

# Eco-evolutionary dynamics lead to functionally robust and redundant communities

Lorenzo Fant,<sup>1,\*</sup> Iuri Macocco,<sup>1,\*</sup> and Jacopo Grilli<sup>2,†</sup>

<sup>1</sup>*International School for Advanced Studies (SISSA), Via Bonomea 265, 34136 Trieste, Italy*

<sup>2</sup>*The Abdus Salam International Centre for Theoretical Physics (ICTP), Strada Costiera 11, 34014 Trieste, Italy*

Microbial communities are functionally stable and taxonomically variable. Species abundance greatly fluctuates over time and across communities. The functional composition, measured with the gene content of metagenomic data, is instead robust and conserved over time and across similar environments. These observations together suggest that species are functionally redundant: the same function is performed by many species, and one can assemble many communities with different species and the same functional composition. The clarity of this observation does not parallel with a theoretical understanding of its origin. Here we study a broad class of consumer-resource models with cross-feeding under an eco-evolutionary framework, in presence of a metabolic tradeoff. We show that the eco-evolutionary trajectories separate in a fast and a slow dynamics. The fast dynamics are determined by niche differences and make the system converge to a set of solutions uniquely determined by resource input, which we identify analytically. The slow dynamics occur entirely within that set of solutions and are determined by the interplay between fitness differences and niche differences. Interestingly, the set of solutions determined by the fast eco-evolutionary dynamics uniquely corresponds to the functional composition of the community, characterized by the fraction of individuals able to perform a given function, and not by the species content. We finally show then that a small variation of the fitness differences determines only a variation of the species composition, leaving the functional profile unvaried.

<sup>1</sup>

## Main

2 The staggering taxonomic diversity of microbial communities parallels with their remarkable functional robustness [1, 2]. At the species and strain level, their taxonomic composition is in fact highly variable across communities  
3 with similar environmental conditions and over time. This variability is also observed in microcosmos experiments,  
4 under very controlled conditions [3]. On the other hand, the functional composition of communities, estimated for  
5 instance using metagenomic data [2, 4], appears highly reproducible and stable over time. This — only apparent —  
6 contradiction strongly suggests that microbial taxa are highly functional redundant: since many species can perform  
7 the same functions, there exist multiple species combinations corresponding to the same functional profile.  
8

9 While this observation is robust and replicated across ecosystems, including — at least some — laboratory ex-  
10 periments, its origin is unknown. This lack of understanding is, in part, due to the fact that the vast majority of  
11 the modelling frameworks focus on species composition. Population abundances are in fact the standard degrees of  
12 freedom of mathematical models of community dynamics.

13 Consumer-resource models are the main modelling framework for microbial communities. Their origin goes back  
14 to the classic work of MacArthur and Levins [5], which has been extensively studied and discussed in the following  
15 decades [6, 7], mostly to describe the coexistence of a handful of species.

---

\*These two authors contributed equally

†Electronic address: [jgrilli@ictp.it](mailto:jgrilli@ictp.it)

16 Recently, they have been furtherly extendend to consider facilitation through cross-feeding [3, 8, 9], where species  
17 change resource availability not only by consumption, but also because they release in the environment the waste  
18 products of their metabolism. Such improvements effectively describe experimental results [3] and have the flexibility  
19 to reproduce patterns observed in empirical microbial communities [10].

20 Once the parameters of the model are set and an initial pool of species is chosen, populations converge for large  
21 times to an equilibrium point. Under some mild conditions, identified over decades of theoretical work [11, 12],  
22 consumer-resource models are characterized by a globally stable equilibrium: the steady state is independent of the  
23 initial population abundances and resource concentrations. The competitive exclusion principle — one of the most  
24 fundamental results of theoretical ecology — limits the number of species that can possibly coexists in a stable  
25 equilibrium: diversity cannot exceed the number of resources. While this bound is hard, it is often not realized, as  
26 only fewer species are able to coexist.

27 The number and identity of the species coexisting at equilibrium is in fact determined not only by the ecological  
28 dynamics, but also by the initial pool of species. This pool of species is often interpreted as the metacommunity  
29 diversity: the ecological dynamics unfolds in a local community linked to the metacommunity by a small migration  
30 rate. Most of the recent progresses in understanding the assembly of large ecological communities have been driven  
31 by the assumption of “random” species pools [13–15]. This common choice implicitly assumes a separation of spatial  
32 and temporal scales: the ecological dynamics determining the community composition in the local community occurs  
33 independently of the evolutionary processes determining the pool of diversity of the metacommunity.

34 Instead of assuming an externally fixed species pool, one can let it evolve dynamically. Classic work in adaptive  
35 dynamics [16] has shown how, starting from a clonal population, diversification can evolve under general conditions  
36 on frequency dependent selection. Several works have then studied eco-evolutionary dynamics of interacting popula-  
37 tions [17, 18], by allowing their parameters to be subject to mutations and be inherited by the following generation.  
38 Dynamic adaptation of metabolic preferences allows many species to coexist and respond optimally to environmental  
39 change [19]. A critical aspect of these models is the variation of parameters correspondig to fitness and niche differ-  
40 ences, both contributing to the community evolution [20]. A key question is how the interplay between niche and  
41 fitness difference determine the species and functional composition of communities.

42 Here we consider the broad framework of consumer-resource-crossfeeding models under an explicit eco-evolutionary  
43 dynamics, where strains differ in their resource preference and in their intrinsic fitness. Higher resource intakes  
44 are balanced by a lower efficiency (or equivalently, higher mortality) implemented by a metabolic trade-off [21, 22].  
45 We show that, when fitness differences are small, the evolutionary dynamics converges rapidly to a stationary and  
46 reproducible functional composition — here defined as the fraction of individuals able to grow on a given resource  
47 — which we analytically predict. Interestingly, we show that, once the functional attractor is reached, the strain  
48 dynamics is then dominated by fitness differences, implying that functional composition is robust (independent of  
49 small fitness differences) and redundant (is obtained under multiple strain compositions).

50

## Eco-evolutionary dynamics

51 In our framework, individuals are characterized by a resource preference vector which determines the intake rate of  
52 each of the  $R$  resources available in the environment (relative to a maximum). An individual with preference  $a_i = 0$

53 will not consume resource  $i$ , while an individual with preference  $a_i = 1$  will consume it with the maximum intake  
54 rate  $\nu_i$ . Consumed resources are converted into biomass with finite efficiency (equivalent to an inverse yield). We  
55 assume that the yield (or equivalently the death rate, see Materials and Methods) depends linearly on the number  
56 of resources consumed: the more resources an individual can grow on, the less efficiently it grows. Individuals also  
57 differ in their intrinsic fitness, modelled as an additional — strain dependent — term in the efficiency (or equivalently  
58 mortality).

59 Resource dynamics is described explicitly. Resources are introduced in the system with a resource-specific rate  $h_i$   
60 and consumed by individuals. Their concentration also vary because of cross-feeding. A fraction  $1 - \ell$  of resources  
61 consumed by an individual is used for growth, while a fraction  $\ell$  is transformed to different resources and released  
62 again in the environment [8]. The cross-feeding matrix, with elements  $D_{ij}$ , specifies the relative rates of resource  
63 transformation (see Materials and Methods for its parameterization).

64 In our model, resource preferences are subject to mutations and evolution. After reproduction, the new individual  
65 can differ from the parent in its resource consumption. We consider different implementation of mutations (see  
66 Materials and Methods) which, however, do not affect the results. We restrict the analysis to the most conservative  
67 case of rare mutation, where offspring differs in the resource preference vector for at most one resource.

68 Mutated individuals also differ from parents in their intrinsic fitness, which is drawn at random independently of the  
69 parent fitness (see Materials and Methods). We restrict our analysis to the case of “small” fitness differences, where  
70 their magnitude is smaller than the typical effect of mutations in the resource preference in absence of competition  
71 (see Materials and Methods).

The per-capita growth rate of a strain  $\mu$  — where with “strain” we identify the group of individuals with equal  
resource preference  $a_i$  — will therefore be a function  $g_\mu(\underline{c})$  of the resource concentration  $\underline{c}$ , which, in turn, depends  
dynamically on the population abundance because of consumption and cross-feeding (see Materials and Methods).  
We consider the following choice

$$g_\mu(\underline{c}) = \eta_\mu \left( \sum_{i=1}^R a_{\mu i} r_i(c_i) - \delta \left( 1 + \chi \sum_{i=1}^R a_{\mu i} \right) (1 + \epsilon_\mu) \right). \quad (1)$$

72 The values of  $\eta_\mu$  and  $\delta$  are arbitrary and their choice does not affect the results. The functional form  $r_i(c_i)$  encodes  
73 the functional response. Both linear and saturating (Monod-like) functional responses produce the same results. The  
74 parameter  $\chi$  quantifies the fitness cost of consuming one resource and implements the metabolic trade-off. The form  
75 of the trade-off generalizes the case considered in [21, 22], which corresponds to the limit  $\chi \rightarrow \infty$ . The intrinsic fitness  
76 differences are determined by the value of  $\epsilon_\mu$ .

## 77 Convergence to the functional attractor

78 Fig 1 shows the resulting eco-evolutionary trajectories. Starting from a clonal population, a diverse community is  
79 rapidly assembled. Strain abundances changes abruptly following successful invasion events and keep changing over  
80 the whole duration of the simulations (over  $10^5$  invasion attempts).

81 The final community structure is remarkably simple if, instead of concentrating on its species abundances, we  
82 focus on its functional composition (see Materials and Methods). We define functional occurrence  $F_i$  as the fraction  
83 of individuals able to grow on  $i$  (i.e., with  $a_i = 1$ ). After a short transient (of about  $10^2$  invasion attempts) the

84 functional occurrences and with the total biomass  $N$  relax to the respective stationary values  $F_i^*$  and  $N^*$  which are  
 85 finely reproducible across different realizations.

86 Two phases characterize therefore the eco-evolutionary dynamics. The first phase is an initial-condition dependent  
 87 transient, where the community structure is mainly shaped by rapid invasions. In the second phase, conversely, the  
 88 community has converged to a stable functional composition, which we will refer to as “functional attractor”.

89 The eco-evolutionary dynamics of the second phase unfolds while keeping the total biomass constant, affecting  
 90 therefore the relative abundance of strains. The sequence of invasions and extinctions of strains is determined by  
 91 the interplay of fitness differences  $\epsilon_\mu$  and niche differences, related to the dissimilarity of the resource preferences.  
 92 Importantly, the trajectories of species abundances are effectively restrained to occur on the low(er)-dimensional  
 93 manifold determined by the constraint enforced through the functional occurrences  $F_i^*$ . In this second phase the  
 94 community has thus reached a “functional maturity” and the subsequent evolution only affects strain composition  
 95 while leaving unaltered the functional one.

## 96 Ecological dynamics under infinite species pool predicts the functional attractor

97 The stability and reproducibility of the functional attractor suggest that it is possible to predict its properties. We  
 98 considered a simplified setting which aims at mimicking the effective exploration of the phenotypic space performed  
 99 by mutations. In particular we looked at the same ecological dynamics defined above but with an initial infinitely  
 100 large species pool, which encompasses all the possible strains (i.e., the  $2^R$  possible resource preferences). A similar  
 101 approach has been considered to study a simpler version of the model [21] (corresponding to the limit  $\chi \rightarrow \infty$  and no  
 102 crossfeeding). We further postulate that the small intrinsic fitness differences do not affect substantially the functional  
 103 attractor.

The consumer-resource-crossfeeding model with infinite species pool and no intrinsic fitness differences can be  
 analytically solved by further assuming that the timescales of species and resource dynamics are separated. In the  
 Materials and Methods we show that the stationary functional occurrences  $F_i^*$  and  $N^*$  are given by

$$F_i^* = \min\left\{\frac{h_i^{eff}}{\chi} \frac{1}{N^*}, 1\right\} \quad (2)$$

and

$$N^* = \frac{\sum_i h_i^{eff}}{R(1 + \chi \sum_i F_i^*)}. \quad (3)$$

104 The parameter  $h_i^{eff}$  is the effective resource inflow in the system, which is given by the combination of resources that  
 105 are externally supplied and the ones produced via crossfeeding. This quantity is in a simple linear relation to the  
 106 inflow rate of externally provided resource  $h_i$  trough the crossfeeding matrix  $D$  (see Materials and Methods).

107 Figure 1 shows that the predictions of eq. 2 and eq. 3 accurately describe the outcomes of the eco-evolutionary  
 108 dynamics. Resources can be partitioned in two groups based on their effective influx rate  $h_i^{eff}$ . If the influx rate  
 109 is larger than a critical value  $h_c$ , then the ability to metabolize that resource is a “core” function, shared by each  
 110 individual in the community (i.e.,  $F_i^* = 1$ ). The value of  $h_c$  depends on both the spread of resource quality (the  
 111 variability among the  $h_i$ ) and the metabolic cost  $\chi$ . The higher the metabolic cost and the variability, the higher the  
 112 critical quality threshold  $h_c$  and, consequently, the fewer the core resources.

113 Not all the individuals consume resources with quality lower than the critical threshold. A linear relation links the  
114 functional occurrence  $F_i^*$  with the resource quality  $h_i^{eff}$ . The slope of this relation is simply linked to the metabolic  
115 cost and the total biomass, being equal to  $(\chi N^*)^{-1}$  (see Materials and Methods). Figure 1B shows that the analytical  
116 expression for  $N^*$  (as a function of the metabolic cost  $\chi$ ) is correctly matched by numerical simulations.

117 **Functional stability and robustness**

118 These results connecting resource quality with functional profile and total biomass can be generalized to a wider  
119 range of assumptions. For instance, one can go beyond dichotomous resources preferences, and consider species able  
120 to consume resources at different rates. As long as the species pool is large and variable enough, the results presented  
121 previously still hold. One can also include intrinsic differences between resources. For instance, some resources  
122 could have higher energy content or the maximal intake rate could be resource-dependent. These differences between  
123 resources can be easily reabsorbed in the quality parameters  $h_i^{eff}$  (see Materials and Methods).

124 An emerging feature of the present framework is that the functional composition of communities is extremely  
125 robust to fitness differences. We further explored this robustness by considering fitness variation across space and  
126 time. Figure 3 shows how species composition and functional profile change across communities or over time as species  
127 fitness differences change. Species composition varies widely: the small fluctuations in relative fitness determine which  
128 species go extinct and the relative abundance of those who survive. The functional composition is instead extremely  
129 robust to these fitness differences, closely reproducing the phenomenology observed in microbial communities [2].

130 **Conclusion**

131 Our results shed light on the composition of large ecological communities. When the species pool is not a-priori  
132 constrained, but is allowed to evolve, the complex ecological dynamics can be decomposed in a fast, predictable,  
133 component and a slow one, contingent to the (small, yet relevant) fitness differences. The community composition  
134 rapidly converges to a manifold of solution, determined by resource availability (see Figure 3). The dynamics on that  
135 manifold is governed by difference in relative fitness. Remarkably, this separation of fast and slow components maps  
136 directly into functional and taxonomical composition: the former is robust and governed only by resource quality, the  
137 latter is constrained by function, but free to move along functionally equivalent directions.

138 The functional reproducibility of communities is also affected by fitness differences. We have shown that a typical  
139 fitness difference of 1% does not alter substantially the functional composition. On the other hand, much smaller  
140 fitness differences substantially skew the taxonomic composition. The fitness differences, responsible for taxonomic  
141 variation over space and time, could be due by multiple factors (from abiotic factors to phages). Our framework  
142 provide a baseline to assess the role of this variation, and could be extended to include explicitly those factors  
143 contributing to fitness differences.

144 Here we assumed that intrinsic fitness was not heritable, and it was drawn from a constant distribution for every new  
145 mutant. The alternative choice [20], with heritable fitness differences, would likely produce the same results, provided  
146 that intrinsic fitness differences were not evolving rapidly. In this scenario, we expect again that the eco-evolutionary  
147 dynamics is divided in two phases. In the first, the community to converge to the functional attractor. In the second,

148 the strains keep evolving and the intrinsic fitnesses keep increasing, with the constraint imposed by the functional  
149 composition.

150 A remarkable aspect of our framework is that functional composition — as opposed to taxonomic composition  
151 — naturally emerges as the relevant, reproducible observable to characterize an ecological community evolved in a  
152 constant environment. This property is likely to hold more generally and not to be restricted to consumer-resource  
153 systems or microbial communities. We expect that a similar approach could be developed to study mutualistic  
154 communities or pathogen dynamics.

155 **Materials and Methods**

156 **Definition of the model**

We consider a consumer-resource model in presence of crossfeeding [8], which describes the dynamics of population biomasses  $n_\sigma$  (for  $\sigma \in \mathcal{S}$ ) and resources concentration  $c_i$  (for  $i \in R$ ). Changes in population abundance are defined by

$$\frac{dn_\sigma}{dt} = n_\sigma \left( \eta_\sigma \sum_{i \in R} \mathcal{E}_{i\sigma}^g - \delta_\sigma \right) . \quad (4)$$

where  $\delta_\sigma$  is a death term and  $\eta_\sigma$  is the efficiency of the conversion of energy into biomass.  $\mathcal{E}_{i\sigma}^g$  is the energy flux used for species  $\sigma$  to grow from metabolite  $i$ . The total energy flux into a cell of type  $\sigma$  is given by  $\mathcal{E}_{i\sigma}^{in} = \mathcal{E}_{i\sigma}^g + \mathcal{E}_{i\sigma}^{out}$ , where  $\mathcal{E}_{i\sigma}^{out}$  are the energy fluxed of secreted metabolites. The associated dynamics of resource concentration  $c_i$  is defined by

$$\frac{dc_i}{dt} = h_i(c_i) - \frac{1}{w_i} \sum_{\sigma \in \mathcal{S}} n_\sigma \mathcal{E}_{i\sigma}^{in} + \frac{1}{w_i} \sum_{\sigma \in \mathcal{S}} n_\sigma \mathcal{E}_{i\sigma}^{out} , \quad (5)$$

157 were  $w_i$  defines the conversion between energy and concentration of resource  $i$ . The function  $h_i(c_i)$  specify the  
158 dynamics of resource concentration in absence of consumers.

159 We assume that the energy fluxes used for growth are a constant fraction  $1 - \ell$  of the total ones:  $\mathcal{E}_{i\sigma}^g = (1 - \ell) \mathcal{E}_{i\sigma}^{in}$ .  
160 The energy fluxed from secreted metabolites is given by  $\mathcal{E}_{i\sigma}^{out} = \ell \sum_{j \in R} D_{ij} \mathcal{E}_{j\sigma}^{in}$ . The crossfeeding matrix element  $D_{ij}$   
161 defines energy conversion between resource  $j$  and resource  $i$ . Energy conservation implies  $\sum_i D_{ij} = 1$ .

The energy flux  $\mathcal{E}_{i\sigma}^{in}$  is takes the form

$$\mathcal{E}_{i\sigma}^{in} = w_i \nu_i a_{\sigma i} r_i(c_i) , \quad (6)$$

162 where  $r_i(c_i)$  is a non-decreasing function of the concentration of resource  $i$ , and  $\nu_i$  is the maximal intake rate of  
163 resource  $i$ . The elements  $a_{\sigma i} \in [0, 1]$  measures the intake rate of metabolite  $i$  species  $\sigma$  relative to the maximum  $\nu_i$ .  
164 Here we focus on the case of externally supplied resources  $h_i(c_i) = h_i$ .

For the death rates and yield we assume the expression

$$\frac{\delta_\sigma}{\eta_\sigma} = \frac{1}{\tau} \left( 1 + \chi \sum_{j \in R} a_{\sigma j} \right) . \quad (7)$$

165 Without loss of generality we can set the timescale  $\tau = 1$ . In the simple setting of  $a_{\sigma j} \in \{0, 1\}$ , the parameter  $\chi$   
166 measure the cost of being able to metabolize aeh metabolite ( $\chi R$  is the fitness cost paid to be a generalist).

167

## Metabolic tradeoffs and functional attractor

If the dynamics of resources is much faster than population dynamics, one can effectively describe the dynamics of populations as

$$\frac{dn_\sigma}{dt} = n_\sigma \left( \eta_\sigma \sum_{i \in R} a_{\sigma i} \nu_i \frac{h_i^{eff}}{\sum_{\mu \in S} n_\mu a_{\mu i} \nu_i} - \delta_\sigma \right), \quad (8)$$

where  $h_i^{eff} = (1 - \ell) \sum_{j \in R} B_{ij} h_j w_j$  and the matrix  $B = (I - \ell D)^{-1}$ . It is known [21] that

$$L(\{n\}) = \sum_\sigma \frac{\delta_\sigma}{\eta_\sigma} n_\sigma - \sum_i h_i^{eff} \log \left( \sum_\sigma \nu_i a_{\sigma i} n_\sigma \right), \quad (9)$$

is a Lyapunov function. With our choice for the metabolic tradeoff (7), it can also be written as

$$L(\{n\}) = \sum_\sigma n_\alpha \left( 1 + \chi \sum_{j \in R} a_{\sigma j} \right) - \sum_i h_i \log \left( \sum_\beta \nu_i a_{\beta i} n_\beta \right) \quad (10)$$

We then introduce the total population size  $N = \sum_{\sigma \in S} n_\sigma$  and define the functional abundances  $F_i$  as

$$F_i = \sum_{\sigma \in S} a_{\sigma i} \frac{n_\sigma}{N}, \quad (11)$$

which corresponds to the fraction of individuals that are able to metabolize metabolite  $i$ . Interestingly, and surprisingly, the Lyapunov function can then be rewritten as function of  $N$  and  $\{F\}$  alone:

$$L(N, \{F\}) = N \left( 1 + \chi \sum_{j \in R} F_j \right) - \sum_{j \in R} h_j \log (N \nu_j F_j). \quad (12)$$

By minimizing the function over  $F_i$  in  $[0, 1]$  we obtain

$$F_i^* = \min\left\{1, \frac{1}{N^*} \frac{h_i}{\chi}\right\}, \quad (13)$$

where the total biomass is the solution of

$$N^* = \frac{\sum_{j \in R} h_j}{\left( 1 + \chi \sum_{j \in R} F_j^* \right)}. \quad (14)$$

168 We can now notice that the equilibrium solutions is primarily identified by its belonging to a manifold described by  
 169 equations (13,14). For a given system, ordering the resources based on their  $h_i$  we can identify a set “core” resources.  
 170 These are the most abundant ones that must to be metabolised in order to survive ( $F_i = 1$ ).

171 This simpler prescription with time-scale separation between populations and resources allows us to analytically  
 172 derive the properties of the stability manifold. Nonetheless, even if we were not able to find a close general expression  
 173 for the Lyapunov in presence of resources dynamics and cross feeding, the computational results behave according to  
 174 the same modelling developed in the simpler case.

175

### A. Eco-Evolutionary dynamics

176 We assume a separation of timescales between the evolutionary and the ecological dynamics: once a mutant appears  
 177 in the system, the ecological dynamics relaxes to its equilibirum before a new mutant appears.

178 We further assume that mutations are small. A mutant differ from the parent for the resource preference for only  
 179 one resource. In the following, with a small abuse of notation, we will use the term “genome” to indicate the array of  
 180 resource preference  $a_{\sigma i}$  for strain  $\sigma$ , “locus”  $i$  to indicate a position in the genome (corresponding to the consumption  
 181 of resource  $i$ ), and “allele” to indicate whether a resource  $i$  is consumed or not ( $a_{\sigma i} = 1$  or 0).

The mutation probability of each locus depends on the allele in that locus. The rate at which an allele 1 mutate to become 0 (deletion) is constant, independent of the locus, and equal to 1. The rate at which an allele 0 mutate to become 1 (addition) equals to  $P_h F_i + P_{dn}$ . The quantity  $P_h$  is the probability that an addition happens because of horizontal transfer, while  $P_{dn}$  the rate of “de novo” mutations. The rate of horizontal transfer is proportional to the frequency of the allele in the population  $F_i$ . Note that these choices make the mutation rate genome dependent: the probability that genome  $\mu$  mutates equals

$$p_\mu = x_\mu \left( \sum_{i|a_{\mu i}=0} U_{+,i} + \sum_{i|a_{\mu i}\neq 0} U_{-,i} \right) \quad (15)$$

182 where  $x_\mu$  is the frequency of genome  $\mu$  in the population.

Since the mutant  $\tilde{\mu}$  appears at a small frequency in the population, its survival probability of the new species is given by

$$p_{\tilde{\mu}}^{surv} = 1 - \min \left( 1, \frac{(1 + \chi \sum_i a_{\tilde{\mu} i})(1 + \epsilon \delta)}{(1 - l) \sum_i a_{\tilde{\mu} i} r_i(c_i)} \right) \quad (16)$$

Calculating all these quatities for all possible mutations of all existing genomes one can create a set of rates at which species  $\mu$  mutates gene  $i$

$$\underline{P}^{inv} = \{p_\mu p_{i|\mu} P_{\tilde{\mu}}^{surv}\}_{i,\mu} \quad (17)$$

183 We eventually select a random mutation using  $\underline{P}^{inv}$  as probabilities and considering  $T = \frac{1}{\sum_{i,\mu} \underline{P}^{inv}}$  as the mean time  
 184 necessary for the evolution to take place.

## 185 Numerical Simulations

186 Numerical simulations were performed on an average desktop, using a custom Python package (available upon  
 187 request). The integration of the coupled differential equations (4) and (5) was achieved with the standard solve\_ivp  
 188 function from the SciPy.integrate package. In particular we set a flag to stop the integration once  $|\max_\sigma g_\sigma(\underline{c})| < 10^{-4}$ ,  
 189 i.e. the populations are no more changing significantly. In the case of assembly, at each iteration a new patch is created  
 190 from the previous one by dropping the taxa with relative abundance lower than a threshold ( $10^{-5}$ ) and introducing a  
 191 new strain with abundance equal to the threshold value itself. the system undergoes a fast initial transient and reaches  
 192 the functional manifold within the very few first integration time steps; it then slowly evolves to the final equilibrium  
 193 state. The latter is ensured to satisfy the stability criterium defined above. On the other hand, in presence of a time  
 194 dependent noise all phenotypes were kept “available”, since even a small change in the noise can make a phenotype  
 195 enter the pool of relevant species.

196

## Choice of parameters

197 The results presented in this paper were achieved using generic parameters, whose details can affect the distribution  
198 of taxa or relaxation time but not the macroscopic observables that characterise the stability manifold. Referring  
199 to eq. 4 and 5, for simplicity we selected  $\eta_\sigma = \nu_i = w_i = 1$ . The time scale  $\tau$  was also set to 1. For the fitness  
200 difference, we used a gaussian noise of amplitude  $\epsilon = 0.005$ . A larger values of noise (of the order 0.1) often disrupts  
201 the properties of the manifold. For instance species not consuming a core resources is still able to survive because of  
202 a low death rate. On the other hand, for  $\epsilon < 0.001$  the simulation became very long as the small fitness difference  
203 means a very slow decay of the less fit species. The cross-feeding matrix  $D$  has been chosen following [10]. Entries  
204 were extracted according to a Dirichlet distribution, where resources have been dividend in three classes. We chose  
205 an effective sparsity of  $s = 0.1$ . The fraction of resources remaining in the same class was  $f_s = 0.7$  while the ones  
206 going to the waste class is  $f_w = 0.28$ . The results of the simulations are independent of the choices of  $P_h$  and  $P_{dn}$ ,  
207 which we considered equal to  $P_h = 0.9$  and  $P_{dn} = 0.1$

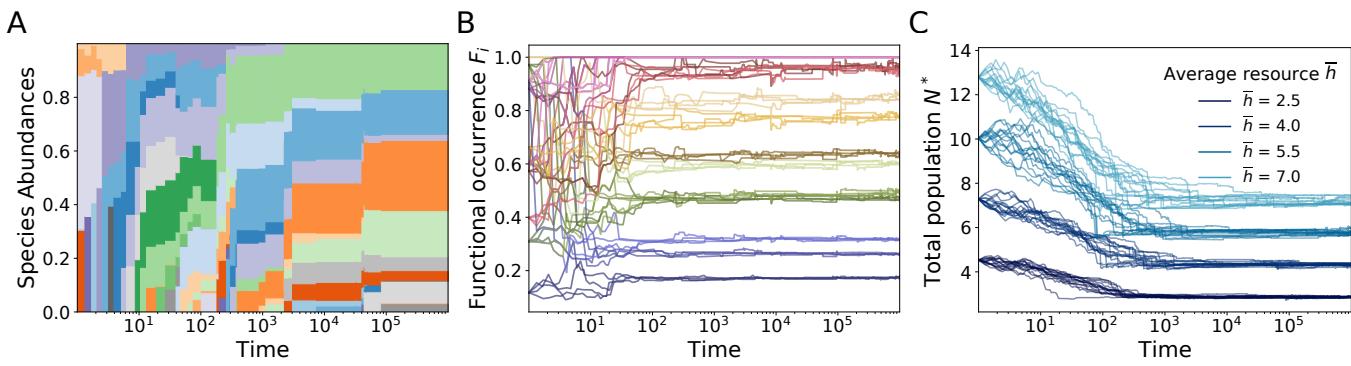


FIG. 1: Communities evolving under a consumer-resource model. the system is initialized with a low number of initial random species. The former are chosen so that each gene is present at least once. The system evolves in a chemostat with a fixed resources input. When equilibrium is reached, one mutant is added to the batch. The chemostat then equilibrates to a new fixed point and the procedure is repeated until function and biomass reach stability. **A:** Time evolution of the species relative abundances for one realisation of the system. **B:** Time evolution of functional occurrences for three different realisations of the system. 15 resource are given. **C:** Time evolution of the total biomass of the system. 20 realisations of the system are shown for each value of average resources income.

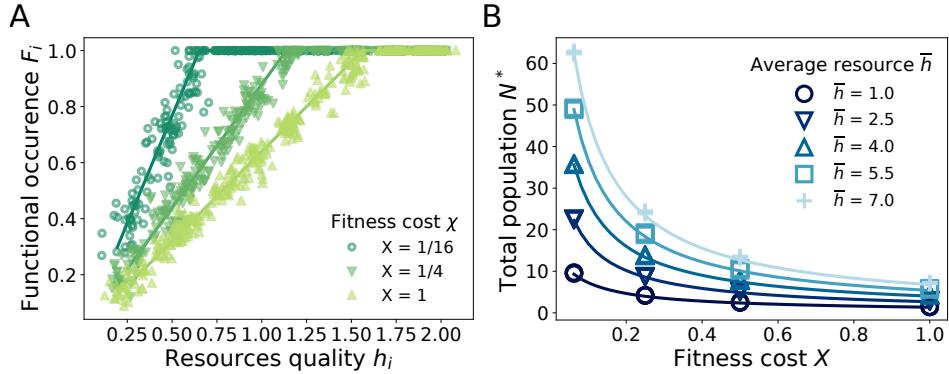


FIG. 2: The theoretical predictions given by eq. 13, 14 (solid lines) are reproduced by numerical integration of eq. 4, 5 (markers). **A:** Occurrence of the phenotypes  $F_i$  as a function of the resource income rates  $h_i$ . According to equation 13 the most abundant resources (core) are consumed by all species ( $F_i=1$ ) while the remaining ones only by a fraction  $F_i = \frac{h_i}{N^* \chi}$ . Notice that increasing  $\chi$  reflects in a decrease of the number of core resources. **C:** Dependence of the total equilibrium population  $N^*$  (biomass) on the value of  $\chi$ . Unlike the the quantities in **A** here we here find a dependence on average value of the resources incomes  $\bar{h}$ . In each figure are represented 20 different noise realizations solutions of the system for each  $\chi$  (**A** and each **B**)

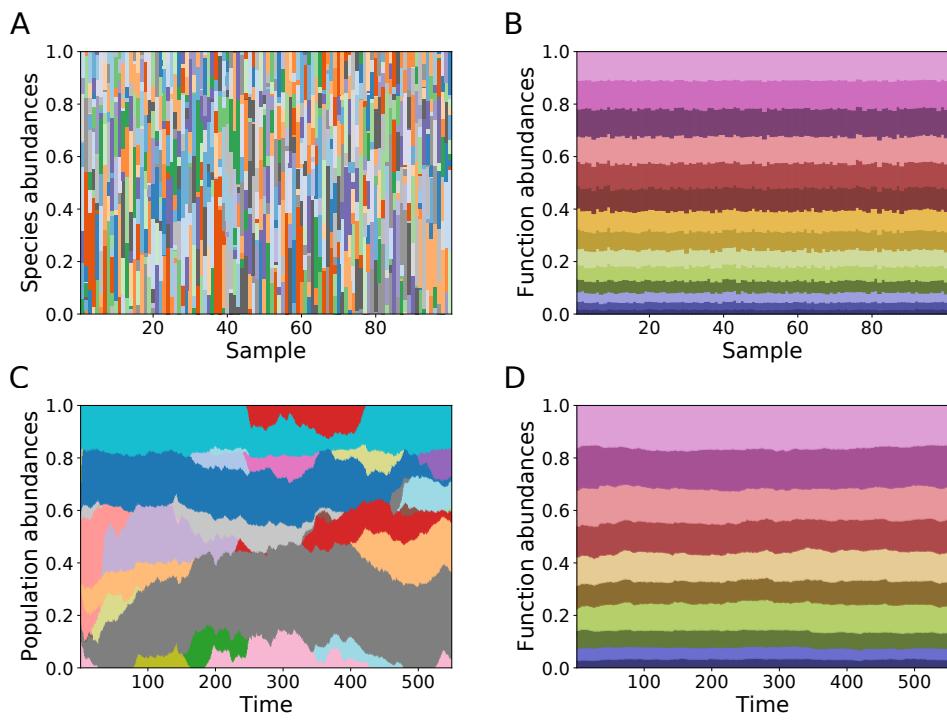


FIG. 3: Fitness differences allow to demonstrate how functional stability is already encoded within the model. While the populations/taxonomies becomes highly variable and heterogeneous, the functional composition is preserved and unaffected by stochasticity. **A,B** show a collection of equilibrium configurations of systems with different realisations of a static noise. **C,D** show the evolution of the composition of a system with a dynamically varying noise .

- 
- 208 [1] Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S. & Thomas, T. Bacterial community assembly based on functional genes  
209 rather than species. *Proceedings of the National Academy of Sciences* **108**, 14288–14293 (2011).
- 210 [2] Louca, S. *et al.* Function and functional redundancy in microbial systems (2018).
- 211 [3] Goldford, J. E. *et al.* Emergent simplicity in microbial community assembly. *Science* **361**, 469–474 (2018).
- 212 [4] Louca, S., Parfrey, L. W. & Doebeli, M. Decoupling function and taxonomy in the global ocean microbiome. *Science* **353**,  
213 1272–1277 (2016).
- 214 [5] MacArthur, R. & Levins, R. The Limiting Similarity, Convergence, and Divergence of Coexisting Species. *The American  
215 Naturalist* **101**, 377–385 (1967).
- 216 [6] Tilman, D. A consumer-resource approach to community structure. *American Zoologist* **26**, 5–22 (1986).
- 217 [7] Chesson, P. MacArthur's consumer-resource model. *Theoretical Population Biology* **37**, 26–38 (1990).
- 218 [8] Marsland, R. *et al.* Available energy fluxes drive a transition in the diversity, stability, and functional structure of microbial  
219 communities. *PLoS Computational Biology* **15** (2019). 1805.12516.
- 220 [9] Butler, S. & ODwyer, J. P. Stability criteria for complex microbial communities. *Nature communications* **9**, 1–10 (2018).
- 221 [10] Marsland, R., Cui, W., Goldford, J. & Mehta, P. The community simulator: A python package for microbial ecology. *Plos  
222 one* **15**, e0230430 (2020).
- 223 [11] Case, T. J. & Casten, R. G. Global stability and multiple domains of attraction in ecological systems. *The American  
224 Naturalist* **113**, 705–714 (1979).
- 225 [12] Harrison, G. W. Global stability of predator-prey interactions. *Journal of Mathematical Biology* **8**, 159–171 (1979).
- 226 [13] Bunin, G. Ecological communities with lotka-volterra dynamics. *Physical Review E* **95**, 042414 (2017).
- 227 [14] Advani, M., Bunin, G. & Mehta, P. Statistical physics of community ecology: a cavity solution to macarthurs consumer  
228 resource model. *Journal of Statistical Mechanics: Theory and Experiment* **2018**, 033406 (2018).
- 229 [15] Cui, W., Marsland III, R. & Mehta, P. Effect of resource dynamics on species packing in diverse ecosystems. *Physical  
230 Review Letters* **125**, 048101 (2020).
- 231 [16] Geritz, S. A., Mesze, G., Metz, J. A. *et al.* Evolutionarily singular strategies and the adaptive growth and branching of  
232 the evolutionary tree. *Evolutionary ecology* **12**, 35–57 (1998).
- 233 [17] Ackland, G. & Gallagher, I. Stabilization of large generalized lotka-volterra foodwebs by evolutionary feedback. *Physical  
234 review letters* **93**, 158701 (2004).
- 235 [18] DeLong, J. P. & Gibert, J. P. Gillespie eco-evolutionary models (gem s) reveal the role of heritable trait variation in  
236 eco-evolutionary dynamics. *Ecology and evolution* **6**, 935–945 (2016).
- 237 [19] Pacciani-Mori, L., Giometto, A., Suweis, S. & Maritan, A. Dynamic metabolic adaptation can promote species coexistence  
238 in competitive communities. *PLoS computational biology* **16**, e1007896 (2020).
- 239 [20] Good, B. H., Martis, S. & Hallatschek, O. Adaptation limits ecological diversification and promotes ecological tinkering  
240 during the competition for substitutable resources. *Proceedings of the National Academy of Sciences* **115**, E10407–E10416  
241 (2018).
- 242 [21] Tikhonov, M. Community-level cohesion without cooperation. *Elife* **5**, e15747 (2016).
- 243 [22] Posfai, A., Taillefumier, T. & Wingreen, N. S. Metabolic Trade-Offs Promote Diversity in a Model Ecosystem. *Physical  
244 Review Letters* **118**, 028103 (2017).