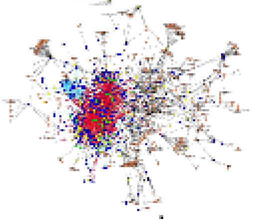
Graphexes and Multi-Graphexes: A Deeper Look

Christian Borgs
Microsoft Research New England

Joint work with
J.T. Chayes, H. Cohn, S. Dhara, L.M. Lovasz, S. Sen and V. Veitch



Summary Part 1:
Graphs and Graphexes

<u>Graph</u> G over a set of vertices V

- Adjacency matrix $A: V \times V \to \{0,1\}$

↓ Sampling Convergence

↑ Graphex Process

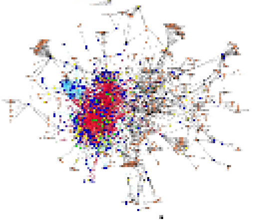
<u>Graphex</u> $\mathbb{W}$ over a σ-finite feature space (Ω, μ)

- Graphon $W: \Omega \times \Omega \to [0,1]$
- Star function $S: \Omega \to \mathbb{R}_+$
- Dust intensity $I \in \mathbb{R}_+$

This Talk

1. Sampling Convergence for Multi-Graphs [BCDS'18]
2. Sampling Convergence for the Configuration Model [BCDS'18]
3. Weak Kernel Metric [BCCL'18]
4. Uniqueness/Identifiability [BCCL'18]

B=Borgs, C=Chayes, C=Cohn, D=Dhara, L=L.M.Lovasz, S=Sen



1. Sampling Convergence for Multi-Graphs

Samples from a Multi-Graph: $\mathrm{Smpl}_k(G)$

Given $k \in \mathbb{N}$ and a finite multi-graph G

- Choose a random map $\phi: [k] \to V(G): i \mapsto x_i$
- connect i and j with the same multiplicity as $x_i x_j \in E(G)$

Poisson Sampling: $\mathrm{Poiss}(G, \kappa)$

- Choose $k = \mathrm{Poiss}(\kappa |V(G)|)$
- Return $\mathrm{Smpl}_k(G)$ after removing isolated vertices

Def: A sequence of multi-graph G_n with $e(G_n)$ non-loop edges is called sampling convergent, if $\mathrm{Poiss}(G_n, t/\sqrt{2e(G_n)})$ converges in distribution for all $t > 0$

Q: What is the limit?

Sampling Convergence for Multi-Graphs

A1: A process (G_t) of random multi-graphs

A2: An extremal, exchangeable random measure ξ on $\mathbb{R}_+^2$

Construction of ξ

- Couple $\mathrm{Pois}(G_n, t/\sqrt{2e(G_n)})$ for different t by using a Poisson process (t_i, v_i) with rate $1/\sqrt{2e(G_n)}$ on $\mathbb{R}_+ \times V(G_n)$, and using the vertices with birthtime $t_i \le t$ to generate $\mathrm{Poiss}(G_n, t/\sqrt{2e(G_n)})$
- Define $\xi(G_n)$ as the random counting measure

$$\xi(G_n) = \sum_{i,j} m_{ij} \delta_{t_i t_j}$$

with m_{ij} being the multiplicity of the edge $v_i v_j$ in G_n

- Take the limit $n \to \infty$

Sampling Convergence for Multi-Graphs

Q: Can we again use Kallenberg's theorem to represent the limiting distribution of random multi-graphs in a compact form?

A: Yes, via multi-graphexes

Def: A multi-graphex $\mathbb{W}$ over a σ-finite measure space $(\Omega, \mathcal{F}, \mu)$ is a triple (W, S, I) consisting of

- a multi-graphon, defined as a symmetric map W from $\Omega \times \Omega$ to the set of probability distributions on $\mathbb{N}_0$, $(x, y) \mapsto W(x, y)$
- a sequence $S = (S_k)_{k \geq 1}$ of star functions $S_k : \Omega \to \mathbb{R}_+$ with $\sum_k S_k(x) < \infty$
- a sequence $I = (I_k)$ of dust intensities $I_k \in \mathbb{R}_+$ with $\sum_k I_k < \infty$

Sampling Convergence for Multi-Graphs

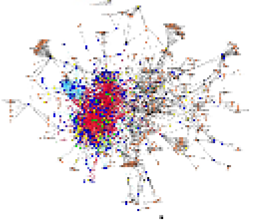
Multi-Graphex Process $G_t(\mathbb{W})$

Create a Poisson process $x_1, x_2, \ldots \in \Omega$ of intensity $t\mu$

- Join i and j with k edges, where $k \sim W(x_i, x_j)$
- For each i and k, add a star with $\mathrm{Poiss}(tS_k(x_i))$ many edges of multiplicity k, $(k \geq 1)$
- Independently, add isolated edges of multiplicity k at rate $t^2 I_k, (k \geq 1)$

Remove all isolated vertices

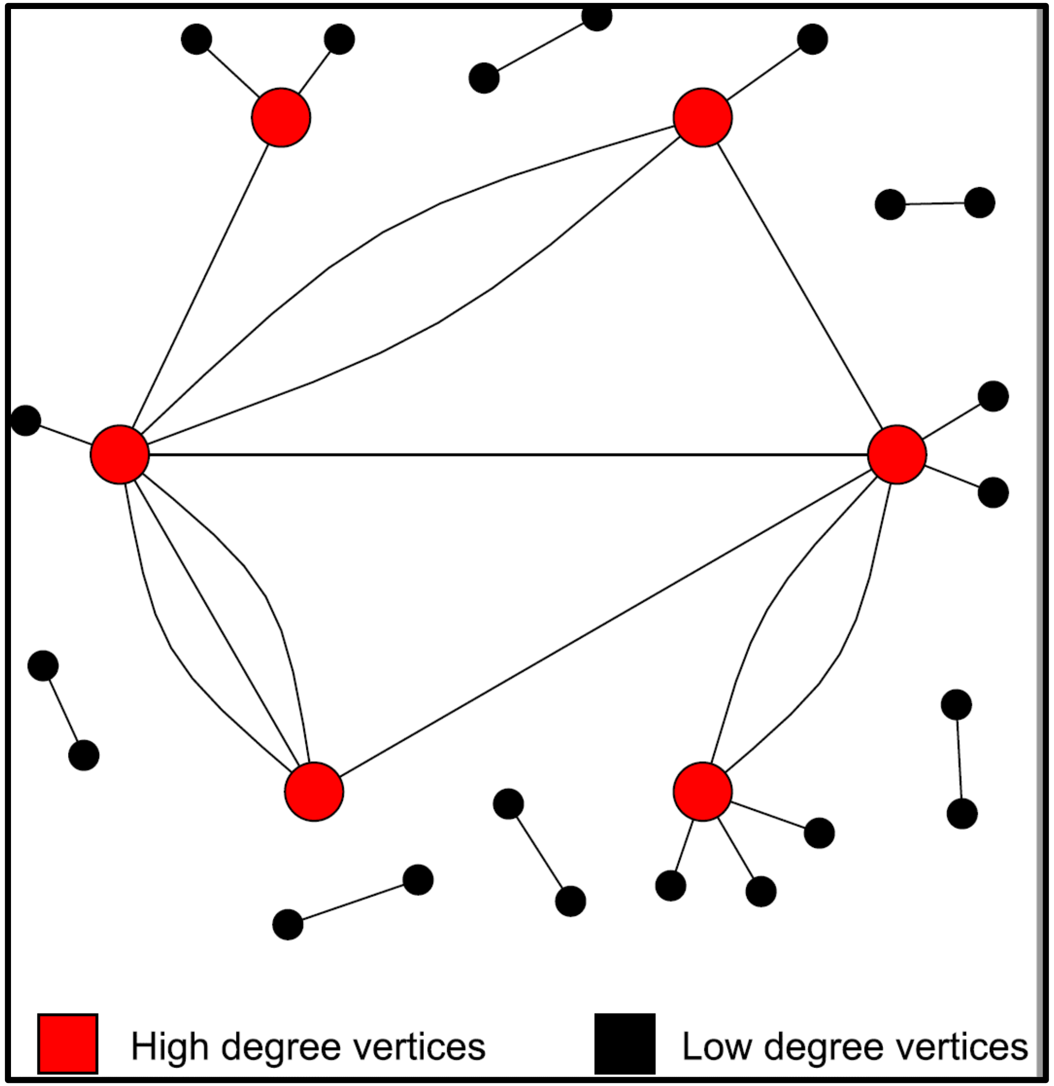
Thm [BCSS'18]: If G_n is sampling convergent there exists an multi-graphex $\mathbb{W}$ s.th. for all t, the limiting distribution of $\mathrm{Poiss}(G_n, \frac{t}{\sqrt{2|E(G_n)|}})$ is equal to that of $G_t(\mathbb{W})$



Sampling Convergence for Multi-Graphs

Resulting Graph Structure

- A core of high degree vertices joined by multi-edges generated by the multi-graphon
- High degree stars with leaves not connected to anything
- A "dust component" of isolated edges



Sampling Convergence for Multi-Graphs

Resulting Graph Structure

- A core of high degree vertices joined by multi-edges generated by the multi-graphon
- High degree stars with leaves not connected to anything
- A "dust component" of isolated edges

Couple $G_t(\mathbb{W})$ for different t so that $G_s(\mathbb{W})$ is an induced subgraph of $G_t(\mathbb{W})$ if $s < t$, and consider $G_\infty(\mathbb{W}) = \bigcup_{t \geq 0} G_t(\mathbb{W})$

- The vertices in the core have infinite degrees,
- The stars become stars with infinitely many leaves,
- The only finite degree vertices have degree 1, and are either a leaf of a star, or an endpoint of an isolated edge.

Sampling Convergence for Multi-Graphs

Q: Why does sampling lead to this strange limit?

Heuristic Picture: We want to analyze $\mathrm{Poiss}(G,p)$, where $G=(V,E)$ has m edges and $p=t/\sqrt{2m}$

Decompose V into a

- core C of vertices of degrees of order $\sqrt{m}$,
- a low degree part L with vertices of degrees $o(\sqrt{m})$,
- a high degree part H of vertices of order $\omega(\sqrt{m})$

Let $C_p \subseteq C$, $L_p \subseteq L$, $H_p \subseteq H$ be the subsets left after sampling

- H contains $o(\sqrt{m})$ vertices, so H_p will be empty
- If C contains $o(\sqrt{m})$ vertices, C_p will be empty as well

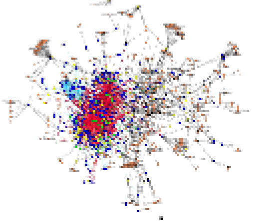
Sampling Convergence for Multi-Graphs

Let $C_p \subseteq C$, $L_p \subseteq L$, $H_p \subseteq H$ be the subsets left after sampling

- If C contains $\theta(\sqrt{m})$ vertices, C_p will contain $p\theta(\sqrt{m}) = \theta(t)$ vertices in expectation, each of expected degrees $\theta(t)$;
- The probability that a vertex $i \in L$ has degree ≥ 2 after sampling is $O\left(p^3 d_i^2\right) = o(p^3 d_i \sqrt{m}) = o(p^2 d_i)$, hence in expectation, the number of vertices of degree ≥ 2 in L_p is of order

$$o(p^2)\sum d_i = o(p^2)m = o(1)$$

- In the limit, the edges with two vertices in L_p will have two endpoints with degree 1 contributing to I, those between L_p and C_p will be stars contributing to S, and those with both endpoints in the core C_p will contribute to W



2. Sampling Convergence for the Configuration Model

Configuration Model: Given n and a sequence $\boldsymbol{d} = (d_i)_{i\in[n]}$ with $\ell_n = \sum d_i$ even, define $\mathbf{CM}_n(\boldsymbol{d})$ as the random multi-graph obtained by randomly pairing all half-edges corresponding to $\boldsymbol{d}$

Degree measure and Levi tail density

$$\rho_n = \frac{1}{\sqrt{\ell_n}} \sum_i \delta_{d_i/\sqrt{\ell_n}} \quad \text{and} \quad \bar{\rho}_n(x) = \int_x^\infty d\rho_n(x')$$

Lemma: In distribution, the "half-degrees" of $\mathrm{Poiss}(\mathbf{CM}_n(\boldsymbol{d}), t/\sqrt{\ell_n})$ can be obtained by

- taking a Poisson process $(x_i)_{i\in\mathbb{N}}$ of intensity t on $\mathbb{R}_+$
- discarding all points not in $[0, n/\sqrt{\ell_n}]$
- setting $d_i = \sqrt{\ell_n}\,(\bar{\rho}_n)^{-1}(x_i)$

Sampling Convergence for the Configuration Model

Proof:

- taking a Poisson process $(x_i)_{i\in\mathbb{N}}$ of intensity t on $\mathbb{R}_+$ and discarding all points not in $[0, n/\sqrt{\ell_n}]$ gives $\mathrm{Poiss}(tn/\sqrt{\ell_n})$ many vertices, as required

$$\rho_n = \frac{1}{\sqrt{\ell_n}} \sum_i \delta_{d_i/\sqrt{\ell_n}} \quad \text{and} \quad \bar{\rho}_n(x) = \int_x^\infty d\rho_n(x')$$

- $\bar{\rho}_n$ is a step function that falls from $\bar{\rho}_n(0) = n/\sqrt{\ell_n}$ to $\bar{\rho}_n(\infty) = 0$
- If there are k vertices with degree $d_i = d$, then at $x = d/\sqrt{\ell_n}$, this step function takes a step of size $-k$
- If we choose $U \in [0, n/\sqrt{\ell_n}]$ uniformly at random, $\sqrt{\ell_n}\bar{\rho}_n^{-1}(U)$ is the degree of a uniformly chosen vertex, as required

Sampling Convergence for the Configuration Model

<u>Thm1</u>: Assume that $\boldsymbol{d} = (d_i)_{i\in[n]}$ is uniformly tail regular and that $\ell_n = \Omega(\log n)$ and $\max d_i = o(\ell_n)$. Then $\mathbf{CM}_n(\boldsymbol{d})$ is sampling convergent if and only if $\rho_n \to \rho$ vaguely*

In this case, the limit is the multi-graphon

$$W(x,y) = \begin{cases} \mathrm{Poiss}(\bar{\rho}^{-1}(x)\ \bar{\rho}^{-1}(y)) & \text{for } x \neq y \\ \mathrm{Poiss}\left(\frac{1}{2}\left(\bar{\rho}^{-1}(x)\right)^2\right) & \text{for } x = y \end{cases}$$

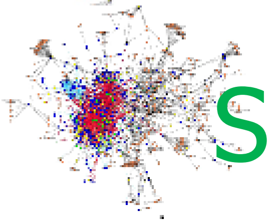
<u>Rem</u>: This gives a "derivation" of the model considered in [CF'14]

---

*As usual, ρ_n is said to converges vaguely to ρ if

$$\int f(x)d\rho_n(dx) \to \int f(x)d\rho(x)$$

for all continuous f with compact support in $(0,\infty)$



Sampling Convergence for the Configuration Model

Proof Sketch:

- If S, S' are disjoint sets of half-edges then the number of edges in $\mathbf{CM}_n(\boldsymbol{d})$ that join half-edges in S and S' is asymptotically distributed according to $\mathrm{Poiss}(|S||S'|/\ell_n)$
- Thus the distribution of the number of edges between i and j behaves asymptotically like $\mathrm{Poiss}(d_i d_j/\ell_n)$
- Recall that the degree distribution of $\mathrm{Poiss}(\mathbf{CM}_n(\boldsymbol{d}), t/\sqrt{\ell_n})$ can be obtained by taking a Poisson process $(x_i)_{i\in\mathbb{N}}$ of intensity t on $\mathbb{R}_+$ and setting $d_i = \sqrt{\ell_n}\,(\bar{\rho}_n)^{-1}(x_i)$
- Asymptotically, the number of edges between i and j thus behaves like $\mathrm{Poiss}(\bar{\rho}^{-1}(x_i)\,\bar{\rho}^{-1}(x_j)) = W(x_i, x_j)$

Sampling Convergence for the Configuration Model

Proof Sketch:

- In a similar way, the multiplicity of the loops at i behaves like

$$\mathrm{Poiss}\left(\tfrac{1}{2}\left(\bar{\rho}^{-1}(x_i)\right)^2\right) = W(x_i, x_i)$$

Sampling Convergence for the Configuration Model

Q: What happens if we don't assume uniform tail regularity, i.e., if there are too many small degree vertices?

Thm2: Assume that $\ell_n = \Omega(\log n)$ and $\max d_i = o(\ell_n)$. Then $\mathbf{CM}_n(\boldsymbol{d})$ is sampling convergent if and only if $\rho_n \to \rho$ vaguely and the limit $a = \lim_{n\to\infty} \int \min\{x,1\}\, d\rho_n(x) - \int \min\{x,1\}\, d\rho(x)$ exists

In this case, the limiting multi-graphex is given by

$$W(x,y) = \begin{cases} \text{Poiss}(\bar{\rho}^{-1}(x)\ \bar{\rho}^{-1}(y)) & \text{for } x \neq y \\ \text{Poiss}(\frac{1}{2}(\bar{\rho}^{-1}(x))^2) & \text{for } x = y \end{cases}$$

$$S_k(x) = a\bar{\rho}^{-1}(x)\delta_{k,1} \text{ and } I_k = \frac{a^2}{2}\delta_{k,1}$$

Sampling Convergence for the Configuration Model

Proof Sketch:

- $\rho_n = \frac{1}{\sqrt{\ell_n}} \sum_i \delta_{d_i/\sqrt{\ell_n}}$, so $\int x d\rho_n(x) = \frac{1}{\ell_n} \sum_i d_i = 1$
- By Fatou, $\int x d\rho(x) \leq 1$, so large x don't contribute much to $\int \min\{x, 1\}\, d\rho_n(x)$ or $\int \min\{x, 1\}\, d\rho(x)$
- Thus $a = \lim_{n\to\infty} \int \min\{x, 1\}\, d\rho_n(x) - \int \min\{x, 1\}\, d\rho(x)$ represents the fraction of half-edges with degrees $o(\sqrt{\ell_n})$
- The terms $S_k(x) = a\bar{\rho}^{-1}(x)\delta_{k,1}$ and $I_k = \frac{a^2}{2}\delta_{k,1}$ represent the edges generated between the core and the low degree vertices, and the edges joining two low degree vertices, resp.

3. Weak Kernel Metric

Question: Is there a metric describing sampling convergence

A1: Yes, for abstract reasons:

- Sampling convergence is equivalent to convergence of the graphex process generated by the stretched empirical graphon
- This is equivalent to convergence of the corresponding locally finite random measures on $\mathbb{R}_+^2$
- By the abstract theory of random measures, this can be metricized to make the space a complete, separable metric space
- Following [J'17] and [VR'16], we call this the metric of GP convergence δ_{GP}

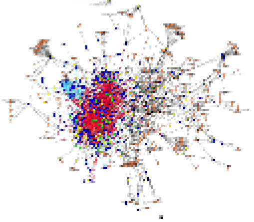
Weak Kernel Metric

<u>Question</u>: Can we relate δ_{GP} to a more concrete metric on graphexes, comparable to $\delta_{\square}$, and then establish things like a regularity lemma, a counting lemma, and a sampling lemma?

<u>Step 1</u>: For the graphon part, use the kernel norm instead of the cut-norm: Given a σ-finite measure space $(\Omega, \mathcal{F}, \mu)$, and a function $W: \Omega \times \Omega \to \mathbb{R}$ define

$$\|W\|_{2\to 2} = \sup_{g,f} \frac{\int f(x) W(x,y) g(y) d\mu(x) d\mu(y)}{\|f\|_2 \|g\|_2}$$

which is the operator norm if we consider W as the kernel of an integral operator on $L^2(\Omega)$



Weak Kernel Metric

Step 2: Given a σ-finite measure space $(\Omega, \mathcal{F}, \mu)$, and two graphexes $\mathbb{W}_1 = (\lambda_1, S_1, W_1)$ and $\mathbb{W}_2 = (\lambda_2, S_2, W_2)$ let

$$d_{2\to2}(\mathbb{W}_1, \mathbb{W}_2) = \|W_1 - W_2\|_{2\to2} + \|S_1 - S_2\|_2 + |\lambda_1 - \lambda_2|$$

More precisely, take the max. of

$$\|W_1, -W_2\|_{2\to2}, \sqrt{\left\|D_{\mathbb{W}_1} - D_{\mathbb{W}_2}\right\|_2} \text{ and } \sqrt[3]{\left|\|\mathbb{W}_1\|_1 - \|\mathbb{W}_1\|_1\right|}$$

where

$$D_{\mathbb{W}_i}(x) = S(x) + \int W_i(x,y)\, d\mu_i(y)$$

Weak Kernel Metric

Step 3: Use coupling for two graphexes $\mathbb{W}_1 = (\lambda_1, S_1, W_1)$ and $\mathbb{W}_2 = (\lambda_2, S_2, W_2)$ over different spaces with total measure ∞.

- a coupling of $(\Omega_1, \mathcal{F}_1, \mu_1)$ and $(\Omega_2, \mathcal{F}_2, \mu_2)$ is a measure μ on $\Omega_1 \times \Omega_2$ such that $\mu(A \times \Omega_2) = \mu_1(A)$ and $\mu(\Omega_1 \times B) = \mu_2(B)$
- use coordinate projections $\pi_i : \Omega_1 \times \Omega_2 \to \Omega_i$ to pull $\mathbb{W}_i$ back to $\Omega_1 \times \Omega_2$: $\mathbb{W}_i^{\pi_i} = \left(\lambda_i, S_i^{\pi_i}, W_i^{\pi_i}\right)$ where $S_i^{\pi_i}((x_1, x_2)) = S_i\left(\pi_i(x_1, x_2)\right) = S_i(x_i)$ and similarly for $W_i^{\pi_i}$
- Define the kernel metric as an infimum over all couplings

$$\delta_{2\to2}(\mathbb{W}_1, \mathbb{W}_2) = \inf_{\mu} \mathrm{d}_{2\to2}(\mathbb{W}_1^{\pi_1}, \mathbb{W}_2^{\pi_2})$$

Weak Kernel Metric

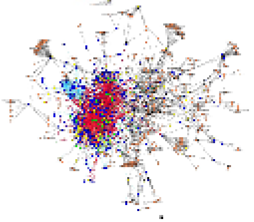
Step 4: If $(\Omega_1, \mathcal{F}_1, \mu_1)$ or $(\Omega_2, \mathcal{F}_2, \mu_2)$ has finite measure, extend it by appending an arbitrary σ-finte space of infinite measure, e.g., $\mathbb{R}_+$

Problem: $\delta_{2\to 2}(\mathbb{W}_1, \mathbb{W}_2)$ is not necessarily finite and even when restricted to graphexes where it is finite, it gives a different topology than δ_{GP}

Definition: Given two graphexes $\mathbb{W}_1$ and $\mathbb{W}_2$ over $(\Omega_1, \mathcal{F}_1, \mu_1)$ and $(\Omega_2, \mathcal{F}_2, \mu_2)$, define the weak kernel distance $\delta(\mathbb{W}_1, \mathbb{W}_2)$ as the infimum over all ϵ s.th. there exists measures μ_i' with

$$d_{TV}(\mu_i, \mu_i') \le \epsilon \text{ and } \delta_{2\to 2}(\mathbb{W}_1', \mathbb{W}_2') \le \epsilon$$

where $\mathbb{W}_i'$ is obtained from $\mathbb{W}_i$ by replacing μ_i with μ_i'



Weak Kernel Metric

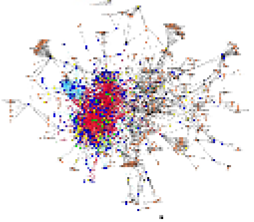
<u>Thm</u> [BCCL'17]: The weak kernel metric δ metric is equivalent to δ_{GP}, the metric of graphex process convergence

<u>Proof Strategy:</u>

Establish

- a regularity lemma (showing compactness of the space of graphexes with $\|\mathbb{W}\|_1 \leq 1$ equipped with the weak kernel metric)
- a counting lemma (proving that convergence of W_n w.r.t. δ implies convergence in distribution of $G_t(W_n)$ for all t)
- a sampling lemma (proving the converse)

Details are complicated and require jumping back and forth between δ and $\delta_{2\to 2}$



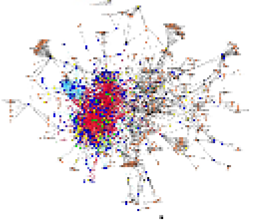
4. Uniqueness/Identifiability

Q1: Given the distribution of $G_t(\mathbb{W})$ for all t, can we determine $\mathbb{W}$ or the underlying space $(\Omega,,\mathcal{F},\mu)$?

A1: No! Take a second σ-finite measure space $(\Omega',,\mathcal{F}',\mu')$ and consider a measure preserving map $\phi\colon\Omega'\to\Omega$ and the pullback $\mathbb{W}^{\phi}=\left(\lambda,S^{\phi},W^{\phi}\right)$ where $S^{\phi}(x)=S(\Phi(x))$ and $W^{\phi}(\mathrm{x},\mathrm{y})=\mathrm{W}\big(\phi(x),\phi(y)\big)$.

Then $G_t(\mathbb{W})=_d G_t(\mathbb{W}')$

Q2: Is it enough to consider (maybe a sequence of) pullbacks?



Uniqueness/Identifiability

Def: Given a graphex $\mathbb{W} = (\lambda, S, W)$ define its degree support as

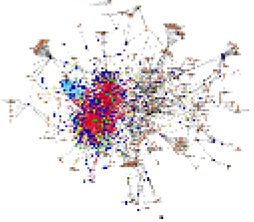
$$dsupp(\mathbb{W}) = \{x \in \Omega : S(x) + \int W(x,y)d\mu(y) > 0\}$$

Thm: The following are equivalent

- $G_t(\mathbb{W}) =_d G_t(\mathbb{W}')$ for all $t > 0$
- $\delta(\mathbb{W}, \mathbb{W}') = 0$
- The restrictions of $\mathbb{W}$ and $\mathbb{W}'$ to their degree supports are pullbacks of a third graphon $\mathbb{U}$: $\mathbb{W}|_{dsupp(\mathbb{W})} = \mathbb{U}^{\phi}$ and $\mathbb{W}'|_{dsupp(\mathbb{W}')} = \mathbb{U}^{\phi'}$ for two measure preserving transformations from $dsupp(\mathbb{W})$ and $dsupp(\mathbb{W}')$ to the space $\mathbb{U}$ is defined on

Summary

- Sampling convergence generalizes the notion of left convergence from dense to sparse graphs and multi-graphs
- The limit objects for simple graphs are graphexes, consisting of a graphon over a σ-finite measure space, a star function, and a dust intensity
- For multi-graphs, the limiting object is a multi-graphex, consisting of a multi-graphon (describing a probability distribution on $\mathbb{N}_+$), a sequence of star functions, and a sequence of dust intensities
- For the configuration model, the limiting graphex takes a simple product form, describing connections within a high degree core, the core and low degree vertices, and between low degree vertices



Summary

- We introduce a new metric, the weak kernel metric, which metricizes the topology of sampling convergence for simple graphs
- The limit of a sampling convergent sequence is unique up to measure preserving transformations, once we restrict ourselves to the degree support of the limiting graphex

Thank you

for your attention!