Multiclass generalized quasirandom properties

Marianna Bolla
Institute of Mathematics
Budapest University of Technology and Economics
Partly joint work with **Bojan Mohar**

December 1, 2016, Budapest

Notation

$G = (V, \mathbf{A})$ edge-weighted graph, $|V| = n$,
$\mathbf{A} = (a_{ij})$: weighted adjacency matrix
$a_{ij} = a_{ji} \geq 0$ $(i \neq j)$ and $a_{ii} = 0$ (i=1,...,n).

$d_i := \sum_{j=1}^{n} a_{ij}$ $(i = 1, \dots, n)$ generalized degrees
$\mathbf{d} := (d_1, \dots, d_n)^T$: degree vector, $\sqrt{\mathbf{d}} := (\sqrt{d_1}, \dots, \sqrt{d_n})^T$

$\mathbf{D} := \mathtt{diag}\,(d_1, \dots, d_n)$: degree matrix

w.l.g. $\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} = 1$ will often be assumed

Modularity matrices

$\mathbf{M} = \mathbf{A} - \mathbf{d}\mathbf{d}^T$: modularity matrix
$m_{ij} = a_{ij} - d_i d_j$: discrepancy
$\mathbf{M}$ is usually indefinite and for simple graphs (B, BSM students, Friedl, LAA (2015)): $\mathbf{M}$ is negative semidefinite $\Leftrightarrow G = K_{n_1,\dots,n_k}$.

$\mathbf{M}_D = \mathbf{D}^{-1/2}\mathbf{M}\mathbf{D}^{-1/2} = \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2} - \sqrt{\mathbf{d}}\sqrt{\mathbf{d}}^T$: normalized modularity matrix (B, Phys. Rev. E (2011)), $\text{Spec}\,(\mathbf{M}_D) \in [-1, 1]$.
1 cannot be an eigenvalue if G is connected ($\mathbf{A}$ is irreducible), and 0 is always an eigenvalue with eigenvector $\sqrt{\mathbf{d}}$.

The spectral gap of G: $1 - \|\mathbf{M}_D\|$ (spectral norm).

Eigenvectors and representation

$\mathbf{u}_1, \dots, \mathbf{u}_k \in \mathbb{R}^n$: unit-norm, pairwise orthogonal eigenvectors corresponding to the k largest absolute value eigenvalues of $\mathbf{A}$.

$$(\mathbf{u}_1, \dots, \mathbf{u}_k) = \begin{pmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \vdots \\ \mathbf{r}_n^T \end{pmatrix},$$

where $\mathbf{r}_1, \dots, \mathbf{r}_n \in \mathbb{R}^k$: k-dimensional vertex representatives.
k-variance of them over $(U_1, \dots, U_k) \in \mathcal{P}_k$:

$$S_k^2 = \min_{(U_1,\dots,U_k)\in\mathcal{P}_k} \sum_{i=1}^{k} \sum_{v\in U_i} \|\mathbf{r}_v - \mathbf{c}_i\|^2, \quad \mathbf{c}_i = \frac{1}{|U_i|} \sum_{v\in U_i} \mathbf{r}_v .$$

Minimizer: k-means algorithm. Ostrovsky et. al., J. ACM (2012): if $S_k^2 \le \epsilon^2 S_{k-1}^2$, then there is a PTAS.

Weighted k-variance, subspace distances

$\mathbf{u}_1, \dots, \mathbf{u}_{k-1}$: unit-norm, pairwise orthogonal eigenvectors corresponding to the $k-1$ largest absolute value eigenvalues of $\mathbf{M}_D$.

$$(\mathbf{D}^{-1/2}\mathbf{u}_1, \dots, \mathbf{D}^{-1/2}\mathbf{u}_{k-1}) = \begin{pmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \vdots \\ \mathbf{r}_n^T \end{pmatrix}.$$

$$\tilde{S}_k^2 = \min_{(U_1,\dots,U_k)\in\mathcal{P}_k} \sum_{i=1}^{k} \sum_{v\in U_i} d_v \|\mathbf{r}_v - \mathbf{c}_i\|^2$$

weighted k-variance of them, where $\mathbf{c}_i = \frac{1}{\mathtt{Vol}\,(U_i)} \sum_{v\in U_i} d_v \mathbf{r}_v$, $\mathtt{Vol}\,(U_i) = \sum_{v\in U_i} d_v$. Minimizer: weighted k-means algorithm.
S_k^2 and $\tilde{S}_k^2$: squared distance between the spectral subspace and the one spanned by step-vectors over $\mathcal{P}_k$ (ANOVA fact).

Discrepancy

ACTUAL–EXPECTED connection between $X, Y \subset V$:

$$\sum_{i \in X} \sum_{j \in Y} (a_{ij} - d_i d_j) = a(X, Y) - \mathtt{Vol}\,(X)\mathtt{Vol}\,(Y),$$

Definition

The multiway discrepancy of the edge-weighted graph $G = (V, \mathbf{A})$ in the proper k-partition $U_1, \dots, U_k$ of its vertices is

$$\mathtt{md}(G; U_1, \dots, U_k) = \max_{\substack{1 \le i \le j \le k \\ X \subset U_i,\, Y \subset U_j}} \mathtt{md}(X, Y; U_i, U_j)$$

$$\mathtt{md}(X, Y; U_i, U_j) = \frac{|a(X, Y) - \rho(U_i, U_j)\mathtt{Vol}\,(X)\mathtt{Vol}\,(Y)|}{\sqrt{\mathtt{Vol}\,(X)\mathtt{Vol}\,(Y)}}$$

$$= |\rho(X, Y) - \rho(U_i, U_j)|\sqrt{\mathtt{Vol}\,(X)\mathtt{Vol}\,(Y)}$$

Minimum k-way discrepancy

where $a(X,Y) = \sum_{i \in X} \sum_{j \in Y} a_{ij}$: weighted cut between X and Y. If $a_{ij} = 0/1$, then it is the number of cut-edges $e(X,Y)$, with edges counted twice in $X \cap Y$ if it is not empty;
and $\rho(X,Y) = \frac{a(X,Y)}{\mathtt{Vol}\,(X)\mathtt{Vol}\,(Y)}$: density between X and Y.

Definition

The minimum k-way discrepancy of G is

$$\mathtt{md}_k(G) = \min_{(U_1,\dots,U_k)\in\mathcal{P}_k} \mathtt{md}(G; U_1,\dots,U_k)$$

B, DAM (2016)

Remark

$\mathtt{md}(G; U_1, \dots, U_k)$ is the smallest α such that for every U_i, U_j pair and for every $X \subset U_i$, $Y \subset U_j$:

$$|a(X,Y) - \rho(U_i,U_j)\mathtt{Vol}\,(X)\mathtt{Vol}\,(Y)| \le \alpha\sqrt{\mathtt{Vol}\,(X)\mathtt{Vol}\,(Y)}.$$

In the minimizer $\mathcal{P}_k$, every U_i, U_j pair is α-volume regular, and this is the smallest possible discrepancy that can be attained with proper k-partitions of the vertices of G.
See the volume regular pairs of Alon et al. (2010) and the ε-regular pairs of the Szemerédi regularity lemma (1976) (simple graph, equitable partition, cardinality instead of volume, too "small" X, Y's are excluded).

Generalized random graphs

Definition

We are given a model graph H on k vertices with vertex-weights $r_1, \dots, r_k$ ($r_i > 0$, $\sum_{i=1}^k r_i = 1$) and edge-weights $p_{ij} = p_{ji}$, $1 \le i \le j \le k$ (entries of the $k \times k$ symmetric probability matrix $\mathbf{P}$ of rank k, where $0 \le p_{ij} \le 1$, $1 \le i \le j \le k$). G_n is a the general term of a generalized random graph sequence on the model graph H if

- it has n vertices;
- to each vertex v a membership $c_v \in \{1, \dots, k\}$ is assigned according to the probability distribution $r_1, \dots, r_k$;
- given the memberships, each pair $v \neq u$ is connected with probability $p_{c_v c_u}$;
- further, all these decisions are made independently.

Remark

Let $(U_1, \dots, U_k)$ be the so obtained k-partition (clustering) of the vertices (they also depend on n, however we will not denote this dependence, unless necessary).
The definition implies the following **strong balancing condition** on the growth of the cluster sizes $n_i = |U_i|$, $i = 1, \dots, k$ $(\sum_{i=1}^{k} n_i = n)$: if $n \to \infty$, then $\frac{n_i}{n} \to r_i$ $(i = 1, \dots, k)$.

Lovász, Sós, J. Comb. Theory B (2008)
Abbe, Sandon, FOCS (2015)

In another context:
Holland, Bickel, Coja-Oghlan, Karrer, McSherry, Rohe et al.:
stochastic block-model, planted partition model.

Erdo"s–Rényi random graph: $k = 1$, $G_n(p)$.

Properties of generalized random graphs

Theorem

Let (G_n) be a generalized random graph sequence on the model graph H; G_n has n vertices with vertex-classes $U_1, \dots, U_k$ of sizes $n_1, \dots n_k$. Let H, and so k be kept fixed, i.e., the $k \times k$ probability matrix $\mathbf{P}$ of rank k and the "blow-up" ratios $r_1, \dots, r_k$ are fixed, while $n \to \infty$ under the strong balancing condition. Then the following hold almost surely for the homorphism densities of simple graphs in G_n, for the adjacency matrix $\mathbf{A}_n = (a_{ij}^{(n)})$, the normalized modularity matrix $\mathbf{M}_{D,n}$, the multiway disccrepancies, and the within- and between-cluster codegrees of G_n.

Property 0.

$G_n \to W_H$ as $n \to \infty$ under the strong balancing conditions, where W_H is the step-function graphon corresponding to H, and the convergence is meant in the sense of the convergence of homomorphism densities of any simple graph F into G_n.

B, Kói, Krámli, DAM (2012)
Idea: $\mathbf{A}_n = \mathbf{B}_n + \mathbf{W}_n$, where
$\mathbf{B}_n$: blown-up of $\mathbf{P}$ with sizes $n_1, \dots, n_k$
$\mathbf{W}_n$: Wigner-noise (uniformly bounded, independent entries in and above the main diagonal, of 0 expectation)
$\|W_{\mathbf{W}_n}\|_\square \to 0$ as $n \to \infty$

Property 1.

$\mathbf{A}_n$ has exactly k structural eigenvalues that are $\Theta(n)$ in absolute value, while the remaining eigenvalues are $\mathcal{O}(\sqrt{n})$. Further, the k-variance $S^2_{k,n}$ of the k-dimensional vertex representatives, based on the eigenvectors corresponding to the structural eigenvalues of $\mathbf{A}_n$, is $\mathcal{O}(\frac{1}{n})$.

B, LAA (2005)
Idea: perturbation results for spectra and spectral subspaces
$\mathbf{A}_n = \mathbf{B}_n + \mathbf{W}_n$, where
$\mathbf{B}_n$ has as many non-zero eigenvalues of order n as the rank of $\mathbf{P}$ (in our case, k) with stepwise constant eigenvectors;
for the Wigner-noise $\mathbf{W}_n$: $\|\mathbf{W}_n\| = \mathcal{O}(\sqrt{n})$ almost surely.

Spectral norm of a Wigner-noise

Füredi, Komlós, Combinatorica (1981):

$$\|\mathbf{W}_n\| = \max_{1\le i\le n} |\lambda_i(\mathbf{W}_n)| \le 2\sigma\sqrt{n} + \mathcal{O}(n^{1/3}\log n)$$

with probability tending to 1 as $n \to \infty$.

Sharp concentration theorem

Theorem

$\mathbf{W}$ *is an* $n \times n$ *real symmetric matrix, its entries in and above the main diagonal are independent random variables with absolute value at most 1.* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$*: eigenvalues of* $\mathbf{W}$*.*
For any $t > 0$*:*

$$\mathbb{P}\left(|\lambda_i - \mathbb{E}(\lambda_i)| > t\right) \leq \exp\left(-\frac{(1-o(1))t^2}{32i^2}\right) \quad \textit{when} \quad i \leq \frac{n}{2},$$

and the same estimate holds for the probability

$$\mathbb{P}\left(|\lambda_{n-i+1} - \mathbb{E}(\lambda_{n-i+1})| > t\right).$$

Alon, Krivelevich, Vu, Israel J. Math. (2002)

Consequence

Lemma

There exist positive constants C_1 and C_2, depending only on the common bound K for the entries of the Wigner-noise $\mathbf{W}_n$, such that

$$\mathbb{P}\left(\|\mathbf{W}_n\| > C_1 \cdot \sqrt{n} \right) \leq \exp(-C_2 \cdot n)$$

with probability tending to 1 *as $n \to \infty$.*

Borel–Cantelli lemma$\Rightarrow$
The spectral norm of $\mathbf{W}_n$ is $\mathcal{O}(\sqrt{n})$ almost surely.

Property 2.

There exists a positive constant $0 < \delta < 1$ independent of n (it only depends on k) such that $\mathbf{M}_{D,n}$ has exactly $k-1$ structural eigenvalues of absolute value greater than δ, while all the other eigenvalues are $\mathcal{O}(n^{-\tau})$ in absolute value, for every $0 < \tau < \frac{1}{2}$. Further, the weighted k-variance $\tilde{S}^2_{k,n}$ of the $(k-1)$-dimensional vertex representatives, based on the transformed eigenvectors corresponding to the structural eigenvalues of $\mathbf{M}_{D,n}$, is $\mathcal{O}(n^{-2\tau})$, for every $0 < \tau < \frac{1}{2}$.

B, DM (2008)
B, Friedl, Krámli, JMVA (2010), for rectangular matrices of nonnegative entries

Property 3.

There is a constant $0 < \theta < 1$ (independent of n) such that $\mathtt{md}_1(G_n) > \theta$, ..., $\mathtt{md}_{k-1}(G_n) > \theta$, and the k-way discrepancy $\mathtt{md}(G_n; U_1, \dots, U_k)$ is $\mathcal{O}(n^{-\tau})$, for every $0 < \tau < \frac{1}{2}$, where $U_1, \dots, U_k$ are the vertex-classes in the definition of G_n.

B, DAM (2016)
estimates between multiway discrepancy and normalized modularity spectra

Property 4.

For every $1 \le i \le j \le k$ and $u \in U_i$:

$$N_1(u; U_j) = \sum_{v \in U_j} a_{uv}^{(n)} = p_{ij} n_j + o(n) \qquad \text{(degrees)}.$$

For every $1 \le i \le j \le k$ and $u, v \in U_i$:

$$N_2(u, v; U_j) = \sum_{t \in U_j} a_{ut}^{(n)} a_{vt}^{(n)} = p_{ij}^2 n_j + o(n) \qquad \text{(codegrees)}.$$

Proof: by Bernstein inequality.

Conclusion: The induced subgraph of G_n, induced by U_i and denoted by $G_{ii,n}$, is the general term of an Erdo"s–Rényi type random graph sequence with edge probability p_{ii}, for every $i = 1, \dots, k$. The induced bipartite subgraph of G_n, induced by the U_i, U_j pair and denoted by $G_{ij,n}$, is the general term of a bipartite random graph sequence with edge probability p_{ij} for $i \neq j$.

Generalized quasirandom graphs

Definition

We have a model graph graph H on k vertices with vertex-weights $r_1, \dots, r_k$ and edge-weights $p_{ij} = p_{ji}$, $1 \le i \le j \le k$, entries of $\mathbf{P}$. Then (G_n) is H-quasirandom if $G_n \to W_H$ as $n \to \infty$.

Lovász, Sós, J. Comb. Theory B (2008)

The authors also proved that the vertex set V of a generalized quasirandom graph G_n can be partitioned into $U_1, \dots, U_k$ in such a way that $\frac{|U_i|}{|V|} \to r_i$, $i = 1, \dots, k$ (strong balancing condition) and the subgraph of $G_{ii,n}$ of G_n induced by U_i is the general term of a quasirandom graph sequence with edge-density tending to p_{ii} $(i = 1, \dots, k)$, whereas the bipartite subgraph $G_{ij,n}$ between U_i and U_j is the general term of a quasirandom bipartite graph sequence with edge-density tending to p_{ij} $(i \neq j)$ as $n \to \infty$.

Revisiting the notion of graph convergence

Lovász, Szegedi, J. Comb. Theory B (2006)
Borgs et al., Ann. Math. (2012)
$G_n \to W_H$ means that for any simple graph F:

$$\frac{\mathrm{hom}(F, G_n)}{|V(G_n)|^{|V(F)|}} \to \mathrm{hom}(F, H) = \sum_{\psi: V(F) \to V(H)} \prod_{i \in V(F)} r_{\psi(i)} \prod_{ij \in E(F)} p_{\psi(i)\psi(j)}.$$

If $|V(F)| = m$, then

$$\mathrm{hom}(F, H) = \mathrm{hom}(F, W_H) = \int_{[0,1]^m} \prod_{\{i,j\} \in E(F)} W(x_i, x_j)\, dx_1 \ldots dx_m.$$

Construction of a generalized quasirandom graph

Given k, $\mathbf{P}$, and vertex-weights of the model graph H: consider the instance when there are k sets $U_1, \dots, U_k \subset V$ of sizes $n_1, \dots, n_k$ such that $\frac{n_i}{n} = r_i$ $(i = 1, \dots, k)$. Let us choose the independent irrational numbers α_{ij} $(1 \le i \le j \le k)$.
Then the subgraph on the vertex-set U_i is constructed as follows:

$$u \sim v \Leftrightarrow \{(u-v)^2 \alpha_{ii}\} \le p_{ii}, \quad i = 1, \dots, k.$$

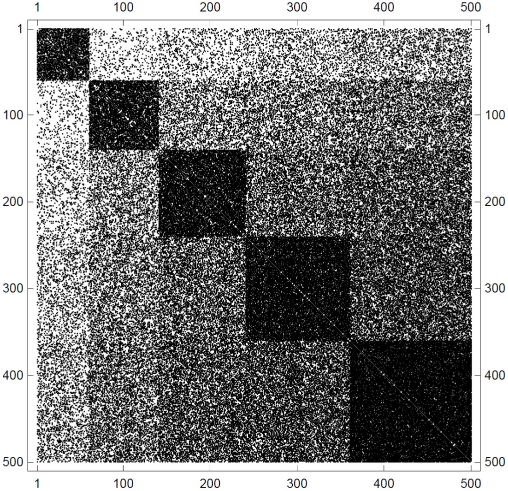
The bipartite subgraph between U_i and U_j: $v \in U_i$ and $u \in U_j$

$$u \sim v \Leftrightarrow \{(u-v)^2 \alpha_{ij}\} \le p_{ij}, \quad 1 \le i < j \le k.$$

Analytical number theoretical considerations guarantee that the above fractional parts are symmetrically well-distributed over $[0,1]^2$ if $n \to \infty$ and $\frac{n_i}{n} \to r_i$ $(i = 1, \dots, k)$. V. T. Sós, Pinch, G. Kiss

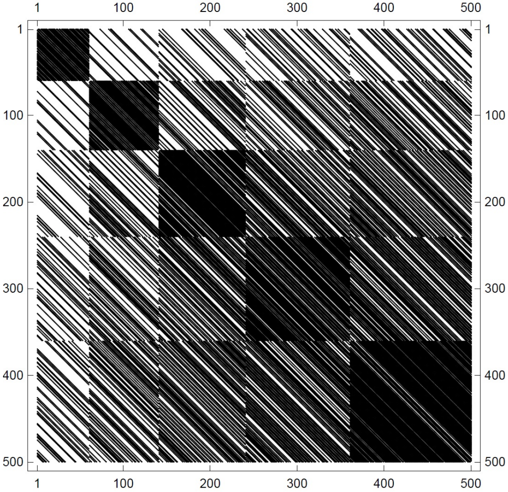
Generalized random graph with $k = 5$

Ev's of $\mathbf{M}_D$: 0.304, 0.214, 0.17, 0.153, -0.097, -0.094, -0.093, ...

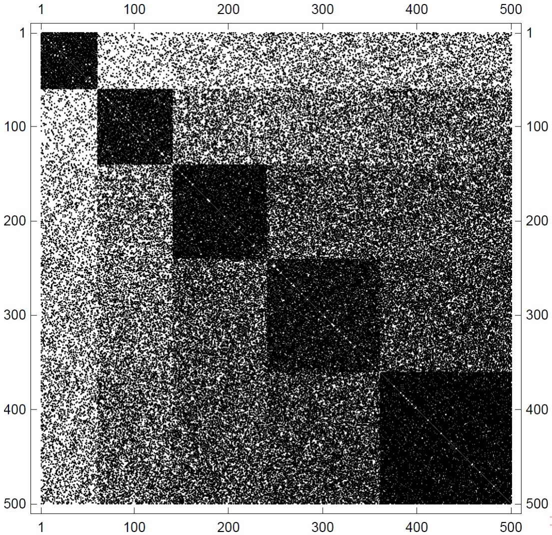


Generalized quasirandom graph with $k = 5$

Ev's of $\mathbf{M}_D$: 0.318, 0.207, 0.154, 0.115, -0.100, -0.099, -0.091, ...



The same with appropriately mixing the vertices



Equivalent properties

$G_n = (V_n, \mathbf{A}_n)$ with $|V_n| = n \to \infty$ is an expanding family of graphs and k be a fixed positive integer.
Then, for them, we consider the following properties, irrespective of stochastic models.

P0. There exists a vertex- and edge-weighted graph H on k vertices, with probability matrix $\mathbf{P}$, $\mathtt{rank}\,\mathbf{P} = k$, such that $G_n \to W_H$ as $n \to \infty$ in terms of the homomorphism densities.

PI. $\mathbf{A}_n$ has k *structural* eigenvalues $\lambda_{1,n}, \dots, \lambda_{k,n}$ such that the normalized eigenvalues converge: $\frac{1}{n}|\lambda_{i,n}| \to q_i$ as $n \to \infty$ $(i = 1, \dots, k)$ with some positive reals $q_1, \dots, q_k$, and the remaining eigenvalues are $o(n)$ in absolute value. The k-variance $S^2_{k,n}$ of the optimal vertex representatives is $o(n)$. For the k-partition $(U_1, \dots, U_k)$ minimizing this k-variance the strong balancing condition holds.

continued

PII. G_n has no dominant vertices, and there exists a constant $0 < \delta < 1$ (independent of n, it only depends on k) such that $\mathbf{M}_{D,n}$ has $k-1$ *structural* eigenvalues that are greater than δ in absolute value, while the remaining eigenvalues are $o(1)$. Further, the weighted k-variance $\tilde{S}^2_{k,n}$ of the optimal vertex representatives is $o(1)$. For the k-partition $(U_1, \dots, U_k)$ minimizing this k-variance the strong balancing condition holds.

PIII. There are vertex-classes $(U_1, \dots, U_k)$ satisfying the strong balancing condition and there exists a constant $0 < \theta < 1$ (independent of n, it only depends on k) such that $\mathtt{md}_1(G_n), \dots, \mathtt{md}_{k-1}(G_n) > \theta$, and $\mathtt{md}_k(G_n; U_1, \dots, U_k) = o(1)$.

continued

PIV. There are vertex-classes $U_1, \dots, U_k$ of cluster sizes $n_1, \dots, n_k$ obeying the strong balancing condition, and there is a $k \times k$ symmetric probability matrix $\mathbf{P} = (p_{ij})$ of rank k such that, with them, the following holds:

$$\sum_{u,v \in U_i} |N_2(u,v;U_j) - p_{ij}^2 n_j| = o(p_{ij}^2 n_i^2 n_j) = o(n^3), \quad \forall i,j = 1, \dots, k,$$

where $N_2(u,v;U_j)$ denotes the number of common neighbors of u, v in U_j.

Then P0 is equivalent to PIV, and they imply PI and PII; further, PII implies PIII.

P0$\Longrightarrow$PIV

We use the results of Chung–Graham–Wilson, Combinatorica (1989); Lovász–T. Sós, JCTB (2008) in view of which:
The vertex set of the generalized quasirandom graph G_n (defined by P0) can be partitioned into classes $U_1, \dots, U_k$ in such a way that $\frac{|U_i|}{n} \to r_i$ $(i = 1, \dots, k)$, that gives the strong balancing; the subgraph $G_{ii,n}$ is the general term of a quasirandom graph sequence with edge-density tending to p_{ii} $(i = 1, \dots, k)$, whereas $G_{ij,n}$ is the general term of a bipartite quasirandom graph sequence with edge-density tending to p_{ij} $(i \neq j)$ as $n \to \infty$. Therefore, for the subgraphs, the equivalent statements of Chung–Graham–Wilson of the usual (1-class) quasirandomness are applicable, and similar considerations can be made for the bipartite subgraphs as well.

Chung–Graham–Wilson: Quasi-random graphs, $k = 1$, $p = \frac{1}{2}$

- $P_1(s)$: for all graphs $M(s)$ on s vertices,

$$N^*_{G_n}(M(s)) = (1+o(1))n^s(\frac{1}{2})^{\binom{s}{2}} \quad \text{labelled induced subgraphs.}$$

$$N^*_{G_n}(M(s)) = (1+o(1))n^s p^{|E(M(s))|}(1-p)^{\binom{s}{2}-|E(M(s))|}.$$

- $P_2(t)$: $e(G_n) \geq (1+o(1))\frac{n^2}{4}$, $N_{G_n}(C_t)) \leq (1+o(1))n^t(\frac{1}{2})^t$.

$$2e(G_n) \geq (1+o(1))pn^2, \quad \mathtt{hom}(C_t, G_n) \leq (1+o(1))n^t p^t.$$

Chung–Graham–Wilson continued

- P_3: $e(G_n) \geq (1+o(1))\frac{n^2}{4}$, $\lambda_1 = (1+o(1))\frac{n}{2}$, $\lambda_2 = o(n)$.

  $$2e(G_n) \geq (1+o(1))pn^2, \quad \lambda_1 = (1+o(1))pn, \quad \lambda_2 = o(n).$$

- $P_4 : \forall S \subset V, e(S) = \frac{1}{4}|S|^2 + o(n^2)$.
  $\forall X \subset V : \quad e(X,X) = p|X|^2 + o(n^2)$.
- P_7: $\sum_{u,v} |N_2(u,v) - \frac{n}{4}| = o(n^3)$,
  $\sum_{u,v} |N_2(u,v) - p^2 n| = o(n^3)$.

Then for $s \geq 4$ and $t \geq 4$ even,

$$P_2(4) \Rightarrow P_2(t) \Rightarrow P_1(s) \Rightarrow P_3 \Rightarrow \cdots \Rightarrow P_7 \Rightarrow P_2(4).$$

Quasirandom graph: satisfies any (all) of the above properties.

Lemma 1.

Lemma

If $(G_{ii,n})$ is quasirandom, then

$$\sum_{u,v\in U_i} N_2(u,v;U_i) \ge (1+o(1))p_{ii}^2 n_i^3, \quad i=1,\dots,k.$$

Proof: We drop the index n of the adjacency entries.

$$\sum_{u,v\in U_i} N_2(u,v;U_i) = \sum_{u,v\in U_i}\sum_{t\in U_i} a_{ut}a_{vt}$$

$$= \sum_{t\in U_i}\sum_{u\in U_i} a_{ut}\sum_{v\in U_i} a_{vt} = \sum_{t\in U_i}[N_1(t;U_i)]^2 \ge \frac{1}{n_i}[\sum_{t\in U_i} N_1(t;U_i)]^2$$

$$= \frac{1}{n_i}[2e(U_i)]^2 \ge \frac{1}{n_i}[(1+o(1))p_{ii}n_i^2]^2 = (1+o(1))p_{ii}^2 n_i^3,$$

Proof of Lemma 1, continued

where $N_1(t;U_i)$ denotes the number of neighbors of t in U_i, and $e(U_i)$ is the number of edges within the induced subgraph $G_{ii,n}$ of G_n, induced by U_i. In the first inequality we used the Cauchy–Schwarz, and in the second one, the first part of the equivalent quasirandom property P_2 of Chung–Graham–Wilson.

Lemma 2.

Lemma

If $(G_{ij,n})$ is bipartite quasirandom, then

$$\sum_{u,v\in U_i} N_2(u,v;U_j) = (1+o(1))p_{ij}^2 n_i^2 n_j, \quad i\neq j.$$

Proof:

$$\sum_{u,v\in U_i} N_2(u,v;U_j) = \sum_{u,v\in U_i}\sum_{t\in U_j} a_{ut}a_{vt}$$

$$= \sum_{t\in U_j}\sum_{u\in U_i} a_{ut}\sum_{v\in U_i} a_{vt} = \sum_{t\in U_j}[N_1(t;U_i)]^2 \geq \frac{1}{n_j}[\sum_{t\in U_j} N_1(t;U_i)]^2$$

$$= \frac{1}{n_j}[e(U_i,U_j)]^2 \geq \frac{1}{n_j}[(1+o(1))p_{ij}n_in_j]^2 = (1+o(1))p_{ij}^2n_i^2n_j,$$

Proof of Lemma 2, continued

where $e(U_i, U_j)$ is the number of cut-edges between U_i and U_j, i.e., the number of edges in the induced bipartite subgraph $G_{ij,n}$ of G_n, induced by the U_i, U_j pair. Here, in the first inequality we used the Cauchy–Schwarz, and in the second one, the equivalent quasirandom property of bipartite quasirandom graphs.

P0$\Longrightarrow$PIV

In view of the lemmas and the the Cauchy–Schwarz inequality:

$$\left[\sum_{u,v\in U_i} |N_2(u,v;U_j) - p_{ij}^2 n_j|\right]^2 \le n_i^2 \sum_{u,v\in U_i} |N_2(u,v;U_j) - p_{ij}^2 n_j|^2$$
$$= n_i^2 \left\{ \sum_{u,v\in U_i} [N_2(u,v;U_j)]^2 - 2p_{ij}^2 n_j \sum_{u,v\in U_i} N_2(u,v;U_j) + n_i^2 (p_{ij}^2 n_j)^2 \right\}$$
$$\le n_i^2 \left\{ (1+o(1)) p_{ij}^4 n_i^2 n_j^2 - 2(1+o(1)) p_{ij}^4 n_i^2 n_j^2 + p_{ij}^4 n_i^2 n_j^2 \right\}$$
$$= n_i^2 o(1) p_{ij}^4 n_i^2 n_j^2 = o(p_{ij}^4 n_i^4 n_j^2),$$

Proof continued

$\sum_{u,v \in U_i} [N_2(u,v;U_j)]^2 \sim \mathrm{hom}(C_4, G_{ij,n})$

- $i = j$: by $P_2(4)$,

$$\mathrm{hom}(C_4, G_{ii,n}) \leq (1+o(1)) p_{ii}^4 n_i^4 .$$

- $i \neq j$: by Lovász–Sós (bipartite quasirandom graphs),

$$\frac{\mathrm{hom}(C_4, G_{ij,n})}{n_i^2 n_j^2} = (1+o(1)) p_{ij}^4 .$$

Only 4-cycles in the above bipartition have to be considered; these 4-cycles have 2 vertices from U_i and 2 from U_j, and any 2 of the common neighbors of $u, v \in U_i$ in U_j are possible candidates to close a (labelled) 4-cycle with them.

Proof of PIV$\Longrightarrow$P0

By C-G-W $P_7 \Rightarrow P_1(s)$, the subgraphs $G_{ii,n}$ are quasirandom. Likewise, if $i \neq j$, the bipartite subgraphs $G_{ij,n}$ are bipartite quasirandom.
Therefore, G_n is built of quasirandom and bipartite quasirandom blocks, so under the strong balancing condition, they together form a generalized quasirandom graph sequence on k classes and model graph H, the vertex-weights of which are $r_1, \dots, r_k$ of the strong balancing condition, and the edge-weights are entries of the probability matrix $\mathbf{P} = (p_{ij})$.

For P0$\Longrightarrow$PII we use the following

Theorem (B, EJC (2014))

$G_n = (V_n, \mathbf{A}_n) \to W$, G_n connected with edge-weights in [0,1] and the vertex-weights are the generalized degrees. Assume that there are no dominant vertices. $|\mu_{n,1}| \geq |\mu_{n,2}| \geq \cdots \geq |\mu_{n,n}| = 0$ is the spectrum of $\mathbf{M}_{D,n}$.
Let $\mu_i(P_{\mathbb{W}})$ be the i-th largest absolute value eigenvalue of the integral operator $P_{\mathbb{W}} : L^2(\xi') \to L^2(\xi)$ taking conditional expectation with respect to the joint measure $\mathbb{W}$ embodied by the normalized limit graphon W, and ξ, ξ' are identically distributed random variables with the marginal distribution of their symmetric joint distribution $\mathbb{W}$.
Then for every $i \geq 1$: $\mu_{n,i} \to \mu_i(P_{\mathbb{W}})$ as $n \to \infty$.

Theorem (B, EJC (2014))

Assume that there are constants $0 < \varepsilon < \delta \leq 1$ *such that*

$$|\mu_{n,1}| \geq \cdots \geq |\mu_{n,k-1}| \geq \delta > \varepsilon \geq |\mu_{n,k}| \geq \cdots \geq |\mu_{n,n}| = 0.$$

Then the subspace spanned by the transformed eigenvectors $\mathbf{D}_n^{-1/2}\mathbf{u}_{n,1}, \dots, \mathbf{D}_n^{-1/2}\mathbf{u}_{n,k-1}$ *converges to the corresponding* $(k-1)$*-dimensional subspace of* $P_{\mathbb{W}}$. *More exactly, if* $\mathbf{P}_{n,k-1}$ *denotes the projection onto the subspace spanned by the transformed eigenvectors belonging to* $k-1$ *largest absolute value eigenvalues of* $\mathbf{M}_{D,n}$, *and* $\mathbf{P}_{k-1}$ *denotes the projection onto the analogous eigen-subspace of* $P_{\mathbb{W}}$, *then* $\|\mathbf{P}_{n,k-1} - \mathbf{P}_{k-1}\| \to 0$ *as* $n \to \infty$.

Note: $\tilde{S}^2_{k,n}$ is the distance between these subspaces.

For PII$\Longrightarrow$PIII we use the following

Theorem (B, DAM (2016))

Let $G = (V, \mathbf{A})$ be an edge-weighted, undirected graph, $\mathbf{A}$ is irreducible. Then for any integer $1 \le k < \mathtt{rank}\,(\mathbf{A})$,

$$|\mu_k| \le 9\mathtt{md}_k(G)(k + 2 - 9k \ln \mathtt{md}_k(G))$$

holds, provided $0 < \mathtt{md}_k(G) < 1$, where μ_k is the k-th largest absolute value eigenvalue of the normalized modularity matrix $\mathbf{M}_D$ of G.

Converse: $k = 1$

Theorem (Chung-Graham, RSA (2008), expander mixing lemma for irregular graphs)

$$\mathtt{md}_1(G) \le \|\mathbf{M}_D\| = |\mu_1|.$$

Theorem (B, EJC (2014), $k \geq 1$, developed version)

Let $G = (V, \mathbf{A})$ be a graph on n vertices, with degrees $d_1, \dots, d_n$ and degree-matrix $\mathbf{D}$. Assume that G is connected, and there are no dominant vertices: $d_v = \Theta(n)$ except for $o(n)$ vertices. Let the eigenvalues of the normalized modularity matrix $\mathbf{M}_D$ of G, enumerated in decreasing absolute values, be

$$|\mu_1| \geq \dots \geq |\mu_{k-1}| > \varepsilon \geq |\mu_k| \geq \dots \geq |\mu_n| = 0.$$

The partition $(U_1, \dots, U_k)$ of V is defined so that it minimizes the weighted k-variance $s^2 = \tilde{S}_k^2$ of the optimal vertex representatives. Assume that the k-partition $(U_1, \dots, U_k)$ satisfies the strong balancing condition. Then

$$\mathtt{md}(G; U_1, \dots, U_k) = \mathcal{O}(\sqrt{2ks} + \varepsilon).$$

P0$\Longrightarrow$PI

We use Theorem 6.7 of Borgs et al., Ann. Math. (2012), where the authors prove that if the sequence (W_{G_n}) of graphons converges to the limit graphon W, then both ends of the spectra of the integral operators, induced by W_{G_n}'s as kernels (these are the numbers $\frac{1}{n}\lambda_{i,n}$), converge to the ends of the spectrum of the integral operator induced by W as kernel. We apply this argument for the limit graphon W_H of (G_n). The same argument as in P0 $=>$ PII can be applied to the convergence of the spectral subspaces, so the convergence of the k-variances is also obtained.
The steps are proportional to r_i's $\Longrightarrow$ strong balancing.
PI does not necessarily implies P0!

Strengthening of PI

PI+: $\mathbf{A}_n$ has k *structural* eigenvalues $\lambda_{1,n}, \dots, \lambda_{k,n}$ such that the normalized eigenvalues converge: $\frac{1}{n}\lambda_{i,n} \to q_i$ as $n \to \infty$ $(i = 1, \dots, k)$ with some non-zero reals $q_1, \dots, q_k$, and the remaining eigenvalues are $o(\sqrt{n})$. Further, the k-variance $S^2_{k,n}$ of the k-dimensional vertex representatives, based on the eigenvectors corresponding to the structural eigenvalues of $\mathbf{A}_n$, is $o(\frac{1}{n})$. The k-partition $P_{k,n} = (U_{1n}, \dots, U_{kn})$ of the vertices of G_n minimizing this k-variance satisfies: $\frac{|U_{in}|}{n} \to r_i$ with some r_i $(i = 1, \dots, k)$. Also assume that there is a $k \times k$ symmetric probability matrix $\mathbf{P} = (p_{ij})$ of rank k such that

$$d(U_{in}, U_{jn}) := \frac{e(U_{in}, U_{jn})}{|U_{in}||U_{jn}|} \to p_{ij} \quad (1 \le i \le j \le k), \quad n \to \infty. \quad (1)$$

(I.e., the within- and between-cluster edge densities converge to the entries of $\mathbf{P}$.)

PI+$\Longrightarrow$P0

By B, LAA (2005) we are able to find a blown-up matrix $\mathbf{B}_n$ of rank k and an error-matrix $\mathbf{E}_n$ with $\|\mathbf{E}_n\| = o(\sqrt{n})$ such that $\mathbf{A}_n = \mathbf{B}_n + \mathbf{E}_n$ $(n = k, k+1, \dots)$. Say $\mathbf{B}_n$ is the blown-up matrix of the $k \times k$ pattern matrix $\mathbf{P}_n$, the ij entry $p_{ij}^{(n)}$ of which is the common entry of the $U_{in} \times U_{jn}$ block of $\mathbf{B}_n$.
Then using the relation between the cut-norm of a graphon and a matrix, further, between the cut-norm and the spectral norm of a matrix, and the transformation of a graph into graphon, we get that

$$\|W_{\mathbf{E}_n}\|_\Box \le \frac{1}{n^2}\|\mathbf{E}_n\|_\Box \le \frac{1}{n^2} n\|\mathbf{E}_n\| = \frac{1}{n} o(\sqrt{n}) = o(n^{-1/2}),$$

where $\|\mathbf{E}_n\|$ is the spectral-norm, $\|\mathbf{E}_n\|_\Box$ is the matrix cut-norm of $\mathbf{E}_n$, and $W_{\mathbf{E}_n}$ denotes the graphon corresponding to the symmetric matrix $\mathbf{E}_n$ of uniformly bounded entries.

Using the Steiner equality, we get that the squared Frobenius norm of $\mathbf{A}_n - \mathbf{B}_n$, restricted to the ij block, is

$$\|(\mathbf{A}_n - \mathbf{B}_n)_{ij}\|_F^2 = \sum_{u \in U_{in}} \sum_{v \in U_{jn}} (a_{uv}^{(n)} - p_{ij}^{(n)})^2$$
$$= \sum_{u \in U_{in}} \sum_{v \in U_{jn}} (a_{uv}^{(n)} - d(U_{in}, U_{jn}))^2 + |U_{in}||U_{jn}|(d(U_{in}, U_{jn})$$

where the edge-density $d(U_{in}, U_{jn})$ is now viewed as the average of the entries of $\mathbf{A}_n$ in the $U_{in} \times U_{jn}$ block. Then by the inequality between the Frobenius and spectral norms,

$$\|(\mathbf{A}_n - \mathbf{B}_n)_{ij}\|_F^2 \le n\|\mathbf{A}_n - \mathbf{B}_n\|^2 = n\|\mathbf{E}_n\|^2 = no^2(\sqrt{n}).$$

Therefore, for every $1 \le i \le j \le k$ pair: $(d(U_{in}, U_{jn}) - p_{ij}^{(n)})^2 \le \frac{1}{|U_i||U_j|} no^2(\sqrt{n}) = \frac{1}{\frac{|U_{in}|}{n}\frac{|U_{jn}|}{n}} n(\frac{o(\sqrt{n})}{n})^2 = no^2(n^{-1/2})$ as $\frac{|U_{in}|}{n} \to r_i$ when $n \to \infty$ $(i = 1, \dots, k)$.

Eventually, we prove the $G_n \to W_H$ convergence by proving that the cut-distance between the corresponding graphons tends to 0.

H is a model graph with vertex-weights r_i's and edge-weights p_{ij}'s in the PI+ conditions.

Using the triangle inequality, we get

$$\|W_{G_n} - W_H\|_\square \le \|W_{G_n} - W_{\mathbf{B}_n}\|_\square + \|W_{\mathbf{B}_n} - W_{G_n/P_{k,n}}\|_\square + \|W_{G_n/P_{k,n}} - W_H\|_\square$$

were $G_n/P_{k,n}$ is the factor graph of G_n with respect to the k-partition $P_{k,n}$. This is an edge- and vertex-weighted graph on k vertices, with vertex-weights $\frac{|U_{in}|}{n}$ and edge-weights $d(U_{in}, U_{jn})$, $i, j = 1, \dots, k$.

The first term is $\|W_{\mathbf{E}_n}\|_\square = o(n^{-1/2})$. To estimate the second term, observe that because $\mathbf{B}_n$ is the blown-up matrix of $\mathbf{P}_n$ with respect to the k-partition $P_{k,n}$, after conveniently permuting its rows (and columns, accordingly). The gaphon $W_{\mathbf{B}_n}$ is also stepwise constant over the unit square, where the sides are divided into k parts: the interval I_j has lengths $\frac{|U_{jn}|}{n}$ $(j = 1, \dots, k)$, and over $I_i \times I_j$ the stepfunction takes on the value $p_{ij}^{(n)}$. By its nature, the graphon $W_{G_n/P_{k,n}}$ is stepwise constant with the same subdivision of the unit square, and over $I_i \times I_j$ it takes on the value $d(U_{in}, U_{jn})$, $i, j = 1, \dots, k$. But in view of the above, $\|W_{\mathbf{B}_n} - W_{G_n/P_{k,n}}\|_\square = \sqrt{n}o(n^{-1/2}) = o(1)$. The third term is $o(1)$, because of the assumptions $\frac{|U_{in}|}{n} \to r_i$ $(i = 1, \dots, k)$ and $d(U_{in}, U_{jn}) \to p_{ij}$, $i, j = 1, \dots, k$. Therefore, $\|W_{G_n} - W_H\|_\square = o(1)$ and so, $G_n \to H$, which finishes the proof.

Define PII+ and PIII+ as PII and PIII together with

- There is a $k \times k$ symmetric probability matrix $\mathbf{P} = (p_{ij})$ of rank k such that

$$d(U_i, U_j) = p_{ij} + o(1) \quad (1 \leq i \leq j \leq k), \quad n \to \infty$$

- for every $1 \leq i \leq j \leq k$ and $u \in U_i$:

$$N_1(u; U_j) = (1 + o(1)) p_{ij} n_j$$

hold.

Lemma

Under P0, the following holds for except $o(n_i)$ vertices $u \in U_i$, and for every $i = 1, \dots, k$: $N_1(u; U_i) = (1 + o(1))p_{ii}n_i$.
Under P0, the following holds for except $o(n_i)$ vertices $u \in U_i$, and for every $1 \le i < j \le k$: $N_1(u; U_j) = (1 + o(1))p_{ij}n_j$.

The statement follows from the $P_1(s)$ $(\forall s)$ $=>$ P_0' implication of C-G-W and its bipartite analogue.

The subgraphs ara almost-regular, the bipartite subgraphs are almost-biregular: weaker than quasirandomness.

P0 => PII+ => PIII+ => PIV => P0, so they are all equivalent.

Proof of P0 => PII+ => PIII+

By Lovász–Sós, JCTB (2008) and the Lemma, P0 implies the extras of PII and PIII too.

Proof of PIII+ => PIV

We are able to prove that P_4 of C-G-W, and the analogous statement of Theorem 2 of Thomason, DM (1989) hold, whenever PIII+ holds.

Proof in the $i = j$ case

Let $(U_1, \dots, U_k)$ be the k-partition, guaranteed by PIII+, such that $\mathtt{md}_k(G_n; U_1, \dots U_k) = o(1)$. Then by the extra conditions of PIII+, for $X \subset U_i$, $\mathtt{Vol}\,(X) = |X|(1 + o(1)) \sum_{\ell=1}^{k} p_{i\ell} n_\ell$, and so,

$$e(X,X) - p_{ii}|X|^2 = e(X,X) - [d(U_i,U_i) + o(1)]|X|^2$$

$$= e(X,X) - \frac{e(U_i,U_i)}{\frac{\mathtt{Vol}^2(U_i)}{(1+o(1))^2(\sum_{\ell=1}^{k} p_{i\ell} n_\ell)^2}} \frac{\mathtt{Vol}^2(X)}{(1+o(1))^2(\sum_{\ell=1}^{k} p_{i\ell} n_\ell)^2} - o(1)|X|^2$$

$$= [e(X,X) - \rho(U_i,U_i)\mathtt{Vol}^2(X)] - o(1)\rho(U_i,U_i)\mathtt{Vol}^2(X) - o(1)|X|^2$$

$$\le \mathtt{md}_k(G_n; U_1, \dots U_k)\sqrt{\mathtt{Vol}^2(X)} - o(1)e(U_i,U_i)\left(\frac{\mathtt{Vol}\,(X)}{\mathtt{Vol}\,(U_i)}\right)^2 - o(n^2)$$

$$= o(n^2).$$

Then P_4 implies P_7 of Chung–Graham–Wilson, that is our PIV.