Graph limit and graphons

A short statistical introduction

Charles Dufour

charles.dufour@epfl.ch

September 30, 2022

Chair of Statistical Data Science (SDS)

Outline

1. Graph limit
2. Graphon in statistics
3. Estimation
4. Summary

Inspired by [Janson, 2010, Glasscock, 2015, Keriven, 2020]

Why bother ?

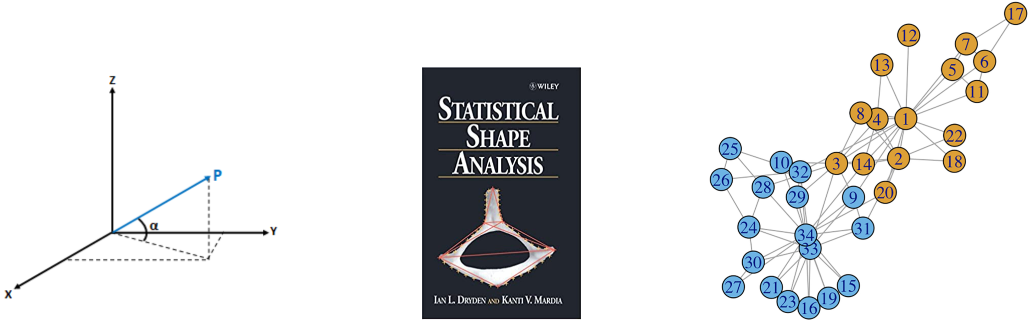


Figure 1: Different levels of Euclideanity

Graph limit

1. Graph limit
2. Graphon in statistics
3. Estimation
4. Summary

”Simple” problems

How many 4-cycles must a graph with edge density at least $1/2$ have?

"Simple" problems

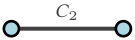
How many 4-cycles must a graph with edge density at least $1/2$ have?

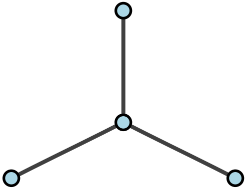
- $d(H,G) := \frac{\#\text{ copies of } H \text{ in } G}{\binom{n_G}{n_H}}$.
- $d(C_2, G)$: edge density
- $d(C_4, G)$: 4-cycles density

"Simple" problems

How many 4-cycles must a graph with edge density at least $1/2$ have?

- $d(H,G) := \frac{\#\text{ copies of } H \text{ in } G}{\binom{n_G}{n_H}}$.
- $d(C_2, G)$: edge density
- $d(C_4, G)$: 4-cycles density



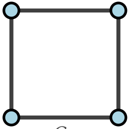


$d(C_2, G) = 0.5$

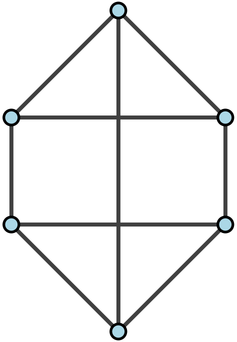
"Simple" problems

How many 4-cycles must a graph with edge density at least $1/2$ have?

- $d(H,G) := \frac{\#\text{ copies of } H \text{ in } G}{\binom{n_G}{n_H}}$.
- $d(C_2,G)$: edge density
- $d(C_4,G)$: 4-cycles density



C_4

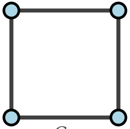


$d(C_4,G) = 3/\binom{6}{4} = 3/15$

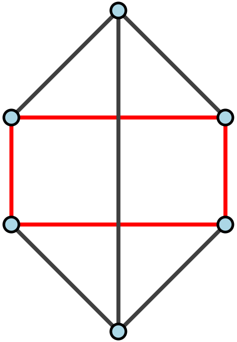
"Simple" problems

How many 4-cycles must a graph with edge density at least $1/2$ have?

- $d(H, G) := \frac{\#\text{ copies of } H \text{ in } G}{\binom{n_G}{n_H}}$.
- $d(C_2, G)$: edge density
- $d(C_4, G)$: 4-cycles density



C_4

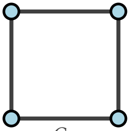


$d(C_4, G) = 3/\binom{6}{4} = 3/15$

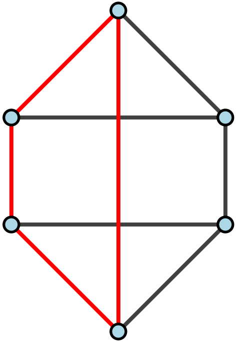
"Simple" problems

How many 4-cycles must a graph with edge density at least $1/2$ have?

- $d(H,G) := \frac{\#\text{ copies of } H \text{ in } G}{\binom{n_G}{n_H}}$.
- $d(C_2, G)$: edge density
- $d(C_4, G)$: 4-cycles density



C_4

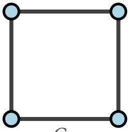


$$d(C_4, G) = 3/\binom{6}{4} = 3/15$$

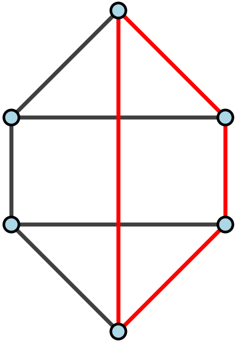
"Simple" problems

How many 4-cycles must a graph with edge density at least $1/2$ have?

- $d(H,G) := \frac{\#\text{ copies of } H \text{ in } G}{\binom{n_G}{n_H}}$.
- $d(C_2, G)$: edge density
- $d(C_4, G)$: 4-cycles density



C_4

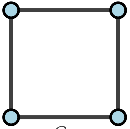


$$d(C_4, G) = 3/\binom{6}{4} = 3/15$$

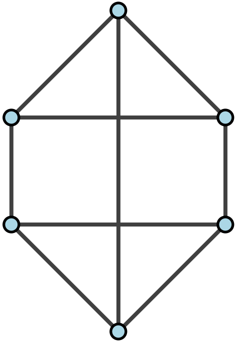
"Simple" problems

How many 4-cycles must a graph with edge density at least $1/2$ have?

- $d(H,G) := \frac{\#\text{ copies of } H \text{ in } G}{\binom{n_G}{n_H}}$.
- $d(C_2, G)$: edge density
- $d(C_4, G)$: 4-cycles density



C_4



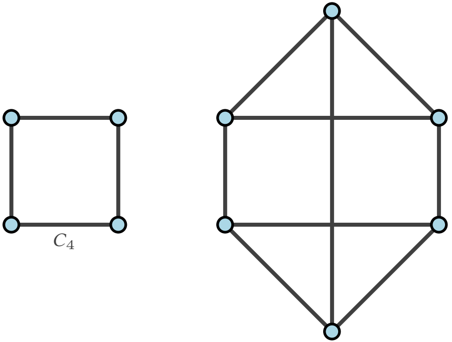
$d(C_4, G) = 3/\binom{6}{4} = 3/15$

- Minimize $d(C_4, G)$ over all finite graphs G with $d(C_2, G) = 1/2$.

"Simple" problems

How many 4-cycles must a graph with edge density at least $1/2$ have?

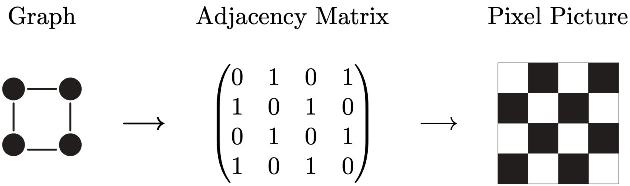
- $d(H,G) := \frac{\#\text{ copies of } H \text{ in } G}{\binom{n_G}{n_H}}$.
- $d(C_2,G)$: edge density
- $d(C_4,G)$: 4-cycles density



$d(C_4,G) = 3/\binom{6}{4} = 3/15$

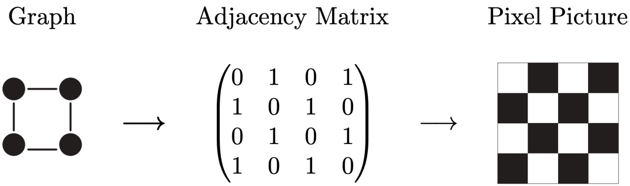
- Minimize $d(C_4,G)$ over all finite graphs G with $d(C_2,G) = 1/2$.
- Minimize $x^3 - 6x$ over all rational numbers x satisfying $x \geq 0$.

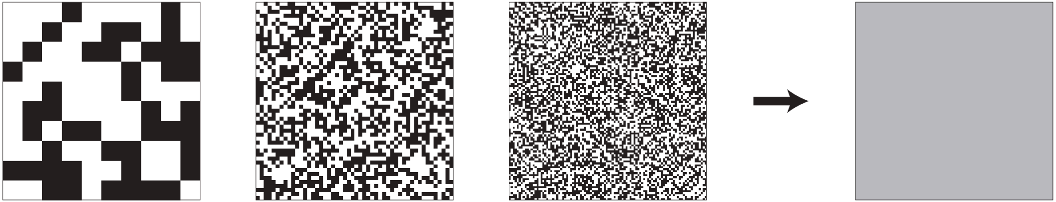
Graph limit and graphon: intuition



Images from [Glasscock, 2015]
Graph limits were introduced by [Lovász and Szegedy, 2006]

Graph limit and graphon: intuition





Images from [Glasscock, 2015]
Graph limits were introduced by [Lovász and Szegedy, 2006]

Labelling of the nodes gets in the way

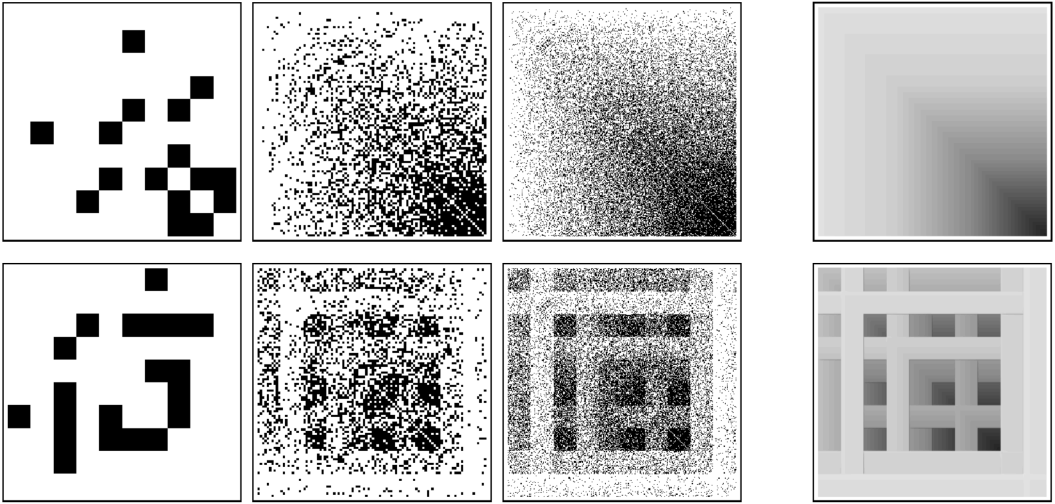


Image from [Orbanz and Roy, 2014]

Labelling of the nodes gets in the way

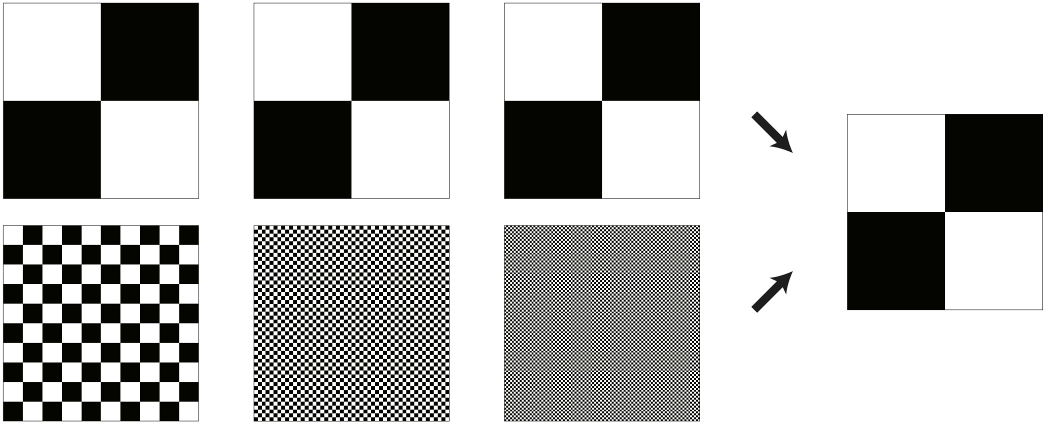


Image from [Glasscock, 2015]

Graph limit and graphon: definition

- **Graphon** [Lovász and Szegedy, 2006]:

$$W : [0,1]^2 \mapsto [0,1] \text{ Lebesque-measurable}$$

Graph limit and graphon: definition

- **Graphon** [Lovász and Szegedy, 2006]:

$$W : [0,1]^2 \mapsto [0,1] \text{ Lebesque-measurable}$$

- **Cut-norm**:

$$\delta_\square(W,U) := \inf_{\varphi,\psi} \sup_{\substack{S,T\subseteq[0,1] \\ \text{measurable}}} \left| \int_{S\times T} W^\varphi(x,y) - U^\psi(x,y)\mathrm{d}x\mathrm{d}y \right|,$$

φ, ψ measure preserving transformations of $[0,1]$.

Graph limit and graphon: definition

- **Graphon** [Lovász and Szegedy, 2006]:

$$W : [0,1]^2 \mapsto [0,1] \text{ Lebesque-measurable}$$

- **Cut-norm**:

$$\delta_{\square}(W,U) := \inf_{\varphi,\psi} \sup_{\substack{S,T\subseteq[0,1] \\ \text{measurable}}} \left| \int_{S\times T} W^{\varphi}(x,y) - U^{\psi}(x,y)\mathrm{d}x\mathrm{d}y \right|,$$

φ, ψ measure preserving transformations of $[0,1]$.

Space of graphons equipped with $\delta_{\square}$ is **compact**

Graph limit and graphon: definition

- **Graphon** [Lovász and Szegedy, 2006]:

$$W : [0,1]^2 \mapsto [0,1] \text{ Lebesque-measurable}$$

- **Cut-norm**:

$$\delta_\square(W,U) := \inf_{\varphi,\psi} \sup_{\substack{S,T\subseteq[0,1] \\ \text{measurable}}} \left| \int_{S\times T} W^\varphi(x,y) - U^\psi(x,y)\mathrm{d}x\mathrm{d}y \right|,$$

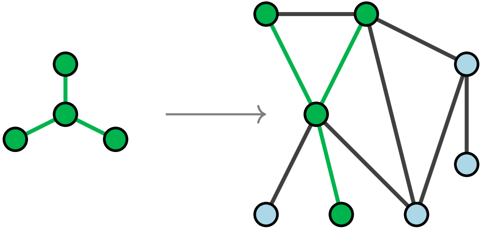
φ, ψ measure preserving transformations of $[0,1]$.

Space of graphons equipped with $\delta_\square$ is **compact**

- **Graph limit**: equivalence class

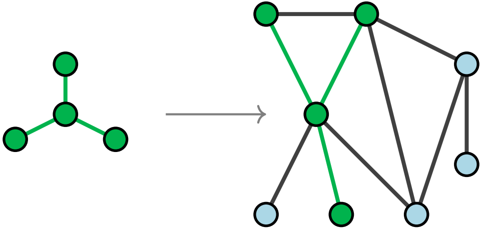
$$[W] = \left\{ W^\varphi : (x,y) \mapsto W(\varphi(x),\varphi(y)) \mid \begin{array}{c}\varphi \text{ an invertible, measure} \\ \text{preserving transformation of } [0,1]\end{array} \right\}$$

Convergence ”in parameters”



Finitely forcible graphons: [Lovász and Szegedy, 2011]

Convergence "in parameters"



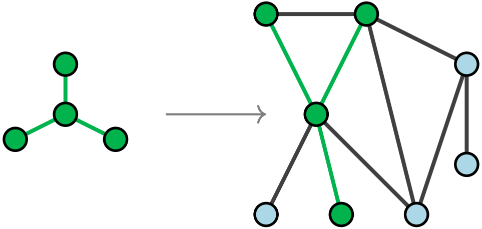
Homomorphism densities: $F = (V, E)$ a simple finite graph, W a graphon

$$t(F, W) := \int_{[0,1]^{|V|}} \prod_{ij \in E_F} W\left(x_i, x_j\right) \prod_{i \in V} \mathrm{d}x_i$$

$\approx$ Expected density of F in W

Finitely forcible graphons: [Lovász and Szegedy, 2011]

Convergence "in parameters"



Homomorphism densities: $F = (V, E)$ a simple finite graph, W a graphon

$$t(F, W) := \int_{[0,1]^{|V|}} \prod_{ij \in E_F} W\left(x_i, x_j\right) \prod_{i \in V} \mathrm{d}x_i$$

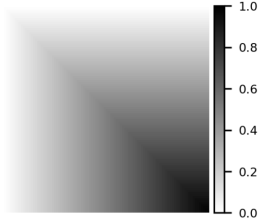
$\approx$ Expected density of F in W

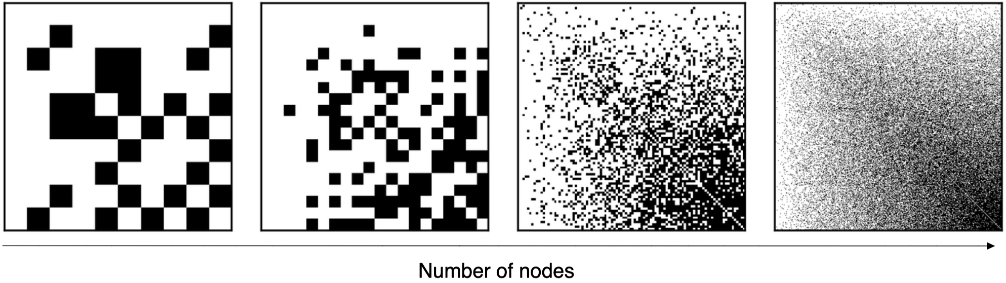
$W_n \to W \Leftrightarrow t(F, W_n) \to t(F, W)$ for all simple graphs F

Finitely forcible graphons: [Lovász and Szegedy, 2011]

Graphon in statistics

When we consider networks, what is the sample size ?





We consider exchangeable simple graphs

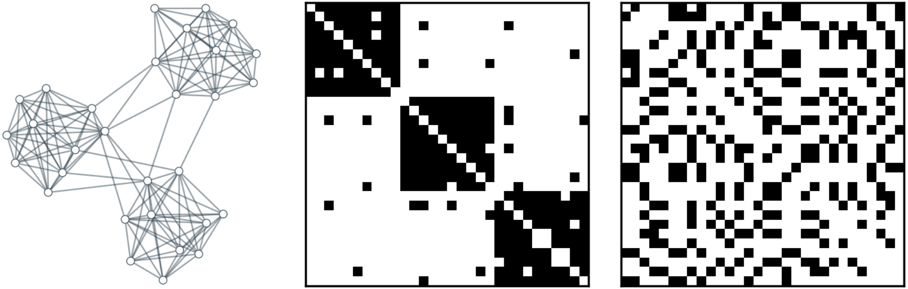


Figure 2: Graph generated from a stochastic block model with 3 groups. Relabelling the nodes hides the structure in the adjacency matrix.

$$\boxed{(A_{ij}) \overset{\mathrm{d}}{=} (A_{\sigma(i)\sigma(j)}) \text{ for all finite permutations } \sigma.}$$

Exchangeability gives us representation

From Aldous-Hoover theorem:

$$U_i \overset{iid}{\sim} \text{Uniform}[0, 1]$$

$$A_{ij} | U_i, U_j \overset{iid}{\sim} \text{Bernoulli}\left(W(U_i, U_j)\right) \quad \text{for } i < j,$$

where $W : [0, 1]^2 \mapsto [0, 1]$ is symmetric, measurable $\rightarrow$ **graphon**.

Exchangeability gives us representation

From Aldous-Hoover theorem:

$$U_i \overset{iid}{\sim} \mathrm{Uniform}[0,1]$$

$$A_{ij} | U_i, U_j \overset{iid}{\sim} \mathrm{Bernoulli}\left(W(U_i, U_j)\right) \quad \text{for } i < j,$$

where $W : [0,1]^2 \mapsto [0,1]$ is symmetric, measurable $\rightarrow$ **graphon**.

All exchangeable graphs come from a graphon.

How to generate a graph given a graphon

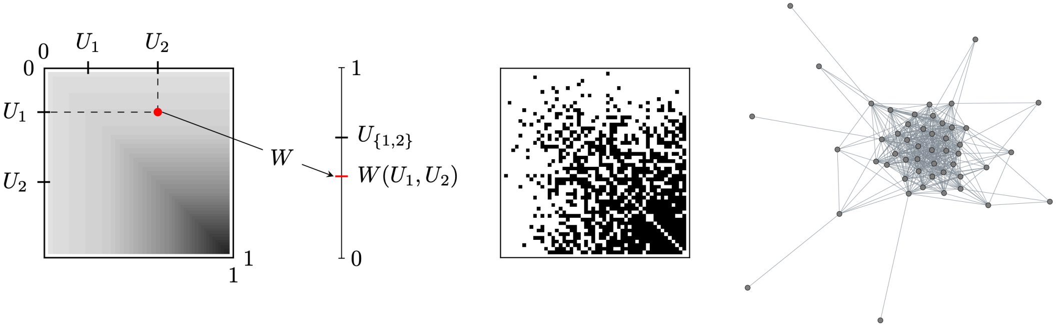
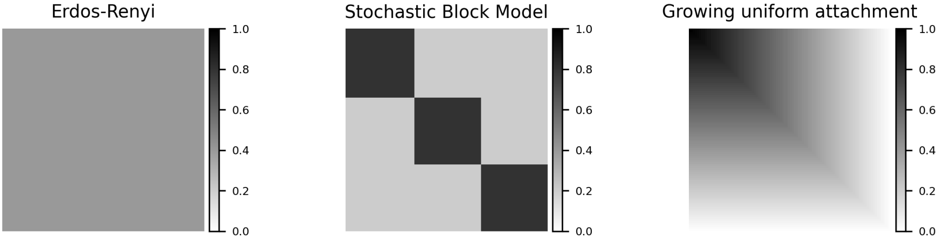


Figure 3: Sampling procedure for node exchangeable graphs. Image from [Orbanz and Roy, 2014].

Example of graphons for common graph models



"Vanilla" graphons cannot generate sparse graphs

Basic graphon theory only covers **dense graphs**:

$$\boxed{\text{If } |E_{G_n}| = o(n^2) \text{(i.e., sparse), then } G_n \to \text{empty graph.}}$$

"Vanilla" graphons cannot generate sparse graphs

Basic graphon theory only covers **dense graphs**:

If $|E_{G_n}| = o(n^2)$(i.e., sparse), then $G_n \to$ empty graph.

solution

$\downarrow$

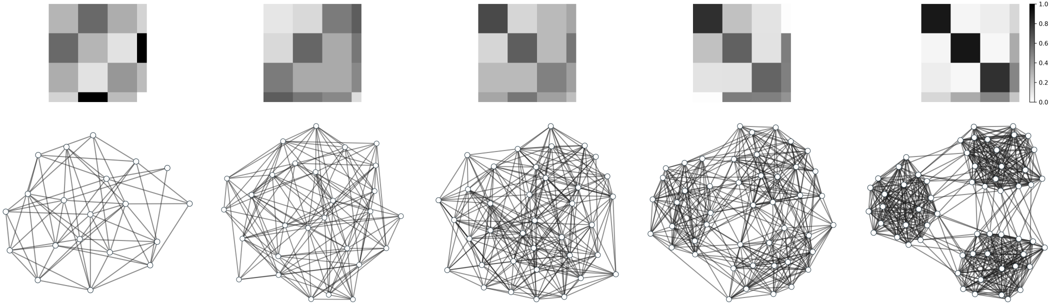
Scaled graphon [Bickel and Chen, 2009]

$$A_{ij}|U_i, U_j \overset{iid}{\sim} \text{Bernoulli}\left(\rho_\text{n} f(U_i, U_j)\right) \quad \text{for } 0 < i < j \leq n,$$

where $f : [0,1]^2 \mapsto \mathbb{R}$ with $\iint_{[0,1]^2} f(x,y)\mathrm{d}x\mathrm{d}y = 1.$

Estimation

Step function approximation



Step function on $[0,1]^2 \to$ Stochastic Block model

How can we estimate graphons ? (network histogram)

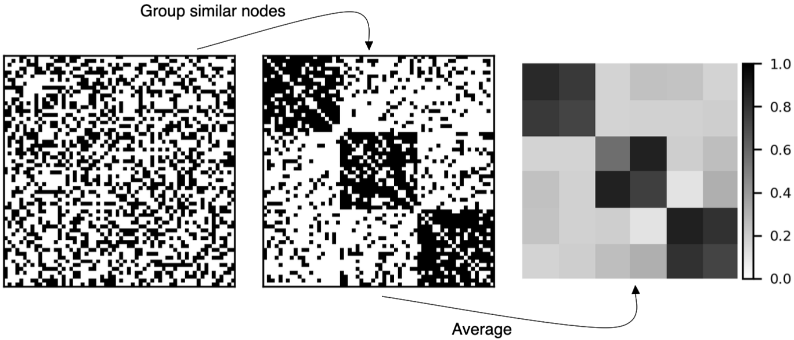


Figure 4: The 2 steps in graphon estimation using **network histogram** [Olhede et al., 2014]. Group sizes are computed from the data.

How can we estimate graphons ? (method of moments)

Method of moments: compare theoretical and empirical homomorphism densities of a set of finite simple graphs to find parameters of SBM.

$$t(F, W_1) = t(F, W_2) \quad \forall F \Rightarrow W_1, W_2 \text{ parametrize the same random graph.}$$

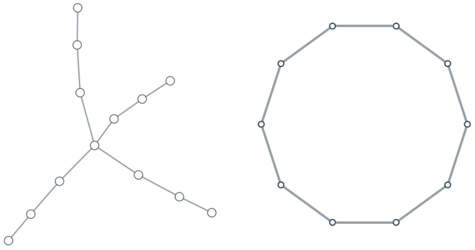


Figure 5: On the left k-l-wheel ($k = 3, l = 4$) used in [Bickel et al., 2011], on the right cycle used in ours.

Summary

1. Graph limit

2. Graphon in statistics

3. Estimation

4. Summary

OK, so what was this talk about?

Three POVs on graphons

- Generative models
- Limits of convergent graph sequences
- Objects identifying family of structurally similar graphs (hom. densities)

OK, so what was this talk about?

Three POVs on graphons

- Generative models
- Limits of convergent graph sequences
- Objects identifying family of structurally similar graphs (hom. densities)

General summary

- Graphons $\rightarrow$ **closure** of space of **exchangeable** finite simple graphs
- Graphons $\leftrightarrow$ exchangeable graphs
- Mostly **theoretical** (e.g., used to study properties of GNN [Keriven et al., 2020, Ruiz et al., 2021b, Ruiz et al., 2021a])

References [1]

Bickel, P. J. and Chen, A. (2009).
A nonparametric view of network models and newman–girvan and other modularities.
Proceedings of the National Academy of Sciences, 106(50):21068–21073.

Bickel, P. J., Chen, A., and Levina, E. (2011).
The method of moments and degree distributions for network models.
The Annals of Statistics, 39(5):2280–2301.

Glasscock, D. (2015).
What is... a graphon.
Notices of the AMS, 62(1).

Janson, S. (2010).
Graphons, cut norm and distance, couplings and rearrangements.
arXiv preprint arXiv:1009.2376.

Keriven, N. (2020).
A short introduction to graphons.
`https://nkeriven.github.io/files/tuto_graphon.pdf`.
Accessed: 2022–14-06.

Keriven, N., Bietti, A., and Vaiter, S. (2020).
Convergence and stability of graph convolutional networks on large random graphs.
Advances in Neural Information Processing Systems, 33:21512–21523.

Lovász, L. and Szegedy, B. (2006).
Limits of dense graph sequences.
Journal of Combinatorial Theory, Series B, 96(6):933–957.

References [2]

Lovász, L. and Szegedy, B. (2011).
Finitely forcible graphons.
Journal of Combinatorial Theory, Series B, 101(5):269–301.

Olhede, S. C., Wolfe, P. J., and Bickel, P. J. (2014).
Network histograms and universality of blockmodel approximation.
Proceedings of the National Academy of Sciences of the United States of America, 111(41):14722–14727.

Orbanz, P. and Roy, D. M. (2014).
Bayesian models of graphs, arrays and other exchangeable random structures.
IEEE transactions on pattern analysis and machine intelligence, 37(2):437–461.

Ruiz, L., Chamon, L. F., and Ribeiro, A. (2021a).
Graphon signal processing.
IEEE Transactions on Signal Processing, 69:4961–4976.

Ruiz, L., Wang, Z., and Ribeiro, A. (2021b).
Graphon and graph neural network stability.
In *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5255–5259. IEEE.

Homomorphism definition

Homomorphisms are adjacency preserving maps from motif $F = (V', E')$ into graph $G = (V, E)$

$$\beta : V' \to V \text{ such that } (i, j) \in E' \text{ implies } (\beta(i), \beta(j)) \in E$$

$\hom(F, G)$ represent the number of homomorphisms between motif F and graph G.

$$t(F, G) = \frac{\hom(F, G)}{n^{n'}}$$

Dense and sparse

Dense

$$\frac{|E(g_m)|}{|V(g_m)|^2} = O(1)$$

Sparse

$$\frac{|E(g_m)|}{|V(g_m)|^2} = o(1)$$

sparse $\Rightarrow \iint_{[0,1]^2} w(x,y)dxdy = 0 \Rightarrow w(x,y) = 0$ a.e.

Network Histogram

$$\hat{z} = \underset{z \in \mathcal{Z}_k}{\operatorname{argmax}} \sum_{i<j} \left\{ A_{ij} \log \bar{A}_{z_i z_j} + (1 - A_{ij}) \log \left(1 - \bar{A}_{z_i z_j}\right) \right\}$$

where for all $1 \leq a, b \leq k$ they define the histogram bin heights

$$\bar{A}_{ab} = \frac{\sum_{i<j} A_{ij} 1_{\{\hat{z}_i = a\}} 1_{\{\hat{z}_j = b\}}}{\sum_{i<j} 1_{\{\hat{z}_i = a\}} 1_{\{\hat{z}_j = b\}}},$$

where $\bar{A}$ an be seen as the estimated connection matrix of the block model conditional on the assignment vector z