

# The Tangled Nature model of Evolutionary Ecology: An overview

Simon Laird, Daniel Lawson, and Henrik Jeldtoft Jensen

Department of Mathematics, Imperial College London, South Kensington campus, SW7 2AZ, U.K.

**Summary.** We present a review of the Tangled Nature model. The model is developed to focus on the effect of evolution and multiple interactions on ecological and evolutionary observables. The model is individual based and ecological structures, such as species, are emergent quantities. The dynamics consists of a simplistic mutation prone multiplication in which the probability of producing an offspring is determined by the occupancy in genotype space. The macroscopic long time dynamics is intermittent and exhibit a slow decrease in the macroscopic extinction rate. Ecological quantities such as the Species Abundance Distribution and the Species Area Relation compare qualitatively well with observations, as does the relation between interaction and diversity. The effect of correlations between parents and mutants has been studied as has the effect of a conserved resource.

**Key words:** evolution, ecology, network of interactions, Species-Area Relations, Species-Abundance Distribution, SAR, SAD

## 1.1 Introduction

Can we establish a general framework for the description of ecosystems consisting of many interacting and evolving organisms? We seek a description which will enable us to bridge the span from microscopic to macroscopic time. The model is to be defined at the level of individuals so that ecological and taxonomic structures emerge as a result of the dynamics. The description should be logically simple and cover broad classes of observed facts.

Much work is done on the regime of strongly separated timescales, using e.g. ‘adaptive dynamics’ [1] or game theory [2, 3]. One can incorporate evolution by considering that evolutionary dynamics occurs much more slowly than ecological dynamics, so that at all times the system is in an equilibria (e.g. the Webworld model, [4][5]). However, the timescales may overlap when long-lived species interact with short-lived, and additionally there is evidence [6] that the evolutionary dynamics can affect ecological stability, and vice versa. These are not insurmountable problems, e.g. the ‘Streetcar theory of evolution’ [7]

addresses this problem, but assumes that the system equilibrates between mutations. Such models are essential to understand individual cases but we wish to also address the broad picture. Individual-based models [8] are an appropriate tool to combine fast and *usually* predictable ecological processes with the highly stochastic evolutionary behaviour in a controlled and realistic manner. Our model falls into the category (e) in DeAngelis and Mooij’s classification system [9]: “Genetic Variability and evolution”, although arguably the most important feature gained from an individual based approach at such a simple level is ‘demographic stochasticity’, that is, representing the inherent noise in the population in an accurate way. Our approach is inspired by statistical mechanics and complex systems theory, looking for overall general structures instead of specific details of the organisms comprising the system.

Our model consists of asexually reproducing individuals all subject to the same killing probability (per time unit). The multitude of interactions between co-existing organisms is included through a weight function which determines the reproduction rate of individuals. Our philosophy is that a simple statistical model may help to identify the important mechanisms behind macroscopic observed ecological measures.

## 1.2 A simple introductory model

The Tangled Nature model is an individual based model of evolutionary ecology. We give a brief outline of the model here, with more details available in [10, 11, 12, 13]. This version of the model is an attempt to address systems with many interactions between species in the simplest possible way, with detail and realism added in stages. We start with the bare model, to which we will add spatial effects and consider more realistic forms of the fitness.

### 1.2.1 Uncorrelated non-spatial model

An individual is represented by a vector  $\mathbf{S}^\alpha = (S_1^\alpha, S_2^\alpha, \dots, S_L^\alpha)$  in the genotype space  $\mathcal{S}$ , where the ‘genes’  $S_i^\alpha$  may take the values  $\pm 1$ , i.e.  $\mathbf{S}^\alpha$  denotes a corner of the  $L$ -dimensional hypercube. In the present paper we take  $L = 20$ , giving a reasonable sized space (over a million genotypes) whilst not being computationally prohibitive. We think of the genotype space  $\mathcal{S}$  as containing all possible ways of combining the genes into genome sequences. Many sequences may not correspond to viable organisms. The viability of a genotype is determined by the evolutionary dynamics. All possible sequences are available for evolution to select from. We will see that a natural species concept arises from the dynamics, in which each species is separated in genotype space.

The system consists of  $N(t)$  individuals, and a time step consists of *one* annihilation attempt followed by *one* reproduction attempt. A reproduction

event is successful with varying probability  $p_{off}$ , defined below, and an annihilation attempt is successful with constant probability  $p_{kill}$ <sup>1</sup>. One generation consists of  $N(t)/p_{kill}$  time steps, which is the time taken (on average) to kill all currently living individuals. The dynamics lead to an (approximately) constant population size, on short timescales.

The ability of an individual to reproduce,  $p_{off}$ , is ultimately controlled by a weight function  $H(\mathbf{S}^\alpha, t)$ :

$$H(\mathbf{S}^\alpha, t) = \frac{c}{N(t)} \left( \sum_{\mathbf{S} \in \mathcal{S}} \mathbf{J}(\mathbf{S}^\alpha, \mathbf{S}) n(\mathbf{S}, t) \right) - \mu N(t), \quad (1.1)$$

where  $c$  controls the strength of interaction (large  $c$  means large interaction),  $N(t)$  is the total number of individuals at time  $t$ , the sum is over the  $2^L$  locations in  $\mathcal{S}$  and  $n(\mathbf{S}, t)$  is the number of individuals (or occupancy) at position  $\mathbf{S}$ . Two positions  $\mathbf{S}^a$  and  $\mathbf{S}^b$  in genome space are coupled with fixed but random strength  $\mathbf{J}(\mathbf{S}^a, \mathbf{S}^b)$  ( $= J^{ab}$  in matrix notation) which can be either positive, negative or zero. This link exists (in both directions) with probability  $\theta$  ( $= 0.2$  in Sect. 1.2.1), i.e.  $\theta$  is simply the probability that any two sites are interacting. If the link exists, then  $\mathbf{J}(\mathbf{S}^a, \mathbf{S}^b)$  and  $\mathbf{J}(\mathbf{S}^b, \mathbf{S}^a)$  are both generated random and independent  $\in (-1, 1)$ . To study the effects of interactions *between* species, we exclude self-interaction so that  $\mathbf{J}(\mathbf{S}^a, \mathbf{S}^a) = 0$ .

The conditions of the physical environment are simplistically described by the term  $\mu N(t)$  in Eq. (1.1), where  $\mu$  determines the average sustainable total population size, i.e. the carrying capacity of the environment. An increase in  $\mu$  corresponds to harsher physical conditions. We use asexual reproduction consisting of one individual being replaced by two copies mimicking the process of binary fission seen in bacteria. Successful reproduction occurs with a probability per unit time given by

$$p_{off}(\mathbf{S}^\alpha, t) = \frac{\exp[H(\mathbf{S}^\alpha, t)]}{1 + \exp[H(\mathbf{S}^\alpha, t)]} \in [0, 1]. \quad (1.2)$$

This function is chosen for convenience, the specific functional form having no effect on the dynamics of the model - any smoothly varying function that maps  $H(\mathbf{S}^\alpha, t)$  to the interval  $[0, 1]$  will do. We allow for mutations in the following way: with probability  $p_{mut}$  per gene we perform a change of sign  $S_i^\alpha \rightarrow -S_i^\alpha$  during reproduction.

Eq. (1.1) can be understood as the average interaction for an individual with all others, with a normalisation condition given by the  $\mu N$  term, which determines the total population and controls fluctuations. The interaction strength  $c$  gives the magnitude of the total interactions, regardless of density; i.e. we choose to represent systems in which the population size has saturated<sup>2</sup>.

<sup>1</sup>The restriction of constant killing probability can be shown to be qualitatively irrelevant in this model, as we do not include individual aging.

<sup>2</sup>One could easily consider the case of density dependent interactions by allowing  $c = c(N)$ .

We can tune the effective ‘resource’ density (and hence population density) with the parameters  $c$  and  $\mu$ . The total population remains approximately constant over ecological timescales (and actually increases over evolutionary timescales). Setting self-interaction to zero is equivalent to considering that all types interact equally with their own species (one can rescale  $p_{kill}$  and  $\mu$  to accommodate this); we relax this constraint in Sect. 1.2.3.

Initially, we place  $N(0) = 500$  individuals with randomly chosen genotype. Their initial location in genotype space does not affect the nature of the dynamics. A two-phase switching dynamic is seen, consisting of long periods of relatively stable configurations (quasi-Evolutionary Stable Strategies or q-ESSs) (Fig. 1.1) interrupted by brief spells of reorganisation of occupancy called transitions. Transition periods are terminated when a new q-ESS is found, as discussed in [10]. A ‘species’ can be well defined as the highly occupied genotype points called ‘wildtypes’, which are separated in genotype space. Each wildtype is surrounded by a ‘cloud’ of mutant genotypes with low occupancy. Thus we can take a natural definition of diversity: the number of wildtypes in the system.

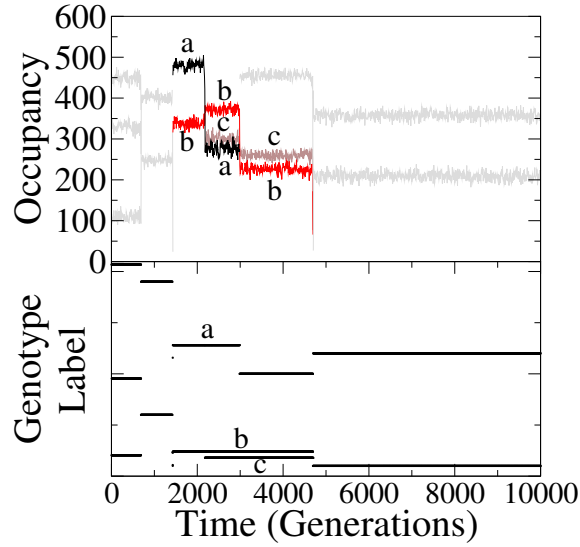
The q-ESS themselves consist of a number of wildtype species, which are ecologically stable and stable to mutations from the neighbouring genotype space. The absolute stability of a q-ESS depends on both the stability against invasion by mutants, and against ecological ‘accidents’ such as the extinction of a keystone species. The two cannot be separated in our model, as, for example, the probability of an accident eliminating a wildtype will depend on the interaction properties of all individuals in the system.

The observed Species Abundance Distribution is log-normal like [10] with a log-series like tail, consistent with many observations and similar to that found in neutral theory [14]. The log-normal portion is made up of wildtype species only. These have evolved so that the number of births in a given species exactly cancels the number of deaths (and mutations). The log-series section is made up of ‘mutant’ sub-species, that is, species who have experienced deleterious mutations from a wildtype. They are less successful and short lived, as their population is only supported by a constant influx of mutants from the neighbouring wildtype.

The long term dynamics of this model are essentially the same as the extension in Sect. 1.2.3. The stability of the q-ESS found increases slowly with time, as does the mean total population  $\langle N(t) \rangle$ . The waiting time for extinction events (occurring at transitions) decreases with time (close to  $1/t$ ), and the species lifetime distribution compares well with much fossil data [10]. Results which support these conclusions are presented in [15, 16] for a very similar model but with simplified dynamics.

### 1.2.2 Uncorrelated spatial model

To add spatial extension to the Tangled Nature model, we consider a number of sites on a spatial lattice, with a number of individuals in each. We now

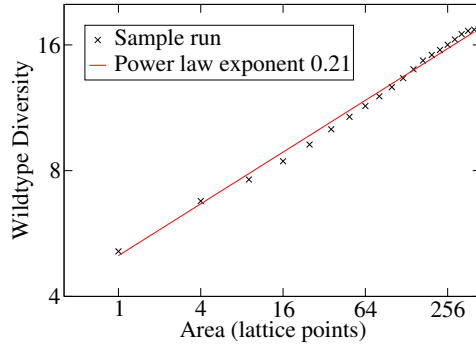


**Fig. 1.1.** q-ESS in a single run, both graphs have the same time axis. Top: Occupancy of wildtype species vs. time. Horizontal lines constitute q-ESS and the changes between q-ESS are transitions, too fast to see on this timescale. Labels are placed only on species that survive q-ESS transitions. Bottom: A sample plot of species existence vs time. The Genotype Label axis is not meaningfully ordered, and is included for comparison to the correlated phenotype space model in Sect. 1.3 (Fig. 1.4).

consider that individuals interact only with other individuals at the same site, and hence compete only locally for space via the  $\mu$  term. Migration moving an individual to a randomly chosen neighbouring lattice point and occurs with constant probability  $p_{move}$ . Spatial boundaries are periodic to prevent the complication of edge effects.

Each site on spatial lattice behaves similarly to the non-spatial model, spending most time in a q-ESS. However, in the spatial case a small number of migrants are present in each site. This acts as an additional pressure to enter a transition phase, **unless** the site is in the same type of q-ESS as its neighbours. This gives rise to spatial patches of q-ESS types. These patches grow and shrink on a very slow timescale, except when the migration probability  $p_{move}$  is high.

Patches form in sizes that appear to be power-law distributed, and close to power-law Species-Area Curves are observed (Fig. 1.2). The species area curve has a slight s-shape produced by the periodic boundary. This same shape was found in a ‘neutral’ (non-interacting) model[17] of Durrett and Levin. In their neutral model the z-value (the slope in the log-log plot of diversity v.s. area) decreases with decreasing speciation rate. However, their speciation rate can be thought of as a migration rate from an external pool as there is no genotype



**Fig. 1.2.** Sample Species Area Relation (from a single run), which is very close to a power law, with a slight S shape due to the periodic boundaries.

space. In our model with interactions and explicit genotype space, we find that  $z$ -value decreases with *increasing* migration rate inside the system. This is because mutation occurs at constant speed, therefore increasing migration rate increases the competition faced by a new species.

In general it is logical that *immigration* (from outside the system) and *migration* (within the system) are correlated. However, increased internal migration rate reduces the chances of mutations surviving, and so produces the inverse effect of the immigration rate (i.e. mutation rate in models without genotype space) of new species from outside the system. High mobility (i.e. migration and immigration rates) for a family of species may mean better mixing and so less chance for spatial segregation of species within a single family - the standard explanation for why birds families generally have lower  $z$ -values than land-dwelling families. Conversely, e.g. on islands, it allows species from elsewhere to arrive, so possibly increasing diversity (as argued in [17]). Which effect dominates will depend on the geography in question - i.e. the size of the local groups of individuals, and the separation between them.

Magurran and Henderson [18], noted that permanent fish species have log-normal Species Abundance Distribution whilst transient species have a log-series distribution. Our local  $q$ -ESS has the same distribution, with a log-normal like distribution for the wildtypes and a log-series like for mutants and migrants. For low mutation rates and high migration rates, clearly migrants will outnumber local mutants and we will observe the exact same distribution near the  $q$ -ESS patch borders. In our spatial model, the distinction between the two types is of fitness - the wildtypes with a log-normal like Species Abundance Distribution are all equally fit in that they have a reproduction rate exactly balancing the death rate; the migrants with a log-series like Species Abundance Distribution are all less fit and rely on repopulation from an external pool (and, hence, are transient).

Full details for the spatial model are available in [19].

### 1.2.3 Diversity and Fitness in the non-spatial model

The restriction that all species are uniform with respect to their own members is approximately valid for many circumstances but is in general unrealistic. We therefore introduce different values of ‘intrinsic fitness’ to each genotype, which can be formed in many ways. The cases are considered separately for clarity, but are easily combined.

A fitness landscape can be uncorrelated or correlated - the correlation we choose is a type of *Fujiyama* landscape (because it has a single, large peak) [2] defined as follows. One type  $\alpha$  has a fitness of 1, and with each mutational step away from this type we subtract  $\Delta$  ( $= 0.1$  in simulations), down to a minimum of 0. An uncorrelated landscape is generated with each type having a fitness drawn uniformly from  $(0, 1)$ . This results in the modified weight functions:

- 1: Density Dependent fitness, which is the varying interaction of an individual with its own type, defined by:

$$H_d(\mathbf{S}^\alpha, t) = H(\mathbf{S}^\alpha, t) + \frac{\epsilon}{N(t)} n(\mathbf{S}^\alpha, t) E(\mathbf{S}^\alpha) \quad (1.3)$$

Here,  $\epsilon$  is the magnitude of the density dependent part of the ‘intrinsic fitness strength’ and  $\epsilon E(\mathbf{S}^\alpha)$  is the intrinsic fitness of individual  $\alpha$ .  $E(\mathbf{S}^\alpha)$  is determined according to the case studied. 1(a): Uncorrelated, density dependent intrinsic fitness landscape, 1(b): Correlated, density dependent intrinsic fitness using a ‘Fujiyama’ landscape.

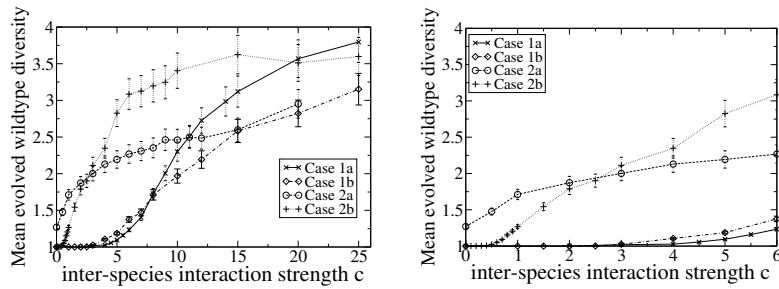
- 2: Density Independent fitness, defined by:

$$H_i(\mathbf{S}^\alpha, t) = H(\mathbf{S}^\alpha, t) + \epsilon_I E(\mathbf{S}^\alpha) \quad (1.4)$$

Here,  $\epsilon_I$  is the magnitude of the density independent part of the intrinsic fitness strength.  $E_i$  is again determined by the case studied. 2(a): Uncorrelated, density dependent intrinsic fitness landscape, 2(b): Correlated, density dependent intrinsic fitness using a ‘Fujiyama’ landscape.

Fig. 1.3 shows the behaviour of the wildtype diversity. Cases 1(a), 1(b) and 2(b) are qualitatively the same, with a rise in density above some characteristic interaction strength  $c$ . The density independent case produces diverse states at much smaller values of  $c$  because only the *difference* in fitness needs to be overcome; in the density dependent case, it is the absolute fitness that provides a barrier to diversification. In the uncorrelated density independent case, there are many species with (approximately) the same high fitness within a couple of mutation steps. This means that, for low  $c$ , neutral drift can occur between the numerous fit types, and for larger  $c$  interaction is the dominant form of selection; in each case, multiple species can be supported (although not in q-ESS for the case of drift).

Interestingly, case 1(b) allows the fittest species to be replaced at lower interaction than for which a diverse state can be supported. However, case 2(b) allows diverse states at lower interaction than for which replacement of



**Fig. 1.3.** left:  $c$  dependence of the average wildtype diversity of an evolved system, taken for 40000 – 50000 generations and 500 runs per data point for the separate cases of density dependent (case 1) and density independent (case 2) fitness, both on a rugged random fitness landscape (cases (a)) and a Fujiyama landscape (cases (b)). Right: A closer look at the low- $c$  region, showing the qualitative similarity in  $D$  vs  $c$  for all cases except case 2(a).

a fittest type is possible. These claims can all be shown with a simple mean field approximation of our model; see [20] for details.

### 1.3 Correlated non-spatial model

In the original model, mutated offspring acquired interaction properties that were uncorrelated with those of the parent. This was unrepresentative of a real biological system where correlations are vital to the evolutionary process. Efforts have since been made to rectify the issue [21] but the limited size and hypercubic geometry of the genotype space have proven to be very restrictive. The problems were finally resolved by disregarding the hypercubic geometry entirely. Correlations were then successfully incorporated so that mutated offspring had interactions comparable with those of the parent [22]. To achieve this goal we use a phenotype description of  $L = 16$  traits,  $\mathbf{T}^a = (T_1^a, T_2^a, \dots, T_L^a)$ , with each trait taking a value from the periodically bounded range,  $[0, 99999]$ . A proportion,  $\theta$ , of the entries of the greatly enlarged interaction matrix,  $\mathbf{J}(\mathbf{T}^a, \mathbf{T}^b)$  are assigned normally distributed values that are locally correlated within the  $\mathbf{J}$ -matrix. All other values of the remaining proportion of the matrix,  $1 - \theta$ , are assigned zeroes which are treated as a lack of interaction between the two relevant phenotypes. As well as being locally correlated in value, the non-zero entries, as a set, are also distributed with a local correlation. By this we mean the  $\mathbf{J}$ -matrix exhibits a clustering of non-zero entries so correlated phenotypes will interact with similar sets of other phenotypes. The result is that given a mutation of one trait value we have an exponential decay in the correlation between parent and offspring interaction set values that is dependent upon the distance mutated in the trait value,  $\Delta(\mathbf{T}^\alpha, \mathbf{T}^\beta)$ ,

$$\mathbf{c}(\mathbf{J}(\mathbf{T}^\alpha, \mathbf{T}^\gamma), \mathbf{J}(\mathbf{T}^\beta, \mathbf{T}^\gamma)) = \exp[-\Delta(\mathbf{T}^\alpha, \mathbf{T}^\beta)/\xi] \in (0, 1], \quad (1.5)$$



Here,  $\xi$  is the correlation length and,  $\alpha$ ,  $\beta$  and  $\gamma$ , indices are used to represent individuals as opposed to,  $a$  and  $b$ , which would indicate points in phenotype space - multiple individuals may exist with the same phenotype vector. As phenotypes are essentially defined by their interaction sets we shall herein refer to the correlation measure between two phenotypes as,  $\mathbf{C}(\mathbf{T}^\alpha, \mathbf{T}^\beta)$  ie. as an entry in the correlation matrix,  $\mathbf{C}$ . The clustering of non-zero  $\mathbf{J}$ -matrix entries also exhibits a short range correlation length similar to (1.5). The probability of a random phenotype,  $\alpha$  having an interaction with another phenotype,  $\gamma$  is  $\theta$  but if  $\alpha$  has a vector similar to another,  $\beta$  that has an interaction with  $\gamma$  then this probability is elevated by virtue of the correlation (and therefore we have clustering). Importantly, interacting uncorrelated phenotypes take values  $\mathbf{J}(\mathbf{T}^\alpha, \mathbf{T}^\beta)$  and  $\mathbf{J}(\mathbf{T}^\beta, \mathbf{T}^\alpha)$  that are uncorrelated thus permitting any interaction type (predator-prey, mutualistic etc.) to exist in principle. The process with which we achieve these quantifiable correlations is quite involved so we have elected to give an overview here whilst referring the reader to the original paper for a fuller explanation [22].

In order to more truly represent evolution in a real ecosystem several changes were made to the original model. These changes can best be described via the modifications made to the weight function shown previously in (1.1),

$$H(\mathbf{T}^\alpha, t) = a_1 \frac{\sum_{\mathbf{T} \in \mathcal{T}} \mathbf{J}(\mathbf{T}^\alpha, \mathbf{T}) n(\mathbf{T}, t)}{\sum_{\mathbf{T} \in \mathcal{T}} \mathbf{C}(\mathbf{T}^\alpha, \mathbf{T}) n(\mathbf{T}, t)} - a_2 \sum_{\mathbf{T} \in \mathcal{T}} \mathbf{C}(\mathbf{T}^\alpha, \mathbf{T}) n(\mathbf{T}, t) - a_3 \frac{N(t)}{R(t)}. \quad (1.6)$$

The sums are made over the points in phenotype space,  $\mathcal{T}$ , and the occupancies (population associated with each phenotype),  $n(\mathbf{T}, t)$  are used to account for the multiplicity of individuals with the same phenotype vector. We consider here a well mixed system of constant spatial size, although spatial extent is not explicitly considered. For clarity we reiterate at this point that the phenotype space is a pre-defined, complete set of all possible phenotypes and it is evolution and contingency that select the actualised phenotypes in the evolved system. The  $\mathbf{J}$ -matrix is similarly a pre-defined complete set of all possible interactions for all possible phenotypes that may exist *in potentia*.

The correlation measure is used in the first term of the weight function to restrict the impact of the interaction sum. It represents the fact that interactions are shared amongst members of the same species. For example, a tree may provide a volume of fruit to a solo member of a species but the provision must be shared with reduced efficacy if there are many members. So the overall effect of this denominator, on an individual, is to dampen its interaction sum as a whole thus representing the effect of distributing all interaction effects amongst the individual's own species members. Another example could be a wildebeest in proximity to a lion. The negative predatory effect (or predation probability) of the single lion on that specific wildebeest is decreased if there are many wildebeest about to select from. The interaction is damped by the presence of other members of the individuals own species. This aspect is not

ubiquitous in species interactions but does feature in many cases. It is less applicable at low population densities but our system generally maintains significant species populations even at low resource levels. Lowering the resource tends to reduce diversity rather than distinct species populations so this limitation was deemed to be acceptable. In recognition of the phenotypic variation inherent in a species, the sum over the correlation values,  $\mathbf{C}(\mathbf{T}^a, \mathbf{T}^b)$ , accounts for the fact that species members have different but similar phenotypes. This can be seen as a species description in itself.

When referring to diversity we specifically mean the wildtype diversity but all interaction effects from all extant phenotypes, mutants included, are accounted for in every interaction sum. The mutant cloud itself is in fact very sparse as we have elected to use a low mutation rate,  $P_{mut} = 0.0002$ . As a consequence, the phenotype distribution is essentially a set of delta points of high occupancy with infrequent mutants existing with low occupancy. This makes the recognition of the wildtype diversity very simple as each species is massively dominated by the population of the wildtype.

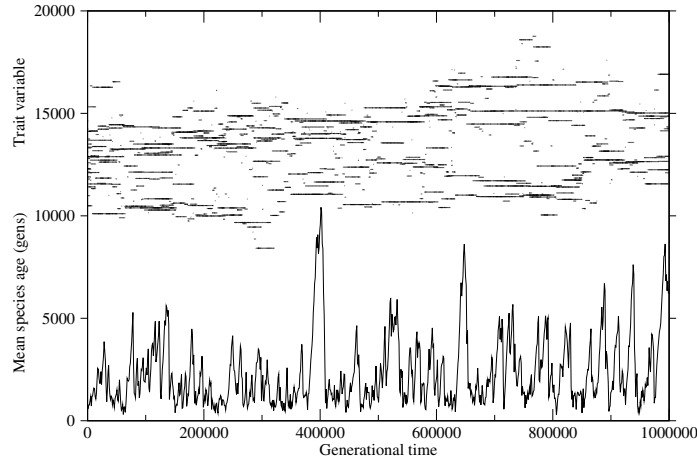
The second term of the weight function represents intra-specific competition and uses the same correlation measure as before. Similar but distinct phenotypes are likely to be in competition for resources, space etc. that are specific to their niche. The correlation measure accounts for this similarity. The original model operated without such a term, but it was deemed essential here to allow sufficient wildtype diversity to develop. As otherwise mutualistic pairs of phenotypes dominate the system, competitively excluding all others.

The third term represents competition for a conserved vital resource that all phenotypes require for survival. Any successful reproduction event produces an offspring that assumes a unit of resource from the bath,  $R(t)$ . The conservation requirement,  $R(t) + N(t) = constant$  means we have a carrying capacity for the system as a whole. It's functional form represents the number of system members competing per unit of available resource.

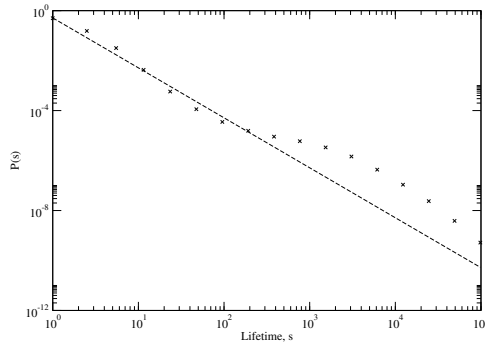
The parameters,  $a_1 = 0.5$ ,  $a_2 = 0.01$ ,  $a_3 = 0.2$  are the selection, conspecific competition and resource competition parameters respectively. These are subjectively chosen to allow interaction controlled dynamics and a sufficient wildtype diversity to develop. A value of  $\theta = 0.05$  is used throughout the simulations.

### 1.3.1 Lifetimes and extinctions

The macroscopic dynamics of the system have similarities with the original model as intermittency occurs but the phases of stability are not so disjoint. Fig. 1.4 shows the evolution of the occupancy in phenotype space where phenotypes can be seen to drift, speciate and often go extinct. The transitions are there to some degree as demonstrated in the accompanying plot of the mean age of the system wildtype phenotypes. There are regular collapses of the system where long term phenotypes go extinct and the extant set becomes partially reset. Many phenotypes survive these transitions, however.



**Fig. 1.4.** The upper plot represents the time evolution of the extant phenotype distribution projected onto one of the 16 trait variables. The lower plot represents the mean age of the extant phenotypes weighted according to population. Stable periods can be seen where the mean age increases near linearly until a collapse occurs and new species are introduced.

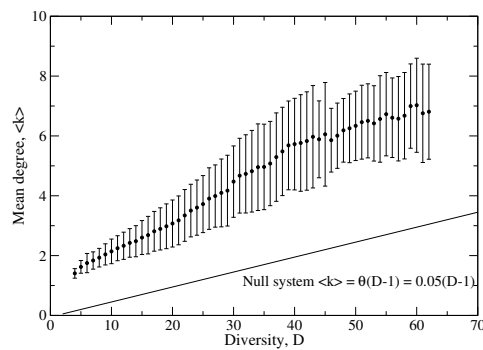


**Fig. 1.5.** An example of a lifetime distribution for a low resource system. The form is nearly a power-law with a distinct deviation. Dotted line represents a functional power law,  $P(s) \sim s^{-\gamma}$  with exponent,  $\gamma=2$ .

The lifetime distribution associated with these extinctions has a near power-law form,  $P(s) \sim s^{-\gamma}$  as can be seen in Fig. 1.5. It has often been suggested that the such real system distributions are power laws with proposed exponents around  $\gamma \simeq 2$  but this still an open debate. Our data implies a near power law form that is exhibited both in other models and the original Tangled Nature model itself, [23] [24] [25] [26] (and also [27] in reference to [28]). By increasing the resource in our system we can shift the mean lifetime

to lower values which is a consequence of an increase in the mean population density. This phenomenon arises as a population density increase leads to an elevation in the rate of mutant offspring per generational time step resulting in the system destabilising more frequently. It has recently been discovered that species diversity may negatively correlate with average species duration so implying a greater species turnover rate [29]. The population density and wildtype diversity are positively correlated in our model due to the presence of the intra-specific competition term in (1.6). As this aspect is an arguably realistic portrayal of niches within an ecosystem we can provide agreement that wildtype diversity does incur a greater species turnover rate.

### 1.3.2 Species interaction network



**Fig. 1.6.** Plot of mean degree,  $\langle k \rangle$ , against wildtype diversity,  $D$ . Each point is an ensemble average of data taken at each instance of each diversity achieved at the resource level,  $R = 30000$ , with error bars representing the standard error. Fifty simulation runs of a million generations each were made, so given the fluctuations involved, a large range of diversities were repeatedly encountered. The lower line is the mean degree of the null system  $\langle k \rangle = \theta(D - 1) = 0.05(D - 1)$ .

The species interaction network is defined by the realised adjacency matrix of the wildtype diversity,  $D$ . The structural formation of the network arises from evolutionary processes occurring at the level of individuals but it has dynamic global properties as a result of this. The evolutionary pressure to achieve positive interactions leads to a global increase in the mean degree,  $\langle k \rangle$  of the extant species set - here the mean degree,  $\langle k \rangle$  is the number of interactions of any type a species has with other species, averaged over all species. We can compare the evolved system to a ‘null’ one consisting of a *random* species set. For this set, the connectance values will conform on average to the value of  $\theta$ , the proportion of non-zero interactions in the  $\mathbf{J}$ -matrix. Fig. 1.6 shows the mean degree versus wildtype diversity taken at the highest resource level,  $R = 30000$  (all other resource levels exhibit similar behaviour). In

comparison to the null system mean degree,  $\langle k \rangle_\theta = \theta(D-1) = 0.05(D-1)$ , it is obvious that the network achieves a status that far exceeds this value. It is generally unclear whether greater stability is achieved through mean degrees that are either lower than the network percolation threshold or higher [30] [31] [32] [33], but here the system naturally evolves towards the latter. The mean interaction strength also increases to a positive value that is diversity-dependent, but around one standard deviation of the normal distribution values of the null system. Our system therefore has a global response whereby it evolves to assume networks involving greater numbers of positive interactions. This process is strongly enhanced by mutualistic interaction types and whilst the interactions are generally asymmetric their strengths are more positively biased in the evolved systems.

The correlations in inheritance have an unusual effect on the degree distributions of the assumed networks. They all conform closely to exponential,  $P(k) \sim e^{-k/\langle k \rangle}$ , as opposed to the null system binomial distributions. We attribute this degree form to the inheritance process which directs the evolution of the network. Simulations performed with zero correlations lead to networks that display enhanced mean degrees but without any deviation from the binomial degree distribution. The correlation appears to be a necessary requirement for these longer tailed distributions to appear.

## 1.4 Overview

Our focus has been on emergence and large numbers of interdependent organisms. We have described a set of minimalistic models in which reproduction probability is the only explicit phenotypical property of individuals, and is not defined as an intrinsic property but rather determined for each type of organism through its interaction with other extant types. Basic properties, such as the existence and formation of separate species, and the existence of quasi-stable states separated by periods of mutation, are emergent from the dynamics.

In the simple model in Sect. 1.2.1, transitions between q-ESS are rare events, even though mutations are common. On transition, all species perform an evolutionary random walk *with an effective selection gradient due to interactions*, meaning the distance covered is much larger than predicted by a standard random walk of equal time (i.e. the jumps resemble Levy Flights [34]).

The Tangled Nature model, and similar approaches, should be considered as complimentary to more detailed models such as those in adaptive dynamics, in which ecosystems are established with many shared features with observations on real biological systems. The sort of questions which can be asked of our model are very different. In particular, we do not intend to model a specific situation, but to capture features common to a great many evolving systems, such as the effect of interaction on the stability of an ecosystem, and

the time dependence of properties such as stability, average population, the form of interactions, etc.

By allowing evolution to select the network of interacting species from an *in potentia* configuration space we are able to study issues concerning functionality and stability. Our approach allow us to consider the network properties of the ecosystem as collectively selected quantities.

## References

1. P. A. Abrams. ‘adaptive dynamics’ vs. ‘adaptive dynamics’. *J. Evol. Biol.*, 18:1162–1165, 2005.
2. Barbara Drossel. Biological evolution and statistical physics. *Advances in Physics*, 50(2):209–295, 2001.
3. J. Maynard Smith. *Evolution and the theory of games*. Cambridge UP, The Edinburgh Building, Cambridge CB2 2RU, 1982.
4. Barbara Drossel, Paul G. Higgs, and Alan J. McKane. The influence of predator-prey population dynamics on the long-term evolution of food web structure. *J. Theor. Biol.*, 208:91–107, 2001.
5. G. Caldarelli, P.G. Higgs, and A.J. McKane. Modelling coevolution in multi-species communities. *J. theor. Biol.*, 193:345–358, 1998.
6. John N. Thompson. The evolution of species interactions. *Science*, 284:2116–2118, 1999.
7. Peter Hammerstein. Darwinian adaptation, population genetics and the street-car theory of evolution. *J. Math. Biol.*, 34:511–532, 1996.
8. D.L. DeAngelis and L.J. Gross. *Individual-based models and approaches in ecology : populations, communities and ecosystems*. Chapman & Hall, New York, 1992.
9. Donald L. DeAngelis and Wolf M. Mooij. Individual-based modelling of ecological and evolutionary processes. *Ann. Rev. of Ecol., Evol., and Sys.*, 36:147–168, 2005.
10. 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.
11. Paul Anderson and Henrik Jeldtoft Jensen. Network properties, species abundance and evolution in a model of evolutionary ecology. *J. Theor. Biol.*, 232:551–558, 2005.
12. Simone Avogadro di Collobiano, Kim Christensen, and Henrik Jeldtoft Jensen. The tangled nature model as an evolving quasi-species model. *J. Phys A*, 36:883–891, 2003.
13. Matt Hall, Kim Christensen, Simone A. di Collobiano, and Henrik Jeldtoft Jensen. Time-dependent extinction rate and species abundance in a tangled-nature model of biological evolution. *Phys. Rev. E*, 66(011904), 2002.
14. Stephen Hubbell. *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton UP, 41 William Street, Princeton, New Jersey 08540, 2001.
15. Per Arne Rikvold and R.K.P. Zia. Punctuated equilibria and 1/f noise in a biological coevolution model with individual-based dynamics. *Phys. Rev. E*, 68(031913), 2003.

16. R.K.P. Zia and Per Arne Rikvold. Fluctuations and correlations in an individual-based model of evolution. *J. Phys. A*, 37:5135–5155, 2004.
17. Rick Durrett and Simon Levin. Spatial models for species-area curves. *J. theor. Biol.*, 179:119–127, 1996.
18. Anne E. Magurran and Peter A. Henderson. Explaining the excess of rare species in natural species abundance distributions. *Nature*, 422:714–716, 2003.
19. Daniel Lawson and Henrik Jeldtoft Jensen. The species-area relationship and evolution. *arXiv:q-bio.PE*, (0412024), 2005.
20. Daniel Lawson, Henrik Jeldtoft Jensen, and Kunihiko Kaneko. Diversity as a product of interspecific interactions. *arXiv:q-bio.PE*, (0505019), 2005.
21. Volkan Sevim and Per A. Rikvold. Effects of correlated interactions in a biological coevolution model with individual-based dynamics. *arXiv:q-bio.PE*, (0507040), 2005.
22. Simon Laird and Henrik Jeldtoft Jensen. The tangled nature model with inheritance and constraint: Evolutionary ecology restricted by a conserved resource. *arXiv:q-bio.PE*, (0510008), 2005.
23. Russell K. Standish. Ecolab: where to now? *Complexity International*, 3, 1996. (See <http://parallel.hpc.unsw.edu.au/rks/docs/newman-model/newman-model.html> for a plot of the full lifetime distribution.)
24. 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.
25. Debashish Chowdhury, Dietrich Stauffer, and Ambarish Kunwar. Unification of small and large time scales for biological evolution: deviations from power law. *Phys. Rev. Lett.*, 90(6):068101, 2003.
26. Debashish Chowdhury and Dietrich Stauffer. Food web based unified model of macro- and microevolution. *Phys. Rev. E*, 68:041901, 2003.
27. Debashish Chowdhury and Dietrich Stauffer. Sole-manrubia model of biological evolutions: some new insights. *Physica A*, 318:461–468, 2003.
28. Ricard V. Sole and Susanna C. Manrubia. Extinction and self-organised criticality in a model of large scale evolution. *Phys. Rev. E*, 54(1):42–45, 1996.
29. Brent C. Emerson and Niclas Kolm. Species diversity can drive speciation. *Nature*, 434:1015–1017, 2005.
30. Robert M. May, editor. *Stability and complexity in model ecosystems: monographs in population biology*. Princeton UP, 1974.
31. Kevin Shear McCann. The diversity stability debate. *Nature*, 405:228–233, 2000.
32. Daniel T. Haydon. Maximally stable model ecosystems can be highly connected. *Ecology*, 81(9):2631–2636, 2000.
33. Jennifer A. Dunne, Richard J. Williams, and Neo D. Martinez. Food web structure and network theory: the role of connectance and size. *Proc. Natl. Acad. Sci. USA*, 99(20):12917–12922, 2002.
34. Ralf Metzler and Joseph Klafter. The random walk’s guide to anomalous diffusion: a fractional dynamics approach. *Physics Reports*, 339:1–77, 2000.