(Sparse) exchangeable random graphs

Victor Veitch

based on joint work with:

Daniel M. Roy, Christian Borgs, Jennifer Chayes, Henry Cohn

Big Picture

Motivating Problem

Need families of random graphs for modelling network structures

Example

- network: friendships among n users of a social network
- model: family of random graphs $(G_n)_{n \in \mathbb{N}}$

Sparse Networks

Real world networks are sparsely connected

Random graph models should be *sparse*: $o(n^2)$ edges as number of observed nodes n becomes large.

Problem

No general framework for the statistical analysis of sparsely connected networks.

Sparse Networks

Real world networks are sparsely connected

Random graph models should be *sparse*: $o(n^2)$ edges as number of observed nodes n becomes large.

Problem

No general framework for the statistical analysis of sparsely connected networks.

Sparse Graphs

We have no satisfactory answers to some fundamental questions:

1. how should we parameterize the space of distributions on sparse graphs?
2. what can we learn about a large graph if we observe only a small subgraph? and what do we mean when we say "observe"?

Results (VR 2015)

Results

We derive and study a general class of random graphs suitable for modelling network structures.

Special cases

- All dense (graphon) models
- Sparse graphs with e.g. small world and power law behaviour
- Caron & Fox models

Results (VR 2015)

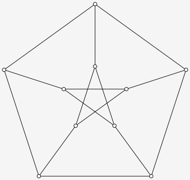
Results

We derive and study a general class of random graphs suitable for modelling network structures.

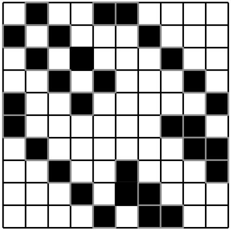
Special cases

- All dense (graphon) models
- Sparse graphs with e.g. small world and power law behaviour
- Caron & Fox models

Graphs, Adjacency, and Pixel Pictures



$$
\begin{matrix}
0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0
\end{matrix}
$$



[Lov12]

Dense Graphs

Graph models

- Basic object: infinite random binary matrix (X_{ij})
- Random graphs: $(G_n)_{n \in \mathbb{N}}$ defined by taking adjacency matrix to be upper left $n \times n$ submatrix

Joint exchangeability of infinite random matrices

$(X_{ij}) \stackrel{d}{=} (X_{\sigma(i)\sigma(j)})$ for all permutations $\sigma \in S_\infty$ of the positive integers

Dense Graphs

Graph models

- Basic object: infinite random binary matrix (X_{ij})
- Random graphs: $(G_n)_{n\in\mathbb{N}}$ defined by taking adjacency matrix to be upper left $n \times n$ submatrix

Joint exchangeability of infinite random matrices

$(X_{ij}) \stackrel{d}{=} (X_{\sigma(i)\sigma(j)})$ for all permutations $\sigma \in S_\infty$ of the positive integers

Graphon Generative Model

Aldous-Hoover (translated)

Every infinite exchangeable array has a representation in terms of a *graphon* $W : [0,1]^2 \to [0,1]$

Given a graphon $W : [0,1]^2 \to [0,1]$, sample a random graph by:

1. Assign each vertex i an iid $U[0,1]$ latent random variable U_i
2. Include each edge (i,j) independently with probability $W(U_i, U_j)$

Representation Theorem

Recipe for constructing statistical models

- Assume a probabilistic symmetry on some infinite random structure
- Associated representation theorem picks out privileged family of distributions

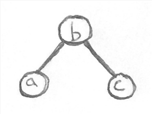
Examples

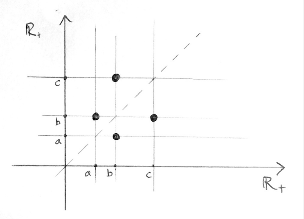
- de Finetti's representation theorem for exchangeable sequences
- Aldous-Hoover-Kallenberg theorem for exchangeable arrays

Caron & Fox 2014: (Sparse) Exchangeable Graphs

Key insights

- adjacency matrix $\rightarrow$ point process on $\mathbb{R}_+^2$
- array joint exchangeability $\rightarrow$ point process joint exchangeability





Graph edges correspond to points on $\mathbb{R}_+^2$

(Sparse) Graph Representation Theorem

Setup

- Random structure: point process on $\mathbb{R}_+^2$
- Finite graph Γ_s: truncate to $[0, s]^2$
- Symmetry: joint exchangeability of point process

Representation theorem*

Distribution characterized by a *graphex*: (I, S, W) where $I \in \mathbb{R}_+$, $S : \mathbb{R}_+ \to \mathbb{R}_+$, and the *graphon* $W : \mathbb{R}_+^2 \to [0, 1]$.
W is symmetric, and S and W satisfy certain weak integrability conditions (integrability suffices).

For this talk, restrict to $(0, 0, W)$ with W integrable.

(Sparse) Graph Representation Theorem

Setup

- Random structure: point process on $\mathbb{R}_+^2$
- Finite graph Γ_s: truncate to $[0, s]^2$
- Symmetry: joint exchangeability of point process

Representation theorem*

Distribution characterized by a *graphex*: (I, S, W) where $I \in \mathbb{R}_+$, $S : \mathbb{R}_+ \to \mathbb{R}_+$, and the *graphon* $W : \mathbb{R}_+^2 \to [0, 1]$.
W is symmetric, and S and W satisfy certain weak integrability conditions (integrability suffices).

For this talk, restrict to $(0, 0, W)$ with W integrable.

(Sparse) Graph Representation Theorem

Setup

- Random structure: point process on $\mathbb{R}_+^2$
- Finite graph Γ_s: truncate to $[0, s]^2$
- Symmetry: joint exchangeability of point process

Representation theorem*

Distribution characterized by a *graphex*: (I, S, W) where $I \in \mathbb{R}_+$, $S : \mathbb{R}_+ \to \mathbb{R}_+$, and the *graphon* $W : \mathbb{R}_+^2 \to [0, 1]$.
W is symmetric, and S and W satisfy certain weak integrability conditions (integrability suffices).

For this talk, restrict to $(0, 0, W)$ with W integrable.

Graphex Model

Generative Model

Given W, an infinite random graph is sampled by:

1. Sample a (latent) unit rate Poisson process Π on $\theta \times \vartheta$.
2. For each pair of points (θ_i, ϑ_i), $(\theta_j, \vartheta_j) \in \Pi$ include edge (θ_i, θ_j) with probability $W(\vartheta_i, \vartheta_j)$.
3. Include θ_i as a vertex whenever θ_i participates in at least one edge.

Finite graph Γ_s is given by restricting to $[0, s]^2$

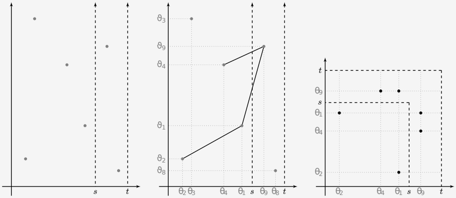


Figure: Graphex Model

Sampling Distribution Results

Given graphon W we know:

1. the expected number of vertices and edges as a function of the size s
2. the asymptotic degree distribution for certain families of graphexes
3. the asymptotic connectivity structure for certain families of graphexes.

Punchline

These models include a wide range of interesting graphs.

Example - Dense Exchangeable Graphs

Let $\widetilde{W} : [0,1]^2 \to [0,1]$ and let

$$W(x,y) = \begin{cases} \widetilde{W}(x,y) & x \leq 1, y \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Generative model for Γ_s

$$N_s \sim \mathrm{Poi}(s) \tag{1}$$

$$\{\theta_i\} \mid N_s \overset{iid}{\sim} \mathrm{Uni}[0,s] \tag{2}$$

$$\{\vartheta_i\} \mid N_s \overset{iid}{\sim} \mathrm{Uni}[0,1] \tag{3}$$

$$(\theta_i, \theta_j) \mid \widetilde{W}, \vartheta_i, \vartheta_j \overset{ind}{\sim} \mathrm{Bern}(\widetilde{W}(\vartheta_i, \vartheta_j)). \tag{4}$$

Example - Slow Decay

$$W(x,y) = \begin{cases} 0 & x = y \\ (x+1)^{-2}(y+1)^{-2} & \text{otherwise} \end{cases}$$

Properties

$$\mathbb{E}[e_s] = \frac{1}{2}s^2$$

$$\mathbb{E}[v_s] \sim \sqrt{\frac{\pi}{3}s^{3/2}}, s \to \infty$$

$$\mathrm{P}(D_s = k \mid \Gamma_s) \xrightarrow{\mathrm{p}} \frac{\Gamma\left(-\frac{1}{2}+k\right)}{2\sqrt{\pi}k!}, s \to \infty$$

$$\sim k^{-3/2}, k \to \infty$$

Example - Fast Decay

$$W(x,y) = \begin{cases} 0 & x = y \\ e^{-x}e^{-y} & \text{otherwise} \end{cases}$$

Properties

$$\mathbb{E}\left[e_s\right] = \frac{1}{2}s^2$$

$$\mathbb{E}\left[v_s\right] \sim s\log s \ , s \to \infty$$

$$\mathrm{P}\left(D_s \leq s^\beta \mid \Gamma_s\right) \xrightarrow{\mathrm{p}} \beta \ , s \to \infty$$

VR16: Sampling and Estimation

Sampling

Let $\mathscr{G}(\Gamma_s)$ be the unlabeled graph associated with Γ_s

Definition (p-sampling)

A p-sampling of a graph G is a random subgraph given by selecting each vertex of G independently with probability p, and then returning the induced edge set.

Theorem

Let $(\Gamma_s)_{s\in\mathbb{R}_+}$ be generated by W. If G_r is an r/s-sampling of $\mathscr{G}(\Gamma_s)$, then

$$G_r \overset{d}{=} \mathscr{G}(\Gamma_s)$$

Sampling

Let $\mathscr{G}(\Gamma_s)$ be the unlabeled graph associated with Γ_s

Definition (p-sampling)

A p-sampling of a graph G is a random subgraph given by selecting each vertex of G independently with probability p, and then returning the induced edge set.

Theorem

Let $(\Gamma_s)_{s\in\mathbb{R}_+}$ be generated by W. If G_r is an r/s-sampling of $\mathscr{G}(\Gamma_s)$, then

$$G_r \overset{d}{=} \mathscr{G}(\Gamma_s)$$

Estimation

Problem

How can we estimate a graphon?

Setup

- Observation: $G_s = \mathscr{G}(\Gamma_s)$ (and s)
- Estimator: $\widehat{W}_{(G_s, s)} : \mathbb{R}_+^2 \to [0, 1]$

Estimation

Problem

How can we estimate a graphon?

Setup

- Observation: $G_s = \mathscr{G}(\Gamma_s)$ (and s)
- Estimator: $\widehat{W}_{(G_s,s)} : \mathbb{R}_+^2 \to [0,1]$

Dense Graph Estimation

Empirical Graphon

Let $\widetilde{W}_G : [0,1]^2 \to \{0,1\}$ be the step function corresponding to the (arbitrarily permuted) adjacency matrix of G.

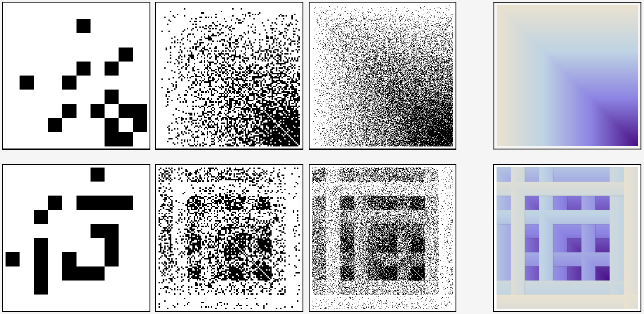


Figure: Convergence of Empirical Graphon (from Orbanz Roy 2015)

Formalization

Definition

Write $\mathrm{uKEG}(W, s) = \mathrm{P}(\mathscr{G}(\Gamma_s) \in \cdot)$ for the distribution of the unlabeled size-s graph generated by W.

Definition

Write $W_k \to_{\mathrm{GP}} W$ as $k \to \infty$, when $\mathrm{uKEG}(W_k, s) \to \mathrm{uKEG}(W, s)$ weakly as $k \to \infty$, for all $s \in \mathbb{R}_+$.

Theorem (Kallenberg 1999)

Let $(\Gamma_s)_{s \in \mathbb{R}_+}$ be generated by compactly supported W, let $s_1, s_2, \ldots \uparrow \infty$ and let $G_k = \mathscr{G}(\Gamma_{s_k})$ for all $k \in \mathbb{N}$. Then $\widetilde{W}_{G_k} \to_{\mathrm{GP}} W$ a.s., as $k \to \infty$.

Formalization

Definition

Write $\mathrm{uKEG}(W, s) = \mathrm{P}(\mathscr{G}(\Gamma_s) \in \cdot)$ for the distribution of the unlabeled size-s graph generated by W.

Definition

Write $W_k \to_{\mathrm{GP}} W$ as $k \to \infty$, when $\mathrm{uKEG}(W_k, s) \to \mathrm{uKEG}(W, s)$ weakly as $k \to \infty$, for all $s \in \mathbb{R}_+$.

Theorem (Kallenberg 1999)

Let $(\Gamma_s)_{s \in \mathbb{R}_+}$ be generated by compactly supported W, let $s_1, s_2, \ldots \uparrow \infty$ and let $G_k = \mathscr{G}(\Gamma_{s_k})$ for all $k \in \mathbb{N}$. Then $\widetilde{W}_{G_k} \to_{\mathrm{GP}} W$ a.s., as $k \to \infty$.

Estimation with Known Sizes

Definition (Dilated empirical graphon)

The s-dilated empirical graphon of G, $\widehat{W}_{(G,s)} : [0, \frac{v_s}{s}] \to \{0,1\}$, is defined by

$$\widehat{W}_{(G,s)}(x,y) = \widetilde{W}_G(\frac{v_s}{s}x, \frac{v_s}{s}y)$$

Observations

Let $(\Gamma_s)_{s\in\mathbb{R}_+}$ be generated by compactly supported W, let $s_1, s_{2,} \ldots \uparrow \infty$ and let $G_k = \mathscr{G}(\Gamma_{s_k})$ for all $k \in \mathbb{N}$.

Theorem (Estimation with known sizes)

$\widehat{W}_{(G_k,s_k)} \to_{\mathrm{GP}} W$ in probability as $k \to \infty$.

Estimation with Known Sizes

Definition (Dilated empirical graphon)

The s-dilated empirical graphon of G, $\widehat{W}_{(G,s)} : [0, \frac{v_s}{s}] \to \{0,1\}$, is defined by

$$\widehat{W}_{(G,s)}(x,y) = \widetilde{W}_G(\frac{v_s}{s}x, \frac{v_s}{s}y)$$

Observations

Let $(\Gamma_s)_{s\in\mathbb{R}_+}$ be generated by compactly supported W, let $s_1, s_2, \ldots \uparrow \infty$ and let $G_k = \mathscr{G}(\Gamma_{s_k})$ for all $k \in \mathbb{N}$.

Theorem (Estimation with known sizes)

$\widehat{W}_{(G_k,s_k)} \to_{\mathrm{GP}} W$ in probability as $k \to \infty$.

Convergence of Dilated Empirical Graphon

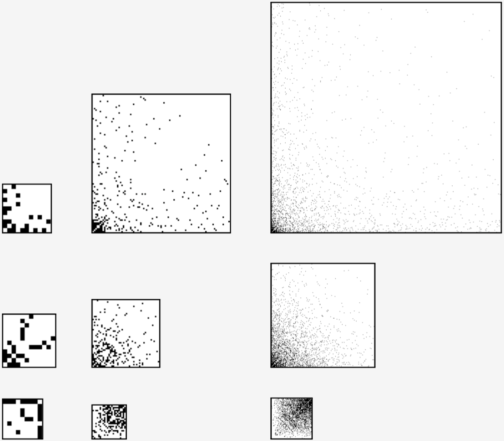


Figure: Convergence of the dilated empirical graphons. Generating graphon in rightmost column. Each square shows a realization of the s_k-dilated empirical graphon of G_{s_k}.

Estimation with Unknown Sizes

What if the observations are just $G_1, G_2, \ldots$? i.e., what if the sizes $(s_k)_{k \in \mathbb{N}}$ aren't included?

Fact

If $W^c(\cdot,\cdot) = W(\cdot/c, \cdot/c)$ then $\mathrm{uKEG}(W, s) = \mathrm{uKEG}(W^c, s/c)$.

Corollary

The dilation of W is not identifiable.

Estimation with Unknown Sizes

What if the observations are just $G_1, G_2, \ldots$? i.e., what if the sizes $(s_k)_{k \in \mathbb{N}}$ aren't included?

Fact

If $W^c(\cdot,\cdot) = W(\cdot/c, \cdot/c)$ *then* $\mathrm{uKEG}(W, s) = \mathrm{uKEG}(W^c, s/c)$.

Corollary

The dilation of W is not identifiable.

Estimation with Unknown Sizes

We need a notion of estimation that is 'dilation invariant':

Definition

The graph sequence $\mathbb{G}(\Gamma)$ of Γ is the ordered sequence of all distinct unlabeled graph structures of $(\mathscr{G}(\Gamma_s))_{s\in\mathbb{R}_+}$

Definition

Write $W_k \to_{\mathrm{GS}} W$ as $k \to \infty$ when $\mathbb{G}(\Gamma^k) \xrightarrow{d} \mathbb{G}(\Gamma)$ as $k \to \infty$, for Γ^k generated by W_k and Γ generated by W

Estimation with Unknown Sizes

We need a notion of estimation that is 'dilation invariant':

Definition

The graph sequence $\mathbb{G}(\Gamma)$ of Γ is the ordered sequence of all distinct unlabeled graph structures of $(\mathscr{G}(\Gamma_s))_{s\in\mathbb{R}_+}$

Definition

Write $W_k \to_{\mathrm{GS}} W$ as $k \to \infty$ when $\mathbb{G}(\Gamma^k) \xrightarrow{d} \mathbb{G}(\Gamma)$ as $k \to \infty$, for Γ^k generated by W_k and Γ generated by W

Estimation with Unknown Sizes

Intuitively, our estimator is the dilated empirical graphon up to dilation:

Theorem (Estimation with unknown sizes)

$\widetilde{W}_{G_k} \to_{\mathrm{GS}} W$ *in probability as* $k \to \infty$

BCCV16+: Sampling Perspectives on Exchangeability and Graph Limits

Background

In the dense graph setting, graphons were independently discovered two different ways:

- The representation theorem work of Aldous–Hover–Kallenberg, and
- The graph limits of Lovasz, Sós, Szegedy, Borgs, Chayes, Vesztergombi, Schrijver, Freedman

These perspectives were unified by Diaconis and Janson (2008)

Concurrent with Veitch–Roy 2015,
Borgs, Chayes, Cohn, and Holden introduced a generalization of dense graph limits where the limit object are the same graphons as VR15.
Also based on Caron–Fox, and has the same graphs as point sets flavour

Background

In the dense graph setting, graphons were independently discovered two different ways:

- The representation theorem work of Aldous–Hover–Kallenberg, and
- The graph limits of Lovasz, Sós, Szegedy, Borgs, Chayes, Vesztergombi, Schrijver, Freedman

These perspectives were unified by Diaconis and Janson (2008)

Concurrent with Veitch–Roy 2015,
Borgs, Chayes, Cohn, and Holden introduced a generalization of dense graph limits where the limit object are the same graphons as VR15.
Also based on Caron–Fox, and has the same graphs as point sets flavour

Questions

1. Is there an analogous unification in this setting?
2. There are several sparse graph limit theories — why this one?
3. Why do we represent random graphs as point processes?
4. What does exchangeability mean in our setting?

Contribution

recast the core ideas—graph limits, network modeling, and exchangeability—in terms of p-sampling.

Questions

1. Is there an analogous unification in this setting?
2. There are several sparse graph limit theories — why this one?
3. Why do we represent random graphs as point processes?
4. What does exchangeability mean in our setting?

Contribution

recast the core ideas—graph limits, network modeling, and exchangeability—in terms of p-sampling.

Sampling Convergence

Definition

A sequence of graphs $G_1, G_2, \ldots$ is sampling convergent if the $r/\sqrt{2e_j}$-samplings of G_j converge in distribution as $j \to \infty$, for all $r \in \mathbb{R}_+$

Theorem

If $G_1, G_2, \ldots$ is a sampling convergent sequence then there is some graphex $\mathscr{W}$ such that the limiting distribution of the $r/\sqrt{2e_j}$ is $\mathrm{uKEG}(\mathscr{W}, r)$

Theorem

Convergence to W in the BCCH16 sense implies sampling convergence to W

Sampling Convergence

Definition

A sequence of graphs $G_1, G_2, \ldots$ is sampling convergent if the $r/\sqrt{2e_j}$-samplings of G_j converge in distribution as $j \to \infty$, for all $r \in \mathbb{R}_+$

Theorem

If $G_1, G_2, \ldots$ is a sampling convergent sequence then there is some graphex $\mathscr{W}$ such that the limiting distribution of the $r/\sqrt{2e_j}$ is $\mathrm{uKEG}(\mathscr{W}, r)$

Theorem

Convergence to W in the BCCH16 sense implies sampling convergence to W

Sampling Convergence

Definition

A sequence of graphs $G_1, G_2, \dots$ is sampling convergent if the $r/\sqrt{2e_j}$-samplings of G_j converge in distribution as $j \to \infty$, for all $r \in \mathbb{R}_+$

Theorem

If $G_1, G_2, \dots$ is a sampling convergent sequence then there is some graphex $\mathscr{W}$ such that the limiting distribution of the $r/\sqrt{2e_j}$ is $\mathrm{uKEG}(\mathscr{W}, r)$

Theorem

Convergence to W in the BCCH16 sense implies sampling convergence to W

Exchangeability and Graph Limits

- Map graphs $G_1, G_2, \ldots$ to point processes $\mathrm{Lbl}(G_1), \mathrm{Lbl}(G_2), \ldots$ by independently labeling each vertex of G_j by a $\mathrm{Uni}[0, \sqrt{2e_j}]$ random variable
    - These are obviously (finitely) exchangeable
- Notice that each vertex has a label in $[0, r]$ independently with probability $r/\sqrt{2e_j}$
    - Sampling convergence implies that $\mathrm{Lbl}(G_j)(\cdot \cap [0, r]^2)$ converges in distribution for all r
    - This is enough for convergence in distribution of $\mathrm{Lbl}(G_j)$
- Limiting point process is exchangeable, with distribution characterized by same graphex $\mathscr{W}$ that is the sampling convergent limit

Graphexes from p-sampling

- We can get graphex models from graph limits, without any direct appeal to point processes or the associated exchangeability
- Can we get these models directly from p-sampling, without an appeal to graph limits?

Graphexes from p-sampling

Definition

Call $(G_s)_{s\in\mathbb{R}_+}$ an unlabeled random graph process indexed by $\mathbb{R}_+$ if, for all s, G_s is a finite unlabeled graph, and, for all $s \leq t$, it holds that $G_s \subseteq G_t$ in the sense that there is some subgraph of G_t that is isomorphic to G_s.

Theorem

Let $(G_s)_{s\in\mathbb{R}_+}$ be an unlabeled random graph process such that $e_s \uparrow \infty$ a.s. as $s \to \infty$. For each $s \in \mathbb{R}_+$ and $p \in (0,1)$, let $\mathrm{Smpl}_p(G_s)$ be a p-sampling of G_s. If for all $s \in \mathbb{R}_+$ and $p \in (0,1)$ it holds that

$$\mathrm{Smpl}_p(G_s) \overset{d}{=} G_{ps},$$

then there is some (possibly random, possibly non-integrable) almost surely non-zero graphex $\mathscr{W}$ such that, for all $s \in \mathbb{R}_+$,
$G_s \mid \mathscr{W} \sim \mathrm{uKEG}(\mathscr{W}, s)$.

Summary

VR15

- Representation theorem for sparse random graphs
- Extends dense (graphon) theory to sparse graphs
- Formulas for sampling distribution properties in terms of graphex

VR16

- *p*-sampling
- Estimation via the (dilated) empirical graphon

BCCV16+

- Sampling convergence (graph limits)
- Graphex models from *p*-sampling
- Point process construction can be purged from definitions