"GENERAL" MODEL

$$\pi_n(g) = \frac{\exp\left\{ n^2 \left[\frac{\alpha}{n^2} \sum_i \sum_j g_{ij} + \frac{\beta}{2n^2} \sum_i \sum_j g_{ij} g_{ji} + \frac{\gamma}{n^3} \sum_i \sum_j \sum_{k \neq i,j} g_{ij} g_{jk} \right] \right\}}{c(\alpha, \beta, \gamma, \mathcal{G}_n)}$$

$\frac{1}{n^2} \sum_i \sum_j g_{ij}$ = direct link density

$\frac{1}{n^2} \sum_i \sum_j g_{ij} g_{ji}$ = reciprocity density

$\frac{1}{n^2} \sum_i \sum_j \sum_{k \neq i,j} g_{ij} g_{jk}$ = density of directed two-paths

$\longleftarrow$ $c(\alpha, \beta, \gamma, \mathcal{G}_n)$

- Connection with the exponential family in statistics
- Smooth behaviour if $c(\cdot)$ is analytic in $\theta = [\alpha, \beta, \gamma]'$
- unstable or phase transition if $c(\cdot)$ is not analytic in θ.

CONVERGENT GRAPH SEQUENCES

H, G simple graphs (~~undirected~~, unweighted, no loops, no multiple edges)

Homomorphism: arc-preserving map from $V(H)$ to $V(G)$

$$t(H,G) = \frac{|\hom(H,G)|}{|V(G)|^{|V(H)|}}$$

homomorphism density

(= prob. that a random mapping is a homomorphism)

Let $G_n \to G$ as $n \to \infty$, then for any simple H

$$t(H, G_n) \to t(H, G)$$

Need a graphon!

GRAPHONS

Let $h \in W$, W = set of all measurable (~~symmetric~~) functions from $[0,1]^2$ to $[0,1]$

Interpretation

h represents a "continuum" graph, i.e. a graph s.t. $|V(G)|$ is non denumerable

- $(X,Y) \sim \text{Uniform}([0,1]^2)$
- $h(x,y)$ = (conditional) probability that x and y are connected

Let H be s.t. $V(H) = \{1, 2, \dots, k\}$

$$t(H,h) = \int_{[0,1]^k} \prod_{(i,j)\in E(H)} h(x_i, x_j)\, dx_1, dx_2, \dots, dx_k$$ = probability of homomorphism from H to $G = h$

Let $V(H) = \{1, 2\}$, then

$$t(H, h) = \int_{[0,1]^2} h(x_1, x_2)\, dx_1\, dx_2$$

Any finite simple graph can be represented as a graphon.

Let $V(G_n) = \{1, 2, \dots, n\}$

$$f^G(x,y) = \begin{cases} 1 & \text{if } (\lceil nx \rceil, \lceil ny \rceil) \in E(G) \\ 0 & \text{otherwise} \end{cases}$$

EXCHANGEABILITY

X = adjacency matrix

Exchangeability $\iff$ $(X_{ij}) \overset{d}{=} (X_{\sigma(i)\sigma(j)})$

Notice that if node permutations are equally likely, then any permutation is a measure preserving transformation

CUT DISTANCE

Let $f, g \in \mathcal{W}$

$$d_\square(f,g) = \sup_{S,T \subseteq [0,1]} \left| \int_{S\times T} f(x,y) - g(x,y) \, dx \, dy \right|$$

Let Σ be the space of all the measure preserving bijections

$$\sigma : [0,1] \to [0,1] \quad \text{s.t.} \quad \mu(A) = \mu(\sigma^{-1} A)$$

$f(x,y) \sim g(x,y)$ if $f(x,y) = g_\sigma(x,y) := g(\sigma x, \sigma y)$ for some $\sigma \in \Sigma$

$\tilde{g}$ represents the closure of $\{\tilde{g}\}$ in $(W, d_\square)$

$\tilde{W}$ denotes the quotient space and $\tau: g \to \tilde{g}$

We can define a natural distance on $\tilde{W}$:

$$\delta_\square(\tilde{f}, \tilde{g}) := \inf_\sigma d_\square(f, g_\sigma) = \inf_\sigma d_\square(f_\sigma, g) = \inf_{\sigma_1, \sigma_2} d_\square(f_{\sigma_1}, g_{\sigma_2})$$

$\tilde{W}$ is a compact space

$t(H, \cdot)$ is continuous for any finite H

LARGE DEVIATIONS

let $$I_p(u) := \frac{1}{2} u \log \frac{u}{p} + \frac{1}{2}(1-u) \log \frac{1-u}{1-p}$$

and

$$I_p(h) = \int_0^1 \int_0^1 I_p\big(h(x,y)\big)\, dx, dy, \qquad h \in W$$

$$I_p(\tilde{h}) := I_p(h), \quad \tilde{h} \in \tilde{W} \quad \text{and } h \sim \tilde{h}$$

Erdös Renyi $G(n,p)$ induces $\mathbb{P}_{n,p}$ measure on W through the map $G \to f^G$, hence

$$\tilde{\mathbb{P}}_{n,p} \text{ on } \tilde{W} \text{ through } G \to f^G \to \tilde{f}^G = \tilde{G}$$

LARGE DEVIATION PRINCIPLE FOR $\tilde{\mathbb{P}}_{n,p}$ ON $(\tilde{\mathcal{W}}, \delta_\square)$

For each fixed $p \in (0,1)$, the sequence $\tilde{\mathbb{P}}_{n,p}$ obeys a large deviation principle in the space $\tilde{\mathcal{W}}$ (equipped with the cut metric) with rate function I_p. This means that, for any closed set $\tilde{F} \subseteq \tilde{\mathcal{W}}$

$$\limsup_{n\to\infty} \frac{1}{n^2} \log \tilde{\mathbb{P}}_{n,p}(\tilde{F}) \le -\inf_{\tilde{h}\in\tilde{F}} I_p(\tilde{h})$$

and for any open set $\tilde{U} \subseteq \tilde{\mathcal{W}}$

$$\liminf_{n\to\infty} \frac{1}{n^2} \log \tilde{\mathbb{P}}_{n,p}(\tilde{U}) \ge -\inf_{\tilde{h}\in\tilde{U}} I_p(\tilde{h})$$

MAIN RESULT

Let $T : \widetilde{W} \to \mathbb{R}$ a bounded continuous function on $(\widetilde{W}, \delta_\square)$

$\mathcal{G}_n$ = set of simple graphs with n vertices

$$P_n(G) := e^{n^2 (T(\widetilde{G}) - \psi_n)} \qquad \widetilde{G} \in \widetilde{W}$$

$$\psi_n = \frac{1}{n^2} \log \sum_{G \in \mathcal{G}_n} e^{n^2 T(\widetilde{G})}$$

Let $I(u) = \frac{1}{2} u \log(u) + \frac{1}{2}(1-u) \log(1-u)$

and

$$I(\widetilde{h}) = \int_0^1 \int_0^1 I(h(x,y)) \, dx \, dy \qquad h \sim \widetilde{h}$$

Then

$$\psi := \lim_{n\to\infty} \psi_n = \sup_{\tilde{h}\in\tilde{W}} \left(T(\tilde{h}) - I(\tilde{h}) \right)$$

The maximiser might be

- a scalar $\Rightarrow$ ψ analytic function of the parameters

  asymptotically E-R graph

- scalar, but not unique

  still E-R asymptotically but with phase transition
  small variations in the parameters determine radically different behaviours (from nearly complete to sparse graphs)

- a function

  no more E-R asymptotically