Effects of phylogeny on coexistence in model communities

Carlos A. Serván[1*], José A. Capitán[2], Zachary R. Miller[1] & Stefano Allesina[1,3]

[1]Department of Ecology & Evolution, University of Chicago, Chicago, IL, USA
[2]Complex Systems Group. Department of Applied Mathematics, Universidad Politécnica de Madrid, Av. Juan de Herrera, 6, 28040 Madrid, Spain
[3]Northwestern Institute on Complex Systems, Evanston, IL, USA

*corresponding author e-mail: cmarceloservan@uchicago.edu

Abstract

A species' traits influence the way in which it interacts with the environment. Thus, we expect traits to play a role in determining whether a given set of species coexists. Traits are, in turn, the outcome of an eco-evolutionary process summarized by a phylogenetic tree. Therefore, the phylogenetic tree associated with a set of species should encode information about the assembly properties of the community. Many studies have highlighted the potentially complex ways in which phylogenetic information is translated into species' ecological properties. However, much less emphasis has been placed on developing expectations for community properties under a particular hypothesis.

In this work, we couple a simple model of trait evolution on a phylogenetic tree with local community dynamics governed by Lotka-Volterra equations. This allows us to derive properties of the community of coexisting species as a function of the number of traits, tree topology and the size of the species pool. Our results highlight how phylogenies and traits, in concert, affect the coexistence of a set of species.

In this way, our work provides new baseline expectations for the ways in which phylogenetic information is reflected in the structure of and coexistence within local communities.

Introduction

Gause's pioneering work [15] provided the first clear empirical evidence for the principle of competitive exclusion, which states that two species competing for a unique resource cannot coexist. In the context of niche theory, this principle resonates in the concept of limiting

similarity: In a community shaped only by biotic interactions, species with similar niches are less likely to coexist [26]. Making stronger assumptions, one can draw a direct link between evolutionary relatedness among the members of an ecological community and their co-occurrence patterns. In particular, if one is willing to assume that species' traits are well-described by phylogeny, and that similarity in traits maps into strength of competition between species, one can connect the phylogenetic structure of an ecological community with coexistence [37]. While this hypothesis has found mixed support [11], it has served as one of the cornerstones of the budding field of community phylogenetics [35, 38]. In recent years, several tools have been developed to test whether a given mechanism of community assembly (e.g., competitive exclusion or environmental filtering) has acted on a community, by analyzing the signal it is expected to leave in the community's phylogenetic structure [14]. However, some authors have noted that phylogenetic relatedness might affect community patterns in a variety of ways, obscuring a link between phylogenetic and co-occurrence patterns [11, 28].

Here we take a step back and analyze a model in which we incorporate an explicit link between phylogenetic relatedness and ecological interactions. In particular, we connect phylogeny to species' traits, and then similarity in traits to the strength of interaction between any two species [4, 29]. Given a phylogenetic tree representing the evolutionary history of a regional pool of n species, we assume that species interactions are determined by a set of $\ell \geq n$ traits, which have evolved independently on the tree via Brownian motion [18]. Species are assumed to have a baseline competitive effect on each other, which is then modified according their trait covariance. In this way, species that are more closely related tend to interact, on average, more strongly with each other than with distantly-related species. As we will show, the variance of the distribution of interaction strengths is controlled by the number of traits ℓ.

Clearly, species' intrinsic growth rates could also reflect their evolutionary history (e.g., closely related species with similar traits might find similar environments to be harsh or hospitable). To clearly separate the effect of phylogeny on interspecific interactions from its effect on growth rates, we therefore assume that all species have the same intrinsic growth rate. [5]. This assumption severs any connection between phylogeny and environmental filtering.

Having established our model for trait evolution and the link between trait values and species interactions, we analyze the case in which all species in the pool are present at arbitrary

initial conditions, and dynamics follow the Generalized Lotka-Volterra model. Contrary to previous simulation-based studies [14, 23] we develop an analytical framework to characterize the resulting community of coexisting species, as a function of both the number of traits, ℓ, and the tree structure. In particular, we show that when the number of traits is large enough relative to the number of species in the pool, coexistence of all species is guaranteed by the tree-induced interaction structure. Furthermore, the abundance distribution of the community reflects the structure of the tree. On the other hand, while $\ell = n$ is a well-known *necessary* condition for coexistence [24, 41], we find that full coexistence is almost never achieved in this case (see also [13]). Yet, even when coexistence of all n species is precluded, one typically observes coexisting communities of moderate size, as expected if interactions were purely random [10, 33]. Differently from the purely random case, here we find that the probability that a particular species survives is determined by its position in the tree.

Our model shows that phylogenetic relatedness, modulated by the number of traits controlling species interactions, affects multiple aspects of the local community. The explicit incorporation of community dynamics allows us to move from pairwise comparisons to global aspects of community structure. Furthermore, we advance the growing body of literature on random interaction models [3, 6, 10, 33] by analyzing a case in which the correlations between interaction strengths are controlled by phylogenetic relatedness.

Model

Consider a regional pool $\mathcal{R} = \{s_i\}$ of n species indexed by $1 \leq i \leq n$, and assume that a species' identity is defined by its $\ell \geq n$ trait values. For a given trait $1 \leq j \leq \ell$, collect the values of j for all members of the pool in the trait vector $\tau_j \in \mathbb{R}^n$. We sample each τ_j independently from a multivariate normal distribution $\mathcal{N}(0, \Sigma)$. This choice implies that: (a) the values for distinct traits of a given species are independent, and thus we are not considering trade-offs between traits; (b) the processes leading to the correlation structure Σ are statistically equivalent for distinct traits; (c) lastly, if $\Sigma_{ii} = \sigma$ for all i, then the distribution of trait values *within* a species is independent of species identity. For an example of an evolutionary process consistent with the assumptions above, consider the case in which $T_{\mathcal{R}}$ is the phylogenetic tree for the species in the regional pool, and each trait j starts at an

ancestral mean value of 0, and evolves independently on the tree via Brownian motion. Then the value of trait j at the n tips, τ_j, follows the normal distribution $\mathcal{N}(0, \Sigma)$ with Σ induced by the tree structure, and called the variance-covariance matrix of $T_{\mathcal{R}}$ [18]. In Σ, the covariance between two species is given by the shared branch length on $T_{\mathcal{R}}$ [9]. As such, whenever $T_{\mathcal{R}}$ is ultrametric, then $\Sigma_{ii} = 1$ for all i. Unless otherwise specified, here Σ is always assumed to originate from an ultrametric, rooted phylogenetic tree (see Figures 1 and 2).

In this setting, each realization of the ℓ traits defines a species pool $\mathcal{R}$. For a given pool, we imagine that the following experiment is performed [33]: all the species from the pool are introduced in the local habitat *at the same time* and at *arbitrary initial densities*. Population dynamics, as determined by the species' interactions and growth rates, will lead the community to an *asymptotic* state in which some of the species are extinct, while others coexist. Our aim is to characterize the resulting community of coexisting species in terms of the parameters ℓ, n and Σ.

To this end, we consider dynamics governed by the Generalized Lotka-Volterra (GLV) model. Species are assumed to differ only in their interactions, so that the growth of each species in isolation is the same:

$$\frac{dx_i}{dt} = x_i(r - \sum_j (\mu + A_{ij})x_j). \tag{1}$$

Here, x_i is the density of species i and r is the common intrinsic growth rate. The interaction coefficients are modeled as deviations from a "mean-field" competition value $\mu > 0$. These deviations are controlled by trait similarity between species. More precisely, the deviations are modeled as the sample covariance matrix resulting from the trait sampling process, so that competition between two species is strengthened if their trait vectors are positively correlated and weakened otherwise:

$$\begin{aligned} A_{ij} &= \frac{1}{\ell} \sum_l \tau_l^i \tau_l^j, \\ G &= (\tau_j^i), \\ A &= \frac{1}{\ell} G G^T. \end{aligned} \tag{2}$$

In the supplementary information (section S1), we show how this model arises by assuming

a separation of time scales for consumer-resource models in which consumers share the same attack and death rates, but differ in their preferences for resources.

Under these assumptions, A is a symmetric and stable matrix, and a member of the *Wishart ensemble* [30, 40]:

$$A \sim \mathcal{W}_n(\frac{1}{\ell}\Sigma, \ell) \tag{3}$$

The Wishart distribution describes the probability with which a given *sample covariance matrix* is observed when sampling from a multivariate normal distribution. Given its many applications in statistics and other fields, the Wishart distribution has been studied extensively, allowing us to draw upon a large body of results [7, 8, 22, 30].

Since A is stable, the community reaches a unique, globally-stable equilibrium, and the sub-community of coexisting species is characterized by a feasibility and non-invasibility condition [20]. Importantly, in this case one can prove that the effect of the mean interaction strength μ on the resulting community is relatively straightforward: μ does not affect the identity of the coexisting species, and rescales their biomasses by a constant (see supplementary information, section S6, for details). Similarly, any choice of $r > 0$ only rescales the equilibrium biomasses. Thus, without loss of generality, we can assume $\mu = 0$, and $r = 1$, so that the regional pool is completely characterized by A.

To describe the statistical properties of the community of coexisting species, we need to condition the distribution of the variables of interest on the *unique* feasible and non-invasible sub-community for a given species pool $\mathcal{R}$. We focus on the following properties: the distribution of the number of coexisting species, the total biomass of the community, and the relative abundance distribution of the coexisting species.

To illustrate our results with an empirical tree structure, we take the phylogeny of the clade *Senna* (Fabales) as an example [39]. The tree contains a total of 94 species and we use the outlier group comprising the species *Senna silvestris var guarantica, Senna siamea, Senna polyantha* and *Senna galeottiana* to root the subtree containing the remaining 90 species.

Notice that, as shown by [32] the final community composition reached in each of our theoretical experiments is the same as would be reached under sequential, one-at-a-time species invasions. Thus, our results map directly to the process of bottom-up assembly of ecological

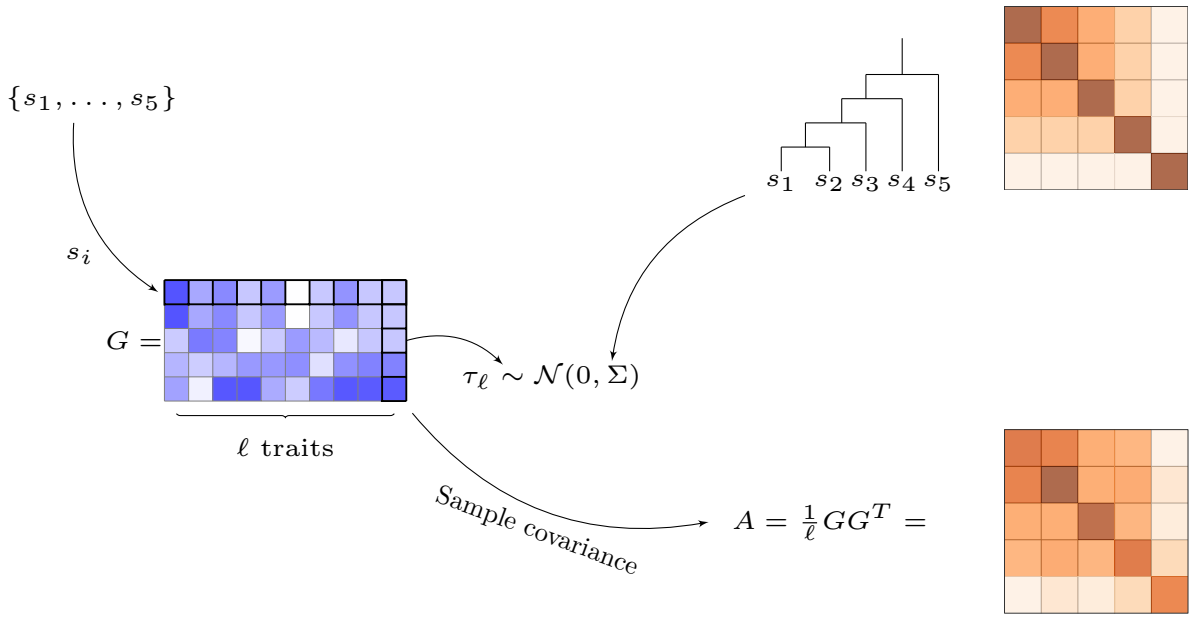

Figure 1: **Construction of the regional pool $\mathcal{R}$ and interaction matrix A.** Each species in the pool $\mathcal{R}$ is assigned ℓ trait values. The vector containing the values for trait j ($\tau_j \in \mathbb{R}^n$) of all members of the pool is sampled independently from $\mathcal{N}(0, \Sigma)$. This is equivalent to a neutral model of trait evolution for each j on a phylogenetic tree $T_\mathcal{R}$. The model relates the structure of $T_\mathcal{R}$ to the interactions between the species in the pool: the matrix Σ measures the shared evolutionary history between any two species s_i and s_j on $T_\mathcal{R}$ (in our example $\Sigma_{12} > \Sigma_{13} > \ldots > \Sigma_{15}$). In turn, the number of traits ℓ and Σ determine the interactions between species, stored in the matrix A.

communities. In particular, our results can be used to infer properties of the assembly graph $\mathcal{G}$ associated with each regional pool [12].

Results

Deterministic Limit. First, consider the case where the number of traits, ℓ, is very large relative to the number of species, n. Let $\gamma = \frac{\ell}{n}$ be their ratio. In the limit $\gamma \to \infty$ we find that $A \to \Sigma$ (i.e., the sample covariance matrix converges to the population covariance matrix). Thus, the properties of the community are determined solely by Σ. The simplest case to study is if $\Sigma = I_n$ (the identity matrix), which represents the covariance matrix induced by the degenerate n-star tree with 0 shared branch length among all species (see Figure 2). This covariance structure corresponds to an evolutionary scenario where all species diverge

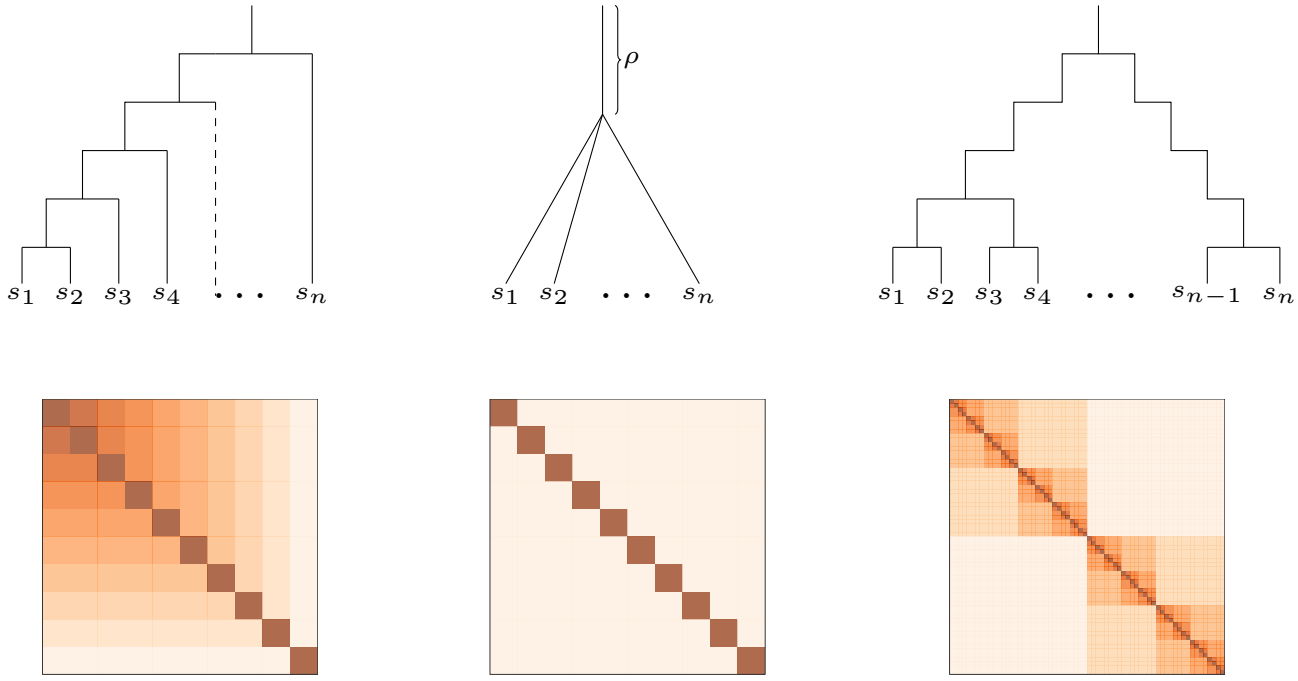

Figure 2: **Examples of ultrametric rooted phylogenies and its induced covariance matrices**. The perfectly unbalanced tree (left) has $n-1$ branching times $0 < t_1 < \ldots < t_{n-1}$ for a pool of n species, where each new branching happens to the "left" and creates a new pair of species. We call the times between branching events, $t_i - t_{i-1}$, *inter-branching times*. The star tree (middle) displays a unique branching event which generates all the n species. For the perfectly balanced tree (right) we have n branching *times* at each of which all the tips present up to that point generate two new species. Proceeding recursively, n branching times generate 2^n species and we have $n+1$ distinct inter-branching times. The covariance matrix associated with each tree is constructed as follows: For any s_i take γ_i to be the path "backwards" in time to the ancestral species at the root of the tree, then for any two s_i, s_j, let $t(i,j)$ be the time at with γ_i and γ_j merge, i.e., the coalescence time between s_i and s_j [36]. Then, $\Sigma_{ij} = 1 - t(i,j)$. In particular, Σ_{ij} is the total time for which the evolutionary processes for s_i and s_j are completely linked. For example, in the star tree $\Sigma_{ij} = \rho$ for any $i \neq j$ and $\Sigma_{ii} = 1$, given that each tree is ultrametric.

immediately at time 0. In this case, complete coexistence of $\mathcal{R}$ and any of its sub-communities $S \subseteq \mathcal{R}$ follows trivially, since species do not interact with each other. Remarkably, the same behavior (full coexistence) is shared by any Σ induced by a tree T. This can be proved inductively using the following observation: If t_1 is the time at which the first split happens in the phylogenetic tree, then "cutting" the tree at this branching point generates non-interacting sub-trees T_i, which we assume to have fully coexisting equilibria under the induction hypothesis. Pasting these sub-trees together at their roots gives us a degenerate tree for which the induced covariance matrix is a block diagonal matrix $\tilde{\Sigma}$. This operation preserves coexistence, since the sub-trees are still non-interacting. We recover T by attaching a branch length t_1 to the root. In terms of the variance-covariance matrix, Σ is obtained by adding a constant to $\tilde{\Sigma}$. Assuming equal growth rates, this transformation does not affect the feasibility of the system; hence Σ has a feasible equilibrium (see fig S1 and supplementary information S2 for a more detailed argument). Thus, for $\gamma \gg 1$ (i.e., if $A \approx \Sigma$), we have full coexistence regardless of the tree topology. Moreover, the assembly graph $\mathcal{G}$ for the species pool $\mathcal{R}$ contains all possible assembly histories (c.f. [32] and [42]); in other words, any sub-community can be built by starting with a single species and adding the remaining members sequentially in any order.

Total biomass and abundance distribution. Under our model, phylogeny strongly influences the biomass and relative abundance distribution of a community. As illustrative examples, consider the two extreme tree topologies given by the "perfectly unbalanced" tree and the "perfectly balanced" tree (Figure 2). Assuming equal inter-branching times, the total biomass of the system, $W(n)$, for a pool of n species is given by $W(n) \approx \sqrt{n} - 1/4$ in the perfectly unbalanced case, and $W(n) = \frac{\log_2(n)+1}{2-1/n}$ in the perfectly balanced case (see section S2 for details). Similarly, we are able to derive expressions for the individual biomass of each species s_i, where the index corresponds to the position in the ordered tips of the tree (see Figure 2). For the perfectly balanced case, the abundance distribution is trivial, since each species necessarily has the same abundance. On the other hand, the hierarchical nature of the perfectly unbalanced tree is reflected in the individual biomasses, with species that split from the rest early on having much higher abundances. Fig 3 and S2 show that the results are qualitatively unchanged if we sample the inter-branching times from appropriately normalized exponential or uniform distributions. The uneven distribution of abundances for the unbalanced tree helps explain the difference in total biomass: in the perfectly unbalanced

case, as n grows there is a fraction of species (outliers) that interact less and less strongly with the rest of the community, so that their abundance approaches the limit 1 (obtained for non-interacting species). In contrast, in the perfectly balanced case the abundance of all species is the same, and approximately $\frac{\log_2(n)}{n}$.

To compare these results with predicted abundances using a more complicated tree structure, we repeated the calculation using the phylogenetic tree of the *Senna* clade (Fabales) [39]. We considered two cases: either we include the branch length information, or we set all branch lengths to be equal, so that only the effects due to the shape of the tree are considered. The average total biomass $W(n)$ for sub-communities of different sizes (Figure 3) shows that for both cases at small enough sizes, $W(n)$ behaves as predicted by the perfectly unbalanced model—which reflects the hierarchical low-level structure of the tree (Figure 6). But as the size of the sub-community increases, $W(n)$ either reaches values even smaller than the perfectly unbalanced tree, or settles in the middle of the two—showing that, under equal inter-branching times, the perfectly balanced and perfectly unbalanced tree represent the two extreme topologies. The species' abundance distribution, as for the perfectly unbalanced tree, reflects the tree structure: the abundance profile shows peaks at each of the outliers within clades, and an overall decreasing trend toward more deeply nested parts of the tree (see also Figure 6).

Star phylogenies. Classical results in theoretical ecology have extended the principle of competitive exclusion to the case of multiple resources/regulating factors, showing that a necessary condition to observe a non-degenerate coexisting community of n species in our model is $\ell \geq n$ [24, 41]. We have shown above that, for a fixed size of the pool, n, coexistence is guaranteed in the $\ell \to \infty$. To characterize the cases in between $\ell = n$ and $\ell \to \infty$, we exploit the fact that A follows the Wishart distribution; as such we can make use of tools developed in statistics and economics to explore how the limit of full coexistence is approached (see section S3). To start, let Σ be induced by a star-tree with shared root of length ρ (see Figure 1). In this setting, there is a constant correlation ρ among the species in the pool. We find that for $\gamma \approx 1$, full coexistence is never achieved for large enough communities (Fig S3). Nevertheless, the community does not collapse completely, and a non-vanishing fraction of species is observed to coexist (Figure 4). The effect of increasing correlation among species is, as expected, to reduce the proportion of coexisting species, $\wp$. In particular, to observe at

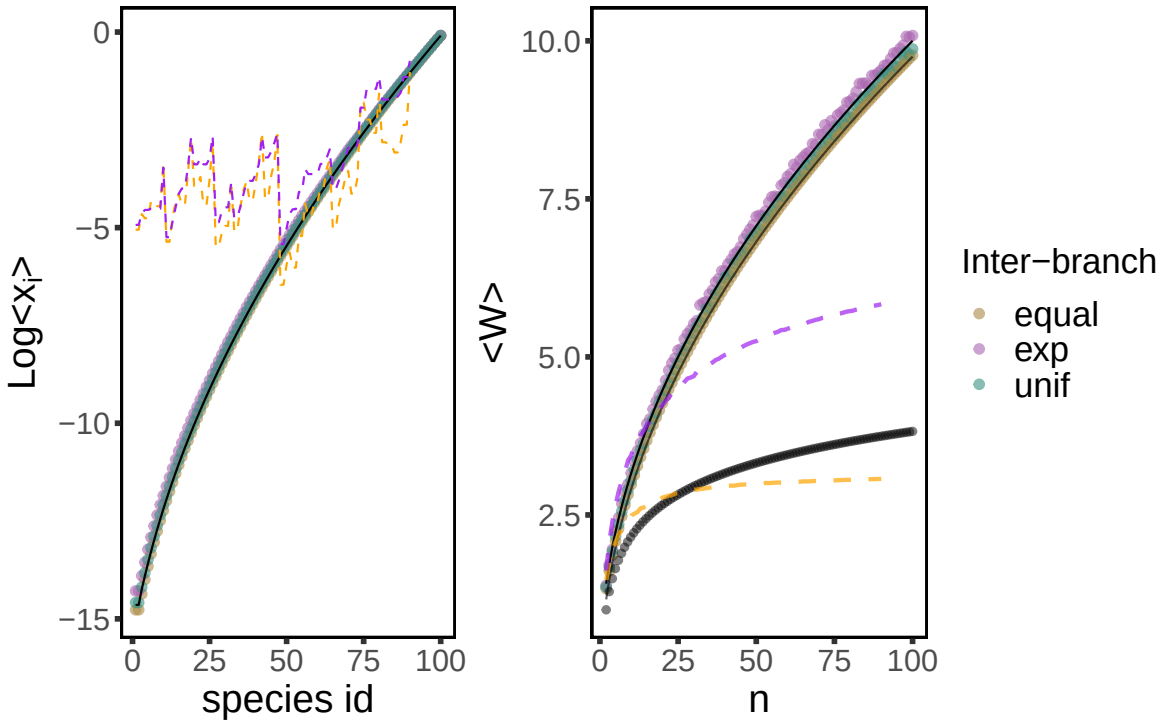

Figure 3: **Individual and total abundance for the deterministic limit**. Log individual abundance (left) and total abundance (right) for the communities in the deterministic limit of a perfectly unbalanced tree. Dots mark the average values when sampling the branch lengths from an exponential distribution with rate 1, a uniform $[0, 1]$ distribution, and the case of equal branch lengths. The total branch length is renormalized to 1 in all cases. Solid lines are the analytic predictions under equal branch length $1/n$. In the right panel, the two solid lines are given by $\sqrt{n}$ and $\sqrt{n} - 1/4$, the black dots represent the analytic formula for a perfectly balanced tree which shows logarithmic growth $\sim \frac{\log_2(n)+1}{2}$. The dashed lines on the right panel are the scaling with size of sub-trees of the *Senna* phylogenetic tree, and the dashed lines on the left plot are the abundance distribution for the full tree (compare with Figure 6). In both cases the purple line is for the case of equal inter-branching times and the orange includes the branch length information.

least half of the species to coexist (in expectation), we need:

$$2\gamma \geq 1 + \frac{n\rho}{\pi(1-\rho)} \tag{4}$$

The quantity $\zeta = \frac{\rho}{1-\rho}$ could be interpreted in the framework of population genetics as the ratio of shared to private mutations for each species. It is a key quantity, in the sense that two distinct pools $\mathcal{R}$ and $\mathcal{R}'$, of sizes n and n' will yield the same mean fraction of coexisting species, for a given γ, whenever $n\zeta = n'\zeta'$.

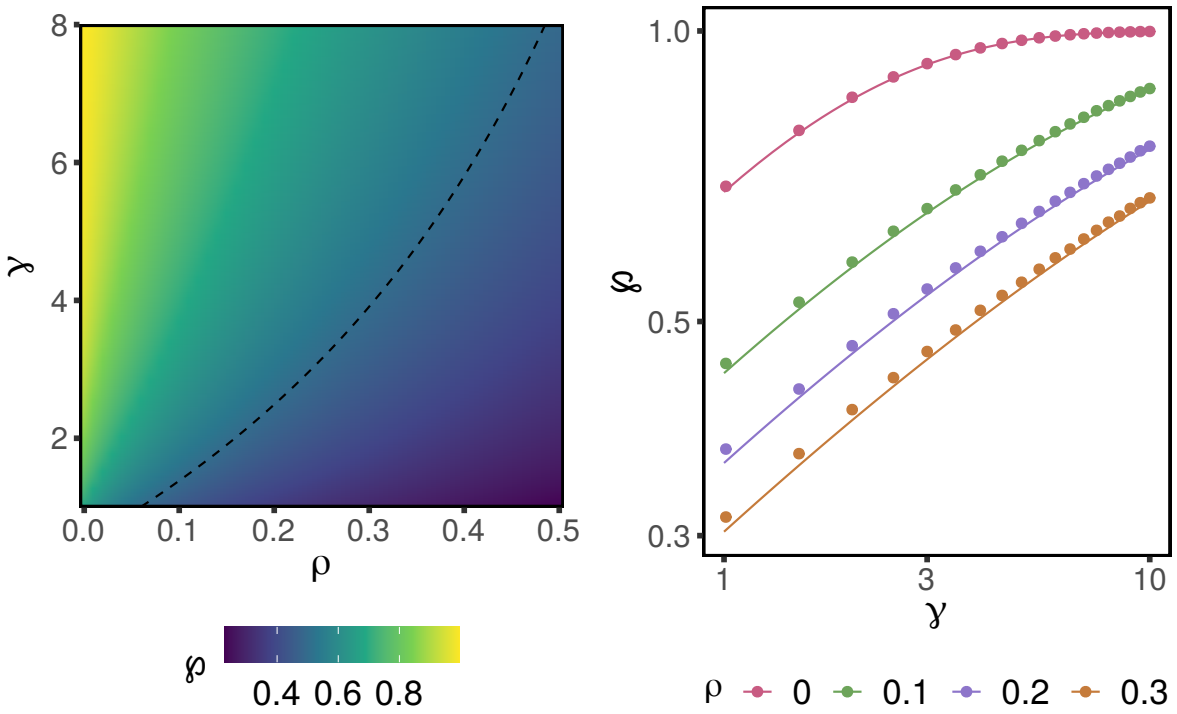

Figure 4: **Proportion of coexisting species $\wp$ as a function of ρ and γ.** In the right panel, we compare our analytical approximations (solid lines) with simulations (dots) for a regional pool of 50 species (log-log scale). The left panel explores in more detail the parameter space (γ, ρ). The dashed line marks parameters for which we expect half of the species to coexist. As suggested by eq. (4), the value of γ giving a fixed $\wp$ increases sharply with ρ.

The distribution of total biomass, W, for the community of coexisting species is influenced by γ and ρ in two different ways: the parameters affect both the distribution of the number of coexisting species, and the conditional distribution of W for a given community size. Assuming that the distribution of the number of coexisting species is highly peaked at the mode, we

derive an approximation for the mean of W that closely matches results from simulations (see section S4 and figure S5 for exact results and full distribution). For small enough γ the variance of the interactions allows the possibility of positive interactions that enhance W, as γ increases the interaction matrix converge to the purely competitive interaction matrix given by Σ. This convergence explains the decrease of W with γ depicted in Figure 5.

Using the same strategy, we are able to derive approximations (see section S5 for exact formula) for the survival function of the relative abundance distribution under distinct values of ρ and γ. In particular, the distribution becomes very peaked as γ increases, while increasing ρ tends to make the distribution flatter (Figure 5).

Beyond constant correlation. Imposing a more general covariance structure Σ is challenging from a mathematical standpoint, due to the breaking of the statistical equivalence among species—species in distinct parts of the tree have now different statistical properties. In the general case, the identities of the species matter, and instead of looking at the total number of coexisting species, we focus on how the probability that a particular species survives (p_s) changes with its position in the tree. Simulations for the phylogenetic tree of the *Senna* clade show that the model recreates the phenomenon of phylogenetic over-dispersion: for a group of closely related species, p_s peaks at the outliers of the clade. Furthermore p_s reflects the tree structure in the same manner as the total abundance distribution (compare Figures 1 and 6).

To further explore this relationship, we are able to analytically compute the probability of observing each sub-community in a three-species community (Figure 7). For $n = 3$, there is only one possible tree topology, and we consider the case where all branch lengths are equal. We find that sub-communities containing the outlier species, s_3, are always more likely to be observed than sub-communities of the same size in which s_3 is absent. More generally, the formulas in section S3 can be evaluated numerically to find the probability of observing a particular sub-community under an arbitrary, not necessarily ultrametric, tree structure.

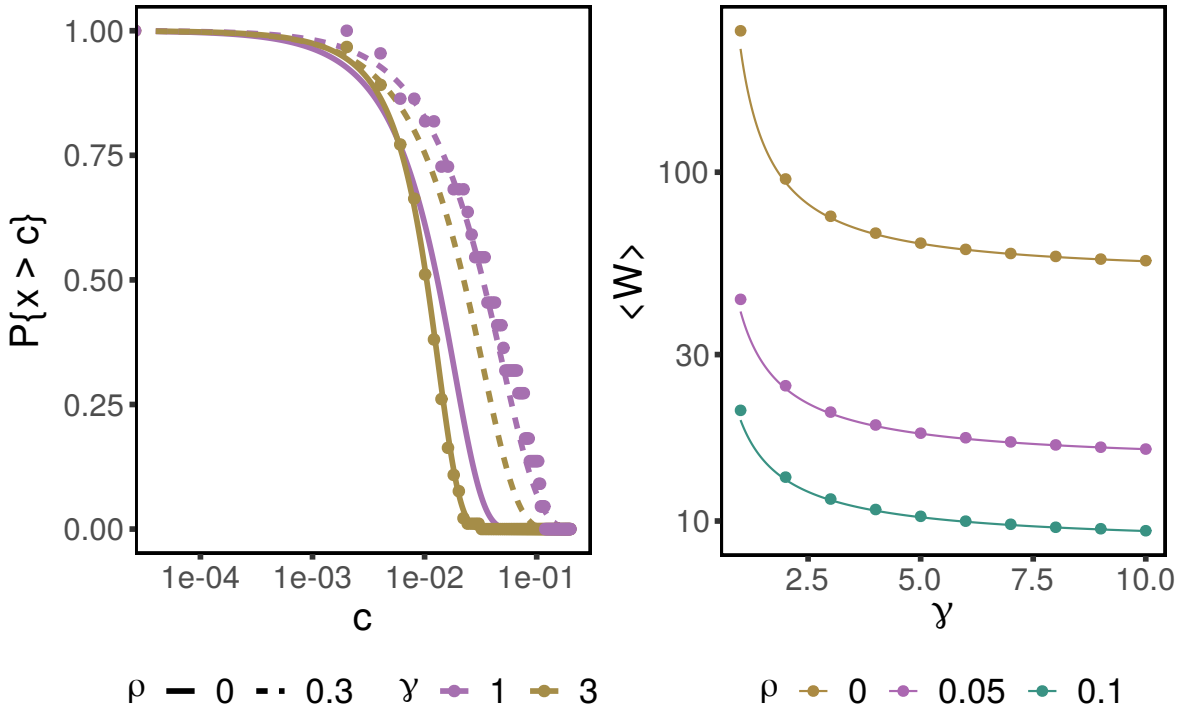

Figure 5: **Mean total biomass and relative abundance distribution**. The panel on the right shows (note the log-transformation for the y-axis) the mean total biomass for the community of coexisting species; the points represent simulations, and solid lines the corresponding analytical approximations for a pool of 50 species (see section S6 for the effect of changing μ). The total biomass decreases as γ grows, because the overall strength of interaction between species decreases. The survival function for the relative abundance values of the community is plotted on the left panel (note the log x-axis), where again points stand for simulations and lines for analytical predictions for distinct γ and ρ values, and a pool of 100 species. For clarity, we just show simulations for the parameters $(\rho, \gamma) \in \{(0, 3), (0.3, 1)\}$. In particular we have that as γ increases the distribution becomes more and more peaked (as expected) while increasing ρ flattens the distribution.

.

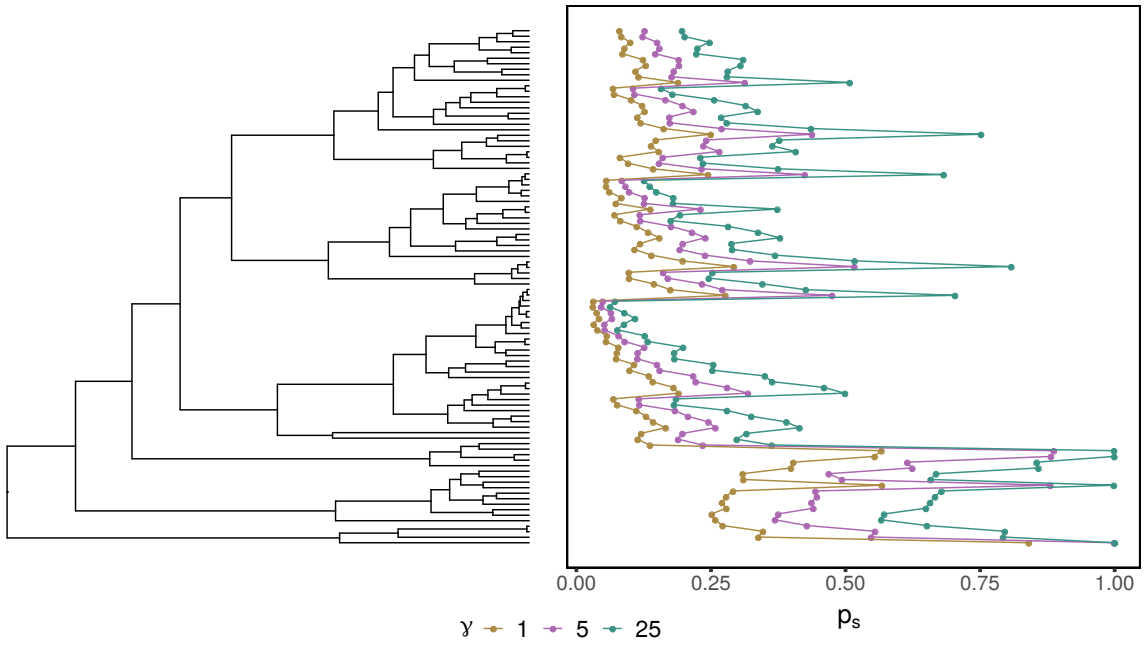

Figure 6: **Probability of individual species survival for an empirical tree**. The probability that a species is observed in the community of coexisting species, p_s, out of 5000 simulations, is shown alongside the phylogenetic tree (*Senna* clade) where the outermost group is used to set the root. The values p_s reflects the tree structure and the abundance distribution showed in Figure 1: The peaks in p_s correspond to outliers within a group of closely related species and p_s has a decreasing trend towards the most nested parts of the tree (upward direction). In particular, the model produces phylogenetic over-dispersion.

Discussion

By considering local community dynamics in a trait-based interaction model, our results provide a clear link between the phylogeny of the regional species pool and many aspects of species coexistence. Importantly, while the tree structure is reflected in the local community patterns, the number of traits controlling interspecific interactions modulates the outcomes.

We found that, when phylogenetic relatedness completely controls interactions, i.e., when the number of traits is sufficiently high compared to the number of species, full coexistence of any sub-community is guaranteed. This result requires both the tree structure (which induces a particular interaction matrix) and the assumption that all species have equal growth rates.

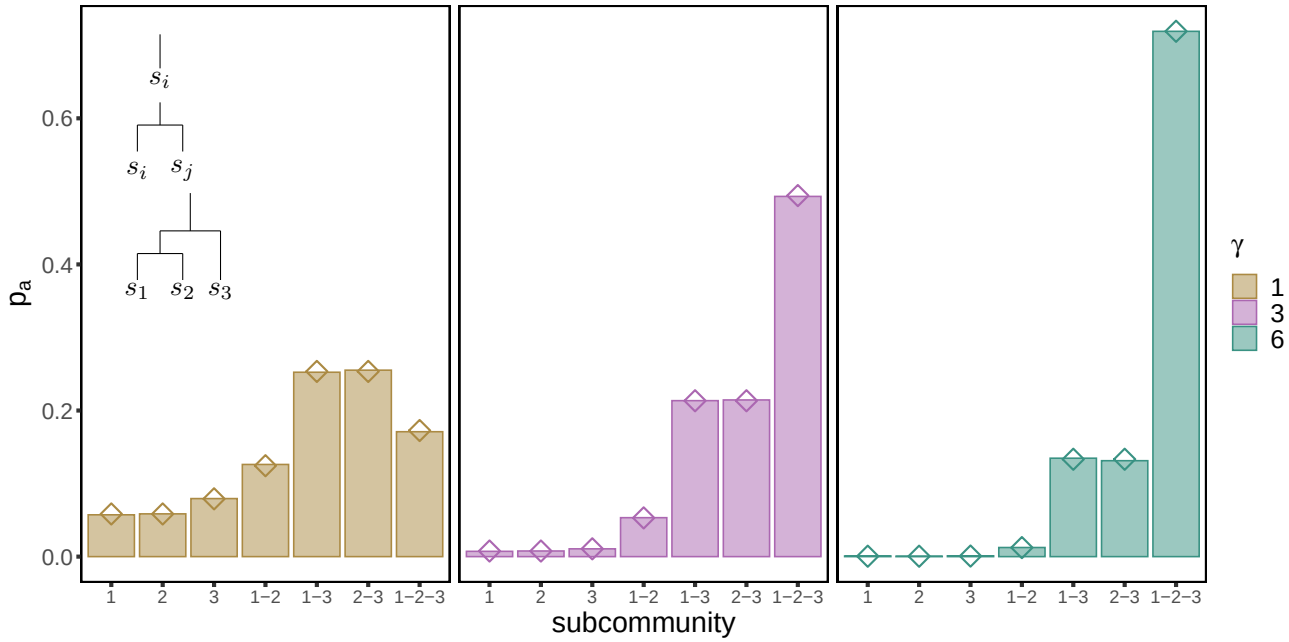

Figure 7: **Sub-communities of perfectly unbalanced tree** Probability of observing a given sub-community of the three-species tree with equal branch lengths. The inset shows the tree sub-structures for each of the sub-communities. Bars represent frequencies over 50000 simulations, and dots the analytical predictions.

While we expect this result to hold qualitatively for small deviations from the assumption of identical growth rates (see section S7), it is false in general when these requirements are not satisfied. Under the same assumptions, the abundance distribution of the community reflects the tree structure at distinct levels: high biomass is observed for the outliers within each clade (local tree structure), and one expects an overall decreasing trend towards more nested parts of the tree (coarser structure).

When the number of traits is comparable to the number of species, our model is an instance of a Lotka-Volterra model with random interactions. The analysis of models considering random interactions between species has a long tradition in ecology [1, 17, 27], and in recent years the field has moved beyond questions concerning the stability and feasibility of the whole system, focusing more closely on the properties of sub-communities that coexist through the dynamics [3, 6, 10, 33]. In these models, one must usually assume that species interactions are independent of species identity (but see [2, 16]). The star-tree case studied above satisfies this assumption, but with a stronger correlation structure than has been previously considered.

This case behaves much like other random interaction models: full coexistence of many species is extremely unlikely, but we expect a moderate number of species to coexist. We have derived an approximation for the mean number of coexisting species, which depends on the ratio of traits to species, γ, and on the expected ratio of shared to private mutations for each species $n\zeta$. As long as these two quantities are the same, pools of distinct sizes will yield the same distribution for the number of coexisting species. Contrary to previous studies [33], the analytic tractability of the model allows us to derive exact expressions for the total biomass and relative abundance distribution of the system.

The general case of an arbitrary, tree-induced correlation structure provides a biologically-meaningful way to relax the statistical equivalence between species. Taking advantage of the vast literature on the Wishart ensemble in fields ranging from economics to statistics [8, 22, 30], we are able to derive exact integral formulas to compute the probability of survival for any sub-community under arbitrary tree structure. In this way, one can measure properties of the system (conditioning on a final sub-community) by numerically evaluating the integral expression. For small enough communities and simple enough phylogenies, this approach can be replicated on each sub-system to compute the marginal distribution of the properties of interest. However, as the number of species grows, these calculations become burdensome. As such, devising new analytical techniques to tackle the general case would be an important step toward studying more general random interaction models, and also advance our understanding of the effects of phylogenies on communities.

Our approach can be extended in a variety of ways, and we briefly discuss some of the most promising avenues.

First, instead of assuming that the same tree structure controls the evolution of all ℓ traits, we can partition the traits into m classes and assume that the evolution of each class is determined by a distinct phylogenetic tree. These type of processes are studied in population genetics when either admixture or incomplete lineage sorting lead to traits that cannot be explained by a single tree [31]. In such cases, A would no longer follow the Wishart distribution but would rather be a sum of (possibly degenerate) Wishart matrices.

Second, our assumption of equal growth rates among species allowed us to examine how phylogenetic relatedness influences coexistence in a purely interaction-driven model. When we include variation in growth rates, we expect our results to hold only for sufficiently small

variance. In this case, the restriction on $l \geq n$ can be lifted, provided that there is a background competitive effect μ strong enough to prevent divergence of the dynamics (see section S7). It would be interesting to consider models where growth rates vary under the influence of phylogeny; by modulating how strongly evolutionary relatedness affects both growth rates and interactions, one could investigate the duality between "competition" and "filtering" that is frequently discussed in the literature [14, 28, 38].

Lastly, our approach assumes an explicit separation between evolutionary processes at the regional level (which give rise to the phylogenetic structure) and ecological interactions (at the local level). To remove this separation, one could model the tree generation process and ecological dynamics concurrently. For example, as done by Maynard et al. [29], one could "run" the dynamics after each new speciation event, thereby pruning the community to a coexisting sub-community. One would then take the sub-tree of that community as the starting point for the new speciation event. In this setting, in a similar manner to studies of community assembly [32] and the framework of adaptive dynamics [21], we have a separation of time-scales between the speciation events and the local community dynamics. Traits evolve on the tree between each pruning event. In this regard, our results provide baseline comparisons, and even suggests patterns that would emerge from the process: assuming that the number of traits is a constant ℓ, the community cannot reach more than ℓ species, yet at the early steps of the process the ratio of traits to species could be extremely high—hence we expect that most speciation events occurring early on would not cause extinctions; in this case, the bulk of the phylogenetic structure would be built at the beginning of the process. Perturbing the growth rates slightly, one could compare the structure of this tree with the structure of the tree found by simply letting the tree generation process run, and after having the same number of speciation events let the species interact and get a coexisting sub-community.

While there has been extensive discussion of the potential ways in which phylogeny could affect ecological differences, and thus interactions, among species [11], much less has been said about the patterns one would observe under a particular hypothesis. In this work, by linking phylogenies to a simple model of trait evolution and local community dynamics, we were able to fully characterize many global aspects of the community. We showed that the phylogenetic structure of the species pool and the the number of traits determining competition affect the results in concert. Our results provide a useful baseline prediction for the effect of phylogeny

on community dynamics and coexistence.

Acknowledgments. We thank P. Lemos and M.O. Carlson for comments on the manuscript. J.A.C. acknowledges financial support from the Spanish 'Ministerio de Economía y Competitividad' project PGC2018-096577-B-I00. Z.R.M. acknowledges support from the National Science Foundation Graduate Research Fellowship Program under Grant No.(DGE-1746045).

Supplementary information

1 Motivation

From consumer-resource dynamics to covariances

We start with a model of consumer-resource dynamics in which the consumers differ only in the relative preference of each resource and the resources have an homogenous growth rate. Let $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{y} \in \mathbb{R}^\ell$ be vectors denoting the density of predators and resources. We model the dynamics as the MacArthur's consumer-resource model [25]:

$$\begin{aligned} \frac{d\boldsymbol{x}}{dt} &= \boldsymbol{x} \circ (-d\mathbf{1}_n + \alpha\tilde{G}\boldsymbol{y}), \\ \frac{d\boldsymbol{y}}{dt} &= \boldsymbol{y} \circ (r\mathbf{1}_\ell - \boldsymbol{y} - \beta\tilde{G}^T\boldsymbol{x}), \end{aligned} \tag{5}$$

where $\circ$ stands for the Hadamard (component-wise) matrix product, and $\mathbf{1}_k = (1, \ldots, 1)^T \in \mathbb{R}^k$ is a notation for a column vector whose entries are exactly k ones.

By our assumptions, matrix $\tilde{G} \in \mathbb{R}_+^{n\times\ell}$ encodes the preference distribution (alternatively, the time allocation distribution) of the predators over the resources, so that $\tilde{G}\mathbf{1}_\ell = \mathbf{1}_n$. Then by a separation of time scales, which implies that resource densities remain at equilibrium, we can model the competition between the consumers as following competitive Lotka-Volterra dynamics [25]:

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{x} \circ (\alpha r\tilde{G}\mathbf{1}_\ell - d\mathbf{1}_n - \alpha\beta\tilde{G}\tilde{G}^T\boldsymbol{x}) = \boldsymbol{x} \circ ((\alpha r - d)\mathbf{1}_n - \alpha\beta\tilde{G}\tilde{G}^T\boldsymbol{x}). \tag{6}$$

As long as $n \leq \ell$ (besides measure zero sets) we have that matrix $\tilde{A} := \tilde{G}\tilde{G}^T$ is positive definite. This property of $\tilde{A}$ allows one to further transform the system (6) without affecting

the set of coexisting species. In particular we can perform the following operations (see section 6 for a more detailed discussion):

(a) Rescale the growth rate, $\boldsymbol{v} = (\alpha r - d)\mathbf{1}_n$, by any positive constant.

(b) Multiply $\tilde{A}$ by a positive, constant diagonal matrix.

(c) Multiply both $\tilde{A}$ and $\boldsymbol{v}$ by a positive diagonal matrix.

Following this operations we reduce the system to

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{x} \circ (\mathbf{1}_n - \tilde{G}\tilde{G}^T\boldsymbol{x}). \tag{7}$$

To distinguish the effect of the mean of $\tilde{G}$, write $\tilde{G} = G + \frac{1}{n}\mathbf{1}_n\mathbf{1}_\ell^T$. Notice that this decomposition, together with the restriction $\tilde{G}\mathbf{1}_\ell = \mathbf{1}_n$, implies that $G\mathbf{1}_\ell = \mathbf{0}_n$, which means that the entries of G have zero mean —here $\mathbf{0}_k = (0, \ldots, 0)^T$ stands for a column vector formed by k zeros. Then matrix $\tilde{A}$ can be decomposed as $\tilde{A} = GG^T + \mathbf{1}_n\mathbf{1}_n^T$. Because the system in (7) has constant growth rates then one can show (section 6) that, as long as $\ell > n$ (the strict inequality arising due to G having rank $\ell - 1$), the set of coexisting species for (7) is invariant to the shift $\mathbf{1}_n\mathbf{1}_n^T$. Therefore the system reduces to:

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{x} \circ (\mathbf{1}_n - GG^T\boldsymbol{x}) = \boldsymbol{x} \circ (\mathbf{1}_n - A\boldsymbol{x}), \tag{8}$$

where we have defined $A := GG^T$. This is the competitive, deterministic dynamics that we have assumed for consumers throughout this contribution. Observe that the set of coexisting species remains unchanged if we define interaction matrix $A = \frac{1}{\ell}GG^T$, as in the main text, because of the aforementioned invariant operations.

Modelling the covariance matrix

From (8) we see that the interactions between species A_{ij} are fully determined by the row vectors $\boldsymbol{G}_i$. Because each row $\tilde{\boldsymbol{G}}_i$ of matrix $\tilde{G}$ is a preference vector, then it lies on the standard $\ell - 1$ dimensional simplex $\Delta^{\ell-1} = \{\tilde{\boldsymbol{G}}_i \in \mathbb{R}^\ell | \sum_{j=1}^{\ell} \tilde{G}_{ij} = 1, \text{ for } i = 1, \ldots, n\}$, which implies that $\boldsymbol{G}_i$ lies on a bounded subset of a linear subspace of $\mathbb{R}^\ell$ defined by the restrictions $\sum_{j=1}^{\ell} G_{ij} = 0$ for $i = 1, \ldots, n$. By choosing a suitable (linear) coordinate system $\{\boldsymbol{w}_j\}_{j=1}^{\ell}$ we

can express

$$\begin{aligned} \boldsymbol{G}_i &= \sum_{j=1}^{\ell} c_i^j \boldsymbol{w}_j, \\ A_{ij} &= \boldsymbol{G}_i \boldsymbol{G}_j^T = \sum_{k=1}^{\ell} c_i^k c_j^k. \end{aligned} \tag{9}$$

Therefore, the entries of A are fully determined by the coordinates of row vectors $\boldsymbol{G}_i$ on the basis $\{\boldsymbol{w}_j\}_{j=1}^{\ell}$.

To model coordinates c_i^j we assume that each (rescaled) preference vector $\boldsymbol{G}_i$ is the result of a diffusion process starting at the origin of this space (this maps back to our $\tilde{\boldsymbol{G}}$ matrix as saying that every consumer has an *homogeneous* preference for any resource). Assuming that each coordinate is independent and letting the diffusion time be small enough, then coefficients c_i^j are normally distributed, $c_i^j \sim \mathcal{N}(0, \sigma)$. The invariant properties of the model allow us to forget about the deviation σ and simply model $c_i^j \sim \mathcal{N}(0, 1)$. This shows that A satisfies the assumptions of model (8) up to a change of number of traits from ℓ to $\ell - 1$.

2 Deterministic limit

Full coexistence

We provide more details for the proof that, in the deterministic limit, every subcommunity of the pool is feasible. Since every subcommunity has an interaction matrix induced by a tree, it is enough to show that feasibility is guaranteed whenever this is the case.

We proceed by induction on the number of species. For $n = 1$ the claim holds trivially. Let T be a phylogenetic tree (not necessarily ultrametric) for n species, and Σ its respective covariance matrix. Let t_1 be the time at which the first split happens, so that at t_1 the ancestral branch splits into $m \geq 2$ lineages (L_i, with $i = 1, \ldots, m$) where each L_i contains at most $n - 1$ species. Lineages are defined by the condition that species $j, k \in L_i$ if and only if the shared branch length between both species $t(j, k)$ satisfies $t(j, k) > t_1$. That is, each lineage contains the subset of species whose shared evolutionary time is strictly greater than t_1. For each L_i, take T_i to be the subtree induced by L_i. Consider $\tilde{T}$, the tree obtained by shrinking the segment between the root and t_1 to a point (see Fig. 8), then $\tilde{T}$ is a phylogenetic

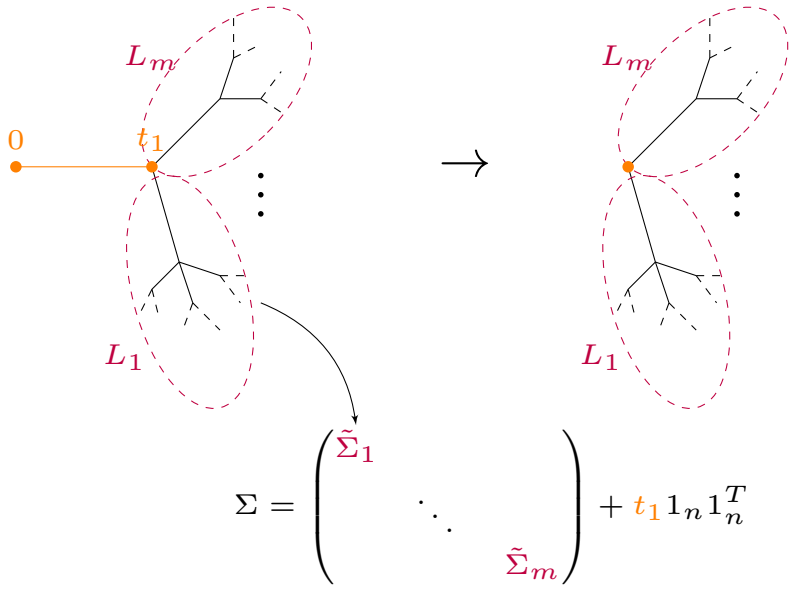

Figure 8: **Schematic representation of the inductive step on the proof of full coexistence**. Starting with the tree T (left), we shrink the ancestral branch up to the first splitting time t_1 to have a degenerate tree $\tilde{T}$ (on the right). $\tilde{T}$ splits at time 0 into m distinct subtrees induced by the lineages L_i for $i = 1, \ldots, m$. The covariance matrix for T, Σ, is obtained from the covariance matrix $\tilde{\Sigma}$ of $\tilde{T}$ by "adding back" the ancestral branch. This amounts to a constant rank-one update of $\tilde{\Sigma}$ which preserves feasibility.

tree, for which the covariance matrix $\tilde{\Sigma}$ is block diagonal and given by diagonal blocks $\tilde{\Sigma}_i$. Each $\tilde{\Sigma}_i$ is the covariance matrix of the tree $\tilde{T}_i$ which is obtained from T_i by shrinking the root branch by t_1. By induction it follows that each block $\tilde{\Sigma}_i$ is feasible, hence $\tilde{\Sigma}$ is also feasible. Observe that, going backwards, T is obtained from $\tilde{T}$ by adding a root segment of length t_1. In particular this says that the shared evolutionary times of all species increases by t_1, i.e. $\Sigma = \tilde{\Sigma} + t_1 \mathbf{1}_n \mathbf{1}_n^T$, so that Σ is a constant rank-one update of $\tilde{\Sigma}$. Then by section 6, the equilibrium associated to Σ is feasible.

Perfectly hierarchical trees

Consider a perfectly hierarchical tree T_n with n tips and branching times $t_0 = 0 < t_1 < \ldots < t_n < 1$ (see figures 1-2 of the main text), and let Σ_n be its covariance matrix. Then it follows trivially that

$$\Sigma_n = \begin{pmatrix} \tilde{\Sigma}_{n-1} & \mathbf{0}_{n-1} \\ \mathbf{0}_{n-1}^T & s_1 \end{pmatrix} + t_1 \mathbf{1}_n \mathbf{1}_n^T, \tag{10}$$

where $s_i := \sum_{j=i+1}^{n} \Delta t_j$, for $\Delta t_j = t_j - t_{j-1}$ the time between two branching events— the *inter-branching time*. In this subsections we find accurate bounds for the total biomass and analyze the expected abundance distribution.

Define the vector of abundances $\boldsymbol{x}_n = (x_n^i)$ for a hierarchical tree T_n with n tips. In the deterministic limit, this vector satisfies the linear system

$$\Sigma_n \boldsymbol{x}_n = \mathbf{1}_n. \tag{11}$$

As in the proof of feasibility, $\boldsymbol{x}_n$ is given recursively by the updated equilibrium abundances $\tilde{\boldsymbol{x}}_{n-1}$and s_1^{-1} of the non-interacting subtrees $\tilde{T}_{n-1}$ and the one formed by the first species, respectively. Indeed, if we look for solutions of the form $\boldsymbol{x}_n = \begin{pmatrix} a\tilde{\boldsymbol{x}}_{n-1} \\ x_n^n \end{pmatrix}$, where the vector of abundances $\tilde{\boldsymbol{x}}_{n-1}$ satisfies $\tilde{\Sigma}_{n-1}\tilde{\boldsymbol{x}}_{n-1} = \mathbf{1}_{n-1}$, $\tilde{\Sigma}_{n-1}$ being the covariance matrix of the subtree $\tilde{T}_{n-1}$, the equilibrium condition (11) for $\boldsymbol{x}_n$ reduces to a linear system for a and x_n^n:

$$\begin{cases} a + at_1 \mathbf{1}_{n-1}^T \tilde{\boldsymbol{x}}_{n-1} + t_1 x_n^n = 1, \\ at_1 \mathbf{1}_{n-1}^T \tilde{\boldsymbol{x}}_{n-1} + (s_1 + t_1) x_n^n = 1. \end{cases} \tag{12}$$

The solution is $a = s_1 x_n^n$, with $x_n^n = (s_1 + t_1 + s_1 t_1 \mathbf{1}_{n-1}^T \tilde{\boldsymbol{x}}_{n-1})^{-1}$. Let $\tilde{W}_{n-1} := \sum_{i=1}^{n-1} \tilde{x}_{n-1}^i = \mathbf{1}_{n-1}^T \tilde{\boldsymbol{x}}_{n-1}$. Then $\boldsymbol{x}_n$ can be written in terms of $\tilde{W}_{n-1}$, $\tilde{\boldsymbol{x}}_{n-1}$, $s_0 = s_1 + t_1$, and s_1 as

$$\begin{aligned} x_n^n &= \frac{1}{s_0 + t_1 \tilde{W}_{n-1} s_1}, \\ x_n^i &= \frac{s_1 \tilde{x}_{n-1}^i}{s_0 + t_1 \tilde{W}_{n-1} s_1}, \quad 1 \le i < n. \end{aligned} \tag{13}$$

In particular, this implies the following recurrence for the total biomass, W_n:

$$W_n = \frac{1 + \tilde{W}_{n-1} s_1}{s_0 + t_1 \tilde{W}_{n-1} s_1}. \tag{14}$$

In the case of equal inter-branching times, $\Delta t_i = \frac{1}{n}$ for all $i = 1, 2, \ldots, n$, observe that $s_0 = 1$, $s_1 = \frac{n-1}{n}$ and $\tilde{\Sigma}_{n-1} = \frac{n-1}{n}\Sigma_{n-1}$. Hence $\boldsymbol{x}_{n-1} = s_1 \tilde{\boldsymbol{x}}_{n-1}$ and $W_{n-1} = s_1 \tilde{W}_{n-1}$, so

Eqs. (13) and (14) above reduce to:

$$\begin{aligned} x_n^n &= \frac{n}{n+W_{n-1}}, \\ x_n^i &= \frac{n x_{n-1}^i}{n+W_{n-1}}, \quad 1 \le i < n, \\ W_n &= \frac{n(1+W_{n-1})}{n+W_{n-1}}. \end{aligned} \tag{15}$$

The following proposition provides accurate upper and lower bounds for total biomass in the limit of large number of species.

Proposition 1. *Let*

$$\varphi(n) := \frac{4n-1-\sqrt{16n^2+1-8n\sqrt{n-1}}}{4\sqrt{n-1}}. \tag{16}$$

Then, for equal branching times, it holds that $\sqrt{n}-\varphi(n) > W_n > \sqrt{n}-1/4$ *for* $n \ge 2$ *and* $\varphi(n) \to 1/4$ *in the limit* $n \to \infty$.

Proof. Direct computation shows that the inequality holds at $n = 2$ so we proceed by induction on n.

Consider first the lower bound. Suppose it holds at $n-1$, then:

$$W_n = \frac{n(1+W_{n-1})}{n+W_{n-1}} = n\left(1-\frac{n-1}{n+W_{n-1}}\right) > \frac{n(\sqrt{n-1}+3/4)}{n+\sqrt{n-1}-1/4}.$$

If the claim were not satisfied at n we would have

$$\sqrt{n}-1/4 \ge \frac{n(\sqrt{n-1}+3/4)}{n+\sqrt{n-1}-1/4}.$$

Rearranging terms, this gives the following chain of equivalent inequalities:

$$\begin{aligned} n\sqrt{n}+\sqrt{n-1}\sqrt{n}+\frac{1}{16} &\ge n\sqrt{n-1}+n+\frac{1}{4}(\sqrt{n-1}+\sqrt{n}), \\ n(\sqrt{n}-1)+\sqrt{n-1}\sqrt{n}(1-\sqrt{n})+\frac{1}{16} &\ge \frac{1}{4}(\sqrt{n-1}+\sqrt{n}), \\ \sqrt{n}(\sqrt{n}-1)(\sqrt{n}-\sqrt{n-1})+\frac{1}{16} &\ge \frac{1}{4}(\sqrt{n-1}+\sqrt{n}). \end{aligned} \tag{17}$$

Multiplying both sides by $\sqrt{n-1}+\sqrt{n}$ we get

$$\sqrt{n}(\sqrt{n}-1)+\frac{1}{16}(\sqrt{n-1}+\sqrt{n})\geq\frac{1}{4}(\sqrt{n-1}+\sqrt{n})^2=\frac{1}{4}(2n-1+2\sqrt{n-1}\sqrt{n}). \quad (18)$$

The last inequality implies

$$\frac{3}{4}\geq\frac{7}{8}\sqrt{n},$$

which says $n\leq 1$. This is a contradiction and we are done.

We proceed in the similar way for the upper bound. By induction hypothesis at $n-1$ we have

$$W_n<\frac{n(\sqrt{n-1}+1-\varphi(n))}{n+\sqrt{n-1}-\varphi(n)}.$$

If the inequality is not satisfied at n then, a similar chain of inequalities yields

$$n-\sqrt{n}+\varphi(n)^2(\sqrt{n}+\sqrt{n-1})\leq\varphi(n)(2n-1+2\sqrt{n-1}\sqrt{n}). \quad (19)$$

Note that the above restriction is exactly the same as (18) with the inequality reversed and changing $\varphi(n)$ instead of $1/4$. Using that $\sqrt{n}>\sqrt{n-1}$, the last inequality implies

$$n-\sqrt{n}+2\sqrt{n-1}\varphi(n)^2-(4n-1)\varphi(n)\leq 0.$$

In particular, this means that $\varphi(n)\leq u$ for u the smaller root of the above quadratic equation,

$$u:=\frac{4n-1-\sqrt{16n^2-8n+1-8n\sqrt{n-1}+8\sqrt{n-1}\sqrt{n}}}{4\sqrt{n-1}},$$

but with this definition and (16) it is easy to see that

$$u>\frac{4n-1-\sqrt{16n^2+1-8n\sqrt{n-1}}}{4\sqrt{n-1}}=\varphi(n),$$

which is again a contradiction and this completes the proof for the upper bound.

We have just proved that $\sqrt{n}-\varphi(n)>W_n>\sqrt{n}-1/4$. In particular, this implies that $\varphi(n)<1/4$. Taking the limit in the numerator of expression (16) it is easy to see that the

leading order is

$$\lim_{n\to\infty} 4n-1-\sqrt{16n^2+1-8n\sqrt{n-1}} = \lim_{n\to\infty} \frac{(4n-1)^2-(16n^2+1-8n\sqrt{n-1})}{4n-1+\sqrt{16n^2+1-8n\sqrt{n-1}}} = \lim_{n\to\infty} \sqrt{n-1},$$

which shows that

$$\lim_{n\to\infty} \varphi(n) = \frac{1}{4} \tag{20}$$

and the proof is complete. □

Note that, for large communities, a very good approximation for the total biomass in a perfectly hierarchical tree is given by the formula $W_n = \sqrt{n} - \frac{1}{4}$.

The recursions in (15) for individual abundances can be easily solved in terms of total biomass W_n as

$$x_n^i = \prod_{j=i}^{n} \frac{j}{j+W_{j-1}}. \tag{21}$$

This formula gives the abundance of the i-th species (in increasing order of the tips) for $i \geq 2$ (observe that the first two species have the same abundance). Alternatively,

$$\log(x_n^i) = \sum_{j=i}^{n} \log\left(\frac{j}{j+W_{j-1}}\right) = -\sum_{j=i}^{n} \log\left(1+\frac{W_{j-1}}{j}\right).$$

Approximating W_{j-1} by its lower bound, $W_{j-1} \approx \sqrt{j-1} - 1/4$, we find

$$\log(x_n^i) \approx -\sum_{j=i}^{n} \log\left(1+\frac{\sqrt{j-1}-1/4}{j}\right). \tag{22}$$

Cutting the series for $\log(1+x)$ at second order and considering only the leading term, with respect to j for the quadratic term, we get:

$$\log(x_n^k) \approx -\sum_{j=k}^{n} \frac{\sqrt{j-1}}{j} - \frac{1}{4j} - \frac{1}{2}\frac{j-1}{j^2} \approx -\sum_{j=k}^{n} \frac{1}{\sqrt{j}} - \frac{3}{4j}. \tag{23}$$

By the Euler-Maclaurin formula we obtain:

$$\log(x_n^k) \approx 2(\sqrt{n} - \sqrt{j-1}) + \frac{3}{4}(\log(n) - \log(j-1)). \tag{24}$$

and we can further refine the first terms x_n^k for k small by replacing the actual value W_j.

Perfectly balanced tree

The total biomass for perfectly balanced trees is easier to derive because the covariance matrix has constant row sums in that case. To show this statement, order tree splits by the time they happen ($t_1 < \ldots < t_q$). At each time t_i, the number of lineages doubles, so we get a total of $n = 2^q$ species. As species interact by their shared evolutionary time, in this case each species shares the time with 2^{q-k} other species. Now let $s_k = \sum_{i=1}^{k} \Delta t_i$, Δt_i being the inter-branching time —compare the different notation for s_k here and in the previous subsection. Summing over all possible split times we get the sum over any row of A (observe that $A_{ii} = 1$),

$$r_q = \sum_{j=1}^{n} A_{ij} = 1 + \sum_{k=1}^{q} 2^{q-k} s_k, \tag{25}$$

which is independent of i. Because row sums are constant, the vector or equilibrium abundances can be written as $\boldsymbol{x}_n = x\mathbf{1}_n$, and substitution into $\Sigma_n \boldsymbol{x}_n = \mathbf{1}_n$ yields $r_q x = 1$. Therefore, individual abundances at equilibrium are constant and given by $x = r_q^{-1}$. Consequently, the total biomass at equilibrium, W_q, is simply given by

$$W_q = \frac{2^q}{1 + \sum_{k=1}^{q} 2^{q-k} s_k}. \tag{26}$$

By our assumption of ultrametric trees, we have $s_k < 1$ (we need to add the tip lengths to sum up to one). In the particular case of equal inter-branching times, $\Delta t_i = \frac{1}{q+1}$, then $s_k = \frac{k}{q+1}$ and

$$r_q = 1 + \frac{2^{q-1}}{q+1} \sum_{k=1}^{q} \frac{k}{2^{k-1}}. \tag{27}$$

Observe that

$$\sum_{k=1}^{q} \frac{k}{2^{k-1}} = \frac{\partial}{\partial x} \left(\frac{1 - x^{q+1}}{1 - x} \right) \Bigg|_{x=\frac{1}{2}} = 4 \left(1 - \frac{1}{2^q} \left(q + 1 - \frac{q}{2} \right) \right). \tag{28}$$

Thus,

$$r_q = 1 + \frac{2^{q+1} - q - 2}{q+1} = \frac{2^{q+1} - 1}{q+1}, \tag{29}$$

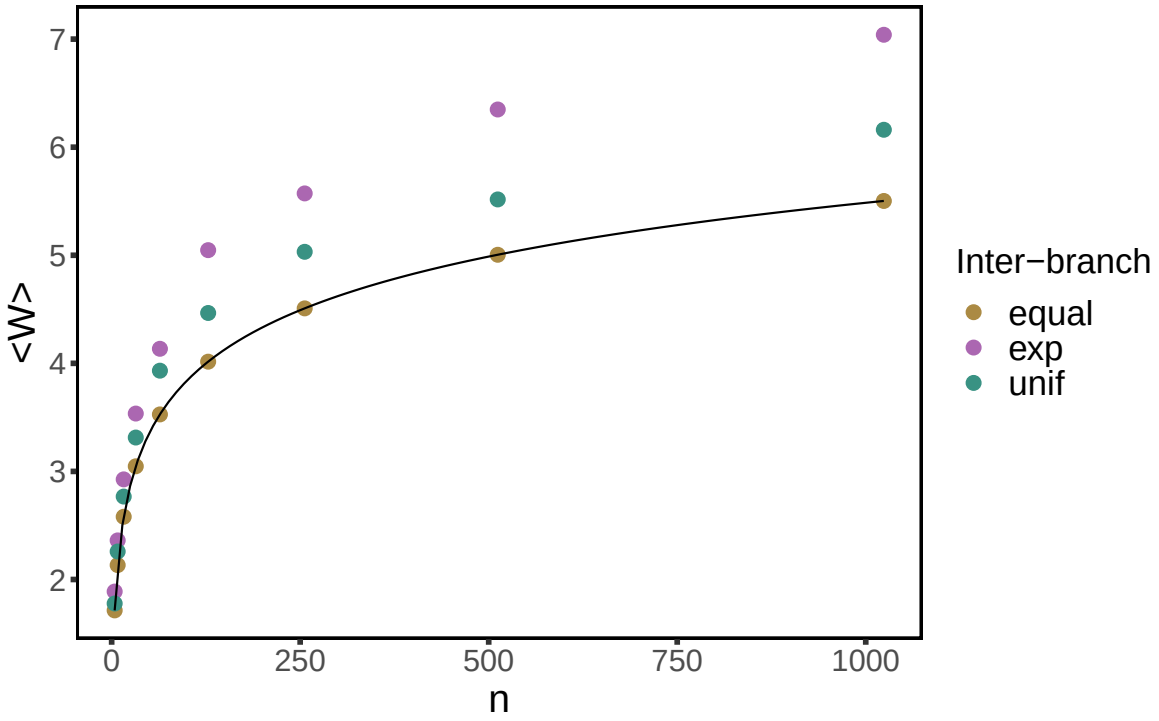

Figure 9: **Total biomass for the perfectly balanced tree**. Dots mark the average values over simulations when sampling branch lengths from an exponential distribution with rate 1, a uniform $[0, 1]$ distribution, and the case of equal branch lengths, for which the analytical prediction (31) is shown with a solid line.

and the total biomass reads

$$W_q = \frac{q+1}{2-2^{-q}}. \tag{30}$$

Let $n = 2^q$ be the number of species, then the number of tree splits is $q = \log_2(n)$. In terms of the number of species, the formula is given by

$$W_n = \frac{\log_2(n)+1}{2-1/n}, \tag{31}$$

which grows logarithmically with n. Fig. 9 compares the case of perfectly balanced tress for equal branching times with two cases, in which sampling times are drawn from exponential and uniform distributions.

3 Number of coexisting species

We have shown above that, in the $\ell \to \infty$ limit, full coexistence is guaranteed. To study species coexistence for finite $\ell \geq n$ we use the fact that A follows the Wishart distribution. As in [33], first we will compute the probability of the equilibrium point being feasible, i.e., where all species survive. Second, since the attractor is unique (it is the only saturated equilibrium point that appears), we can calculate the probability that the equilibrium point cannot be invaded by the remaining species in the pool. Then we will show that the probability of feasibility and non-invasibility factors into the corresponding product, which yields the distribution of the number of species that coexist, as well as the expected number of species that survive.

Because matrix $A = GG^T$ is symmetric and positive definite, it is diagonally-stable [19], which implies that generalized Lotka-Volterra dynamics exhibits a single, globally stable fixed point [19], so there is a unique endpoint for the dynamics. Let us write the equilibrium abundances of the attractor, formed by m survivors, as

$$\boldsymbol{x}_n = \begin{pmatrix} \boldsymbol{x}_m \\ \boldsymbol{0}_{n-m} \end{pmatrix}, \tag{32}$$

where, without loss of generality, we have located the survivors as the first m species. Let $\{S\}_m$ denote the set of species that survive (i.e., the support of the endpoint). Therefore, the attractor can be fully characterized by two conditions [33]:

- Define the vector $\boldsymbol{z}_n = \boldsymbol{1}_n - A\boldsymbol{x}_n = (x_n^i)$ with components z_n^i. Then it holds: first, $z_n^i = 0$ for all species $i \in \{S\}_m$, which simply states that equilibrium abundances of survivors satisfy the linear system $A_m \boldsymbol{x}_m = \boldsymbol{1}_m$, for A_m the submatrix of A restricted to the support $\{S\}_m$. Second, it also holds that $z_n^i < 0$ for all species $i \notin \{S\}_m$, i.e., the fixed point *cannot be invaded* by the remaining species outside the endpoint. We have, therefore, a fixed point that cannot be invaded.

- The equilibrium point hast to be *feasible*, i.e., $\boldsymbol{x}_m > \boldsymbol{0}_m$ —here we use the notation that vectors $\boldsymbol{a} > \boldsymbol{b}$ if all inequalities are satisfied component-wise.

Since matrix A belongs to the Wishart ensemble, these two conditions are to be understood in statistical terms. In the following subsections we are going to compute exact formulae for the

probability that all the species in the pool form a *feasible* attractor, and the probability that an endpoint formed by m species remains *non-invasible*. Using the properties of the Wishart ensemble [30], we will calculate separately the probabilities of feasibility and non-invasibility, and with them we will obtain the distribution of the number of species that survive.

Probability of feasibility

Let n be the number of species in the community and ℓ the number of traits, and define $\gamma := \ell/n$ as the ratio between the number of traits and the size of the pool. An equilibrium point for the system such that all species coexist satisfies:

$$A\boldsymbol{x}_n = \mathbf{1}_n, \text{ with } x_n^i > 0 \text{ for all } i = 1, \dots, n. \tag{33}$$

The probability of feasibility is then the probability that $A^{-1}\mathbf{1}_n$ has all entries greater than 0. Observe that interaction matrix is defined as $A = \frac{1}{\ell}GG^T$ in the main text. Since rescaling by a positive constant in A does not affect the condition for feasibility, we can forget about the rescaling by the number of traits ℓ.

Let $A \sim \mathcal{W}_n(\Sigma, \ell)$ and $L_{n-1} = (I_{n-1}, \mathbf{0}_{n-1})$ be a rectangular $(n-1) \times n$ matrix with 0 in its last column, I_k being the $k \times k$ identity matrix. Then equation (2) of [22] (similarly stated in the proof of Theorem 1 in [8]) implies that

$$\widetilde{\boldsymbol{x}} := \frac{L_{n-1}A^{-1}\mathbf{1}_n}{\mathbf{1}_n^T A^{-1}\mathbf{1}_n} \sim t_{n-1}\left(\ell - n + 2, \frac{L_\ell \Sigma^{-1}\mathbf{1}_n}{\mathbf{1}_n^T\Sigma^{-1}\mathbf{1}_n}, \frac{L_{n-1}R_1L_{n-1}^T}{(\ell-n+2)\mathbf{1}_n^T\Sigma^{-1}\mathbf{1}_n}\right), \tag{34}$$

where $t_p(\nu, \boldsymbol{\mu}, \Lambda)$ is a multivariate, p-dimensional t distribution with ν degrees of freedom, localization vector $\boldsymbol{\mu}$ and dispersion matrix Λ [34]. Matrix R_1 is given by

$$R_1 = \Sigma^{-1} - \frac{\Sigma^{-1}\mathbf{1}_n\mathbf{1}_n^T\Sigma^{-1}}{\mathbf{1}_n^T\Sigma^{-1}\mathbf{1}_n}. \tag{35}$$

Up to a normalization by a positive constant (which is precisely the total biomass, $\mathbf{1}_n^T A^{-1}\mathbf{1}_n$, given that A is positive definite), vector $\widetilde{\boldsymbol{x}} = (\tilde{x}_i)$ precisely gives the abundances of the *first* $n-1$ species. Moreover, the last (normalized) abundance is expressed as $1 - \mathbf{1}_{n-1}^T\widetilde{\boldsymbol{x}}$, so the

probability of feasibility turns out to be

$$P_{\mathrm{f}}(n) = \int d^{n-1}\widetilde{\boldsymbol{x}} f(\widetilde{\boldsymbol{x}})\Theta(1 - \mathbf{1}^T\widetilde{\boldsymbol{x}}) \prod_{i=1}^{n-1} \Theta(\tilde{x}_i), \tag{36}$$

for $f(\widetilde{\boldsymbol{x}})$ the probability density function of the multivariate t distribution defined in (34).

Because a multivariate t distribution is the ratio between a multivariate Gaussian and the square root of a chi-square distribution, it holds that if $\widetilde{\boldsymbol{x}} \sim t_p(\nu, \boldsymbol{\mu}, \Lambda)$, then we have that $\widetilde{\boldsymbol{x}} = \boldsymbol{y}/\sqrt{u/\nu} + \boldsymbol{\mu}$, where $\boldsymbol{y} \sim \mathcal{N}(\mathbf{0}, \Lambda)$ is a multivariate Gaussian and $u \sim \chi^2_\nu$, which is independent of $\boldsymbol{y}$. Therefore, conditioning on u, we find that $\boldsymbol{y}_u := \widetilde{\boldsymbol{x}}|u \sim \mathcal{N}(\boldsymbol{\mu}, \nu\Lambda/u)$ and we can transform the integral above to get

$$P_{\mathrm{f}}(n) = \int_0^\infty du\, g(\nu, u)\Pr(\boldsymbol{y}_u > \mathbf{0}_{n-1}, \mathbf{1}_{n-1}^T\boldsymbol{y}_u < 1), \tag{37}$$

where $u \sim \chi^2_\nu$, $g(\nu, u)$ is the corresponding pdf with $\nu = \ell - n + 2$, and the random variable $\boldsymbol{y}_u$ is distributed as a multivariate normal,

$$\boldsymbol{y}_u \sim \mathcal{N}\left(\frac{L_{n-1}\Sigma^{-1}\mathbf{1}_n}{\mathbf{1}_n^T\Sigma^{-1}\mathbf{1}_n}, \frac{L_{n-1}R_1L_{n-1}^T}{u\mathbf{1}_n^T\Sigma^{-1}\mathbf{1}_n}\right). \tag{38}$$

In this way, all the dependence in the number of traits ℓ remains included in the chi-square distribution. Eqs. (37) and (38) yield the probability of feasibility for an arbitrary covariance matrix Σ. An explicit calculation of the probability of feasibility amount to evaluating the probability $\Pr(\boldsymbol{y}_u > \mathbf{0}_{n-1}, \mathbf{1}_{n-1}^T\boldsymbol{y}_u < 1)$. This can be done explicitly for the case of constant, non-negative correlation.

Constant, non-negative correlation

Consider the covariance matrix $\Sigma = (1-\rho)I_n + \rho\mathbf{1}_n\mathbf{1}_n^T$ with $\rho \geq 0$. Then (38) simplifies to:

$$y_u \sim \mathcal{N}\left(\frac{1}{n}\mathbf{1}_{n-1}, \frac{1-\rho+n\rho}{un(1-\rho)}\left(I_{n-1} - \frac{1}{n}\mathbf{1}_{n-1}\mathbf{1}_{n-1}^T\right)\right). \tag{39}$$

Let us define

$$\alpha_u := \frac{1-\rho+n\rho}{un(1-\rho)} \text{ and } \beta_u := \frac{\alpha_u}{n}. \tag{40}$$

In this way, the covariance matrix Σ_u in (39) can be expressed as $\Sigma_u = \alpha_u I_{n-1} - \beta_u \mathbf{1}_{n-1}\mathbf{1}_{n-1}^T$. Σ_u has two eigenvalues, α_u and $\alpha_u + (n-1)\beta_u$. The first has multiplicity $n-1$, and the second 1. Hence the determinant follows immediately,

$$|\Sigma_u| = \alpha_u^{n-2}(\alpha_u - (n-1)\beta_u). \tag{41}$$

The inverse can be easily calculated:

$$\Sigma_u^{-1} = \frac{1}{\alpha_u}\left(I + \frac{\beta_u}{\alpha_u - (n-1)\beta_u}\mathbf{1}_{n-1}\mathbf{1}_{n-1}^T\right). \tag{42}$$

Therefore we can write the pdf for the random variable y_u as

$$f_u(\boldsymbol{y}) = Ke^{-\frac{1}{2}\left(\boldsymbol{y}-\frac{1}{n}\mathbf{1}_{n-1}\right)^T\Sigma_u^{-1}\left(\boldsymbol{y}-\frac{1}{n}\mathbf{1}_{n-1}\right)} = Ke^{-\frac{1}{2\alpha_u}\left(\left\|\boldsymbol{y}-\frac{1}{n}\mathbf{1}_{n-1}\right\|^2+\frac{\beta_u}{\alpha_u-(n-1)\beta_u}(\mathbf{1}_{n-1}^T(\boldsymbol{y}-\frac{1}{n}\mathbf{1}_{n-1}))^2\right)} \tag{43}$$

for $K = (2\pi)^{-(n-1)/2}|\Sigma_u|^{-1/2}$. First we have to compute the probability

$$p(u) := \Pr(\boldsymbol{y}_u > \mathbf{0}_{n-1}, \mathbf{1}_{n-1}^T\boldsymbol{y}_u < 1) = \int_{\mathbb{R}^{n-1}} d^{n-1}\boldsymbol{y} f_u(\boldsymbol{y})\Theta(1-\mathbf{1}_{n-1}^T\boldsymbol{y})\prod_{i=1}^{n-1}\Theta(y_i), \tag{44}$$

with $\Theta(x)$ the Heaviside step function, defined as $\Theta(x) = 1$ if $x \geq 0$ and $\Theta(x) = 0$ if $x < 0$. Thus after a change of variables $\boldsymbol{y}' = \boldsymbol{y} - \frac{1}{n}\mathbf{1}_{n-1}$, we have

$$p(u) = K\int_{\mathbb{R}^{n-1}} d^{n-1}\boldsymbol{y}e^{-\frac{1}{2\alpha_u}(\|\boldsymbol{y}\|^2+(\mathbf{1}_{n-1}^T\boldsymbol{y})^2)}\Theta\left(\frac{1}{n}-\mathbf{1}_{n-1}^T\boldsymbol{y}\right)\prod_{i=1}^{n-1}\Theta\left(y_i+\frac{1}{n}\right), \tag{45}$$

where we have omitted primes to ease notation and we have used (40) to see that

$$\frac{\beta_u}{\alpha_u - (n-1)\beta_u} = 1. \tag{46}$$

To simplify the term $(\mathbf{1}_{n-1}^T\boldsymbol{y})^2$ in the exponential, we introduce a Dirac's delta function,

$$p(u) = K\int_{\mathbb{R}^{n-1}} d^{n-1}\boldsymbol{y}\int_{\mathbb{R}} d\omega e^{-\frac{1}{2\alpha_u}(\|\boldsymbol{y}\|^2+\omega^2)}\delta(\omega-\mathbf{1}_{n-1}^T\boldsymbol{y})\Theta\left(\frac{1}{n}-\omega\right)\prod_{i=1}^{n-1}\Theta\left(y_i+\frac{1}{n}\right), \tag{47}$$

and use its integral representation,

$$\delta(\omega - \mathbf{1}_{n-1}^T \boldsymbol{y}) = \frac{1}{2\pi} \int_{\mathbb{R}} d\xi e^{-i\xi(\omega - \mathbf{1}_{n-1}^T \boldsymbol{y})}. \tag{48}$$

This transformation, together with an interchange in the order of integration, yields the following expression for $p(u)$:

$$p(u) = \frac{K}{2\pi} \int_{\mathbb{R}} d\omega \int_{\mathbb{R}} d\xi \int_{\mathbb{R}^{n-1}} d^{n-1}\boldsymbol{y} e^{-\frac{1}{2\alpha_u}\left(\|\boldsymbol{y}\|^2 + \omega^2\right) + i(\mathbf{1}_{n-1}^T \boldsymbol{y} - \omega)\xi} \Theta\left(\frac{1}{n} - \omega\right) \prod_{i=1}^{n-1} \Theta\left(y_i + \frac{1}{n}\right). \tag{49}$$

Apparently we are increasing the complexity of the integral, but rearranging terms we observe that

$$p(u) = \frac{K}{2\pi} \int_{\mathbb{R}} d\xi \int_{\mathbb{R}} d\omega e^{-\frac{\omega^2}{2\alpha_u} - i\omega\xi} \Theta\left(\frac{1}{n} - \omega\right) \int_{\mathbb{R}^{n-1}} d^{n-1}\boldsymbol{y} e^{-\frac{\|\boldsymbol{y}\|^2}{2\alpha_u} + i\xi \mathbf{1}_{n-1}^T \boldsymbol{y}} \prod_{i=1}^{n-1} \Theta\left(y_i + \frac{1}{n}\right), \tag{50}$$

and the integral over $\boldsymbol{y}$ factorizes,

$$p(u) = \frac{K}{2\pi} \int_{\mathbb{R}} d\xi \int_{-\infty}^{1/n} d\omega e^{-\frac{\omega^2}{2\alpha_u} - i\omega\xi} \left(\int_{-1/n}^{\infty} dy e^{-\frac{y^2}{2\alpha_u} + iy\xi} \right)^{n-1}. \tag{51}$$

Now, in the integral over ω, change to the variable $\omega' = -\omega$ to get

$$p(u) = \frac{K}{2\pi} \int_{\mathbb{R}} d\xi \int_{-1/n}^{\infty} d\omega e^{-\frac{\omega^2}{2\alpha_u} + i\omega\xi} \left(\int_{-1/n}^{\infty} dy e^{-\frac{y^2}{2\alpha_u} + iy\xi} \right)^{n-1} = \frac{K}{2\pi} \int_{\mathbb{R}} d\xi \left(\int_{-1/n}^{\infty} dy e^{-\frac{y^2}{2\alpha_u} + iy\xi} \right)^{n}. \tag{52}$$

Let

$$\Phi(x) := \frac{1}{2}\left(1 + \operatorname{erf}(x/\sqrt{2})\right) \tag{53}$$

be the cdf of the standard Gaussian distribution, which can be extended to the complex plane. Then it holds that

$$\int_{-1/n}^{\infty} dy e^{-\frac{y^2}{2\alpha_u} + iy\xi} = \sqrt{2\pi\alpha_u}\, e^{-\frac{\alpha_u \xi^2}{2}} \Phi\left(\frac{1/n + i\alpha_u \xi}{\sqrt{\alpha_u}}\right). \tag{54}$$

Therefore, the sought probability can be written as

$$p(u) = \frac{K(2\pi\alpha_u)^{n/2}}{2\pi} \int_{\mathbb{R}} d\xi e^{-\frac{n\alpha_u \xi^2}{2}} \Phi\left(\frac{1/n + i\alpha_u \xi}{\sqrt{\alpha_u}}\right)^n. \tag{55}$$

An alternative way to express the integral over ξ it is to consider a path Γ in the complex plane such that $\Gamma = \{z \in \mathbb{C} | z = x_0 + i\xi\}$ and then reducing the result to the limit $x_0 \to 0$, so that the integral over the imaginary axis is well defined. In practice, this amounts to change to the variable $\zeta = i\xi$. Consequently, an equivalent form of writing this equation is

$$p(u) = -i\sqrt{\frac{n\alpha_u}{2\pi}} \int_{\Gamma} d\zeta e^{\frac{n\alpha_u \zeta^2}{2}} \Phi\left(\frac{1/n + \alpha_u \zeta}{\sqrt{\alpha_u}}\right)^n, \tag{56}$$

where we have used that $K = \sqrt{n}(2\pi\alpha_u)^{-(n-1)/2}$ in this case. Finally, according to (37), in the case of constant, positive correlation the probability of feasibility is given by a two dimensional integral,

$$P_{\rm f}(n) = -i\sqrt{\frac{n}{2\pi}} \int_0^\infty du\, g(\nu, u)\sqrt{\alpha_u} \int_{\Gamma} d\zeta e^{\frac{n\alpha_u \zeta^2}{2}} \Phi\left(\frac{1/n + \alpha_u \zeta}{\sqrt{\alpha_u}}\right)^n, \tag{57}$$

where $g(\nu, u)$ is the pdf of the chi-square distribution with $\nu = \ell - n + 2$ degrees of freedom. Fig. 3 compares this exact formula with numerical simulation for different values of the correlation.

Probability of non-invasibility

In this subsection we compute the probability that an attractor formed by $m \leq n$ species cannot be invaded by the remaining $n - m$ species. Let $A \sim W_n(\Sigma, \ell)$. Observe that for invasibility the rescaling of interaction matrix as $A = \frac{1}{\ell} GG^T$ does not matter. Partition matrices A and Σ in four blocks as follows:

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \tag{58}$$

where Σ_{11} refers to the species that belong to the support $\{S\}_m$ of the attractor, Σ_{22} is related to those species outside the attractor, and off-diagonal matrices are formed by the

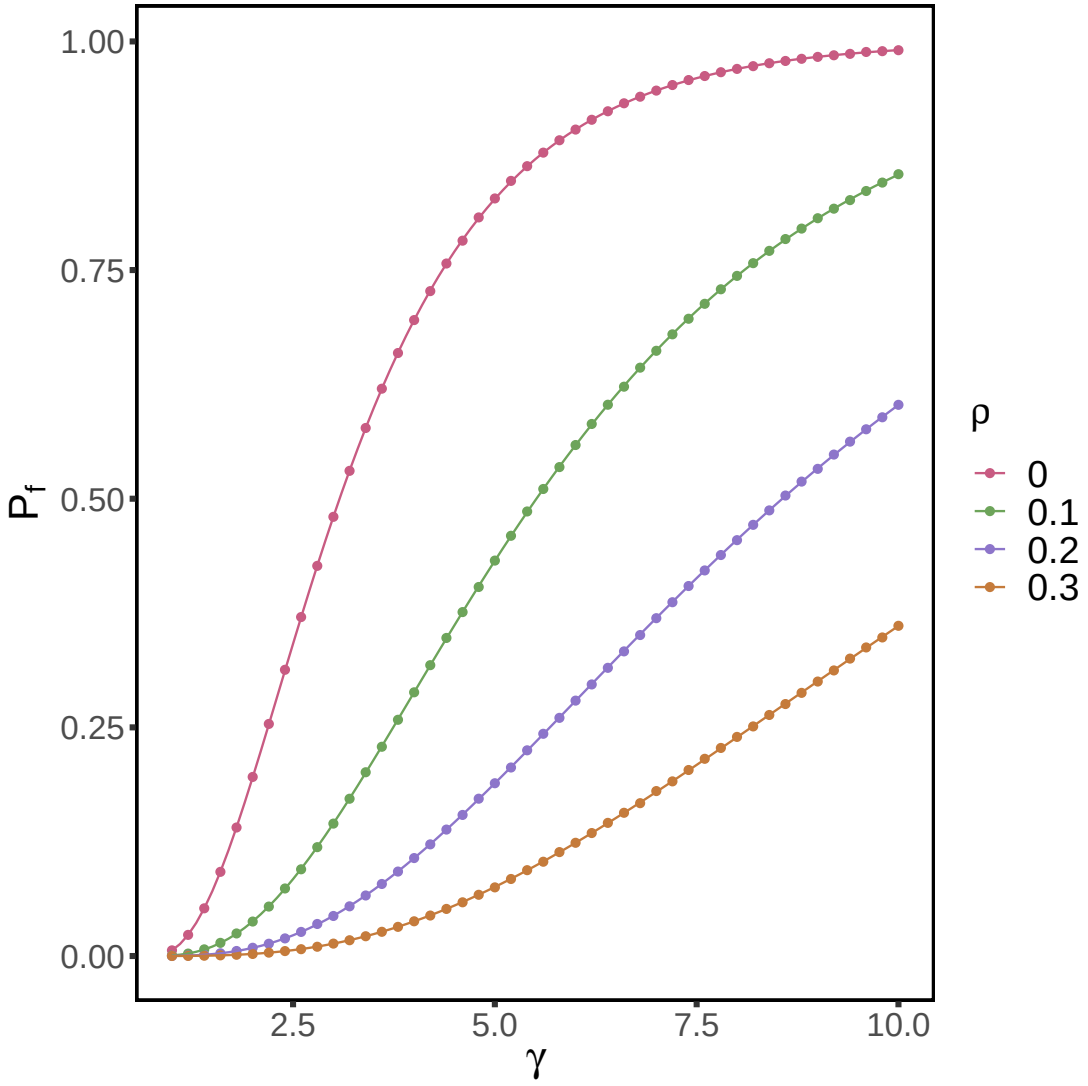

Figure 10: **Probability of feasibility as a function of the ratio γ of number of traits to number of species for different *constant* correlation matrices.** The simulations were done with $n = 10$ species. Dots are simulations, solid lines are numerical evaluations of the exact formula (57). The larger the correlation, the slower curves approach to one in the deterministic limit $\gamma \to \infty$.

corresponding rows and columns in $\{S\}_m$ and $\{S\}_n \setminus \{S\}_m$, and *vice versa*. The exact same notation applies to blocks in A.

Then by theorem 3.2.10 of [30] we have that

$$A_{21}|A_{11} \sim \mathcal{N}(\Sigma_{21}\Sigma_{11}^{-1}A_{11}, \Sigma_{22.1} \otimes A_{11}), \tag{59}$$

where $\Sigma_{22.1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$ is the Schur complement of Σ_{22}, $\otimes$ is the tensor product of matrices, and the normal distribution appearing is meant to be understood as the distribution of the *flatten* matrix A_{21}. By the properties of the normal distribution it follows that

$$\begin{aligned} A_{21}A_{11}^{-1}|A_{11} &\sim \mathcal{N}(\Sigma_{21}\Sigma_{11}^{-1}, \Sigma_{22.1}\otimes A_{11}^{-1}), \\ A_{21}A_{11}^{-1}\mathbf{1}_m|A_{11} &\sim \mathcal{N}(\Sigma_{21}\Sigma_{11}^{-1}\mathbf{1}_m, \mathbf{1}_m^T A_{11}^{-1}\mathbf{1}_m\Sigma_{22.1}). \end{aligned} \tag{60}$$

In order to get the last line, we first transpose the matrix, then notice that the $\mathbf{1}_m^T$ operator acts on the vector of elements of the matrix as $I_m \otimes \mathbf{1}^T$. Hence by the property $(A\otimes B)(C\otimes D) = AC\otimes BD$ of the tensor product the second statement above follows.

As mentioned at the begining of Sec. 3, the probability that the attractor cannot be invaded by any species in $\{S\}_n \setminus \{S\}_m$ coincides with the probability that $\boldsymbol{z} = \mathbf{1}_{n-m} - A_{21}A_{11}^{-1}\mathbf{1}_m < \mathbf{0}_{n-m}$. Define $W := \mathbf{1}_m^T A_{11}^{-1}\mathbf{1}_m$ and $f_W(w)$ as the pdf of the random variable W, which is non-negative. Then

$$\begin{aligned} P_{\rm ni}(m,n) &= \int_0^\infty dw f_W(w)\,\Pr(\boldsymbol{z}<\mathbf{0}|W=w) \\ &= \int_0^\infty dw f_W(w)\int_{\mathcal{V}_w^+} dA_{11}\Pr(A_{11}|W=w)\Pr(\boldsymbol{z}<\mathbf{0}|A_{11},W=w), \end{aligned} \tag{61}$$

where $\mathcal{V}^+$ is the set of positive definite symmetric matrices and $\mathcal{V}_w^+$ the set conditional to $W = \mathbf{1}_m^T A_{11}^{-1}\mathbf{1}_m = w$. Using that $\boldsymbol{z} = \mathbf{1}_{n-m} - A_{21}A_{11}^{-1}\mathbf{1}_m$ and (60), the conditional variable $\boldsymbol{z}|A_{11}, W=w$ is distributed as

$$\boldsymbol{z}|A_{11},W=w \sim \mathcal{N}\left(\mathbf{1}_{n-m} - \Sigma_{21}\Sigma_{11}^{-1}\mathbf{1}_m, w\Sigma_{22.1}\right), \tag{62}$$

which does not depend explicitly on A_{11}. Neither does $\Pr(\boldsymbol{z}<\mathbf{0}|A_{11},W=w)$, so we can factor this probability out of the integration over A_{11}. In this way, we can write

$$P_{\rm ni}(m,n) = \int_0^\infty dw f_W(w) Q_{n-m}^-\big(\mathbf{1}_{n-m} - \Sigma_{21}\Sigma_{11}^{-1}\mathbf{1}_m, w\Sigma_{22.1}\big), \tag{63}$$

because $\int_{\mathcal{V}_w^+} dA_{11}\Pr(A_{11}|W=w) = 1$. In (63) we have defined Q_p^- as the probability that a multivariate Gaussian variable with the specified parameters is contained in the fully negative

orthant,

$$Q_p^-\left(\boldsymbol{\mu},\Lambda\right) := (2\pi)^{-p/2}|\Lambda|^{-1/2}\int_{\mathbb{R}_-^n} d\boldsymbol{y} e^{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu})^T\Lambda^{-1}(\boldsymbol{y}-\boldsymbol{\mu})}. \tag{64}$$

Corollary 3.2.6 in [30] implies that $A_{11} \sim \mathcal{W}_m(\Sigma,\ell)$. Therefore, theorem 3.2.12 in the same reference holds, which ensures that

$$W^{-1}\mathbf{1}_m^T\Sigma_{11}^{-1}\mathbf{1}_m = \frac{\mathbf{1}_m^T\Sigma_{11}^{-1}\mathbf{1}_m}{\mathbf{1}_m^TA_{11}^{-1}\mathbf{1}_m} \sim \chi^2_{\ell-m+1}. \tag{65}$$

This means that

$$g(\nu',w) = -w^{-2}\mathbf{1}_m^T\Sigma_{11}^{-1}\mathbf{1}_m f_W\left(w^{-1}\mathbf{1}_m^T\Sigma_{11}^{-1}\mathbf{1}_m\right), \tag{66}$$

for $g(\nu,w)$ the pdf of a $\chi^2_{\nu'}$ distribution with $\nu' = \ell - m + 1$ degrees of freedom. Now, making the change of variable $w' = w^{-1}\mathbf{1}_m^T\Sigma_{11}^{-1}\mathbf{1}_m$ in (63) we finally get

$$P_{\text{ni}}(m,n) = \int_0^\infty dw g(\nu',w) Q^-_{n-m}\left(\mathbf{1}_{n-m} - \Sigma_{21}\Sigma_{11}^{-1}\mathbf{1}_m, w^{-1}\mathbf{1}_m^T\Sigma_{11}^{-1}\mathbf{1}_m\Sigma_{22.1}\right). \tag{67}$$

As for the case of feasibility, (67) is an exact formula for the probability that an endpoint composed by m species cannot be invaded by the remaining $n-m$ species. Similarly, the multidimensional integral associated to Q^-_{n-m} can be reduced to a single integral in the case of constant, non-negative correlation, as we show in the following subsection. Thus, in that particular case, the probability of non-invasibility is expressed as a double integral.

Constant, non-negative correlation

In the case of constant, non-egative correlation, (67) simplifies to:

$$P_{\text{ni}}(m) = \int_0^\infty dw g(\nu',w) Q^-_{n-m}(\boldsymbol{\mu},\Sigma_w) \tag{68}$$

with

$$\begin{aligned}\boldsymbol{\mu} &= \frac{1-\rho}{1-\rho+m\rho}\mathbf{1}_{n-m},\\ \Sigma_w &= \frac{m(1-\rho)}{w(1-\rho+m\rho)}\left(I_{n-m} + \frac{\rho}{1-\rho+m\rho}\mathbf{1}_{n-m}\mathbf{1}_{n-m}^T\right).\end{aligned} \tag{69}$$

Now focus on the probability Q^-_{n-m}. Making the substitution $\boldsymbol{y}' = k\boldsymbol{y}$ in (64) it is easy to show that

$$Q^-_p(\boldsymbol{\mu}, \Lambda) = Q^-_p(\boldsymbol{\mu}/k, \Lambda/k^2). \tag{70}$$

Therefore, for $k = \frac{m(1-\rho)}{1-\rho+m\rho}$ we recover Eq. (84) with $\boldsymbol{\mu}$ and Λ given by

$$\boldsymbol{\mu} = \frac{1}{m}\mathbf{1}_{n-m}, \quad \Sigma_w = \frac{1-\rho+m\rho}{mw(1-\rho)}\left(I_{n-m} + \frac{\rho}{1-\rho+m\rho}\mathbf{1}_{n-m}\mathbf{1}^T_{n-m}\right). \tag{71}$$

Now let us write $\Sigma_w := \alpha_w I_{n-m} + \beta_w \mathbf{1}_{n-m}\mathbf{1}^T_{n-m}$, with $\alpha_w := \frac{1-\rho+m\rho}{mw(1-\rho)}$, $\beta_w := \frac{\rho\alpha_w}{1-\rho+m\rho}$. As we did for the probability of feasibility, the probability Q^-_{n-m} can be written as a one-dimensional integral. For that is crucial that, contrary to what happened in the case of feasibility, correlations given by Σ_w are positive —notice the plus sign in (71). This is due to the special structure of Σ_w, which implies that the correlation between any two distinct y_i, y_j in (64) is constant and given by $\lambda = \frac{\rho}{1+m\rho} \geq 0$. Hence, the following result of Tong [34] (section 8.2.5) applies:

Proposition 2. *Let $\boldsymbol{x}$ be distributed according to $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$ such that covariance matrix entries satisfy $\Sigma_{ii} = \sigma_i^2$ and $\Sigma_{ij} = \sigma_i\sigma_j\lambda$. Then, the joint probability that $\boldsymbol{x} \in C := \{\boldsymbol{x} \in \mathbb{R}^n | b_i \leq x_i \leq a_i, i = 1, \ldots, n\}$, where $-\infty \leq b_i < a_i \leq \infty$ for i=1,..., n, is expressed as*

$$\Pr(\boldsymbol{x} \in C) = \int_{-\infty}^{\infty} dy\phi(y) \prod_{i=1}^{n}\left[\Phi\left(\frac{(a_i-\mu_i)/\sigma_i + \sqrt{\lambda}y}{\sqrt{1-\lambda}}\right) - \Phi\left(\frac{(b_i-\mu_i)/\sigma_i + \sqrt{\lambda}y}{\sqrt{1-\lambda}}\right)\right] \tag{72}$$

for $\phi(z)$ and $\Phi(z)$ the pdf and cdf, respectively, of a univariate standard normal distribution.

In our particular case $\sigma_i^2 = \frac{1+m\rho}{wm(1-\rho)}$, $\lambda = \frac{\rho}{1+m\rho}$, $b_i = -\infty$, $a_i = 0$ and, according to (71), $\mu_i = \frac{1}{m}$ for $i = 1, \ldots, n-m$. Therefore, putting all the pieces together, we can write

$$P_{\text{ni}}(m, n) = \int_0^{\infty} dw g(\nu', w) \int_{-\infty}^{\infty} dy\phi(y)\Phi\left(\frac{-1/m + y\sqrt{\beta_w}}{\sqrt{\alpha_w}}\right)^{n-m}. \tag{73}$$

As for the probability of feasibility, in the case of constant, non-negative correlation we can reduce it to a two-dimensional integral.

Notice the resemblance between the expressions for feasibility and non-invasibility — Eqs. (57) and (73). In the case of $\rho > 0$, by changing $y \to y'\frac{\alpha_w}{\sqrt{\beta_w}}$, we can make the resemblance

stronger:

$$P_{\rm ni}(m,n) = \sqrt{\frac{1-\rho+m\rho}{2\pi\rho}} \int_0^\infty dw g(\nu', w)\sqrt{\alpha_w} \int_{-\infty}^\infty dy e^{-\frac{(1-\rho+m\rho)\alpha_w y^2}{2\rho}} \Phi\left(\frac{-1/m + y\alpha_w}{\sqrt{\alpha_w}}\right)^{n-m}. \tag{74}$$

Observe that the number of degrees of freedom of the $\chi^2_{\nu'}$ distribution here is $\nu' = \ell - m + 1$. Notice also that the change of variables leading to (74) does not apply for $\rho = 0$. This case is trivial, however, and will not be discussed explicitly.

Sign independence of feasibility and invasibility

In this section we show that the joint probability of feasibility and non-invasibility factors into the product of the two probabilities calculated above. For that purpose, it suffices to show that

$$\Pr\big(\boldsymbol{z} < \mathbf{0}_{n-m} | A_{11}^{-1}\mathbf{1}_m > \mathbf{0}_m\big) = \Pr(\boldsymbol{z} < \mathbf{0}_{n-m}). \tag{75}$$

For that purpose we can calculate

$$\begin{aligned}\Pr\big(\boldsymbol{z} < \mathbf{0}_{n-m} | A_{11}^{-1}\mathbf{1}_m > \mathbf{0}_m\big) &= \int_0^\infty dw\, g_W(w) \Pr\big(\boldsymbol{z} < \mathbf{0}_{n-m} | A_{11}^{-1}\mathbf{1}_m > \mathbf{0}_m, W = w\big) \\ &= \int_0^\infty dw\, g_W(w) \int_{\mathcal{G}_w^+} dA_{11} \Pr\big(\boldsymbol{z} < \mathbf{0}_{n-m} | A_{11}, W = w\big) \Pr\big(A_{11} | A_{11}^{-1}\mathbf{1}_m > \mathbf{0}_m, W = w\big), \end{aligned} \tag{76}$$

where $W = \mathbf{1}_m^T A_{11}^{-1}\mathbf{1}_m$ as for the calculation of $P_{\rm ni}$, and g_W is the pdf of the random variable $W|A_{11}^{-1}\mathbf{1}_m > \mathbf{0}_m$. In the second line we have introduced an integral over the set $\mathcal{G}_w^+$ of symmetric matrices and positive definite that verify the conditions $A_{11}^{-1}\mathbf{1}_m > \mathbf{0}_m$ and $W = \mathbf{1}_m^T A_{11}^{-1}\mathbf{1}_m = w$. As before, by (62) we can factor the probability $\Pr\big(\boldsymbol{z} < \mathbf{0}_{n-m} | A_{11}, W = w\big)$ out, so we get

$$\Pr\big(\boldsymbol{z} < \mathbf{0}_{n-m} | A_{11}^{-1}\mathbf{1}_m > \mathbf{0}_m\big) = \int_0^\infty dw g_W(w) Q_{n-m}^-\big(\mathbf{1}_{n-m} - \Sigma_{21}\Sigma_{11}^{-1}\mathbf{1}_m, w\Sigma_{22.1}\big), \tag{77}$$

which coincides with (67) except for the probability density g_W. In the last step we have used the normalization condition $\int_{\mathcal{G}_w^+} dA_{11} \Pr(A_{11} | A_{11}^{-1}\mathbf{1}_m > \mathbf{0}_m, W = w) = 1$.

Observe that the condition $A_{11}^{-1}\mathbf{1}_m > \mathbf{0}_m$ is equivalent to the conditions $\mathbf{1}_{m-1}^T \widetilde{\boldsymbol{x}} < 1$ and $\widetilde{\boldsymbol{x}} > \mathbf{0}_{m-1}$, for $\widetilde{\boldsymbol{x}}$ the vector of the first $m-1$ relative abundances defined in (34). Let

$R := \{\boldsymbol{v} \in \mathbb{R}^{m-1} | \mathbf{1}_{m-1}^T \boldsymbol{v} < 1, \boldsymbol{v} > \mathbf{0}_{m-1}\}$ the set of vectors satisfying the two last conditions. Then it is easy to see that

$$g_W(w) = \frac{d}{dw}\Pr\big(W < w | A_{11}^{-1}\mathbf{1}_m > \mathbf{0}_m\big) \\ = \frac{d}{dw}\Pr\big(W < w | \widetilde{\boldsymbol{x}} \in R\big) = \frac{d}{dw}\Pr(W < z) = f_W(w). \quad (78)$$

The last equality in the chain above follows because W and $\widetilde{\boldsymbol{x}}$ are independent random variables —see the proof of theorem 1 in [8].

This shows that the probability of observing and endpoint with m survivors can be factored as the probability of feasibility (37) times the probability (67) that the attractor cannot be invaded by the remaining $n - m$ species in the pool.

Distribution of the number of coexisting species

Due to the independence shown in the previous section, the probability that the system settles in a subset $\{S\}_m \subset \{1, \ldots, n\}$ formed by m species is simply

$$\Pr(\{S\}_m | n, \ell, \Sigma) = \binom{n}{m} P_{\mathrm{a}}(m, n) = \binom{n}{m} P_{\mathrm{f}}(m) P_{\mathrm{ni}}(m, n), \quad (79)$$

because all subsets with cardinality m are statistically equivalent.

Assuming constant and non-negative correlation, in Fig. S5 we compare numerical integration of Eqs. (57) and (73) appearing in (79) with simulations.

Average number of species

In this section we will focus on the case of constant correlation. Our aim is to approximate the integrals for feasibility and invasibility in the large number of species limit by a saddle point technique. With these approximations, we provide an analytical way to compute the probability of coexistence $\Pr(\{S\}_m | n, \ell, \rho)$ —cf. Eq. (79)— as well as an approximation for the average fraction of species

$$\wp(n, \ell, \rho) := \frac{1}{n} \sum_{m=0}^{n} \binom{n}{m} m P_{\mathrm{a}}(m, n). \quad (80)$$

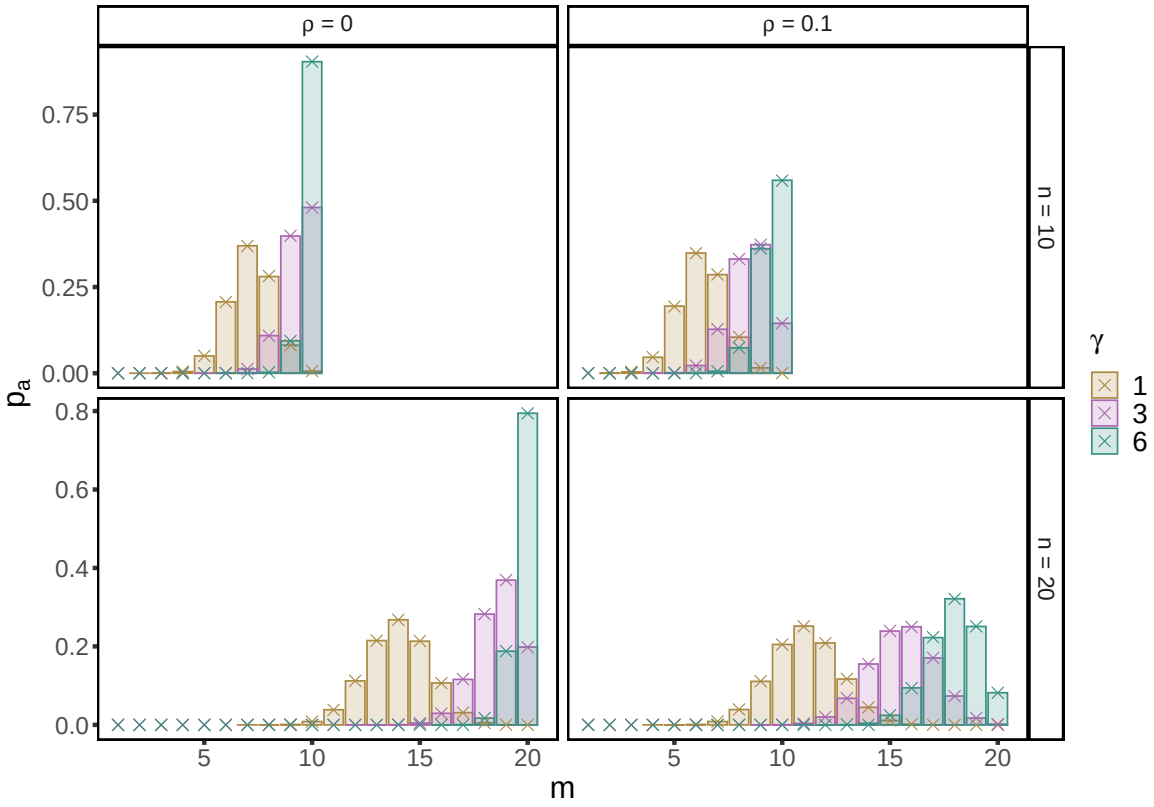

Figure 11: **Distribution of the set of coexisting species as a function of the ratio γ of number of traits to number of species for different *constant* correlation matrices.** The simulations were done with $n = 10$ and 20 species. Bar are simulations, crosses are numerical evaluations of formula (79).

We distinguish the cases $\rho > 0$ and $\rho = 0$ for invasibility. For $\rho > 0$ we use expression (74). Let us define $q := m/n$ as the fraction of survivors, and recall that $\ell = n\gamma$. Also let

$$\lambda_q := mw\alpha_w = 1 + \frac{m\rho}{1-\rho} = 1 + \frac{nq\rho}{1-\rho}. \tag{81}$$

In terms of λ_q, the probability of non-invasibility reads

$$P_{\text{ni}}(m,n) = \frac{\lambda_q}{\sqrt{2\pi(\lambda_q - 1)}} \int_0^\infty dw g(\nu, w) w^{-1/2} \int_{-\infty}^\infty dy e^{-\frac{y^2\lambda_q^2}{2w(\lambda_q-1)}} \Phi\left(-\sqrt{\frac{w}{m\lambda_q}} + y\sqrt{\frac{\lambda_q}{mw}}\right)^{n-m}. \tag{82}$$

Now we make a change of variables,

$$\begin{aligned} w' &= \sqrt{\frac{w}{m}}, \\ \frac{y'}{w'} &= \frac{y}{\sqrt{wm}}. \end{aligned} \tag{83}$$

Then the integral becomes

$$P_{\rm ni}(m,n) = \frac{2\lambda_q}{\sqrt{2\pi(\lambda_q-1)}} \int_0^\infty dw m^{3/2} g(\nu', mw^2) \int_{-\infty}^\infty dy e^{-\frac{my^2\lambda_q^2}{2w^2(\lambda_q-1)}} \Phi\left(-\frac{w}{\sqrt{\lambda_q}} + \frac{y}{w}\sqrt{\lambda_q}\right)^{n-m}. \tag{84}$$

Recall that the probability density function $g(\nu', x)$, for $\nu' = \ell - m + 1$, is:

$$g(\nu, x) = \frac{x^{(\ell-m-1)/2} e^{-x/2}}{2^{(\ell-m+1)/2}\Gamma((\ell-m+1)/2)} \tag{85}$$

Hence the integral (84) is

$$\begin{aligned} P_{\rm ni}(m,n) =& \frac{\lambda_q m}{\sqrt{\pi(\lambda_q-1)}} \frac{(m/2)^{(\ell-m)/2}}{\Gamma((\ell-m+1)/2)} \int_0^\infty dw w^{\ell-m-1} e^{-mw^2/2} \\ &\times \int_{-\infty}^\infty dy e^{-\frac{my^2\lambda_q^2}{2w^2(\lambda_q-1)}} \Phi\left(-\frac{w}{\sqrt{\lambda_q}} + \frac{y}{w}\sqrt{\lambda_q}\right)^{n-m} \\ =& \frac{\lambda_q m}{\sqrt{\pi(\lambda_q-1)}} \frac{(m/2)^{(\ell-m)/2}}{\Gamma((\ell-m+1)/2)} \int_0^\infty dw w^{-1} \int_{-\infty}^\infty dy e^{nF_{\rm ni}(w,y)}, \end{aligned} \tag{86}$$

where the exponent $F_{\rm ni}(w,y)$ has been defined as

$$F_{\rm ni}(w,y) := (\gamma - q)\log(w) - \frac{qw^2}{2} - \frac{qy^2\lambda_q^2}{2w^2(\lambda_q-1)} + (1-q)\log\Phi\left(-\frac{w}{\sqrt{\lambda_q}} + \frac{y}{w}\sqrt{\lambda_q}\right). \tag{87}$$

Now we evaluate the double integral in the limit $n \to \infty$ via a saddle-point technique. For that purpose, since the exponential becomes peaked around the maximum of the exponent, we calculate the equations to be satisfied by the critical point. Taking derivatives of the exponent

we get

$$\begin{aligned}\frac{\partial F_{\rm ni}}{\partial y} &= -\frac{qy\lambda_q^2}{w^2(\lambda_q-1)} + \frac{(1-q)\sqrt{\lambda_q}}{w}\frac{\phi\Big(-\frac{w}{\sqrt{\lambda_q}}+\frac{y}{w}\sqrt{\lambda_q}\Big)}{\Phi\Big(-\frac{w}{\sqrt{\lambda_q}}+\frac{y}{w}\sqrt{\lambda_q}\Big)},\\ \frac{\partial F_{\rm ni}}{\partial w} &= \frac{\gamma-q}{w} - qw + \frac{qy^2\lambda_q^2}{w^3(\lambda_q-1)} - (1-q)\left(\frac{1}{\sqrt{\lambda_q}}+\frac{y\sqrt{\lambda_q}}{w^2}\right)\frac{\phi\Big(-\frac{w}{\sqrt{\lambda_q}}+\frac{y}{w}\sqrt{\lambda_q}\Big)}{\Phi\Big(-\frac{w}{\sqrt{\lambda_q}}+\frac{y}{w}\sqrt{\lambda_q}\Big)}.\end{aligned} \tag{88}$$

Therefore at a critical point $(w^\star, y^\star)$ we have the following conditions:

$$\begin{aligned}-\frac{qy\lambda_q^{3/2}}{w(\lambda_q-1)} + (1-q)\frac{\phi\Big(-\frac{w}{\sqrt{\lambda_q}}+\frac{y}{w}\sqrt{\lambda_q}\Big)}{\Phi\Big(-\frac{w}{\sqrt{\lambda_q}}+\frac{y}{w}\sqrt{\lambda_q}\Big)} &= 0,\\ \gamma - q - qw^2 - \frac{qy\lambda_q}{\lambda_q-1} &= 0.\end{aligned} \tag{89}$$

Similarly we can rewrite the integral for the probability that an endpoint formed by m species is feasible, see Eq. (57), as

$$P_{\rm f}(m) = -i\sqrt{\frac{\lambda_q}{2\pi}}\int_0^\infty du g(\nu,u)u^{-1/2}\int_\Gamma d\zeta e^{\frac{\lambda_q\zeta^2}{2u}}\Phi\left(\sqrt{\frac{u}{m\lambda_q}}+\zeta\sqrt{\frac{\lambda_q}{mu}}\right)^m, \tag{90}$$

where now the number of degrees of freedom is $\nu = \ell - m + 2$.

Following essentially the same procedure as before, i.e. making a change of variables and replacing the density function for the χ^2_ν distribution we get

$$P_{\rm f}(m) = -im^{3/2}\sqrt{\frac{\lambda_q}{2\pi}}\frac{(m/2)^{(\ell-m)/2}}{\Gamma((\ell-m)/2+1)}\int_{-\infty}^\infty du\int_\Gamma d\zeta e^{nF_{\rm f}(u,\zeta)}, \tag{91}$$

with the exponent

$$F_{\rm f}(u,\zeta) := (\gamma-q)\log(u) - \frac{qu^2}{2} + \frac{q\lambda_q\zeta^2}{2u^2} + q\log\Phi\left(\frac{u}{\sqrt{\lambda_q}}+\frac{\zeta}{u}\sqrt{\lambda_q}\right). \tag{92}$$

Similarly, the conditions satisfied by the critical point $(u^\star, \zeta^\star)$ are

$$\begin{aligned} \frac{\zeta\sqrt{\lambda_q}}{u} + \frac{\phi\Big(\frac{u}{\sqrt{\lambda_q}} + \frac{\zeta}{u}\sqrt{\lambda_q}\Big)}{\Phi\Big(\frac{u}{\sqrt{\lambda_q}} + \frac{\zeta}{u}\sqrt{\lambda_q}\Big)} &= 0, \\ \gamma - q - qu^2 - q\zeta &= 0. \end{aligned} \tag{93}$$

Notice that the product of the densities of the χ^2 distributions in each integral —Eqs. (86) and (91)— introduce an extra term which scales exponentially with $m = nq$, namely

$$\frac{m^{\ell-m}}{2^{\ell-m}\Gamma((\ell-m)/2+1)\Gamma((\ell-m)/2+1/2)} = \frac{m^{\ell-m}}{\Gamma(\ell-m+1)}. \tag{94}$$

Using the Stirling's asymptotic form of the gamma function we get

$$\frac{m^{\ell-m}}{\Gamma(\ell-m+1)} \sim \frac{e^{n(\gamma-q)(1+\log q - \log(\gamma-q))}}{\sqrt{2\pi n(\gamma-q)}}. \tag{95}$$

Let

$$F_{\mathrm{e}}(q) := (\gamma-q)(1+\log q - \log(\gamma-q)) \tag{96}$$

and

$$F_{\mathrm{c}}(q) := -q\log q - (1-q)\log(1-q), \tag{97}$$

$F_{\mathrm{c}}(q)$ being the exponent appearing in Stirling's asymptotic formula for the binomial coefficient $\binom{n}{nq}$. Consequentely the probability that the system settles in an endpoint with $m = nq$ species is given, up to a normalization factor, by:

$$\Pr(\{S\}_m|n,\ell,\rho) = \binom{n}{m} P_{\mathrm{a}}(m,n) \sim \exp\{n(F_{\mathrm{f}}(u^\star,\zeta^\star,q) + F_{\mathrm{ni}}(w^\star,y^\star,q) + F_{\mathrm{e}}(q) + F_{\mathrm{c}}(q))\}. \tag{98}$$

Observe that critical point coordinates $u^\star$, $\zeta^\star$, $w^\star$ and $y^\star$ depend implicitly on q through (89) and (93). Observe that one can use the asymptotic expansion (98) to obtain numerically the distribution of the number of survivors, $\Pr(\{S\}_m|n,\ell,\rho)$, up to a normalization factor. The calculation amounts to solve numerically the non-linear systems (89) and (93).

We are now ready to provide an analytical approximation for the mean fraction of survivors $\wp$, cf. Eq. (80). In the limit of large pool size n, we can approximate the mean of the

distribution $\Pr(\{S\}_m|m,\ell,\rho)$ by its mode, which is easier to compute. In fact, to calculate the mode of the distribution q in the large n limit we need to find the $q^\star$ value that maximizes the exponent in (98). Due to the critical point conditions for $(u^\star,\zeta^\star)$ and $(w^\star,y^\star)$, $q^\star$ satisfies

$$\frac{\partial F_{\rm f}}{\partial q}+\frac{\partial F_{\rm ni}}{\partial q}+\frac{\partial F_{\rm e}}{\partial q}+\frac{\partial F_{\rm c}}{\partial q}=0. \tag{99}$$

Evaluated at the critical points $(u^\star,\zeta^\star)$ and $(w^\star,y^\star)$, the derivatives read

$$\begin{aligned}
\frac{\partial F_{\rm ni}}{\partial q}&=-\log(w)-\frac{w^2}{2}-\frac{y^2\lambda_q}{2w^2}+\frac{y}{2}-\log\Phi\left(-\frac{w}{\sqrt{\lambda_q}}+\frac{y}{w}\sqrt{\lambda_q}\right),\\
\frac{\partial F_{\rm f}}{\partial q}&=-\log(u)-\frac{u^2}{2}+\lambda_q\frac{\zeta^2}{2u^2}+\frac{\zeta(\lambda_q-1)}{2\lambda_q}+\log\Phi\left(\frac{u}{\sqrt{\lambda_q}}+\frac{\zeta}{u}\sqrt{\lambda_q}\right),\\
\frac{\partial F_{\rm e}}{\partial q}&=\log\left(\frac{\gamma-q}{q}\right)+\frac{\gamma-q}{q}=\log\left(\frac{\gamma-q}{q}\right)+\frac{u^2}{2}+\frac{w^2}{2}+\frac{q\zeta}{2}+\frac{qy\lambda_q}{2(\lambda_q-1)},\\
\frac{\partial F_{\rm c}}{\partial q}&=\log(1-q)-\log q.
\end{aligned} \tag{100}$$

Therefore the condition for $q^\star$ reduces to

$$-\log\left(\frac{qwu}{\gamma-q}\right)+\frac{\lambda_q}{2}\left(\frac{\zeta^2}{u^2}-\frac{y^2}{w^2}\right)+\frac{2\lambda_q-1}{2}\left(\frac{y}{\lambda_q-1}+\frac{\zeta}{\lambda_q}\right)+\log\frac{(1-q)\Phi\left(\frac{u}{\sqrt{\lambda_q}}+\frac{\zeta}{u}\sqrt{\lambda_q}\right)}{q\Phi\left(-\frac{w}{\sqrt{\lambda_q}}+\frac{y}{w}\sqrt{\lambda_q}\right)}=0. \tag{101}$$

A direct calculation shows that, at $wu=\frac{\gamma-q}{q}$, the terms up to the last logarithm vanish. We now show that the last one can be written as $\left(wu-\frac{\gamma-q}{q}\right)h$ for some function h.

Indeed, using conditions (93) and (89) we have

$$\frac{(1-q)\phi(-w,-y,q)}{q\Phi(-w,-y,q)}-\frac{\phi(u,\zeta,q)}{\Phi(u,\zeta,q)}=\frac{(u+w)\sqrt{\lambda_q}}{uw}\left(\frac{\gamma-q}{q}-uw\right), \tag{102}$$

where we have used the abbreviations $\Phi(u,\zeta,q):=\Phi\left(\frac{u}{\sqrt{\lambda_q}}+\frac{\zeta}{u}\sqrt{\lambda_q}\right)$ and $\phi(u,\zeta,q):=\phi\left(\frac{u}{\sqrt{\lambda_q}}+\frac{\zeta}{u}\sqrt{\lambda_q}\right)$ to simplify notation. Therefore,

$$\frac{(1-q)\Phi(u,\zeta,q)}{q\Phi(-w,-y,q)}=\frac{\phi(u,\zeta,q)}{\phi(-w,-y,q)}+\frac{(u+w)\Phi(u,\zeta,q)\sqrt{\lambda_q}}{uw\phi(-w,-y,q)}\left(\frac{\gamma-q}{q}-uw\right). \tag{103}$$

Letting $\mu_q := (\gamma - q)/q$, it holds that

$$\frac{\phi(u,\zeta,q)}{\phi(-w,-y,q)} = e^{(\mu_q^2-(uw)^2)((\lambda_q-1)^2u^2-\lambda_q^2w^2)/(2\lambda_q u^2 w^2)}. \tag{104}$$

Now, due to the series representation of the exponential function we have

$$\frac{\phi(u,\zeta,q)}{\phi(-w,-y,q)} = 1 + (\mu_q - uw)h(u,w), \tag{105}$$

where

$$h(u,w) := \frac{q(u+w)\Phi(u,\zeta,q)\sqrt{\lambda_q}}{uw\phi(-w,-y,q)} + \sum_{j=1}^{\infty} \frac{1}{j!}(\mu_q - uw)^{j-1}\left((\mu_q+uw)\frac{(\lambda_q-1)^2u^2-\lambda_q^2w^2}{2\lambda_q u^2 w^2}\right)^j. \tag{106}$$

Thus, the claim follows by using the series expansion of $\log(1+x)$. Therefore, all the terms in (101) vanish at $uw = \mu_q$.

We have just shown that the last logarithm in (101) is equal to zero. Consequently $q^\star$ satisfies

$$\frac{(1-q)\Phi\left(\frac{u}{\sqrt{\lambda_q}} + \frac{\zeta}{u}\sqrt{\lambda_q}\right)}{q\Phi\left(-\frac{w}{\sqrt{\lambda_q}} + \frac{y}{w}\sqrt{\lambda_q}\right)} = 1. \tag{107}$$

At the point $uw = \mu_q$ we can write

$$\frac{u}{\sqrt{\lambda_q}} + \frac{\zeta}{u}\sqrt{\lambda_q} = \frac{\lambda_q w - (\lambda_q - 1)u}{\sqrt{\lambda_q}} = \frac{w}{\sqrt{\lambda_q}} - \frac{y}{w}\sqrt{\lambda_q}, \tag{108}$$

which in turn implies that

$$\Phi\left(\frac{\lambda_q w - (\lambda_q-1)u}{\sqrt{\lambda_q}}\right) = q^\star. \tag{109}$$

Let $\hat{q} := \Phi^{-1}(q^\star) = \sqrt{2}\mathrm{erf}^{-1}(2q^\star - 1)$, for erf^{-1} the inverse error function. Then it holds that

$(\lambda_q w - (\lambda_q - 1)u)/\sqrt{\lambda_q} = \hat{q}$ and using eq. (93) we can solve for $u^\star, w^\star$ in terms of $\hat{q}$, yielding

$$\begin{aligned} u^\star &= \sqrt{\lambda_q}\left(\frac{\phi(\hat{q})}{q^\star} + \hat{q}\right), \\ w^\star &= \frac{1}{\sqrt{\lambda_q}}\left((\lambda_q - 1)\frac{\phi(\hat{q})}{q^\star} + \lambda_q \hat{q}\right). \end{aligned} \tag{110}$$

The final condition for $q^\star$ at the saddle point reduces to substitute the expressions above into the condition $uw = \mu_q$, which finally reads

$$\frac{\gamma}{q^\star} = 1 + \left(\frac{\phi(\Phi^{-1}(q^\star))}{q^\star} + \Phi^{-1}(q^\star)\right)\left(\frac{\phi(\Phi^{-1}(q^\star))}{q^\star}(\lambda_{q^\star} - 1) + \Phi^{-1}(q^\star)\lambda_{q^\star}\right). \tag{111}$$

The case $\rho = 0$ for invasibility is similar, and simpler.

Level Curves

Eq. (111) gives a very good approximation to the level curves on the (ρ, γ) plane mapping to constant mean fraction of survivors $q = m/n$. This implicit condition can be rewritten equivalently as

$$\gamma = q + \Phi^{-1}(q)H(q) + \frac{n\rho H(q)^2}{1-\rho}, \tag{112}$$

where $H(q) := \phi(\Phi^{-1}(q)) + q\Phi^{-1}(q)$. This condition is compared with simulation results in Fig. 3 of the main text (right panel).

4 Total biomass distribution at endpoints

The proof of independence of invasibility and feasibility (section 3) also shows that, for any fixed size m of a subset of species and total biomass w, we have that $\Pr(\boldsymbol{z}_{n-m} < \mathbf{0}_{n-m} | \boldsymbol{x}_m > \mathbf{0}_m, W = w) = \Pr(\boldsymbol{z}_{n-m} < \mathbf{0}_{n-m} | W = w)$. This remark, together with the independence of W and $\boldsymbol{x}_m > \mathbf{0}_m$ (feasibility), helps us derive the distribution of total biomass. To simplify notation we do not rescale the interaction matrix by ℓ (as shown in section 6 this would amount to a rescaling of total biomass $w \to \ell w$). The cdf for the random variable W is precisely

$$\Pr(W < w) = \sum_{m=0}^{n} \binom{n}{m} P_{\mathrm{a}}(m, n) \Pr(W < w | m), \tag{113}$$

where $\Pr(W < w|m)$ is the probability that $W < w$ conditional on the m-species endpoint is feasible and non-invasible. Thus,

$$\begin{aligned}\Pr(W < w|m) &= \frac{\Pr(W < w, \boldsymbol{x}_m > \mathbf{0}_m, \boldsymbol{z}_{n-m} < \mathbf{0}_{n-m})}{P_{\mathrm{a}}(m,n)} \\ &= \frac{\Pr(W < w, \boldsymbol{z}_{n-m} < \mathbf{0}_{n-m}|\boldsymbol{x}_m > \mathbf{0}_m)P_{\mathrm{f}}(m)}{P_{\mathrm{a}}(m,n)} \\ &= \frac{\Pr(W < w, \boldsymbol{z}_{n-m} < \mathbf{0}_{n-m})P_{\mathrm{f}}(m)}{P_{\mathrm{a}}(m,n)},\end{aligned} \tag{114}$$

the last equality following from the statement in the paragraph above. Now, using the notations introduced in the last section, it holds that

$$\begin{aligned}\Pr(W < w, \boldsymbol{z}_{n-m} < \mathbf{0}_{n-m}) = \int_0^\infty dug(\nu', u)\Theta(u - w^{-1}\mathbf{1}_m^T\Sigma_{11}^{-1}\mathbf{1}_m) \\ \times Q_{n-m}^-(\mathbf{1}_{n-m} - \Sigma_{21}\Sigma_{11}^{-1}\mathbf{1}_m, u^{-1}\mathbf{1}_m^T\Sigma_{11}^{-1}\mathbf{1}_m\Sigma_{22.1}).\end{aligned} \tag{115}$$

Hence, using (114) and $P_{\mathrm{a}}(m,n) = P_{\mathrm{f}}(m)P_{\mathrm{ni}}(m,n)$, the probability density function of the biomass distribution can be expressed as

$$\begin{aligned}g_{\mathrm{a}}(w) &= \sum_{m=0}^n \binom{n}{m} P_{\mathrm{f}}(m)\frac{\partial \Pr(W < w, \boldsymbol{z}_{n-m} < \mathbf{0}_{n-m})}{\partial w} \\ &= \sum_{m=0}^n \binom{n}{m}\frac{\tilde{w}}{w}P_{\mathrm{f}}(m)g(\nu', \tilde{w})Q_{n-m}^-(\mathbf{1}_{n-m} - \Sigma_{21}\Sigma_{11}^{-1}\mathbf{1}_m, \tilde{w}^{-1}\mathbf{1}_m^T\Sigma_{11}^{-1}\mathbf{1}_m\Sigma_{22.1}),\end{aligned} \tag{116}$$

where $\tilde{w} := w^{-1}\mathbf{1}_m^T\Sigma_{11}^{-1}\mathbf{1}_m$. Fig. 12 shows the comparison of (116) with simulations for the constant correlation case in the case in which the interaction matrix is rescaled by the number of traits.

Going back to re-scaling the interaction matrix by ℓ, total biomass transforms as $w \to \ell w$. By the above calculation and a change of variables $w \to \tilde{w}$, the moments of the distribution of ℓW conditional to m coexisting species are given by

$$\begin{aligned}\mathbb{E}[(\ell W)^k|m] = \int_0^\infty dw(\ell w)^k g_{\mathrm{a}}(w|m) = \frac{1}{P_{\mathrm{ni}}(m,n)}\int_0^\infty dwg(\nu', w) \\ \times (\ell w^{-1}\mathbf{1}_m^T\Sigma_{11}^{-1}\mathbf{1}_m)^k Q_{n-m}^-(\mathbf{1}_{n-m} - \Sigma_{21}\Sigma_{11}^{-1}\mathbf{1}_m, w^{-1}\mathbf{1}_m^T\Sigma_{11}^{-1}\mathbf{1}_m\Sigma_{22.1}).\end{aligned} \tag{117}$$

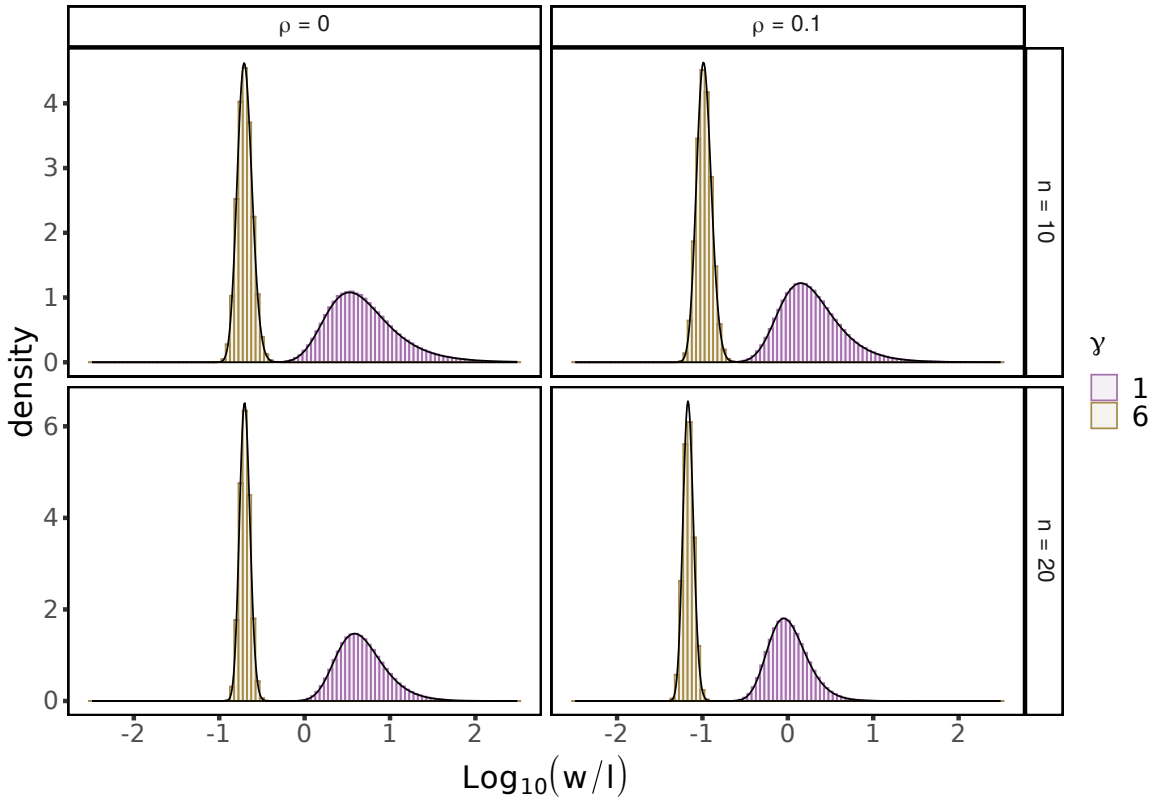

Figure 12: **Distribution of the total biomass w of the survival community as a function of the ratio γ of number of traits k to number of species n for different *constant* correlation matrices.** The simulations were done with $n = 10, 20$ species. Histograms are simulations and black lines are the numerical integration of (116).

By the saddle point calculation done while computing the expected number of survivors we can approximate the mean of $\ell W|m$ for $\rho \geq 0$, $m = nq$ and $\ell = \gamma n$ as follows: the above integral satisfies (86) up to a multiplication by $\frac{\gamma}{w^2 q}\mathbf{1}_m^T \Sigma_{11}^{-1}\mathbf{1}_m$ —observe the reescaling in (83). Hence the exponent in the integral does not change so, at the solution (y_0, w_0) of (89), neglecting all but the leading order terms we can approximate

$$\mathbb{E}[\ell W|m] \approx \frac{\ell}{(1-\rho+\rho m)w_0^2}. \tag{118}$$

Assuming that the distribution of survivors is highly peaked at the mode, we can approximate the mean of W by the mean conditional at the mode, which we get from Eq. (111):

$$\mathbb{E}[\ell W] \approx \frac{\ell}{(1-\rho+\rho q^\star n)w_0(q^\star)^2}. \tag{119}$$

5 Relative abundances

For an equilibrium attractor $\boldsymbol{x}_m$ with m species, let $\boldsymbol{v} := \boldsymbol{x}_m / \sum_{i=1}^m x_m^i$ be the relative abundance vector. In particular, $v_m = 1 - \sum_{i=1}^{m-1} v_i =: 1 - \tilde{v}$. By section 3, Eq. (34), we know that $\tilde{v}$ follows a multivariate t distribution, so we can write the distribution function for v_m conditional on $\boldsymbol{x}$ being feasible as

$$\begin{aligned}\Pr(v_m < c | \boldsymbol{x}_m > \boldsymbol{0}_m) &= 1 - \Pr(v_m > c | \boldsymbol{x} > \boldsymbol{0}_m) \\ &= 1 - \frac{1}{P_{\mathrm{f}}(m)} \int_0^\infty du g(\nu, u) \Pr(\boldsymbol{y}_u > \boldsymbol{0}_{m-1}, \boldsymbol{1}_{m-1}^T \boldsymbol{y}_u < 1 - c) \quad (120)\end{aligned}$$

with $\nu = \ell - m + 2$. The independence of $\tilde{v}$ and invasibility gives us the distribution of v conditional to $\boldsymbol{x}_m$ being an attractor of the system with m out of n survivors. Let $\boldsymbol{z}_{n-m}$ be defined as in section 3. Then

$$\begin{aligned}\Pr(v_m < c | m) &= \frac{\Pr(v_m < c, \boldsymbol{x} > \boldsymbol{0}_m, \boldsymbol{z}_{n-m} < \boldsymbol{0}_{n-m})}{P_{\mathrm{a}}(m,n)} \\ &= \frac{\Pr(\boldsymbol{z}_{n-m} < \boldsymbol{0}_{n-m} | \boldsymbol{x} > \boldsymbol{0}_m, v_m < c) \Pr(v_m < c | \boldsymbol{x} > \boldsymbol{0}_m)}{P_{\mathrm{ni}}(m,n)} = \Pr(v_m < c | \boldsymbol{x} > \boldsymbol{0}_m), \quad (121)\end{aligned}$$

where we have used the independence of feasibility and invasibility, $P_{\mathrm{a}}(m,n) = P_{\mathrm{f}}(m) P_{\mathrm{ni}}(m,n)$.

In case of a constant correlation $\rho \geq 0$, all species are equivalent so any surviving species i has the same distribution as x_m. Applying the same derivation as for the feasibility case, and using the notation of the saddle point calculation with $m = qn$ (see Eq. (90)), we get

$$\begin{aligned}\Pr(v_m < c | m) = 1 - \frac{i\sqrt{\lambda_q}}{\sqrt{2\pi} P_{\mathrm{f}}(m)} \int_0^\infty du g(\nu, u) u^{-1/2} \int_\Gamma d\zeta e^{\frac{\lambda_q \zeta^2}{2u}} \\ \times \Phi\left(\sqrt{\frac{u}{n\lambda_q}} + \zeta\sqrt{\frac{\lambda_q}{nu}}\right)^{m-1} \Phi\left(\sqrt{\frac{u}{n\lambda_q}} - c\sqrt{\frac{nu}{\lambda_q}} + \zeta\sqrt{\frac{\lambda_q}{nu}}\right). \quad (122)\end{aligned}$$

Letting $\tilde{c} = cn$, the integral above can be approximated by the same saddle point calculation we did for feasibility (section 3) up to a multiplication factor given by

$$\frac{\Phi\left(\frac{u}{\sqrt{\lambda_q}}(1 - \tilde{c}q) + \frac{\zeta}{u}\sqrt{\lambda_q}\right)}{\Phi\left(\frac{u}{\sqrt{\lambda_q}} + \frac{\zeta}{u}\sqrt{\lambda_q}\right)}. \quad (123)$$

Thus, for (u, ζ) satisfying the system of equations (93) with ζ real, we get an approximation for the distribution function by neglecting all but the leading terms:

$$\Pr(v_m < c|m) = 1 - \frac{\Phi\left(\frac{u}{\sqrt{\lambda_q}}(1-\tilde{c}q) + \frac{\zeta}{u}\sqrt{\lambda_q}\right)}{\Phi\left(\frac{u}{\sqrt{\lambda_q}} + \frac{\zeta}{u}\sqrt{\lambda_q}\right)}. \tag{124}$$

This distribution was compared to simulations in the main text (Fig 4, left panel).

6 Invariant Lotka-Volterra operations

In this section we detail the operations that can be performed in a symmetric stable GLV system without changing the subset of coexisting species.

Let $\boldsymbol{r} \in \mathbb{R}^n$ be the vector of growth rates, and $A \in \mathbb{R}^n$ a symmetric and positive definite interaction matrix. Let $\{S\}_m \subset \{1, \ldots, n\}$ be the *unique* subset of m species that form the attractor, with vector of densities $\boldsymbol{x} = (x_i)$. Then $\boldsymbol{x}$ satisfies:

$$\begin{cases} x_i > 0, & i \in \{S\}_m, \\ x_i(A\boldsymbol{x} + \boldsymbol{r})_i = 0, & \text{for all i}, \\ (A\boldsymbol{x} + \boldsymbol{r})_i < 0, & i \notin \{S\}_m. \end{cases} \tag{125}$$

Then we can easily see the effect of the following operations on A and $\boldsymbol{r}$ on the attractor $\boldsymbol{x}$. Let $\kappa > 0$ and D a positive diagonal matrix. The operations that maintain the identity of the species in the endpoint are:

(a) $A \to \kappa A$: then $\boldsymbol{x} \to \kappa^{-1}\boldsymbol{x}$.

(b) $\boldsymbol{r} \to \kappa \boldsymbol{r}$: Then $\boldsymbol{x} \to \kappa \boldsymbol{x}$.

(c) $A \to DAD, \boldsymbol{r} \to D\boldsymbol{r}$: Then $\boldsymbol{x} \to D^{-1}\boldsymbol{x}$.

After any of these operations, the set of coexisting species remains *unchanged*.

Additionally, in the case of $\boldsymbol{r} = \kappa \mathbf{1}_n$, for $\kappa > 0$, we can perform an additional operation:

$$A \to B = A + \mu \mathbf{1}_n \mathbf{1}_n^T. \tag{126}$$

Then shifting

$$\boldsymbol{x} \to \boldsymbol{y} = \frac{\kappa \boldsymbol{x}}{1 + \mu \mathbf{1}_n^T \boldsymbol{x}}, \tag{127}$$

by direct computation of conditions (125) we see that $\boldsymbol{y}$ is a non-invasible equilibrium. If we additionally restrict $\mu > 0$, $\boldsymbol{y}$ satisfies the feasibility property and B is positive definite so again the support $\{S\}_m$ of the attractor is unchanged.

7 Varying growth rates

In this section we analyze the effect that growth rates are not equal for all species. By continuity, we expect our results to hold when $\boldsymbol{r} = \mathbf{1}_n + \boldsymbol{\epsilon}_n$ and $\|\boldsymbol{\epsilon}_n\| \ll 1$ if $\ell \geq n$. In case $\ell < n$, the matrix A is singular and the solutions of the system can be unbounded. To correct for that, assume that $A = A + \mu \mathbf{1}_n \mathbf{1}_n^T$ where μ is a sufficiently large enough perturbation so that $A_{ij} + \mu > 0$. In this case $-(A + \mu \mathbf{1}_n \mathbf{1}_n^T)$ is negative semidefinite and dissipative [19], so the solutions are always bounded. Still, the solutions can be degenerate in the sense that there is a hyperplane of non-invasible equilibria towards which the system converges. By perturbing the growth rates we can correct for that. Assume now that $\boldsymbol{r} = \mathbf{1}_n + \mathcal{N}(0, \sigma^2)$, where $\sigma \ll 1$ and that $\hat{\boldsymbol{x}}$ is a saturated rest point of the system (which exists because $A_{ij} + \mu > 0$). Without lost of generality, we can assume that the first m species survive. Then, we have

$$A\hat{\boldsymbol{x}} + \boldsymbol{r} = \begin{pmatrix} \mathbf{0}_m \\ \boldsymbol{z} \end{pmatrix}. \tag{128}$$

For $\boldsymbol{z} \in \mathbb{R}_-^{n-m}$, if any $z_i = 0$, then for the system considering only the species $\{1, \ldots, m\} \cup \{i\}$ we have that the restriction of $\boldsymbol{r}$ to this subsystem is contained on a plane of dimension $m < m + 1$. Since the distribution of $\boldsymbol{r}$ is continuous, the probability of this event is 0 almost surely. Hence $z_i < 0$ for any i so that invasibility is *strict*. Furthermore, the same argument shows that the rank of A restricted to the survivor subset must be m, i.e. the restriction of matrix A to the set of coexisting species is *full rank*.

Apply the usual Lyapunov function for the system [19],

$$V(\boldsymbol{x}) = -\sum_{i=1}^{n} (\hat{x}_i \log x_i - x_i). \tag{129}$$

Defined for any $\boldsymbol{x} \in \mathbb{R}^n_+$, with a global minimum at $\boldsymbol{x} = \hat{\boldsymbol{x}}$ and radially unbounded, then we have

$$\dot{V}(\boldsymbol{x}) = -\sum_{ij} A_{ij}(x_i - \hat{x}_i)(x_j - \hat{x}_j) + \sum_i (x_i - \hat{x}_i)\Big(r_i + \sum_j a_{ij}\hat{x}_j\Big). \tag{130}$$

The first sum is non-negative since the matrix is negative semidefinite, and the second is non-positive and is negative unless $x_i = 0$ for any $i > m$. Given that the restriction of A to the survivors subset is full rank then $\dot{V} = 0$ only at $\hat{\boldsymbol{x}}$, which implies that $\hat{\boldsymbol{x}}$ is globally stable and, in particular, is unique [19].

In these cases, while our previous analyses are not exact because of the perturbations introduced in the vector of rates $\boldsymbol{r}$ and in interaction coefficients ($A \to A + \mu \mathbf{1}_n \mathbf{1}_n^T$), we can apply the same machinery that we have developed to provide approximations. This works because we know that the shift of $A \to A + \mu \mathbf{1}_n \mathbf{1}_n^T$ does not change properties like feasibility or invasibility (see section 6). What changes is that the rank of A goes up by one (see the observation below). Forgetting about this, we can use the same machinery as in the non-degenerate case: for feasibility this follows because only full rank subsets are considered, and the restriction of a singular Wishart to a block of $m \leq \ell$ subsets is a Wishart matrix. Further, the conditional distribution of blocks used for the derivation of the probability of non-invasibility holds in the non-degenerate case too [7].

Observation. The rank of $B = A + \mu \mathbf{1}_n \mathbf{1}_n^T$ is equal to the rank of A plus one. Indeed, let $\boldsymbol{w} \in \ker B$, then $\boldsymbol{w}^T B \boldsymbol{w} = \boldsymbol{w}^T A \boldsymbol{w} + \mu(\mathbf{1}_n^T \boldsymbol{w})^2 = 0$, hence $\boldsymbol{w} \in \ker A \cap \mathbf{1}_n^\perp$, and similarly any $\boldsymbol{w} \in \ker A \cap \mathbf{1}_n^\perp$ is in the kernel of B, hence $\ker B = \ker(A \cap 1^\perp)$. Unless $\ker A \subset \mathbf{1}_n^\perp$, $\dim(\ker B) = \dim(\ker A) - 1$, so the rank increases by one.

Consider then $A = CC^T$ for $C \in \mathbb{R}^{n \times \ell}$, and let $\{\boldsymbol{C}_i\}$ be the set of columns of matrix C. Then $\ker A$ is simply $U^\perp = \{\boldsymbol{C}_i\}^\perp$. As each column $\boldsymbol{C}_i$ is sampled independently from a continuous distribution then $W = \{\boldsymbol{C}_1, \ldots, \boldsymbol{C}_\ell, \mathbf{1}_n\}$ is a linearly independent set almost surely, then $\dim W^\perp = n - \ell - 1$. Since $W^\perp = U^\perp \cap \mathbf{1}_n^\perp$, and $\dim U^\perp = n - \ell$ then $U^\perp$ cannot be contained in $\mathbf{1}_n^\perp$.

Observe that the restriction on the size of the subsystems set $\gamma + 1/n$ as an upper bound for the mode q^*. In the singular case it may happen that q^* satisfying eq. (111) is bigger than $\gamma + 1/n$. Given that we expect the function to be unimodal and increasing with q, then our approximation for the mode in those cases is simply $\gamma + 1/n$.

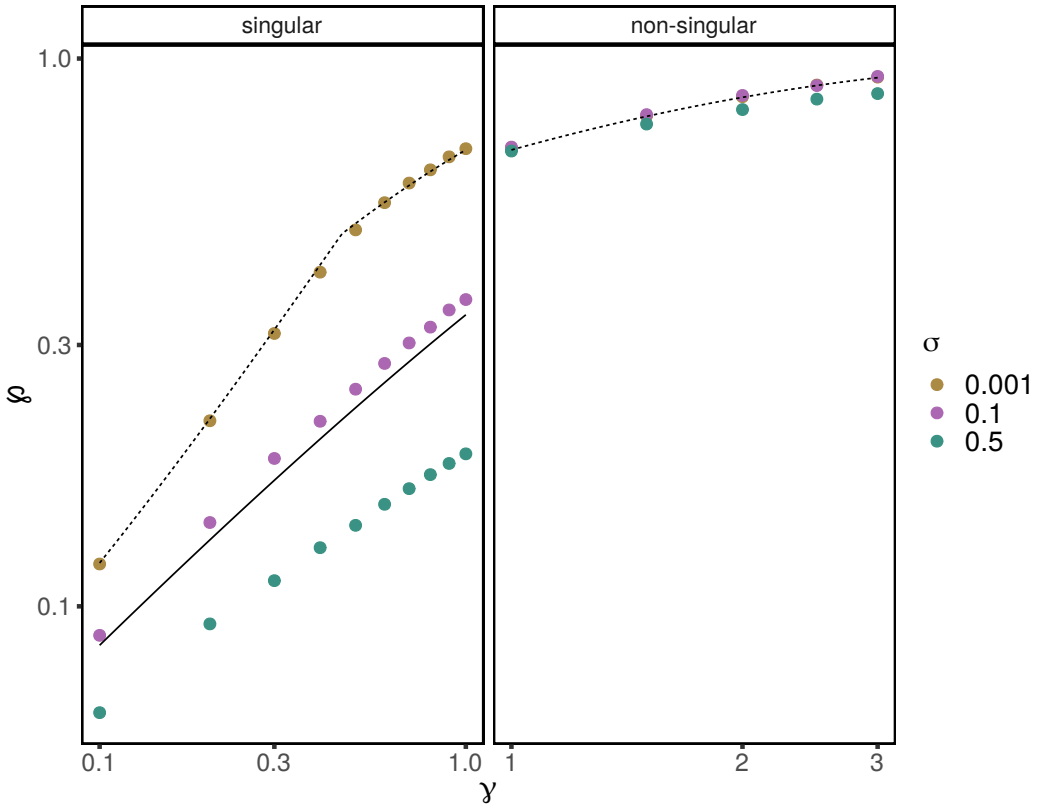

Figure 13: **Fraction of survivors under distinct levels of growth rate variability**. Dots mark the average values over simulations with $r \sim \mathcal{N}(1, \sigma^2)$ and $A \sim \mathcal{W}_\ell(I_n, n)$. In the singular case, the matrix A was perturbed by $A \to A + (b + 0.01)\mathbf{1}_n\mathbf{1}_n^T$ for $b = -\min(A)$. Dotted lines represent our analytical predictions assuming $\sigma = 0$. By Section 6 the shift in A does not affect $\wp$ when $\sigma = 0$. The initial decrease of $\wp$ in the singular case is due to this property not holding when $\sigma \neq 0$. The solid line is our analytical prediction for $\sigma = 0$, when $A \sim \mathcal{W}_\ell(\Sigma, n)$. Σ is a constant correlation matrix with $\rho = \frac{2\sigma_\ell + 0.01}{1 + 2\sigma_\ell + 0.01}$ and $\sigma_\ell = \sqrt{\mathbb{V}(A_{ij})}$ for $i \neq j$ which in this case is simply $\frac{1}{\sqrt{\ell}}$.

References

[1] S. Allesina and S. Tang. The stability–complexity relationship at age 40: a random matrix perspective. *Population Ecology*, 57(1):63–75, 2015.

[2] S. Allesina, J. Grilli, G. Barabás, S. Tang, J. Aljadeff, and A. Maritan. Predicting the stability of large structured food webs. *Nature communications*, 6(1):1–6, 2015.

[3] M. Barbier, J.-F. Arnoldi, G. Bunin, and M. Loreau. Generic assembly patterns in complex ecological communities. *Proceedings of the National Academy of Sciences*, 115 (9):2156–2161, 2018.

[4] U. Bastolla, M. Lässig, S. C. Manrubia, and A. Valleriani. Biodiversity in model ecosystems, ii: species assembly and food web structure. *Journal of theoretical biology*, 235(4): 531–539, 2005.

[5] L. R. Belyea and J. Lancaster. Assembly rules within a contingent ecology. *Oikos*, pages 402–416, 1999.

[6] G. Biroli, G. Bunin, and C. Cammarota. Marginally stable equilibria in critical ecosystems. *New Journal of Physics*, 20(8):083051, 2018.

[7] T. Bodnar and Y. Okhrin. Properties of the singular, inverse and generalized inverse partitioned wishart distributions. *Journal of Multivariate Analysis*, 99(10):2389–2405, 2008.

[8] T. Bodnar and Y. Okhrin. On the product of inverse wishart and normal distributions with applications to discriminant analysis and portfolio theory. *Scandinavian Journal of Statistics*, 38(2):311–331, 2011.

[9] H. C. Bravo, S. Wright, K. Eng, S. Keles, and G. Wahba. Estimating tree-structured covariance matrices via mixed-integer programming. In *Artificial Intelligence and Statistics*, pages 41–48, 2009.

[10] G. Bunin. Ecological communities with lotka-volterra dynamics. *Physical Review E*, 95 (4):042414, 2017.

[11] M. W. Cadotte, T. J. Davies, and P. R. Peres-Neto. Why phylogenies do not always predict ecological differences. *Ecological Monographs*, 87(4):535–551, 2017.

[12] J. A. Capitán, J. A. Cuesta, and J. Bascompte. Statistical mechanics of ecosystem assembly. *Physical Review Letters*, 103(16):168101, 2009.

[13] J. A. Capitán, S. Cuenda, and D. Alonso. How similar can co-occurring species be in the presence of competition and ecological drift? *Journal of The Royal Society Interface*, 12 (110):20150604, 2015. doi: 10.1098/rsif.2015.0604.

[14] M. A. Freilich and S. R. Connolly. Phylogenetic community structure when competition and environmental filtering determine abundances. *Global ecology and biogeography*, 24 (12):1390–1400, 2015.

[15] G. F. Gause. Experimental studies on the struggle for existence: I. mixed population of two species of yeast. *Journal of experimental biology*, 9(4):389–402, 1932.

[16] J. Grilli, T. Rogers, and S. Allesina. Modularity and stability in ecological communities. *Nature communications*, 7(1):1–10, 2016.

[17] J. Grilli, M. Adorisio, S. Suweis, G. Barabás, J. R. Banavar, S. Allesina, and A. Maritan. Feasibility and coexistence of large ecological communities. *Nature communications*, 8: 14389, 2017.

[18] L. Harmon. *Phylogenetic comparative methods: learning from trees*. Self published under a CC-BY-4.0 license, 2018. URL `https://lukejharmon.github.io/pcm`.

[19] J. Hofbauer and K. Sigmund. *Evolutionary games and population dynamics*. Cambridge University Press, 1998.

[20] J. Hofbauer, K. Sigmund, et al. *Evolutionary games and population dynamics*. Cambridge university press, 1998.

[21] C. Hui, H. O. Minoarivelo, and P. Landi. Modelling coevolution in ecological networks with adaptive dynamics. *Mathematical Methods in the Applied Sciences*, 41(18):8407–8422, 2018. doi: 10.1002/mma.4612.

[22] I. Kotsiuba and S. Mazur. On the asymptotic and approximate distributions of the product of an inverse wishart matrix and a gaussian vector. *Theory of Probability and Mathematical Statistics*, 93:103–112, 2016.

[23] N. Kraft, W. Cornwell, C. Webb, and D. Ackerly. Trait evolution, community assembly, and the phylogenetic structure of ecological communities. *The American Naturalist*, 170 (2):271–283, 2007.

[24] S. A. Levin. Community equilibria and stability, and an extension of the competitive exclusion principle. *The American Naturalist*, 104(939):413–423, 1970.

[25] R. Mac Arthur. Species packing, and what competition minimizes. *Proceedings of the National Academy of Sciences*, 64(4):1369–1371, 1969.

[26] R. MacArthur and R. Levins. The limiting similarity, convergence, and divergence of coexisting species. *The American Naturalist*, 101(921):377–385, 1967.

[27] R. M. May. Will a large complex system be stable? *Nature*, 238(5364):413–414, 1972.

[28] M. M. Mayfield and J. M. Levine. Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology letters*, 13(9):1085–1093, 2010.

[29] D. S. Maynard, C. A. Serván, and S. Allesina. Network spandrels reflect ecological assembly. *Ecology letters*, 21(3):324–334, 2018.

[30] R. J. Muirhead. *Aspects of multivariate statistical theory*, volume 197. John Wiley & Sons, 2009.

[31] R. Nichols. Gene trees and species trees are not the same. *Trends in Ecology & Evolution*, 16(7):358–364, 2001.

[32] C. A. Serván and S. Allesina. Tractable models of ecological assembly. *bioRxiv*, 2020. doi: 10.1101/2020.09.02.279943.

[33] C. A. Serván, J. A. Capitán, J. Grilli, K. E. Morrison, and S. Allesina. Coexistence of many species in random ecosystems. *Nature Ecology and Evolution*, 2(8):1237–1242, 2018.

[34] Y. L. Tong. *The multivariate normal distribution*. Springer Science & Business Media, 2012.

[35] C. Violle, D. R. Nemergut, Z. Pu, and L. Jiang. Phylogenetic limiting similarity and competitive exclusion. *Ecology letters*, 14(8):782–787, 2011.

[36] J. Wakely. *Coalescent Theory: An Introduction*. Macmillan Learning, 2016. ISBN 9780974707754.

[37] C. O. Webb. Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *The American Naturalist*, 156(2):145–155, 2000.

[38] C. O. Webb, D. D. Ackerly, M. A. McPeek, and M. J. Donoghue. Phylogenies and community ecology. *Annual review of ecology and systematics*, 33(1):475–505, 2002.

[39] M. G. Weber and A. A. Agrawal. Defense mutualisms enhance plant diversification. *Proceedings of the National Academy of Sciences*, 111(46):16442–16447, 2014. ISSN 0027-8424.

[40] J. Wishart. The generalised product moment distribution in samples from a normal multivariate population. *Biometrika*, pages 32–52, 1928.

[41] P. Yodzis. *Introduction to Theoretical Ecology*. Cambridge: Harper & Row, 1989.

[42] N. Zhao, S. Saavedra, and Y.-Y. Liu. The impact of colonization history on the composition of ecological systems. *bioRxiv*, 2020. doi: 10.1101/2020.02.26.965715.