Method of cumulants and mod-Gaussian convergence of the graphon models

Pierre-Loïc Méliot

(Joint work with Valentin Féray and Ashkan Nikeghbali)

2017, May 11th

University Paris-Sud

When looking at a sum $S_n = \sum_{i=1}^{n} A_i$ of centered i.i.d. random variables, the fluctuations are universally predicted by the **central limit theorem**

$$\frac{S_n}{\sqrt{n\,\mathrm{Var}(A_1)}} \rightharpoonup \mathcal{N}(0,1).$$

This is not the whole story:

- **large deviations** (Cramér, 1938): $\log\left(\mathbb{P}[S_n \geq nx]\right) \simeq -n\,I(x)$.
- **speed of convergence** (Berry, 1941; Esseen, 1945):

$$\sup_{s\in\mathbb{R}} \left| \mathbb{P}\left[\frac{S_n}{\sqrt{n\,\mathrm{Var}(A_1)}} \leq s\right] - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{s} \mathrm{e}^{-\frac{t^2}{2}}\,dt \right| \leq \frac{3\,\mathbb{E}[|A_1|^3]}{(\mathrm{Var}(A_1))^{3/2}\sqrt{n}}.$$

- **local limit theorem** (Gnedenko, 1948; Stone, 1965): if A_1 is non-lattice distributed and $\mathrm{Var}(A_1) = 1$, then

$$\sqrt{n}\;\mathbb{P}\big[S_n \in (\sqrt{n}x, \sqrt{n}x + h)\big] \simeq \frac{\mathrm{e}^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\,h.$$

Many other sequences of random variables are asymptotically normal: functionals of Markov chains, martingales, *etc.*

Idea: there is a renormalisation theory of random variables that allows one to go beyond the central limit theorem, and to prove in one time the CLT *and* the other limiting results.

Definition (Mod-Gaussian convergence)

A sequence of real random variables $(X_n)_{n\in\mathbb{N}}$ is mod-Gaussian with parameters $t_n \to +\infty$ and limit $\psi(z)$ if, locally uniformly on a domain $D \subset \mathbb{C}$,

$$\mathbb{E}[\mathrm{e}^{zX_n}]\,\mathrm{e}^{-\frac{t_n z^2}{2}} = \psi_n(z) \to \psi(z)$$

with ψ continuous on D and $\psi(0) = 1$.

For a sum of i.i.d. S_n, one looks at $X_n = \frac{S_n}{n^{1/3}}$; $t_n = n^{1/3}\,\mathrm{Var}(A_1)$ and $\psi(z) = \exp(\frac{\mathbb{E}[(A_1)^3]\,z^3}{6})$.

Example: let $X_n = \mathrm{Re}(\log\det(I_n - M_n))$, with $M_n \sim \mathrm{Haar}(\mathrm{U}(n))$. One has the mod-Gaussian convergence

$$\mathbb{E}[\mathrm{e}^{zX_n}]\,\mathrm{e}^{-\frac{(\log n)\,z^2}{4}} \to \frac{G(1+\frac{z}{2})^2}{G(1+z)}, \quad G = \text{ Barnes' function.}$$

Later: Markov chains, random graphs, random permutations, *etc.*

Remark: one can replace the exponent $\frac{z^2}{2}$ of the Gaussian distribution by the exponent $\eta(z)$ of any infinitely divisible distribution.

Objectives:

1. Explain the consequences of mod-Gaussian convergence.
2. Describe general conditions which ensure the mod-Gaussian convergence.
3. Prove the mod-Gaussian convergence of a large class of models of random graphs.

Mod-Gaussian convergence and bounds on cumulants

Method of cumulants

If X is a random variable with convergent Laplace transform, its cumulants are:

$$\kappa^{(r)}(X) = \frac{d^r}{dz^r}\left(\log \mathbb{E}[\mathrm{e}^{zX}]\right)\Big|_{z=0}.$$

So, $\log \mathbb{E}[\mathrm{e}^{zX}] = \sum_{r=1}^{\infty} \frac{\kappa^{(r)}(X)}{r!}\, z^r$. The first cumulants are

$$\kappa^{(1)}(X) = \mathbb{E}[X] \quad ; \quad \kappa^{(2)}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \mathrm{Var}(X) \quad ;$$
$$\kappa^{(3)}(X) = \mathbb{E}[X^3] - 3\,\mathbb{E}[X^2]\,\mathbb{E}[X] + 2\,(\mathbb{E}[X])^3.$$

The Gaussian distribution $\mathcal{N}(m, \sigma^2)$ is characterized by $\kappa^{(1)}(X) = m$, $\kappa^{(2)}(X) = \sigma^2$, $\kappa^{r\geq 3}(X) = 0$.

Idea: characterize similarly the mod-Gaussian convergence of a sequence $(X_n)_{n\in\mathbb{N}}$.

Definition (Method of cumulants)

A sequence of random variables $(S_n)_{n\in\mathbb{N}}$ satisfies the hypotheses of the method of cumulants with parameters (D_n, N_n, A) if:

(MC1) One has $N_n \to +\infty$ and $\frac{D_n}{N_n} \to 0$.

(MC2) The first cumulants satisfy

$$\kappa^{(1)}(S_n) = 0;$$
$$\kappa^{(2)}(S_n) = (\sigma_n)^2 N_n D_n;$$
$$\kappa^{(3)}(S_n) = L_n\, N_n (D_n)^2$$

with $\lim_{n\to\infty}(\sigma_n)^2 = \sigma^2 > 0$ and $\lim_{n\to\infty} L_n = L$.

(MC3) All the cumulants satisfy

$$|\kappa^{(r)}(S_n)| \le N_n\,(2D_n)^{r-1}\, r^{r-2}\, A^r.$$

Mod-Gaussian convergence and its consequences

If $(S_n)_{n\in\mathbb{N}}$ satisfies the hypotheses MC1-MC3, then

$$X_n = \frac{S_n}{(N_n)^{1/3}(D_n)^{2/3}}$$

is mod-Gaussian convergent, with $t_n = (\sigma_n)^2 \left(\frac{N_n}{D_n}\right)^{1/3}$ and $\psi(z) = \exp\left(\frac{Lz^3}{6}\right)$.

Consequences:

1. Central limit theorem: if $Y_n = \frac{S_n}{\sqrt{\mathrm{Var}(S_n)}}$, then $Y_n \rightharpoonup \mathcal{N}(0,1)$.
2. Speed of convergence:

$$d_{\mathrm{Kol}}(Y_n\,,\,\mathcal{N}(0,1)) \leq \left(\frac{3A}{\sigma_n}\right)^3 \sqrt{\frac{D_n}{N_n}}.$$

   This inequality relies on the general estimate

$$d_{\mathrm{Kol}}(\mu\,,\,\nu) \leq \frac{1}{\pi}\int_{-T}^{T}\left|\frac{\widehat{\mu}(\xi)-\widehat{\nu}(\xi)}{\xi}\right| d\xi + \frac{24}{\pi T}\left\|\frac{d\nu(x)}{dx}\right\|_{\infty}.$$

3. **Normality zone and moderate deviations:** if $y \ll \left(\frac{N_n}{D_n}\right)^{1/6}$, then

$$\mathbb{P}[Y_n \geq y] = \mathbb{P}[\mathcal{N}(0,1) \geq y]\,(1+o(1)).$$

   If $1 \ll y \ll \left(\frac{N_n}{D_n}\right)^{1/4}$, then

$$\mathbb{P}[Y_n \geq y] = \frac{\mathrm{e}^{-\frac{y^2}{2}}}{y\sqrt{2\pi}}\, \exp\left(\frac{Ly^3}{6\sigma^3}\sqrt{\frac{D_n}{N_n}}\right)\,(1+o(1)).$$

   This estimate relies on the Berry–Esseen inequality and an argument of change of measure.

4. **Local limit theorem:** for any exponent $\varepsilon \in (0, \frac{1}{2})$,

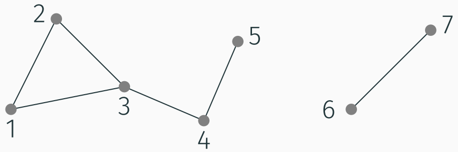
$$\lim_{n\to\infty} \left(\frac{N_n}{D_n}\right)^{\varepsilon} \mathbb{P}\left[Y_n - y \in \left(\frac{D_n}{N_n}\right)^{\varepsilon}(a,b)\right] = \frac{\mathrm{e}^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\,(b-a).$$

   Thus, Y_n is normal between the two scales $\left(\frac{N_n}{D_n}\right)^{-1/2}$ and $\left(\frac{N_n}{D_n}\right)^{1/6}$.

Joint cumulants and dependency graphs

Let $S = \sum_{v \in V} A_v$ be a sum of random variables, and $G = (V, E)$ a **dependency graph** for $(A_v)_{v \in V}$: if V_1 and V_2 are two disjoint subsets of V without edge $e = \{v_1, v_2\}$ between $v_1 \in V_1$ and $v_2 \in V_2$, then $(A_v)_{v \in V_1}$ and $(A_v)_{v \in V_2}$ are independent.

Example:



$(A_1, A_2, \ldots, A_5) \perp (A_6, A_7)$, but one has also $(A_1, A_2, A_3) \perp A_5$.

Parameters of the graph:

$$D = \max_{v \in V}(\deg v + 1),$$
$$N = \mathrm{card}(V),$$
$$A = \max_{v \in V} \|A_v\|_\infty.$$

Theorem (Bound on cumulants; Féray–M.–Nikeghbali, 2013)

If S is a sum of random variables with a dependency graph of parameters (D, N, A), then for any $r \geq 1$,

$$|\kappa^{(r)}(S)| \leq N\,(2D)^{r-1}\,r^{r-2}\,A^r.$$

Corollary: if $S_n = \sum_{i=1}^{N_n} A_{i,n}$ with the $A_{i,n}$'s bounded by A and a sparse dependency graph of maximal degree $D_n \ll N_n$, then MC3 is satisfied.

The proof of the bound relies on the notion of **joint cumulant**:

$$\kappa(A_1, A_2, \ldots, A_r) = \frac{d^r}{dz_1 dz_2 \cdots dz_r}\left(\log \mathbb{E}[\mathrm{e}^{z_1A_1+z_2A_2+\cdots+z_rA_r}]\right)\Big|_{z_1=\cdots=z_r=0}$$

$$= \sum_{\pi_1 \sqcup \pi_2 \sqcup \cdots \sqcup \pi_{\ell(\pi)} = [\![1,r]\!]} (-1)^{\ell(\pi)-1}(\ell(\pi)-1)!\,\prod_{i=1}^{\ell(\pi)} \mathbb{E}\left[\prod_{j\in\pi_i} A_j\right].$$

Properties of joint cumulants

1. For any random variable X, $\kappa^{(r)}(X) = \kappa(X, X, \ldots, X)$ (r occurrences).
2. The joint cumulants are multilinear and invariant by permutation.
3. If $\{A_1, A_2, \ldots, A_r\}$ can be split in two independent families, then $\kappa(\{A_1, \ldots, A_r\}) = 0$.

Consider a sum $S = \sum_{v \in V} A_v$ with a dependency graph G of parameters (D, N, A).

$$\kappa^{(r)}(S) = \sum_{v_1, v_2, \ldots, v_r} \kappa(A_{v_1}, A_{v_2}, \ldots, A_{v_r})$$

and the sum can be restricted to families $\{v_1, v_2, \ldots, v_r\}$ such that the induced multigraph $H = G[v_1, v_2, \ldots, v_r]$ is connected. Actually,

$$|\kappa(A_{v_1}, A_{v_2}, \ldots, A_{v_r})| \leq A^r\, 2^{r-1}\, \mathrm{ST}_H,$$

where ST_H is the number of spanning trees of H.

Sketch of proof of the bound

1. In the expansion of $\kappa(A_1, \ldots, A_r)$, many set partitions yield the same moment $M_\pi = \prod_{i=1}^{\ell(\pi)} \mathbb{E}[\prod_{j \in \pi_i} A_j]$, so

$$\kappa(A_1, \ldots, A_r) = \sum_{\pi'} M_{\pi'} \left(\sum_{\pi \to_H \pi'} \mu(\pi) \right)$$

$$|\kappa(A_1, \ldots, A_r)| \leq A^r \sum_{\pi'} \left| \sum_{\pi \to_H \pi'} \mu(\pi) \right|.$$

2. The functional $F_{H/\pi'} = \sum_{\pi \to_H \pi'} \mu(\pi)$ depends only on the contraction H/π' of H along π', and one can show that is up to a sign the bivariate Tutte polynomial

$$|F_{H/\pi'}| = T_{H/\pi'}(1, 0) \leq T_{H/\pi'}(1, 1) = \mathrm{ST}_{H/\pi'}.$$

3. A pair $(\pi', T \in \mathrm{ST}_{H/\pi'})$ can be associated to a bicolored spanning tree of H, hence

$$\sum_{\pi'} \mathrm{ST}_{H/\pi'} \leq 2^{r-1}\, \mathrm{ST}_H.$$

The bound on the cumulant of the sum S follows by noticing that:

- given a vertex v_1 and a Cayley tree T, the number of lists $(v_2, \dots, v_r)$ such that T is contained in $H = G[v_1, \dots, v_r]$ is smaller than D^{r-1};
- the number of pairs $(v_1 \in V,\ T$ Cayley tree) is $N\, r^{r-2}$.

The proof leads to the notion of **weighted dependency graph**.

Weighted dependency graphs

Definition (Weighted dependency graph; Féray, 2016)

A sum $S = \sum_{v \in V} A_v$ admits a weighted dependency graph $G = (V, E)$ of parameters $(\mathrm{wt} : E \to \mathbb{R}_+, A)$ if, for any family $\{v_1, v_2, \ldots, v_r\}$,

$$|\kappa(A_{v_1}, A_{v_2}, \ldots, A_{v_r})| \leq A^r \sum_{T \in \mathrm{ST}_{G[v_1,\ldots,v_r]}} \left(\prod_{(v_i,v_j) \text{ edge of } T} \mathrm{wt}(v_i, v_j) \right).$$

The same proof gives:

$$|\kappa(S)| \leq N\,(2D)^{r-1}\, r^{r-2}\, A^r$$

with $N = \mathrm{card}(V)$ and $D = \frac{1}{2}\left(1 + \max_{v \in W}\left(\sum_{w \sim v} \mathrm{wt}(v, w)\right)\right)$.

Sums of weakly dependent random variables

Let $S_n = \sum_{i=1}^{N_n} A_{i,n}$ be a sum of random variables, with $|A_{i,n}| \leq A$ a.s. and a dependency graph of maximal degree D_n. We suppose that

$$\frac{D_n}{N_n} \to 0 \quad ; \quad \frac{\mathrm{Var}(S_n)}{N_n D_n} \to \sigma^2 > 0 \quad ; \quad \frac{\kappa^{(3)}(S_n)}{N_n (D_n)^2} \to L.$$

Then, $S_n - \mathbb{E}[S_n]$ satisfies the hypotheses of the method of cumulants, and all its consequences. Moreover, one has the concentration inequality:

$$\mathbb{P}[|S_n - \mathbb{E}[S_n]| \geq \varepsilon] \leq 2 \exp\left(-\frac{\varepsilon^2}{9\,(\sum_{i=1}^{N_n} \mathbb{E}[|A_i|])D_n\,A}\right)$$
$$\leq 2 \exp\left(-\frac{\varepsilon^2}{9\,N_n D_n\,A^2}\right).$$

Functionals of ergodic Markov chains

Let $(X_n)_{n\in\mathbb{N}}$ be a reversible **ergodic Markov chain** on a finite state space $\mathfrak{X}$ of size M, and $f : \mathfrak{X} \to \mathbb{R}$. We set $S_n(f) = \sum_{i=1}^{n} f(X_i)$, and we denote π the stationary distribution, P the transition matrix, and

$$\theta_P = \max\{|z| \,|\, z \neq 1,\ z \text{ eigenvalue of } P\}.$$

The sequence $(S_n(f))_{n\in\mathbb{N}}$ has a weighted dependency graph and satisfies the hypotheses of the method of cumulants, with parameters $D_n = \frac{1+\theta_P}{2(1-\theta_P)}$, $N_n = n$, and $A = 2\|f\|_\infty\sqrt{M}$.

Remarks:

1. If $f = 1_{X_i=a}$, then one can take $A = 2$.
2. One can remove the hypothesis of reversibility if

$$\lim_{n\to\infty} \frac{\mathrm{Var}(S_n(f))}{n} = \mathrm{Var}_\pi(f) + 2\sum_{i=1}^{\infty} \mathrm{cov}_\pi(f(X_0), f(X_i)) \neq 0.$$

Magnetisation of the Ising model

Consider the Ising model on $\Lambda \subset \mathbb{Z}^d$, which is the probability measure on spin configurations $\sigma \in \{\pm 1\}^\Lambda$ proportional to $\exp(-\mathcal{H}^\Lambda_{\beta,h}(\sigma))$, with

$$\mathcal{H}^\Lambda_{\beta,h}(\sigma) = -\beta \sum_{i \sim j \in \Lambda} \sigma_i \sigma_j - h \sum_{i \in \Lambda} \sigma_i.$$

If $h \neq 0$ or $\beta < \beta_c(d)$, then the Ising model has a unique limiting probability measure $\mu^{\mathbb{Z}^d}_{\beta,h}$ on $\mathbb{Z}^d$.

Let $(\Lambda_n)_{n \in \mathbb{N}}$ be a growing sequence of boxes, and $M_n = \sum_{i \in \lambda_n} \sigma_i$ be the magnetization. Under $\mu^{\mathbb{Z}^d}_{\beta,h}$, $(M_n - \mathbb{E}[M_n])_{n \in \mathbb{N}}$ has a weighted dependency graph and satisfies the hypotheses of the method of cumulants if

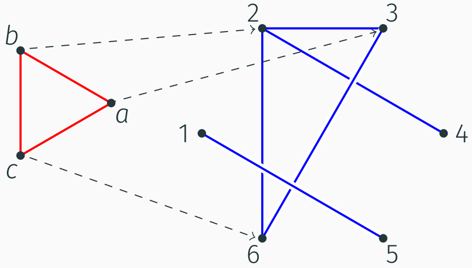
- $h \neq 0$ (non-zero ambient magnetic field);
- $h = 0$ and $\beta < \beta_1(d) < \beta_c(d)$ (very high temperature).

Subgraph counts in graphon models

Subgraph counts and subgraph densities

If $G = (V_G, E_G)$ is a finite graph, one says that $F = (V_F, E_F)$ is a subgraph of G if there is a map $\psi : V_F \to V_G$ such that

$$\forall e = \{x, y\} \in E_F, \ \ \{\psi(x), \psi(y)\} \in E_G.$$



Density of F in G: $t(F, G) = \frac{|\hom(F,G)|}{|V_G|^{|V_F|}} = \frac{6}{6^3} = \frac{1}{36}$.

Objective: establish the mod-Gaussian convergence of $t(F, G_n)$ for some models $(G_n)_{n \in \mathbb{N}}$ of random graphs.

Graph functions and graphons

A graph function is a measurable function $g : [0,1]^2 \to [0,1]$ that is symmetric: $g(x,y) = g(y,x)$ almost everywhere. If F is a graph on k vertices and g is a graph function, the density of F in g is

$$t(F,g) = \int_{[0,1]^k} \left(\prod_{\{i,j\} \in E_F} g(x_i, x_j) \right) dx_1\, dx_2 \cdots dx_k.$$

Let $\mathscr{F}$ be the set of graph functions, and $\mathscr{G} = \mathscr{F} / \sim$ its quotient by the relation:

$$g \sim h \iff \exists\, \sigma \text{ Lebesgue isomorphism of } [0,1], \text{ with } h(x,y) = g(\sigma(x), \sigma(y)).$$

Definition (Graphon; Lovász–Szegedy, 2006)

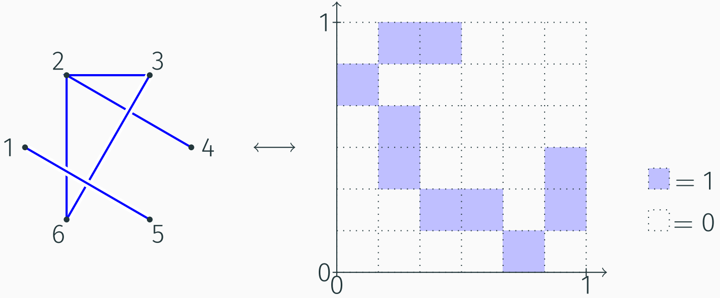
A graphon is an element $\gamma = [g]$ of the quotient space $\mathscr{G}$. Endowed with the topology of convergence of all the observables $t(F,\cdot)$, $\mathscr{G}$ is a compact metrisable space.

From graphons to random graphs

To any graphon $\gamma = [g]$, one can associate a random graph $G_n(\gamma)$ on n vertices:

1. One chooses n independent uniform variables $X_1, \ldots, X_n$ in $[0,1]$.
2. One connects i to j in $G_n(\gamma)$ according to a Bernoulli variable of parameter $g(X_i, X_j)$, independently for each pair $\{i, j\}$.

Conversely, to any graph G on n vertices, one can associate a graph function g:



Convergence of graphon models

Theorem (Lovász–Szegedy, 2006)

If γ is the graphon associated to a graph G, then $t(F,G) = t(F,\gamma)$ for any finite graph F. If $\gamma_n(\gamma)$ is the random graphon associated to the random graph $G_n(\gamma)$, then $\mathbb{E}[t(F,\gamma_n(\gamma))] = t(F,\gamma)$ and

$$\gamma_n(\gamma) \to_{\mathbb{P}} \gamma.$$

We introduce the algebra $\mathscr{O}$ of finite graphs F, endowed with the degree $\deg F = \mathrm{card}(V_F)$ and with the product $F_1 \times F_2 = F_1 \sqcup F_2$. One evaluates an **observable** $f \in \mathscr{O}$ by linear extension of the rule $F(\gamma) = t(F,\gamma)$. The convergence of graphon models amounts to:

$$\forall \gamma \in \mathscr{G},\ \ \forall f \in \mathscr{O},\ \ f(\gamma_n(\gamma)) \to_{\mathbb{P}} f(\gamma).$$

Dependency graphs for densities of subgraphs

Let γ be a graphon, F a finite graph on k vertices, $N_{n,k} = n^k$ and

$$\begin{aligned} S_n(F,\gamma) &= n^k\, t(F, G_n(\gamma)) \\ &= \sum_{\psi:[\![1,k]\!]\to[\![1,n]\!]} 1_{\psi \text{ is a morphism from } F \text{ to } G_n(\gamma)} = \sum_{\psi:[\![1,k]\!]\to[\![1,n]\!]} A_\psi. \end{aligned}$$

Given independent uniform random variables $(X_i)_{1\le i\le n}$ and $(U_{i,j})_{1\le i<j\le n}$, one can write :

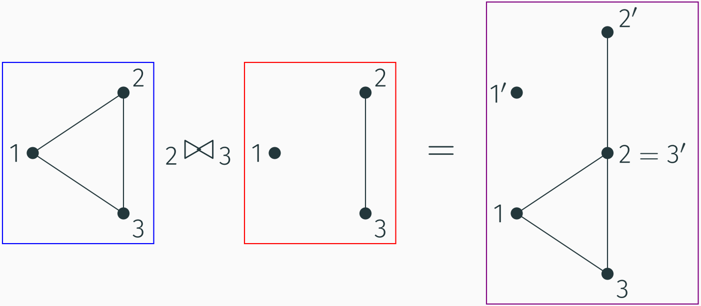
$$A_\psi = \prod_{\{i<j\}\in E_F} 1_{U_{\psi(i),\psi(j)} \le g(X_{\psi(i)}, X_{\psi(j)})}.$$

If ψ and ϕ have disjoint images, then A_ψ and A_ϕ are independent. Therefore, for any $n \in \mathbb{N}$, $\gamma \in \mathscr{G}$, $f \in \mathscr{O}_k$, $S_n(f,\gamma)$ is a sum of random variables with a dependency graph of parameters

$$D_{n,k} = k^2\, n^{k-1};\ \ N_{n,k} = n^k;\ \ A = \|f\|_{\mathscr{O}_k}.$$

Asymptotics of the first cumulants

The computation of the limits $\sigma^2(f,\gamma)$ and $L(f,\gamma)$ involves the operation of junction of graphs. If F and G are finite graphs of size k, $a \in V_F$ and $b \in V_G$, we denote $(F \bowtie G)(a,b)$ the graph on $2k-1$ vertices obtained by identifying $a \in V_F$ with $b \in V_G$.



$$\lim_{n\to\infty} \frac{\mathrm{cov}(S_n(F_1,\gamma),S_n(F_2,\gamma))}{n^{2k-1}} = \sum_{1\leq a,b\leq k} t((F_1 \bowtie F_2)(a,b),\gamma) - t(F_1,\gamma)\,t(F_2,\gamma).$$

Mod-Gaussian convergence of the graphon models

Theorem (Féray–M.–Nikeghbali, 2016)

Fix $\gamma \in \mathscr{G}$, $f \in \mathscr{O}_k$, and define

$$\kappa_2(F,G) = \frac{1}{k^2} \sum_{1\leq a,b \leq k} (F \bowtie G)(a,b) - F \cdot G;$$

$$\kappa_3(F,G,H) = \frac{1}{k^4} \sum_{1\leq a,b,c\leq k} \begin{array}{l}(F \bowtie G \bowtie H)(a,b,c) + 2\, F\cdot G\cdot H - (F \bowtie G)(a,b)\cdot H \\ -(G \bowtie H)(b,c)\cdot F - (F\bowtie H)(a,c)\cdot G\end{array}$$

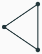
$$+ \frac{1}{k^4} \sum_{\mathbb{Z}/3\mathbb{Z}} \sum_{1\leq a,b\neq c,d \leq k} \begin{array}{l}(F \bowtie G \bowtie H)(a,b;c,d) + F\cdot G\cdot H \\ -(F\bowtie G)(a,b)\cdot H - (G\bowtie H)(c,d)\cdot F.\end{array}$$

If $\kappa_2(f,f)(\gamma) \neq 0$, then $S_n(f,\gamma)$ satisfies MC1-MC3 with parameters $D_{n,k} = k^2\, n^{k-1}$, $N_{n,k} = n^k$ and $A = \|f\|_{\mathscr{O}_k}$. Moreover,

$$\sigma^2 = \kappa_2(f,f)(\gamma)$$
$$L = \kappa_3(f,f,f)(\gamma).$$

Numbers of triangles

So, any subgraph count of a random graph $G_n(\gamma)$ stemming from any graphon $\gamma \in \mathscr{G}$ is generically mod-Gaussian convergent.

Example: If $K_3 =$  and $H =$ , then the density of triangles $t(K_3, G_n(\gamma))$ satisfies the central limit theorem:

$$Y_n = \sqrt{n}\,\frac{t(K_3, G_n(\gamma)) - t(K_3, \gamma)}{3\,\sqrt{t(H, \gamma) - t(K_3, \gamma)^2}} \rightharpoonup \mathcal{N}(0, 1),$$

assuming that the denominator is positive. Furthermore, one has

$$d_{\mathrm{Kol}}(Y_n, \mathcal{N}(0, 1)) \leq \frac{81}{(t(H, \gamma) - t(K_3, \gamma)^2)^{\frac{3}{2}}\sqrt{n}}$$

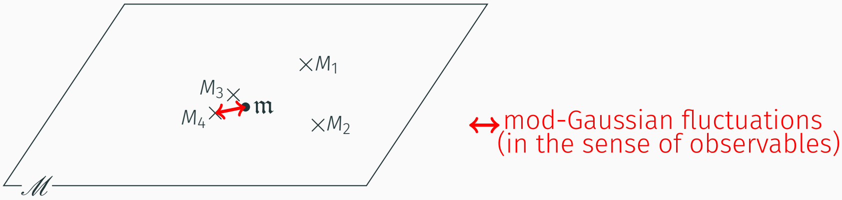
for n large enough; the concentration inequality

$$\mathbb{P}\left[|t(K_3, G_n(\gamma)) - t(K_3, \gamma)| \geq \varepsilon\right] \leq 2\,\exp\left(-\frac{n\varepsilon^2}{3}\right);$$

as well as a moderate deviation result and a local limit theorem.

Mod-Gaussian moduli spaces

We consider a compact metrisable space $\mathscr{M}$, where convergence is controled by a graded algebra of observables $\mathscr{O}$.



Informal definition: each parameter $\mathfrak{m} \in \mathscr{M}$ generates its own random perturbations $(M_n(\mathfrak{m}))_{n\in\mathbb{N}}$, and for any observable $f \in \mathscr{O}$, the sequence $(f(M_n(\mathfrak{m})))_{n\in\mathbb{N}}$ is mod-Gaussian convergent after appropriate renormalisation, assuming $\kappa_2(f,f)(\mathfrak{m}) \neq 0$.

One can prove that:

- the space of probability measures on a compact space;
- the space of permutons;
- the Thoma simplex

are mod-Gaussian moduli spaces for the following observables and random variables:

- polynomial functionals of empirical measures of random sequences;
- counts of motives in random permutations;
- random characters values associated to random integer partitions.

Informal conjecture: if one approximates a continous object by a random discrete one, the observables of the model usually have mod-Gaussian fluctuations (example: the Gromov–Hausdorff–Prohorov space).

The end