from field studies.

The species abundance distribution (SAD) evolves and becomes a closer fit to a log-normal at later times. Thus the single cluster of highly interdependent genotypes produces a similar SAD to those observed by ecologists.

Low connectivity → unrealistic SAD

High connectivity → realistic SAD