COMPLEX NETWORKS: STRUCTURE AND FUNCTIONALTY

II. Equivalence

Frank den Hollander
Mathematical Institute, Leiden University

12th MSJ-SI,
Fukuoka, Japan, 31/07–09/08, 2019.

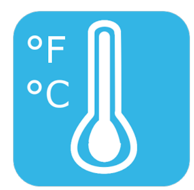
§ STATISTICAL PHYSICS

Systems consisting of a very large number of interacting particles can be described by statistical ensembles, i.e., probability distributions on spaces of configurations.

Two important examples are:

I.	micro-canonical ensemble
II.	canonical ensemble

The former fixes the energy of the system, the latter fixes the average energy of the system, with temperature as the control parameter.

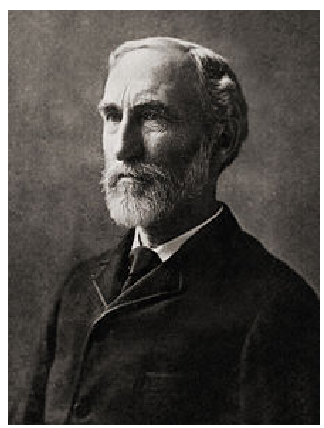


The two ensembles capture physically different microscopic situations. For both the entropy is maximal subject to the constraint.

The canonical ensemble is easier to compute with than the micro-canonical ensemble, because the constraint is soft rather than hard.

In textbooks of statistical physics the two ensembles are assumed (!) to be thermodynamically equivalent, i.e., to have the same macroscopic behaviour. Here the idea is that for large systems the energy is typically close to its average value.

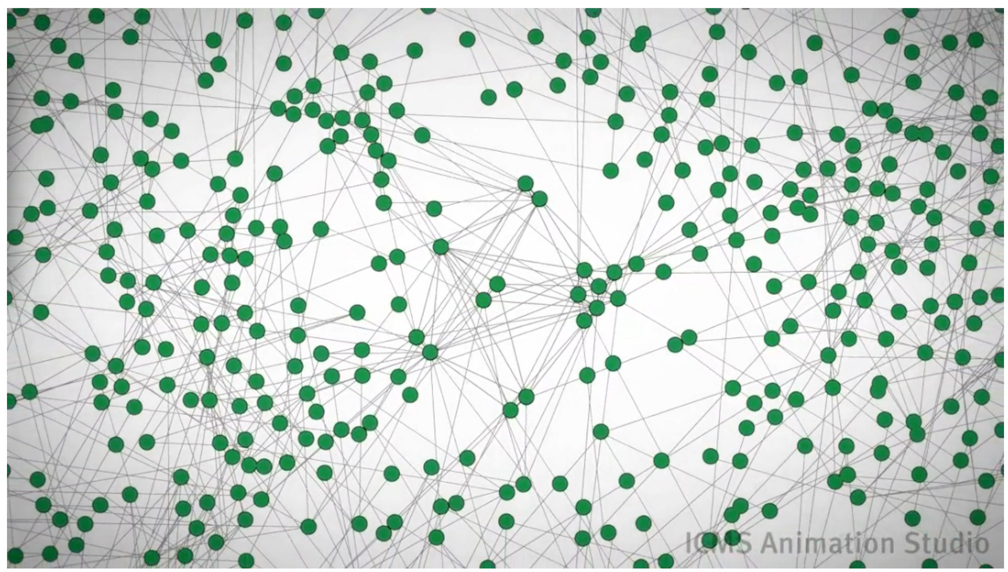
This assumption is certainly reasonable for systems with interactions that are short-ranged. But, counterexamples have been found for systems with interactions that are long-ranged.



Gibbs

§ COMPLEX NETWORKS

In this talk we will be interested in large random graphs, i.e., the two ensembles live on the set $\mathcal{S}_N$ of all simple graphs with N vertices where $N \to \infty$.



A realisation of a large random graph

DEFINITIONS

Given are a vector-valued function $\vec{C}$ on $\mathcal{S}_N$, and a specific vector $\vec{C}^*$ called the constraint.

I. The micro-canonical ensemble is defined by

$$P_N^{\text{mic}}(G) = \begin{cases} 1/\Omega_{\vec{C}^*} & \text{if } \vec{C}(G) = \vec{C}^*, \\ 0 & \text{else,} \end{cases}$$

where $\Omega_{\vec{C}^*} = |\{G \in \mathcal{S}_N : \vec{C}(G) = \vec{C}^*\}|$.

II. The canonical ensemble is defined by

$$P_N^{\text{can}}(G) = \frac{1}{\mathcal{N}(\vec{\theta}^*)} \, \mathrm{e}^{-\vec{\theta}^* \cdot \vec{C}(G)},$$

where $\mathcal{N}(\vec{\theta}^*)$ is the normalising constant and $\vec{\theta}^*$ is to be chosen such that $\sum_{G \in \mathcal{S}_N} \vec{C}(G) P_N^{\text{can}}(G) = \vec{C}^*$.

INTERPRETATION

- P_N^{mic} models a random graph of which no information is available other than the constraint.
- P_N^{can} models a random graph of which no information is available other than the average constraint.

Which of the two should be used to model a real-world network depends on the a priori knowledge that is available about the network.

ENSEMBLE EQUIVALENCE   Touchette 2015

P_N^{mic} and P_N^{can} are said to be equivalent when their relative entropy per vertex defined by

$$s_N\left(P_N^{\text{mic}} \mid P_N^{\text{can}}\right) = \frac{1}{N} \sum_{G \in \mathcal{S}_N} P_N^{\text{mic}}(G) \log\left(\frac{P_N^{\text{mic}}(G)}{P_N^{\text{can}}(G)}\right)$$

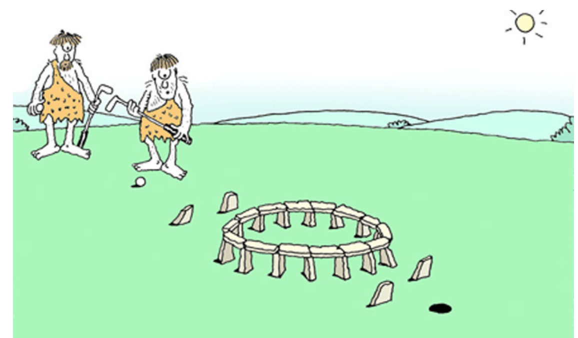
tends to zero as $N \to \infty$.

Because in both ensembles all $G \in \mathcal{S}_N$ such that $\vec{C}(G) = \vec{C}^*$ have the same probability, we get the simpler formula

$$s_N\left(P_N^{\text{mic}} \mid P_N^{\text{can}}\right) = \frac{1}{N}\log\left(\frac{P_N^{\text{mic}}(G^*)}{P_N^{\text{can}}(G^*)}\right)$$

for any G^* such that $\vec{C}(G^*) = \vec{C}^*$. This greatly simplifies the computation, since we need not carry out the sum over $\mathcal{S}_N$ and only need to compute with a single graph G^*.

In the remainder of the lecture we illustrate breaking of ensemble equivalence via a number of examples.

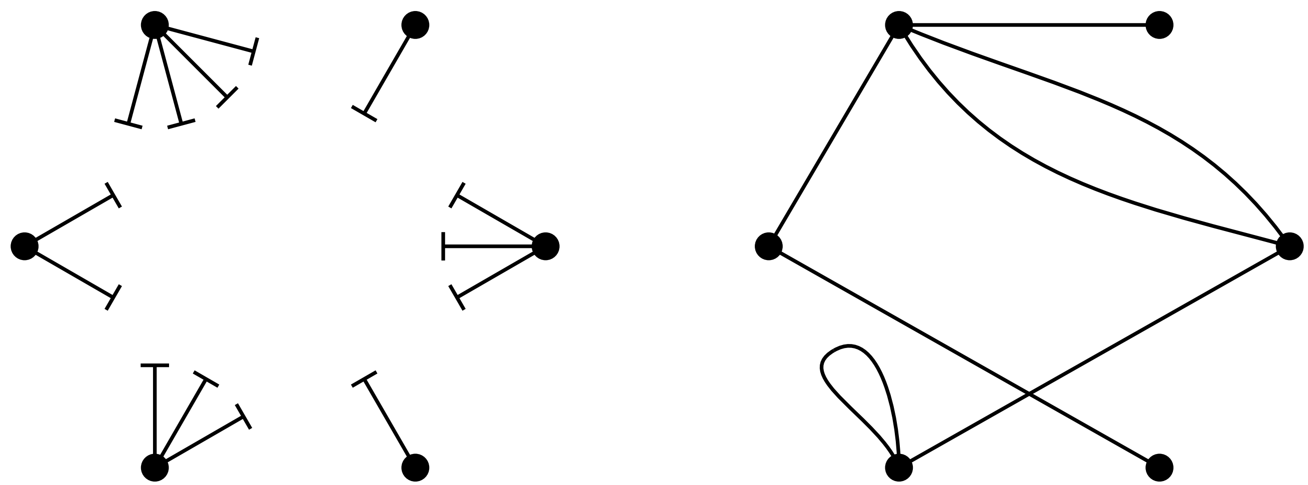


What follows is joint work with

Diego Garlaschelli	(Leiden & Lucca)
Michel Mandjes	(Amsterdam)
Andrea Roccaverde	(Leiden)
Tiziano Squartini	(Lucca)
Nicos Starreveld	(Amsterdam)

§ CONSTRAINT ON THE DEGREE SEQUENCE

Each vertex gets a prescribed number of half-edges, which are paired off randomly to form edges.



Example with $N = 6$ and $\vec{d}_N = (1, 3, 1, 3, 2, 4)$

Consider a graph $G = (V, E)$ with vertex set $V = \{1, \ldots, N\}$ and edge set E such that all the vertices have prescribed degrees. In other words, consider the constraint

$$\vec{C}^* = \vec{d}_N^* = (d_1^*, \ldots, d_N^*) \in \mathbb{N}_0^N.$$

Suppose that the degrees are moderate, corresponding to what is called the sparse regime:

$$\max_{1 \leq i \leq N} d_i^* = o(\sqrt{N}), \qquad N \to \infty.$$

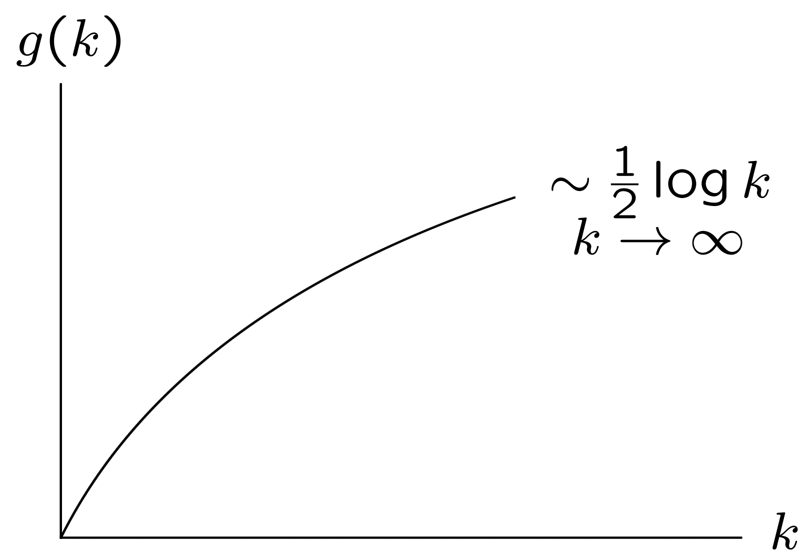
sparseness

Let

$$f_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{d_i^*} = \text{empirical degree distribution}.$$

Define

$$g(k) = \log\left(\frac{k!}{k^k e^{-k}}\right), \qquad k \in \mathbb{N}_0.$$



THEOREM 1:

Suppose that

$$\lim_{N\to\infty} \|f_N - f\|_{\ell^1(g)} = 0$$

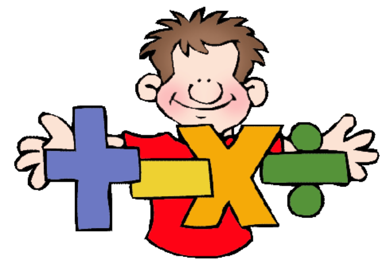
for some limiting degree distribution f. Then

$$s_\infty = \lim_{N\to\infty} s_N\left(P_N^{\text{mic}} \mid P_N^{\text{can}}\right) = \|f\|_{\ell^1(g)}.$$

Interpretation: Each vertex with degree k contributes an amount $g(k)$ to the relative entropy.

There is breaking of ensemble equivalence for all $f \neq \delta_0$.

The proof is based on graph counting (micro-canonical) and percolation theory (canonical).



It turns out that $g(k)$ is the relative entropy of Dirac(k) with respect to Poisson(k). What this says is that, in the limit as $N \to \infty$,

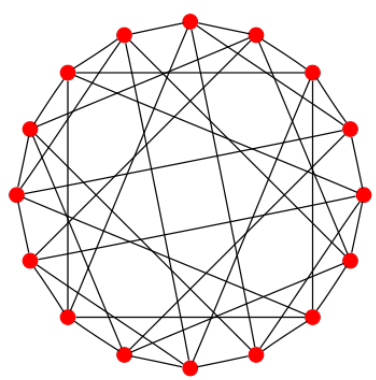
- Micro-canonical ensemble:
  vertices have a fixed degree.
- Canonical ensemble:
  vertices have a random degree.

Example 1:

$f_N = \delta_k$ with $k = o(\sqrt{N})$.

For k-regular graphs:

$$s_\infty = g(k) > 0.$$



5-regular graph

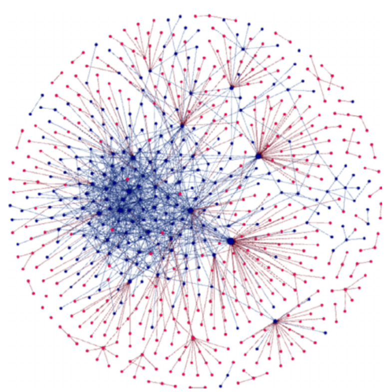
Example 2:

$f_N(k) = C_N \, k^{-\tau}$, $1 \le k \le k_{\text{cutoff}}(N)$,

with $k_{\text{cutoff}}(N) = o(\sqrt{N})$ and $\tau \in (1, \infty)$ a tail exponent.

For scale-free graphs:

$$s_\infty \approx \frac{1}{2(\tau - 1)} + \tfrac{1}{2} \log(2\pi) > 0.$$



graph with hubs

§ CONSTRAINT ON THE TOTAL NUMBER OF EDGES AND TRIANGLES

Interesting behaviour shows up when we pick

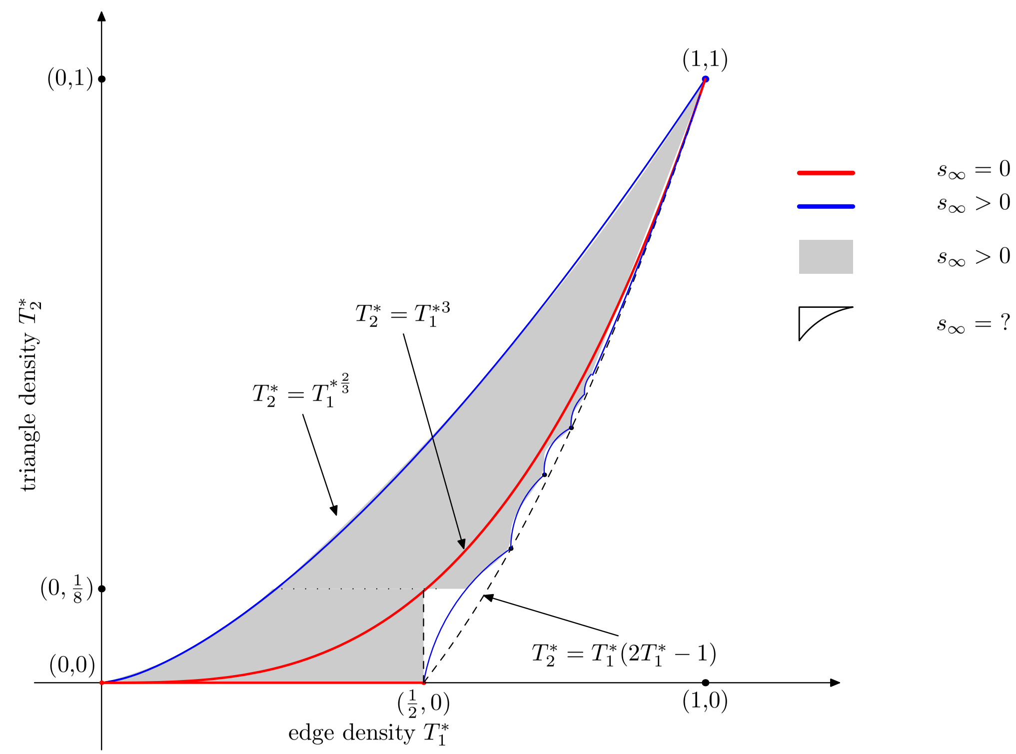
$$\vec{C}^* = (\text{number of edges}, \text{number of triangles})$$
$$= \left(T_1^* \binom{N}{2}, T_2^* \binom{N}{3}\right), \qquad T_1^*, T_2^* \in [0,1].$$

This corresponds to the so-called dense regime, in which the number of edges per vertex is of order N. The quantity of interest is now

$$s_\infty = \lim_{N\to\infty} \frac{1}{N^2} \log \left(\frac{P_N^{\text{mic}}(G^*)}{P_N^{\text{can}}(G^*)}\right),$$

where we scale by N^2 instead of N.

THEOREM 2:



Between the blue curves the edge-triangle densities are admissible.

Radin and Sadun 2015

Breaking of ensemble equivalence occurs when (T_1^*, T_2^*) is frustrated.

The proof is based on the theory of graphons, which are continuum limits of adjacency matrices of graphs.

Borgs, Chayes, Lovász $\geq$ 2008

We derive a variational formula for s_∞ with the help of the large deviation principle for graphons associated with the Erdős-Rényi random graph.

Chatterjee and Varadhan 2011

What happens close the line $T_2^* = T_1^{*3}$? It turns out that anomalous behaviour shows up:

THEOREM 3:

For $T_1^* \in (0,1)$,

$$\lim_{\epsilon \downarrow 0} \epsilon^{-1} s_\infty(T_1^*, T_1^{*3} + \epsilon) = C^+ \in (0,\infty),$$
$$\lim_{\epsilon \downarrow 0} \epsilon^{-2/3} s_\infty(T_1^*, T_1^{*3} - \epsilon) = C^- \in (0,\infty),$$

where C^+, C^- *are computable functions of* T_1^* *that are, however, not so easy to identify.*

§ CONCLUSION

We have obtained a complete classification of breaking of ensemble equivalence in random graphs with constraints on the degree sequence, respectively, the total number of edges and triangles.

Breaking occurs when the number of constraints is extensive or when the constraints are frustrated.

§ LITERATURE

D. Garlaschelli, F. den Hollander, J. de Mol, T. Squartini,
Phys. Rev. Lett. 115 (2015) 268701.

D. Garlaschelli, F. den Hollander, A. Roccaverde,
Nieuw Archief voor Wiskunde 5/16 (2015) 207–209.

D. Garlaschelli, F. den Hollander, A. Roccaverde,
J. Phys. A: Math. Theor. 50 (2017) 015001.

F. den Hollander, M. Mandjes, A. Roccaverde, N.J. Starreveld,
Electronic J. Prob. 2018.

D. Garlaschelli, F. den Hollander, A. Roccaverde,
J. Stat. Phys. 2018.

F. den Hollander, M. Mandjes, A. Roccaverde, N.J. Starreveld,
Random Struct. Alg. 2019.


Thank You