Testability of minimum balanced multiway cut densities

Marianna Bolla [a,*], Tamás Kói [a], András Krámli [b]

[a] *Institute of Mathematics, Budapest University of Technology and Economics*
[b] *Bolyai Institute, University of Szeged*

Abstract

Testable weighted graph parameters and equivalent notions of testability are investigated based on [4]. We prove that certain balanced minimum multiway cut densities are testable. Using this fact, quadratic programming techniques are applied to approximate some of these quantities. The problem is related to cluster analysis and statistical physics. Convergence of special noisy graph sequences is also discussed.

Key words: Weighted graphs, Testable graph parameters, Minimum balanced multiway cuts, Quadratic programming, Wigner-noise
1991 MSC: 05C35, 62H30, 68R10

1 Introduction

A typical problem of contemporary cluster analysis is to find relatively small number of homogeneous groups of data that do not differ significantly in size. To make inferences on the separation that can be achieved for a given number of clusters, some types of minimum cut densities are investigated.

In a fairly general setup of [4], the objects to be classified are vertices of a weighted graph whose edges and vertices both have nonnegative, real weights. Edge-weights are similarities between the vertices normalized in such a way that 0 is the minimum and 1 is the maximum similarity, while vertex-weights

* Corresponding author. Research supported in part by the Hungarian National Research Grants OTKA 76481 and OTKA-NKTH 77778.

Email addresses: marib@math.bme.hu (Marianna Bolla), koitomi@math.bme.hu (Tamás Kói), kramli@informatika.ilab.sztaki.hu (András Krámli).

reflect individual values of the vertices. Classical (simple) graphs have vertex-weights all equal to 1 and edge-weights 0 or 1.

For given number n of vertices and for a fixed integer $0 < q \leq n$ we define three types of minimum q-way cut densities, each being the minimum of the weight-sum of between-cluster edges, occasionally adjusted with a factor characterizing within-cluster densities, over all or over balanced q-partitions of the vertices. The limit of these densities is considered as $n \to \infty$. If this limit exists for any convergent graph sequence, we say, that the q-way cut density in question is a testable graph parameter. In fact, the subsequent terms of such a convergent graph sequence (G_n) become more and more similar in their global structure, which fact can be formulated in terms of convergence of the homomorphism densities of injective maps $F \to G_n$ for any simple graph F.

Hence, testable parameters measure statistical properties of a large graph that are indifferent to minor changes in the edge- and vertex-weights. It will be proved that certain balanced q-way cut densities are testable. To this end, notions of testability are extended to weighted graphs, and we prove equivalent statements of testability by means of large deviation results of Lovász and coauthors [4]. Roughly speaking, these propositions state that if a smaller simple graph is selected – by an appropriate randomization – based on a large weighted graph, the testable parameter of the randomized one is very close to that of the whole graph with high probability.

The organization of the paper is as follows. In Section 2, notion of a convergent graph sequence and that of a graphon is introduced based on [4]. In Section 3, equivalent statements of testability are discussed for weighted graphs. In Section 4, testability of different kinds of minimum multiway cut densities is investigated. For non testable ones counterexamples are presented, while for testable ones theorems of [5] based on statistical physics are applied. In Section 5, continuous extensions of testable weighted graph parameters to graphons are constructed that gives rise to a quadratic programming task. In Section 6, special graph sequences (blown up structures burdened with a very general kind of noise) are analyzed utilizing the fact that the cut-norm of a so-called Wigner-noise tends to zero as its size tends to infinity.

2 Preliminaries

Let $G = G_n$ be a weighted graph on the vertex set $V(G) = \{1, \ldots, n\} = [n]$ and edge set $E(G)$. Both the edges and vertices have weights: the edge-weights are pairwise similarities $\beta_{ij} = \beta_{ji} \in [0, 1]$, $i, j \in [n]$, while the vertex-weights $\alpha_i > 0$ ($i \in [n]$) indicate relative significance of the vertices. It is important that the edge-weights are nonnegative (zero means no connection at all), the normalization into the [0,1] interval is for the sake of treating them later as probabilities for random sampling. Let $\mathcal{G}$ denote the set of all such weighted graphs.

The *volume* of $G \in \mathcal{G}$ is defined by $\alpha_G = \sum_{i=1}^n \alpha_i$, while that of the vertex-subset T by $\alpha_T = \sum_{i \in T} \alpha_i$. Further,

$$e_G(S,T) = \sum_{s \in S} \sum_{t \in T} \alpha_s \alpha_t \beta_{st}$$

denotes the *weighted cut* between the (not necessarily disjoint) vertex-subsets S and T.

Lovász and coauthors [4] define the homomorphism density between the simple graph F (on vertex set $V(F) = [k]$) and the above weighted graph G. With the notations

$$\alpha_\Phi = \prod_{i=1}^k \alpha_{\Phi(i)}, \qquad \mathrm{inj}_\Phi(F,G) = \prod_{ij \in E(F)} \beta_{\Phi(i)\Phi(j)}$$

the homomorphism density between F and G is defined by

$$t(F,G) = \frac{1}{(\alpha_G)^k} \sum_{\Phi: V(F) \to V(G)} \alpha_\Phi \cdot \mathrm{inj}_\Phi(F,G). \tag{1}$$

For a simple graph G, $t(F,G)$ is the probability that a random map $V(F) \to V(G)$ is a homomorphism. Similarly, $t_{inj}(F,G)$ and $t_{ind}(F,G)$ are defined in such a way that for a simple G, they are the probabilities that a random injective map $V(F) \to V(G)$ is adjacency preserving and results in an induced subgraph of F in G, respectively. With the notation

$$\mathrm{ind}_\Phi(F,G) = \prod_{ij \in E(F)} \beta_{\Phi(i)\Phi(j)} \prod_{ij \in E(\bar{F})} (1 - \beta_{\Phi(i)\Phi(j)}),$$

let

$$t_{\mathrm{inj}}(F,G) := \frac{1}{k!(\alpha)_k} \sum_{\Phi \text{ inj.}} \alpha_\Phi \cdot \mathrm{inj}_\Phi(F,G)$$

and

$$t_{\mathrm{ind}}(F,G) := \frac{1}{k!(\alpha)_k} \sum_{\Phi \text{ inj.}} \alpha_\Phi \cdot \mathrm{ind}_\Phi(F,G),$$

where $(\alpha)_k$ denotes the kth elementary symmetric polynomial of $\alpha_1, \dots, \alpha_n$. Latter one resembles to the likelihood function of taking a sample – that is a simple graph on k vertices – from the weighted graph G in the following way: k vertices are chosen *with replacement* with respective probabilities α_i / α_G $(i = 1, \dots, n)$. Given the vertex-subset $\{\Phi(1), \dots, \Phi(k)\}$, the edges come into existence conditionally independently, with probabilities of the edge-weights.

Such a random graph is denoted by $\xi(k, G)$. Obviously,

$$\mathbb{P}(\xi(k,G) = F) = \frac{1}{(\alpha_G)^k} \sum_{\Phi : V(F) \to V(G)} \alpha_\Phi \mathrm{ind}_\Phi(F, G),$$

since we may get back F, even if Φ is not injective. As most maps into a large graph are injective, the above probability is very close to $t_{\mathrm{ind}}(F, G)$, and $t(F, G)$ is very close to $t_{\mathrm{inj}}(F, G)$. Further, $t_{\mathrm{ind}}(F, G)$ has a well-defined relation to $t_{inj}(F, G)$ that will be formulated in Section 3. In the sequel only the $k \ll n$ case makes sense, and this is the situation we need: k is kept fixed, while n tends to infinity.

Definition 1 *We say that the weighted graph sequence (G_n) is (left-)convergent, if the sequence $t(F, G_n)$ converges for any simple graph F $(n \to \infty)$.*

As other kinds of convergence are not discussed here, in the sequel the word left will be omitted, and we simply use convergence.

Authors in [4] also construct the limit object that is a symmetric, bounded, measurable function $W : [0, 1] \times [0, 1] \to \mathbb{R}$ and they call it *graphon*. Let $\mathcal{W}$ denote the set of these functions. The interval [0,1] corresponds to the vertices and the values $W(x, y) = W(y, x)$ to the edge-weights. In view of the conditions imposed on the edge-weights, the range is also the [0,1] interval. The set of symmetric, measurable functions $W : [0, 1] \times [0, 1] \to [0, 1]$ is denoted by $\mathcal{W}_{[0,1]}$. The stepfunction graphon $W_G \in \mathcal{W}_{[0,1]}$ is assigned to the weighted graph $G \in \mathcal{G}$ in the following way: the sides of the unit square are divided into intervals $I_1, \ldots, I_n$ of lengths $\alpha_1/\alpha_G, \ldots, \alpha_n/\alpha_G$, and over the rectangle $I_i \times I_j$ the stepfunction takes on the value β_{ij}.

The so-called *cut-distance* between the graphons W and U is

$$\delta_\square(W, U) = \inf_\nu \|W - U^\nu\|_\square \tag{2}$$

where the *cut-norm* of the graphon W is defined by

$$\|W\|_\square = \sup_{S,T \subset [0,1]} \left| \iint_{S \times T} W(x, y) \, dx \, dy \right|,$$

and the infimum in (2) is taken over all measure preserving bijections $\nu : [0, 1] \to [0, 1]$, while U^ν denotes the transformed U after performing the same measure preserving bijection ν on both sides of the unit square. An equivalence relation is defined over the set of graphons: two graphons belong to the same class if they can be transformed into each other by a measure preserving map, i.e., their $\delta_\square$-distance is zero. In the sequel, we consider graphons modulo measure preserving maps, and under graphon we understand the whole equivalence class. By Theorem 5.1 of [9], the classes of $\mathcal{W}_{[0,1]}$ form a compact metric space with the $\delta_\square$ metric.

We will intensively use the following reversible relation between convergent weighted graph sequences and graphons.

Theorem 2 *(Corollary 3.9 of [4]). For any convergent sequence (G_n) of weighted graphs with uniformly bounded edge-weights there exists a graphon such that $\delta_\square(W_{G_n}, W) \to 0$. Conversely, any graphon W can be obtained as the limit of a sequence of weighted graphs with uniformly bounded edge-weights. The limit of a convergent graph sequence is essentially unique: If $G_n \to W$, then also $G_n \to W'$ for precisely those graphons W' for which $\delta_\square(W, W') = 0$.*

Authors of [4] also define the $\delta_\square$-distance of two weighted graphs and that of a graphon and a graph. Without going into details, we just cite the following facts: for the weighted graphs G, G', and for the graphon W

$$\delta_\square(G, G') = \delta_\square(W_G, W_{G'}) \quad \text{and} \quad \delta_\square(W, G) = \delta_\square(W, W_G).$$

They prove (Theorem 2.6) that a sequence of weighted graphs with uniformly bounded edge-weights is convergent if and only if it is a Cauchy sequence in the metric $\delta_\square$.

A simple graph on k vertices can be sampled based on W in the following way: k uniform random numbers, $X_1, \dots, X_k$ are generated on [0,1] independently. Then we connect the vertices corresponding to X_i and X_j with probability $W(X_i, X_j)$. For the so obtained simple graph $\xi(k, W)$ the following large deviation result is proved.

Theorem 3 *(Theorem 4.7 of [4], part (ii)). Let k be a positive integer and $W \in \mathcal{W}_{[0,1]}$ be a graphon. Then with probability at least $1 - e^{-k^2/(2\log_2 k)}$, we have*

$$\delta_\square(W, \xi(k, W)) \le \frac{10}{\sqrt{\log_2 k}}. \tag{3}$$

Fixing k, the inequality (3) holds uniformly for any graphon $W \in \mathcal{W}_{[0,1]}$, especially for W_G. Further, the sampling from W_G is identical to the previously defined sampling with replacement from G, that is $\xi(k, G) = \xi(k, W_G)$. In fact, this argument is relevant in the $k \le |V(G)|$ case.

3 Testable weighted graph parameters

A function $f : G \to \mathbb{R}$ is called a *graph parameter* if it is invariant under isomorphism. In fact, a graph parameter is a statistic evaluated on the graph, and hence, we are interested in weighted graph parameters that are not sensitive to minor changes in the weights of the graph.

The testability results of [4] for simple graphs remain valid if we consider weighted graph sequences (G_n) with *no dominant vertex-weights*, that is

$$\max_i \frac{\alpha_i(G_n)}{\alpha_{G_n}} \to 0, \qquad n \to \infty.$$

Definition 4 *A weighted graph parameter f is testable if for every $\varepsilon > 0$ there is a positive integer k such that if $G \in \mathcal{G}$ satisfies*

$$\max_i \frac{\alpha_i(G)}{\alpha_G} \le \frac{1}{k},$$

then

$$\mathbb{P}(|f(G) - f(\xi(k, G))| > \varepsilon) \le \varepsilon, \tag{4}$$

where $\xi(k, G)$ is a random simple graph on k vertices selected randomly from G with replacement as described in Section 2.

Consequently, such a graph parameter can be consistently estimated based on a fairly large sample. As the randomization depends only on the $\alpha_i(G)/\alpha_G$ ratios, it is not able to distinguish between weighted graphs whose vertex-weights differ only in a constant factor. Thus, a testable weighted graph parameter is invariant under scaling the vertex-weights. Now, we introduce some equivalent statements of the testability, indicating that a testable parameter depends continuously on the whole graph. This is the generalization of Theorem 6.1 of [4] applicable for simple graphs.

Theorem 5 *For the weighted graph parameter f the following are equivalent:*

(a) f is testable.
(b) For every $\varepsilon > 0$ there is a positive integer k such that for every weighted graph $G \in \mathcal{G}$ satisfying the node-condition $\max_i \alpha_i(G)/\alpha_G \le 1/k$,

$$|f(G) - \mathbb{E}(f(\xi(k, G)))| \le \varepsilon.$$

(c) For every convergent weighted graph sequence (G_n) with $\max_i \alpha_i(G_n)/\alpha_{G_n} \to 0$, $f(G_n)$ is also convergent ($n \to \infty$).
(d) f can be extended to graphons such that the graphon functional $\tilde{f}$ is continuous in the cut-norm and $\tilde{f}(W_{G_n}) - f(G_n) \to 0$, whenever $\max_i \alpha_i(G_n)/\alpha_{G_n} \to 0$ ($n \to \infty$).
(e) For every $\varepsilon > 0$ there is an $\varepsilon_0 > 0$ real and an $n_0 > 0$ integer such that if G_1, G_2 are weighted graphs satisfying $\max_i \alpha_i(G_1)/\alpha_{G_1} \le 1/n_0$, $\max_i \alpha_i(G_2)/\alpha_{G_2} \le 1/n_0$, and $\delta_\square(G_1, G_2) < \varepsilon_0$, then $|f(G_1) - f(G_2)| < \varepsilon$.

To prove the theorem we need three lemmas that are partly generalizations of results in [4] stated for simple graphs.

Lemma 6 *If* (G_n) *is a weighted graph sequence with no dominant vertex-weights, then for any simple graph* F

$$|t(F,G_n) - t_{inj}(F,G_n)| \to 0, \qquad n \to \infty.$$

PROOF. The difference between between $t(F,G_n)$ and $t_{inj}(F,G_n)$ is essentially obtained by the summation in (1) over the non injective maps. As the sum is increased if we take the non-zero β_{ij}'s 1, it suffices to prove that

$$\sum_{\Phi \text{ non inj.}} \frac{\alpha_\Phi(G_n)}{\alpha_{G_n}^k} \to 0,$$

where the left hand side is the probability that there are repetitions in the vertices. As $\max_i \alpha_i(G_n)/\alpha_{G_n} < \varepsilon_n$, this probability is less than

$$1 - 1 \cdot (1-\varepsilon_n) \dots (1-(k-1)\varepsilon_n) \le 1 - (1-(k-1)\varepsilon_n)^k$$

that tends to 0, since $\varepsilon_n \to 0$, provided k is fixed.

Lemma 7

$$t_{inj}(F,G) = \sum_{F' \supseteq F} t_{ind}(F',G) \quad \textit{and} \quad t_{ind}(F,G) = \sum_{F' \supseteq F} (-1)^{|E(F') \setminus E(F)|} t_{inj}(F',G)$$

where F' *is a simple super-graph of* F *(on the same vertex-set, and edge-set containing the edge-set of* F*).*

PROOF. To verify the first statement, it suffices to prove that for any injective map $\Phi : V(F) \to V(G)$,

$$\text{inj}_\Phi(F,G) = \sum_{F' \supseteq F} \text{ind}_\Phi(F',G). \tag{5}$$

Suppose that ℓ edges are missing from F to be a complete graph. As $F' \supseteq F$ is a super-graph of F, it can easily be seen that $\text{ind}_\Phi(F',G)$ contains the multiplicative factor $\text{inj}_\Phi(F,G)$. Hence, the right hand side of (5) can be factorized as $\text{inj}_\Phi(F,G) \cdot S_\ell$, where S_ℓ depends on F, G, and Φ, but for the sake of simplicity we omit these ones. We show – by reverse induction on the number of edges – that $S_\ell = 1$. If F is a complete graph on k vertices, then by the definition of inj_Φ and ind_Φ, $S_0 = 1$. If j edges are missing from F to be a complete graph, denote by $\beta_1, \dots, \beta_j$ the weights of their Φ-images. Further, denote by $\sigma = (\sigma_1, \dots, \sigma_j)$ a $\{0,1\}$ sequence of length j and $a_i^{(0)}(x) := x$, $a_i^{(1)}(x) := 1-x$, $i = 1, \dots, j$. With this notation $S_j = \sum_{\sigma \in \{0,1\}^j} \prod_{i=1}^j a_i^{(\sigma_i)}(\beta_i)$. Coupling the sequences which differ only in the first coordinate, S_j reduces to S_{j-1}, etc.

By inclusion-exclusion, the second statement also follows.

Lemma 8 *(Lemma 5.3 of [4]). Let (G_n) be a sequence of weighted graphs with uniformly bounded edge-weights, and no dominant vertex-weights. If $\delta_\square(U, W_{G_n}) \to 0$ for some $U \in \mathcal{W}$, then the graphs in the sequence (G_n) can be relabeled in such a way that the resulting sequence (G'_n) of labeled graphs converges to U in the cut-norm: $\|U - W_{G'_n}\|_\square \to 0$.*

Now, we are able to prove the main theorem (Theorem 5).

PROOF. The idea of the proof is analogous to that of Theorem 6.1 of [4].

First we prove that (a),(b),(c),(e) are equivalent:

$(a) \Rightarrow (b)$: The statement is obvious, as due to the boundedness of f, (4) implies that the difference is small on average.

$(b) \Rightarrow (c)$: Let (G_n) be a convergent sequence of weighted graphs with no dominant vertex-weights. Let $\varepsilon > 0$ be arbitrary, and k is chosen corresponding to ε as in statement (b). If n is large enough, then $|f(G_n) - \mathbb{E}(f(\xi(k, G_n)))| \le \varepsilon$. On the other hand, by the definition of convergence it follows that $t(F, G_n)$ is convergent for all simple graphs F on k vertices. Using Lemmas 6, 7, $t_{\mathrm{ind}}(F, G_n)$ tends to a limit value denoted by $t_{\mathrm{ind}}(F)$. This means that $\mathbb{P}\left(\xi(k, G_n) = F\right) \to t_{\mathrm{ind}}(F)$ and so

$$\mathbb{E}(f(\xi(k, G_n))) \to \sum_{F:\,|V(F)|=k} t_{\mathrm{ind}}(F) \cdot f(F) = a_k,$$

since the number of simple graphs on k vertices is finite. In summary,

$$|f(G_n) - a_k| \le |f(G_n) - \mathbb{E}(f(\xi(k, G_n)))| + |\mathbb{E}(f(\xi(k, G_n))) - a_k| \le 2\varepsilon$$

provided n is large enough.

$(c) \Rightarrow (e)$: Suppose that (e) does not hold. In this case there exist $\varepsilon > 0$, further sequences (G_n) and $(G_n^{'})$ of weighted graphs, such that the dominant vertex-weights of both sequences tend to 0, $\delta_\square(G_n, G_n^{'}) \to 0$, and $|f(G_n) - f(G_n^{'})| \ge \varepsilon$. Using the compactness of $\mathcal{W}_{[0,1]}$ we can assume that both sequences are convergent. For this reason, the merged sequence $G_1, G_1^{'}, G_2, G_2^{'}, \ldots$ is also convergent. For the above merged sequence, by (c), the sequence $f(G_1), f(G_1^{'}), f(G_2), f(G_2^{'}), \ldots$ is covergent, that contradicts to $|f(G_n) - f(G_n^{'})| \ge \varepsilon$.

$(e) \Rightarrow (a)$: Suppose that (a) does not hold. In this case there exist $\varepsilon > 0$ and a sequence (G_n) such that $\max_i \frac{\alpha_i(G_n)}{\alpha_{G_n}} \le \frac{1}{n}$, and with probability at least ε the inequality $|f(G_n) - f(\xi(n, G_n))| > \varepsilon$ holds for all n. To this ε choose the corresponding n_0 and ε_0 as in the statement (e). Furtheron, because of

(3), the sequence $\delta_\Box(G_n, \xi(n, G_n))$ tends to 0 in probability. In particular, $\mathbb{P}\left(\delta_\Box(G_n, \xi(n, G_n)) < \varepsilon_0\right) \geq 1 - \frac{\varepsilon}{2}$. Using the definition of ε_0 and n_0 we get that $\mathbb{P}\left(|f(G_n) - f(\xi(n, G_n))| < \varepsilon\right) \geq 1 - \frac{\varepsilon}{2}$. This contradicts to the fact that with probability at least ε the opposite is true.

Now we prove that the statement (d) is also equivalent to the testability.

$(c), (e) \Rightarrow (d)$: Let $W \in \mathcal{W}_{[0,1]}$ be an arbitrary graphon. By Theorem 2 we can find a sequence (G_n) of weighted graphs with no dominant vertex-weights tending to W. Let $\tilde{f}(W)$ be the limit of $f(G_n)$. Because of (c) the limit exists, and due to the statement (e) this definition is correct. First we prove the continuity. Let $\varepsilon > 0$ be arbitrary. Using the statement (e), to $\frac{\varepsilon}{3}$ we assign the corresponding ε' and n_0. We show, that $||W - W'||_\Box \leq \frac{\varepsilon'}{3}$ implies $|\tilde{f}(W) - \tilde{f}(W')| \leq \varepsilon$. For this purpose let (G_n) be a sequence of weighted graphs with no dominant vertex-weights tending to W. We can choose a G from (G_n) such that the dominant vertex-weight of G is smaller than $\frac{1}{n_0}$; further, $\delta_\Box(G, W) < \frac{\varepsilon'}{3}$ and $|f(G) - \tilde{f}(W)| \leq \frac{\varepsilon}{3}$. Similarly, we can choose a G' from the sequence (G'_n) tending to W' with analogous properties. In this case $\delta_\Box(G, G') \leq \delta_\Box(G, W) + \delta_\Box(W, W') + \delta_\Box(W', G') \leq \varepsilon'$. By (e), $|f(G) - f(G'| \leq \frac{\varepsilon}{3}$, and hence,

$$|\tilde{f}(W) - \tilde{f}(W')| \leq |\tilde{f}(W) - f(G)| + |f(G)| - f(G')| + |f(G') - \tilde{f}(W')| \leq \varepsilon.$$

It remains to show that $\left|\tilde{f}(W_{G_n}) - f(G_n)\right| \to 0$, whenever $|V(G_n)| \to \infty$ with no dominant vertex-weights. On the contrary, suppose that there exists a sequence (G_n) with no dominant vertex-weights such that $\tilde{f}(W_{G_n}) - f(G_n)$ does not tend to 0. For the sake of simplicity we can assume that for some ε: $\left|\tilde{f}(W_{G_n}) - f(G_n)\right| > \varepsilon$ for all n. We can also assume that (G_n) converges to some graphon W in the $\delta_\Box$ metric. By Lemma 8, there is a sequence (G'_n) isomorphic to (G_n) such that $||W_{G'_n} - W||_\Box \to 0$. Using the statement (c), $\lim_{n\to\infty} f(G_n) = \lim_{n\to\infty} f(G'_n) = \tilde{f}(W)$. In addition, $\tilde{f}$ is continuous, and for this reason, $\lim_{n\to\infty} \tilde{f}(W_{G_n}) = \lim_{n\to\infty} \tilde{f}(W_{G'_n}) = \tilde{f}(W)$. But this is a contradiction.

$(d) \Rightarrow (c)$: Let (G_n) be a convergent sequence of weighted graphs with no dominant vertex-weights. Let W be its limit. So $\delta_\Box(W_{G_n}, W) \to 0$. In this way, by Lemma 8, we can relabel (G_n) into (G'_n) in such a way that $||W_{G'_n} - W||_\Box \to 0$. Therefore, using the continuity of $\tilde{f}$ we get $\tilde{f}(W_{G'_n}) - \tilde{f}(W) \to 0$. Since $f(G_n) - \tilde{f}(W_{G'_n}) = f(G_n) - \tilde{f}(W_{G_n})$, the last term tends to 0 because of the statement (d). Thus, $f(G_n)$ is convergent.

Remark 9 *The original testability theorem for simple graphs in [4] was formulated in terms of sampling without replacement. In the most important case, when the size of the sample is small compared to the size of the underlaying graph, the two sampling methods are approximately the same. Usually, this* *is the case in practical applications. In our definition of the testability of a weighted graph parameter we use sampling with replacement, but the testability could be defined by any randomization for which a large deviation result similar to that of Theorem 3 holds. However, equivalent statements (c), (d), (e) do not depend on the randomization, and we may expect their equivalence to statemants (a),(b) under an appropriate sampling with likelihood function strongly connected to* $t_{ind}(F, G)$ *and satisfying (3).*

4 Balanced multiway cuts

Lovász and coauthors [5] proved the testability of the maximum cut density. The minimum cut density is somewhat different. E.g., if a single vertex is loosely connected to a dense part, the minimum cut density of the whole graph is small, however, randomizing a smaller sample, with high probability, it will come from the dense part with a large minimum cut density.

To prove the testability of certain balanced minimum multiway cut densities we use the notions of statistical physics in the same way as in [5]. Most of these notions are self-explanatory. However, to be self-contained, we included some definitions for clarification together with the notion of a factor graph.

Let $G \in \mathcal{G}$ be a weighted graph on n vertices with vertex-weights $\alpha_1, \dots, \alpha_n$ and edge-weights β_{ij}'s. Let $q \leq n$ be a fixed positive integer, and $\mathcal{P}_q$ denote the set of q-partitions $P = (V_1, \dots, V_q)$ of the vertex set V. The non-empty, disjoint vertex-subsets sometimes are referred to as clusters or states. The *factor graph* or *q-quotient* of G with respect to the q-partition P is denoted by G/P and it is defined as the weighted graph on q vertices with vertex- and edge-weights

$$\alpha_i(G/P) = \frac{\alpha_{V_i}}{\alpha_G} \quad (i \in [q]) \quad \text{and} \quad \beta_{ij}(G/P) = \frac{e_G(V_i, V_j)}{\alpha_{V_i}\alpha_{V_j}} \quad (i, j \in [q]),$$

respectively. Let $\hat{\mathcal{S}}_q(G)$ denote the set of all q-quotients of G. The Hausdorff distance between $\hat{\mathcal{S}}_q(G)$ and $\hat{\mathcal{S}}_q(G')$ is defined by

$$d^{\mathrm{Hf}}(\hat{\mathcal{S}}_q(G), \hat{\mathcal{S}}_q(G')) = \max\{ \sup_{H \in \hat{\mathcal{S}}_q(G)} \inf_{H' \in \hat{\mathcal{S}}_q(G')} d_1(H, H') \, , \sup_{H' \in \hat{\mathcal{S}}_q(G')} \inf_{H \in \hat{\mathcal{S}}_q(G)} d_1(H, H') \},$$

where

$$d_1(H, H') = \sum_{i,j \in [q]} \left| \frac{\alpha_i(H)\alpha_j(H)\beta_{ij}(H)}{\alpha_H^2} - \frac{\alpha_i(H')\alpha_j(H')\beta_{ij}(H')}{\alpha_{H'}^2} \right| + \sum_{i \in [q]} \left| \frac{\alpha_i(H)}{\alpha_H} - \frac{\alpha_i(H')}{\alpha_{H'}} \right|$$

is the l_1-distance between two weighted graphs H and H' on the same number of vertices. Here especially, H and H' are factor graphs, and hence, $\alpha_H = \alpha_{H'} = 1$, therefore the denominators can be omitted.

Given the real symmetric $q \times q$ matrix $\mathbf{J}$ and the vector $\mathbf{h} \in \mathbb{R}^q$, the partitions $P \in \mathcal{P}_q$ also define a spin system on the weighted graph G. The so-called *ground state energy* of such a spin configuration is

$$\hat{\mathcal{E}}_q(G, \mathbf{J}, \mathbf{h}) = -\max_{P \in \mathcal{P}_q} \left(\sum_{i \in [q]} \alpha_i(G/P) h_i + \sum_{i,j \in [q]} \alpha_i(G/P)\alpha_j(G/P)\beta_{ij}(G/P) J_{ij} \right).$$

Here $\mathbf{J}$ is the so-called coupling-constant matrix, where J_{ij} represents the strength of interaction between states i and j, and $\mathbf{h}$ is the magnetic field. They carry physical meaning. We shall use only special $\mathbf{J}$ and $\mathbf{h}$, especially $\mathbf{h} = \mathbf{0}$.

Sometimes, we need balanced q-partitions to regulate the proportion of the cluster volumes. A slight balancing between the cluster volumes is achieved by fixing a positive real number c ($c \leq 1/q$). Let $\mathcal{P}_q^c$ denote the set of q-partitions of V such that $\frac{\alpha_{V_i}}{\alpha_G} \geq c \quad i \in [q]$, or equivalently, $c \leq \frac{\alpha_{V_i}}{\alpha_{V_j}} \leq \frac{1}{c}$ ($i \neq j$).

A more accurate balancing is defined by fixing a vector $\mathbf{a} = (a_1, \ldots, a_q)$ with components forming a probability distribution over $[q]$: $a_i > 0 \ i \in [q]$, $\sum_{i=1}^q a_i = 1$. Let $\mathcal{P}_q^{\mathbf{a}}$ denote the set of q-partitions of V such that $\left(\frac{\alpha_{V_1}}{\alpha_G}, \ldots, \frac{\alpha_{V_q}}{\alpha_G}\right)$ is approximately $\mathbf{a}$-distributed, that is

$$\left| \frac{\alpha_{V_i}}{\alpha_G} - a_i \right| \leq \frac{\alpha_{\max}(G)}{\alpha_G} \quad (i = 1, \ldots, q),$$

the right hand side tending to 0 as $|V(G)| \to \infty$ for weighted graphs with no dominant vertex-weights.

The *microcanonical ground state energy* of G given $\mathbf{a}$ and $\mathbf{J}$ ($\mathbf{h} = \mathbf{0}$) is

$$\hat{\mathcal{E}}_q^{\mathbf{a}}(G, \mathbf{J}) = -\max_{P \in \mathcal{P}_q^{\mathbf{a}}} \sum_{i,j \in [q]} \alpha_i(G/P)\alpha_j(G/P)\beta_{ij}(G/P) J_{ij}.$$

Remark 10 *In Theorem 2.14 of [5] it is proved that the convergence of the weighted graph sequence (G_n) with no dominant vertex-weights is equivalent to the convergence of its microcanonical ground state energies for any q, $\mathbf{a}$, and $\mathbf{J}$. Also, it is equivalent to the convergence of its q-quotients in Hausdorff distance for any q.*

Remark 11 *Under the same conditions, Theorem 2.15 of [5] states that the convergence of the above (G_n) implies the convergence of its ground state energies for any q, $\mathbf{J}$, and $\mathbf{h}$; further the convergence of the spectrum of (G_n).*

Using these facts, we investigate the testability of some special multiway cut densities defined in the forthcoming definitions.

Definition 12 *The minimum q-way cut density of G is*

$$f_q(G) = \min_{P \in \mathcal{P}_q} \frac{1}{\alpha_G^2} \sum_{i=1}^{q-1} \sum_{j=i+1}^{q} e_G(V_i, V_j),$$

the minimum c-balanced q-way cut density of G is

$$f_q^c(G) = \min_{P \in \mathcal{P}_q^c} \frac{1}{\alpha_G^2} \sum_{i=1}^{q-1} \sum_{j=i+1}^{q} e_G(V_i, V_j), \tag{6}$$

and the minimum **a***-balanced q-way cut density of G is*

$$f_q^{\mathbf{a}}(G) = \min_{P \in \mathcal{P}_q^{\mathbf{a}}} \frac{1}{\alpha_G^2} \sum_{i=1}^{q-1} \sum_{j=i+1}^{q} e_G(V_i, V_j).$$

Occasionally, we want to penalize cluster volumes that wildly differ. For this purpose we herein introduce the notions of weighted minimum cut densities.

Definition 13 *The minimum weighted q-way cut density of G is*

$$\mu_q(G) = \min_{P \in \mathcal{P}_q} \sum_{i=1}^{q-1} \sum_{j=i+1}^{q} \frac{1}{\alpha_{V_i} \cdot \alpha_{V_j}} \cdot e_G(V_i, V_j),$$

the minimum weighted c-balanced q-way cut density of G is

$$\mu_q^c(G) = \min_{P \in \mathcal{P}_q^c} \sum_{i=1}^{q-1} \sum_{j=i+1}^{q} \frac{1}{\alpha_{V_i} \cdot \alpha_{V_j}} \cdot e_G(V_i, V_j),$$

and the minimum weighted **a***-balanced q-way cut density of G is*

$$\mu_q^{\mathbf{a}}(G) = \min_{P \in \mathcal{P}_q^c} \sum_{i=1}^{q-1} \sum_{j=i+1}^{q} \frac{1}{\alpha_{V_i} \cdot \alpha_{V_j}} \cdot e_G(V_i, V_j).$$

Proposition 14 *$f_q(G)$ is testable for any $q \leq |V(G)|$.*

PROOF. Observe that $f_q(G)$ is a special ground state energy:

$$f_q(G) = \hat{\mathcal{E}}_q(G, \mathbf{J}, \mathbf{0}),$$

where the magnetic field is $\mathbf{0}$ and the $q \times q$ symmetric matrix $\mathbf{J}$ is the following: $J_{ii} = 0$ $i \in [q]$, further $J_{ij} = -1/2$ $(i \neq j)$. By Remark 10 and the equivalent statement (c) of Theorem 5, the minimum q-way cut density is testable for any q.

However, this statement is of not much use, since $f_q(G_n) \to 0$, in the lack of dominant vertex-weights. In fact, the minimum q-way cut density is trivially estimated from above by

$$f_q(G_n) \leq (q-1)\frac{\alpha_{max}(G_n)}{\alpha_{G_n}} + \binom{q-1}{2}\left(\frac{\alpha_{max}(G_n)}{\alpha_{G_n}}\right)^2$$

that tends to 0 provided $\alpha_{\max}(G_n)/\alpha_{G_n} \to 0$ as $n \to \infty$.

Proposition 15 *$f_q^{\mathbf{a}}(G)$ is testable for any $q \leq |V(G)|$ and distribution $\mathbf{a}$ over $[q]$.*

PROOF. Choose $\mathbf{J}$ as in the proof of Proposition 14. In this way, $f_q^{\mathbf{a}}(G)$ is a special microcanonical ground state energy:

$$f_q^{\mathbf{a}}(G) = \hat{\mathcal{E}}_q^{\mathbf{a}}(G, \mathbf{J}). \tag{7}$$

Hence, by Remark 10, the convergence of (G_n) is equivalent to the convergence of $f_q^{\mathbf{a}}(G_n)$ for any q and any distribution $\mathbf{a}$ over $[q]$. Therefore, by the equivalent statement (c) of Theorem 5, the testability of the minimum $\mathbf{a}$-balanced q-way cut density also follows.

Proposition 16 *$f_q^c(G)$ is testable for any $q \leq |V(G)|$ and $c \leq 1/q$.*

PROOF. Theorem 4.7 and Theorem 5.5 of [5] imply that for any two weighted graphs G, G'

$$|\hat{\mathcal{E}}_q^{\mathbf{a}}(G, \mathbf{J}) - \hat{\mathcal{E}}_q^{\mathbf{a}}(G', \mathbf{J})| \leq (3/2 + \kappa) \cdot d^{\mathrm{Hf}}(\hat{S}_q(G), \hat{S}_q(G')), \tag{8}$$

where $\kappa = o(\min\{|V(G)|, |V(G')|\})$ is a negligible small constant, provided the number of vertices of G and G' is sufficiently large. By Remark 11 we know that if (G_n) converges, its q-quotients also converge in Hausdorff distance, consequently form a Cauchy-sequence. This means that for any $\varepsilon > 0$ there is an N_0 such that for $n, m > N_0$: $d^{\mathrm{Hf}}(\hat{S}_q(G_n), \hat{S}_q(G_m)) < \varepsilon$. We want to prove that for $n, m > N_0$: $|f_q^c(G_n) - f_q^c(G_m)| < 2\varepsilon$. On the contrary, suppose that there are $n, m > N_0$ such that $|f_q^c(G_n) - f_q^c(G_m)| \geq 2\varepsilon$. Say, $f_q^c(G_n) \geq f_q^c(G_m)$. Let $A := \{\mathbf{a} : a_i \geq c\,,\ i = 1, \ldots, q\}$ is the subset of special c-balanced distributions over $[q]$. On the one hand,

$$f_q^c(G_m) = \min_{\mathbf{a} \in A} f_q^{\mathbf{a}}(G_m) = f_q^{\mathbf{a}^*}(G_m)$$

for some $\mathbf{a}^* \in A$. On the other hand, by (7) and (8), $f_q^{\mathbf{a}^*}(G_n) - f_q^{\mathbf{a}^*}(G_m) \leq (\frac{3}{2} + \kappa)\varepsilon$, that together with the indirect assumption implies that $f_q^c(G_n) - f_q^{\mathbf{a}^*}(G_n) \geq (\frac{1}{2} - \kappa)\varepsilon > 0$ for this $\mathbf{a}^* \in A$. But this contradicts to the fact that $f_q^c(G_n)$ is the minimum of $f_q^{\mathbf{a}}(G_n)$'s over A. Thus, $f_q^c(G_n)$ is also a Cauchy sequence, and being a real sequence, it is also convergent.

Concerning the penalized densities, trivially,

$$\mu_q(G) = \min_{P \in \mathcal{P}_q} \sum_{i=1}^{q-1} \sum_{j=i+1}^{q} \beta_{ij}(G/P).$$

In fact, $\mu_q(G)$ is not testable as we can show an example where $\mu_q(G_n) \to 0$, but randomizing a sufficiently large part of G_n, the weighted minimum q-way cut density of that part is constant. The example is for $q = 2$ and for a simple graph on n vertices such that order of $\sqrt{n}$ vertices are connected with a single edge to the remaining vertices that form a complete graph. Then $\mu_2(G_n) \to 0$, but randomizing a sufficiently large part of the graph, with high probability, it will be a subgraph of the complete graph, whose minimum 2-way cut density is of constant order.

Proposition 17 *$\mu_q^{\mathbf{a}}(G)$ is testable for any $q \leq |V(G)|$ and distribution $\mathbf{a}$ over $[q]$.*

PROOF. By the definition of Hausdorff distance, the convergence of q-quotients guarantees the convergence of

$$\mu_q^{\mathbf{a}}(G) = \min_{P \in \mathcal{P}_q^{\mathbf{a}}} \sum_{i=1}^{q-1} \sum_{j=i+1}^{q} \beta_{ij}(G/P) \tag{9}$$

for any $\mathbf{a}$ and q in the following way. Let $\hat{\mathcal{S}}_q^{\mathbf{a}}(G)$ denote the set of factor graphs of G with respect to partitions in $\mathcal{P}_q^{\mathbf{a}}$. As a consequence of Lemma 4.5 and Theorem 5.4 of [5], for any two weighted graphs G, G'

$$\max_{\mathbf{a}} d^{Hf}(\hat{\mathcal{S}}_q^{\mathbf{a}}(G), \hat{\mathcal{S}}_q^{\mathbf{a}}(G')) \leq (3 + \kappa) \cdot d^{Hf}(\hat{\mathcal{S}}_q(G), \hat{\mathcal{S}}_q(G')), \tag{10}$$

where $\kappa = o(\min\{|V(G)|, |V(G')|\})$.

By Remark 10, for a convergent graph-sequence (G_n), the sequence $\hat{\mathcal{S}}_q(G_n)$ converges, and by the inequality (10), $\hat{\mathcal{S}}_q^{\mathbf{a}}(G_n)$ also converges in Hausdorff distance for any distribution $\mathbf{a}$ over $[q]$. As they form a Cauchy sequence, $\forall \varepsilon$ $\exists N_0$ such that for $n, m > N_0$

$$d^{Hf}(\hat{\mathcal{S}}_q^{\mathbf{a}}(G_n), \hat{\mathcal{S}}_q^{\mathbf{a}}(G_m)) < \varepsilon$$

uniformly for any $\mathbf{a}$. In view of the Hausdorff distance's definition, this means that for any q-quotient $H \in \hat{\mathcal{S}}_q^{\mathbf{a}}(G_n)$ there exists (at least one) q-quotient $H' \in \hat{\mathcal{S}}_q^{\mathbf{a}}(G_m)$, and vice versa, for any $H' \in \hat{\mathcal{S}}_q^{\mathbf{a}}(G_m)$ there exists (at least one) $H \in \hat{\mathcal{S}}_q^{\mathbf{a}}(G_n)$ such that $d_1(H, H') < \varepsilon$. (In fact, the maximum distance between the elements of the above pairs is less than ε. Note that the symmetry in the definition of the Hausdorff distance is important: the pairing exhausts the sets even if they have different cardinalities.)

Using the fact that the vertex-weights of such a pair H es H' are almost the same (the coordinates of the vector $\mathbf{a}$), by the notation $a = \min_{i\in[q]} a_i$, the following argument is valid for n, m large enough:

$$\begin{aligned} 2a^2 \sum_{i\neq j} |\beta_{ij}(H) - \beta_{ij}(H')| &\leq \sum_{i,j=1}^{q} a^2 |\beta_{ij}(H) - \beta_{ij}(H')| \leq \\ \leq \sum_{i,j=1}^{q} |\alpha_i(H)\alpha_j(H)\beta_{ij}(H) &- \alpha_i(H')\alpha_j(H')\beta_{ij}(H')| = d_1(H, H') < \varepsilon. \end{aligned} \tag{11}$$

Therefore

$$|\sum_{i=1}^{q-1}\sum_{j=i+1}^{q} \beta_{ij}(H) - \sum_{i=1}^{q-1}\sum_{j=i+1}^{q} \beta_{ij}(H')| < \frac{\varepsilon}{2a^2} := \varepsilon',$$

and because $\sum_{i=1}^{q-1}\sum_{j=i+1}^{q} \beta_{ij}(H)$ and $\sum_{i=1}^{q-1}\sum_{j=i+1}^{q} \beta_{ij}(H')$ are individual terms behind the minimum in (9), the above inequality holds for their minima over $\mathcal{P}_q^{\mathbf{a}}$ as well:

$$|\mu_q^{\mathbf{a}}(G_n) - \mu_q^{\mathbf{a}}(G_m)| < \varepsilon'. \tag{12}$$

Consequently, the sequence $\mu_q^{\mathbf{a}}(G_n)$ is a Cauchy sequence, and being a real sequence, it is also convergent. Thus $\mu_q^{\mathbf{a}}$ is testable.

Remark 18 *The testability of μ_q, apparently, does not follow in the same way due to presence of distinct vetex-weights in H and H'. Thus, the smallness of $d_1(H, H')$ does not imply the closeness of their edge-weights.*

However, as the testability of $f_q^{\mathbf{a}}$ implied the testability of f_q^c, the testability of $\mu_q^{\mathbf{a}}$ also implies the testability of μ_q^c.

Proposition 19 *$\mu_q^c(G)$ is testable for any $q \leq |V(G)|$ and $c \leq 1/q$.*

PROOF. The proof is analogous to that of Proposition 16 using equation (12) instead of equation (8). By the pairing argument of the proof of Proposition 17, the real sequence $\mu_q^c(G_n)$ is a Cauchy sequence, and therefore, convergent. This immediately implies the testability of μ_q^c.

By Remark 11, the convergence of (G_n) also implies the convergence of the spectra, though the convergence of the spectrum itself is weaker than the convergence of the graph sequence. Without going into details, we remark that in [2], $f_q(G)$ and $\mu_q(G)$ were bounded from below by the q smallest Laplacian eigenvalues of G. An upper estimate can also be constructed and we conjecture that in case of testable parameters an asymptotic estimate is also valid.

5 Minimum cut as a quadratic programming problem

In Section 4, we proved that f_q^c is a testable weighted graph parameter. Now, we extend it to graphons.

Proposition 20 *Let us define the graphon functional $\tilde{f}_q^c$ in the following way:*

$$\tilde{f}_q^c(W) := \inf_{Q\in\mathcal{Q}_q^c} \sum_{i=1}^{q-1}\sum_{j=i+1}^{q} \iint_{S_i\times S_j} W(x,y)\,dx\,dy = \inf_{Q\in\mathcal{Q}_q^c} \tilde{f}_q(W;S_1,\dots,S_q) \quad (13)$$

where the infimum is taken over all the c-balanced Lebesgue-measurable partitions $Q=(S_1,\dots,S_q)$ of [0,1]. For these, $\sum_{i=1}^q \lambda(S_i)=1$ and $\lambda(S_i)\geq c$ ($i\in[q]$), where λ denotes the Lebesgue-measure, and $\mathcal{Q}_q^c$ denotes the set of c-balanced q-partitions of [0,1]. We state that $\tilde{f}_q^c$ is the extension of f_q^c in the following sense: If (G_n) is a convergent weighted graph sequence with uniformly bounded edge-weights and no dominant vertex-weights, then denoting by W the essentially unique limit graphon of the sequence (see Theorem 2), $f_q^c(G_n)\to\tilde{f}_q^c(W)$ as $n\to\infty$.

PROOF. First we show that $\tilde{f}_q^c$ is continuous in the cut-norm. As $\tilde{f}_q^c(W)$ is insensitive to measure preserving maps of W, it suffices to prove that to any ε we can find ε' such that for any two graphons W,U with $\|W-U\|_\square<\varepsilon'$, the relation $|\tilde{f}_q^c(W)-\tilde{f}_q^c(U)|<\varepsilon$ also holds. By the definition of the cut-norm, for any Lebesgue-measurable q-partition $(S_1,\dots,S_q)$ of [0,1], the relation

$$|\iint_{S_i\times S_j}(W(x,y)-U(x,y))\,dx\,dy|\leq\varepsilon' \qquad (i\neq j)$$

holds. Summing up for the $i\neq j$ pairs

$$|\sum_{i=1}^{q-1}\sum_{j=i+1}^{q}\iint_{S_i\times S_j}W(x,y)\,dx\,dy-\sum_{i=1}^{q-1}\sum_{j=i+1}^{q}\iint_{S_i\times S_j}U(x,y)\,dx\,dy|\leq\binom{q}{2}\varepsilon'. \quad (14)$$

Therefore

$$\inf_{(S_1,\dots,S_q)\in\mathcal{Q}_q^c}\sum_{i=1}^{q-1}\sum_{j=i+1}^{q}\iint_{S_i\times S_j}W(x,y)\,dx\,dy\geq\inf_{(S_1,\dots,S_q)\in\mathcal{Q}_q^c}\sum_{i=1}^{q-1}\sum_{j=i+1}^{q}\iint_{S_i\times S_j}U(x,y)\,dx\,dy-\binom{q}{2}\varepsilon'$$

and vice versa,

$$\inf_{(S_1,\dots,S_q)\in\mathcal{Q}_q^c}\sum_{i=1}^{q-1}\sum_{j=i+1}^{q}\iint_{S_i\times S_j}U(x,y)\,dx\,dy\ge\inf_{(S_1,\dots,S_q)\in\mathcal{Q}_q^c}\sum_{i=1}^{q-1}\sum_{j=i+1}^{q}\iint_{S_i\times S_j}W(x,y)\,dx\,dy-\binom{q}{2}\varepsilon'.$$

Consequently the absolute difference of the two infima is bounded from above by $\binom{q}{2}\varepsilon'$. Thus, $\varepsilon'=\varepsilon/\binom{q}{2}$ will do.

Let (G_n) be a convergent weighted graph sequence with uniforly bounded edge-weights and no dominant vertex-weights. By Theorem 2, there is an essentially unique graphon W such that $G_n\to W$, i.e., $\delta_\square(W_{G_n},W)\to 0$ as $n\to\infty$. By the continuity of $\tilde{f}_q^c$,

$$\tilde{f}_q^c(W_{G_n})\to\tilde{f}_q^c(W),\qquad n\to\infty. \tag{15}$$

Suppose that

$$\tilde{f}_q^c(W_{G_n})=\tilde{f}_q(W_{G_n};S_1^*,\dots,S_q^*),$$

that is the infimum in (13) is attained at the c-balanced Lebesgue-measurable q-partition $(S_1^*,\dots,S_q^*)$ of [0,1].

Let G_{nq}^* be the q-fold blown-up of G_n with respect to $(S_1^*,\dots,S_q^*)$. It is a weighted graph on at most nq vertices defined in the following way. Let $I_1,\dots,I_n$ be consecutive intervals of [0,1] such that $\lambda(I_j)=\alpha_j(G_n)$, $j=1,\dots,n$. The weight of the vertex labeled by ju of G_{nq}^* is $\lambda(I_j\cap S_u^*)$, $u\in[q]$, $j\in[n]$, while the edge-weights are $\beta_{ju,iv}(G_{nq}^*)=\beta_{ji}(G_n)$. Trivially, the graphons W_{G_n} and $W_{G_{nq}^*}$ essentially define the same stepfunction, hence $\tilde{f}_q^c(W_{G_n})=\tilde{f}_q^c(W_{G_{nq}^*})$. Therefore, by (15),

$$\tilde{f}_q^c(W_{G_{nq}^*})\to\tilde{f}_q^c(W),\qquad n\to\infty. \tag{16}$$

As $\delta_\square(G_n,G_{nq}^*)=\delta_\square(W_{G_n},W_{G_{nq}^*})=0$, by part (e) of Theorem 5 it follows that

$$|f_q^c(G_{nq}^*)-f_q^c(G_n)|\to 0,\qquad n\to\infty. \tag{17}$$

Finally, by the construction of G_{nq}^*, $\tilde{f}_q^c(W_{G_{nq}^*})=f_q^c(G_{nq}^*)$, and hence,

$$|f_q^c(G_n)-\tilde{f}_q^c(W)|\le|f_q^c(G_n)-f_q^c(G_{nq}^*)|+|\tilde{f}_q^c(W_{G_{nq}^*})-\tilde{f}_q^c(W)|$$

that, in view of (16), (17), implies the required statement.

Corollary 21 *In Section 3, while proving Theorem 5, an essentially unique extension of a testable graph parameter to graphons was given. By Proposition 20, the above* $\tilde{f}_q^c$ *is the desired extension of* f_q^c*, therefore part (d) of Theorem 5 is also applicable to it: For a weighted graph sequence* (G_n) *with* $\max_i\frac{\alpha_i(G_n)}{\alpha_{G_n}}\to 0$*, the limit relation* $\tilde{f}_q^c(W_{G_n})-f_q^c(G_n)\to 0$ *also holds as* $n\to\infty$*.*

Corollary 21 gives rise to approximate the minimum c-balanced q-way cut density of a weighted graph on "many" vertices with no dominant vertex weights by the extended c-balanced q-way cut density of the stepfunction graphon assigned to the graph. In this way, the discrete optimization problem can be formulated as a quadratic programming task with linear equality and inequality constraints.

To this end, let us investigate a fixed weighted graph G on n vertices (n is large). To simplify notation we drop the subscript n, and G in the arguments of the vertex- and edge-weights. As $f_q^c(G)$ is invariant under the scale of the vertices, we can suppose that $\alpha_G = \sum_{i=1}^n \alpha_i = 1$. As $\beta_{ij} \in [0,1]$, W_G is uniformly bounded by 1. Recall that $W_G(x,y) = \beta_{ij}$, if $x \in I_i$, $y \in I_j$, where $\lambda(I_j) = \alpha_j$ $(j = 1, \dots, n)$ and $I_1, \dots, I_n$ are consecutive intervals of [0,1].

For fixed q and $c \le 1/q$, $f_q(G; V_1, \dots, V_q) = \frac{1}{\alpha_G^2} \sum_{i=1}^{q-1} \sum_{j=i+1}^q e_G(V_i, V_j)$ is a function taking on discrete values over c-balanced q-partitions $P = (V_1, \dots, V_q) \in \mathcal{P}_q^c$ of the vertices of G. As $n \to \infty$, by Corollary 21, this function approaches $\tilde{f}_q(W_G; S_1, \dots, S_q)$ that is already a continuous function over c-balanced q-partitions $Q = (S_1, \dots, S_q) \in \mathcal{Q}_q^c$ of [0,1]. In fact, this continuous function can be regarded as a multilinear function of the variable

$$\mathbf{x} = (x_{11}, \dots, x_{1n}, x_{21}, \dots, x_{2n}, \dots, x_{q1}, \dots, x_{qn})^T \in \mathbb{R}^{nq}$$

where the coordinate indexed by ij is

$$x_{ij} = \lambda(S_i \cap I_j), \quad j = 1, \dots, n; \quad i = 1, \dots q.$$

Hence,

$$\tilde{f}_q(W_G; S_1, \dots, S_q) = \tilde{f}_q(\mathbf{x}) = \sum_{i=1}^{q-1} \sum_{i'=i+1}^{q} \sum_{j=1}^{n} \sum_{j'=1}^{n} x_{ij} x_{i'j'} \beta_{jj'} = \frac{1}{2} \mathbf{x}^T (\mathbf{A} \otimes \mathbf{B}) \mathbf{x},$$

where – denoting by $\mathbf{1}_{q\times q}$ and $\mathbf{I}_{q\times q}$ the $q \times q$ all 1's and the identity matrix, respectively – the eigenvalues of the $q \times q$ symmetric matrix $\mathbf{A} = \mathbf{1}_{q\times q} - \mathbf{I}_{q\times q}$ are the number $q-1$ and -1 with multiplicity $q-1$, while those of the $n \times n$ symmetric matrix $\mathbf{B} = (\beta_{ij})$ are $\lambda_1 \ge \dots \ge \lambda_n$. Latter one being a Frobenius-type matrix, $\lambda_1 > 0$. The eigenvalues of the Kronecker-product $\mathbf{A} \otimes \mathbf{B}$ are the numbers $(q-1)\lambda_i$ $(i = 1, \dots, n)$ and $-\lambda_i$ with multiplicity $q-1$ $(i = 1, \dots, n)$. Therefore the above quadratic form is indefinite.

Hence, we have the following *quadratic programming* task:

$$\begin{aligned} &\text{minimize} \quad && \tilde{f}_q(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T (\mathbf{A} \otimes \mathbf{B}) \mathbf{x} \\ &\text{subject to} \quad && \mathbf{x} \ge 0; \quad \sum_{i=1}^q x_{ij} = \alpha_j \quad (j \in [n]); \quad \sum_{j=1}^n x_{ij} \ge c \quad (i \in [q]). \end{aligned} \tag{18}$$

The feasible region is the closed convex polytope of (18), and it is, in fact, in an $n(q-1)$-dimensional hyperplane of $\mathbb{R}^{nq}$. The gradient of the objective function $\nabla \tilde{f}_q(\mathbf{x}) = (\mathbf{A} \otimes \mathbf{B})\mathbf{x}$ cannot be $\mathbf{0}$ in the feasible region, provided the weight matrix $\mathbf{B}$, and hence $\mathbf{A} \otimes \mathbf{B}$ is non singular.

The arg-min of the quadratic programming task (18) is one of the Kuhn–Tucker points (giving relative minima of the indefinite quadratic form over the feasible region), that can be found by numerical algorithms (by tracing back the problem to a linear programming task), see [1].

Eventually, we give the extension of the testable weighted graph parameter μ_q^c to graphons.

Proposition 22 *Let us define the graphon functional $\tilde{\mu}_q^c$ in the following way:*

$$\tilde{\mu}_q^c(W) := \inf_{Q \in \mathcal{Q}_q^c} \sum_{i=1}^{q-1} \sum_{j=1}^{q} \iint_{S_i \times S_j} \frac{1}{\lambda(S_i)\lambda(S_j)} W(x,y)\, dx\, dy.$$

We state that $\tilde{\mu}_q^c$ is the extension of μ_q^c in the following sense: If (G_n) is a convergent weighted graph sequence with uniforly bounded edge-weights and no dominant vertex-weights, then denoting by W the essentially unique limit graphon of the sequence (see Theorem 2), $\mu_q^c(G_n) \to \tilde{\mu}_q^c(W)$ as $n \to \infty$.

The proof is analogous to that of Proposition 20, after we have proved that $\tilde{\mu}_q^c$ is continuous in the cut-norm. In fact, with estimates, analogous to (14), $\varepsilon' = \varepsilon c^2 / \binom{q}{2}$ will do.

Consequently, $\tilde{\mu}_q^c(W_G) - \mu_q^c(G) \to 0$ as $V(G) \to \infty$ with no dominant vertex-weights. This fact also gives rise to approximate the minimum c-balanced weighted q-way cut density of a large graph by quadratic programming methods.

6 Convergence of noisy graph sequences

Now, we use the above theory for perturbations. If not stated otherwise, the vertex-weights are equal (say 1), and a weighted graph G on n vertices is identified with its $n \times n$ symmetric weight matrix $\mathbf{A}$. Let $G_{\mathbf{A}}$ denote the weighted graph with unit vertex-weights and edge-weights that are entries of $\mathbf{A}$.

Definition 23 *Let w_{ij} $(1 \le i \le j \le n)$ be independent random variables defined on the same probability space, and $w_{ji} = w_{ij}$. $\mathbb{E}(w_{ij}) = 0$ $(\forall i,j)$ and the w_{ij}'s are uniformly bounded, i.e., there is a constant $K > 0$ – that does* *not depend of n – such that $|w_{ij}| \le K$, $\forall i,j$. The $n \times n$ symmetric real random matrix $\mathbf{W} = (w_{ij})_{1\le i\le n,\, 1\le j\le n}$ is called a Wigner-noise.*

The corresponding edge-weighted graph $G_{\mathbf{W}}$ is called a *Wigner-graph*. To indicate that the size n is expanding, we use the notations $\mathbf{W}_n$ and $G_{\mathbf{W}_n}$.

Definition 24 *The $n \times n$ symmetric real matrix $\mathbf{B}$ is a blown-up matrix, if there is a $q \times q$ symmetric so-called pattern matrix $\mathbf{P}$ with entries $0 < p_{ij} < 1$, and there are positive integers $n_1, \dots, n_q$ with $\sum_{i=1}^q n_i = n$, such that – after rearranging its rows and columns – the matrix $\mathbf{B}$ can be divided into $q \times q$ blocks, where block (i,j) is an $n_i \times n_j$ matrix with entries all equal to p_{ij} $(1 \le i, j \le n)$.*

Fix $\mathbf{P}$, blow it up to an $n \times n$ matrix $\mathbf{B}_n$, and consider the noisy matrix $\mathbf{A}_n = \mathbf{B}_n + \mathbf{W}_n$ as $n_1, \dots, n_q \to \infty$ at the same rate.

Remark 25 *While perturbing $\mathbf{B}_n$ by $\mathbf{W}_n$, for the uniform bound of the entries of $\mathbf{W}_n$ the condition*

$$K \le \min\{\min_{i,j\in[q]} p_{ij}\;,\; 1 - \max_{i,j\in[q]} p_{ij}\} \tag{19}$$

is satisfied. In this way, the entries of $\mathbf{A}_n$ are in the [0,1] interval, and hence, $G_{\mathbf{A}_n} \in \mathcal{G}$.

We remark that $G_{\mathbf{W}_n} \notin \mathcal{G}$, but $W_{G_{\mathbf{W}_n}} \in \mathcal{W}$ and the theory of bounded graphons applies to it. By adding an appropriate Wigner-noise to $\mathbf{B}_n$, we can achieve that $\mathbf{A}_n$ becomes a 0-1 matrix: its entries are equal to 1 with probability p_{ij} and 0 otherwise within the block of size $n_i \times n_j$ (after rearranging its rows and columns). In this case, the corresponding noisy graph $G_{\mathbf{A}_n}$ is a random simple graph.

As $\operatorname{rank}(\mathbf{B}_n) = q$ and $\|\mathbf{W}_n\| = \mathcal{O}(\sqrt{n})$ almost surely $(n \to \infty)$, the noisy matrix $\mathbf{A}_n$ almost surely has q protruding eigenvalues (of order n), and all the other eigenvalues are of order $\sqrt{n}$, there is a spectral gap between the q largest and the other eigenvalues $\mathbf{A}_n$.

Let $\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_q)$ be the $n \times q$ matrix containing the eigenvectors belonging to the q protruding eigenvalues of $\mathbf{A}_n$ in its columns. The rows of $\mathbf{X}_n$, that is the vectors $\mathbf{x}^1, \dots, \mathbf{x}^n \in \mathbb{R}^q$ are regarded as q-dimensional representatives of the vertices of $G_{\mathbf{A}_n}$. The q-variance of the representatives is

$$S_q^2(\mathbf{X}_n) = \sum_{i=1}^{q} \sum_{j\in V_i} \|\mathbf{x}^j - \bar{\mathbf{x}}^i\|^2,$$

where $\bar{\mathbf{x}}^i = \frac{1}{n_i}\sum_{j\in V_i} \mathbf{x}^j$.

In [3] we proved that

$$S_q^2(\mathbf{X}_n) = \mathcal{O}\left(\frac{1}{n}\right)$$

almost surely, under the growth condition $n_i/n \geq c$ $(i = 1, \dots, q)$.

In the other direction: for sufficiently large n, under some conditions, we can separate an $n \times n$ symmetric "error-matrix" $\mathbf{E}$ from $\mathbf{A}$, such that $\|\mathbf{E}\| = \mathcal{O}(\sqrt{n})$ and the remaining matrix $\mathbf{A} - \mathbf{E}$ is a blown-up matrix $\mathbf{B}$ of "low rank". Consequently, $G_{\mathbf{B}}$ is a weighted graph with homogeneous edge-densities within the clusters (determined by the blow-up). It resembles to the weak Szemerédi-partition, cf. [9], but the error-term is bounded in spectral norm, instead of the cut-norm. However, by large deviations, we can prove that the cut-norm of a Wigner-graph tends to zero almost surely as $n \to \infty$.

Theorem 26 *For any sequence* $(G_{\mathbf{W}_n})$ *of Wigner-graphs*

$$\lim_{n\to\infty} \|W_{G_{\mathbf{W}_n}}\|_\square = 0 \qquad (n \to \infty)$$

almost surely.

To prove the theorem, we need a proposition that is an easy consequence of Azuma's martingale inequality, see Theorem 5.3 of [6].

Proposition 27 *Let* $X_1, X_2, \dots, X_N$ *be i.i.d. random variables with zero mean and* $|X_i| \leq 1$, $i = 1, \dots, n$. *Then*

$$\mathbb{P}\left(\left|\sum_{j=1}^{N} X_j\right| > \gamma N\right) < 2\exp\left(-\frac{N\gamma^2}{2}\right), \qquad 0 < \gamma < 1. \tag{20}$$

PROOF. Now we are ready to prove Theorem 26. By the definition of the cut-norm of a stepfunction graphon and [8],

$$\|W_{G_{\mathbf{W}_n}}\|_\square = \frac{1}{n^2} \max_{U,T \subset [n]} \left|\sum_{i\in U}\sum_{j\in T} w_{ij}\right|. \tag{21}$$

We remark that $\max_{U,T\subset[n]} \left|\sum_{i\in U}\sum_{j\in T} w_{ij}\right|$ is the cut-norm of the matrix $\mathbf{W}_n$ defined in [7].

To make the entries behind the double sum of (21) independent, we use formulas (7.2), (7.3) of [4]:

$$\|W_{G_{\mathbf{W}_n}}\|_\square \leq 6 \max_{U\subset[n]} \frac{1}{n^2} \left|\sum_{i\in U}\sum_{j\in[n]\setminus U} w_{ij}\right|. \tag{22}$$

Apply Proposition 27 for a subsequence of length N of entries of $\mathbf{W}_n$ which does not contain w_{ij} and w_{ji} simultaneously. Namely, $i \in U$, $j \in [n] \setminus U$, $N = |U| \cdot (n - |U|)$. Remark that $n - 1 \le N \le n^2/4$.

We distinguish between two cases.

- *Case 1.* Suppose that $N \le n^{3/2}$ and apply Proposition 27 for the right hand side of (22) with $\gamma = n^{4/10}$:

$$\mathbb{P}\left(\frac{1}{n^2}\left|\sum_{i\in U}\sum_{j\in[n]\setminus U} w_{ij}\right| > n^{-1/10}\right) = \mathbb{P}\left(\left|\sum_{i\in U}\sum_{j\in[n]\setminus U} w_{ij}\right| > n^{3/2}n^{4/10}\right) \le$$
$$\le \mathbb{P}\left(\left|\sum_{i\in U}\sum_{j\in[n]\setminus U} w_{ij}\right| > Nn^{4/10}\right) < 2\exp\left(-\frac{n^{9/5}}{2}\right). \tag{23}$$

  In the first inequality we used the condition $N \le n^{3/2}$, while the second inequality follows from (20) and the fact that $N \ge n - 1$.
- *Case 2.* Now suppose that $N > n^{3/2}$ and apply Proposition 27 for the right hand side of (22) with $\gamma = 4n^{-1/10}$:

$$\mathbb{P}\left(\frac{1}{n^2}\left|\sum_{i\in U}\sum_{j\in[n]\setminus U} w_{ij}\right| > n^{-1/10}\right) = \mathbb{P}\left(\left|\sum_{i\in U}\sum_{j\in[n]\setminus U} w_{ij}\right| > n^2 n^{-1/10}\right) \le$$
$$\le \mathbb{P}\left(\left|\sum_{i\in U}\sum_{j\in[n]\setminus U} w_{ij}\right| > 4Nn^{-1/10}\right) < 2\exp\left(-8n^{13/10}\right). \tag{24}$$

  In the first inequality we used the fact $4N \le n^2$, while the second inequality follows from (20) and the condition $N > n^{3/2}$.

As for large values of n the right hand side of (24) is greater than that of (23), the probability $\mathbb{P}\left(\max_{U\subset[n]}\left|\sum_{i\in U}\sum_{j\in[n]\setminus U} w_{ij}\right| > n^{-1/10}\right)$ can be bounded by the number of possible 2-partitions of [n] times the right hand side of (24):

$$\mathbb{P}\left(\|W_{G_{\mathbf{W}_n}}\|_\square > 6 \cdot (n)^{-1/10}\right) < 2^{n+1}\exp\left(-8n^{13/10}\right). \tag{25}$$

As the right hand side of (25) is a general term of a convergent series, the statement of the theorem follows by the Borel-Cantelli Lemma.

Remark 28 *Let $\mathbf{A}_n := \mathbf{B}_n + \mathbf{W}_n$ and $n_1, \dots, n_q \to \infty$ in such a way that $\lim_{n\to\infty} \frac{n_i}{n} = r_i$ $(i = 1, \dots, q)$, $n = \sum_{i=1}^q n_i$; further, for the uniform bound K of the entries of the "noise" matrix $\mathbf{W}_n$ the condition (19) is satisfied. Under these conditions, Theorem 26 implies that the "noisy" graph sequence $(G_{\mathbf{A}_n}) \subset \mathcal{G}$ converges almost surely in the $\delta_\square$ metric. It is easy to see that the almost sure limit is the stepfunction W_H, where the factor graph $H = G_{\mathbf{B}_n}/P$*

does not depend on n, as P is the q-partition of the vertices of $G_{\mathbf{B}_n}$ with resepect to the blow-up (with cluster sizes $n_1, \dots, n_q$). Actually, the vertex- and edge-weights of the weighted graph H are

$$\alpha_i(H) = r_i \quad (i \in [q]), \qquad \beta_{ij}(H) = \frac{n_i n_j p_{ij}}{n_i n_j} = p_{ij} \quad (i, j \in [q]).$$

Remark 29 *Under the conditions of Remark 28, as $(G_{\mathbf{A}_n}) \subset \mathcal{G}$ converges almost surely and f_q, f_q^c, $f_q^{\mathbf{a}}$, μ_q^c, $\mu_q^{\mathbf{a}}$ are testable graph parameters, the sequences $f_q(G_{\mathbf{A}_n})$, $f_q^c(G_{\mathbf{A}_n})$, $f_q^{\mathbf{a}}(G_{\mathbf{A}_n})$, $\mu_q^c(G_{\mathbf{A}_n})$, $\mu_q^{\mathbf{a}}(G_{\mathbf{A}_n})$ also converge almost surely. However, the almost sure limits are the corresponding extended $\tilde{f}_q$- or $\tilde{\mu}_q$-values of the graphon W_H and not the f_q- or μ_q-values of H. For example, in Section 4, we have shown that $f_q(G_{\mathbf{A}_n}) \to 0$ $(n \to \infty)$, but $f_q(H) = \sum_{i=1}^{q-1} \sum_{j=i+1}^{q} p_{ij} \neq 0$.*

Acknowledgements

We are indebted to László Lovász and Katalin Friedl for inspiring discussions.

References

[1] Bazaraa, M. S., Shetty, C. M., Nonlinear Programming, Theory and Algorithms, Wiley, New York (1979).

[2] Bolla, M., Tusnády, G., Spectra and Optimal Partitions of Weighted Graphs, Discrete Mathematics 128 (1994), 1-20.

[3] Bolla, M., Recognizing Linear Structure in Noisy Matrices, Lin. Alg. Appl 402 (2005), 228-244.

[4] Borgs, C., Chayes, J. T., Lovász, L., Sós, V. T., Vesztergombi, K., Convergent Sequences of Dense Graphs I: Subgraph Frequences, Metric Properties and Testing, Advances in Math. 219 (2008), 1801-1851.

[5] Borgs, C., Chayes, J. T., Lovász, L., Sós, V. T., Vesztergombi, K., Convergent Sequences of Dense Graphs II: Multiway Cuts and Statistical Physics (2007), http://www.cs.elte.hu/∼lovasz/ConvRight.pdf

[6] Chung, F., Lu, L., Concentration Inequalities and Martingale Inequalities: A Survay, Internet Mathematics Vol. 3, No. I (2005), 79-127.

[7] Frieze, A., Kannan, R., Quick approximation to matrices and applications, Combinatorica 19 (1999), 175-220.

[8] Lovász, L., Szegedy, B., Limits of Dense Graph Sequences, J. Comb. Theory B 96 (2006), 933-957.

[9] Lovász, L., Szegedy, B., Szemerédi's Lemma for the Analyst, preprint (2006), http://research.microsoft.com/users/lovasz/analyst.pdf