Limits of discrete structures and their applications

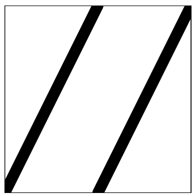
Dan Král'

THE UNIVERSITY OF WARWICK

September 26, 2014 Bedlewo, Poland

Why limits?

- asymptotic properties of large (discrete) objects
  we implicitly use limits in our considerations anyway
- How does the seq. $1, 3, \ldots, 2n-1, 2, 4, \ldots, 2n$ look like?
  How does the adjacency matrix of $K_{n,n}$ look like?
  How does the adjacency matrix of $K_{n,n+1}$ look like?
- convergence of a sequence of discrete objects vs.
  formal analytic representation of its limit



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

Overview of the talk

- Limits of dense structures

  Survey of main results in the area

  Limits of permutations and dense graphs

- The flag algebra method

  Applications in extremal combinatorics

- Limits of sparse structures

  Various concepts, less understood

Permutations

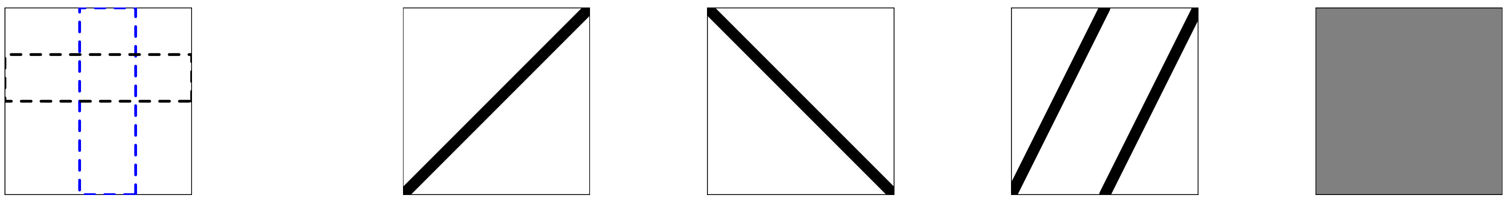
- permutation of order n: order on numbers $1, \ldots, n$
  subpermutation: $4\underline{53}21\underline{6} \quad \longrightarrow \quad 213$
- density of a permutation π in a permutation Π:

$$d(\pi, \Pi) = \frac{\#\ \text{subpermutations of } \Pi \text{ that are } \pi}{\#\ \text{all subpermutations of order } \pi}$$

- $(\Pi_j)_{j \in \mathbb{N}}$ convergent if $\exists \lim\limits_{j \to \infty} d(\pi, \Pi_j)$ for every π

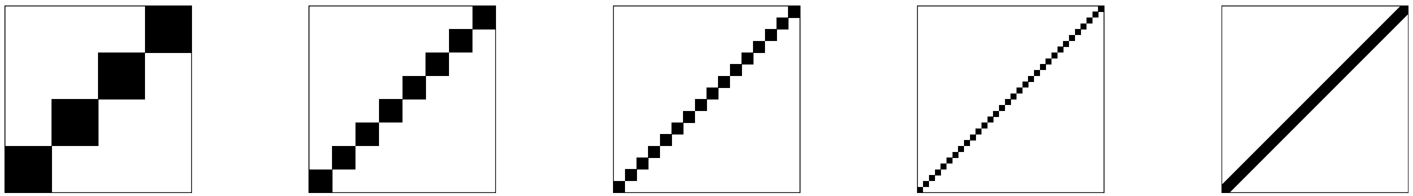
Representation of a limit

- probability measure μ on $[0,1]^2$ with unit marginals
  $\mu([a,b]\times[0,1]) = \mu([0,1]\times[a,b]) = b-a$
  Hoppen, Kohayakawa, Moreira, Ráth and Sampaio
  similar ideas in work of Presutti and Stromquist
- μ-random permutation
  choose n random points, x- and y-coordinates



Existence and uniqueness

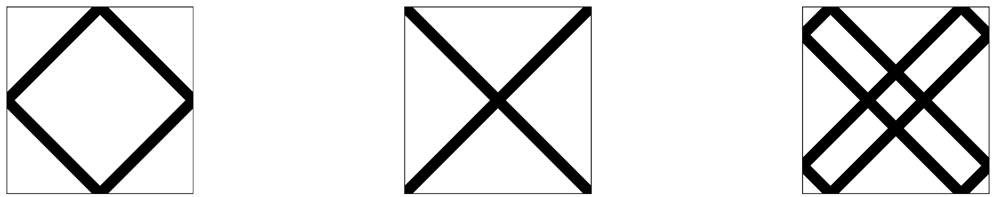
- existence: associate Π_j with a measure μ_j
  observe $|d(\pi, \Pi_j) - d(\pi, \mu_j)| \leq O(|\pi|^2 \cdot |\Pi_j|^{-1})$
  set $\mu(A) := \lim_{j \to \infty} \mu_j(A)$ for $A \subseteq [0,1]^2$



- uniqueness: $\mu(A) = \lim_{n \to \infty} \mathbb{E} \frac{|\{m, [\frac{m}{2^n}, \frac{a_m}{2^n}] \in A\}|}{2^n}$
  where $a_1, \ldots, a_{2^n}$ is μ-random and $A \subseteq [0,1]^2$
  $\forall\, \pi\ d(\pi, \mu) = d(\pi, \mu') \implies \mu = \mu'$

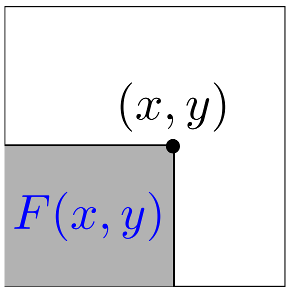
Application: quasirandomness

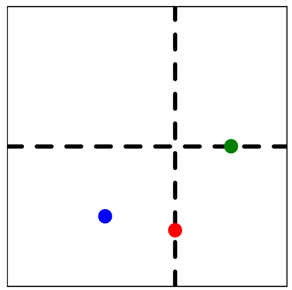
- property $P(k)$ of a sequence $(\Pi_j)_{j\in\mathbb{N}}$:
  $d(\pi, \Pi_j) \to 1/k!$ for every $\pi \in S_k$
- Question (Graham): Is there k_0 such $\forall k\ P(k_0) \Rightarrow P(k)$?
- Theorem (K., Pikhurko): yes, $k_0 = 4$; best possible
  density of 123 is 1/8 in the left and 1/4 in the middle
  a μ-random permutation based on the right measure



PROOF $P(5) \Rightarrow$ QUASIRANDOMNESS

- relate integrals and permutation densities
  $\int x \, \mathrm{d}x\mathrm{d}y = \frac{d(12,\mu)+d(21,\mu)}{2} = \frac{1}{2} = \int x \, \mathrm{d}x$
  prob. $x_2 \le x_1$ for two random points $[x_1, y_1]$ and $[x_2, y_2]$



$\int F(x,y) =$  $=$

123 123 132
132 213 213
231 312 321

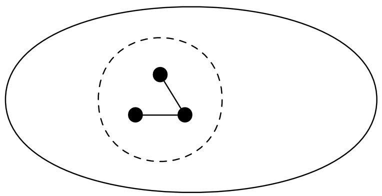
$$\int F(x,y) \, \mathrm{d}x\mathrm{d}y = \frac{d(123)+d(132)+d(213)}{3} + \frac{d(231)+d(312)+d(321)}{6}$$

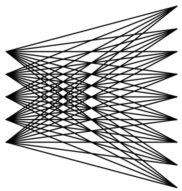
$$\frac{1}{81} = \left(\int F(x,y)xy \, \mathrm{d}x\mathrm{d}y\right)^2 \le \frac{1}{9} \int F(x,y)^2 \mathrm{d}x\mathrm{d}y = \frac{1}{81}$$

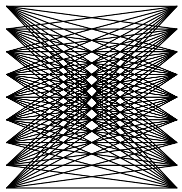
Questions?

Dense graph convergence

- Borgs, Chayes, Lovász, Sós, Szegedy and Vesztergombi
- convergence for dense graphs ($|E| = \Omega(|V|^2)$)
- $d(H, G)$ = probability $|H|$-vertex subgraph of G is H
- a sequence $(G_n)_{n\in\mathbb{N}}$ of graphs is L-convergent
  if $d(H, G_n)$ converges for every H
- extendable to other discrete structures

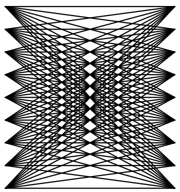




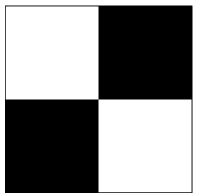


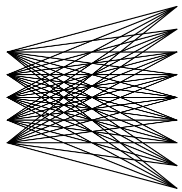
Limit object: graphon

- graphon $W : [0,1]^2 \to [0,1]$, s.t. $W(x,y) = W(y,x)$
- W-random graph of order n
  random points $x_i \in [0,1]$, edge probability $W(x_i, x_j)$
- $d(H,W)$ = prob. $|H|$-vertex W-random graph is H
- W is a limit of $(G_n)_{n\in\mathbb{N}}$ if $d(H,W) = \lim_{n\to\infty} d(H,G_n)$

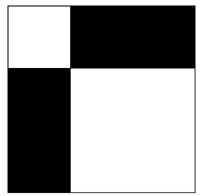


0	0	0	0	0	0	1	1	1	1	1	1
0	0	0	0	0	0	1	1	1	1	1	1
0	0	0	0	0	0	1	1	1	1	1	1
0	0	0	0	0	0	1	1	1	1	1	1
0	0	0	0	0	0	1	1	1	1	1	1
0	0	0	0	0	0	1	1	1	1	1	1
1	1	1	1	1	1	0	0	0	0	0	0
1	1	1	1	1	1	0	0	0	0	0	0
1	1	1	1	1	1	0	0	0	0	0	0
1	1	1	1	1	1	0	0	0	0	0	0
1	1	1	1	1	1	0	0	0	0	0	0
1	1	1	1	1	1	0	0	0	0	0	0





0	0	0	0	1	1	1	1	1	1	1	1
0	0	0	0	1	1	1	1	1	1	1	1
0	0	0	0	1	1	1	1	1	1	1	1
0	0	0	0	1	1	1	1	1	1	1	1
1	1	1	1	0	0	0	0	0	0	0	0
1	1	1	1	0	0	0	0	0	0	0	0
1	1	1	1	0	0	0	0	0	0	0	0
1	1	1	1	0	0	0	0	0	0	0	0
1	1	1	1	0	0	0	0	0	0	0	0
1	1	1	1	0	0	0	0	0	0	0	0
1	1	1	1	0	0	0	0	0	0	0	0
1	1	1	1	0	0	0	0	0	0	0	0



Limit object: graphon

- graphon $W : [0,1]^2 \to [0,1]$, s.t. $W(x,y) = W(y,x)$
- W-random graph of order n
  random points $x_i \in [0,1]$, edge probability $W(x_i, x_j)$
- $d(H,W)$ = prob. $|H|$-vertex W-random graph is H
- W is a limit of $(G_n)_{n\in\mathbb{N}}$ if $d(H,W) = \lim_{n\to\infty} d(H,G_n)$
- every L-convergent sequence of graphs has a limit
- W-random graphs converge to W with probability one

W-RANDOM GRAPHS

- the density of a graph H in a graphon W:

$$\frac{|H|!}{|\mathrm{Aut}(H)|}\int_{[0,1]^{|H|}}\prod_{v_iv_j}W(x_i,x_j)\prod_{\overline{v_iv_j}}(1-W(x_i,x_j))\,\mathrm{d}x_1\cdots x_n$$

- a sequence $(G_n)_{n\in\mathbb{N}}$ of W-random graphs, $|G_n|=n$

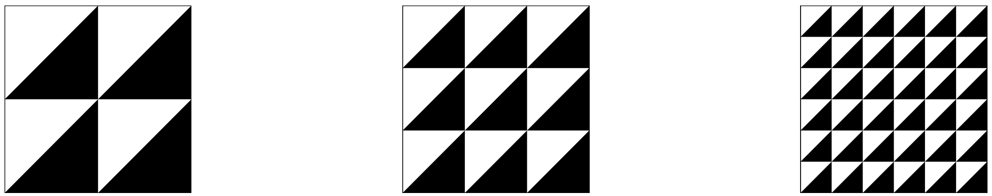
  the expectation of $d(H,G_n)$ conditioned on $x_1,\ldots,x_i$
  Azuma's ineq.: $\mathbb{P}[|d(H,G_n)-d(H,W)|\geq\varepsilon]\leq e^{-O(\varepsilon^2\cdot n)}$
  Borel-Cantelli $\Rightarrow$ $(G_n)_{n\in\mathbb{N}}$ converges with prob. one

Construction of the limit

- sequence of mutually refining regularity partitions
  removal lemma ⇒ subgraph counts
- interpret the partitions as functions $[0,1]^2 \to [0,1]$
  the pointwise limit is the sought graphon
  existence by martingale convergence

Uniqueness of the limit

- $W^\varphi(x,y) := W(\varphi(x),\varphi(y))$ for $\varphi : [0,1] \to [0,1]$
- $d(H,W) = d(H,W^\varphi)$ if φ is measure preserving
- Theorem (Borgs, Chayes, Lovász)
  If $d(H,W_1) = d(H,W_2)$ for all graphs H,
  then there exist measure preserving maps φ_1 and φ_2
  such that $W_1^{\varphi_1} = W_2^{\varphi_2}$ almost everywhere.



Space of typical vertices

- vertices of a graphon W: its rows $f_x(y) := W(x,y)$
  $R(W) = \{f \in L_1[0,1] \mid \exists x \; f = f_x\}$
  $R(W)$ is naturally equipped with measure μ

- typical vertices $T(W)$ = the support of μ in $L_1[0,1]$
  topology can be given by L_1 or by similarity distance

$$d_W(f,f') = \mathbb{E}_{g_x} \left| \int_y (f(y) - f'(y)) g_x(y) \right|$$

Weak regularity partitions

- Frieze and Kannan (1999)
  an ε-regular partition of a graph G is $V_1, \dots, V_k$

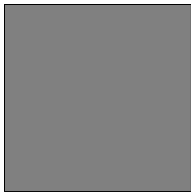
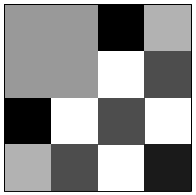
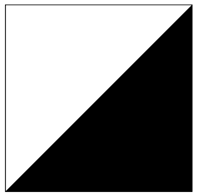
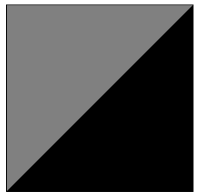
$$\left| e(S,T) - \sum_{i,j} d_{ij} \cdot |S \cap V_i| \cdot |T \cap V_j| \right| \le \varepsilon$$

- $\exists$ ε-regular partitions with k parts $\Leftrightarrow$
  $\exists$ cover of typical vertices with k ε-balls except for ε

- finite dimension $\Rightarrow$ polynomial weak regularity partition

Finitely forcible graphons

- a graphon W is finitely forcible if there exist $H_1, \ldots, H_k$ and $d_1, \ldots, d_k$ such that W is the only graphon with the expected density of H_i equal to d_i
- $\Leftrightarrow$ the only graphon minimizing $\sum \alpha_j d(H'_j, W)$
- Lovász and Sós (2008),
  Every step function is finitely forcible.

Finitely forcible graphons

- Conjecture (Lovász and Szegedy):
  Every extremal problem $\min \sum \alpha_j d(H_j, W)$
  has a finitely forcible optimal solution.

- extremal graph theory problem $\rightarrow$
  finitely forcible optimal solution $\rightarrow$
  simple structure gives new bounds on old problems
  (bounded dimension, few kinds of typical vertices)

Finitely forcible graphons

- Conjectures (Lovász and Szegedy):
  The space $T(W)$ of a finitely forcible W is compact.
  The space $T(W)$ has finite dimension.
- Theorem (Glebov, K., Volec):
  $T(W)$ can fail to be locally compact
- Theorem (Glebov, Klimošová, K.):
  $T(W)$ can have a part homeomorphic to $[0,1]^\infty$
- Theorem (Cooper, Kaiser, K., Noel):
  $\exists$ finitely forcible W such that every ε-regular partition has at least $2^{\varepsilon^{-2}/\log\log\varepsilon^{-1}}$ parts.

Parameter testing

- graph parameter $\mathcal{P}$: graphs $\to \mathbb{R}$
- large input data, not possible to process
  providing an estimate based on a small sample
- $\mathcal{P}$ is testable if there exists a randomized algorithm that estimates the parameter $\mathcal{P}$ within the additive error ε based on a sample of size $f(\varepsilon)$ with probability $\geq 1 - \varepsilon$
- $\mathcal{P}$ is testable $\Leftrightarrow$ $\mathcal{P}$ is continuous on the graphon space

Questions?

Overview of the talk

- Limits of dense structures

  Survey of main results in the area

  Limits of permutations and dense graphs

- The flag algebra method

  Applications in extremal combinatorics

- Limits of sparse structures

  Various concepts, less understood

Flag algebras: homomorphisms

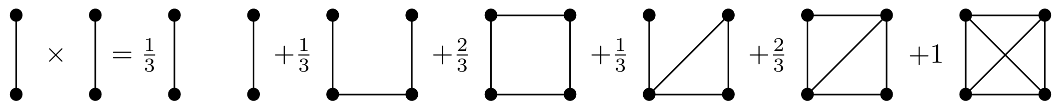
- algebra $\mathcal{A}$ on formal linear combinations of graphs
  addition and multiplication by a scalar

- if W is a graphon, define $f(\sum \alpha_i H_i) := \sum \alpha_i d(H_i, W)$
  the function f is a homomorphism from $\mathcal{A}$ to $\mathbb{R}$
  we define multiplication in $\mathcal{A}$ on the next slide

- $\mathrm{Ker}(f)$ always contains certain elements
  e.g. $K_2 - \frac{1}{3}\overline{K_{1,2}} - \frac{2}{3}K_{1,2} - \frac{3}{3}K_3$
  $f(K_2) = \frac{1}{3}f(\overline{K_{1,2}}) + \frac{2}{3}f(K_{1,2}) + \frac{3}{3}f(K_3)$

Flag algebras: multiplication

- picking two pairs of vertices ≈ picking a quadruple assuming the considered graph G is huge
- $d(K_2, G) \times d(K_2, G) = \frac{1}{3} d(K_2 \cup K_2, G) + \frac{1}{3} d(P_4, G) + \cdots$
- this leads to a definition of a product of two graphs

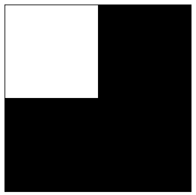


- If W is the unique graphon with $d(H_i, W) = \alpha_i$, it is the unique graphon with $\sum_i (d(H_i, W) - \alpha_i)^2 = 0$.

$$\sum_i (H_i - \alpha_i K_1)^2 = \sum_j \beta_j H'_j$$

Rooted homomorphisms

- consider a graph G with a distinguish vertex (root)
  a random sample always includes the root
- algebra $\mathcal{A}^\bullet$ on combinations of rooted graphs
- rooted graph $\to$ a homomorphism from $\mathcal{A}^\bullet$ to $\mathbb{R}$
  random choice of the root $\to$ probability distribution
  on homomorphisms $f^\bullet$ from $\mathcal{A}^\bullet$ to $\mathbb{R}$



$$f^\bullet(K_2^\bullet) = 1/2,\ f^\bullet(\overline{K_2^\bullet}) = 1/2,\ f^\bullet(K_3^\bullet) = 0,\ \ldots$$

$$f^\bullet(K_2^\bullet) = 1,\ f^\bullet(\overline{K_2^\bullet}) = 0,\ f^\bullet(K_3^\bullet) = 3/4,\ \ldots$$

Averaging operator

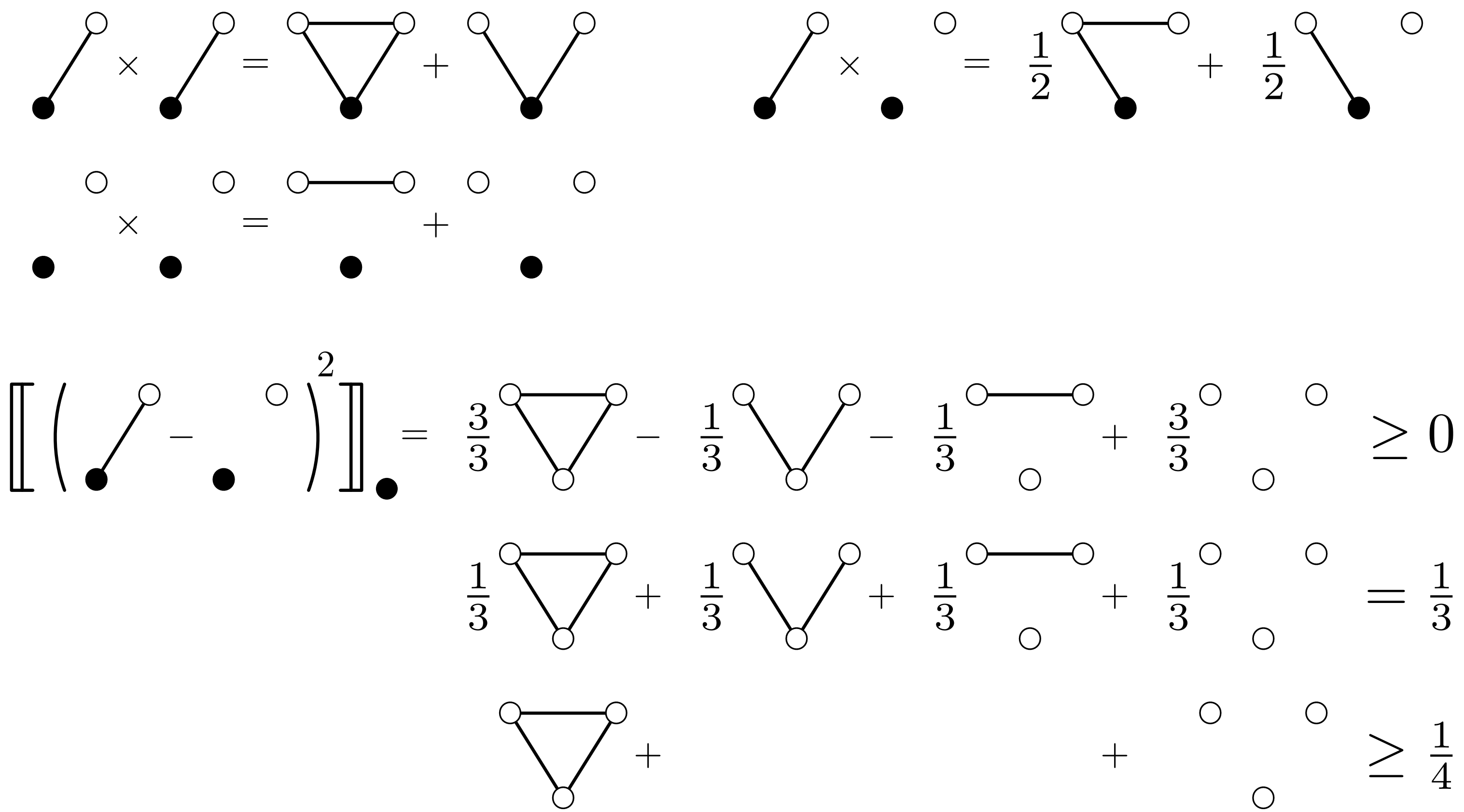
- goal: $f(\llbracket x \rrbracket_\bullet) = \mathbb{E}_\bullet f^\bullet(x)$ for $x \in \mathcal{A}^\bullet$

- expressing degree constraints:
  $f(\llbracket (K_2^\bullet - 1/3)^2 (K_2^\bullet - 2/3)^2 \rrbracket_\bullet) = 0$

- if A is PSD $n \times n$-matrix and $x \in \mathcal{A}^n$, $f(x^T A x) \geq 0$
  example: $f((\alpha K_2 - \beta \overline{K_2})^2) = f(\sum \alpha_i H_i) \geq 0$
  analogously for vectors x with entries from $\mathcal{A}^\bullet$

- search for such inequalities can be computer assisted

Flag algebras: example



Quasirandomness of graphs

- Theorem (Thomason):

  $d_{\text{hom}}(H,G) = d(H,G_{n,p})$ for $H = K_2, C_4 \Rightarrow G \approx G_{n,p}$

- $d(H,W) = d(H,W_p)$ for all 4-vertex $H \Rightarrow W = W_p$

  $$\forall\, x: \ \int W(x,y)\,\mathrm{d}y = p \Longleftrightarrow [\![(K_2^{\bullet} - p)^2]\!]_{\bullet} = 0$$

  $$\forall\, x,x': \ \int W(x,y)W(x',y)\,\mathrm{d}y = p^2$$
  $$\Longleftrightarrow [\![(K_3^{\bullet\bullet} + K_{1,2}^{\bullet\bullet} - p^2)^2]\!]_{\bullet\bullet} = 0$$
  $$\Longrightarrow \forall\, x: \ \int W(x,y)^2\,\mathrm{d}y = p^2$$

  $$\forall\, x: \ \int W(x,y)\,\mathrm{d}y = p \ \wedge \ \int W(x,y)^2\,\mathrm{d}y = p^2$$
  $$\Longrightarrow \forall\, x,y: \ W(x,y) = p$$

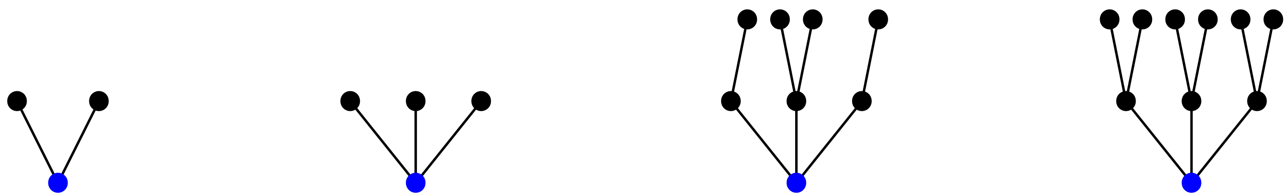
Questions?

Sparse graph convergence

- convergence of graphs with bounded degree
  trivially converging to the zero graphon
- need of a different notion of convergence
  several notions, each having some cons
- absence of understood analytic representation
  Aldous and Lyons Conjecture, relation to group theory
  Does every graphing has a sequence converging to it?

Left convergence

- introduced by Benjamini and Schramm in 2001
- bounded number of types of d-neighborhoods
  convergence of statistic of d-neighborhoods
- cons: connected vs. disconnected
  bipartite vs. non-bipartite graphs



Local-global convergence

- introduced by Hatami, Lovász and Szegedy in 2012
- types of d-neighborhoods k-vertex-colored graphs
  convergence of statistic of such d-neighborhoods
  attainable by a k-vertex-coloring of graphs
- $(G_i)_{i\in\mathbb{N}} \to (A_i)_{i\in\mathbb{N}}$ where $A_i \subseteq \mathbb{R}^K$ and K is # of types
  $\forall\varepsilon > 0\ \exists\, n\ \forall\, i, j > n, x \in A_i\ \exists y \in A_j\ ||x - y|| \leq \varepsilon$
- almost bipartite vs. non-bipartite graphs
  local-global convergence $\Rightarrow$ left convergence

Graphings

- graphing G is a graph with $V(G) = [0,1]$
  bounded maximum degree, Borel measurable edge-set
  mass preservation: $\int_A \deg_B(x)\,\mathrm{d}x = \int_B \deg_A(y)\,\mathrm{d}y$
- Theorem (Elek, 2007)
  Every BS-convergent sequence has a graphing.
  Theorem (Hatami, Lovász, Szegedy, 2012)
  Every LG-convergent sequence has a graphing.
- Conjecture (Aldous, Lyons)
  Every graphing is a BS-limit of a graph sequence.

Right convergence

- developed by Borgs, Chayes, Kahn and Lovász in 2013

$$\lim_{i\to\infty} \frac{\log \hom(G_i, H)}{|G_i|}$$

$$\hom(G,H) = \sum_{f:G\to H} \prod_{v\in V} w(f(v)) \prod_{vv'\in E} w(f(v)f(v'))$$

- cons: connected vs. disconnected
$\frac{\log \hom(G,H)}{|G|} = \frac{\log \hom(G\cup G,H)}{|G\cup G|}$

Right $\Rightarrow$ left convergence

- complex proof by Borgs, Chayes, Kahn and Lovász
- use LLL to count $\hom(G, H_k)$ for $H_k = K_k \setminus K_2$

  if $p \le x_i \prod_{j \sim i} (1 - x_j)$, then $\ge \prod_i (1 - x_i)$

  $p = \frac{2}{k^2} \le (\frac{2}{k^2} + \frac{c}{k^4})(1 - \frac{2}{k^2} - \frac{c}{k^4})^{2\Delta}$

  $\hom(G, H_k) \ge k^{|G|} \cdot (1 - \frac{2}{k^2} - \frac{c}{k^4})^{|E(G)|}$
- upper bound given by Suen's inequality

  $\hom(G, H_k) \le k^{|G|} \cdot e^{\frac{2|E(G)|\Delta}{k^3}} (1 - \frac{2}{k^2})^{|E(G)|}$
- $\frac{\log \hom(G, H_k)}{|G|} = k - \frac{|E(G)|}{|G| \cdot k^2} + O\left(\frac{\Delta^2}{k^3}\right)$

Partition convergence

- introduced by Bollobás and Riordan in 2011
- statistic of a k-partition: $(a_1, \ldots, a_k, d_{11}, d_{12}, \ldots, d_{kk})$
  convergence of attainable statistic of k-partitions
- $n \times C_4$ vs. $n \times C_6$: same attainable statistic
  if $k = 2$: $a_1 + a_2 = 1$, $a_1 = d_{11} + d_{12}/2$, $a_2 = d_{22} + d_{12}/2$

Large deviation convergence

- introduced by Borgs, Chayes and Gamarnik in 2013
- counting k-partitions with a statistic $x \in \mathbb{R}^{k+\binom{k+1}{2}}$

$$r(x) = \lim_{\varepsilon \to 0} \lim_{i \to \infty} - \frac{\log \frac{\#\ k\text{-partitions } \varepsilon\text{-close to } x}{k^{|G_i|}}}{|G_i|}$$

possible values: $[0, \log k] \cup \{\infty\}$

number of k-partitions with statistic $x \approx k^{|G|} \cdot e^{-r(x)|G|}$

- $\Rightarrow$ partition convergence (the limit is finite)

  $\Rightarrow$ right convergence (weighted partitions)

Large deviation $\Rightarrow$ right converg.

- $\hom(G,H) = \sum_{f:G\to H} \prod_{v\in V} w(f(v)) \prod_{vv'\in E} w(f(v)f(v'))$
- if $k = |H|$, determined by the statistic $x \in \mathbb{R}^{k+\binom{k+1}{2}}$

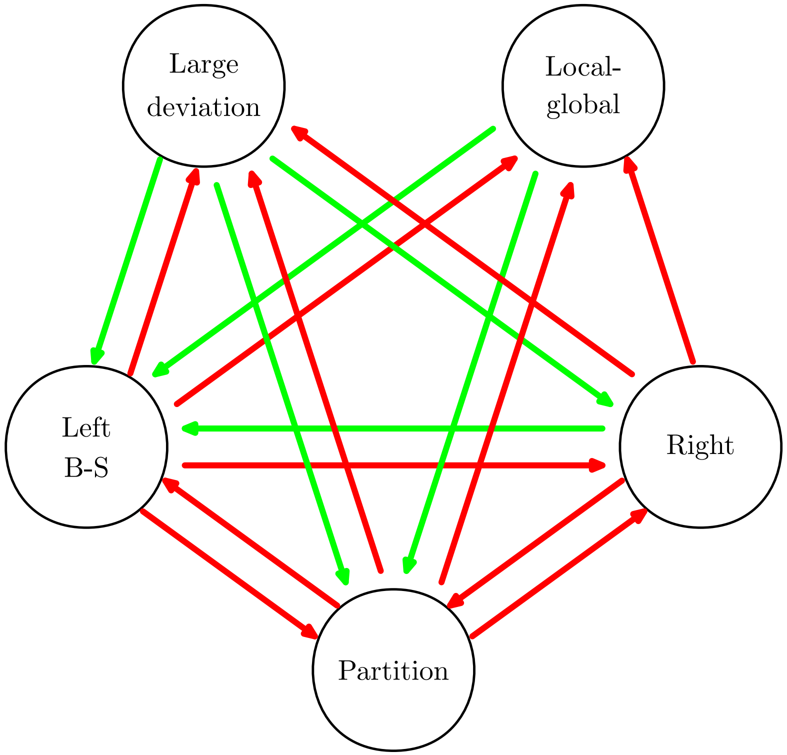
$$\approx \sum_x k^{|G|} e^{-r(x)|G|} \prod_{v_i} w(v_i)^{x_i|G|} \prod_{v_i v_j} w(v_i v_j)^{x_{ij}|G|}$$

- What is $\lim_{i\to\infty} \frac{\log \hom(G_i,H)}{|G_i|}$?

$$\sup_x \; \log k - r(x) + \sum_{v_i} x_i \log w(v_i) + \sum_{v_i v_j} x_i x_j \log w(v_i v_j)$$

convergence of $r(x) \Rightarrow$ right convergence

Mutual relations



Questions?

Exercises

- Compute the density of 123 in the permutons below.
- Describe the limit of $K_{n/3} \cup K_{2n/3}$ and show that it is finitely forcible.
- Show that every BS-converegent sequence of 2-regular graphs has a graphing.