MORE ON QUASI-RANDOM GRAPHS, SUBGRAPH COUNTS AND GRAPH LIMITS

SVANTE JANSON AND VERA T. SÓS

Abstract. We study some properties of graphs (or, rather, graph sequences) defined by demanding that the number of subgraphs of a given type, with vertices in subsets of given sizes, approximatively equals the number expected in a random graph. It has been shown by several authors that several such conditions are quasi-random, but that there are exceptions. In order to understand this better, we investigate some new properties of this type. We show that these properties too are quasi-random, at least in some cases; however, there are also cases that are left as open problems, and we discuss why the proofs fail in these cases.

The proofs are based on the theory of graph limits; and on the method and results developed by Janson (2011), this translates the combinatorial problem to an analytic problem, which then is translated to an algebraic problem.

1. Introduction

Consider a sequence of graphs (G_n), with $|G_n| \to \infty$ as $n \to \infty$. Thomason [17; 18] and Chung, Graham and Wilson [4] showed that a number of different 'random-like' properties of the sequence (G_n) are equivalent, and we say that (G_n) is *quasi-random*, or more precisely *p-quasi-random*, if it satisfies these properties. (Here $p \in [0, 1]$ is a parameter.) Many other equivalent properties of different types have later been added by various authors. We say that a property of sequences (G_n) of graphs (with $|G_n| \to \infty$) is a *quasi-random property* (or more specifically a *p-quasi-random property*) if it characterizes quasi-random (or p-quasi-random) sequences of graphs.

One of the quasi-random properties considered by Chung, Graham and Wilson [4] is based on subgraph counts, see (2.2) below. Further quasi-random properties based on restricted subgraph count properties have been found by Chung and Graham [3], Simonovits and Sós [15; 16], Shapira [11], Shapira and Yuster [12; 13], Yuster [19], Janson [6], Huang and Lee [5], see Section 2.

The purpose of the present paper is to continue the study of such properties by considering some further cases not treated earlier; in particular

Date: 26 May, 2014.

2010 *Mathematics Subject Classification.* 05C99.

SJ partly supported by the Knut and Alice Wallenberg Foundation.

VTS partly supported by OTKA 101535.2012.

(Theorems 2.11 and 2.12), we prove that some further properties of this type are quasi-random. Our main purpose is not to just add to the already long list of quasi-random properties; we hope that this study will contribute to the understanding of this type of quasi-random properties, and in particular explain why the case in Theorem 2.12 is more difficult than the one in Theorem 2.11. (See also Section 9 for a discussion of further similar properties.)

We use the method of Janson [6] based on graph limits. We assume that the reader is familiar with the basics of the theory of graph limits and graphons developed in e.g. Lovász and Szegedy [8] and Borgs, Chayes, Lovász, Sós and Vesztergombi [1]; otherwise, see Janson [6] (for the present context) or the comprehensive book by Lovász [7]. As is well-known, there is a simple characterization of quasi-random sequences in terms of graph limits: a sequence (G_n) with $|G_n| \to \infty$ is p-quasi-random if and only if $G_n \to W_p$, where W_p is the graphon that is constant with $W_p = p$ [1; 2; 8], see also [7, Section 1.4.2 and Example 11.37]. (Indeed, quasi-random graphs form one of the roots of graph limit theory.)

The idea of the method is to use this characterization to translate the property of graph sequences to a property of graphons, and then show that only constant graphons satisfy this property. It turns out that this leads to both analytic (Section 4) and algebraic (Section 6) problems, which we find interesting in themselves. We have only partly succeeded to solve these problems, so we leave several open problems.

Remark 1.1. Many of the references above use Szemerédi's regularity lemma as their main tool to study quasi-random properties; it has been known since [14] that quasi-randomness can be characterized using Szemerédi partitions. It is also well-known that there are strong connections between Szemerédi's regularity lemma and graph limits, see [1; 9; 7], so on a deeper level the methods are related although they superficially look very different. (It thus might be possible to translate arguments of one type to the other, although it is far from clear how this might be done.) Both methods lead also to the same (sometimes difficult) algebraic problems. As discussed in [6], the method used here eliminates the many small error terms in the regularity lemma approach; on the other hand, it leads to analytic problems with no direct counterpart in the other approach. It is partly a matter of taste what type of arguments one prefers.

Acknowledgement. This research was begun during the workshop *Graph limits, homomorphisms and structures II* at Hraniční Zámeček, Czech Republic, 2012; parts were also done during the workshop *Combinatorics and Probability* at Mathematisches Forschungsinstitut Oberwolfach, 2013. We thank the organisers for providing us with these opportunities.

2. Notation, background and main results

All graphs in this paper are finite, undirected and simple. The vertex and edge sets of a graph G are denoted by $V(G)$ and $E(G)$. We write $|G| := |V(G)|$ for the number of vertices of G, and $e(G) := |E(G)|$ for the number of edges. As usual, $[n] := \{1, \dots, n\}$.

All unspecified limits in this paper are as $n \to \infty$, and $o(1)$ denotes a quantity that tends to 0 as $n \to \infty$. We will often use $o(1)$ for quantities that depend on some subset(s) of a vertex set $V(G)$; we then always implicitly assume that the convergence is uniform for all choices of the subsets. We interpret $o(a_n)$ for a given sequence a_n similarly.

Let F and G be labelled graphs. For convenience, we assume throughout the paper (when it matters) that $V(F) = [|F|] = \{1, \dots, |F|\}$. We generally let $m = |F|$.

Definition 2.1. (i) $N(F, G)$ is the number of labelled copies of F in G (not necessarily induced); equivalently, $N(F, G)$ is the number of injective maps $\varphi : V(F) \to V(G)$ that are graph homomorphisms (i.e., if i and j are adjacent in F, then $\varphi(i)$ and $\varphi(j)$ are adjacent in G).

(ii) If $U_1, \dots, U_{|F|}$ are subsets of $V(G)$, let $N(F, G; U_1, \dots, U_{|F|})$ be the number of labelled copies of F in G with the ith vertex in U_i; equivalently, $N(F, G; U_1, \dots, U_{|F|})$ is the number of injective graph homomorphisms $\varphi : F \to G$ such that $\varphi(i) \in U_i$ for every $i \in V(F)$. (Note that we consider a fixed labelling of the vertices of F and count the number of copies where vertex i is in U_i, so the labelling and the ordering of $U_1, \dots, U_{|F|}$ are important.)

(iii) We also define a symmetrized version $\widetilde{N}(F, G; U_1, \dots, U_{|F|})$ by taking the average over all labellings of F; equivalently,

$$\widetilde{N}(F, G; U_1, \dots, U_{|F|}) := \frac{1}{|F|!} \sum_{\sigma} N(F, G; U_{\sigma(1)}, \dots, U_{\sigma(|F|)}), \tag{2.1}$$

summing over all permutations σ of $\{1, \dots, |F|\}$.

In (ii) and (iii), we are often interested in the case when $U_1, \dots, U_{|F|}$ are pairwise disjoint, and then $\widetilde{N}(F, G; U_1, \dots, U_{|F|})$ is the number of labelled copies of F in G with one vertex in each set U_i (in any order), divided by $1/|F|!$.

Remark 2.2. If either $U_1 = \cdots = U_{|F|}$ or $F = K_m$ for some m, then $\widetilde{N}(F, G; U_1, \dots, U_{|F|}) := N(F, G; U_1, \dots, U_{|F|})$, and the symmetrized version $\widetilde{N}$ is equal to N.

One of the several equivalent definitions of quasi-random graphs by Chung, Graham and Wilson [4] is the following using the subgraph counts $N(F, G)$:

Theorem 2.3 (Chung, Graham and Wilson [4]). *A sequence of graphs (G_n) with $|G_n| \to \infty$ is p-quasi-random if and only if, for every graph F,*

$$N(F, G_n) = (p^{e(F)} + o(1))|G_n|^{|F|}. \tag{2.2}$$

□

It is not necessary to require (2.2) for all graphs F; in particular, it suffices to use the graphs K_2 and C_4 [4]. However, it is not enough to require (2.2) for just one graph F. As a substitute, Simonovits and Sós [15] considered the hereditary version of (2.2), i.e. the condition $N(F, G; U, \dots, U)$ for subsets U.

We note first that for quasi-random graphs, it is shown in [15] and [11] that the restricted subgraph count $N(F, G; U_1, \dots, U_{|F|})$ is asymptotically the same as it is for random graphs, for any subsets $U_1, \dots, U_{|F|}$. (For a proof using graph limits, see Janson [6, Lemma 4.2].)

Lemma 2.4 ([15] and [11])**.** *Suppose that (G_n) is a p-quasi-random sequence of graphs, where $0 \leqslant p \leqslant 1$, and let F be any fixed graph with $e(F) > 0$. Then, for all subsets $U_1, \dots, U_{|F|}$ of $V(G_n)$,*

$$N(F, G_n; U_1, \dots, U_{|F|}) = p^{e(F)} \prod_{i=1}^{|F|} |U_i| + o\big(|G_n|^{|F|}\big). \tag{2.3}$$

and

$$\widetilde{N}(F, G_n; U_1, \dots, U_{|F|}) = p^{e(F)} \prod_{i=1}^{|F|} |U_i| + o\big(|G_n|^{|F|}\big). \tag{2.4}$$

□

Note that (2.4) is an immediate consequence of (2.3) by the definition (2.1).

Conversely, Simonovits and Sós [15] showed that (2.3) implies that (G_n) is p-quasi-random. Actually, they considered only the symmetric case $U_1 = \dots = U_{|F|}$ and proved the following stronger result. (In this case, (2.4) is obviously equivalent to (2.3), see Remark 2.2.)

Theorem 2.5 (Simonovits and Sós [15])**.** *Suppose that (G_n) is a sequence of graphs with $|G_n| \to \infty$. Let F be any fixed graph with $e(F) > 0$ and let $0 < p \leqslant 1$. Then (G_n) is p-quasi-random if and only if, for all subsets U of $V(G_n)$, (2.3) holds with $U_1 = \dots = U_{|F|} = U$.* □

Remark 2.6. The case $F = K_2$, when $N(K_2, G_n; U)$ is twice the number of edges with both endpoints in U, is one of the original quasi-random properties in Chung, Graham and Wilson [4].

Remark 2.7. Theorem 2.5 obviously fails when $e(F) = 0$, since then (2.3) holds trivially for any G_n. It fails also if $p = 0$; for example, if $F = K_3$ and G_n is the complete bipartite graph $K_{n,n}$.

In other words, Theorem 2.5 says that, if $e(F) > 0$ and $0 < p \leqslant 1$, then (2.3) and (2.4) (for arbitrary $U_1, \dots, U_{|F|}$) are both p-quasi-random properties, and this holds also if we restrict $U_1, \dots, U_{|F|}$ to $U_1 = \dots = U_{|F|}$.

Several authors have considered other restrictions on $U_1, \dots, U_{|F|}$ and shown that (2.3) or (2.4) still is a quasi-random property.

Shapira [11] and Yuster [19] continued to consider $U_1 = \cdots = U_{|F|}$, and assumed further that $|U_1| = \lfloor \alpha |G_n| \rfloor$ for some fixed α with $0 < \alpha < 1$; they showed ([11] for $\alpha = 1/(|F|+1)$ and [19] in general) that (2.3) for such $U_1, \dots, U_{|F|}$ is a quasi-random property. (The case $F = K_2$ and $\alpha = 1/2$ is in Chung, Graham and Wilson [4].) Note that for such $U_1, \dots, U_{|F|}$, (2.4) is equivalent to (2.3) by Remark 2.2.

The case when $U_1, \dots, U_{|F|}$ are disjoint and furthermore have the same size is considered by Shapira [11] and Shapira and Yuster [12]; they show that (2.4) with this restriction also is a quasi-random property. (As a consequence, (2.3) with this restriction is a quasi-random property.) Moreover, by combining Shapira [11, Lemma 2.2] and the result of Yuster [19] just mentioned, it follows that it suffices to consider disjoint $U_1, \dots, U_{|F|}$ with the same size $\lfloor \alpha |G_n| \rfloor$, for any fixed $\alpha < 1/|F|$.

We introduce some more notation.

Definition 2.8. Let F be a graph, $m := |F|$ and $(\alpha_1, \dots, \alpha_m)$ a vector of positive numbers with $\sum_{i=1}^m \alpha_i \leqslant 1$; let further $p \in [0,1]$. We define the following properties of graph sequences (G_n). (For convenience, we omit p from the notations.)

(i) Let F be labelled. Then $\mathcal{P}(F; \alpha_1, \dots, \alpha_m)$ is the property that (2.3) holds for all disjoint subsets $U_1, \dots, U_m$ of $V(G_n)$ with $|U_i| = \lfloor \alpha_i |G_n| \rfloor$, $i = 1, \dots, m$.

(ii) Let F be unlabelled. Then $\mathcal{P}'(F; \alpha_1, \dots, \alpha_m)$ is the property that $\mathcal{P}(F; \alpha_1, \dots, \alpha_m)$ holds for every labelling of F.

(iii) Let F be unlabelled. Then $\widetilde{\mathcal{P}}(F; \alpha_1, \dots, \alpha_m)$ is the property that (2.4) holds for all $U_1, \dots, U_m$ as in (i).

Of course, we can use $\mathcal{P}'$ and $\widetilde{\mathcal{P}}$ also for a labelled F by ignoring the labelling.

Remark 2.9. If $F = K_m$, then all labellings of F are equivalent, and the three properties $\mathcal{P}(F; \alpha_1, \dots, \alpha_m)$, $\mathcal{P}'(F; \alpha_1, \dots, \alpha_m)$ and $\widetilde{\mathcal{P}}(F; \alpha_1, \dots, \alpha_m)$ are equivalent. In general, $\mathcal{P}'(F; \alpha_1, \dots, \alpha_m) \implies \widetilde{\mathcal{P}}(F; \alpha_1, \dots, \alpha_m)$ by the definition of $\widetilde{N}$ as an average of N over all labellings of F, but we do not know whether the converse implication always holds.

Furthermore, for a fixed labelling of F, $\mathcal{P}'(F; \alpha_1, \dots, \alpha_m)$ is equivalent to the conjunction of $\mathcal{P}(F; \alpha_{\sigma(1)}, \dots, \alpha_{\sigma(m)})$ for all permutations $(\alpha_{\sigma(1)}, \dots, \alpha_{\sigma(m)})$ of $(\alpha_1, \dots, \alpha_m)$. In particular, if $\alpha_1 = \cdots = \alpha_m$, then $\mathcal{P}'(F; \alpha_1, \dots, \alpha_m)$ equals $\mathcal{P}(F; \alpha_1, \dots, \alpha_m)$, for any labelling.

In general, trivially $\mathcal{P}'(F; \alpha_1, \dots, \alpha_m) \implies \mathcal{P}(F; \alpha_1, \dots, \alpha_m)$ for a labelled graph F, but we do not know whether the converse holds. Nor do we know any general implications between $\mathcal{P}(F; \alpha_1, \dots, \alpha_m)$ and $\widetilde{\mathcal{P}}(F; \alpha_1, \dots, \alpha_m)$.

See further Remark 2.14.

Using this notation, it thus follows from Shapira [11] and Yuster [19] that, for any graph F with $e(F) > 0$ and $0 < p \leqslant 1$, $\widetilde{\mathcal{P}}(F;\alpha,\dots,\alpha)$ is a quasi-random property for every $\alpha < 1/|F|$. This can also be proved by the methods of Janson [6], where the somewhat weaker statement that $\mathcal{P}(F;\alpha,\dots,\alpha)$ is a quasi-random property for every $\alpha < 1/|F|$ is shown [6, Theorem 3.6]. We show here a more general statement in Theorem 2.11 below.

Example 2.10. For $F = K_2$, $\widetilde{\mathcal{P}}(K_2,\alpha_1,\alpha_2) = \mathcal{P}(K_2,\alpha_1,\alpha_2)$ says that (asymptotically) the number of edges $e(U_1,U_2)$ is as expected in $G(n,p)$ for any two disjoint sets U_1,U_2 with $U_i = \lfloor \alpha_i |G_n| \rfloor$. Chung and Graham [3] showed that the cut property $\mathcal{P}(K_2;\alpha,1-\alpha)$ is a quasi-random property for every fixed $\alpha \in (0,1)$ except $\alpha = 1/2$, when it is not; see further Janson [6, Section 9]. Simonovits and Sós [15] showed that $\mathcal{P}(K_2,1/3,1/3)$ is a quasi-random property.

Shapira and Yuster [13, Proposition 14] showed (as a consequence of related results for cuts in hypergraphs) that $\mathcal{P}(K_m,\alpha_1,\dots,\alpha_m)$ is a quasi-random property, for every $m \geqslant 2$ and $(\alpha_1,\dots,\alpha_m) \neq (1/m,\dots,1/m)$ with $\sum_{i=1}^m \alpha_i = 1$. This can easily be extended to subgraph counts for arbitrary graphs F with $e(F) > 0$; we give a proof using our methods in Section 6.

Theorem 2.11. *Let F be a graph with $e(F) > 0$, and let $0 < p \leqslant 1$. Further, let $(\alpha_1,\dots,\alpha_m)$ be a vector of positive numbers of length $m = |F|$ with $\sum_{i=1}^m \alpha_i \leqslant 1$.*

(i) *If $(\alpha_1,\dots,\alpha_m) \neq (1/m,\dots,1/m)$, then $\widetilde{\mathcal{P}}(F;\alpha_1,\dots,\alpha_m)$ and the stronger $\mathcal{P}'(F;\alpha_1,\dots,\alpha_m)$ are quasi-random properties.*
(ii) *If $\sum_{i=1}^m \alpha_i < 1$, then $\mathcal{P}(F;\alpha_1,\dots,\alpha_m)$ is a quasi-random property.*

The exceptional case $\alpha_1 = \dots = \alpha_m = 1/m$ is more complicated; Shapira and Yuster [13] showed that the related hypergraph cut property used by them to prove Theorem 2.11 fails in this case; nevertheless, Huang and Lee [5] showed that also $\mathcal{P}(K_m,1/m,\dots,1/m)$ is a quasi-random property for any $m \geqslant 3$. (For $m = 2$ it is not, see Example 2.10.)

We give a new proof of their theorem in Section 7 and extend the result to counts of several other subgraphs. With our methods using graph limits, the crucial fact is that while the central analytic Lemma 4.1 does not generalize to the case $(\alpha_1,\dots,\alpha_m) = (1/m,\dots,1/m)$, there is a weaker version Lemma 4.3 that holds in this case, and this is sufficient to draw the conclusion with some extra algebraic work. We have so far not succeded to extend the final, algebraic, part to all graphs F, but we can prove the following, see Section 7. (Section 7 contains also some further examples of small graphs F for which the conclusion holds.)

Theorem 2.12. *Let F be a graph with $e(F) > 1$ and $m = |F|$. Let also $0 < p \leqslant 1$. If F is either a regular graph or a star, or disconnected, then $\mathcal{P}(F;1/m,\dots,1/m)$ and the weaker $\widetilde{\mathcal{P}}(F;1/m,\dots,1/m)$ are quasi-random properties.*

One indication that this theorem is more complicated than Theorem 2.11 is that the conclusion is false for $F = K_2$ by Example 2.10, and slightly more generally when $e(F) \leqslant 1$. We conjecture that this is the only counterexample.

Conjecture 2.13. *Theorem 2.12 holds for any graph F with $e(F) > 1$.*

Remark 2.14. When $F \neq K_m$, the relation between the properties $\mathcal{P}$ (non-averaged) and $\widetilde{\mathcal{P}}$ (averaged) is not completely clear. (For $F = K_m$, these properties coincide, see Remark 2.9.)

Consider first $\alpha_1 = \cdots = \alpha_m = 1/m$ as in Theorem 2.12. Then $\mathcal{P} = \mathcal{P}' \implies \widetilde{\mathcal{P}}$. (See Remark 2.9 again.) For a graph F such that Theorem 2.12 applies, the theorem implies that the properties are equivalent, but as said above, we do not know whether that holds in general. In principle, it should be easier to show that the property $\mathcal{P}(F; 1/m, \dots, 1/m)$ is p-quasi-random than to show that the weaker (averaged) property $\widetilde{\mathcal{P}}(F; 1/m, \dots, 1/m)$ is; it is even conceivable that there exists a counterexample to Conjecture 2.13 such that nevertheless $\mathcal{P}(F; 1/m, \dots, 1/m)$ is p-quasi-random. However, our method of proof uses Lemma 4.3 below which assumes that the function f there is symmetric, and hence our proofs use the symmetric $\widetilde{\mathcal{P}}(F; 1/m, \dots, 1/m)$ and we are not able to use the extra power of $\mathcal{P}(F; 1/m, \dots, 1/m)$. For example, we cannot answer the following question. (Cf. Section 5 for $\widetilde{\mathcal{P}}(F; 1/m, \dots, 1/m)$.) A *2-type graphon* is a graphon that is constant on the sets $S_i \times S_j$, $i, j \in \{1, 2\}$, for some partition $[0, 1] = S_1 \cup S_2$ into two disjoint sets; we can without loss of generality assume that the sets S_i are intervals. (Equivalently, we may regard W as a graphon defined on a two-point probability space.)

Problem 2.15. If F is such that $\mathcal{P}(F; 1/m, \dots, 1/m)$ is not p-quasi-random, is there always a 2-type graphon witnessing this?

For other sequences $\alpha_1, \dots, \alpha_m$, we note first that if $\sum_{i=1}^m \alpha_i < 1$, then Theorem 2.11 shows that both $\mathcal{P}$ and $\widetilde{\mathcal{P}}$ are quasi-random properties, and thus equivalent. Similarly, if $\sum_{i=1}^m \alpha_i = 1$ but $(\alpha_1, \dots, \alpha_m) \neq (1/m, \dots, 1/m)$, then $\widetilde{\mathcal{P}}$ is quasi-random by Theorem 2.11, and thus $\widetilde{\mathcal{P}} \implies \mathcal{P}$. However, we do not know whether the converse holds:

Problem 2.16. Suppose that F is a labelled graph with $e(F) > 0$, that $0 < p \leqslant 1$ and that $\sum_{i=1}^m \alpha_i = 1$ but $(\alpha_1, \dots, \alpha_m) \neq (1/m, \dots, 1/m)$. Is then $\mathcal{P}(F; \alpha_1, \dots, \alpha_m)$ a quasi-random property?

If there is any case such that the answer to this problem is negative, we can ask the same question as in Problem 2.15:

Problem 2.17. If F and $(\alpha_1, \dots, \alpha_m)$ are such that $\mathcal{P}(F; \alpha_1, \dots, \alpha_m)$ is not p-quasi-random, is there always a 2-type graphon witnessing this?

Example 2.18. Let $F = P_3 = K_{1,2}$, for definiteness labelled with edges 12 and 13, and consider the property $\mathcal{P}(F; \alpha_1, \alpha_2, \alpha_3)$. If $\alpha_1 + \alpha_2 + \alpha_3 < 1$, then the property is quasi-random by Theorem 2.11; thus assume $\alpha_1+\alpha_2+\alpha_3=1$. In the case $\alpha_1=\alpha_2=\alpha_3=1/3$, the property is quasi-random by Theorem 2.12. We can show this also in the case $\alpha_2\neq\alpha_3$, using the symmetry of P_3, see Remark 6.1. However, we do not know if this extends to $\alpha_2=\alpha_3$, for example in the following case:

Problem 2.19. Is (with the labelling above) $\mathcal{P}(P_3,\frac12,\frac14,\frac14)$ a quasi-random property?

Remark 2.20. We have considered the subgraph counts $N(F,G_n;U_1,\dots,U_m)$ and $\widetilde{N}(F,G_n;U_1,\dots,U_m)$ in two cases: either $U_1=\dots=U_m$ (as in [15]) or $U_1,\dots,U_m$ are disjoint. It also seems interesting to consider other, intermediate, cases of restrictions. This is discussed in Section 9, where we in particular consider, as a typical example, the case $U_1=U_2$ and $U_1\cap U_3=\emptyset$.

Remark 2.21. We consider in this paper not necessarily induced copies of a fixed graph F. There are also similar results for counts of induced copies of F, but these are more complicated and less complete, see Simonovits and Sós [16], Shapira and Yuster [12] and Janson [6]. We hope to return to the induced case, but leave it for now as an open problem:

Problem 2.22. Are there analogues of Theorems 2.11 and 2.12 for the induced case?

3. Transfer to graph limits

We introduce some further notation:

The support of a function ψ is the set $\operatorname{supp}(\psi):=\{x:\psi(x)\neq 0\}$.

λ denotes Lebesgue measure.

All functions are supposed to be (Lebesgue) measurable.

If F is a labelled graph and W a graphon, we define

$$\Psi_{F,W}(x_1,\dots,x_{|F|}):=\prod_{ij\in E(F)}W(x_i,x_j). \tag{3.1}$$

If f is a function on $[0,1]^m$ for some m, we let $\tilde f$ denote its symmetrization defined by

$$\tilde f(x_1,\dots,x_m):=\frac{1}{m!}\sum_{\sigma\in\mathfrak{S}_m}f\bigl(x_{\sigma(1)},\dots,x_{\sigma(m)}\bigr), \tag{3.2}$$

where $\mathfrak{S}_m$ is the symmetric group of all $m!$ permutations of $\{1,\dots,m\}$.

The connection between the subgraph count properties and properties of graph limits is given by the following lemma.

Lemma 3.1. *Suppose that $G_n\to W$ for some graphon W. Let F be a fixed graph, let $m:=|F|$ and let $\gamma\geqslant 0$ and $\alpha_1,\dots,\alpha_m\in(0,1)$ be fixed numbers with $\sum_{i=1}^m\alpha_i\leqslant 1$. Then the following are equivalent:*

(i) *For all disjoint subsets* $U_1, \dots, U_{|F|}$ *of* $V(G_n)$ *with* $|U_i| = \lfloor \alpha_i |G_n| \rfloor$,

$$N(F, G_n; U_1, \dots, U_{|F|}) = \gamma \prod_{i=1}^{|F|} |U_i| + o\big(|G_n|^{|F|}\big). \tag{3.3}$$

(ii) *For all disjoint subsets* $A_1, \dots, A_{|F|}$ *of* $[0,1]$ *with* $\lambda(A_i) = \alpha_i$,

$$\int_{A_1 \times \dots \times A_{|F|}} \Psi_{F,W}(x_1, \dots, x_{|F|}) = \gamma \prod_{i=1}^{|F|} \lambda(A_i). \tag{3.4}$$

The same holds if we replace N *in* (i) *and* $\Psi_{F,W}$ *in* (ii) *by the symmetrized versions* $\widetilde{N}$ *and* $\widetilde{\Psi}_{F,W}$.

Proof. The case with N and $\Psi_{F,W}$ and with $\alpha_1 = \dots = \alpha_m < 1/|F|$ is part of Janson [6, Lemma 7.2]. The case of general $\alpha_1, \dots, \alpha_m$, and the symmetrized version with $\widetilde{N}$ and $\widetilde{\Psi}_{F,W}$ are proved in exactly the same way. ☐

With this lemma in mind, we make the following definitions corresponding to Definition 2.8.

Definition 3.2. Let, as in Definition 2.8, F be a graph, $m := |F|$, $(\alpha_1, \dots, \alpha_m)$ a vector of positive numbers with $\sum_{i=1}^m \alpha_i \leqslant 1$, and $p \in [0,1]$. We define the following properties of graphons W.

(i) $\mathcal{P}_*(F; \alpha_1, \dots, \alpha_m)$ is the property that

$$\int_{A_1 \times \dots \times A_m} \Psi_{F,W}(x_1, \dots, x_m) = p^{e(F)} \prod_{i=1}^{m} \lambda(A_i), \tag{3.5}$$

for all disjoint subsets $A_1, \dots, A_m$ of $[0,1]$ with $\lambda(A_i) = \alpha_i$, $i = 1, \dots, m$.

(ii) $\mathcal{P}'_*(F; \alpha_1, \dots, \alpha_m)$ is the property that $\mathcal{P}_*(F; \alpha_1, \dots, \alpha_m)$ holds for every labelling of F.

(iii) $\widetilde{\mathcal{P}}_*(F; \alpha_1, \dots, \alpha_m)$ is the property that

$$\int_{A_1 \times \dots \times A_m} \widetilde{\Psi}_{F,W}(x_1, \dots, x_m) = p^{e(F)} \prod_{i=1}^{m} \lambda(A_i) \tag{3.6}$$

for all $A_1, \dots, A_m$ as in (i).

Definition 3.3. A property of graphons W is *quasi-random* if every graphon W that satisfies it is a.e. equal to a constant. Furthermore, the property is *p-quasi-random* if it is satisfied only by graphons W that are a.e. equal to p.

We can now use standard arguments to translate our problem from graph sequences to graphons. Recall that $m := |F|$.

Lemma 3.4. *For any given graph* F, $p \in [0,1]$ *and* $\alpha_1, \dots, \alpha_m \in (0,1)$ *with* $\sum_{i=1}^m \alpha_i \leqslant 1$, *the property* $\mathcal{P}(F; \alpha_1, \dots, \alpha_m)$ *(of graph sequences) is*

p-quasi-random if and only if the property $\mathcal{P}_*(F;\alpha_1,\dots,\alpha_m)$ (of graphons) is.

Similarly, the property $\mathcal{P}'(F;\alpha_1,\dots,\alpha_m)$ is p-quasi-random if and only if the property $\mathcal{P}'_*(F;\alpha_1,\dots,\alpha_m)$ is, and $\widetilde{\mathcal{P}}(F;\alpha_1,\dots,\alpha_m)$ is p-quasi-random if and only if $\widetilde{\mathcal{P}}_*(F;\alpha_1,\dots,\alpha_m)$ is.

Proof. Suppose that $\mathcal{P}(F;\alpha_1,\dots,\alpha_m)$ is p-quasi-random, and let W be a graphon satisfying $\mathcal{P}_*(F;\alpha_1,\dots,\alpha_m)$. Let (G_n) be any sequence of graphs converging to W. By assumption, Lemma 3.1(ii) holds with $\gamma = p^{e(F)}$, and thus Lemma 3.1 shows that (3.3) holds for all disjoint $U_1,\dots,U_m$ with $|U_i| = \lfloor \alpha_i |G_n| \rfloor$. In other words, (G_n) satisfies the property $\mathcal{P}(F;\alpha_1,\dots,\alpha_m)$, and since this property was assumed to be p-quasi-random, the sequence (G_n) is p-quasi-random, and thus $G_n \to W_p$, where $W_p = p$ everywhere. Since $G_n \to W$, this implies $W = W_p = p$ a.e.

Conversely, suppose that $\mathcal{P}_*(F;\alpha_1,\dots,\alpha_m)$ is p-quasi-random, and let (G_n) be a graph sequence satisfying $\mathcal{P}(F;\alpha_1,\dots,\alpha_m)$. This means that Lemma 3.1(i) holds with $\gamma = p^{e(F)}$. Consider a subsequence of (G_n) that converges to some graphon W. Lemma 3.1 then shows that (3.4) holds for all disjoint $A_1,\dots,A_m$ with $\lambda(A_i) = \alpha_i$. In other words, W satisfies the property $\mathcal{P}_*(F;\alpha_1,\dots,\alpha_m)$, and since this property was assumed to be p-quasi-random, $W = p$ a.e. Consequently, every convergent subsequence of (G_n) converges to the constant graphon $W_p = p$. Since every subsequence has convergent subsubsequences, it follows that the full sequence (G_n) converges to W_p, i.e., (G_n) is p-quasi-random.

The same proof works for $\widetilde{\mathcal{P}}(F;\alpha_1,\dots,\alpha_m)$ and $\widetilde{\mathcal{P}}_*(F;\alpha_1,\dots,\alpha_m)$. $\square$

In the rest of the paper we analyze the graphon properties $\mathcal{P}_*(F;\alpha_1,\dots,\alpha_m)$ and $\widetilde{\mathcal{P}}_*(F;\alpha_1,\dots,\alpha_m)$.

4. The analytic part

Janson [6] proved the following lemma:

Lemma 4.1 ([6, Lemma 7.3])**.** *Let $m \geqslant 1$ and $\alpha \in (0,1)$. Suppose that f is an integrable function on $[0,1]^m$ such that $\int_{A_1\times\cdots\times A_m} f = 0$ for all sequences $A_1,\dots,A_m$ of measurable subsets of $[0,1]$ such that $\lambda(A_1) = \cdots = \lambda(A_m) = \alpha$. Then $f = 0$ a.e.*

Moreover, if $\alpha < m^{-1}$, it is enough to consider disjoint $A_1,\dots,A_m$.

It was remarked in [6, Remark 7.4] that the second part (disjoint subsets) of this lemma fails when $\alpha = 1/m$, i.e., when we consider partitions of $[0,1]$ into m disjoint sets of equal measure $1/m$ (we call these *equipartitions*); a simple counterexample is provided by the following lemma.

Lemma 4.2. *Let $m \geqslant 1$. Suppose that*

$$f(x_1,\dots,x_m) = g(x_1) + \cdots + g(x_m) \tag{4.1}$$

for some integrable function g on $[0,1]$ with $\int_0^1 g = 0$. Then f is a symmetric integrable function on $[0,1]^m$ and

$$\int_{A_1\times\cdots\times A_m} f = 0 \tag{4.2}$$

for all partitions $\{A_1,\dots,A_m\}$ of $[0,1]$ into m disjoint measurable subsets such that $\lambda(A_1)=\cdots=\lambda(A_m)=1/m$. □

Proof. If $\{A_1,\dots,A_m\}$ is an equipartition of $[0,1]$, then

$$\begin{aligned}\int_{A_1\times\cdots\times A_m} f(x_1,\dots,x_m) &= \sum_{i=1}^m \Big(\frac{1}{m}\Big)^{m-1}\int_{A_i} g(x_i)\,\mathrm{d}x_i\\ &= m^{1-m}\sum_{i=1}^m\int_{A_i} g(x)\,\mathrm{d}x = m^{1-m}\int_0^1 g(x)\,\mathrm{d}x = 0. \quad (4.3)\end{aligned}$$

□

Moreover, it was shown in [6, Proof of Lemma 9.4 and the comments after it], see also [6, Lemma 10.3], that if $m=2$ and f is symmetric with $\int_{A_1\times A_2} f = 0$ for every equipartition $\{A_1,A_2\}$, then f has to be of the form (4.1) a.e. We shall here extend this to any m, thus showing that the converse to Lemma 4.2 holds.

Lemma 4.3. *Let $m \geqslant 1$. Suppose that $f : [0,1]^m \to \mathbb{C}$ is a symmetric integrable function such that*

$$\int_{A_1\times\cdots\times A_m} f = 0 \tag{4.4}$$

for all partitions $\{A_1,\dots,A_m\}$ of $[0,1]$ into m disjoint measurable subsets such that $\lambda(A_1)=\cdots=\lambda(A_m)=1/m$. Then

$$f(x_1,\dots,x_m) = g(x_1)+\cdots+g(x_m) \qquad \textit{a.e.} \tag{4.5}$$

for some integrable function g on $[0,1]$ with $\int_0^1 g = 0$.

Proof. The lemma is trivial when $m=1$. The case $m=2$ is, as said above, proved in [6], but for completeness, we repeat the argument:

Let $f_1(x) := \int_0^1 f(x,y)\,\mathrm{d}y$. Then, for every subset $A\subset[0,1]$ with $\lambda(A) = 1/2$, (4.4) with $A_1 := A$ and $A_2 := [0,1]\setminus A$ yields

$$\begin{aligned}0 &= \int_{A_1\times A_2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{A\times[0,1]} f(x,y)\,\mathrm{d}x\,\mathrm{d}y - \int_{A\times A} f(x,y)\,\mathrm{d}x\,\mathrm{d}y\\ &= \int_A f_1(x)\,\mathrm{d}x - \int_{A\times A} f(x,y)\,\mathrm{d}x\,\mathrm{d}y\\ &= \int_{A\times A}\big(f_1(x)+f_1(y)-f(x,y)\big)\,\mathrm{d}x\,\mathrm{d}y.\end{aligned} \tag{4.6}$$

The integrand in the last integral is symmetric, and it follows by [6, Lemma 7.6] that it vanishes a.e., which proves (4.5) with $g = f_1$; moreover, arguing as in (4.3), for any equipartition $\{A_1, A_2\}$ of [0,1],

$$0 = \int_{A_1 \times A_2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \frac{1}{2}\int_0^1 g(x)\,\mathrm{d}x, \tag{4.7}$$

and thus $\int_0^1 g = 0$, completing the proof when $m = 2$.

Thus suppose in the remainder of the proof that $m \geqslant 3$.

Step 1: Fix a subset $B \subset [0,1]$ with measure $\lambda(B) = 2/m$, and fix an equipartition of the complement $[0,1] \setminus B$ into $m-2$ sets $A_3, \dots, A_m$ of equal measure $1/m$. Let

$$f_2(x_1, x_2) := \int_{A_3 \times \dots \times A_m} f(x_1, x_2, \dots, x_m)\,\mathrm{d}x_3 \cdots \mathrm{d}x_m. \tag{4.8}$$

Then the assumption (4.4) says that for any equipartition $B = A_1 \cup A_2$ of B into two disjoint subsets of equal measure,

$$\int_{A_1 \times A_2} f_2(x_1, x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 = 0. \tag{4.9}$$

The set B is, as a measure space, isomorphic to $[0, 2/m]$, and by a trivial rescaling, the case $m = 2$ shows that there exists an integrable function h on B with $\int_B h = 0$ such that

$$f_2(x_1, x_2) = h(x_1) + h(x_2), \qquad \text{a.e. } x_1, x_2 \in B. \tag{4.10}$$

This means that if ψ_1 and ψ_2 are bounded functions on $[0,1]$ such that $\int_0^1 \psi_1 = \int_0^1 \psi_2 = 0$ and $\operatorname{supp}(\psi_1) \cup \operatorname{supp}(\psi_2) \subseteq B$, then

$$\begin{aligned}
&\int_{[0,1]^m} f(x_1, \dots, x_m)\psi_1(x_1)\psi_2(x_2)\mathbf{1}_{A_3}(x_3)\cdots\mathbf{1}_{A_m}(x_m)\,\mathrm{d}x_1\cdots\mathrm{d}x_m \\
&\quad= \int_{B \times B} f_2(x_1, x_2)\psi_1(x_1)\psi_2(x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 \\
&\quad= \int_B h(x_1)\psi_1(x_1)\,\mathrm{d}x_1 \int_B \psi_2(x_2)\,\mathrm{d}x_2 + \int_B \psi_1(x_1)\,\mathrm{d}x_1 \int_B h(x_2)\psi_2(x_2)\,\mathrm{d}x_2 \\
&\quad= 0.
\end{aligned} \tag{4.11}$$

Step 2: Let us instead start with two bounded functions ψ_1 and ψ_2 on $[0,1]$ such that $\int_0^1 \psi_1 = \int_0^1 \psi_2 = 0$, and assume that $\lambda(\operatorname{supp}(\psi_1)) + \lambda(\operatorname{supp}(\psi_2)) < 2/m$. Let $B_0 := \operatorname{supp}(\psi_1) \cup \operatorname{supp}(\psi_2)$ and $B_0^{\mathsf{c}} := [0,1] \setminus B_0$. Then $\lambda(B_0) < 2/m$ and $\lambda(B_0^{\mathsf{c}}) = 1 - \lambda(B_0) > (m-2)/m$.

Define

$$f_3(x_3, \dots, x_m) := \int_{[0,1]^2} f(x_1, x_2, \dots, x_m)\psi_1(x_1)\psi_2(x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2. \tag{4.12}$$

For any disjoint sets $A_3, \dots, A_m \subset B_0^{\mathsf{c}}$ with $\lambda(A_3) = \cdots = \lambda(A_m) = 1/m$, we can use Step 1 with $B := [0,1] \setminus \bigcup_3^m A_i \supset B_0$ and conclude by (4.11) that

$$\int_{A_3 \times \cdots \times A_m} f_3(x_3, \dots, x_m) = 0. \tag{4.13}$$

The set B_0^{c} is, as a measure space up to a trivial rescaling of the measure, isomorphic to $[0,1]$. Since $\lambda(B_0^{\mathsf{c}}) > (m-2)/m$, it follows by the second part of Lemma 4.1 that (4.13) (for arbitrary $A_3, \dots, A_m$ as above) implies

$$f_3(x_3, \dots, x_m) = 0, \qquad \text{a.e. } x_3, \dots, x_m \in B_0^{\mathsf{c}}. \tag{4.14}$$

Step 3: Fix bounded functions $\varphi_3, \dots, \varphi_m$ on $[0,1]$. For $B \subseteq [0,1]$, define

$$f_B(x_1, x_2) := \int_{(B^{\mathsf{c}})^{m-2}} f(x_1, \dots, x_m) \varphi_3(x_3) \cdots \varphi_m(x_m). \tag{4.15}$$

If $\lambda(B) > 0$ and ψ_1 and ψ_2 are bounded functions with $\operatorname{supp}(\psi_\nu) \subseteq B$, $\lambda(\operatorname{supp}(\psi_\nu)) < 1/m$ and $\int_0^1 \psi_\nu = 0$, $\nu = 1, 2$, then Step 2 shows, using (4.15), (4.12) and (4.14), since $B_0 \subseteq B$ and thus $B^{\mathsf{c}} \subseteq B_0^{\mathsf{c}}$,

$$\begin{aligned} \int_{[0,1]^2} f_B(x_1, x_2) \psi_1(x_1) \psi_2(x_2) &= \int_{(B^{\mathsf{c}})^{m-2}} f_3(x_1, \dots, x_m) \varphi_3(x_3) \cdots \varphi_m(x_m) \\ &= 0. \end{aligned} \tag{4.16}$$

Now suppose that B is open, and $x_1, x_1', x_2, x_2' \in B$. For small enough $\varepsilon > 0$, the functions

$$\psi_\nu(x) := \frac{1}{2\varepsilon} \big(\mathbf{1}_{(x_\nu - \varepsilon, x_\nu + \varepsilon)}(x) - \mathbf{1}_{(x_\nu' - \varepsilon, x_\nu' + \varepsilon)}(x) \big), \qquad \nu = 1, 2, \tag{4.17}$$

satisfy the conditions above and thus (4.16) holds. Letting $\varepsilon \to 0$, it follows that if $(x_1, x_2), (x_1, x_2'), (x_1', x_2), (x_1', x_2')$ are Lebesgue points of f_B, then

$$f_B(x_1, x_2) - f_B(x_1, x_2') - f_B(x_1', x_2) + f_B(x_1', x_2') = 0. \tag{4.18}$$

Thus, (4.18) holds for a.e. $x_1, x_1', x_2, x_2' \in B$.

Consider now the countable collection $\mathcal{B}$ of sets $B \subset (0,1)$ that are unions of four open intervals with rational endpoints. It follows that for a.e. $x_1, x_1', x_2, x_2' \in [0,1]$, (4.18) holds for every set $B \in \mathcal{B}$ such that $x_1, x_1', x_2, x_2' \in B$.

Consider such a 4-tuple $x_1, x_1', x_2, x_2' \in [0,1]$. There exists a decreasing sequence B_n of sets in $\mathcal{B}$ with $\bigcup_1^\infty B_n = \{x_1, x_1', x_2, x_2'\}$. Then (4.18) holds for each B_n, and by (4.15) and dominated convergence, $f_{B_n}(x, y) \to f_\emptyset(x, y)$ for all $x, y \in [0,1]$; hence,

$$f_\emptyset(x_1, x_2) - f_\emptyset(x_1, x_2') - f_\emptyset(x_1', x_2) + f_\emptyset(x_1', x_2') = 0. \tag{4.19}$$

Step 4: Let $\varphi_1, \dots, \varphi_m$ be bounded functions on $[0,1]$ such that $\int_0^1 \varphi_1 = \int_0^1 \varphi_2 = 0$. Step 3 shows that (4.19) holds for a.e. $x_1, x_1', x_2, x_2' \in [0,1]$. We may thus fix $x_1', x_2' \in [0,1]$ such that (4.19) holds for a.e. x_1, x_2. Then multiply (4.19) by $\varphi_1(x_1) \varphi_2(x_2)$ and integrate over $x_1, x_2 \in [0,1]$. Since

$\int_0^1 \varphi_1 = \int_0^1 \varphi_2 = 0$, the integrals with the last three terms on the left-hand side of (4.19) vanish, and the result is

$$\int_{[0,1]^2} f_\emptyset(x_1, x_2)\varphi_1(x_1)\varphi_2(x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 = 0. \tag{4.20}$$

By the definition (4.15), this says

$$\int_{[0,1]^m} f(x_1, \dots, x_m)\varphi_1(x_1)\varphi_2(x_2)\varphi_3(x_3)\cdots\varphi_m(x_m) = 0. \tag{4.21}$$

Step 5: We may conclude in several ways. The perhaps simplest is to take $\varphi_j(x) = e^{2\pi \mathrm{i} n_j x_j}$, $j = 1, \dots, m$ with $n_j \in \mathbb{Z}$ and $n_1, n_2 \neq 0$. Step 4 then applies and (4.21) says that the Fourier coefficient

$$\widehat{f}(n_1, \dots, n_m) = 0 \tag{4.22}$$

when $n_1, n_2 \neq 0$. Since f is symmetric, it follows that $\widehat{f}(n_1, \dots, n_m) = 0$ as soon as at least two of the indices $n_1, \dots, n_m$ are non-zero.

Furthermore, let

$$g(x_1) := \int_{[0,1]^{m-1}} f(x_1, \dots, x_m)\,\mathrm{d}x_2 \cdots \mathrm{d}x_m - \int_{[0,1]^m} f, \tag{4.23}$$

and note that g is a function on $[0, 1]$ with $\int_0^1 g = 0$, and let

$$h(x_1, \dots, x_m) := \sum_{i=1}^m g(x_i) + \int_{[0,1]^m} f. \tag{4.24}$$

Then $\widehat{h}(n_1, \dots, n_m) = 0 = \widehat{f}(n_1, \dots, n_m)$ as soon as at least two of the indices $n_1, \dots, n_m$ are non-zero. Moreover, when $n_1 \neq 0$,

$$\widehat{h}(n_1, 0, \dots, 0) = \int_{[0,1]^m} h(x_1, \dots, x_n)e^{2\pi \mathrm{i} n_1 x_1} = \widehat{g}(n_1) = \widehat{f}(n_1, 0, \dots, 0) \tag{4.25}$$

and thus by symmetry $\widehat{h}(n_1, \dots, n_m) = \widehat{f}(n_1, \dots, n_m)$ also when exactly one index $n_1, \dots, n_m$ is non-zero. Finally, since $\int_0^1 g(x)\,\mathrm{d}x = 0$,

$$\widehat{h}(0, \dots, 0) = \int_{[0,1]^m} h = \int_{[0,1]^m} f = \widehat{f}(0, \dots, 0). \tag{4.26}$$

Consequently, $\widehat{h}(n_1, \dots, n_m) = \widehat{f}(n_1, \dots, n_m)$ for all $n_1, \dots, n_m$ and thus $h = f$ a.e.

Step 6: We have shown that a.e. $f = h$, given by (4.24). Let $a := \int f$; it remains to show that $a = 0$. This is easy; using (4.24) and Lemma 4.2,

$$\int_{A_1 \times \cdots \times A_m} f = \int_{A_1 \times \cdots \times A_m} h = \int_{A_1 \times \cdots \times A_m} a = a\lambda(A_1) \cdots \lambda(A_m), \tag{4.27}$$

and thus the assumption (4.4) yields $a = 0$. $\square$

Remark 4.4. As remarked in Janson [6, Remark 9.5], it is essential that f is symmetric in Lemma 4.3 (unlike Lemma 4.1). For example, it is easily seen that the condition (4.4) is also satisfied by every anti-symmetric f such that the margin

$$\int_0^1 f(x_1,\dots,x_m)\,\mathrm{d}x_m = 0 \tag{4.28}$$

for a.e. $x_1,\dots,x_{m-1}$; in fact, (4.28) implies, for any partition $\{A_1,\dots,A_m\}$,

$$\int_{A_1\times\cdots\times A_m} f\,\mathrm{d}x_1\cdots\,\mathrm{d}x_m = -\sum_{k=1}^{m-1}\int_{A_1\times\cdots\times A_{m-1}\times A_k} f\,\mathrm{d}x_1\cdots\,\mathrm{d}x_m = 0,$$

since each integral in the sum vanishes by the anti-symmetry. As a concrete example, for any $m\geqslant 2$, we may take the modified discriminant

$$f(x_1,\dots,x_m) = e^{2\pi\mathrm{i}(x_1+\cdots+x_m)}\prod_{j<k}\bigl(e^{2\pi\mathrm{i}x_j}-e^{2\pi\mathrm{i}x_k}\bigr) \tag{4.29}$$

(or its real part).

For $m=2$, it is easy to see that every f satisfying (4.4) for all equipartitions $\{A_1,\dots,A_m\}$ is the sum of a symmetric function satisfying (4.5) and an anti-symmetric function satisfying (4.28), see [6, Remark 9.5]. For $m\geqslant 3$, we do not know any characterization of general f satisfying (4.4), and we leave that as an open problem:

Problem 4.5. Find all integrable functions f on $[0,1]^m$ (not necessarily symmetric) that satisfy (4.4) for all partitions $\{A_1,\dots,A_m\}$ of $[0,1]$ into m disjoint measurable subsets such that $\lambda(A_1)=\cdots=\lambda(A_m)=1/m$.

We end this section with another, much simpler, extension of Lemma 4.1 to $\alpha_1,\dots,\alpha_m$ that may be different and possibly with $\sum_{i=1}^m\alpha_i<1$, with exception only of the exceptional case treated in Lemma 4.3 when all α_i are equal to $1/m$.

Lemma 4.6. *Let $m\geqslant 1$ and let $\alpha_1,\dots,\alpha_m\in(0,1)$ with $\sum_{i=1}^m\alpha_i\leqslant 1$. Suppose that f is an integrable function on $[0,1]^m$ such that*

$$\int_{A_1\times\cdots\times A_m} f = 0 \tag{4.30}$$

for all sequences $A_1,\dots,A_m$ of disjoint measurable subsets of $[0,1]$ such that $\lambda(A_i)=\alpha_i$, $i=1,\dots,m$. Suppose further that either

(i) $\sum_{i=1}^m\alpha_i<1$, *or*
(ii) $\sum_{i=1}^m\alpha_i=1$ *but $(\alpha_1,\dots,\alpha_m)\neq(1/m,\dots,1/m)$, and f is symmetric.*

Then $f=0$ a.e.

Proof. The case $m=1$ is included in Lemma 4.1. (For $m=1$, (ii) cannot occur.) The case (ii) with $m=2$ is [6, Lemma 9.4]. The remaining cases are proved by induction (on m) in the same way as the special case in [6, Lemma 7.3]; we sketch the proof and refer to [6] for omitted details.

We thus assume $m \geqslant 2$, and in the case $m = 2$ that (i) holds. Furthermore, if (ii) holds, we may assume $\alpha_m \neq \alpha_{m-1}$ by permuting the coordinates.

We fix a set A_1 with $\lambda(A_1) = \alpha_1$ and consider the function

$$f_{A_1}(x_2, \dots, x_m) := \int_{A_1} f(x_1, \dots, x_m) \, \mathrm{d}x_1$$

on B^{m-1} where $B := [0,1] \setminus A_1$. B is as a measure space isomorphic to $[0,1]$, after rescaling the measure, and the hypothesis implies that f_{A_1} satisfies a corresponding hypothesis on B^{m-1}; hence $f_{A_1} = 0$ a.e. on B^{m-1} by induction. It follows that (4.30) holds for all disjoint sets $A_1, \dots, A_m$ with $\lambda(A_1) = \alpha_1$ and $\lambda(A_2), \dots, \lambda(A_m)$ arbitrary.

Now instead fix any disjoint $A_2, \dots, A_m$ with $\sum_{i=2}^m \lambda(A_i) < 1 - \alpha_1$, and let $B' := [0,1] \setminus \bigcup_{i=2}^m A_i$. Then (4.30) thus holds for any $A_1 \subset B'$ with $\lambda(A_1) = \alpha_1$, and it follows from the case $m = 1$ applied to $f^{A_2,\dots,A_m}(x) := \int_{A_2 \times \dots \times A_m} f(x, x_2, \dots, x_m) \, \mathrm{d}x_2 \cdots \mathrm{d}x_m$ that $f^{A_2,\dots,A_m}(x) = 0$ a.e.; hence (4.30) holds for all disjoint $A_1, \dots, A_m$ with $\sum_{i=2}^m \lambda(A_i) < 1 - \alpha_1$. It follows that $f(x_1, \dots, x_m) = 0$ for every Lebesgue point $(x_1, \dots, x_m)$ of f with $x_1, \dots, x_m$ distinct. □

Remark 4.7. In Lemma 4.6(ii), the assumption that f is symmetric is essential, as is seen by the counterexample in Remark 4.4.

Remark 4.8. The proof shows that in the case $\sum_{i=1}^m \alpha_i = 1$, it suffices to assume that f is symmetric in the last two variables, provided $\alpha_{m-1} \neq \alpha_m$.

We apply the results above to the property $\widetilde{\mathcal{P}}_*(F; \alpha_1, \dots, \alpha_m)$. By (3.6), this property says that (4.30) holds for $f := \widetilde{\Psi}_{F,W} - p^{e(F)}$ and all disjoint subsets $A_1, \dots, A_m$ of $[0,1]$ with $\lambda(A_i) = \alpha_i$, $i = 1, \dots, m$.

Lemma 4.9. *Let $m \geqslant 1$ and let $\alpha_1, \dots, \alpha_m \in (0,1)$ with $\sum_{i=1}^m \alpha_i \leqslant 1$. Suppose that W is a graphon and $p \in [0,1]$.*

(a) *If $(\alpha_1, \dots, \alpha_m) \neq (1/m, \dots, 1/m)$, then $\widetilde{\mathcal{P}}_*(F; \alpha_1, \dots, \alpha_m)$ holds if and only if*

$$\widetilde{\Psi}_{F,W}(x_1, \dots, x_m) = p^{e(F)} \quad a.e. \tag{4.31}$$

(b) *If $(\alpha_1, \dots, \alpha_m) = (1/m, \dots, 1/m)$, then $\widetilde{\mathcal{P}}_*(F; \alpha_1, \dots, \alpha_m)$ holds if and only if there exists an integrable function h with $\int_0^1 h = p^{e(F)}/m$ such that*

$$\widetilde{\Psi}_{F,W}(x_1, \dots, x_m) = \sum_{i=1}^m h(x_i) \quad a.e. \tag{4.32}$$

Proof. Part (a) follows directly from (3.6) and Lemma 4.6, while (b) follows from Lemmas 4.2 and 4.3, with $h(x) = g(x) + p^{e(F)}/m$. □

We thus see that the exceptional case $\alpha_1 = \dots = \alpha_m = 1/m$ in Theorem 2.12 is more intricate than the cases covered by Theorem 2.11.

We note also a similar result for $\mathcal{P}_*$.

Lemma 4.10. *If $\sum_{i=1}^m \alpha_i < 1$, then $\mathcal{P}_*(F;\alpha_1,\dots,\alpha_m)$ holds if and only if*

$$\Psi_{F,W}(x_1,\dots,x_m) = p^{e(F)} \quad a.e. \tag{4.33}$$

Proof. This too follows from (3.6) and Lemma 4.6. □

In this case we have to assume $\sum_{i=1}^m \alpha_i < 1$ for the proof, because $\Psi_{F,W}$ is (in general) not symmetric, cf. Remarks 4.7 and 4.4.

Problem 4.11. Does Lemma 4.10 hold also if $\sum_{i=1}^m \alpha_i = 1$ with $(\alpha_1,\dots,\alpha_m) \neq (1/m,\dots,1/m)$?

5. Reduction to a two-type graphon

We next reduce the problem by showing that, as for the similar problem considered by Simonovits and Sós [16], if the property $\widetilde{\mathcal{P}}_*(F;\alpha_1,\dots,\alpha_m)$ is *not* quasi-random, then there exists a counterexample with a 2-type graphon. This reduction reduces our problem to an algebraic one, which we consider in the next section.

We state the reduction in a somewhat general form, to be used together with Lemma 4.9, and we give two versions (Theorems 5.2 and 5.3), to handle the two cases in parts (a) and (b) in Lemma 4.9. The proofs are given later in this section.

Remark 5.1. Theorem 5.2 is an extension of Janson [6, Theorem 5.5], where Φ is a multiaffine polynomial, which would be sufficient for our application here. We nevertheless state Theorem 5.2 in order to show the similarities between Theorems 5.2 and 5.3, and because we now can give a more elegant proof of a more general statement than in [6], see Remark 5.9.

If $\Phi\big((w_{ij})_{i<j}\big)$ is a function of the $\binom{m}{2}$ variables w_{ij}, $1 \leqslant i < j \leqslant m$, for some $m \geqslant 2$, and W is a graphon, we define, for $x_1,\dots,x_m \in [0,1]$,

$$\Phi_W(x_1,\dots,x_m) := \Phi\big((W(x_i,x_j))_{i<j}\big). \tag{5.1}$$

Theorem 5.2. *Suppose that $\Phi\big((w_{ij})_{i<j}\big)$ is a continuous function of the $\binom{m}{2}$ variables w_{ij}, $1 \leqslant i < j \leqslant m$, for some $m \geqslant 2$, and let $a \in \mathbb{R}$. Then the following are equivalent.*

(i) *There exists a graphon W such that*
$$\Phi_W(x_1,\dots,x_m) = a \tag{5.2}$$
for a.e. $x_1,\dots,x_m \in [0,1]$, but W is not *a.e. constant.*

(ii) *There exists a 2-type graphon W such that (5.2) holds for all $x_1,\dots,x_m$, but W is* not *constant.*

(iii) *There exist numbers $u,v,s \in [0,1]$, not all equal, such that for every subset $A \subseteq [m]$, if we choose*
$$w_{ij} := \begin{cases} u, & i,j \in A, \\ v, & i,j \notin A, \\ s, & i \in A,\ j \notin A \text{ or conversely}, \end{cases} \tag{5.3}$$

then

$$\Phi((w_{ij})_{i<j}) = a. \tag{5.4}$$

Theorem 5.3. *Suppose that* $\Phi\bigl((w_{ij})_{i<j}\bigr)$ *is a continuous function of the* $\binom{m}{2}$ *variables* w_{ij}, $1 \leqslant i < j \leqslant m$, *for some* $m \geqslant 2$. *Then the following are equivalent.*

(i) *There exists a graphon* W *and a function* h *on* $[0,1]$, *with* h *not a.e.* 0, *such that*

$$\Phi_W(x_1, \dots, x_m) = \sum_{i=1}^{m} h(x_i) \tag{5.5}$$

for a.e. $x_1, \dots, x_m \in [0,1]$, *but* W *is* not *a.e. constant.*

(ii) *There exists a* 2*-type graphon* W *and a function* h *on* $[0,1]$, *with* h *not a.e.* 0, *such that* (5.5) *holds for all* $x_1, \dots, x_m$, *but* W *is* not *constant.*

(iii) *There exist numbers* $u, v, s \in [0,1]$, *not all equal, and* $a, b \in \mathbb{R}$, *not both* 0, *such that for every subset* $A \subseteq [m]$, *if we choose*

$$w_{ij} := \begin{cases} u, & i,j \in A, \\ v, & i,j \notin A, \\ s, & i \in A,\ j \notin A \text{ or conversely}, \end{cases} \tag{5.6}$$

then

$$\Phi((w_{ij})_{i<j}) = a + b|A|. \tag{5.7}$$

Remark 5.4. In part (ii) of Theorems 5.2–5.3, we may further require that the two parts of $[0,1]$ are the intervals $[0, \frac{1}{2}]$ and $(\frac{1}{2}, 1]$. Equivalently, we may regard W as a graphon defined on the two-point probability space $(\{0,1\}, \mu)$, with $\mu\{0\} = \mu\{1\} = \frac{1}{2}$.

Remark 5.5. Theorem 5.3 holds also without the restrictions that h is not a.e. 0, and a, b are not both 0; this follows by the same proof (with some simplifications). Note that the excluded case, when $h = 0$ a.e. and $a = b = 0$, is equivalent to Theorem 5.2. For our purposes, it is essential that the case $a = b = 0$ is excluded, since there are such examples that have to be excluded from our arguments, for example the bipartite example in Remark 2.7, which corresponds to the case $u = v = 0$, $s = 1$ and $\Phi((w_{ij})_{i<j}) = 0$ for any A.

The proofs follow the proof of Janson [6, Theorem 5.5], with some modifications. We prove the more complicated Theorem 5.3 in detail first, and then sketch the similar but simpler proof of Theorem 5.2.

Proof of Theorem 5.3. (ii) $\Longrightarrow$ (i): Trivial.

(iii) $\Longrightarrow$ (ii): Define a 2-type graphon W by

$$W(x,y) := \begin{cases} u, & x, y > \frac{1}{2}, \\ v, & x, y \leqslant \frac{1}{2}, \\ s, & x \leqslant \frac{1}{2} < y \text{ or conversely}, \end{cases} \tag{5.8}$$

and let the function h be

$$h(x) := \begin{cases} a/m, & x \leqslant \frac{1}{2}, \\ a/m + b, & x > \frac{1}{2}. \end{cases} \tag{5.9}$$

Then

$$\Phi_W(x_1, \dots, x_m) = \Phi((w_{ij})_{i<j}) \tag{5.10}$$

where w_{ij} is given by (5.6) with $A := \{i : x_i > \frac{1}{2}\}$, and (5.5) follows from (5.7).

(i) $\implies$ (iii): Suppose that W is a graphon as in (i), but that (iii) does not hold; we will show that this leads to a contradiction. We first use Lemma 5.8 below, which (by replacing W by $\overline{W}$ and h by $\bar{h}$) shows that we may assume that (5.5) holds for *all* $x_1, \dots, x_m \in [0,1]$.

Suppose that $x, y \in [0,1]$. Given $A \subseteq [m]$, let $x_i := x$ for $i \in A$ and $x_i := y$ for $i \notin A$. Then $W(x_i, x_j) = w_{ij}$ as given by (5.6) with $u = W(x,x)$, $v = W(y,y)$, $s = W(x,y)$. Furthermore, (5.5) holds by our assumption, and thus

$$\begin{aligned} \Phi((w_{ij})_{i<j}) = \Phi_W(x_1, \dots, x_m) &= \sum_{i=1}^m h(x_i) = |A|h(x) + (m - |A|)h(y) \\ &= a + b|A| \end{aligned} \tag{5.11}$$

with $a = mh(y)$ and $b = h(x) - h(y)$. Hence, (5.7) holds. Since (iii) does not hold, we must have either $u = v = s$ or $a = b = 0$. Note that $a = b = 0$ if and only if $h(x) = h(y) = 0$. Consequently, we have shown the following property:

$$\text{If } x, y \in [0,1], \text{ then } W(x,x) = W(y,y) = W(x,y) \text{ or } h(x) = h(y) = 0. \tag{5.12}$$

Furthermore, if $W(x,x) = W(y,y)$ then (5.11), with $A = \emptyset$ and $A = [m]$, implies that

$$a = \Phi_W(y, \dots, y) = \Phi_W(x, \dots, x) = a + mb \tag{5.13}$$

and thus $b = 0$ so $h(x) = h(y)$. Consequently, (5.12) implies that

$$x, y \in [0,1] \implies h(x) = h(y). \tag{5.14}$$

In other words, $h(x) = \gamma$ for some constant $\gamma \in \mathbb{R}$.

Note that $\gamma \neq 0$, since otherwise $h(x)$ would be 0 for all x, contrary to the assumption (i). Hence, $h(x) \neq 0$ for all x and (5.12) implies

$$x, y \in [0,1] \implies W(x,x) = W(y,y) = W(x,y). \tag{5.15}$$

Thus W is constant, contradicting the assumption.

This contradiction shows that (iii) holds. □

Proof of Theorem 5.2. We argue as in the proof of Theorem 5.3, with $b = 0$ and $h(x) = a/m$; in the proof of (i) $\implies$ (iii) we use Lemma 5.7 instead of Lemma 5.8, and note directly that (5.11) with $b = 0$, which is (5.4), implies $u = v = s$ since (iii) is assumed not to hold. □

Remark 5.6. In both proofs, the proof of (iii) $\implies$ (ii) also works in the opposite direction and thus shows (iii) $\iff$ (ii) directly; (iii) is just an explicit version of what (ii) means.

The proofs used the following technical lemmas, which both are consequences of a recent powerful general removal lemma by Petrov [10]. Recall that a graphon $\overline{W}$ is a version of W if $W = \overline{W}$ a.e.

Lemma 5.7. *Suppose that $\Phi\bigl((w_{ij})_{i<j}\bigr)$ is a continuous function of the $\binom{m}{2}$ variables $w_{ij} \in [0,1]$, $1 \leqslant i < j \leqslant m$, for some $m \geqslant 2$. Suppose further that $W : [0,1]^2 \to [0,1]$ is a graphon, i.e., a symmetric measurable function, and suppose that*

$$\Phi_W(x_1, \dots, x_m) = a \tag{5.16}$$

for some $a \in \mathbb{R}$ and a.e. $x_1, \dots, x_m \in [0,1]$. Then there is a version $\overline{W}$ of W such that

$$\Phi_{\overline{W}}(x_1, \dots, x_m) = a \tag{5.17}$$

for all $x_1, \dots, x_m \in [0,1]$.

Proof. This is a direct application of [10, Theorem 1(2)], see [10, Example 1]. We let $M := \Phi^{-1}(a) \subseteq [0,1]^{\binom{m}{2}}$ and note that (5.16) can be written $\bigl(W(x_i,x_j)\bigr)_{i<j} \in M$ for a.e. $x_1, \dots, x_m$. By [10, Theorem 1(2)], there exists a version $\overline{W}$ of W such that $\bigl(\overline{W}(x_i,x_j)\bigr)_{i<j} \in M$ for all $x_1, \dots, x_m$, which is (5.17). (Petrov's theorem is stated for an infinite sequence $x_1, x_2, \dots$, for maximal generality, but we can always ignore all but any given finite number of the variables.) $\square$

Lemma 5.8. *Suppose that $\Phi\bigl((w_{ij})_{i<j}\bigr)$ is a continuous function of the $\binom{m}{2}$ variables $w_{ij} \in [0,1]$, $1 \leqslant i < j \leqslant m$, for some $m \geqslant 2$. Suppose further that $W : [0,1]^2 \to [0,1]$ is a graphon, i.e., a symmetric measurable function, and suppose that*

$$\Phi_W(x_1, \dots, x_m) = \sum_{i=1}^m h(x_i) \tag{5.18}$$

for some $h : [0,1] \to \mathbb{R}$ and a.e. $x_1, \dots, x_m \in [0,1]$. Then there is a version $\overline{W}$ of W and a measurable function $\bar{h} : [0,1] \to \mathbb{R}$ such that

$$\Phi_{\overline{W}}(x_1, \dots, x_m) = \sum_{i=1}^m \bar{h}(x_i) \tag{5.19}$$

for all $x_1, \dots, x_m \in [0,1]$.

Proof. We translate (5.18) into the setting of [10] as follows.

By (5.18), for a.e. $x_1, \dots, x_m, y_1, \dots, y_m \in [0,1]$,

$$\Phi_W(x_1, \dots, x_m) - \Phi_W(y_1, \dots, y_m) = \sum_{\ell=1}^m \bigl(\Phi_W(x_\ell, y_1, \dots, \widehat{y_\ell}, \dots, y_m) - \Phi_W(y_1, \dots, \dots, y_m)\bigr), \tag{5.20}$$

where $\widehat{y_\ell}$ means that this variable is omitted. Let $x_{m+i} := y_i$ ($1 \leqslant i \leqslant m$) and $w_{ij} := W(x_i, x_j)$ ($1 \leqslant i, j \leqslant 2m$). Then (5.20) can be written as

$$\widehat{\Phi}\Big((w_{ij})_{i\neq j}\Big) = 0 \tag{5.21}$$

for some continuous function $\widehat{\Phi} : [0,1]^{2m(2m-1)} \to \mathbb{R}$. Let

$$M := \widehat{\Phi}^{-1}(0) \subseteq [0,1]^{2m(2m-1)}. \tag{5.22}$$

Since $\widehat{\Phi}$ is continuous, M is a closed subset, and by (5.21),

$$\big(W(x_i, x_j)\big)_{i\neq j} \in M \tag{5.23}$$

for a.e. $x_1, \dots, x_{2m}$. By [10, Theorem 1(2)], there exists a version $\overline{W}$ of W such that

$$\big(\overline{W}(x_i, x_j)\big)_{i\neq j} \in M \tag{5.24}$$

for all $x_1, \dots, x_{2m}$. This means that $\widehat{\Phi}\big((\overline{W}(x_i, x_j))_{i\neq j}\big) = 0$ for all $x_1, \dots, x_{2m}$, and thus the analogue of (5.20) for $\overline{W}$ holds for all $x_1, \dots, x_m, y_1, \dots, y_m$. Now choose $y_1 = \dots = y_m = 0$. Then this analogue of (5.20) yields (5.19) with $\bar{h}(x) = \Phi_{\overline{W}}(x, 0, \dots, 0) - \frac{m-1}{m}\Phi_{\overline{W}}(0, \dots, 0)$. $\square$

Remark 5.9. Lemma 5.7, which follows from Petrov's removal lemma [10], is a simpler, stronger and more general version of Janson [6, Lemma 5.3]. Similarly, a modification of the proof of [6, Lemma 5.3] can be used to prove a weaker version of Lemma 5.8; however, Petrov's removal lemma enables us to a simpler and stronger statement with a simpler proof.

6. An algebraic condition

It is now easy to prove Theorem 2.11.

Proof of Theorem 2.11. (i): Suppose, in order to get a contradiction, that the property $\widetilde{\mathcal{P}}(F; \alpha_1, \dots, \alpha_m)$ is not p-quasi-random. By Lemma 3.4, also $\widetilde{\mathcal{P}}_*(F; \alpha_1, \dots, \alpha_m)$ is not p-quasi-random. That means that there exists a graphon W that is not a.e. equal to p such that $\widetilde{\mathcal{P}}_*(F; \alpha_1, \dots, \alpha_m)$ holds, and thus by Lemma 4.9(a),

$$\widetilde{\Psi}_{F,W}(x_1, \dots, x_m) = p^{e(F)} \quad \text{a.e.} \tag{6.1}$$

If W a.e. equals a constant, w say, then $\widetilde{\Psi}_{F,W} = w^{e(F)}$ a.e., and thus $w^{e(F)} = p^{e(F)}$ and $w = p$, so $W = p$ a.e. which we have excluded. Hence, W is not a.e. constant.

Note that $\widetilde{\Psi}_{F,W}(x_1, \dots, x_m)$ by (3.1)–(3.2) is a polynomial in $W(x_i, x_j)$, $1 \leqslant i < j \leqslant m$, and thus by (5.1) can be written as $\Phi_W(x_1, \dots, x_m)$ for a suitable polynomial Φ. We apply Theorem 5.2, with $a = p^{e(F)}$. By (6.1), Theorem 5.2(i) holds, and thus Theorem 5.2(iii) holds. Let u, v, s be as there, and define w_{ij} by (5.3).

Choosing $A = [m]$, we have $w_{ij} = u$ for all i and j, and it is easily seen that $\Phi\big((w_{ij})_{i<j}\big) = u^{e(F)}$ (see also Lemma 6.2 below); hence (5.4) yields

$u = p$. Similarly, the case $A = \emptyset$ yields $v = p$. Finally, take $A = \{1\}$, and regard $\Phi\big((w_{ij})_{i<j}\big)$ as a polynomial in s. Since $u, v > 0$ and $e(F) > 0$, this polynomial has non-negative coefficients and at least one non-zero term with a positive power of s; hence the polynomial is strictly increasing in $s > 0$, so (5.4) has at most one root s. However, when $u = v = p$, (5.4) is satisfied by $s = p$, and thus this is the only root. Consequently, $u = v = s = p$, a contradiction, which completes the proof.

(ii): Similar, using Lemmas 3.4 and 4.10 and Theorem 5.2. $\square$

Remark 6.1. Suppose that the graph F contains two vertices that are *twins*, i.e., such that the map interchanging these vertices and fixing all others is an automorphism. Label F such that the twins are vertices $m-1$ and m. The argument in the proof of Theorem 2.11 shows, using Remark 4.8, that $\mathcal{P}(F; \alpha_1, \dots, \alpha_m)$ is quasi-random provided $\alpha_{m-1} \neq \alpha_m$. (We do not know whether this extends to $\alpha_m = \alpha_{m-1}$. Cf. Problem 4.11.) In particular, this applies to $F = P_3$, see Example 2.18 and Problem 2.19.

For Theorem 2.12, the algebra is more complicated, and we analyse the condition (5.7) as follows.

For a subset A of $V(F)$, let $e_F(A)$ be the number of edges in F with both endpoints in A; similarly, if A and B are disjoint subsets of $V(F)$, let $e_F(A, B)$ be the number of edges with one endpoint in A and the other in B. Further, let $A^{\mathsf{c}} := V(F) \setminus A$ be the complement of A.

Lemma 6.2. *Suppose that F is a graph with $|F| = m$ and let W be the 2-type graphon given by (5.8) for some $u, v, s \in [0,1]$. Let $x_1, \dots, x_m \in [0,1]$ and let $k := |\{i \leqslant m : x_i > 1/2\}|$. Then*

$$\widetilde{\Psi}_{F,W}(x_1, \dots, x_m) = \binom{m}{k}^{-1} \sum_{A \subseteq V(F): |A| = k} u^{e_F(A)} v^{e_F(A^{\mathsf{c}})} s^{e_F(A, A^{\mathsf{c}})}. \tag{6.2}$$

Proof. Let $A := \{i \leqslant m : x_i > 1/2\}$. Then by (3.1) and (5.8),

$$\Psi_{F,W}(x_1, \dots, x_m) = u^{e_F(A)} v^{e_F(A^{\mathsf{c}})} s^{e_F(A, A^{\mathsf{c}})}. \tag{6.3}$$

By (3.2), $\widetilde{\Psi}_{F,W}(x_1, \dots, x_m)$ is the average of this over all permutations of $x_1, \dots, x_m$, which means taking the average over the $\binom{m}{k}$ sets $A \subseteq [m]$ with $|A| = k$. $\square$

Lemma 6.3. *Suppose that F is a graph with $|F| = m$. Then the following are equivalent.*

(i) *For some $p \in (0,1]$, $\widetilde{\mathcal{P}}(F; 1/m, \dots, 1/m)$ is* not *p-quasi-random.*
(ii) *For some $p \in (0,1]$, $\widetilde{\mathcal{P}}_*(F; 1/m, \dots, 1/m)$ is* not *p-quasi-random.*
(iii) *There exist numbers $u, v, s \geqslant 0$, not all equal, and some real a and b, not both 0, such that*

$$\sum_{A \subseteq V(F): |A| = k} u^{e_F(A)} v^{e_F(A^{\mathsf{c}})} s^{e_F(A, A^{\mathsf{c}})} = \binom{m}{k}(a + bk), \qquad k = 0, \dots, m. \tag{6.4}$$

(iv) *There exist numbers $u, v, s \geqslant 0$, not all equal, such that the polynomial (in q)*

$$\Lambda_{F;u,v,s}(q) := \sum_{A \subseteq V(F)} u^{e_F(A)} v^{e_F(A^c)} s^{e_F(A,A^c)} q^{|A|}(1-q)^{m-|A|} \tag{6.5}$$

has degree at most 1, but does not vanish identically.

(v) *There exist numbers $u, v, s \geqslant 0$, not all equal, such that the polynomial (in x)*

$$\hat{\Lambda}_{F;u,v,s}(x) := \sum_{A \subseteq V(F)} u^{e_F(A)} v^{e_F(A^c)} s^{e_F(A,A^c)} (x-1)^{|A|} \tag{6.6}$$

is divisible by x^{m-1}, but does not vanish identically.

Note that (for $q \in [0,1]$) $\Lambda_{F;u,v,s}(q)$ is the expectation of $u^{e_F(A)} v^{e_F(A^c)} s^{e_F(A,A^c)}$ if A is the random subset $[m]_q$ of $[m]$ obtain by including each element with probability q, independently of each other.

Proof. (i) $\iff$ (ii): This is contained in Lemma 3.4.

(ii) $\implies$ (iii): If (ii) holds, then there exists a graphon W that is not a.e. constant for which $\widetilde{\mathcal{P}}_*(F; 1/m, \dots, 1/m)$ holds. Then, by Lemma 4.9(b), there exists an integrable function h with $\int_0^1 h \neq 0$ such that (4.32) holds.

As in the proof of Theorem 2.11, $\widetilde{\Psi}_{F,W}(x_1, \dots, x_m)$ can be written as $\Phi_W(x_1, \dots, x_m)$ for a polynomial Φ. Then (4.32) is the same as (5.5) and Theorem 5.3(i) holds. By Theorem 5.3 (and its proof) we may assume that W is a 2-type graphon given by (5.8) for some $u, v, s \in [0,1]$, and then Lemma 6.2 and (5.10) show that (5.7) is equivalent to (6.4), and thus (iii) follows.

(iii) $\implies$ (ii): This is similar but simpler. We may assume that $u, v, s \in [0,1]$, by multiplying them by a small positive number if necessary. Let W be the graphon defined by (5.8). Then Lemma 6.2 and (6.4) yield $\widetilde{\Psi}_{F,W}(x_1, \dots, x_m) = a + bk$ where $k = |\{i : x_i > 1/2\}$, so (4.32) holds with h given by (5.9).

We have assumed that a and b are not both 0, and thus $h(x)$ is not identically 0. Furthermore, (4.32) implies $h(x) \geqslant 0$ a.e., and thus $\int_0^1 h > 0$. Since $\widetilde{\Psi}_{F,W} \leqslant 1$, (4.32) also implies $\int_0^1 h \leqslant 1/m$. Hence there exists $p \in (0,1]$ with $p^{e(F)} = m \int_0^1 h$. (Also in the trivial case $e(F) = 0$, since then $\widetilde{\Psi}_{F,W} = 1$.) Lemma 4.9 now shows that $\widetilde{\mathcal{P}}_*(F; 1/m, \dots, 1/m)$ holds, and since W is not a.e. constant, this yields (ii).

(iii) $\iff$ (iv): By multiplying (6.4) by t^k and summing over k, we see that (6.4) is equivalent to

$$\sum_{A \subseteq V(F)} u^{e_F(A)} v^{e_F(A^c)} s^{e_F(A,A^c)} t^{|A|} = \sum_{k=0}^{m} \binom{m}{k} (a+bk) t^k, \qquad t \in \mathbb{R}. \tag{6.7}$$

Letting $t = q/(1-q)$ and multiplying by $(1-q)^m$, this is equivalent to

$$\sum_{A\subseteq V(F)} u^{e_F(A)} v^{e_F(A^c)} s^{e_F(A,A^c)} q^{|A|}(1-q)^{m-|A|} = \sum_{k=0}^{m} \binom{m}{k}(a+bk)q^k(1-q)^{m-k}$$

where the right hand side equals $a + bmq$ by an elementary calculation (or by the formula for the mean of a binomial distribution). The equivalence follows.

(iv) $\iff$ (v): Take $q = 1/x$, replace A by A^c and interchange u and v to obtain

$$\hat\Lambda_{F;u,v,s}(x) = x^m \Lambda_{F;v,u,s}(1/x). \tag{6.8}$$

□

Remark 6.4. It follows from the proof that the polynomial $\Lambda_{F;u,v,s}(q)$ has degree 0, i.e., is a (non-zero) constant $\iff$ (6.4) holds with $b = 0$ $\iff$ $\widetilde\Psi_{F,W}(x_1,\dots,x_m) = a$ for some (non-zero) a. As shown above in the proof of Theorem 2.11, this happens for some $u, v, s \geqslant 0$, not all equal, only in the trivial case $e(F) = 0$. (This is an equivalent way of stating the algebraic part of the proof of Theorem 2.11, but we preferred to give a direct proof above without the present machinery.) Hence, if $e(F) > 0$ and (iv) holds, then the degree of $\Lambda_{F;u,v,s}$ is exactly 1.

Remark 6.5. $\Lambda_{F;u,v,s}(q)$ is not changed if we add some isolated vertices to F. Hence we may assume that F has no isolated vertices.

We note that the cases $k = 0$ and $k = m$ of (6.4) simply are

$$v^{e(F)} = a, \tag{6.9}$$

$$u^{e(F)} = a + mb. \tag{6.10}$$

In particular, the assumption that not $a = b = 0$ means that not $u = v = 0$. (This case has to be excluded, for any non-bipartite F, cf. Remark 2.7.)

Moreover, if F has degree sequence $d_1,\dots,d_m$, the cases $k = 1$ and $k = m-1$ of (6.4) are

$$\frac{1}{m}\sum_{i=1}^{m} v^{e(F)-d_i} s^{d_i} = a + b, \tag{6.11}$$

$$\frac{1}{m}\sum_{i=1}^{m} u^{e(F)-d_i} s^{d_i} = a + (m-1)b. \tag{6.12}$$

Example 6.6. If $F = K_2$, then by (6.5),

$$\Lambda_{F;u,v,s}(q) = uq^2 + 2sq(1-q) + v(1-q)^2 = v + 2(s-v)q + (u+v-2s)q^2,$$

which has degree 1 if we choose any distinct u and v and let $s = (u+v)/2$. Hence Lemma 6.3 shows that $\widetilde{\mathcal{P}}(K_2; 1/2, 1/2)$ is not quasi-random, as we already know, see Example 2.10.

In this case, $\Psi_{K_2,W}(x_1,x_2) = W(x_1,x_2)$, so Lemma 4.9(b) shows that $\widetilde{\mathcal{P}}_*(K_2;1/2,1/2)$ holds if and only if $W(x,y) = h(x)+h(y)$ for some measurable $h : [0,1] \to [0,1]$ with $\int_0^1 h = p/2$, see further Janson [6, Section 9].

Remark 6.7. We may add some further conditions on u,v,s in Lemma 6.3(iii)–(v). In the trivial case $e(F) = 0$ we can take any u,v,s, so let us assume $e(F) > 0$. By Remark 6.4, we then must have $b \neq 0$, so by (6.9)–(6.10), $u \neq v$. Furthermore, we may interchange u and v (and replace q by $1-q$ in (6.5)), so we may assume $u < v$. In this case, (6.9)–(6.10) yield $b > 0$. By (6.11) and (6.9), this implies $s > v$, and by (6.12) and (6.10), it implies $s < u$. Hence we may assume $v < s < u$.

Suppose $v = 0$. Then $a = 0$ by (6.9). By Remark 6.5, we may assume that F has no isolated vertices. If $d_i < e(F)$ for all i, then (6.11) yields $0 = a+b = b$, which is impossible. Hence we must have $d_i = e(F)$ for some i, which means that F is a star. In the case of a star with $m = |F| \geqslant 3$, $v = a = 0$ in (6.11) yields $s^{m-1} = mb$, while (6.10) yields $u^{m-1} = mb$ so $u = s$, a contradiction. Hence $v = 0$ is impossible and we may assume $v > 0$. (If $m = 2$, so $F = K_2$, $v = 0$ is possible, but we may choose any $v > 0$ and $u > v$ by Example 6.6.)

Consequently, it suffices to consider distinct $u,v,s > 0$, and we may assume $0 < v < s < u$ (or, by symmetry, $0 < u < s < v$).

Furthermore, the equations (6.4) are homogeneous in (u,v,s), so we may assume that any given of them equals 1; for example, we may assume $v = 1$, which implies $a = 1$ by (6.9).

7. Completing the proof of Theorem 2.12

We say that a graph F is *good* if, for every $p \in (0,1]$, $\widetilde{\mathcal{P}}(F;1/m,\dots,1/m)$ is p-quasi-random; otherwise F is *bad*. In this terminology, Lemma 6.3 says (using Remark 6.7) that F is bad if and only if there exist distinct $u,v,s > 0$ such that (6.4) holds, or, equivalently, that $\Lambda_{F;u,v,s}(q)$ in (6.5) has degree at most 1.

An empty graph, i.e., a graph F with $e(F) = 0$, is trivially bad; in this case (6.5) yields $\Lambda_{F;u,v,s}(q) = 1$, so $\Lambda_{F;u,v,s}$ has degree 0. By Remark 6.4, this is the only case when $\deg(\Lambda_{F;u,v,s}) = 0$.

The single edge K_2 is also bad, see Examples 2.10 and 6.6. More generally, any graph F with $e(F) = 1$ is bad by Remark 6.5.

Conjecture 2.13 says that all other graphs are good. We proceed to verify this in the cases given in Theorem 2.12.

Example 7.1 (regular graphs)**.** Suppose that F is d-regular for some $d \geqslant 1$, and that $m = |F| \geqslant 3$. (This includes the case K_m, $m \geqslant 3$, considered by [5].) Then $e(F) = dm/2$.

We use only (6.9)–(6.12); if we further simplify by assuming $v = a = 1$, as we may by Remark 6.7, we obtain, from (6.10)–(6.12),

$$u^{dm/2} = 1 + mb, \tag{7.1}$$

$$s^d = 1 + b, \tag{7.2}$$

$$u^{d(m-2)/2} s^d = 1 + (m-1)b, \tag{7.3}$$

and thus

$$(1 + (m-1)b)^m = (1 + mb)^{m-2}(1 + b)^m. \tag{7.4}$$

However, the function

$$h(x) := (m-2)\log(1+mx) + m\log(1+x) - m\log(1+(m-1)x) \tag{7.5}$$

(defined for $x > -1/m$) has derivative

$$h'(x) = \frac{m(m-1)(m-2)x^2}{(1+x)(1+(m-1)x)(1+mx)} > 0 \tag{7.6}$$

for $x > -1/m$ with $x \neq 0$, and thus $h(x)$ is strictly increasing on $(-1/m, \infty)$ and $h(x) \neq h(0) = 0$ for $x \neq 0$, which shows that (7.4) implies $b = 0$, and thus $s = u = 1 = v$ by (7.1)–(7.3), a contradiction. Consequently, there are no u, v, s satisfying the conditions and thus F is good.

Example 7.2 (stars)**.** Suppose that F is a star $K_{1,m-1}$. Let $A \subseteq V(F)$ and let $k := |A|$. If A contains the centre of F, then $e_F(A) = k - 1$, $e_F(A^{\mathsf{c}}) = 0$ and $e_F(A, A^{\mathsf{c}}) = m - k$; otherwise, $e_F(A) = 0$, $e_F(A^{\mathsf{c}}) = m - k - 1$ and $e_F(A, A^{\mathsf{c}}) = k$. It follows from (6.6) and the binomial theorem that

$$\hat{\Lambda}_{F;u,v,s}(x) = (x-1)\bigl(u(x-1)+s\bigr)^{m-1} + \bigl(s(x-1)+v\bigr)^{m-1}. \tag{7.7}$$

Assume $m \geqslant 3$, and that F is bad. Then, by Lemma 6.3(v) and Remark 6.7, there exist distinct $u, v, s > 0$ such that $\hat{\Lambda}_{F;u,v,s}(x)$ is divisible by x^{m-1}. In particular,

$$0 = \hat{\Lambda}_{F;u,v,s}(0) = -(s-u)^{m-1} + (v-s)^{m-1}. \tag{7.8}$$

Hence $(s-u)^{m-1} = (v-s)^{m-1}$ and thus $|s-u| = |v-s|$, and since u, v, s are real, $s - u = \pm(v - s)$. However, we assume $u \neq v$ and thus $s - u \neq s - v$. Consequently, $s - u = v - s$.

We may further assume $s = 1$, and thus $u = 1 - y$ and $v = 1 + y$ for some $y \neq 0$. Thus, by (7.7),

$$\hat{\Lambda}_{F;u,v,s}(x) = (x-1)\bigl((1-y)x+y\bigr)^{m-1} + \bigl(x+y\bigr)^{m-1}. \tag{7.9}$$

Since $m \geqslant 3$, $\hat{\Lambda}_{F;u,v,s}(x)$ is divisible by x^2, so the derivative $\hat{\Lambda}'_{F;u,v,s}(0) = 0$. Hence,

$$\begin{aligned} 0 = \hat{\Lambda}'_{F;u,v,s}(0) &= y^{m-1} - (m-1)(1-y)y^{m-2} + (m-1)y^{m-2} \\ &= my^{m-1} \neq 0. \end{aligned} \tag{7.10}$$

This is a contradiction, which shows that $F = K_{1,m-1}$ is good when $m \geqslant 3$. (For $m = 2$, $K_{1,1} = K_2$ is bad, as remarked above.)

Example 7.3 (disconnected graphs)**.** Suppose that $F = \bigcup_{i=1}^k F_i$ is disconnected with components $F_1, \dots, F_k$. It follows easily from (6.5) that then

$$\Lambda_{F;u,v,s}(q) = \prod_{i=1}^{k} \Lambda_{F_i;u,v,s}(q). \tag{7.11}$$

A component F_i with $|F_i| = 1$ has $\Lambda_{F_i;u,v,s}(q) = 1$ and can be ignored, as said in Remark 6.5. On the other hand, if $|F_i| \geqslant 2$, and thus $e(F_i) > 0$, then by Remark 6.4, $\Lambda_{F_i;u,v,s}(q)$ has degree at least 1 whenever $u, v, s \geqslant 0$ are not all equal. Consequently, if there are at least 2 components with more than one vertex, then $\Lambda_{F;u,v,s}(q)$ has degree at least 2, and thus F is good.

This ends our (short) list of classes of graphs that are known to be good, and completes the proof of Theorem 2.12. We can give further examples of individual small good graphs F as follows.

Example 7.4 (computer algebra)**.** Fix a graph F and consider again the four equations (6.9)–(6.12). If we set $s = 1$ (see Remark 6.7), we can eliminate a and b and obtain the two equations

$$\sum_{i=1}^{m} u^{e(F)-d_i} = (m-1)u^{e(F)} + v^{e(F)}, \tag{7.12}$$

$$\sum_{i=1}^{m} v^{e(F)-d_i} = u^{e(F)} + (m-1)v^{e(F)}. \tag{7.13}$$

Since these are two polynomial equations in two unknowns, there are plenty of complex solutions (u, v). However, if F is bad, then by Lemma 6.3 and Remark 6.7 there exists a solution with $0 < u < 1 < v$, and by symmetry another solution with $0 < v < 1 < u$. Using computer algebra (in our case `Maple`), we can check this by writing (7.12)–(7.13) as $f_1(u, v) = 0$ and $f_2(u, v) = 0$ and then computing the resultant $R(u)$ of $f_1(u, v)$ and $f_2(u, v)$ as polynomials in v. Then the roots of $R(u)$ are exactly the values u such that (7.12)–(7.13) have a solution (u, v) for some v. Hence, if F is bad, then $R(u)$ has at least one root in the interval $(0, 1)$ and at least one root in $(1, \infty)$. Consequently, if we compute the number of roots of $R(u)$ in $(0, 1)$ and in $(1, \infty)$ (by Sturm's theorem, this can be done using exact integer arithmetic), and one of these numbers is 0, then F is good.

In general, this is perhaps too much to hope for. But even if there are such roots, we can proceed by calculating the roots numerically. If the roots of $R(u)$ in $(0, 1)$ are $u_1, \dots, u_p$ and the roots in $(1, \infty)$ are $v_1, \dots, v_q$, then a solution of (7.12)–(7.13) with $0 < u < 1 < v$ has to be one of (u_i, v_j); hence, if we check the pairs (u_i, v_j) one by one and find that none satisfies both (7.12) and (7.13), then F is good. (Assuming that the computer calculations are done with enough accuracy. It might be possible to find an algorithm using exact arithmetic to test whether (7.12) and (7.13) have a common solution in $(0, 1) \times (1, \infty)$, but we have not investigated that.)

We give some explicit examples where this method succeeds.

Example 7.5 (paths)**.** The path $P_2 = K_2$ is bad, and the path $P_3 = K_{1,2}$ is good by Example 7.2. For $F = P_4$ we have $m = 4$, $e(F) = 3$ and the degree sequence $1, 2, 2, 1$. The equations (7.12)–(7.13) are $2u + 2u^2 = 3u^3 + v^3$ and $2v + 2v^2 = u^3 + 3v^3$, and the resultant $R(u) = -512\,u^9 + 1152\,u^8 + 288\,u^7 - 1872\,u^6 + 288\,u^5 + 976\,u^4 - 112\,u^3 - 192\,u^2 - 16\,u$. In this case, $R(u)$ has no roots in $(0, 1)$, so P_4 is good.

For P_5, the resolvent $R(u)$ (now of degree 16) has a single root in $(0, 1)$, but no root in $(1,\infty)$, so P_5 is good. (As an illustration, the root in $(0, 1)$ is $u = 0.23467\dots$; for this root, (7.12) and (7.13) have a common root $v = -0.65039\dots$, but no common root in $(1,\infty)$.)

We have investigated P_m for $4 \leqslant m \leqslant 20$, and the same pattern holds: For even m, the resolvent has no root in $(0, 1)$ (but one root in $(0,\infty)$). For odd m, the resolvent has no root in $(1,\infty)$ (but one root in $(0, 1)$). In both cases, P_m is good.

We conjecture that this pattern holds for all $m \geqslant 4$.

Example 7.6 (Graphs of size $|F| = 4$)**.** Of the 9 graphs with $m = |F| = 4$ and $e(F) > 1$, 3 are disconnected (Example 7.3), 2 more are regular (Example 7.1), 1 is a star (Example 7.2) and 1 is a path (Example 7.5). The two remaining ones have degree sequences $(3, 2, 2, 1)$ and $(3, 3, 2, 2)$. In both cases, the resolvent $R(u)$ has no root in $(0, 1)$. Thus every F with $|F| = 4$ and $e(F) > 1$ is good.

Example 7.7 (complete bipartite graphs)**.** We have used the method in Example 7.4 to verify that the complete bipartite graphs $K_{2,n}$ ($n \leqslant 8$), $K_{3,n}$ ($n \leqslant 7$), $K_{4,n}$ ($n \leqslant 5$) are good. In all cases, the resolvent $R(u)$ lacks roots in either $(0, 1)$ or $(1,\infty)$, and sometimes in both. (For example, for $K_{2,n}$, there is no root in $(1,\infty)$ for any $n \leqslant 8$, and a root in $(0, 1)$ only for $n = 4$ and $n = 8$. It is not clear whether this extends to larger n.)

Remark 7.8. We have so far not found any example with $e(F) > 1$ where the method in Example 7.4 fails. We thus guess that if $e(F) > 1$, then (7.12)–(7.13) have no common root with $0 < u < 1$ and $1 < v < \infty$. (Equivalently, (6.9)–(6.11) have no common root with $0 < u < s < v$.) However, note that even if there is a graph F for which this fails, F still may be good since, if $m > 3$, there are $m-3$ more equations (6.4) that have to be satisfied, which seems very unlikely. In Examples 7.4–7.7 we consider only the equations that only depend on the degree sequence.

8. More parts than vertices

Shapira and Yuster [13] and Huang and Lee [5] considered also (for $F = K_m$) the case of a partition $U_1, \dots, U_r$ of $V(G_n)$ with $r > m$, where they count the number of copies of K_m with at most one vertex in each part U_i.

We can extend this to arbitrary graphs F (as in [5, Question 5.1]). In our notation this is the same as considering (counting labelled copies and dividing by $m!$, where $m = |F|$)

$$\sum_{i_1<\cdots<i_m\leqslant r} \widetilde{N}(F,G;U_{i_1},\dots,U_{i_m}), \tag{8.1}$$

and we define the property $\widetilde{\mathcal{P}}(F;\alpha_1,\dots,\alpha_r)$ for a sequence (G_n) to mean

$$\sum_{i_1<\cdots<i_m\leqslant r} \widetilde{N}(F,G_n;U_{i_1},\dots,U_{i_m}) = p^{e(F)} \sum_{i_1<\cdots<i_m\leqslant r} \prod_{j=1}^{m} |U_{i_j}| + o\big(|G_n|)^m\big) \tag{8.2}$$

for all disjoint subsets $U_1,\dots,U_r$ of $V(G_n)$ with $|U_i| = \lfloor \alpha_i |G_n| \rfloor$, $1 \leqslant i \leqslant r$. (For $r = m$, this yields the same property as before.)

In the case $0 < p < 1$, $r \geqslant m \geqslant 3$, $F = K_m$ and $\sum_{i=1}^r \alpha_i = 1$. Shapira and Yuster [13] ($(\alpha_1,\dots,\alpha_r) \neq (1/r,\dots,1/r)$) and Huang and Lee [5] ($(\alpha_1,\dots,\alpha_r) = (1/r,\dots,1/r)$) showed that this property is p-quasi-random. We can extend this as follows.

Theorem 8.1. *Let F be a graph with $e(F) > 0$, and let $0 < p \leqslant 1$. Further, let $(\alpha_1,\dots,\alpha_r)$ be a vector of positive numbers of length $r \geqslant m = |F|$ with $\sum_{i=1}^r \alpha_i \leqslant 1$. If either $(\alpha_1,\dots,\alpha_r) \neq (1/r,\dots,1/r)$ or F is as in Theorem 2.12, then $\widetilde{\mathcal{P}}(F;\alpha_1,\dots,\alpha_r)$ is a p-quasi-random property.*

Proof. The case $(\alpha_1,\dots,\alpha_r) = (1/r,\dots,1/r)$ is simple; in this case (and more generally when all α_i are equal), it is easy to see that $\widetilde{\mathcal{P}}(F;\alpha_1,\dots,\alpha_r)$ is the same as $\widetilde{\mathcal{P}}(F_*;\alpha_1,\dots,\alpha_r)$, where F_* is the graph with r vertices obtained by adjoining $r-m$ isolated vertices to F; by Lemma 6.3 and Remark 6.5, this property is p-quasi-random if and only if $\widetilde{\mathcal{P}}(F;1/m,\dots,1/m)$ is, so the result in this case is equivalent to Theorem 2.12.

In general, we note first that Lemmas 3.1 and 3.4 extend (with the same proofs) and show that it is equivalent to consider the property of graphons

$$\sum_{i_1<\cdots<i_m} \int_{A_{i_1}\times\cdots\times A_{i_m}} \widetilde{\Psi}_{F,W}(x_{i_1},\dots,x_{i_m}) = p^{e(F)} \sum_{i_1<\cdots<i_m} \prod_{j=1}^{m} \lambda(A_{i_j}) \tag{8.3}$$

for all disjoint subsets $A_1,\dots,A_r$ of $[0,1]$ with $\lambda(A_i) = \alpha_i$.

Assume this and define

$$\widetilde{\Psi}^*_{F,W}(x_1,\dots,x_r) := \sum_{i_1<\cdots<i_m\leqslant r} \widetilde{\Psi}_{F,W}(x_{i_1},\dots,x_{i_m}) \prod_{k\notin\{i_1,\dots,i_m\}} \alpha_k^{-1}. \tag{8.4}$$

Then (8.3) can be written

$$\int_{A_1\times\cdots\times A_r} \widetilde{\Psi}^*_{F,W}(x_1,\dots,x_r) = p^{e(F)} \sum_{i_1<\cdots<i_m} \prod_{j=1}^{m} \alpha_{i_j} \tag{8.5}$$

for all such subsets $A_1,\dots,A_r$. Suppose now $(\alpha_1,\dots,\alpha_r) \neq (1/r,\dots,1/r)$. Then Lemma 4.6 applies (to $\widetilde{\Psi}^*_{F,W} - \gamma$ for a suitable constant γ) and shows that $\widetilde{\Psi}^*_{F,W}(x_1,\dots,x_r)$ is a.e. constant. Hence, if $n_1,\dots,n_m$ are integers, not all 0, then thus the Fourier coefficient $\big(\widetilde\Psi^*_{F,W}\big)\widehat{\ }(n_1,\dots,n_m,0,\dots,0)=0$. However, it follows easily from (8.4) and symmetry that this Fourier coefficient is a positive multiple of the Fourier coefficient $\big(\widetilde\Psi_{F,W}\big)\widehat{\ }(n_1,\dots,n_m)$. Hence $\big(\widetilde\Psi_{F,W}\big)\widehat{\ }(n_1,\dots,n_m)=0$ when $(n_1,\dots,n_m)\neq(0,\dots,0)$, and thus $\widetilde\Psi_{F,W}$ is a.e. constant; it follows from (8.3) that the constant must be $p^{e(F)}$. By the proof of Theorem 2.11 (or by Lemma 4.9 and Theorem 2.11), this implies $W=p$ a.e. Consequently, (8.3) for disjoint $A_1,\dots,A_r$ with $\lambda(A_i)=\alpha_i$ is a p-quasi-random property, and thus so is $\widetilde{\mathcal P}(F;\alpha_1,\dots,\alpha_r)$. $\square$

Example 8.2 (multicuts)**.** Consider the case $F=K_2$. Then the sum (8.1) is the number of edges with endpoints in two different sets U_i and U_j; we can call this a *multicut*. By Theorem 8.1, as proved already by Shapira and Yuster [13] (see also Huang and Lee [5]), the corresponding multicut property $\widetilde{\mathcal P}(K_2;\alpha_1,\dots,\alpha_r)$ is a p-quasi-random property for any $(\alpha_1,\dots,\alpha_r)\neq(1/r,\dots,1/r)$. However, $\widetilde{\mathcal P}(K_2;1/r,\dots,1/r)$ is not p-quasi-random, which is shown by the same counterexamples as for the case $r=2$ in Example 6.6.

If Conjecture 2.13 holds, then $\widetilde{\mathcal P}(K_2;1/r,\dots,1/r)$ is essentially the only case when $\widetilde{\mathcal P}(F;\alpha_1,\dots,\alpha_r)$ is not p-quasi-random.

9. Less parts than vertices

As said in Remark 2.20, it is interesting to study the subgraph counts $N(F,G_n;U_1,\dots,U_m)$ and $\widetilde N(F,G_n;U_1,\dots,U_m)$ also in situations with other restrictions on the subsets $U_1,\dots,U_m$ than the ones considered above. In particular, we may consider the case when the sets U_i may be repeated, but otherwise are disjoint. (We may also consider even more general situations when sets U_i may overlap partly in some prescribed ways, but that will not be treated here.) This suggests the following general formulation:

Let $r\geqslant 1$ and let $m_1,\dots,m_r$ be given non-negative integers with $m_1+\dots+m_r=m=|F|$, and consider for a sequence of disjoint subsets $U_1,\dots,U_r$ of $V(G)$, the following three subgraph counts:

(i) $N(F,G;U_1^{m_1},\dots,U_r^{m_r})$, defined as $N(F,G;U_1,\dots,U_r)$ where the subset U_i is repeated m_i times. (For a given labelling of F.)

(ii) $\widetilde N(F,G;U_1^{m_1},\dots,U_r^{m_r})$, defined as the average of $N(F,G;U_1^{m_1},\dots,U_r^{m_r})$ over all labellings of F. This equals, up to the constant symmetry factor $|\operatorname{aut}(F)|\prod_i m_i!/m!$, the number of copies of F in G that have exactly m_i vertices in U_i. (For each such copy of F, there are $\prod_i m_i!$ labellings of F for which it is counted, and the total number of labellings of F is $m!/|\operatorname{aut}(F)|$.)

(iii) $\widehat N(F,G;U_1,\dots,U_r;m_1,\dots,m_r)$ defined as the further average of $\widetilde N(F,G;U_1^{m_1},\dots,U_r^{m_r})$ over all permutations of $m_1,\dots,m_r$.

We then define properties $\mathcal P_{m_1,\dots,m_r}(F;\alpha_1,\dots,\alpha_r)$, $\widetilde{\mathcal P}_{m_1,\dots,m_r}(F;\alpha_1,\dots,\alpha_r)$ and $\widehat P_{m_1,\dots,m_r}(F;\alpha_1,\dots,\alpha_r)$ in analogy with Definition 2.8, considering all families of disjoint $U_1, \dots, U_r$ with $|U_i| = \lfloor \alpha_i |G_n| \rfloor$. If $F = K_m$, then $\widetilde{N} = N$ and thus $\mathcal{P}_{m_1,\dots,m_r} = \widetilde{\mathcal{P}}_{m_1,\dots,m_r}$, but in general we do not know any implication, cf. Remark 2.14.

Example 9.1. Note first that this formulation includes the problems studied earlier in the paper:

(a) For $r = m$ and $m_1 = \dots = m_r = 1$, we recover the main subject of the paper, see Section 2. (In this case, $\widehat{N} = \widetilde{N}$.)

(b) For $r > m$ and $m_i = 1$ for $1 \leqslant i \leqslant m$, $m_i = 0$ for $m+1 \leqslant i \leqslant r$, $\widehat{N}$ equals, up to an unimportant constant factor, the sum (8.1) studied in Section 8. Thus $\widehat{P}_{m_1,\dots,m_r}(F; \alpha_1, \dots, \alpha_r) = \widetilde{\mathcal{P}}(F; \alpha_1, \dots, \alpha_r)$.

(c) For $r = 1$ (and thus $m_1 = m$), we consider $N(F, G; U, \dots, U)$ as in Simonovits and Sós [15] (where $|U|$ is unspecified, see Theorem 2.5), Shapira [11] and Yuster [19].

The new case of main interest in the formulation above is thus $1 < r < m$, with $2 \leqslant m_i < m$ for some i; thus some set U_i is repeated, but all are not equal. In the remainder of this section, we consider a simple, but hopefully typical, example of this, viz. $m = 3$, $r = 2$ and $(m_1, m_2) = (2, 1)$.

Thus, assume that $m = |F| = 3$. For $\alpha, \beta > 0$ with $\alpha + \beta \leqslant 1$, the properties $\mathcal{P}_{2,1}(F; \alpha, \beta)$ and $\widetilde{\mathcal{P}}_{2,1}(F; \alpha, \beta)$ mean that (2.3) and (2.4), respectively, hold for all U_1, U_2, U_3 with $U_1 = U_2$ but disjoint from U_3, and $|U_1| = \lfloor \alpha |G_n| \rfloor$, $|U_3| = \lfloor \beta |G_n| \rfloor$. In the case $\alpha + \beta = 1$, we can equivalently assume that $U_1 = U_2 = U$ and $U_3 = V(G_n) \setminus U$ with $|U| = \lfloor \alpha |G_n| \rfloor$. (For $F = K_3$, this means that we count triangles crossing the the cut $(U, V(G_n) \setminus U)$, with exactly two vertices in U.) Are these properties p-quasi-random?

The analogue of Lemma 3.1 holds, and thus we can as in Lemma 3.4 transfer the problem to graphons and the properties defined by (3.5) or (3.6) for all A_1, A_2, A_3 with $A_1 = A_2$ and disjoint from A_3, and $\lambda(A_1) = \alpha$, $\lambda(A_3) = \beta$.

Consider first $\widetilde{\mathcal{P}}_{2,1}(F; \alpha, \beta)$. In the case $\alpha + \beta < 1$, we have the following results, similar to the ones above.

Lemma 9.2. *Let $\alpha, \beta > 0$ with $\alpha + \beta < 1$. Suppose that $f : [0,1]^3 \to \mathbb{C}$ is a symmetric integrable function such that*

$$\int_{A \times A \times B} f = 0 \tag{9.1}$$

for all disjoint subsets A and B of $[0,1]$ such that $\lambda(A) = \alpha$ and $\lambda(B) = \beta$. Then

$$f(x_1, \dots, x_m) = 0 \qquad a.e. \tag{9.2}$$

Proof. A minor variation of the proof of Lemma 4.6, using Janson [6, Lemma 7.6]. We omit the details. ∎

Theorem 9.3. *Let F be a graph with $|F| = 3$ and $e(F) > 0$, let $\alpha, \beta > 0$ with $\alpha + \beta < 1$ and let $0 < p \leqslant 1$. Then $\widetilde{\mathcal{P}}_{2,1}(F; \alpha, \beta)$ is a quasi-random property.*

Proof. Using Lemma 9.2, we argue as in the proof of Theorem 2.11 in Section 6. ∎

The case $\alpha+\beta = 1$, and thus $B = A^{\mathsf{c}}$ in (9.1), is more intricate, and therefore more interesting. We note first that the counterexample in Lemma 4.2 shows that Lemma 9.2 does not hold for $\alpha = 1-\beta = 2/3$. In fact, if $f(x_1,x_2,x_3) = g(x_1)+g(x_2)+g(x_3)$ with $\int_0^1 g = 0$ and $|A| = \alpha$, then

$$\begin{aligned}\int_{A\times A\times A^{\mathsf{c}}} f(x_1,x_2,x_3) &= 2\alpha(1-\alpha)\int_A g + \alpha^2\int_{A^{\mathsf{c}}} g\\ &= 2\alpha(1-\alpha)\int_A g - \alpha^2\int_A g = \alpha(2-3\alpha)\int_A g, \qquad (9.3)\end{aligned}$$

which vanishes for every such A if $\alpha = 2/3$.

Moreover, there is another counterexample for $\alpha = 1-\beta = 1/3$: Now consider $f(x_1,x_2,x_3) = g(x_1,x_2)+g(x_1,x_3)+g(x_2,x_3)$ for a symmetric function g on $[0,1]^2$ such that $\int_0^1 g(x,y)\,\mathrm{d}y = 0$ for every x. Then

$$\begin{aligned}\int_{A\times A\times A^{\mathsf{c}}} f(x_1,x_2,x_3) &= (1-\alpha)\int_{A^2} g + 2\alpha\int_{A\times A^{\mathsf{c}}} g\\ &= (1-\alpha)\int_{A^2} g - 2\alpha\int_{A^2} g = (1-3\alpha)\int_{A^2} g, \qquad (9.4)\end{aligned}$$

which vanishes if $\alpha = 1/3$.

We conjecture that these are the only counterexamples.

Conjecture 9.4. *Let $\alpha \in (0,1)$ and suppose that $f : [0,1]^3 \to \mathbb{C}$ is a symmetric integrable function such that $\int_{A\times A\times A^{\mathsf{c}}} f = 0$ for every $A \subset [0,1]$ with $\lambda(A) = \alpha$.*

(i) *If $\alpha \notin \{\frac13, \frac23\}$, then $f = 0$ a.e.*
(ii) *If $\alpha = \frac13$, then $f(x_1,x_2,x_3) = g(x_1,x_2)+g(x_1,x_3)+g(x_2,x_3)$ a.e. for a symmetric function g on $[0,1]^2$ such that $\int_0^1 g(x,y)\,\mathrm{d}y = 0$ for every x.*
(iii) *If $\alpha = \frac23$, then $f(x_1,x_2,x_3) = g(x_1)+g(x_2)+g(x_3)$ a.e. for a function g on $[0,1]$ such that $\int_0^1 g(x)\,\mathrm{d}x = 0$.*

We leave this (and extensions to $m > 3$) as an open problem. Note that if this conjecture holds, then Theorem 9.3 holds also for $\alpha+\beta = 1$, provided $\alpha \neq \frac13, \frac23$, by the same proof as above. For $\alpha = 1-\beta = \frac23$ we would have the same situation as in Lemmas 4.3 and 6.3; from the discussion in Section 7 follows that Theorem 9.3 would hold if $e(F) \geqslant 2$ (P_2 or K_3), but not for $e(F) \leqslant 1$ ($K_2 \cup K_1$ and the trivial empty graph $K_1\cup K_1\cup K_1$). (Recall that we only consider $m = 3$, as an example.)

For $\alpha = 1-\beta = \frac13$, even if the conjecture holds, it would lead to further open problems: First, is there an analogue of Theorems 5.2 and 5.3 for this case, showing that if the property is not quasi-random, then there is a 2-type graphon counterexample? (This seems likely if Conjecture 9.4 holds, using a suitable analogue of Lemma 5.8 for this case.) Secondly, analysis of a possible 2-type graphon counterexample would lead to a different algebraic problem than the one in Section 6; we leave the formulation and investigations of this as another open problem.

Problem 9.5. Solve these problems for the case $\beta = 1-\alpha$, with particular attention to the cases $\alpha = \frac{1}{3}$ and $\frac{2}{3}$, in particular for $F = K_3$ (crossing triangles). Moreover, consider extensions for $m > 3$.

Remark 9.6. Note that the set of functions satisfying the condition of Lemma 4.1, 4.3 or 9.2, or Conjecture 9.4, is invariant under all measure preserving bijections of $[0,1]$. This suggest the following approach, where we consider only square integrable functions so that we can use Hilbert space theory. Let, for $0 \leqslant k \leqslant m$, $H^{m,k}$ be the subspace of $L^2([0,1]^m)$ consisting of all functions f such that the Fourier coefficient $\hat{f}(n_1,\dots,n_m)$ vanishes unless exactly k indices $n_1,\dots,n_m$ are non-zero. (In particular, $H^{m,0}$ is the space of constant functions.) Let further $L^2_{\mathsf{s}}([0,1]^m)$ be the subspace of symmetric functions in $L^2([0,1]^m)$, and let $H^{m,k}_{\mathsf{s}} := H^{m,k} \cap L^2_{\mathsf{s}}([0,1]^m)$. Then

$$L^2_{\mathsf{s}}([0,1]^m) = \bigoplus_{k=0}^{m} H^{m,k}_{\mathsf{s}} \tag{9.5}$$

and each subspace $H^{m,k}_{\mathsf{s}}$ is invariant under measure preserving bijections of $[0,1]$. We conjecture that every closed subspace of $L^2_{\mathsf{s}}([0,1]^m)$ invariant under all measure preserving bijections of $[0,1]$ is of the form $\bigoplus_{k\in A} H^{m,k}_{\mathsf{s}}$ for some set $A \subseteq \{0,\dots,m\}$.

If this holds, it is easy to verify Conjecture 9.4.

In support of this conjecture, note that a discrete analogue holds: Let $N \geqslant m > 0$ and consider the set $X_{N,m}$ of m-tuples of distinct elements of $[N]$. If $N \geqslant 2m$, then the natural representation of the symmetric group S_N in the $\binom{N}{m}$-dimensional space of all symmetric functions on $X_{N,m}$ has $m+1$ irreducible components, which correspond to the sets $H^{m,k}_{\mathsf{s}}$ above. (This is easily verified by a calculation with the characters of these representations. We omit the details.)

Finally, for the property $\mathcal{P}_{2,1}(F;\alpha,\beta)$, for a directed graph F with $|F| = 3$, we have the same problems as before (unless $F = K_3$), see Remark 2.14. Consider for example $F = P_3$. We may note that in Lemma 9.2, it suffices that f is symmetric in the first two variables; this implies by the argument above that if $F = P_3$ with the central vertex labelled 3, then $\mathcal{P}_{2,1}(F;\alpha,\beta)$ is quasi-random (since then $\Psi_{F,W}$ is symmetric in the first two variables). However, this argument fails for the other labellings of P_3. The case $\alpha+\beta = 1$ seems even more complicated.

Problem 9.7. Is $\mathcal{P}_{2,1}(P_3;\alpha,\beta)$ a quasi-random property for any $\alpha,\beta > 0$ with $\alpha+\beta < 1$, for any labelling of P_3? Does this hold for $\alpha+\beta = 1$?

References

[1] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós & K. Vesztergombi, Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing, *Advances in Math.* **219** (2008), 1801–1851.

[2] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós & K. Vesztergombi, Convergent sequences of dense graphs II. Multiway cuts and statistical physics. *Ann. of Math. (2)* **176** (2012), no. 1, 151–219.

[3] F. R. K. Chung & R. L. Graham, Maximum cuts and quasirandom graphs. *Random graphs, Vol. 2 (Poznań, 1989)*, 23–33, Wiley, New York, 1992.

[4] F. R. K. Chung, R. L. Graham & R. M. Wilson, Quasi-random graphs. *Combinatorica* **9** (1989), no. 4, 345–362.

[5] H. Huang & C. Lee, Quasi-randomness of graph balanced cut properties. *Random Structures Algorithms* **41** (2012), no. 1, 124–145.

[6] S. Janson, Quasi-random graphs and graph limits. *European J. Combin.* **32** (2011), 1054–1083.

[7] L. Lovász, *Large Networks and Graph Limits.* American Mathematical Society, Providence, RI, 2012.

[8] L. Lovász & B. Szegedy, Limits of dense graph sequences. *J. Comb. Theory B* **96** (2006), 933–957.

[9] L. Lovász & B. Szegedy, Szemerédi's lemma for the analyst. *Geom. Funct. Anal.* **17** (2007), no. 1, 252–270.

[10] F. Petrov, General removal lemma. `arXiv:1309.3795v1.`

[11] A. Shapira, Quasi-randomness and the distribution of copies of a fixed graph. *Combinatorica* **28** (2008), 735–745.

[12] A. Shapira & R. Yuster, The effect of induced subgraphs on quasi-randomness. *Random Struct Algorithms* **36** (2010), 90–109.

[13] A. Shapira & R. Yuster, The quasi-randomness of hypergraph cut properties. *Random Structures Algorithms* **40** (2012), no. 1, 105–131.

[14] M. Simonovits & V. T. Sós, Szemerédi's partition and quasirandomness. *Random Structures Algorithms* **2** (1991), no. 1, 1–10.

[15] M. Simonovits & V. T. Sós, Hereditarily extended properties, quasi-random graphs and not necessarily induced subgraphs. *Combinatorica* **17** (1997), no. 4, 577–596.

[16] M. Simonovits & V. T. Sós, Hereditary extended properties, quasi-random graphs and induced subgraphs. *Combin. Probab. Comput.* **12** (2003), no. 3, 319–344.

[17] A. Thomason, Pseudorandom graphs. *Random graphs '85 (Poznań, 1985)*, 307–331, North-Holland, Amsterdam, 1987.

[18] A. Thomason, Random graphs, strongly regular graphs and pseudo-random graphs. *Surveys in Combinatorics 1987 (New Cross, 1987)*, 173–195, London Math. Soc. Lecture Note Ser. **123**, Cambridge Univ. Press, Cambridge, 1987.

[19] R. Yuster, Quasi-randomness is determined by the distribution of copies of a fixed graph in equicardinal large sets. *Approximation, Randomization and Combinatorial Optimization*, 596–601, Lecture Notes in Comput. Sci. **5171**, Springer, Berlin, 2008.

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden
E-mail address: svante.janson@math.uu.se
URL: http://www2.math.uu.se/∼svante/

A. Rényi Institute of Mathematics, Budapest, Hungary
E-mail address: t.sos.vera@renyi.mta.hu