How random are these grasses?

Matthieu Barbier*
(with G Bunin, JF Arnoldi, C de Mazancourt, M Loreau)

December 4, 2019

(* Centre for Biodiversity Theory and Modelling, CNRS, France)

A few steps beyond random

Main claim: Many cases of both ecological and mathematical interest have *very*, but *not completely*, random(-like) interactions.

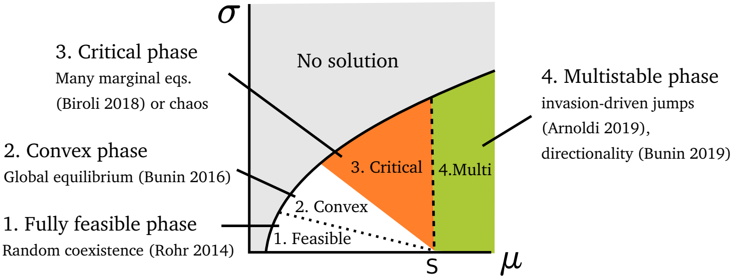
What are the interesting first steps beyond fully-random matrices?

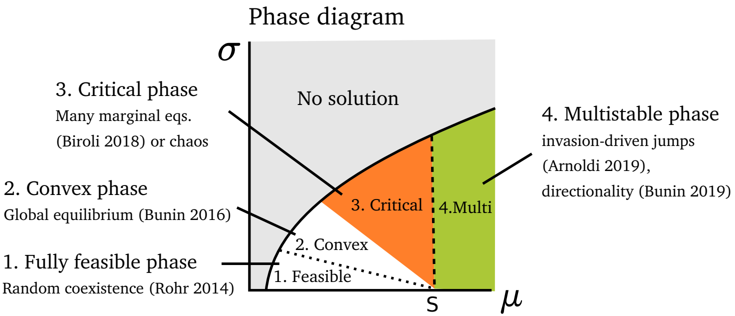
Plan:

1. Minimal structure allowing coexistence
2. Dynamically-relevant extra structure (if time allows)

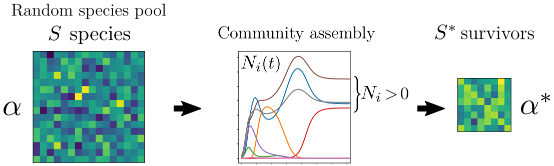
Part 1: Minimal structure

Phase diagram





In region 2, LV dynamics go to globally stable equilibrium, but only $S^* < S$ species survive in it.



Feasible submatrix α^* elements are **not i.i.d.** but they have the *minimal deviation from randomness* that ensures feasibility

Subplan:

- Explain the correlation structure
- Show it in data from grassland experiments

- Our notations for competitive Lotka-Volterra ($\alpha > 0$ = negative interaction)

$$\frac{dN_i}{dt} = \frac{r_i}{K_i} N_i \left(K_i - N_i - \sum_{j \neq i}^{S} \alpha_{ij} N_j \right) \tag{1}$$

- Equilibrium condition for the S^* survivors

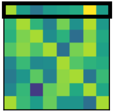
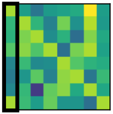
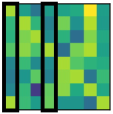
$$N_i = K_i - \sum_{j \neq i}^{S^*} \alpha^*_{ij} N_j \tag{2}$$

- For empirical reasons (explained later), we rather use rescaled abundances $\eta_i = N_i / K_i$

$$\eta_i = 1 - \sum_{j \neq i}^{S^*} \beta^*_{ij} \eta_j \tag{3}$$

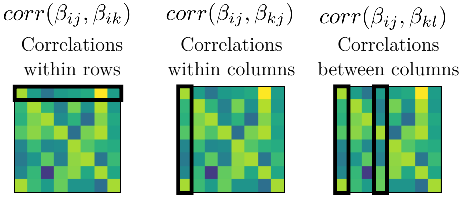
NB: $\beta^*_{ij} = \alpha^*_{ij} K_j / K_i$

Bunin (PRE 2017) showed that β^*_{ij} are not i.i.d.
Species die = rows and columns disappear, creating biases and correlations (shown later).

$corr(\beta_{ij}, \beta_{ik})$	$corr(\beta_{ij}, \beta_{kj})$	$corr(\beta_{ij}, \beta_{kl})$
Correlations within rows	Correlations within columns	Correlations between columns
		

Correlations

Bunin (PRE 2017) showed that β^*_{ij} are not i.i.d.
Species die = rows and columns disappear, creating biases and correlations (shown later).



Surprisingly, same result from simple probabilistic argument:
rather than run dynamics, directly draw β_{ij} from Gaussian with mean $\bar{\beta}$ subject to the S linear constraints

$$0 = 1 - \eta_i - \sum_{j \neq i}^{S} \beta_{ij} \eta_j \tag{4}$$

Correlations

Simple calculation (Lagrange multipliers or rotation) gives:

$$E[\beta_{ij}|\eta_i,\eta_j] = \bar{\beta} + (1-\bar{\beta})\Delta(\eta_i,\eta_j) \tag{5}$$

$$\mathrm{corr}(\beta_{ij},\beta_{ik}|\eta_i,\eta_j,\eta_k) = -\frac{\eta_j\eta_k}{\sum_{m\neq i}\eta_m^2}. \tag{6}$$

where mean interactions are biased to achieve the correct η_i:

$$\Delta(\eta_i,\eta_j) = -\frac{(\eta_i-\eta^*)\eta_j}{\sum_{m\neq i}\eta_m^2}. \tag{7}$$

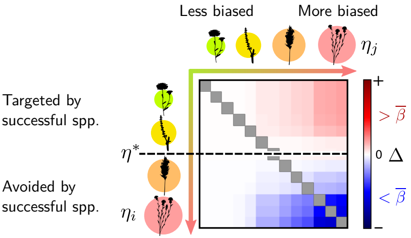
and η^* is baseline abundance obtained if all $\beta_{ij} = \bar{\beta}$

$$\eta^* = \frac{1-\bar{\beta}\sum_i \eta_i}{1-\bar{\beta}}. \tag{8}$$

Correlations

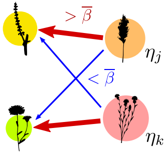
Two statistical patterns among coexisting species

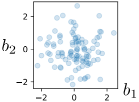
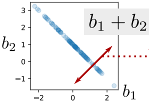
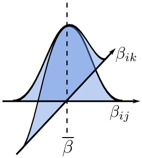
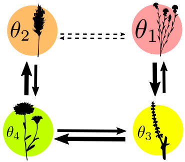
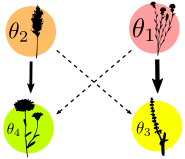
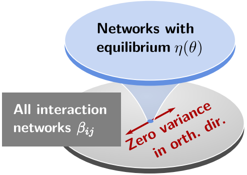
(a) Competitive bias $\Delta(\eta_i, \eta_j) \sim -(\eta_i - \eta^*)\eta_j$



(b) Overlap avoidance $\mathrm{corr}(\beta_{ij}, \beta_{ik}) \sim -\eta_j \eta_k$

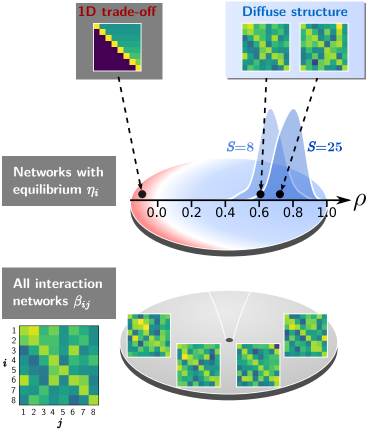
Two successful species do not target the same species.



	Prior distribution	Bias on means	Correlation/covariance	Transformed distribution
Simple example		mean(b_1)=0 mean(b_2)=1	corr(b_1, b_2)=-1 $\mathrm{Cov}_{ij} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$	$b_1 + b_2 - 1 = 0$ 
Clique pattern	Multivariate Gaussian 	$\mathrm{mean}(\beta_{ij}) \sim -A\theta_i\theta_j$ 	$\mathrm{corr}(\beta_{ij}, \beta_{ik}) \sim -A\,\theta_j\theta_k$ 	$\sum_j \beta_{ij}\theta_j - 1 = 0$ 

Interpretation

Many interaction matrices β_{ij} admit the same equilibrium η_i
($O(S^2)$ parameters for only S constraints)



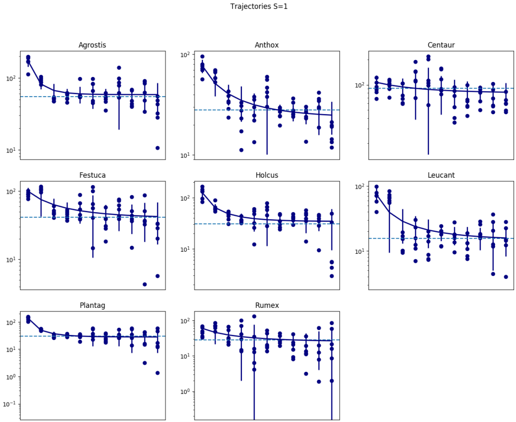
Most (from random prior) will exhibit our predicted correlations,
i.e. minimal deviation from randomness allowing these η_i

Finding it in data

Each plot = some combination of 1, 2... species from a pool of S species

What data?

Data = time series $N_i(t)$ in various combinations of species



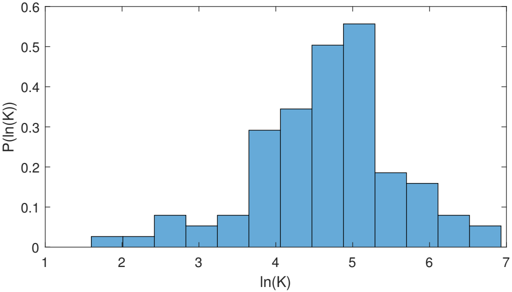
(here Wageningen grassland experiment, $S = 8$, 11 years, 4 replicates)

Inferring interactions

- Somehow estimate equilibrium abundance N_i
- Multilinear fit of interactions α:
  each species combination k = point on hyperplane defined by

$$N_i^{(k)} = K_i - \sum_{j \in k} \alpha_{ij} N_j^{(k)} \tag{9}$$

- Problem: K_i are widely distributed (and α too)

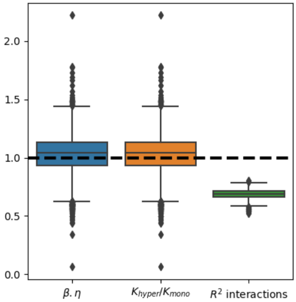


Inferring interactions

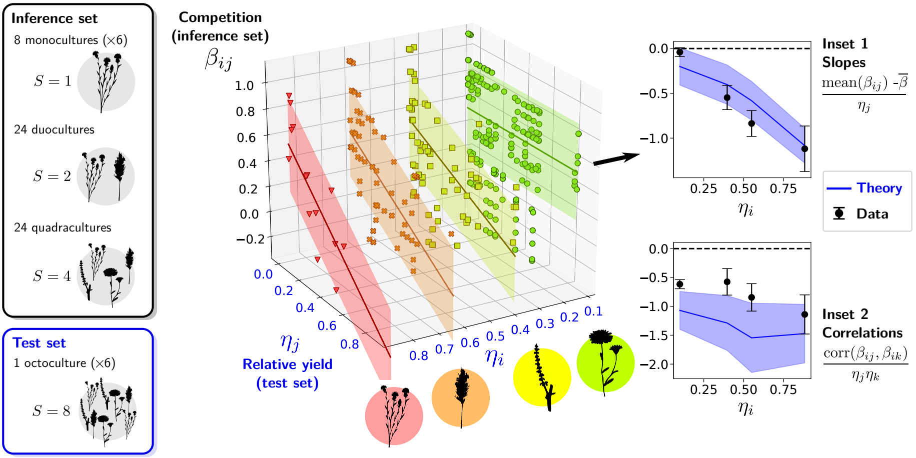
- Solution: take $\eta_i = N_i/K_i$ and fit

$$\eta_i = 1 - \sum_{j \neq i} \beta_{ij} \eta_j \tag{10}$$

- Sanity checks: Wageningen really looks like LV equilibrium

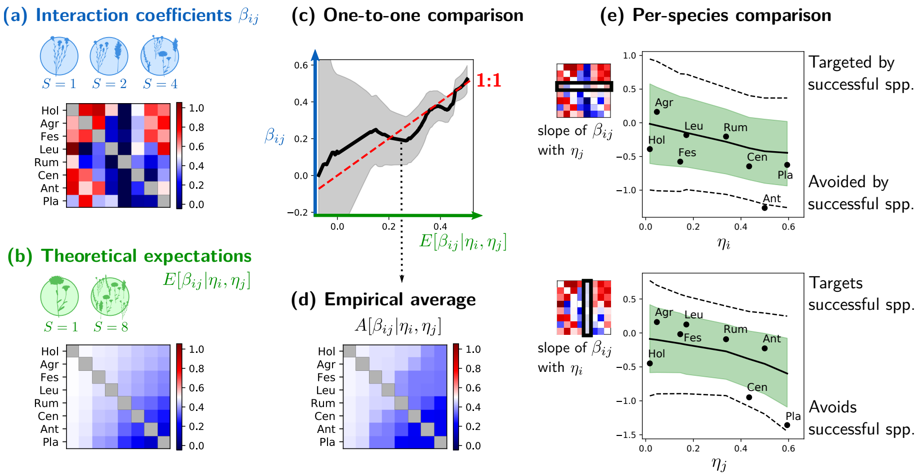


Once we have individual interactions β_{ij}, see if they show the expected trends



(NB: theoretical predictions controlled only by η_i in the full $S = 8$ composition; shaded area = variation of prediction due to error on η_i)

Every way we slice and dice the data, it matches theory



Rightmost panel: shaded area: +- 1 SD of predictions, dashes: 95% CI of predictions

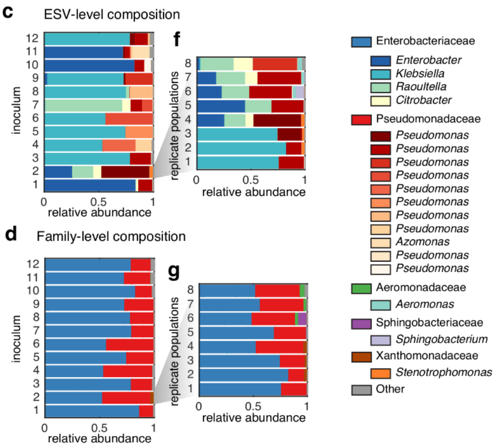
Transition

- Within a group of sufficiently similar grasses, good empirical evidence of minimal deviation from randomness
  (just enough to allow coexistence)

- What about larger ecological networks, how random are they?

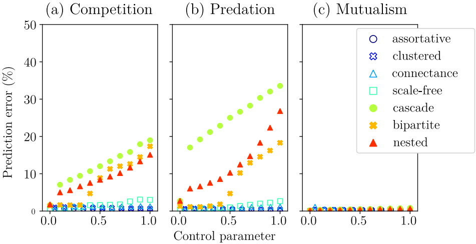
Part 2: Extra structure

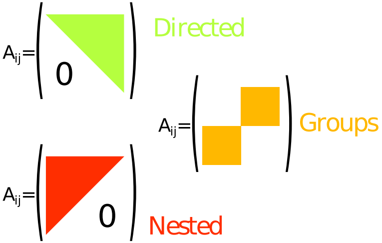
- Random matrix: all species are statistically equivalent, hence we care only about scaled mean μ, SD σ and symmetry γ of interactions.
- Reality: not all statistically equivalent, but there are structures e.g. functional groups



(Goldford et al)

- Not all extra structure is dynamically relevant
- In ecological models, large macrostructures are main causes of departure from mean-field (random prediction)

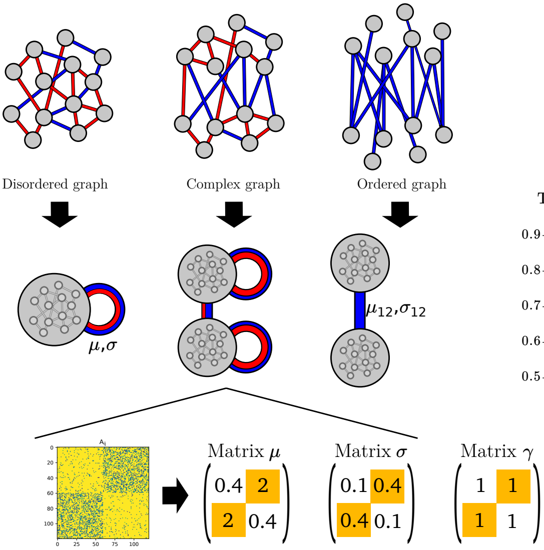


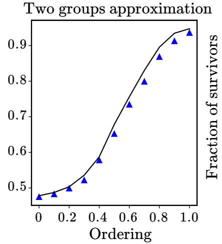


x-axis = degree of order (structure-specific control parameter)
y-axis = difference between abundance distribution for random versus structured interactions

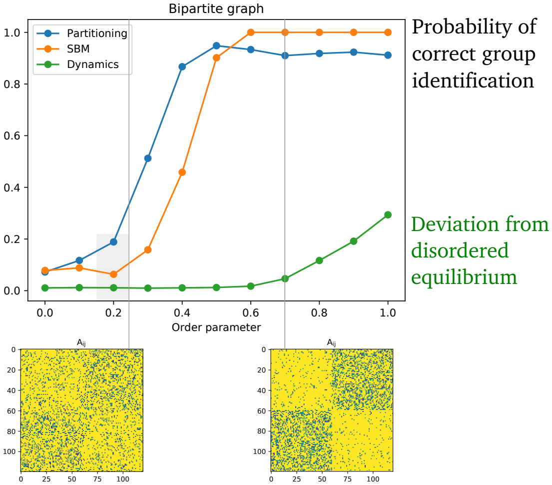
(Barbier, Arnoldi, Bunin and Loreau, PNAS 2018)

Simple extra structure can be captured by a not-completely-random approximation, e.g. matrix of per-group interaction mean μ_{xy} and SD σ_{xy}





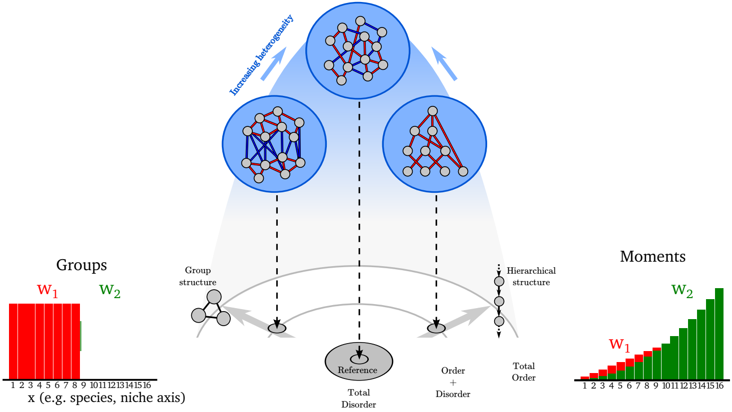
Stochastic Block Models identify the correct structure in the interaction network



(in fact they are a bit overzealous: they identify it *before* it starts to matter in susceptibility V_{ij}, which is more self-averaging)

Directions of deviation from randomness?

Different ways of adding order with more parameters: groups (matrix μ_{xy}), Taylor expansion (function $\mu(x, y)$)...



Technical point of view: low rank structures, perturbed random matrix theory

Conclusions

- Random matrices are not a bad starting point for real systems: empirical evidence of minimal deviation from randomness
- However it is worth looking a bit beyond randomness, into simple order+disorder combinations

"There is a fundamental dichotomy between structure and randomness, which in turn leads to a decomposition of any object into a structured (low-complexity) component and a random (discorrelated) component."

Terence Tao, The dichotomy between structure and randomness, arithmetic progressions, and the primes. 2006 ICM proceedings.