Dependency graphs, upper bounds on cumulants and singular graphons

Pierre-Loïc Méliot

28 May 2021

University Paris-Saclay (Orsay)

Consider independent uniform random variables $(U_{ij})_{i,j\geq 1}$ and $(X_i)_{i\geq 1}$ on the interval $[0,1]$. The **W-random graph** with size n and with graph function g is the random graph on n vertices $1, 2, \ldots, n$ such that

$$(i \text{ is connected to } j) \quad \Leftrightarrow \quad (U_{ij} \leq g(X_i, X_j)).$$

Here, the graph function g is an arbitrary Lebesgue measurable function on $[0,1]^2$ such that $g(x,y) = g(y,x)$ almost everywhere.

Questions: what does $G_n(g)$ look like? What are its fluctuations?

The W-random graphs are generalisations of the Erdős–Rényi random graphs (take $g(x,y) = p$). We shall be interested in the **subgraph densities**

$$t(F, G_n(g)) = \frac{|\operatorname{hom}(F, G_n(g))|}{n^k} \quad \text{with } |V_F| = k.$$

Question: what are the asymptotics of $t(F, G_n(g))$ when $n \to +\infty$?

Dependency graphs and joint cumulants

Sums of random variables

The random variables $t(F, G_n(g))$ are related to the sums $\sum_{\phi:[\![1,k]\!]\to[\![1,n]\!]} A_\phi$, with

$$A_\phi = \prod_{e=\{a,b\}\in E_F} 1_{(\phi(a)\sim\phi(b))} = \begin{cases} 1 & \text{if } \phi \text{ is a morphism of graphs,} \\ 0 & \text{otherwise.} \end{cases}$$

General problem: given a sum $S = \sum_{v\in V} A_v$ of random variables, what can be said about its distribution?

The **Lyapunov central limit theorem** covers most cases of sums of *independent* random variables: if $\sigma^2 = \mathrm{var}(S) = \sum_{v\in V} \mathrm{var}(A_v)$ and $\frac{1}{\sigma^{2+\delta}} \sum_{v\in V} \mathbb{E}[(A_v - \mathbb{E}[A_v])^{2+\delta}] = o(1)$ for some $\delta > 0$, then

$$\frac{S - \mathbb{E}[S]}{\sigma} \rightharpoonup \mathcal{N}(0,1).$$

But what happens for sums of *dependent* random variables?

Cumulants and joint cumulants

The **cumulants** of a random variable X are the coefficients

$$\kappa^{(r)}(X) = r!\,[z^r]\left(\log \mathbb{E}[\mathrm{e}^{zX}]\right)$$

of its log-Laplace transform. Thus, $\log \mathbb{E}[\mathrm{e}^{zX}] = \sum_{r=1}^{\infty} \frac{\kappa^{(r)}(X)}{r!}\, z^r$.

Generalisation: the **joint cumulant** of random variables $X_1, \ldots, X_r$ is

$$\kappa(X_1, \ldots, X_r) = [z_1 z_2 \cdots z_r]\left(\log \mathbb{E}[\mathrm{e}^{z_1 X_1 + z_2 X_2 + \cdots + z_r X_r}]\right).$$

We have $\kappa^{(r)}(X) = \kappa(X^{\otimes r})$ for $r \geq 1$. Other elementary properties:

1. Symmetry, multilinearity.
2. Independence: if $(X_1, \ldots, X_s)$ is independent from $(X_{s+1}, \ldots, X_r)$, then $\kappa(X_1, \ldots, X_r) = 0$.
3. First values: $\kappa(X) = \mathbb{E}[X]$, $\kappa(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]$, and

$$\kappa(X, Y, Z)$$
$$= \mathbb{E}[XYZ] - \mathbb{E}[X]\,\mathbb{E}[YZ] - \mathbb{E}[Y]\,\mathbb{E}[XZ] - \mathbb{E}[Z]\,\mathbb{E}[XY] + 2\,\mathbb{E}[X]\,\mathbb{E}[Y]\,\mathbb{E}[Z].$$

The distribution $\mathcal{N}(m, \sigma^2)$ with density $\frac{1}{\sqrt{2\pi\,\sigma^2}}\,\mathrm{e}^{-\frac{(x-m)^2}{2\sigma^2}}$ is characterised by $\kappa^{(1)}(X) = m$, $\kappa^{(2)}(X) = \sigma^2$, $\kappa^{(r\geq 3)}(X) = 0$.

$\rightsquigarrow$ **method of cumulants** in order to prove asymptotic normality.

Theorem (Leonov–Shiryaev, 1959)

Given random variables $X_1, \ldots, X_r$, we have the cumulant expansion

$$\kappa(X_1, \ldots, X_r) = \sum_{\pi \in \mathfrak{P}(r)} \mu(\pi) \prod_{j=1}^{\ell(\pi)} \mathbb{E}\left[\prod_{i \in \pi_j} X_i\right],$$

with $\mathfrak{P}(r) = \{$set partitions of $[\![1, r]\!]\}$, and $\mu(\pi) = (-1)^{\ell(\pi)-1}(\ell(\pi)-1)!$.

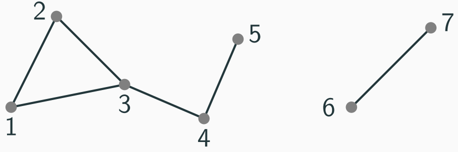
Intuition: if the variables $X_1, \ldots, X_r$ are *weakly dependent*, then their joint cumulant is small.

Dependency graphs and their parameters

A **dependency graph** for a family of variables $(A_v)_{v \in V}$ is a graph $G = (V, E)$ such that, if $(A_v)_{v \in V_1}$ and $(A_w)_{w \in V_2}$ correspond to two disjoint sets of vertices $V_1 \subset V$ and $V_2 \subset V$ without any edge $\{v \in V_1, w \in V_2\} \in E$, then these vectors are independent:

$$\mathbb{E}[f((A_v)_{v \in V_1})\, g((A_w)_{w \in V_2})] = \mathbb{E}[f((A_v)_{v \in V_1})]\, \mathbb{E}[g((A_w)_{w \in V_2})].$$

Example:



$(A_1, A_2, \ldots, A_5) \perp (A_6, A_7)$; $(A_1, A_2, A_3) \perp A_5$.

It is convenient to introduce the **parameters** (D, N, A) of a dependency graph for a family of random variables $(A_v)_{v \in V}$:

- $D = (\max_{v \in V} \deg v) + 1$.
- $N = |V|$.
- $A = \max_{v \in V} \|A_v\|_\infty$ (we shall work with sums of *bounded* variables, using truncation methods if needed).

We consider a family of random variables $(A_v)_{v \in V}$ with dependency graph $G = (V, E)$, and we select random variables $A_{v_1}, \ldots, A_{v_r}$ possibly with repetitions. The **induced dependency graph** $G[v_1, \ldots, v_r]$ is the graph with vertex set $[\![1, r]\!]$, and with one edge between i and j if $v_i = v_j$ or if $\{v_i, v_j\} \in E$.

An upper bound involving spanning trees

Theorem (Féray–M.–Nikeghbali, 2016)

In the aforementioned situation, for any vertices $v_1, \dots, v_r$

$$|\kappa(A_{v_1}, \dots, A_{v_r})| \leq 2^{r-1} A^r \, \mathrm{ST}_{G[v_1,\dots,v_r]},$$

where ST_H denotes the number of spanning trees of a graph H.

If $S = \sum_{v \in V} A_v$, this implies:

$$\begin{aligned} |\kappa^{(r)}(S)| &= \left| \sum_{v_1,\dots,v_r} \kappa(A_{v_1}, \dots, A_{v_r}) \right| \\ &\leq 2^{r-1} A^r \sum_{v_1,\dots,v_r} \mathrm{ST}_{G[v_1,\dots,v_r]} \\ &\leq 2^{r-1} \, r^{r-2} \, N \, D^{r-1} \, A^r. \end{aligned}$$

As $\frac{r^{r-2}}{r!} = O(1)$, this leads to estimates of the Laplace transform of S.

Sketch of proof of the upper bound

Given a graph H and a set partition π of V_H, we write $\pi \perp H$ if any edge $\{i,j\} \in E_H$ connects two distinct parts of π. We then set:

$$\mathrm{F}_H = (-1)^{|V_H|-1} \sum_{\pi \perp H} \mu(\pi).$$

This graph functional plays an important role in the computation of the joint cumulants. One can show that:

$$\mathrm{F}_H = \sum_{\substack{E \subset E_H \\ (V_H,E) \text{ connected}}} (-1)^{1+|E|-|V_H|}.$$

Therefore, $\mathrm{F}_H = \mathrm{T}_H(1,0)$ is a specialisation of the **Tutte polynomial**

$$\mathrm{T}_H(x,y) = \sum_{E \subset E_H} (x-1)^{\mathrm{cc}(E)-\mathrm{cc}(E_H)} (y-1)^{\mathrm{cc}(E)+|E|-|V_H|},$$

and $\mathrm{F}_H = \mathrm{T}_H(1,0) \leq \mathrm{T}_H(1,1) = \mathrm{ST}_H$.

We fix vertices $v_1, \ldots, v_r$, and we write

$$\kappa(A_{v_1}, \ldots, A_{v_r}) = \sum_{\pi \in \mathfrak{P}(r)} \mu(\pi) \prod_{j=1}^{\ell(\pi)} \mathbb{E}\left[\prod_{i \in \pi_j} X_i\right] = \sum_{\pi \in \mathfrak{P}(r)} \mu(\pi)\, M_\pi.$$

If Π is a set partition which can be obtained from another set partition π by splitting parts into independent vectors, then $M_\pi = M_\Pi$. We can therefore gather the set partitions according to the value of

$$\Phi_{G[v_1,\ldots,v_r]}(\pi) = \pi \wedge \{\text{connected components of } G[v_1, \ldots, v_r]\}.$$

So, with $H = G[v_1, \ldots, v_r]$, we get

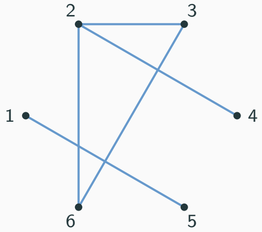
$$\kappa(A_{v_1}, \ldots, A_{v_r}) = \sum_{\Pi \in \mathfrak{P}(r)} M_\Pi \quad \underbrace{\left(\sum_{\Phi_H(\pi)=\Pi} \mu(\pi)\right)}_{\pm 1_{(H[\Pi_i] \text{ is connected for any } i)}\, \mathrm{F}_{H/\Pi}} \quad .$$

Thus, $|\kappa| \leq A^r \sum_\Pi \mathrm{ST}_{H/\Pi} \prod_{i=1}^{\ell(\Pi)} \mathrm{ST}_{H[\Pi_i]} = A^r\, 2^{r-1}\, \mathrm{ST}_H$.

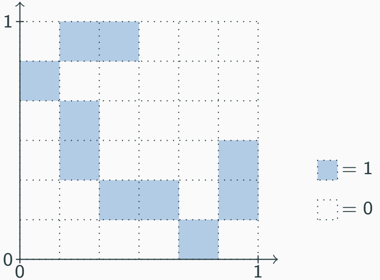
Fluctuations of graphon models

The space of graphons

A **graph function** is a measurable function $g \in \mathrm{L}^\infty([0,1]^2)$ with values in $[0,1]$, and which is symmetric: $g(x,y) = g(y,x)$ almost everywhere. Every finite graph $G = (\llbracket 1, n \rrbracket, E)$ yields a graph function g_G by drawing its adjacency matrix as a function on $[0,1]^2$.



graph G



graph function g_G

$\rightsquigarrow$ convergence of a sequence of graphs. But we would like a notion of convergence independent from the labelling of the vertices.

Two graph functions g and g' are *conjugated* if there exists a Lebesgue isomorphism $\sigma : [0,1] \to [0,1]$ such that $g'(x,y) = g(\sigma(x), \sigma(y))$. On the other hand, a norm on $\mathrm{L}^\infty([0,1]^2)$ is given by:

$$\|g\|_\square = \sup_{S,T\subset[0,1]} \left| \int_{S\times T} g(x,y)\,dx\,dy \right| \quad \text{(equivalent to } \|\cdot\|_{\infty\to 1}\text{)}.$$

The **cut-metric** on the set of graph functions is defined by:

$$d_\square(g,g') = \inf_\sigma \|g^\sigma - g'\|_\square.$$

Two graph functions are *equivalent* if $d_\square(g,g') = 0$. The equivalence classes of graph functions are called **graphons**.

Theorem (Lovász–Szegedy, 2007)

The cut-metric $d_\square$ induces a distance on the space of graphons $\mathscr{G} = \mathrm{L}^\infty([0,1]^2,[0,1])\,/\!\sim$, which makes it a compact metric space.

Topology and observables

If F is a motive with size k, we define for any graph function g

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

In fact, $t(F,g)$ only depends on the graphon $\gamma = [g]$, so we can consider $t(F,\cdot)$ as an **observable** on the space $\mathscr{G}$. These observables generalise the graph densities of the introduction:

$$t(F, g_G) = t(F,G) \quad \text{for any finite graph } G.$$

Theorem (Chayes–Borgs–Lovász–Sós–Vesztergombi, 2008)

The topology of $\mathscr{G}$ is determined by the observables $t(F,\cdot)$:

$$(\gamma_n \to_{d_\square} \gamma) \;\Leftrightarrow\; (\forall F,\; t(F,\gamma_n) \to t(F,\gamma)).$$

The graphon space parametrises all the possible limits of a sequence of graphs (for the convergence of all the densities). For any $\gamma = [g] \in \mathscr{G}$, the law of the W-random graph $G_n(g)$ only depends on γ, so it makes sense to consider random graphs $G_n(\gamma)$ for any $n \in \mathbb{N}$ and any $\gamma \in \mathscr{G}$.

One computes readily:

$$\mathbb{E}[t(F, G_n(\gamma))] = t(F,\gamma) + O\left(\frac{k^2}{n}\right) \qquad ; \qquad \mathrm{var}(t(F, G_n(\gamma))) = O\left(\frac{k^2}{n}\right).$$

So, for any graphon γ, $(G_n(\gamma))_{n\in\mathbb{N}}$ converges in probability towards γ.

Question: what is the asymptotic distribution of

$$Y_n(F,\gamma) = \frac{t(F, G_n(\gamma)) - \mathbb{E}[t(F, G_n(\gamma))]}{\sqrt{\mathrm{var}(t(F, G_n(\gamma)))}}, \quad \gamma \in \mathscr{G},\ n \to \infty.$$

The answer will depend on the size of the variance.

Fluctuations of the generic graphons

We fix $\gamma \in \mathscr{G}$, and a motive F with size k. Let us remark that

$$S_n(F) = n^k\, t(F, G_n(\gamma)) = \sum_{\phi : [\![1,k]\!] \to [\![1,n]\!]} A_\phi$$

is a sum of bounded random variables with dependency graph given by

$$V = [\![1, n]\!]^k\,;$$
$$E = \{\{\phi, \psi\} \mid \phi([\![1, k]\!]) \cap \psi([\![1, k]\!]) \neq \emptyset\}.$$

Therefore, $D \leq k^2\, n^{k-1}$, $N = n^k$ and $A = 1$, so

$$|\kappa^{(r)}(S_n(F))| \leq (2k^2)^{r-1}\, r^{r-2}\, n^{(k-1)r+1};$$
$$|\kappa^{(r)}(t(F, G_n(\gamma)))| \leq (2k^2)^{r-1}\, r^{r-2}\, n^{1-r}.$$

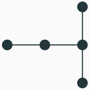
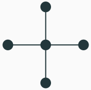
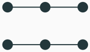
These estimates imply a *possibly degenerate* **central limit theorem**:

$$\begin{aligned}
&\log \mathbb{E}\left[\mathrm{e}^{z\sqrt{n}(t(F,G_n(\gamma))-\mathbb{E}[t(F,G_n(\gamma))])}\right] \\
&= n\,\mathrm{var}(t(F,G_n(\gamma))\,\frac{z^2}{2} + O\left(\sum_{r\geq 3}(2k^2)^{r-1}\,\frac{r^{r-2}}{r!}\,n^{1-\frac{r}{2}}\right) \\
&= n\,\mathrm{var}(t(F,G_n(\gamma))\,\frac{z^2}{2} + O\left(n^{-\frac{1}{2}}\right).
\end{aligned}$$

Asymptotic (co)variance? If $C_n = \mathrm{Cov}(t(F_1,G_n(\gamma)),t(F_2,G_n(\gamma)))$, then:

$$\lim_{n\to\infty} n\,C_n = \kappa_2(F_1,F_2)(\gamma) = \sum_{\substack{i\in V_{F_1}\\ j\in V_{F_2}}} t((F_1\bowtie F_2)(i,j),\gamma) - t(F_1\sqcup F_2,\gamma).$$

Example: with $F = F_1 = F_2 =$ ,

$\kappa_2(F,F) = 4$  $+\ 4$  $+$  $-\ 9$ .

Theorem (Féray–M.–Nikeghbali, 2020)

If $\kappa_2(F,F)(\gamma) = \sigma^2(F,\gamma) \neq 0$, then:

$$\sup_{s\in\mathbb{R}} \left| \mathbb{P}[Y_n(F,\gamma) \leq s] - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{s} \mathrm{e}^{-\frac{x^2}{2}}\, dx \right| = O\left(\frac{k^4}{\sigma^3(F,\gamma)\sqrt{n}} \right);$$

(upper bound on the **Kolmogorov distance***). We also have a* **concentration inequality***:*

$$\mathbb{P}[|t(F,G_n(\gamma)) - \mathbb{E}[t(F,G_n(\gamma))]| \geq x] \leq 2\exp\left(-\frac{nx^2}{9k^2}\right)$$

and **moderate deviation estimates** *of*

$$\mathbb{P}[|t(F,G_n(\gamma)) - \mathbb{E}[t(F,G_n(\gamma))]| \geq x_n], \quad n^{-1/2} \ll x_n \ll n^{-1/4}.$$

Question: what happens if $\kappa_2(F,F)(\gamma) = 0$?

Erdős–Rényi random graphs and singular graphons

Suppose $\gamma = [p]$ with $p \in (0,1)$. Then, $\kappa_2(F,F)(\gamma) = 0$ for any motive F, because $t(F,\gamma) = p^{|E_F|}$ and $|E_{F_1 \bowtie F_2}| = |E_{F_1}| + |E_{F_2}|$. Actually, the observable $n^k\, t(F, G_n(p))$ of a **Erdős–Rényi graph** $G_n(p)$ has a new dependency graph with edge set

$$E' = \{\{\phi,\psi\} \mid \phi(e) = \psi(e) \text{ for some edge } e \in E_F\}.$$

This changes D into $D' = O(n^{k-2})$, and one can prove:

Theorem (Janson–Nowicki, 1988)

For any motive F,

$$n\left(t(F,G_n(p)) - \mathbb{E}[t(F,G_n(p))]\right) \rightharpoonup_{n\to\infty} \mathcal{N}(0, {\sigma^2}'(F,p)).$$

Moreover, the Kolmogorov distance is a $O(n^{-1})$.

The rescaling involves a factor n instead of $\sqrt{n}$.

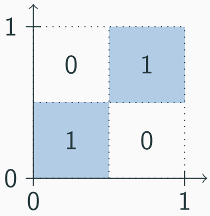
We call a graphon γ **singular** if $\kappa_2(F,F)(\gamma) = 0$ for any motive F. For a long time, we conjectured that the Erdős–Rényi random graphs were the only singular graphons...

... but there are other examples!

Conjecture

We call a graph function g **transitive** if its automorphism group (Lebesgue isomorphisms σ such that $g^\sigma = g$) acts transitively on $[0,1]$. A graphon γ is singular if and only if it is the class of a transitive graph function.

Example:



Let us see why the transitive graph functions yield singular graphons. Fix a motive F and $(i,j) \in [\![1,k]\!]^2$. With $z = x_i = y_j$ and $I = [0,1]$, we have:

$$
\begin{aligned}
&t((F \bowtie F)(i,j), g) \\
&= \int_{I^{2k-1}} \prod_{i \notin \{a,b\}} g(x_a, x_b) \prod_{\{i,k\}} g(z, x_k) \underbrace{\prod_{j \notin \{c,d\}} g(y_c, y_d) \prod_{\{j,l\}} g(z, y_l)}_{f(z)} \, dx_{\setminus i} \, dy_{\setminus j} \, dz.
\end{aligned}
$$

By transitivity, the integral of $f(z)$ with respect to the set of variables $\{y_1, \ldots, y_k\} \setminus \{y_j\}$ does not depend on $z \in [0,1]$, so we can replace $f(z)$ by $\int f(w)\, dw$.

$$
\begin{aligned}
&\int_{I^{2k}} \prod_{i \notin \{a,b\}} g(x_a, x_b) \prod_{\{i,k\}} g(z, x_k) \prod_{j \notin \{c,d\}} g(y_c, y_d) \prod_{\{j,l\}} g(w, y_l) \, dx_{\setminus i} \, dy_{\setminus j} \, dz \, dw \\
&= (t(F,g))^2.
\end{aligned}
$$

Fluctuations of the singular graphons

The singular graphons share many properties with the constant graphons.

Theorem (M., 2021)

Let γ be a singular graphon. For any motive F with size k,

$$|\kappa^{(r)}(t(F, G_n(\gamma)))| \leq 2^{r-1}\,(rk^2)^r\, n^{-r}.$$

Therefore, there exists a **limiting law**

$$n\,(t(F, G_n(\gamma)) - \mathbb{E}[t(F, G_n(\gamma))]) \rightharpoonup_{n\to\infty} \mathcal{L}_{(F,\gamma)}.$$

The probability distribution $\mathcal{L}_{(F,\gamma)}$ is determined by its moments and has subexponential decay.

Chatterjee and Bhattacharya have since proposed an expansion of $X \sim \mathcal{L}_{(F,\gamma)}$ as a series of independent χ^2 variables whose parameters are related to the spectrum of the kernel $g \in \gamma$.

The limiting law $\mathcal{L}_{(F,\gamma)}$ is almost never a Gaussian distribution.

Theorem (M., 2021)

Consider the motive

$$F = K_2 = \bullet\!\!-\!\!\!-\!\!\!-\!\!\bullet\,.$$

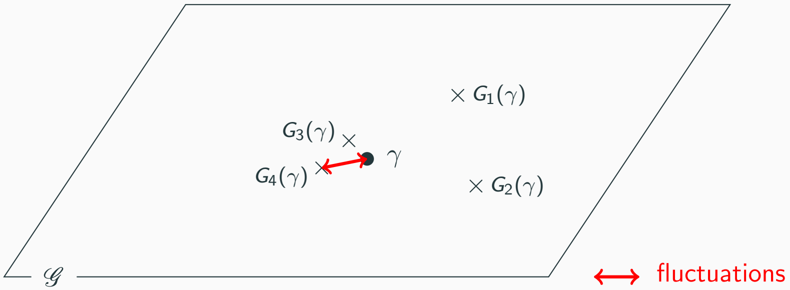
If $\gamma = [g]$ is a singular graphon such that $\mathcal{L}_{(K_2,\gamma)}$ is a normal distribution, then $\gamma = [p]$ with $p = t(K_2, \gamma) = \int_{[0,1]^2} g(x,y)\, dx\, dy$.

Ingredients:

- the vanishing of the term with highest degree in n of the variance implies the vanishing of the term with highest degree in all the higher cumulants (non trivial!).
- we can modify the proof of the generic upper bound on cumulants in order to take into account this vanishing in highest degree.
- matrix-tree theorem.

Mod-Gaussian moduli spaces

Moduli spaces of random models



- a space $\mathscr{G}$ of parameters of models of random combinatorial objects.
- an algebra $\mathscr{O}$ of observables which control the topology.
- for a generic point γ, the fluctuations of the observables $f(G_n(\gamma))$ are asymptotically normal.
- the singular points have smaller fluctuations which are not necessarily asymptotically normal; these points usually correspond to models with additional symmetries.

Random metric spaces

The framework described on the previous slide enables the study of large classes of models of random permutations, of random integer partitions and of random metric spaces.

A **measure metric space** is a complete metric space (X, d) endowed with a probability measure μ. We consider the set $\mathscr{M}$ of all such spaces, up to measure-preserving isometries. A discrete random approximation of (X, d, μ) is (X_n, d_n, μ_n) with:

- $X_n = n$ independent points $x_1, \ldots, x_n$ of X chosen according to μ;
- $d_n = d_{|X_n \times X_n}$;
- $\mu_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i}$.

The space $\mathscr{M}$ is endowed with the **Gromov–Hausdorff–Prohorov** topology, which makes it a polish space.

The observables

$$F(X,d,\mu) = \int_{[0,1]^k} f(\{d(x_i,x_j),\, 1 \leq i < j \leq k\})\, \mu(dx_1)\cdots\mu(dx_k)$$

for $f \in \mathscr{C}_b(\mathbb{R}^{\binom{k}{2}}, \mathbb{R})$ and $k \geq 2$ control the topology of $\mathscr{M}$.

Theorem (De Catelan–M., 2019)

Generically, the fluctuations of $F(X_n, d_n, \mu_n)$ are asymptotically of order $O(n^{-1/2})$ and normal.

1. *The observables of the singular points have fluctuations of order $O(n^{-1})$, with a limiting law which in general is not normal.*
2. *Among the singular points, an important subclass consists in the* **compact homogeneous spaces** *$X = G/K$ with G compact group of measure-preserving isometries of X, and K closed subgroup.*
3. *We conjecture that the only singular points are the compact homogeneous spaces.*

The end