Interplay of structure and randomness in complex communities

Juan Giral Martínez [1] J.F Arnoldi [2,3] Matthieu Barbier [3,4] Silvia De Monte [1,5]

[1]Institut de Biologie ENS (Paris) [2]CNRS (Moulis) [3]INTP (Surba) [4]CIRAD (Montpellier) [5]Max Planck Institute of Evolutionary Biology (Plön)

Motivations

In recent years, great progress has been made in modelling **species-rich microbial communities** with tools from disordered systems. These models are based on the assumption that interactions between different species are chosen at random.

While they allow to reduce parameters to just a few, they cannot encompass information about ecosystem structure. The genericity of their behaviour is therefore uncertain. We investigate a generalization that combines information about global patterns and a random component modelling unknown interaction details.

Approach

Given an ecological network, we construct a model for the evolution of species abundances that embodies the interaction patterns given by the network. We assume that only information about global patterns is available and that fine-grained details are unknown. We model them through a random uncertainty component of intensity σ.

We find that the dynamics of the whole system are driven by a few collective degrees of freedom, stemming from the collective interaction patterns. We call them 'effective functions', since they play the same role as ecological functions but don't necessarily stem from real ecological processes.

The model

The temporal dynamics for the abundances N_i of S interacting species $i = 1, \dots, S$ are modelled with the generalized Lotka-Volterra equations:

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = N_i \left(1 - \frac{N_i}{K_i} + \sum_j \alpha_{ij} N_j\right)$$

K_i is the carrying capacity of species i, it models competitive self-regulation within the species.

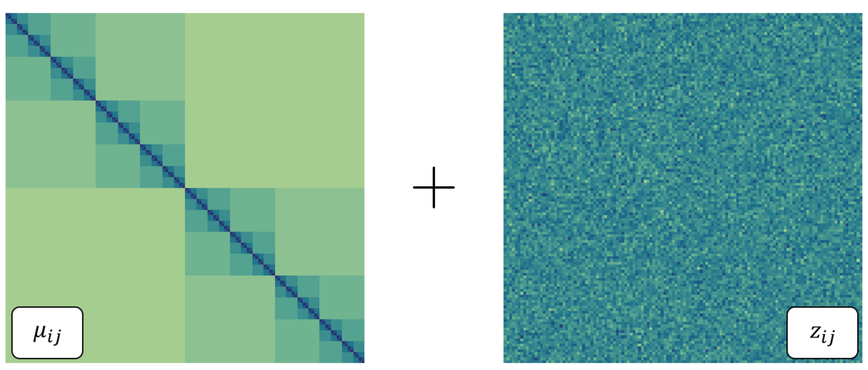
α_{ij} is the effect of species j on species i. We only take pairwise interactions into account.

The interaction matrix is parameterized as the sum of a deterministic structure and a random component:

$$\alpha_{ij} = \frac{\mu_{ij}}{S} + \frac{\sigma}{\sqrt{S}} z_{ij} \qquad \mathbb{E}\left[z_{ij}^2\right] = 1 \qquad \mathbb{E}\left[z_{ij} z_{ji}\right] = \gamma$$

the structure matrix μ_{ij} encodes the modeller's knowledge about the **global interaction patterns**.

σ encodes the **uncertainty** due to the **fine-grained details** of the interactions.



Mapping to effective equations

Using analytical tools from disordered systems, the above equations are equivalent to a set of uncoupled stochastic differential equations:

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = N_i \left(1 - \frac{N_i}{K_i} + \sum_\lambda u_i^\lambda m_\lambda(t) + \sigma \xi_i(t)\right)$$

the uncertainty is replaced by a random noise ξ_i. It has the same statistics for all species.

the relevant features of the community are encompassed by a small set of **'effective functions'** λ such that:

- $m_\lambda(t)$ is the strength of function λ at time t. See §'Species clustering'.
- u_i^λ is the susceptibility of species i to the function λ.

Because each species reacts differently to changes in function strength, each species is different from the rest and the dynamics are high-dimensional. However, because the set of functions is small, such collective dynamics are **driven by a low-dimensional set of variables**.

Fully specified	Our model	Fully disordered
Different species	Different species	Equivalent species
Driven by many quantities	Driven by few functions	Driven by total biomass only

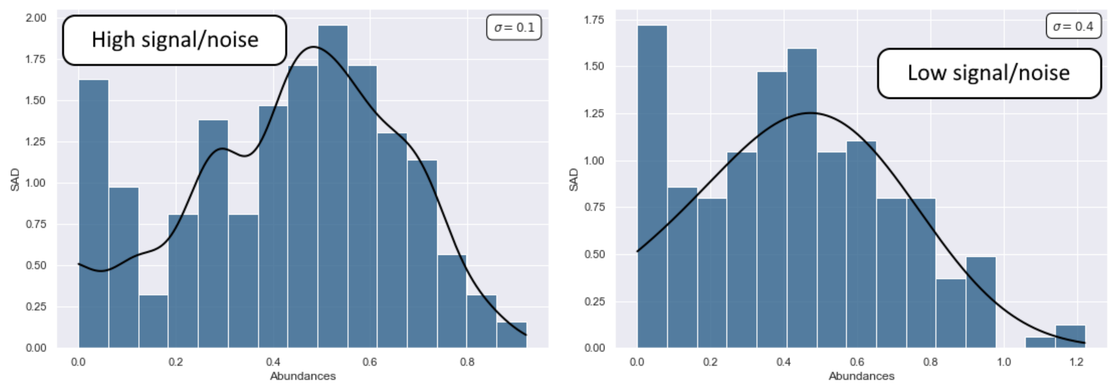
Equilibrium distributions

If the disorder σ is below some threshold, the dynamics reach a globally stable fixed point. The Species Abundance Distribution (SAD) can be caracterized from the knowledge of the effective functions and of σ:

$$x_i = \begin{cases} K_i \left(1 + \sum_\lambda u_i^\lambda m_\lambda + \sigma \xi_i\right) & \text{if positive (extant species)} \\ 0 & \text{otherwise (extinct species)} \end{cases}$$

Because ξ_i are random variables, the abundances of individual species are random. However, **aggregated properties** such as the SAD or function strengths m_λ are not random and **can be predicted**.

The following plots compare SADs from numerical simulations to those predicted by the theory (black curve) for two values of the signal/noise ratio.



The SAD is a simple superposition of a deterministic term stemming from the structure matrix and a random term stemming from interaction uncertainty. Randomness has the effect of smoothing the SAD, hence the stronger it is the more **structure is washed out**.

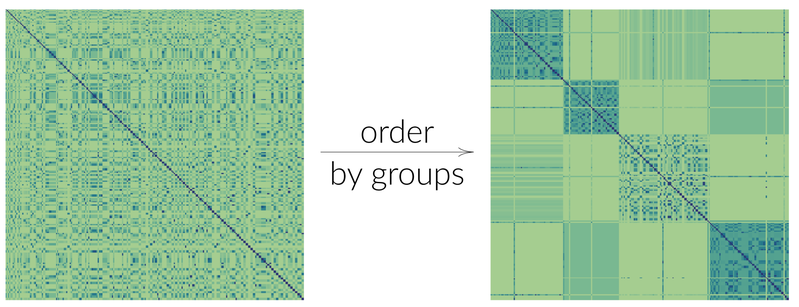
Species clustering

In our model, effective 'functions' emerge as linear combinations of the biomasses of each species

$$m_\lambda(t) = \sum_i v_i^\lambda N_i(t)$$

v_i^λ is the contribution of species i to function λ, e.g. if v_i^λ is the rate of production of some chemical, then $m_\lambda(t)$ is the total rate of production at the community level.

To reconcile this with the traditional view of communities structured in sets of functional groups, we devise an algorithm that clusters species based on their contributions to each function. Species in each group then have a similar **'ecosystem role'**.

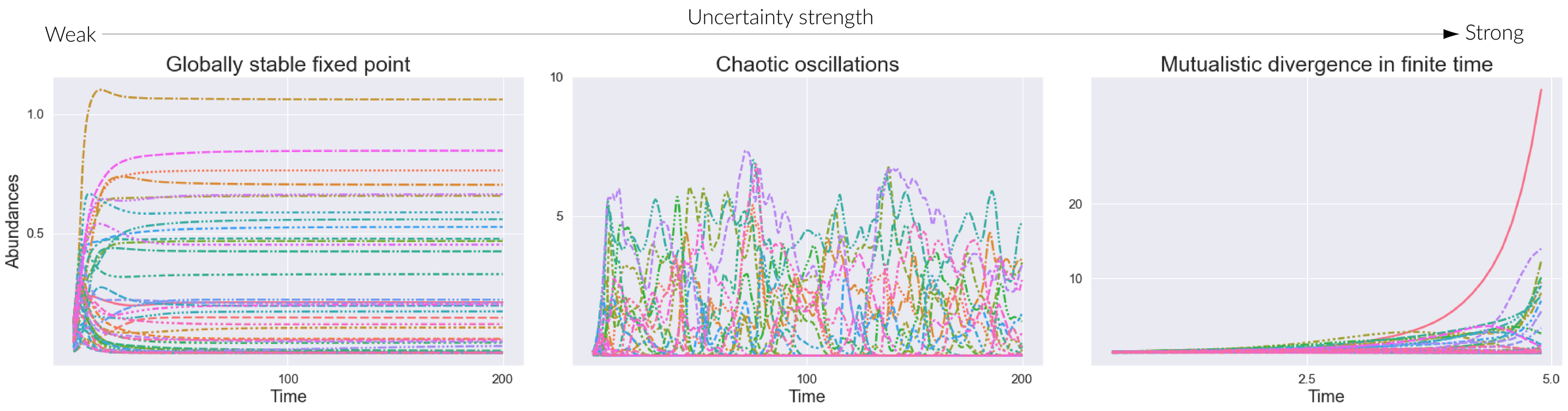


Once a partitioning into groups $C_1, \dots, C_K$ is found, the 'effective functions' can be expressed in terms of the total biomasses of the groups. The dynamics of the system are then approximately driven by changes in these **coarse-grained** biomasses, e.g. changes in individual composition of groups don't matter as long as the total abundance of each group remains stable,

$$M_j(t) = \sum_{i \in C_j} N_i(t)$$

The number K of clusters required for an accurate description of the system is then a measure of its **complexity**: the more species can be aggregated into macroscopic groups, the simpler the community is.

Depending on the amount of uncertainty, the model exhibits different dynamical regimes



jgiral@bio.ens.psl.eu