Limit for sequences of large dense weighted graphs and Probabillity-graphons

Julien Weibel
with Romain Abraham and Jean-François Delmas

Institut Denis Poisson - Université d'Orléans

23 juin 2023, Angers
Journées des Probabilités 2023

Introduction

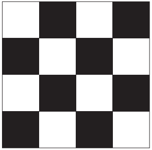
- Large graphs and weighted graphs are ubiquitous
- Theory of real-valued graphons (short for graph functions) developed to study limits of large graphs
- In this talk: adaptation of this theory to large weighted graphs
- To this end, new objects: probability-graphons

Large dense graphs and their representation

A *large dense graph* is a finite (non-directed) graph G with a number of vertices n large, and a number of edges $\Omega(n^2)$.

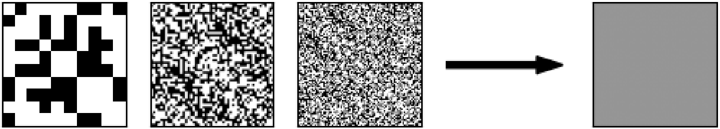
The adjacency matrix is transformed through scaling into a "pixel picture" map on $[0,1]^2$.

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \longrightarrow$$

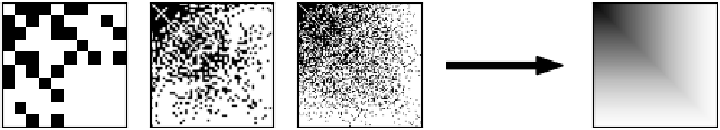


Some examples of convergence

Erdos-Renyi graphs:

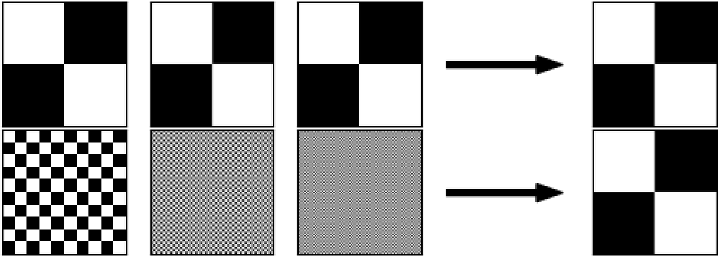


Growing uniform attachement graphs:



The limit may depend on the labeling

Bipartite graphs with different labeling of the vertices:



Real-valued graphons and the cut distance

• A *real-valued graphon* is a symmetric measurable function $w : [0,1]^2 \to [0,1]$.

For two "vertices" $x, y \in [0,1]$, $w(x,y)$ corresponds to the average edge density between x and y.

Real-valued graphons and the cut distance

- A *real-valued graphon* is a symmetric measurable function $w : [0,1]^2 \to [0,1]$.

For two "vertices" $x, y \in [0,1]$, $w(x,y)$ corresponds to the average edge density between x and y.

- The *cut distance* is defined as:

$$\delta_{\Box,\mathbb{R}}(w,u) = \inf_{\varphi,\psi} \sup_{S,T \subset [0,1]} \left| \int_{S \times T} w(\varphi(x),\varphi(y)) - u(\psi(x),\psi(y)) \, \mathrm{d}x\mathrm{d}y \right|,$$

where $\varphi, \psi : [0,1] \to [0,1]$ are measure-preserving maps.

Main results for real-valued graphons

Theorem 1

Every real-valued graphon is a $\delta_{\square,\mathbb{R}}$-limit of a sequence of finite graphs.

Theorem 2

The space of real-valued graphons equipped with the distance $\delta_{\square,\mathbb{R}}$ is compact.

Sampling from real-valued graphons

Let $X_1, \ldots, X_k$ be iid random variables uniformly distributed over $[0, 1]$.
Define the random graph $\mathbb{G}(k, w)$ with k vertices and each edge ij is independently present with probability $w(X_i, X_j)$.

Sampling from real-valued graphons

Let $X_1, \ldots, X_k$ be iid random variables uniformly distributed over $[0, 1]$.
Define the random graph $\mathbb{G}(k, w)$ with k vertices and each edge ij is independently present with probability $w(X_i, X_j)$.

Theorem 3

Let $(w_n)_{n \in \mathbb{N}}$ and w be real-valued graphons. The following properties are equivalent:

1. *For every $k \geq 2$, the sequence of random graphs $(\mathbb{G}(k, w_n))_{n \in \mathbb{N}}$ converges in distribution to $\mathbb{G}(k, w)$.*
2. $\lim_{n \to \infty} \delta_{\square, \mathbb{R}}(w_n, w) = 0.$

Weighted graphs

A *weighted graph* is a complete directed graph (V, E) with a decoration map M that associates to each edge e in E a decoration z in a Polish space $\mathbf{Z}$.

Missing edges may be represented as an edge with decoration ∂, where $\partial \in \mathbf{Z}$ is a cemetery point.

Important cases: $\mathbf{Z} = \mathbb{R}$ or $\mathbb{R}^d$ or $\mathbb{N}$ (here $\partial = 0$).

Definition of probability-graphons

A *probability-graphon* is a map W from $[0,1]^2$ to $\mathcal{M}_1(\mathbf{Z})$ such that:

1. W is a *probability measure* in $\mathrm{d}z$: for every $(x,y) \in [0,1]^2$, $W(x,y;\cdot)$ belongs to the set of probability measures $\mathcal{M}_1(\mathbf{Z})$.
2. W is *measurable* in (x,y): for every measurable set $A \subset \mathbf{Z}$, the function $(x,y) \mapsto W(x,y;A)$ defined on $[0,1]^2$ is measurable.

Definition of probability-graphons

A *probability-graphon* is a map W from $[0,1]^2$ to $\mathcal{M}_1(\mathbf{Z})$ such that:

1. W is a *probability measure* in $\mathrm{d}z$: for every $(x, y) \in [0,1]^2$, $W(x, y; \cdot)$ belongs to the set of probability measures $\mathcal{M}_1(\mathbf{Z})$.
2. W is *measurable* in (x, y): for every measurable set $A \subset \mathbf{Z}$, the function $(x, y) \mapsto W(x, y; A)$ defined on $[0,1]^2$ is measurable.

Denote by $\widetilde{\mathcal{W}}_1$ the space of probability-graphons where we identify W, U if there exist measure-preserving maps $\varphi, \psi : [0,1] \to [0,1]$ such that $W(\varphi(x), \varphi(y); \cdot) = U((\psi(x), \psi(y); \cdot)$ almost everywhere.

Remark: $\mathbf{Z} = \{0,1\}$, $W(x, y; \cdot) = w(x, y)\delta_1 + (1 - w(x, y))\delta_0$

The cut distance for probability-graphons

Let d_{m} be a distance that generates the weak topology on the space of sub-probability measures $\mathcal{M}_{\leq 1}(\mathbf{Z})$.

For $S, T \subset [0,1]$, define the sub-probability measure $W(S,T;\cdot) = \int_{S\times T} W(x,y;\cdot)\, \mathrm{d}x\mathrm{d}y$.

- The *cut distance* for probability-graphons is defined as:

$$\delta_{\Box,\mathrm{m}}(W,U) = \inf_{\varphi,\psi} \sup_{S,T\subset[0,1]} d_{\mathrm{m}}(W(\varphi(S),\varphi(T);\cdot), U(\psi(S),\psi(T);\cdot)),$$

where $\varphi, \psi : [0,1] \to [0,1]$ are measure-preserving maps.

The cut distance for probability-graphons

Let d_{m} be a distance that generates the weak topology on the space of sub-probability measures $\mathcal{M}_{\leq 1}(\mathbf{Z})$.

For $S, T \subset [0,1]$, define the sub-probability measure $W(S,T;\cdot) = \int_{S\times T} W(x,y;\cdot)\,\mathrm{d}x\mathrm{d}y$.

- The *cut distance* for probability-graphons is defined as:

$$\delta_{\square,\mathrm{m}}(W,U) = \inf_{\varphi,\psi} \sup_{S,T\subset[0,1]} d_{\mathrm{m}}(W(\varphi(S),\varphi(T);\cdot), U(\psi(S),\psi(T);\cdot)),$$

where $\varphi, \psi : [0,1] \to [0,1]$ are measure-preserving maps.

Theorem 4 ($\widetilde{\mathcal{W}}_1$ is a Polish space)

If d_{m} is complete, then $(\widetilde{\mathcal{W}}_1, \delta_{\square,\mathrm{m}})$ is a Polish space.

Main results for probability-graphons

A subset of probability-graphons $\mathcal{K}$ is said to be tight if the subset of probability measures $\{M_W : W \in \mathcal{K}\} \subset \mathcal{M}_1(\mathbf{Z})$ is tight, where:

$$M_W(\mathrm{d}z) = W([0,1]^2; \mathrm{d}z) = \int_{[0,1]^2} W(x, y; \mathrm{d}z)\, \mathrm{d}x\mathrm{d}y.$$

Theorem 5 (Compactness theorem for $\widetilde{\mathcal{W}}_1$)

If a sequence of probability-graphons is tight, then it has a subsequence converging for $\delta_{\square,\mathrm{m}}$.
If $\mathbf{Z}$ is compact, then the space $(\widetilde{\mathcal{W}}_1, \delta_{\square,\mathrm{m}})$ is compact.

Main results for probability-graphons

A subset of probability-graphons $\mathcal{K}$ is said to be tight if the subset of probability measures $\{M_W : W \in \mathcal{K}\} \subset \mathcal{M}_1(\mathbf{Z})$ is tight, where:

$$M_W(\mathrm{d}z) = W([0,1]^2; \mathrm{d}z) = \int_{[0,1]^2} W(x, y; \mathrm{d}z)\, \mathrm{d}x\mathrm{d}y.$$

Theorem 5 (Compactness theorem for $\widetilde{\mathcal{W}}_1$)

If a sequence of probability-graphons is tight, then it has a subsequence converging for $\delta_{\square,\mathrm{m}}$.
If $\mathbf{Z}$ is compact, then the space $(\widetilde{\mathcal{W}}_1, \delta_{\square,\mathrm{m}})$ is compact.

Theorem 6 (Equivalence of topologies induced by $\delta_{\square,\mathrm{m}}$ on $\widetilde{\mathcal{W}}_1$)

The topology on $\widetilde{\mathcal{W}}_1$ induced by the distance $\delta_{\square,\mathrm{m}}$ does not depend on the choice of the distance d_m on $\mathcal{M}_{\leq 1}(\mathbf{Z})$, as long as d_m satisfies some mild hypothesis (H).

Sampling from probability-graphons

Let $X_1, \ldots, X_k$ be iid random variables uniformly distributed over $[0, 1]$.
Define the random weighted graph $\mathbb{G}(k, W)$ with k vertices and each edge ij is independently decorated with a random decoration distributed according to $W(X_i, X_j; \cdot)$.

Sampling from probability-graphons

Let $X_1, \ldots, X_k$ be iid random variables uniformly distributed over $[0, 1]$.
Define the random weighted graph $\mathbb{G}(k, W)$ with k vertices and each edge ij is independently decorated with a random decoration distributed according to $W(X_i, X_j; \cdot)$.

Theorem 7 (Characterization of the topology induced by $\delta_{\square,\mathrm{m}}$)

Let $(W_n)_{n\in\mathbb{N}}$ and W be probability-graphons from $\widetilde{\mathcal{W}}_1$. The following properties are equivalent:

1. *For every $k \geq 2$, the sequence of random graphs $(\mathbb{G}(k, W_n))_{n\in\mathbb{N}}$ converges in distribution to $\mathbb{G}(k, W)$.*
2. $\lim_{n\to\infty} \delta_{\square,\mathrm{m}}(W_n, W) = 0$ *for some (and hence for every) choice of the distance d_m on $\mathcal{M}_{\leq 1}(\mathbf{Z})$ that satisfies hypothesis (H).*

Thank you for your attention!