

ECOLOGICAL NETWORKS AND COMPLEX SYSTEMS GUSTAVE EIFFEL UNIVERSITY, 28-29th October, 2024

How do non-Gaussian interactions drive patterns and coexistence in large ecosystems?

Sandro Azaele Department of Physics and Astronomy "G. Galileo", University of Padua & INFN





NATIONAL BIODIVERSITY FUTURE CENTER



Finanziato dall'Unione europea NextGenerationEU



The coexistence problem is as old as the entangled bank

"It is interesting to contemplate an **entangled bank**, clothed with many plants of many kinds, with birds singing on the bushes, with various insects flitting about, and with worms crawling through the damp earth, and to reflect that **these elaborately constructed forms**, so different from each other, and dependent on each other in so complex a manner, **have all been produced by laws acting around us**."

- C. Darwin, On the Origin of Species, 1859 -





multi-species coexistence is an *additive affair,* and all of the coexisting members of a community must also coexist as pairs when isolated from the community context.

Friedman *et al.*, Nat. Ecol. Evol., 2017; Letten *et al.*, Ecol. Letters, 2019; Meroz *et al.*, Nat. Comm., 2020.



multi-species coexistence is an *additive affair*, and all of the coexisting members of a community must also coexist as pairs when isolated from the community context.

Friedman *et al.*, Nat. Ecol. Evol., 2017; Letten *et al.*, Ecol. Letters, 2019; Meroz *et al.*, Nat. Comm., 2020.

Chang-Yu C. et al., Science 2023

coexistence in a multi-species community is a *property of the community*, which is not exhibited by its most elementary units of coexistence, pairs of species in isolation



Friedman *et al.*, Nat. Ecol. Evol., 2017; Letten *et al.*, Ecol. Letters, 2019; Meroz *et al.*, Nat. Comm., 2020.

isolated from the community context.

coexistence in a multi-species community is a *property of the community*, which is not exhibited by its most elementary units of coexistence, pairs of species in isolation

"The best possible knowledge of a *whole* does not necessarily include the best possible knowledge of all its *parts*." — *E. Schrödinger, Discussion of prob. relations...,* 1935 —

Community structure from simple assembly rules?



species will go extinct.

89.5% of trios were predicted by pairwise outcome!

Community context is required for species pairs to coexist

RESEARCH

Science, July 2023

MICROBIOTA

Emergent coexistence in multispecies microbial communities

Chang-Yu Chang^{12,3}*, Djordje Bajić^{12,4}*, Jean C. C. Vila^{1,2}, Sylvie Estrela^{1,2}†, Alvaro Sanchez^{1,2,5}*

Community context is required for species pairs to coexist

RESEARCH

Science, July 2023

MICROBIOTA

Emergent coexistence in multispecies microbial communities

Chang-Yu Chang^{12,3}*, Djordje Bajić^{12,4}*, Jean C. C. Vila^{1,2}, Sylvie Estrela^{1,2}†, Alvaro Sanchez^{1,2,5}*

0.00







In 26.4% of the pairs, one of the two species had become competitively excluded;

In other 45.1% of the pairs, the frequency of one of the species declined, regardless of its initial proportion;

71.6% of the pairs failed to coexist in the absence of the other community members.

Community context is required for species pairs to coexist

RESEARCH

Science, July 2023

MICROBIOTA

Emergent coexistence in multispecies microbial communities

Chang-Yu Chang^{1,2,3}*, Djordje Bajić^{1,2,4}*, Jean C. C. Vila^{1,2}, Sylvie Estrela^{1,2}†, Alvaro Sanchez^{1,2,5}*



71.6% of the pairs failed to coexist in the absence of the other community members.

Given that both hypotheses can be correct in different communities, under which conditions each is most likely to occur?

higher-order interactions? specific network of pairwise interactions? evolution of new species interactions?

...

Can we use the Generalised Lotka-Volterra equations and investigate coexistence?

N interacting species: Generalised Lotka- Volterra (GLV) equations

$$\dot{x}_i(t) = x_i(t) \left(r_i + \sum_{j=1}^N \alpha_{ij} x_j(t) \right)$$

- * you can "describe" competition, mutualism, predator-prey, random interactions...
 - but the *interaction matrix is unknown*;
- ** interactions are direct and pairwise (random encounters);*
- * are interactions mediated by resources?

Can we use the Generalised Lotka-Volterra equations and investigate coexistence?

$$\dot{x}_i(t) = x_i(t) \left(r_i + \sum_{j=1}^N \alpha_{ij} x_j(t) \right)$$

* you can "describe" competition, mutualism, predator-prey, random interactions...

but the interaction matrix is unknown;

- * interactions are direct and pairwise (random encounters);
- * are interactions mediated by resources?

- What kind of properties can we assume for the interaction couplings?
- Inference for the interactions is difficult for large N, but *statistically* one can make some assumptions.
- \bigcirc In natural communities *N* is typically large, so one wonders *what is the "typical behaviour of the GLV as N* → ∞, … *typical patterns*

Is it possible to do a statistical mechanics of GLV?

Can we do a statistical mechanics of GLV?

BULLETIN OF MATHEMATICAL BIOPHYSICS VOLUME 19, 1957

 $\dot{x}_{i}(t) = x_{i}(t) \left(r_{i} + \sum_{j=1}^{N} \alpha_{ij} x_{j}(t) \right)$

A STATISTICAL MECHANICS OF INTERACTING BIOLOGICAL SPECIES

EDWARD H. KERNER PHYSICS DEPARTMENT UNIVERSITY OF BUFFALO, BUFFALO, NEW YORK

The system of differential equations proposed by V. Volterra, describing the variation in time of the populations N, of interacting species in a biological association, admits a Liouville's theorem (when $\log N$ are used as variables) and a universal integral of "motion." Gibbs' microcanonical and canonical ensembles can then provide a thermodynamic description of the association in the large. The "temperature" measures in one number common to all species the mean-square deviations of the N_r from their average values. There are several equipartition theorems. susceptible of direct experimental test, a theorem on the flow of "heat" (the conserved quantity in an isolated association) between two weakly coupled associations at different temperatures, a Dulong-Petit law for the heat capacity, and an analog of the second law of thermodynamics expressing the tendency of an association to decline into an equilibrium state of maximal entropy. The analog of the Maxwell-Boltzmann law is a distribution of intrinsic abundance for each species which has been successfully used by ecologists for interpreting experimental data. A true thermodynamics develops upon introducing the idea of work done on an association through a variation of the variables (such as physical temperature) defining the physical and chemical environment. An ergodic theorem is suggested by the agreement of ensemble and time averages in the one case where the latter may be found explicitly.

Can we do a statistical mechanics of GLV?

BULLETIN OF MATHEMATICAL BIOPHYSICS VOLUME 19, 1957

A STATISTICAL MECHANICS OF INTERACTING BIOLOGICAL SPECIES

EDWARD H. KERNER PHYSICS DEPARTMENT UNIVERSITY OF BUFFALO, BUFFALO, NEW YORK

The system of differential equations proposed by V. Volterra, describing the variation in time of the populations N, of interacting species in a biological association, admits a Liouville's theorem (when $\log N$, are used as variables) and a universal integral of "motion." Gibbs' microcanonical and canonical ensembles can then provide a thermodynamic description of the association in the large. The "temperature" measures in one number common to all species the mean-square deviations of the N_r from their average values. There are several equipartition theorems. susceptible of direct experimental test, a theorem on the flow of "heat" (the conserved quantity in an isolated association) between two weakly coupled associations at different temperatures, a Dulong-Petit law for the heat capacity, and an analog of the second law of thermodynamics expressing the tendency of an association to decline into an equilibrium state of maximal entropy. The analog of the Maxwell-Boltzmann law is a distribution of intrinsic abundance for each species which has been successfully used by ecologists for interpreting experimental data. A true thermodynamics develops upon introducing the idea of work done on an association through a variation of the variables (such as physical temperature) defining the physical and chemical environment. An ergodic theorem is suggested by the agreement of ensemble and time averages in the one case where the latter may be found explicitly.

$$\dot{x}_i(t) = x_i(t) \left(r_i + \sum_{j=1}^N \alpha_{ij} x_j(t) \right)$$

One can write GLV canonically (Hamilton eqs.) and define a *Gibbs ensemble*

the "*temperature*" measures the mean-square deviations of the x_i from their average values

The analog of the *Maxwell-Boltzmann* law is a distribution of abundance for each species which is given by a Gamma distribution

BUT, you can do this only when $\alpha_{ij} = -\alpha_{ji}$

Can we do a statistical mechanics of GLV?

BULLETIN OF MATHEMATICAL BIOPHYSICS VOLUME 19, 1957

A STATISTICAL MECHANICS OF INTERACTING **BIOLOGICAL SPECIES**

EDWARD H. KERNER PHYSICS DEPARTMENT UNIVERSITY OF BUFFALO, BUFFALO, NEW YORK

The system of differential equations proposed by V. Volterra, describing the variation in time of the populations N_{\star} of interacting species

in a biological association, admit used as variables) and a univers canonical and canonical ensemi description of the association in in one number common to all sp N_r from their average values. T susceptible of direct experiments (the conserved quantity in an is coupled associations at differen the heat capacity, and an analo expressing the tendency of an as state of maximal entropy. The a distribution of intrinsic abunda_____

REVIEWS OF MODERN PHYSICS

 $\dot{x}_i(t) = x_i(t) \left| r_i + \sum_{i=1}^N \alpha_{ij} x_j(t) \right|$

One can write GLV canonically (Hamilton eqs.) and define a Gibbs ensemble

VOLUME 43, NUMBER 2, PART 1 APRIL 1971

On the Volterra and Other Nonlinear Models of Interacting Populations*

NARENDRA S. GOEL, SAMARESH C. MAITRA, AND ELLIOTT W. MONTROLL Institute for Fundamental Studies, Department of Physics and Astronomy, The University of Rochester, Rochester, New York 14627

successfully used by ecologists for interpreting experimental data. A true thermodynamics develops upon introducing the idea of work done on an association through a variation of the variables (such as physical temperature) defining the physical and chemical environment. An ergodic theorem is suggested by the agreement of ensemble and time averages in the one case where the latter may be found explicitly.

BUT, you can do this only when $\alpha_{ii} = -\alpha_{ii}$

$$\dot{x}_i(t) = x_i(t) \left(r_i + \sum_{j=1}^N \alpha_{ij} x_j(t) \right)$$

Properties of the interaction couplings for large *N* (*quenched* noise):

$$\alpha_{ij} = \frac{\mu}{N} + \frac{\sigma}{\sqrt{N}} z_{ij}$$

for simplicity $r_i = 1$ and $\alpha_{ii} = -1$.

where

$$\langle z_{ij} \rangle = 0$$
 $\langle z_{ij} z_{lm} \rangle = \delta_{il} \delta_{jm}$ uncorrelated Gaussian r.v.

or more generally

$$\langle z_{ij} z_{ji} \rangle = \gamma$$

... and zero for all other indexes

$$\dot{x}_i(t) = x_i(t) \left(r_i + \sum_{j=1}^N \alpha_{ij} x_j(t) \right)$$

for simplicity $r_i = 1$ and $\alpha_{ii} = -1$.

$$\alpha_{ij} = \frac{\mu}{N} + \frac{\sigma}{\sqrt{N}} z_{ij}$$



$$\mu = -2, \, \sigma = 1$$

$$N = 100, \lambda = 10^{-10}$$

$$\dot{x}_i(t) = x_i(t) \left(r_i + \sum_{j=1}^N \alpha_{ij} x_j(t) \right)$$

for simplicity $r_i = 1$ and $\alpha_{ii} = -1$.

$$\alpha_{ij} = \frac{\mu}{N} + \frac{\sigma}{\sqrt{N}} z_{ij}$$



$$N = 100, \lambda = 10^{-10}$$



$$\dot{x}_{i}(t) = x_{i}(t) \left(r_{i} + \sum_{j=1}^{N} \alpha_{ij} x_{j}(t) \right)$$

for simplicity $r_i = 1$ and $\alpha_{ii} = -1$.

$$\alpha_{ij} = \frac{\mu}{N} + \frac{\sigma}{\sqrt{N}} z_{ij}$$

$$u = -2, \sigma = 3$$

$$N = 100, \lambda = 10^{-10}$$

Dynamical Mean Field Theory of GLV

$$\dot{x}_i(t) = x_i(t) \left(1 - x_i + \sum_{j \neq i}^N \alpha_{ij} x_j(t) \right)$$

$$\alpha_{ij} = \frac{\mu}{N} + \frac{\sigma}{\sqrt{N}} z_{ij}$$

$$\langle z_{ij} \rangle = 0 \quad \langle z_{ij} z_{lm} \rangle = \delta_{il} \delta_{jm}$$

Dynamical Mean Field Theory of GLV

$$\dot{x}_{i}(t) = x_{i}(t) \left(1 - x_{i} + \sum_{j \neq i}^{N} \alpha_{ij} x_{j}(t) \right)$$
$$\alpha_{ij} = \frac{\mu}{N} + \frac{\sigma}{\sqrt{N}} z_{ij} \qquad \qquad \langle z_{ij} \rangle = 0 \quad \langle z_{ij} z_{lm} \rangle = \delta_{il} \delta_{jm}$$
$$N \to \infty$$
$$T. Galla, arxiv.org/abs/2405.14289$$

Equation for a representative species (*neutral* limit with *random self-consistent environment*):

 $\dot{x}(t) = x(t) \left| 1 - x(t) + \mu M(t) + \gamma \sigma^2 \int_0^t dt' G(t, t') x(t') + \eta(t) \right|$

$$\dot{x}(t) = x(t) \left[1 - x(t) + \mu M(t) + \eta(t) \right] \qquad \begin{array}{l} M(t) = \langle x(t) \rangle \\ \langle \eta(t)\eta(t') \rangle = \sigma^2 \langle x(t)x(t') \rangle \end{array}$$

interactions with correlations —> x is non-Markovian

Dynamical Mean Field Theory: the phase diagram

$$\dot{x}_i(t) = x_i(t) \left(r_i + \sum_{j=1}^N \alpha_{ij} x_j(t) \right)$$







Bunin G, PRE (2017); Galla T., EPL (2018); Biroli G. et al., NJP (2018); Roy F. et al., JPA (2019); Altieri A. et al., PRL (2021)...



- *Truncated Gaussians* have not been observed in natural systems;
- Empirical patterns have not such a level of universality (mean and variance are not enough!)
- * There are models in which parameters are non-negative or non-positive by definition;
- sparse interactions;



Species' interactions may be *non-Gaussian*: can we generalise DMFT to this case? *Truncated Gaussians* have not been observed in natural systems;

- Empirical patterns have not such a level of universality (mean and variance are not enough!)
- * There are models in which parameters are non-negative or non-positive by definition;
- sparse interactions;

Metabolic strategies in CR models may *vary with time*: what happens to the CEP?



Species' interactions may be *non-Gaussian*: can we generalise DMFT to this case?

- *Truncated Gaussians* have not been observed in natural systems;
- Empirical patterns have not such a level of universality (mean and variance are not enough!)
- * There are models in which parameters are non-negative or non-positive by definition;
- sparse interactions;

Metabolic strategies in CR models may *vary with time*: what happens to the CEP?



Species' interactions may be *non-Gaussian*: can we generalise DMFT to this case?

- *Truncated Gaussians* have not been observed in natural systems;
- Empirical patterns have not such a level of universality (mean and variance are not enough!)
- There are models in which parameters are non-negative or non-positive by definition;
 sparse interactions;

Metabolic strategies in CR models may *vary with time*: what happens to the CEP?



Species' interactions may be *non-Gaussian*: can we generalise DMFT to this case?

- *Truncated Gaussians* have not been observed in natural systems;
- Empirical patterns have not such a level of universality (mean and variance are not enough!)
- * There are models in which parameters are non-negative or non-positive by definition;
- sparse interactions;

Metabolic strategies in CR models may *vary with time*: what happens to the CEP?

What happens for more general distributions?

$$P_{N}(\alpha) = \frac{1}{\sqrt{2\pi\sigma^{2}/N}} e^{-\frac{(\alpha - \mu/N)^{2}}{2\sigma^{2}/N}} \qquad \qquad \langle \alpha \rangle = \mu/N \langle \alpha^{2} \rangle - \langle \alpha \rangle^{2} = \sigma^{2}/N$$

$$P_N(\alpha) = \frac{|\alpha - \mu(N)|^{-1 + \bar{\delta}}}{2\bar{\beta}^{\bar{\delta}}\Gamma(\bar{\delta})} e^{-|\alpha - \mu(N)|/\bar{\beta}}$$

$$\langle \alpha \rangle = \mu(N) = \mu/N$$
$$\alpha^2 \rangle - \langle \alpha \rangle^2 = \bar{\beta}^2 \bar{\delta}(\bar{\delta} + 1) = \bar{\beta}(N)^2 \bar{\delta}(N)(\bar{\delta}(N) + 1)$$

What happens for more general distributions?

$$P_{N}(\alpha) = \frac{1}{\sqrt{2\pi\sigma^{2}/N}} e^{-\frac{(\alpha - \mu/N)^{2}}{2\sigma^{2}/N}} \qquad \qquad \langle \alpha \rangle = \mu/N \langle \alpha^{2} \rangle - \langle \alpha \rangle^{2} = \sigma^{2}/N$$

$$P_N(\alpha) = \frac{|\alpha - \mu(N)|^{-1 + \bar{\delta}}}{2\bar{\beta}^{\bar{\delta}}\Gamma(\bar{\delta})} e^{-|\alpha - \mu(N)|/\bar{\beta}}$$

$$\langle \alpha \rangle = \mu(N) = \mu/N$$

$$\langle \alpha^2 \rangle - \langle \alpha \rangle^2 = \bar{\beta}^2 \bar{\delta}(\bar{\delta} + 1) = \bar{\beta}(N)^2 \bar{\delta}(N)(\bar{\delta}(N) + 1)$$

what scaling of pars?

- different functions $\overline{\beta}(N)$ and $\overline{\delta}(N)$ give the same $var(\alpha)$. - higher cumulants are no longer zero

JW Baron et al., Phys. Rev. Lett. (2023)

Instead of considering only the first two cumulants, we introduce a distribution of the off-diagonal coefficients, $P_N(\alpha_{ij})$, whose characteristic function scales with *N* as

$$\lim_{N \to +\infty} N \ln \left(\int d\alpha \ P_N(\alpha) e^{-i\alpha z} \right) = F(z)$$

where F(z) is a complex function which is analytic at least at z=0, and F(0)=0. F(z)/N is the cumulant-generating function (because the ch.f. of a pdf is the exponential of its c.g.f.)

Instead of considering only the first two cumulants, we introduce a distribution of the off-diagonal coefficients, $P_N(\alpha_{ij})$, whose characteristic function scales with *N* as

$$\lim_{N \to +\infty} N \ln \left(\int d\alpha \ P_N(\alpha) e^{-i\alpha z} \right) = F(z)$$

where F(z) is a complex function which is analytic at least at z=0, and F(0)=0. F(z)/N is the cumulant-generating function (because the ch.f. of a pdf is the exponential of its c.g.f.)

When $F(z) = -i\mu z - \sigma^2 z^2/2$ (higher cumulants are zero), we recover the assumptions of the standard DMFT:

$$P_N(\alpha) = \frac{1}{\sqrt{2\pi\sigma^2/N}} e^{-\frac{(\alpha - \mu/N)^2}{2\sigma^2/N}}$$

which means

$$\alpha_{ij} = \frac{\mu}{N} + \frac{\sigma}{\sqrt{N}} z_{ij}$$

How do we scale the parameters with *N* in order to get a meaningful F(z)? (Remember *Marcinkiewicz's theorem*: F(z) is either at most a polynomial of degree two or an infinite power series)

$$P_N(\alpha) = \frac{|\alpha - \mu/N|^{-1 + \bar{\delta}(N)}}{2\bar{\beta}(N)\bar{\delta}^{(N)}\Gamma(\bar{\delta}(N))} e^{-|\alpha - \mu/N|/\bar{\beta}(N)}$$



How do we scale the parameters with *N* in order to get a meaningful F(z)? (Remember *Marcinkiewicz's theorem*: F(z) is either at most a polynomial of degree two or an infinite power series)

$$P_N(\alpha) = \frac{|\alpha - \mu/N|^{-1 + \delta(N)}}{2\bar{\beta}(N)\bar{\delta}(N)\Gamma(\bar{\delta}(N))} e^{-|\alpha - \mu/N|/\bar{\beta}(N)}$$

scaling of pars

$$\bar{\delta}(N) = \delta$$
 and $\bar{\beta}(N) = \beta / \sqrt{N}$

then we get

$$F(z) = -i\mu z - \beta^2 \delta(\delta + 1) z^2/2$$

How do we scale the parameters with *N* in order to get a meaningful F(z)? (Remember *Marcinkiewicz's theorem*: F(z) is either at most a polynomial of degree two or an infinite power series)

$$P_{N}(\alpha) = \frac{|\alpha - \mu/N|^{-1 + \bar{\delta}(N)}}{2\bar{\beta}(N)\bar{\delta}^{(N)}\Gamma(\bar{\delta}(N))}}e^{-|\alpha - \mu/N|/\bar{\beta}(N)}$$

scaling of pars
 $\bar{\delta}(N) = \delta$ and $\bar{\beta}(N) = \beta/\sqrt{N}$
 $\bar{\delta}(N) = \delta/N$ and $\bar{\beta}(N) = \beta$
then we get instead we get

 $F(z) = -i\mu z - \frac{\sigma}{2}\log(1 + \beta^2 z^2)$

$$F(z) = -i\mu z - \beta^2 \delta(\delta + 1) z^2/2$$

 $\overline{\delta}$

$$\dot{x}_i(t) = x_i(t) \left(1 - x_i + \sum_{\substack{j \neq i}}^N \alpha_{ij} x_j(t) \right) \qquad N$$

N interacting species $r_i = 1$ and $\alpha_{ii} = -1$.

$$\lim_{N \to +\infty} N \ln \left(\int d\alpha \ P_N(\alpha) e^{-i\alpha z} \right) = F(z) \qquad \qquad N \to \infty \qquad \text{New generalised DMFT}$$

$$\dot{x}_{i}(t) = x_{i}(t) \left(1 - x_{i} + \sum_{j \neq i}^{N} \alpha_{ij} x_{j}(t) \right) \qquad N \text{ interacting species} \\ r_{i} = 1 \text{ and } \alpha_{ii} = -1.$$

$$\lim_{N \to +\infty} N \ln \left(\int d\alpha P_{N}(\alpha) e^{-i\alpha z} \right) = F(z) \qquad N \to \infty \qquad \text{New generalised DMFT}$$

Equation for a representative species:

$$\dot{x}(t) = x(t) \left[1 - x(t) + \mu M(t) + \eta(t) \right]$$

no off-diagonal correlations
Equation for a representative species:

$$\dot{x}(t) = x(t) \left[1 - x(t) + \mu M(t) + \eta(t) \right]$$

$$M(t) = \langle x(t) \rangle$$

$$\langle \prod_{k=1}^{r} \eta(t_k) \rangle_C = i^r b_r r! \langle \prod_{k=1}^{r} x(t_k) \rangle$$

— non-Gaussian coloured noise
— self-consistent *n*-point temporal correlations

where r=1,2,... and b_r is the coefficient of the series

$$F(z) = \sum_{r>0} b_r z^r$$

Assuming that there exists a fixed point in the GLV and in the DMFT eq., i.e., $\lim_{t\to\infty} x(t) = x^*$

$$0 = x^* \left[1 - x^* + \mu M + \eta^* \right]$$

When [...]>0, the stable solution is

$$x^*(\eta^*) = 1 + \mu M + \eta^*$$

and η^* is a (self-consistent) stationary noise that satisfies ($\mu = 0$):

$$P_{\eta}(\eta^*) = \int_{\mathbb{R}} \frac{dz}{2\pi} \exp\left\{iz\eta^* + \int_{-1}^{\infty} d\eta' P_{\eta}(\eta')F(z+z\eta')\right\}$$

The modified GLV equation ($\beta > 1$)

$$\dot{x}_i(t) = x_i(t) \left(1 - x_i^\beta + \sum_{j \neq i}^N \alpha_{ij} x_j(t) \right)$$

 $\alpha_{ij} \sim \alpha$ -stable distribution with exponent α (i.i.d, $i \neq j$) the ch.f. is $\exp(-\gamma |z|^{\alpha})$

The modified GLV equation ($\beta > 1$)

$$\dot{x}_{i}(t) = x_{i}(t) \left(1 - x_{i}^{\beta} + \sum_{j \neq i}^{N} \alpha_{ij} x_{j}(t) \right) \qquad \begin{array}{l} \alpha_{ij} \sim \alpha \text{-stable distribution} \\ \text{with exponent } \alpha \text{ (i.i.d, } i \neq j) \\ \text{the ch.f. is } \exp(-\gamma |z|^{\alpha}) \end{array}$$

The new DMFT predicts (*exact solution*) a distribution of surviving populations:



In the sparse case of the GLV system, we assume that

$$P_N(\alpha) = (1 - \frac{c}{N})\delta(\alpha) + \frac{c}{N}\hat{Q}(\alpha)$$

a species is connected to other *c* species on average and $Q(\alpha)$ is the distribution of the non-zero interactions;

In the sparse case of the GLV system, we assume that

$$P_N(\alpha) = (1 - \frac{c}{N})\delta(\alpha) + \frac{c}{N}\hat{Q}(\alpha)$$

a species is connected to other *c* species on average and $Q(\alpha)$ is the distribution of the non-zero interactions;

when *c* is small and x>1 (1 is the carrying capacity) we obtain

 $P_+(x) \propto \mathcal{Q}(x-1)$

we can statistically infer (some) interactions from a macroscopic pattern

L. Poley et al., *https://arxiv.org/abs/2404.08600; J. II Park et al.*, *https://arxiv.org/abs/2403.15730; F. Aguirre-Lopez, https://arxiv.org/abs/2404.11164; F. Metz, https://arxiv.org/abs/2406.06346*

Brief summary:



- The distribution of *species abundances* is not a Gaussian; shapes are much more realistic
- The distribution of *species abundances*
- depends on all cumulants;
- It can be a power law or other (more realistic) curves;
- There is a link between *microscopic interactions* and *macroscopic patterns* (SAD);
- gDMFT allows to introduce sparse interactions;

Brief summary:

8



The distribution of <i>species abundances</i> is not
a Gaussian; shapes are much more realistic
The distribution of <i>species abundances</i>
depends on all cumulants;
It can be a power law or other (more
realistic) curves;
There is a link between <i>microscopic</i>
interactions and macroscopic patterns (SAD);
gDMFT allows to introduce sparse
interactions;

The phase diagram is more complicated

Brief summary:



The distribution of <i>species abundances</i> is not
a Gaussian; shapes are much more realistic
The distribution of <i>species abundances</i>
depends on all cumulants;
It can be a power law or other (more
realistic) curves;
There is a link between <i>microscopic</i>
interactions and macroscopic patterns (SAD);
gDMFT allows to introduce sparse
interactions;
The phase diagram is more complicated

S.A. & A. Maritan, Phys. Rev. Lett. (2024)



doi:10.1038/nature25504

Fluctuating interaction network and time-varying stability of a natural fish community

Masayuki Ushio^{1,2,3,4}, Chih-hao Hsieh^{5,6,7}, Reiji Masuda⁸, Ethan R Deyle⁹, Hao Ye^{9,10}, Chun-Wei Chang⁶, George Sugihara⁹ & Michio Kondoh¹

LETTER

doi:10.1038/nature25504

Fluctuating interaction network and time-varying stability of a natural fish community

Masayuki Ushio^{1,2,3,4}, Chih-hao Hsieh^{5,6,7}, Reiji Masuda⁸, Ethan R Deyle⁹, Hao Ye^{9,10}, Chun-Wei Chang⁶, George Sugihara⁹ & Michio Kondoh¹



What happens if we introduce a simple stochastic time dependence in the species' interactions?

$$\dot{x}_i(t) = x_i(t) \left(r_i + \sum_{j=1}^N \alpha_{ij}(t) x_j(t) \right)$$

for simplicity $r_i = 1$ and $\alpha_{ii} = -1$.

What happens if we introduce a simple stochastic time dependence in the species' interactions?

$$\dot{x}_i(t) = x_i(t) \left(r_i + \sum_{j=1}^N \alpha_{ij}(t) x_j(t) \right)$$

interaction of

for simplicity $r_i = 1$ and $\alpha_{ii} = -1$.

couplings (annealed noise):
$$\alpha_{ij}(t) = \frac{\mu}{N} + \frac{\sigma}{\sqrt{N}} z_{ij}(t)$$

where $\langle z_{ij}(t) \rangle = 0$ $\langle z_{ij}(t) z_{lm}(t') \rangle = \delta_{i,k} \delta_{j,l} \frac{1 + 2\tau/\tau_0}{2\tau} e^{-\Delta t/\tau}$

11

5

What happens if we introduce a simple stochastic time dependence in the species' interactions?

$$\dot{x}_i(t) = x_i(t) \left(r_i + \sum_{j=1}^N \alpha_{ij}(t) x_j(t) \right)$$

interaction couplings (annealed noise):

where

for simplicity $r_i = 1$ and $\alpha_{ii} = -1$.

(annealed noise):
$$\alpha_{ij}(t) = \frac{\mu}{N} + \frac{\sigma}{\sqrt{N}} z_{ij}(t)$$

 $\langle z_{ij}(t) \rangle = 0 \quad \langle z_{ij}(t) z_{lm}(t') \rangle = \delta_{i,k} \delta_{j,l} \frac{1 + 2\tau/\tau_0}{2\tau} e^{-\Delta t/\tau}$



What happens if we introduce a simple stochastic time dependence in the species' interactions?

$$\dot{x}_i(t) = x_i(t) \left(r_i + \sum_{j=1}^N \alpha_{ij}(t) x_j(t) \right)$$

interaction couplings (annealed noise):

for simplicity
$$r_i = 1$$
 and $\alpha_{ii} = -1$.

$$\alpha_{ij}(t) = \frac{\mu}{N} + \frac{\sigma}{\sqrt{N}} z_{ij}(t)$$

where
$$\langle z_{ij}(t) \rangle = 0$$
 $\langle z_{ij}(t) z_{lm}(t') \rangle = \delta_{i,k} \delta_{j,l} \delta(t - t')$



four time series from the simulation of GLV with *N*=30, $\mu = 0$ and $\sigma = 1$

... no extinctions!



We can apply standard DMFT techniques for obtaining the mean field equation:

$$\dot{x}(t) = x(t) \left[1 - x(t) + \mu M(t) + \eta(t) \right]$$

$$M(t) = \langle x(t) \rangle \qquad \langle \eta(t)\eta(t') \rangle = \sigma^2 \frac{1 + 2\tau/\tau_0}{2\tau} e^{-\Delta t/\tau} \langle x(t)x(t') \rangle$$

We can apply standard DMFT techniques for obtaining the mean field equation:

$$\dot{x}(t) = x(t) \left[1 - x(t) + \mu M(t) + \eta(t) \right]$$

$$M(t) = \langle x(t) \rangle \qquad \langle \eta(t)\eta(t') \rangle = \sigma^2 \frac{1 + 2\tau/\tau_0}{2\tau} e^{-\Delta t/\tau} \langle x(t)x(t') \rangle$$

at stationarity (with some approximations too):

$$P_{\tau}^{*}(x) = \frac{x^{-1+\delta_{\tau}}}{Z} \left(\frac{1}{\bar{\tau}} + x\right) e^{-\frac{x}{D} - \frac{\bar{\tau}}{2D}(x-\bar{x})^{2}}$$

Parameters are calculated self-consistently and depend on μ , τ and σ .



We can apply standard DMFT techniques for obtaining the mean field equation:

$$\dot{x}(t) = x(t) \left[1 - x(t) + \mu M(t) + \eta(t) \right]$$

 $M(t) = \langle x(t) \rangle$ $\langle \eta(t)\eta(t') \rangle = \sigma^2 \langle x(t)^2 \rangle \delta(t-t')$

We can apply standard DMFT techniques for obtaining the mean field equation:

$$\dot{x}(t) = x(t) \left[1 - x(t) + \mu M(t) + \eta(t) \right]$$
$$M(t) = \langle x(t) \rangle \qquad \langle \eta(t) \eta(t') \rangle = \sigma^2 \langle x(t)^2 \rangle \delta(t - t')$$

which leads to a *self-consistent* Fokker-Planck equation:

$$\dot{p}(x,t) = -\partial_x [x(1-x+\mu\langle x(t)\rangle)p(x,t)] + \frac{\sigma^2 \langle x(t)^2 \rangle}{2} \partial_x [x\partial_x xp(x,t)]$$

with stationary distribution (Gamma distribution)

$$p^*(x) = \frac{\beta^{\delta}}{\Gamma(\delta)} x^{-1+\delta} e^{-\beta x} \qquad \beta = \frac{2}{\sigma^2 \langle x^2 \rangle} \qquad \delta = \frac{2(1+\mu \langle x \rangle)}{\sigma^2 \langle x^2 \rangle}$$

$$p^*(x) = \frac{\beta^{\delta}}{\Gamma(\delta)} x^{-1+\delta} e^{-\beta x}$$

$$\beta = \frac{2}{\sigma^2 \langle x^2 \rangle} \qquad \delta = \frac{2(1 + \mu \langle x \rangle)}{\sigma^2 \langle x^2 \rangle}$$



$$p^*(x) = \frac{\beta^{\delta}}{\Gamma(\delta)} x^{-1+\delta} e^{-\beta x}$$

1

$$\beta = \frac{2}{\sigma^2 \langle x^2 \rangle} \qquad \delta = \frac{2(1 + \mu \langle x \rangle)}{\sigma^2 \langle x^2 \rangle}$$

Empirical patterns

Gamma distributions describe well surveys of microbial communities, from Arctic oceans to zebras' guts, and also abundances of species in forest communities...



BCI & Caribbean **Coral Reef**

Xiao Y. et al., Nat.Com. (2017); Grilli J., Nat.Com. (2020); S.A. et al., Rev. Mod. Phys. (2016)

$$p^*(x) = \frac{\beta^{\delta}}{\Gamma(\delta)} x^{-1+\delta} e^{-\beta x}$$

1

$$\beta = \frac{2}{\sigma^2 \langle x^2 \rangle} \qquad \delta = \frac{2(1 + \mu \langle x \rangle)}{\sigma^2 \langle x^2 \rangle}$$

Empirical patterns

Gamma distributions describe well
surveys of microbial communities,
from Arctic oceans to zebras' guts,
and also abundances of species
in forest communities...



BCI & Caribbean Coral Reef

Xiao Y. et al., Nat.Com. (2017); Grilli J., Nat.Com. (2020); S.A. et al., Rev. Mod. Phys. (2016)





Egbert H. van Nes^{a,1} 🕩, Diego G. F. Pujoni^b, Sudarshan A. Shetty^c, Gerben Straatsma^a, Willem M. de Vos^{c,d} 🕩, and Marten Scheffer^{a,1,2} 🕩

Contributed by Marten Scheffer; received December 23, 2022; accepted November 11, 2023; reviewed by Tadashi Fukami and Jonathan M. Levine







Egbert H. van Nes^{a,1} 🕩, Diego G. F. Pujoni^b, Sudarshan A. Shetty^c, Gerben Straatsma^a, Willem M. de Vos^{c,d} 🕩, and Marten Scheffer^{a,1,2} 🕩

Contributed by Marten Scheffer; received December 23, 2022; accepted November 11, 2023; reviewed by Tadashi Fukami and Jonathan M. Levine







Egbert H. van Nes^{a,1} 🕩, Diego G. F. Pujoni^b, Sudarshan A. Shetty^c, Gerben Straatsma^a, Willem M. de Vos^{c,d} 🕩, and Marten Scheffer^{a,1,2} 🕩

Contributed by Marten Scheffer; received December 23, 2022; accepted November 11, 2023; reviewed by Tadashi Fukami and Jonathan M. Levine



The vast bulk of biodiversity is very rare: why?

Apparently, there are no traits that allow abundant species to escape rarity; superiority in competition for resources is not necessary for dominance!





Egbert H. van Nes^{a,1} , Diego G. F. Pujoni^b, Sudarshan A. Shetty^c, Gerben Straatsma^a, Willem M. de Vos^{c,d}, and Marten Scheffer^{a,1,2}

Contributed by Marten Scheffer; received December 23, 2022; accepted November 11, 2023; reviewed by Tadashi Fukami and Jonathan M. Levine

Some evidences:



- in the human microbiome *different* species can take dominant positions over time;
- dominance by a single species is *common but short-lived*;
- rare species may become dominant with *unpredictable alternations* both under controlled laboratory conditions and in natural communities;
- while the presence of a functional group is predictable, the identity of the dominant species is not;





Egbert H. van Nes^{a,1} , Diego G. F. Pujoni^b, Sudarshan A. Shetty^c, Gerben Straatsma^a, Willem M. de Vos^{c,d}, and Marten Scheffer^{a,1,2}

Contributed by Marten Scheffer; received December 23, 2022; accepted November 11, 2023; reviewed by Tadashi Fukami and Jonathan M. Levine

Some evidences:



- in the human microbiome *different* species can take dominant positions over time;
- dominance by a single species is *common but short-lived*;
- rare species may become dominant with *unpredictable alternations* both under controlled laboratory conditions and in natural communities;
- while the presence of a functional group is predictable, the identity of the dominant species is not;

Functional redundancy and near neutrality:



species within a functional group share a single niche and within this niche species are near neutral. This leads to *instability of dominance by functionally similar species; stochastic origin of hyper-dominance:* because fluctuations are proportional to abundances (multiplicative noise), *the rates of change of rare species are slow;* rare species may serve as an *insurance* to maintain function in times of adverse events, because of their different sensitivity to stressors and natural enemies; The MacArthur's consumer-resource model: a more realistic way to account for coexistence

pop density of species σ

biotic res.

$$\dot{n}_{\sigma} = n_{\sigma} \left[\sum_{i=1}^{R} \alpha_{\sigma i} c_{i} - \beta_{\sigma} \right]$$
$$\dot{c}_{i} = c_{i} \left[\mu_{i} (1 - c_{i}) - \sum_{\sigma=1}^{N} n_{\sigma} \alpha_{\sigma i} \right]$$

MacArthur, R. H., *Theor. Pop. Bio.*, **1**, 1-11, 1970

The MacArthur's consumer-resource model

max consumption rate of species σ on the *i*-th resource (metabolic strategy)

pop density of species σ

biotic res.

$$\dot{n}_{\sigma} = n_{\sigma} \left[\sum_{i=1}^{R} \alpha_{\sigma i} \dot{q}_{i} - \beta_{\sigma} \right] \rightarrow \begin{array}{l} \text{death rate} \\ \text{(maintenance} \\$$

The MacArthur's consumer-resource model

max consumption rate of species σ on the *i*-th resource (metabolic strategy)

pop density of species σ

biotic res.

$$\dot{n}_{\sigma} = n_{\sigma} \left[\sum_{i=1}^{R} \alpha_{\sigma i} c_{i} - \beta_{\sigma} \right]$$

$$\dot{c}_{i} = c_{i} \left[\mu_{i} (1 - c_{i}) - \sum_{\sigma=1}^{N} n_{\sigma} \alpha_{\sigma i} \right]$$
MacArthur, R. H., *Theor.*
Pop. Bio., 1, 1-11, 1970

logistic growth of the i-th resource

What's nice about MacArthur's model?

when μ_i are all very large, the dynamics of resources is very fast and we can use the *quasi-stationary assumption*, hence $c_i \simeq 1 - \sum_{\sigma} n_{\sigma} \alpha_{\sigma i} / \mu_i$. The dynamics of the densities can be recast as

$$\dot{n}_{\sigma} = n_{\sigma} \left[k_{\sigma} - \sum_{\rho} a_{\sigma\rho} n_{\rho} \right]$$

where

$$k_{\sigma} = \sum_{i} \alpha_{\sigma i} - \beta_{\sigma}$$
$$a_{\sigma \rho} = \sum_{i} \frac{\alpha_{\sigma i} \alpha_{\rho i}}{\mu_{i}}$$

Chesson P., Theor. Popul. Biol. 37, 26-38 (1990)

we *recover the GLV-equations* at leading order in the quasi-stationary regime, and species' interactions are a measure of the niche overlap.

What's nice about MacArthur's model?

when μ_i are all very large, the dynamics of resources is very fast and we can use the *quasi-stationary assumption*, hence $c_i \simeq 1 - \sum_{\sigma} n_{\sigma} \alpha_{\sigma i} / \mu_i$. The dynamics of the densities can be recast as

$$\dot{n}_{\sigma} = n_{\sigma} \left[k_{\sigma} - \sum_{\rho} a_{\sigma\rho} n_{\rho} \right]$$

where

species' direct interaction $k_{\sigma} = \sum_{i} \alpha_{\sigma i} - \beta_{\sigma}$ f resource i $a_{\sigma \rho} = \sum_{i} \frac{\alpha_{\sigma i} \alpha_{\rho i}}{\mu_{i}}$ Chesson P., Theor. Popul. Biol. 37, 26-38 (1990)

we *recover the GLV-equations* at leading order in the quasi-stationary regime, and species' interactions are a measure of the niche overlap.

... but "all that glitters is not gold"

MacArthur's model at stationarity provides *N* conditions for *R* concentrations

$$\sum_{i=1}^{R} \alpha_{\sigma i} c_i^* = \beta_{\sigma} \qquad \text{for } \sigma = 1, 2, \dots, N$$

This eq. cannot have a solution if N > R.

Competitive Exclusion Principle: The total number of coexisting species cannot exceed the total number of resources on which they feed (at stationarity).



This is not good news for large scale species coexistence! Do we really have one species per niche?



The MacArthur's consumer-resource model with time-dependent metabolic strategies

Now the metabolic strategies depend on time

$$\dot{n}_{\sigma} = n_{\sigma} \left[\sum_{i=1}^{R} \alpha_{\sigma i}(t) c_{i} - \beta_{\sigma} \right]$$
$$\dot{c}_{i} = c_{i} \left[\mu_{i}(1 - c_{i}) - \sum_{\sigma=1}^{N} n_{\sigma} \alpha_{\sigma i}(t) \right]$$

A.R. Batista-Tomàs *et al., Chaos* **31**, 103113 (2021)

The MacArthur's consumer-resource model with time-dependent metabolic strategies

Now the metabolic strategies depend on time

$$\begin{split} \dot{n}_{\sigma} &= n_{\sigma} \left[\sum_{i=1}^{R} \alpha_{\sigma i}(t) c_{i} - \beta_{\sigma} \right] \\ \dot{c}_{i} &= c_{i} \left[\mu_{i}(1 - c_{i}) - \sum_{\sigma=1}^{N} n_{\sigma} \alpha_{\sigma i}(t) \right] \end{split}$$

A.R. Batista-Tomàs *et al.*, *Chaos* **31**, 103113 (2021)

$$\alpha_{\sigma i}(t) = \frac{\bar{\alpha}}{N} + \frac{\Sigma}{\sqrt{N}} Z_{\sigma}^{i}(t)$$
$$\langle Z_{\sigma}^{i}(t) Z_{\sigma'}^{i'}(t') \rangle = \frac{1 + 2\tau/\tau_{0}}{2\tau} e^{-|t-t'|/\tau} \delta_{\sigma,\sigma'} \delta_{i,i'}$$

Now you take the limit $N \rightarrow \infty$ and $R \rightarrow \infty$ such that $\nu = \frac{N}{R}$ is constant and apply DMFT...
$$\dot{n}_{\sigma} = n_{\sigma} \left[\sum_{i=1}^{R} \alpha_{\sigma i}(t) c_{i} - \beta \right]$$
$$\dot{c}_{i} = c_{i} \left[\mu_{i}(1 - c_{i}) - \sum_{\sigma=1}^{N} n_{\sigma} \alpha_{\sigma i}(t) \right]$$

$$\dot{n}(t) = n(t) \left[\frac{\bar{\alpha}\rho_c(t)}{\nu} - \beta + \frac{\Sigma}{\sqrt{\nu}} \xi_n(t) + \frac{\Sigma^2}{\nu} \int_0^t dt' \ q(|t - t'|) \ G_c(t, t')n(t') + h_n(t) \right]$$
$$\dot{c}(t) = c(t) \left[\mu [1 - c(t)] - \bar{\alpha}\rho_n - \Sigma \xi_c(t) - \Sigma^2 \int_0^t dt' \ q(|t - t'|)G_n(t, t')c(t') + h_c(t) \right]$$

where

$$\nu = \lim \frac{N}{R} \qquad \frac{\langle \xi_n(t)\xi_n(t')\rangle = q(|t-t'|)\langle c(t)c(t')\rangle_{paths}}{\langle \xi_c(t)\xi_c(t')\rangle = q(|t-t'|)\langle n(t)n(t')\rangle_{paths}} \qquad \rho_n(t) = \langle n(t)\rangle_{paths}} \\ \rho_c(t) = \langle c(t)\rangle_{paths}$$

When we take the white noise limit ($\tau \rightarrow 0$), we obtain

$$P(n) = Z^{-1}n^{-1+\delta_n}e^{-n/\beta_n}$$

where

$$\delta_n = \frac{2\bar{\alpha}\langle c \rangle - 2\nu\delta}{\Sigma^2 \langle c^2 \rangle} \qquad \beta_n = \frac{\langle c^2 \rangle}{\langle c \rangle}$$

When we take the white noise limit ($\tau \rightarrow 0$), we obtain

$$P(n) = Z^{-1}n^{-1+\delta_n}e^{-n/\beta_n}$$

where

$$\delta_n = \frac{2\bar{\alpha}\langle c \rangle - 2\nu\delta}{\Sigma^2 \langle c^2 \rangle} \qquad \beta_n = \frac{\langle c^2 \rangle}{\langle c \rangle}$$

so we can violate the CEP when $\nu > 1$

$$\langle c \rangle > \frac{\nu \delta}{\bar{\alpha}} \qquad \langle n \rangle < \frac{\mu}{\bar{\alpha}}$$

When we take the white noise limit ($\tau \rightarrow 0$), we obtain

$$P(n) = Z^{-1}n^{-1+\delta_n}e^{-n/\beta_n}$$

comparison simulations and theory: N=200, R=10



A final summary



- * Species abundances can be explained with heterogeneous yet structureless interactions;
- time-varying interactions help coexistence in GLV;
- * time-varying interactions help violating the Competitive Exclusion Principle in CR.

LIPh Lab



LIPh Lab



Alternative ways of explaining coexistence

Classical explanations of coexistence (neo-Darwinism):

- the ecological niche is the range of environmental conditions that allow a population to persist in a given location — i.e., the growth rate is nonnegative (G. H. Hutchinson, 1958-59)
- the split of resources is an outcome of the co-evolution of species in competition (J. Connell, 1975)



D'Andrea et al., Ecology, 2020