

Diversity as a product of inter-specific interactions

Daniel Lawson*, Henrik Jeldtoft Jensen*[†] and Kunihiko Kaneko[‡]

September 19th, 2005

Abstract

We demonstrate diversification rather than optimisation for highly interacting organisms in a well mixed biological system by means of a simple model of coevolution. We find the cause to be the complex network of interactions formed, allowing species that are less well adapted to an environment to succeed, instead of the ‘best’ species. This diversification can be considered as the construction of many co-evolutionary niches by the network of interactions between species. The model predictions are discussed in relation to experimental work on dense communities of the bacteria *Escherichia coli*, which may coexist with their own mutants under certain conditions. We find that diversification only occurs above a certain threshold interaction strength, below which competitive exclusion occurs.

1 Introduction

Understanding how diversity arises through evolution and is sustained in an ecosystem is an important issue. One of the key questions therein is whether interactions between organisms enhance or suppress diversity. If there is no explicit symbiotic interaction, it would be expected that the competition for a given resource leads to exclusion of many types. This results in monodominance, i.e., the survival of the fittest, as determined by Gause’s competitive exclusion principle [1]. In contrast, in the presence of strong interactions, diversification has been shown to occur both in numerous models and in experiment[2][3]. We attempt to understand the relationship between interaction and diversity at a general level, and will relate our work to experimental findings on evolution in *Escherichia coli* [4].

We show that the diversification can indeed be facilitated by the interaction, using a range of different fitness concepts. We do this by adopting a slightly modified version of the Tangled Nature (TaNa) model [5, 6, 7]. In addition to the standard, inter-specific interaction in the TaNa model, we allow types to differ in ‘intrinsic fitness’ - the fitness of a type in the environment, in the absence of other types. A self-supporting, dominant genotype may coexist with, or be displaced by, a number of other genotypes that are less efficient competitors for the resource individually, provided that strong enough interactions are permitted. Diversity is maintained via the complex network of interactions, and we demonstrate a cutoff interaction strength below which monodominance persists. We split the ‘intrinsic fitness’ of a type into *density dependent* (i.e. the interaction with own type) and *density independent* parts, and study them separately. The conditions on the interaction strength are, respectively: (1) the net positive interaction with *other* types is greater than the *density dependent* fitness, (2) the net positive interaction with other types is greater than the *density independent* fitness *difference* between types.

The idea that diverse states can be supported by interaction is not new. Gause’s competitive exclusion principle states, in the general case, [8] that “the dimension of the environmental interaction variable is an upper bound for the number of species that can generically exist at steady state”. Although environmental interactions cannot in general be uniquely identified, and so this dimension is not known, it is still possible

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

[‡]Department of Basic Science, University of Tokyo and ERATO Complex Systems Biology, JST, Komaba, Meguro-ku, Tokyo 153-8902, Japan

[†]Author for correspondence (h.jensen@imperial.ac.uk)

to support high diversities robustly [9][10] (i.e. still supported with a small change in the environmental parameters). This diversity remains finite even in the case of an infinite environment interaction variable [11], as species must be ‘different enough’ [12] to coexist stably. In addition, a greater number of species than environmental factors may be supported by oscillations or chaos, e.g. [13][14][15]. The appearance of diversity in a system with strong interaction is therefore not a surprise by itself, as each interaction contributes to the effective dimensionality. Still, it is important to understand how diversity is mediated by the interaction.

We follow Robert May [16] in using generalised, random interactions. In his book he discusses the ecological implications of such models in detail; we will be looking at the effects of evolution on that stability. Such simple models may most accurately describe molecular replicators (e.g. [17]), and simple bacterial systems. However, because only the net interaction and reproduction probability is considered, there are other biological cases which can be approximated by this approach.

Our model is individual based without any individual aging, considering a generalised system of organisms so that interactions are random. Genotype space is predefined, so that the interactions between all possible organisms are fixed from the start, and mutations are local. In the spirit of other null models, these interactions are not correlated in this version of the model. We consider one reproduction attempt as the basic unit of time, and we allow mutation to occur during the population dynamics. The total population is a result of the dynamics. We will consider an intrinsic fitness landscape in the presence of strong interactions. For general background reading on individual based modelling and for discussion on many basic features the reader is referred to [18], to [19] for a population dynamics perspective, and [20] for a genetics point of view.

The features described above mean that the existence of diversity can be seen to arise in the following way: from an initially monodominant state, we find that evolution forces a search of genotype space for the most stable configurations. Often, these states are diverse, provided the intra-specific competition exceeds inter-specific competition (or, equivalently, the beneficial inter-specific interaction is greater than the intrinsic fitness). Such diverse states do not exist for low interaction strength, and all states are diverse in the limit of very high interaction strength. Stability is determined by the properties of a given configuration in genotype space, and states are, on average, more stable as time progresses. In addition, we find a sharp threshold in interaction strength below which diversity does not occur.

2 Definition of the Model

We now define the Tangled Nature model. Individuals are represented as a vector $\mathbf{S}^\alpha = (S_1^\alpha, S_2^\alpha, \dots, S_L^\alpha)$ in genotype space \mathcal{S} . The S_i^α take the values ± 1 , and we use $L = 20$ throughout, giving $2^{20} = 1048576$ possible types. Each \mathbf{S} string represents an entire type with unique, uncorrelated interactions. The small value of L is necessary for computational reasons as all types exist *in potentia* and have a designated interaction with all other types¹. There are therefore $(2^{20})^2$ interactions to be considered in this model. We consider random interactions for simplicity, which would be correlated in reality. Introducing significant correlation whilst maintaining randomness in this relatively small hypercubic genotype space has proved difficult, and so we consider uncorrelated interactions here. Note that controlled correlations have been achieved in another version of the model [21].

We refer to individuals by Greek letters $\alpha, \beta, \dots = 1, 2, \dots, N(t)$. Points in genotype space are referred to as $\mathbf{S}^a, \mathbf{S}^b, \dots$, and any number of individuals may belong to a point in genotype space \mathbf{S}^a .

In the original TaNa model, individuals α are chosen randomly and allowed to reproduce with probability p_{off} :

¹When discussing the model, we refer to points in genotype space as a type. It is a matter of interpretation whether we consider genotype space to be ‘coarse-grained’ (resulting in each genotype being a different species - valid when k and ϵ are ‘large’ so that genotype differences affect reproduction probability greatly; see Eq. 2 for definitions), or whether we consider genotype space to be a small sample of a much larger space, meaning genotypes are *types* of a base species (which would be valid when k and ϵ are small, and so all genotypes have similar reproduction probabilities). As we operate in neither extreme and reproduction is asexual, the distinction between species and type is difficult.

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

They are then killed with probability p_{kill} , which is a constant parameter. The difference between the original model and the one used here is the definition of the weight function $H(\mathbf{S}^\alpha, t)$. The original version used was:

$$H_0(\mathbf{S}^\alpha, t) = \frac{k}{N(t)} \sum_{\mathbf{S} \in \mathcal{S}} J(\mathbf{S}^\alpha, \mathbf{S}) n(\mathbf{S}, t) - \mu N(t) \quad (2)$$

Here k ($\equiv 1/c$ from previous papers) determines the maximum strength of interactions, $N(t)$ is the total number of individuals at time t and $n(\mathbf{S}, t)$ is the number of individuals with genotype \mathbf{S} at that time. The *interaction matrix* $J(\mathbf{S}^\alpha, \mathbf{S})$ represents all possible couplings between all genotypes, with $J_{ii} = 0$ always and $J_{ij} = J_{ji} = 0$ with probability Θ . If the interaction is not zero, then J_{ij} and J_{ji} are both generated randomly in the range $(-1, 1)$, so that mutualism, predator-prey and competition are all possible, but amensalism and commensalism only occur in the case when one interaction is randomly generated to be very small. Since the functional form of $J(\mathbf{S}^a, \mathbf{S}^b)$ does not affect the dynamics, provided that it is non-symmetric with mean 0, we choose a form of the interaction matrix that speeds computation [5]. In the analysis sections, we will use shorthand versions: J_{ab} as the interaction of an individual from type b on an individual from type a , and n_a as the number of individuals with genotype a .

In the extended model we consider here, we also allow an intrinsic fitness term, representing the different ability of types to survive in the environment. There are at least two possible ways of doing this - either as an *density dependent* fitness term, or a *density independent* fitness. With these fitness concepts, correlations can be introduced easily so we will look at both the case of uncorrelated and correlated landscapes. The correlation we choose is a type of *Fujiyama* landscape² [18] defined as follows. One type α has a fitness of 1, and with each mutational step away from this type we subtract Δ ($= 0.1$ in simulations), down to a minimum of 0. An uncorrelated landscape is generated with each type having a fitness drawn uniformly from $(0, 1)$. See Sect. 3.2 for an explanation and Fig. 2 for results. The modified weight functions take the following forms:

1. Density Dependent fitness (or equivalently, the interaction with an individual's own type), defined by:

$$H_d(\mathbf{S}^\alpha, t) = H_0(\mathbf{S}^\alpha, t) + \frac{\epsilon}{N(t)} n(\mathbf{S}^\alpha, t) E(\mathbf{S}^\alpha) \quad (3)$$

Here, ϵ is the magnitude of the density dependent part of the 'intrinsic fitness strength' and $\epsilon E(\mathbf{S}^\alpha)$ is the intrinsic fitness³ of individual α . $E(\mathbf{S}^\alpha)$ 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:

$$H_i(\mathbf{S}^\alpha, t) = H_0(\mathbf{S}^\alpha, t) + \epsilon_I E(\mathbf{S}^\alpha) \quad (4)$$

Here, ϵ_I is the magnitude of the density independent part of the intrinsic fitness strength. E_i is again determined by the case studied:

²Named because it has a single, large peak.

³We do not merge it with the J term in Eq. 2 in order to allow different values of the strengths.

102 2(a) Uncorrelated, density dependent intrinsic fitness landscape.

103 2(b) Correlated, density dependent intrinsic fitness using a ‘Fujiyama’ landscape.

104 All individuals of the same type will have the same weight function and therefore the same offspring
105 probability at a given time; i.e. if individual α was from type a then $H(S^a) = H(S^\alpha)$. Reproduction occurs
106 asexually, and on a successful reproduction attempt two daughter organisms replace the parent, with each
107 \mathbf{S}_i^α mutated (flipped from 1 to -1, or from -1 to 1) with probability p_{mut} . Thus mutations are equivalent to
108 moving to an adjacent corner of the L -dimensional hypercube in genotype space, as discussed in [5].

109 A time-step consists of choosing an individual⁴ α randomly, and processing according to:

- 110 • α is allowed to reproduce with probability p_{off} .
- 111 • α is killed with probability p_{kill} . (if α reproduced, it is one of the two daughter organisms that is
112 killed).

113 We define a generation as the amount of time for all individuals to have been killed, on average, once.
114 For a stable population size, this is also the time for all individuals to have reproduced once, on average. The
115 diversity is defined as the number of genotypes with occupancy greater than 20 to eliminate unsuccessful
116 mutants from our count, and is called the wildtype diversity. This definition comes from the observed
117 population structure, as discussed in the next section. The total number of genotype points occupied is
118 approximately L times the wildtype diversity.

119 Unless otherwise stated, the parameters used will be: $\Theta = 0.2$, $\mu = 0.01$, $p_{mut} = 0.015$, $\epsilon = 2.0$ and
120 $p_{kill} = 0.1$; see [5] for more details. These parameters are selected to allow the population to remain
121 moderately high (to avoid accidental extinction), and to be well away from the mutation threshold present
122 in this system [6] (c.f. [17]: as the mutation probability is increased, the time spent for the system to find
123 a quasi-stable state - described in the next chapter - increases until it becomes infinity, and the quasi-stable
124 structure described below is lost). The results are robust to moderate parameter changes; that is, the same
125 qualitative behaviour can be found for all small parameter changes by making an appropriate small change
126 in the other parameters. In particular, the cutoff for diversity persists over a region of other parameters,
127 although the value of the cutoff might change slightly.

128 For case 1(a) and case 2(a), the initial conditions are determined by allowing the system to find a
129 monodominant state by running the system for 5000 generations with all interaction disabled ($k = 0$)
130 (starting from a random set of individuals), thus one of the best competitors in the initial set is selected.
131 Then the interaction was enabled by setting k to the desired value. For case 1(b) and 2(b) of the Fujiyama
132 landscape, we simply start the whole population on the fitness peak.

133 2.1 The reproduction equation

134 Eq. 2 consists of two terms: the first is an *average interaction* term, and the second a *resource competition*
135 term with all other individuals. Thus k controls the strength of the average interaction and therefore has
136 relation to a density (as closer individuals will interact more strongly). This makes our model valid for
137 systems in which the population density is roughly constant in time, and individuals compete for a single,
138 fixed quantity of resource (determined by μ). Clearly this is true of all systems in steady state and will be
139 approximately true of many other systems.

140 The introduction of Eq. 1 is in order to turn the infinite ranged H into a probability. One could instead
141 treat H as a stochastic growth rate, and use e.g. the Gillespie algorithm [22][23]; this defines the time step
142 as the expected waiting time to the next event instead of using a fixed time step as we do. We do not
143 choose to do it this way for two reasons. Firstly, there is no reason to assume that the rate of increase of a
144 species (given by $p_{off} - p_{kill}$) will be linear in the quantities defined in H . We have exchanged linearity for

⁴In previous versions a different individual was chosen for reproduction and killing actions. Here we select only one individual and process it for reproduction and killing for code efficiency reasons - above the level of fluctuations the two methods are equivalent.

145 the logical simplicity of having an explicit p_{off} . Secondly, our method is computationally easier. The form
 146 of Eq. 1 was chosen for its simplicity, and the particular form is not essential to the model. Almost any
 147 monotonic continuous mapping of H to p_{off} will give equivalent qualitative results (this is tested for a few
 148 functions, although no proof can be given due to the complexity of the results).

149 Our reproductive form, then, assumes that interactions sum additively only when close to equilibrium,
 150 and that the reproductive advantage gained decreases for additional interactions giving a nonlinear form.
 151 This can be considered as a rule of diminishing returns - if there is a net benefit for an individual, each
 152 additional benefit results in a smaller effect (in p_{off}). The *ordering* of offspring probabilities $p_{off}(\mathbf{S}^\alpha, t)$
 153 is unchanged by this map; only the *differences between* offspring probabilities will change. The effect is
 154 therefore limited to fluctuations as all features of $\{H(\mathbf{S}^\alpha, t)\}_\alpha$ will exist in $\{p_{off}(\mathbf{S}^\alpha, t)\}_\alpha$ as well. The form
 155 of this equation does not appear in many mean field equations - see Sect. 3.2 and [24].

156 Using a constant killing probability p_{kill} is a simplifying approximation, as selection certainly will act
 157 by differential killing as well as differential reproductive success. However, the dynamics in our model are
 158 qualitatively the same without this restriction (provided p_{kill} is not close to 1). This symmetry between
 159 selection (i.e. killing probability) and reproductive ability exists in our model because we do not include
 160 any individual aging. More complex relations are required in models which permit reproduction only for
 161 individuals which have reached a certain age [25].

162 To understand the meaning of the additional density dependent fitness term in case 1 (Eq. 3), we consider
 163 the weight function of a system with only one type a , $H(S^a) = \epsilon E(S^a) - \mu n(S^a)$ since $N = n(S^a)$. If we
 164 assume that the system is in a steady state ($P_{off} = P_{kill}$), then $H(S^a) = H^* = -\ln(\frac{1}{p_{kill}} - 1)$, which is
 165 constant. Thus we find $n(S^a) = \frac{\epsilon E(S^a) + |H^*|}{\mu}$, meaning that $E(S^a)$ determines how numerous type S^a would
 166 be if alone in the system. The same result is obtained for the case 2 - the differences are apparent only when
 167 more than one species is introduced; see Sect. 3.2.

168 3 Results

169 3.1 Observed behaviour

170 As in the basic Tangled Nature model, the system experiences a number of ‘quasi-Evolutionarily Stable
 171 Strategies’ (called q-ESSs⁵ for brevity), during which a single genotype or set of genotypes is present with
 172 constant average occupancy. These q-ESS may end abruptly, leading to a transition phase before a new
 173 q-ESS is found. For the parameter ranges we study, the transition phase usually lasts for tens of generations
 174 and so is instantaneous on an evolutionary timescale. This behaviour is shown in Fig. 1, with some major
 175 events labelled. The qualitative behaviour described here is observed regardless of the form of the intrinsic
 176 fitness.

177 The q-ESS phases have several species (from 1 for small k up to about 6 for large k) with large, stable
 178 populations, and we call these species wildtypes. The wildtypes are generally separated in genotype space
 179 and are surrounded by mutant types with much lower population. These mutants do not have p_{off} large
 180 enough to counteract the death rate, and as such are dependent on the mutations from the wildtype species
 181 for their existence; therefore their species lifetimes are short. The species abundance distribution is log-
 182 normal on average [5] if only wildtypes are considered. Transitions between q-ESS states last for only tens of
 183 generations at these parameter values; however it is still possible for a species to mutate over large distances
 184 in genotype space in this time. When a new type is successful, it increases in number at the expense of the
 185 types it interacts with; by this time another type which benefits from the first may be found, and so there
 186 is an effective ‘selection gradient’ against the dominant species until a q-ESS is reached [5].

187 Fig. 2 shows that fitness of case 2(a) (uncorrelated, density independent landscape) yields a non-unity
 188 diversity at $k = 0$, due to the high level of neutrality in the system. This is because species with similar
 189 (high) fitness are plentiful (as ϵ_I is uniform distributed) and the transition time between them is high, so
 190 average diversity measures pick up diverse states often. This would not be the case if density independent

⁵Named after Maynard Smith’s ‘Evolutionarily Stable Strategies’. ‘Quasi’ refers to the (in)stability of the strategy to collective stochastic fluctuations. See [26] for a more full discussion.

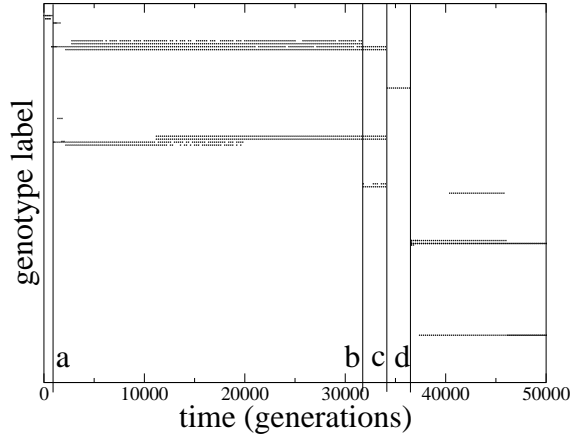


Figure 1: A sample run ($k = 10$, density dependent fitness) showing all genotypes with occupations greater than 20 as an unordered genotype label. Times shown correspond to different cross-over situations. (a) is from the original monodominance to a diverse state, which 100 generations later becomes more diverse again. (b) shows a cross-over from one diverse state to another, which at (c) becomes a new monodominant state. Then at (d) the system returns to a new diverse state.

191 fitness peaks were very sparse and of different heights, in which case Gause's Competitive Exclusion Principle
 192 would act.

193 In contrast, the appearance of diversity in cases 1(a), 1(b) and 2(b) follow another pattern. In these
 194 cases, at low k there is monodominance, and at high k there is high diversity. Also in all cases, the cross-over
 195 region behaves in the same way, with a rapid increase from zero at some characteristic k_{min} , and then a
 196 steady increase towards a saturation diversity. Diverse states occur for lower k values for case 2(b) than
 197 case 1, and the existence and approximate value of a cross over in these cases can be shown by a mean field
 198 argument (see Sect. 3.2).

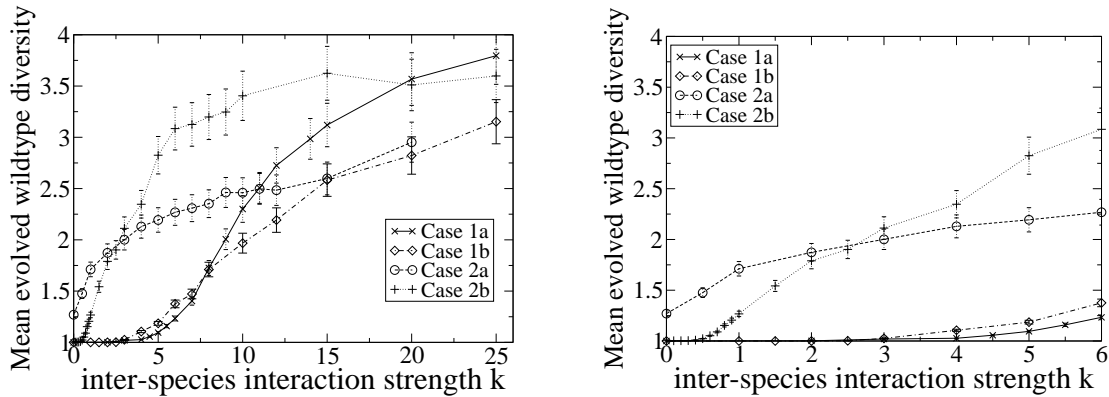


Figure 2: Left: k dependence of the average diversity of an evolved system, taken for 40000 – 50000 generations and 500 runs per data point for the case 1, density dependent ($\epsilon = 2$, $\epsilon_I = 0$) fitness, and case 2, density independent fitness ($\epsilon = 0$, $\epsilon_I = 2$). Case (a) has rugged random fitness landscape and case (b) has a Fujiyama landscape. Right: A closer look at the low- k region. Note that case 1(a), 1(b) and 2(b) display the same qualitative behaviour which can be understood at the mean field level, see Sect. 3.2.

199 For case (b), we consider the proportion of time that the most efficient type is observed in the system
 200 for varying k values in Fig. 3. We see that the most efficient type is always in existence for small k , and
 201 there is a threshold at around $k = 0.8$ (case 2(b)) or $k = 1.8$ (case 1(b)) above which the most efficient type
 202 is no longer always present - it may be entirely replaced or drop in numbers to the point at which it is not
 203 observed during every generation. This means that in case 1 the most efficient type can be replaced, but not
 204 coexist, for a range of k ; for case 2, the most efficient type can coexist without possibility of replacement for

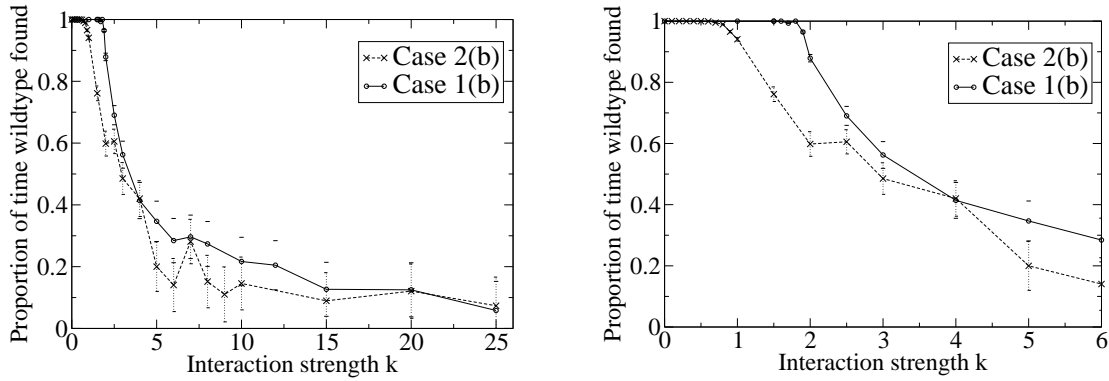


Figure 3: Left: The proportion of generations that the most efficient type is observed in for case 1(b) (density dependent fitness) and 2(b) (density independent fitness), as a function of interaction strength k , averaged over 100 runs per k value. Right: A closer look at the low k region.

We now analyse case 1(a) (uncorrelated density dependent fitness) in more detail. Fig. 4(a) shows the selective drop of E_i with increasing k . The appearance of diversity is clear when considering the ratio of total interactions to intrinsic fitness, $R = \langle \frac{k \sum J_{ij} n_j}{\epsilon E_i n_i} \rangle$ shown in Fig. 4(b). At $k \approx 5.55$, $R = 1$, so the average interaction is greater than the average intrinsic fitness for $k > 5.55$. For $k \gtrsim 10$, $R \sim k$ as interaction becomes the dominant driving force and selection acts to maintain positive interactions. For large k , each term in $\sum J_{ij} n_j$ is already maximised by selection, and is therefore independent of k (again ignoring fluctuations).

For $k \leq 5.55$, interactions do not contribute to fitness of the wildtype. For $k \in (5.55, 10)$, the relative importance of the two selection forces changes. Below $k = 5.55$ diverse q-ESS states are not found⁶ as interaction is always weaker than intrinsic fitness leading to monodominance. Manual examination of the runs confirms that the diverse states found for $k < 5.55$ never have temporal stability, and that diverse, temporally stable states exist for $k = 5.55$ which correspond to q-ESS.

217

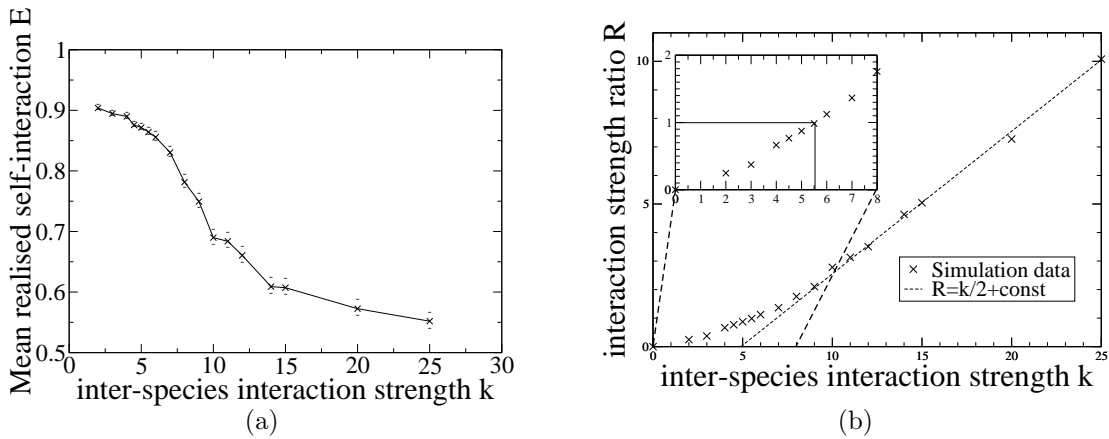


Figure 4: Analysis of case 1(a). Left: $\langle E_\alpha \rangle$ as a function of interaction strength k , which decreases towards the mean value of $1/2$ for large diversities, demonstrating the decline of importance of individual efficiency as a selection factor. Right: Average value of the ratio of the relative interaction strengths R as a function of k , approximately a straight line for $k \gtrsim 10$.

The number of q-ESS switches is higher in high k systems than in low k . At low interaction strengths, monodominant q-ESSs tend to remain for the entire run, with a small possibility of a switch to another

⁶The apparent non-unity diversity below the threshold value appears to be due to occasional mutant fluctuations above the wildtype threshold chosen.

220 monodominant q-ESS with higher intrinsic fitness. As interaction strength increases, the number of q-ESS
 221 switches also increases leading to a greater rate of exploration in genotype space. It was shown in [5] that
 222 the length of q-ESS epochs increases logarithmically with time, as does the average population size. It
 223 would appear that the additional stability of large populations to fluctuations plays a role in determining
 224 the stability of the q-ESS. In addition, the properties of the genotype space near extant types plays a large
 225 role in determining the stability of a state.

226 3.2 Mean field predictions

227 For case 1 and case 2(b), we can show the existence of a cross-over from monodominance to diverse states in
 228 k by a simple argument from the definition of H , although the nature of the cross-over is not determined by
 229 this argument. This mean field result ignores any fluctuations in the system; however, the result provides a
 230 surprising match with observation. We can also simply show that random, density independent fitness will
 231 not allow monodominance.

232 We consider the Weight Function H for the case where species a dominates, and a new species b is added
 233 to the system ($n_a \gg n_b$), and require $n_a + n_b = N \approx n_a$. The requirement for invasion is that $H_b > H_a$ so
 234 that $p_{off}^b > p_{off}^a$. However, when $n_b \gg n_a$ a diverse state must exist if $H_a > H_b$, and competitive exclusion
 235 will occur if $H_a < H_b$ in this case.

236 3.2.1 Case 1: Density dependent fitness

237 For type a , Eq. 3 becomes:

$$H_a = \frac{kn_b J_{ab}}{(n_a + n_b)} + \epsilon_I E_a \frac{n_a}{(n_a + n_b)} - \mu(n_a + n_b) \quad (5)$$

$$\approx \epsilon_I E_a - \mu n_a \quad (6)$$

238 Similarly for type b :

$$H_b = \frac{kn_a J_{ba}}{(n_a + n_b)} + \epsilon_I E_b \frac{n_b}{(n_a + n_b)} - \mu(n_a + n_b) \quad (7)$$

$$\approx kJ_{ba} - \mu n_a \quad (8)$$

239 From above, we require $H_b > H_a$ for invasion; therefore $k > \frac{\epsilon E_a}{J_{ba}}$. For a diverse state, the converse must
 240 be true: $k > \frac{\epsilon E_b}{J_{ab}}$.

241 For case 1(a), E_a has been selected to be high initially, as a was successful on its own; similarly J_{ba} will
 242 be selected to be high to ensure b can proliferate. Thus we can take $E_a = 1$ and $J_{ba} \approx 1$ for both case 1(a)
 243 and 1(b). Thus there is a positive threshold at around $k_{min} \approx \epsilon = 2$. It should be a little less for case 1(a)
 244 as there is a small variation in E_a below unity.

245 The inverse relation provides a different value, as neither E_b nor J_{ab} have been selected for - we take
 246 mean values to get an estimate. E_b is uniform distributed on $(0, 1)$ and thus has mean 0.5, and we take
 247 the mean⁷ $J_{ab} \approx 0.2$. We therefore find that the minimum value of k for a diverse state to exist is around
 248 $k_{min} = 2.5\epsilon = 5$. For $k \in (2, 5)$ invasion is possible coexistence is highly unlikely. Fig. 3 shows that
 249 replacement of the wildtype occurs at just below $k = 2$ and Fig. 2 shows that diverse states exist for
 250 $k > 5.55$.

251 In this case, all interactions should be mutualistic; the ratio $\frac{n_a}{n_b} = \frac{kJ_{ab} - \epsilon E_b}{kJ_{ba} - \epsilon E_a}$ should be positive and of
 252 order 1 for stability; this requires both J_{ab} and J_{ba} are positive.

⁷ J has a mean of zero, but here we are taking the mean of the positive part of the distribution, which has non-zero mean.

253 **3.2.2 Case 2: Density independent fitness**

254 For type a , Eq. 4 becomes:

$$H_a = \frac{kn_b J_{ab}}{(n_a + n_b)} + \epsilon_I E_a - \mu(n_a + n_b) \quad (9)$$

$$\approx \epsilon_I E_a - \mu n_a \quad (10)$$

255 Similarly for type b :

$$H_b = \frac{kn_a J_{ba}}{(n_a + n_b)} + \epsilon_I E_b - \mu(n_a + n_b) \quad (11)$$

$$\approx kJ_{ba} - \mu n_a + \epsilon_I E_b \quad (12)$$

256 Using $H_b > H_a$ we find $k > \frac{\epsilon_I}{J_{ba}}(E_a - E_b)$ for invasion. For diversity the converse must also be true:
 257 $k > \frac{\epsilon_I}{J_{ba}}(E_b - E_a)$. For $k \neq 0$ and case 2(a) of a random fitness landscape, the difference $E_b - E_a$ can be
 258 arbitrarily small and so can always be satisfied for some J_{ab} and J_{ba} . Therefore diverse states always exist
 259 (for large enough genotype spaces). However, diverse states are not always realised depending on whether
 260 such a configuration exists in the local genotype space.

261 For case 2(b) with a Fujiyama fitness landscape, $E_a - E_b = \Delta = 0.1$. Therefore invasion will occur for
 262 $k > \frac{\epsilon_I \Delta}{J_{ba}} \approx 0.2$ (as J_{ba} can be selected to be high). The converse equation requires $k > -\frac{\epsilon_I \Delta}{J_{ba}}$ which can
 263 always be satisfied; therefore diversification can always follow invasion. Only for both $k \gtrsim 1$ and negative
 264 J_{ba} will exclusion of the fittest type occur. Fig. 3 shows that replacement of the wildtype occurs at just
 265 below $k = 1$ and Fig. 2 shows that diverse states exist for $k > 0.5$.

266 Note that in the case 2, diverse states may be parasitic/predatory or mutualistic, depending on the differ-
 267 ence in fitness between types. Simple rearrangement of the above weight function yields $\frac{n_a}{n_b} = \frac{J_{ab} + \frac{\epsilon}{k}(E_a - E_b)}{J_{ba} - \frac{\epsilon}{k}(E_a - E_b)}$.
 268 This ratio must be positive for a diverse state to exist, and of order 1 for stability. Unless $\frac{\epsilon}{k}(E_a - E_j)$ is
 269 small, this implies J_{ba} and J_{ab} are of opposite signs. Thus, both mutualism and predator-prey interactions
 270 are possible.

271 This very simplistic analysis shows that typical species should not be able to coexist below the value
 272 of k_{min} . It is surprising that the system explores many exceptional species and even these cannot remain
 273 in *stable* coexistence below the the mean-field threshold value (approximately, at least). A diverse state at
 274 low k is not stable to invasion from mutations; i.e. neighbouring types in genotype space that satisfy the
 275 above condition that J_{ba} and J_{ab} are both large are not stable to fluctuations. Therefore, these states are
 276 not realised for long periods of time (when compared to q-ESS), and appear infrequently in time average
 277 measures. On a mean field level, a cutoff at some value is inevitable as coexistence is impossible for $k \leq \epsilon$
 278 (case 1) or $k \leq \epsilon \Delta$ (case 2(b)). The analysis required to show the true nature of the cross-over is too complex
 279 for inclusion here and will be studied in future work.

280 **4 Discussion**

281 In our model, we have found that there will be a cross-over from monodominance to a diverse state as
 282 the interaction increases. There appears to be a critical value of interaction strength beyond which the
 283 monodominance is broken down. For rugged fitness landscapes, it is vital that the fitness of individuals
 284 is (to some degree) *density dependent* for a cutoff in diversity to exist. The cutoff appears in the density
 285 dependent part of fitness for very rough fitness landscapes, and also appears in the density independent part
 286 for highly peaked fitness landscapes. The relative importance of each will depend on the specifics of the
 287 system studied. In future work, it will be important to understand the nature of the cross-over theoretically,
 288 beyond the naive estimate of the mean-field type calculation given in Sect. 3.2.

289 Using density dependent intrinsic fitness predicts that replacement of a ‘fitter’ type can occur at lower
290 interaction strength than coexistence. Using density independent intrinsic fitness predicts that coexistence
291 should occur at lower interaction strength than replacement of a ‘fitter’ type. This is certainly a feature of
292 working in a fixed genotype space; it will be important to establish whether this is still true in models with
293 correlated interaction matrices.

294 Each evolutionary course can be different in the simulation. If the initial type has neighbours in genotype
295 space that interact favourably with each other and negatively against the wildtype, then it will quickly go
296 extinct and (possibly several) q-ESS switches are observed. Other initial conditions allow the interactions of
297 local mutants to favour the wildtype, and monodominance continues for a longer time, possibly beyond the
298 timescale of the simulation. On a transition from one q-ESS state to another, our model predicts that at low
299 interaction strengths only monodominant states can occur. However, if the interaction strength $k \geq k_{min}$
300 then all initial wildtypes should be able to diversify eventually via an adaptive walk. If $k \gg k_{min}$ ($> \epsilon$), the
301 contribution to the weight function from the intrinsic fitness becomes negligible and the system reduces to
302 the original Tangled Nature model with the weight function H_0 , meaning all states are diverse.

303 A similar cutoff was observed for diversification of several types of *Escherichia coli* (*e. coli*) by Kashiwagi
304 et al.[4]. In this experiment the culture was well-mixed and fed with glutamate, the sole nutrition source (of
305 nitrogen). Through mutagenesis, evolution of a single gene was studied - the gene for glutamine synthetase
306 production, which synthesises glutamine from glutamate. Since the glutamine synthesis is necessary for the
307 growth of the bacteria in this experiment, those with the higher activity of glutamine synthetase will result
308 in faster growth of the bacteria. Indeed, in a low population density condition, only the fittest type (i.e.,
309 that with highest enzyme activity) survives. However, in a dense condition, multiple types including those
310 with much lower enzyme activity coexist. Interaction is (amongst other things) via leakage of glutamine,
311 and removal of glutamine from the environment confirms survival of the fittest.

312 We can identify our interaction strength k as a surrogate to bacterial density in the experimental setup,
313 as the strength of interaction felt between cells will increase when they are packed together more closely⁸.
314 The comparison is valid for approximately constant population size, which is approximately the case in both
315 experiment and model. Increased density will increase inter-specific interactions more than intra-specific
316 interactions, as the addition of a new substance to the cell will be more significant than the addition of the
317 same amount of an already present substance. Our predictions appear to be consistent with the experiment,
318 and with the subsequent observation that cutting off the interaction prevents the diversification [27].

319 Our theoretical results enable us to probe the underlying factor allowing diversification that is unobserv-
320 able in real systems. Essentially we require both:

- 321 1. The mean realised interactions for all types are equal to or greater than their own intrinsic fitness.
322 ‘Intrinsic fitness’ is an absolute measure if fitness is density dependent, but a relative measure if fitness
323 is density independent. The greater the ratio of the interaction strength to the intrinsic fitness strength,
324 the less selection pressure acts via the type’s efficiency in the environment. Thus all types gain more
325 from each other than they do from the environment.
- 326 2. The possible mutations from the wildtype reinforce themselves, or other types, more than the wildtype.
327 This is a constraint of the local genotype space, and means that the mutant does not interact more
328 positively with the wildtype than with itself.

329 Our model is quite general and so can be considered of relevance to many evolutionary systems. The
330 only fitness concept that fails to give a reasonable interaction vs. diversity graph is uncorrelated density
331 independent fitness; this would not be considered realistic. We conclude that there should be an interaction
332 cutoff below which no diverse states are found, above which diversity can arise in the absence of space with
333 a single resource.

⁸Note that this is an additional density dependence to the one explicitly included in the first term in Eq. 2, which represents the likelihood of an individual α meeting an individual of type \mathbf{S} .

334 Acknowledgements

335 We thank Tetsuya Yomo for discussion on the experiment of [4] and the Engineering and Physical Sciences Research
336 Council (EPSRC) for funding both Daniel Lawson's PhD, and Kunihiko Kaneko's visit to the U.K. We also thank
337 Andy Thomas and Gunnar Pruessner for help processing the model and for setting up the BSD cluster, speeding
338 computations enormously.

339 References

- 340 [1] G. F. Gause. Experimental analysis of vito volterra's mathematical theory of the struggle for existence. *Science*, 79:16–17, 1934.
- 341 [2] Robert B. Helling, Christopher N. Vargas, and Julian Adams. Evolution of *Escherichia coli* during growth in a constant environ-
342 ment. *Genetics*, 116:349–358, 1987.
- 343 [3] Tamaás L. Czárán, Rolf F. Hoekstra, and Ludo Pagie. Chemical warfare between microbes promotes biodiversity. *Proc. Nat.*
344 *Acad. Sci.*, 99:786–790, 2002.
- 345 [4] Akiko Kashiwagi, Wataru Noumachi, Masato Katsuno, Mohammad T. Alam, Itaru Urabe, and Tetsuya Yomo. Plasticity of fitness
346 and diversification process during an experimental molecular evolution. *J. Molec. Evol.*, 52:502–509, 2001.
- 347 [5] Kim Christensen, Simone A. di Collobiano, Matt Hall, and Henrik J. Jensen. Tangled nature: A model of evolutionary ecology.
348 *J. Theor. Biol.*, 216:73–84, 2002.
- 349 [6] Simone Avogadro di Collobiano, Kim Christensen, and Henrik Jeldtoft Jensen. The tangled nature model as an evolving quasi-
350 species model. *J. Phys A*, 36:883–891, 2003.
- 351 [7] Matt Hall, Kim Christensen, Simone A. di Collobiano, and Henrik Jeldtoft Jensen. Time-dependent extinction rate and species
352 abundance in a tangled-nature model of biological evolution. *Phys. Rev. E*, 66(011904), 2002.
- 353 [8] O. Dieckmann, M. Gyllenberg, and J.A.J. Metz. Steady state analysis of structured population models. *Theor. Pop. Biol.*,
354 63:309–338, 2003.
- 355 [9] Kei Tokita and Ayumu Yasutomi. Emergence of a complex and stable network in a model ecosystem with extinction and mutation.
356 *Theor. Pop. Biol.*, 63:131–146, 2003.
- 357 [10] Géza Meszéna, Mats Gyllenberg, Liz Pásztor, and Johan A.J. Metz. Competitive exclusion and limiting similarity: A unified
358 theory. *Theor. Pop. Biol.*, 69:68–87, 2006.
- 359 [11] Mats Gyllenberg. On the impossibility of coexistence of infinitely many strategies. *J. Math. Biol.*, 50:133–160, 2005.
- 360 [12] R. MacArthur and R. Levins. The limiting singularity, convergence, and divergence of coexisting species. *Am. Nat.*, 101:377–385,
361 1967.
- 362 [13] J. Vandermeer, M. A. Evans, P. Foster, T. Höök, M. Reiskind, and M. Wund. Increased competition may promote species
363 coexistence. *Proc. Nat. Acad. Sci.*, 99:8731–8736, 2002.
- 364 [14] Jef Huisman and Franz J. Weissing. Biodiversity of plankton by species oscillations and chaos. *Nature*, 402:407–410, 1999.
- 365 [15] Kunihiko Kaneko and Takashi Ikegami. Homeochaos: Dynamics stability of a symbiotic network with population dynamics and
366 evolving mutation rates. *Physica D*, 56:406–429, 1992.
- 367 [16] Robert M. May. *Stability and complexity in model ecosystems*. Princeton University Press, 3 Market Place, Woodstock, Oxford-
368 shire, UK. OX20 1SY, 1973.
- 369 [17] Manfred Eigen, John McCaskill, and Peter Schuster. Molecular quasi-species. *J. Phys. Chem.*, 92:6881–6891, 1988.
- 370 [18] Barbara Drossel. Biological evolution and statistical physics. *Advances in Physics*, 50(2):209–295, 2001.
- 371 [19] Michel Droz and Andrej Pękaliski. Population dynamics with or without evolution: a physicist's approach. *Physica A*, 336:84–92,
372 2004.
- 373 [20] Massimo Pigliucci and Carl D. Schlichting. On the limits of quantitative genetics for the study of phenotypic evolution. *Acto*
374 *Biotheo.*, 45:143–160, 1997.
- 375 [21] Simon Laird and Henrik Jeldtoft Jensen. The tangled nature model with inheritance and constraint: Evolutionary ecology
376 constricted by a conserved resource. Submitted to Ecological Complexity, www.arxiv.org/q-bio.PE/0510008, 2005.
- 377 [22] D.T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comp.*
378 *Phys.*, 22:403–434, 1976.
- 379 [23] David Bernstein. Simulating mesoscopic reaction-diffusion systems using the gillespie algorithm. *Phys. Rev. E*, 71(041103), 2005.
- 380 [24] Per Arne Rikvold. Self-optimization, community stability, and fluctuations in a class of individual-based models of coevolution.
381 *arXiv:q-bio*, (0508025), 2006.
- 382 [25] Debashish Chowdhury, Dietrich Stauffer, and Ambarish Kunwar1. Unification of small and large time scales for biological evolution:
383 Deviations from power law. *Phys. Rev. Lett.*, 90(068101), 2003.
- 384 [26] Daniel Lawson and Henrik Jeldtoft Jensen. The species-area relationship and evolution. *arXiv:q-bio.PE*, (0412024), 2005.
- 385 [27] Kashiwagi et al. Private communication. 2005.