Quasi-randomness of graph balanced cut properties

Hao Huang [*] Choongbum Lee [†]

Abstract

Quasi-random graphs can be informally described as graphs whose edge distribution closely resembles that of a truly random graph of the same edge density. Recently, Shapira and Yuster proved the following result on quasi-randomness of graphs. Let $k \geq 2$ be a fixed integer, $\alpha_1, \ldots, \alpha_k$ be positive reals satisfying $\sum_i \alpha_i = 1$ and $(\alpha_1, \ldots, \alpha_k) \neq (1/k, \ldots, 1/k)$, and G be a graph on n vertices. If for every partition of the vertices of G into sets $V_1, \ldots, V_k$ of size $\alpha_1 n, \ldots, \alpha_k n$, the number of complete graphs on k vertices which have exactly one vertex in each of these sets is similar to what we would expect in a random graph, then the graph is quasi-random.

However, the method of quasi-random hypergraphs they used did not provide enough information to make any conclusion for the case $(1/k, \ldots, 1/k)$. In their work, Shapira and Yuster asked whether this case also forces the graph to be quasi-random. Janson also posed the same question in his study of quasi-randomness under the framework of graph limits. In this paper, we positively answer this question by proving that the case $(1/k, \ldots, 1/k)$ implies the quasi-randomness of the graph as well.

1 Introduction

The study of random structures has seen a tremendous success in modern combinatorics and theoretical computer science. One example is the Erdős-Rényi random graph $G(n, p)$ proposed in the 1960's and intensively studied thereafter. $G(n, p)$ is the probability space of graphs over n vertices where each pair of vertices forms an edge independently with probability p. Random graphs are not only an interesting object of study on their own but also proved to be a powerful tool in solving numerous open problems. The success of random structures served as a natural motivation for the following question: How can one tell when a given structure behaves like a random one? Such structures are called *quasi-random*. In this paper we study quasi-random graphs, which, following Thomason [17, 18], can be informally defined as graphs whose edge distribution closely resembles that of a random graph (the formal definition will be given later). One fundamental result in the study of quasi-random graphs is the following theorem proved by Chung, Graham and Wilson [2].

Theorem 1.1 *Fix a real $p \in (0, 1)$. For a n-vertex graph G, define $e(U)$ to be the number of edges in the induced subgraph spanned by vertex set U, then the following properties are equivalent.*

$\mathcal{P}_1$*: For any subset of vertices $U \subset V(G)$, we have $e(U) = \frac{1}{2}p|U|^2 \pm o(n^2)$.*

$\mathcal{P}_2(\alpha)$*: For any subset of vertices $U \subset V(G)$ of size αn, we have $e(U) = \frac{1}{2}p|U|^2 \pm o(n^2)$.*

$\mathcal{P}_3$*: $e(G) = \frac{1}{2}pn^2 \pm o(n^2)$ and G has $\frac{1}{8}p^4 n^4 \pm o(n^4)$ cycles of length 4.*

[*]Department of Mathematics, UCLA, Los Angeles, CA, 90095. Email: huanghao@math.ucla.edu.

[†]Department of Mathematics, UCLA, Los Angeles, CA 90095. Email: abdesire@math.ucla.edu. Research supported in part by Samsung Scholarship.

Throught out this paper, unless specified otherwise, when considering a subset of vertices $U \subset V$ such that $|U| = \alpha n$ for some α, we tacitly assume that $|U| = \lfloor \alpha n \rfloor$ or $|U| = \lceil \alpha n \rceil$. Since we mostly consider asymptotic values, this difference will not affect our calculation.

For a positive real δ, we say that a graph G is *δ-close to satisfying* $\mathcal{P}_1$ if $e(U) = \frac{1}{2}p|U|^2 \pm \delta n^2$ for all $U \subset V(G)$, and similarly define it for other properties. The formal definition of equivalence of properties in Theorem 1.1 is as following: for every $\varepsilon > 0$, there exists a δ such that if a graph is δ-close to satisfying one property, then it is ε-close to satisfying another.

We call a graph *p-quasi-random*, or *quasi-random* if the density p is clear from the context, if it satisfies $\mathcal{P}_1$, and consequently satisfies all of the equivalent properties of Theorem 1.1. We also say that a graph property is *quasi-random* if it is equivalent to $\mathcal{P}_1$. Note that the random graph $G(n,p)$ with high probability is p-quasi-random. However, it is not true that all the properties of random graphs are quasi-random. For example, it is easy to check that the property of having $\frac{1}{2}pn^2 + o(n^2)$ edges is not quasi-random (as an instance, there can be many isolated vertices). For more details on quasi-random graphs we refer the reader to the survey of Krivelevich and Sudakov [12]. Quasi-randomness were also studied in many other settings besides graphs, such as set systems [3], tournaments [4] and hypergraphs [5].

The main objective of our paper is to study the quasi-randomness of graph properties given by certain graph cuts. These kind of properties were first studied by Chung and Graham in [3, 6]. For a real $\alpha \in (0,1)$, the cut property $\mathcal{P}_C(\alpha)$ is the collection of graphs G satisfying the following: for any $U \subset V(G)$ of size $|U| = \alpha n$, we have $e(U, V \backslash U) = p\alpha(1-\alpha)n^2 + o(n^2)$. As it turns out, for most values of α, the cut property $\mathcal{P}_C(\alpha)$ is quasi-random. In [3, 6], the authors proved the following beautiful theorem which characterizes the quasi-random cut properties.

Theorem 1.2 *$\mathcal{P}_C(\alpha)$ is quasi-random if and only if $\alpha \neq 1/2$.*

A cut is a partition of a vertex set V into subsets $V_1, \cdots, V_r$, and if for a vector $\vec{\alpha} = (\alpha_1, \cdots, \alpha_r)$, the size of the sets satisfies $|V_i| = \alpha_i |V|$ for all i, then we call this an $\vec{\alpha}$-cut. An $\vec{\alpha}$-cut is called *balanced* if $\vec{\alpha} = (1/r, \cdots, 1/r)$ for some r, and is *unbalanced* otherwise. For a k-uniform hypergraph G and a cut $V_1, \cdots, V_r$ of its vertex set, let $e(V_1, \cdots, V_r)$ be the number of hyperedges which have at most one vertex in each part V_i for all i.

A k-uniform hypergraph G is *p-quasi-random* if for every subset of vertices $U \subset V(G)$, $e(U) = p\frac{|U|^k}{k!} \pm o(n^k)$. Let $\mathcal{P}_C(\vec{\alpha})$ be the following property: for every $\vec{\alpha}$-cut $V_1, \ldots, V_r$, $e(V_1, \cdots, V_r) = (p + o(1))n^k \sum_{S \subset [r], |S| = k} \prod_{i \in S} \alpha_i$. Shapira and Yuster [14] generalized Theorem 1.2 by proving the following theorem.

Theorem 1.3 *Let $k \geq 2$ be a positive integer. For k-uniform hypergraphs, the cut property $\mathcal{P}_C(\vec{\alpha})$ is quasi-random if and only if $\vec{\alpha} \neq (1/r, \ldots, 1/r)$ for some $r \geq k$.*

For a fixed graph H, let $\mathcal{P}_H$ be the following property : for every subset $U \subset V$, the number of copies of H in U is $(p^{|E(H)|} + o(1))\binom{|U|}{|V(H)|}$. In [15], Simonovits and Sós proved that $\mathcal{P}_H$ is equivalent to $\mathcal{P}_1$ and hence is quasi-random. For a fixed graph H, as a common generalization of Chung and Graham's and Simonovits and Sós' Theorems, we can consider the number of copies H having one vertex in each part of a cut. Let us consider the cases when H is a clique of size k.

Definition 1.4 *Let k, r be positive integers such that $r \geq k \geq 2$, and let $\vec{\alpha} = (\alpha_1, \cdots, \alpha_r)$ be a vector of positive real numbers satisfying $\sum_{i=1}^{r} \alpha_i = 1$. We say that a graph satisfies the K_k cut property $\mathcal{C}_k(\vec{\alpha})$ if for every $\vec{\alpha}$-cut $(V_1, \cdots, V_r)$, the number of copies of K_k which have at most one vertex in each of the sets V_i is $(p^{\binom{k}{2}} \pm o(1))n^k \sum_{S \subset [r], |S| = k} \prod_{i \in S} \alpha_i$.*

Shapira and Yuster [14] proved that for $k \geq 3$, $\mathcal{C}_k(\vec{\alpha})$ is quasi-random if $\vec{\alpha}$ is unbalanced (note that $\mathcal{C}_2(\vec{\alpha})$ is quasi-random if and only if $\vec{\alpha}$ is unbalanced). This result is a corollary of Theorem 1.3 by the following argument. For a graph G satisfying $\mathcal{C}_k(\vec{\alpha})$, consider the k-uniform hypergraph G' on the same vertex set where a k-tuple of vertices forms an hyperedge if and only if they form a clique in G. Then G' satisfies $\mathcal{P}_C(\vec{\alpha})$ and thus is quasi-random. By the definition of the quasi-randomness of hypergraphs, this in turn implies that the number of cliques of size k inside every subset of $V(G)$ is "correct", and thus by Simonovits and Sós' result, G is quasi-random.

Note that for balanced $\vec{\alpha}$ this approach does not give any information, since it is not clear if there exists a graph whose hypergraph constructed by the above mentioned process is not quasi-random but satisfies $\mathcal{P}_C(\vec{\alpha})$ (nonetheless as the reader might suspect, the properties $\mathcal{P}_C(\vec{\alpha})$ and $\mathcal{C}_k(\vec{\alpha})$ are closely related even for balanced $\vec{\alpha}$). Shapira and Yuster made this observation and left the balanced case as an open question asking whether it is quasi-random or not. Janson [10] independently posed the same question in his paper that studied quasi-randomness under the framework of graph limits. In this paper, we settle this question by proving the following theorem:

Theorem 1.5 *Fix a real $p \in (0,1)$ and positive integers r, k such that $r \geq k \geq 3$. For every $\varepsilon > 0$, there exists a positive real δ such that the following is true. If G is a graph which has density p and is δ-close to satisfying the K_k balanced cut property $\mathcal{C}_k(1/r, \cdots, 1/r)$, then G is ε-close to being p-quasi-random.*

The rest of the paper is organized as follows. In Section 2 we introduce the notations we are going to use throughout the paper and state previously known results that we need later. In Section 3 we give a detailed proof of the most important base case of Theorem 1.5, triangle balanced cut property, i.e. $\mathcal{C}_3(1/r, \cdots, 1/r)$. In Section 4, we prove the general case as a consequence of the base case. The last section contains some concluding remarks and open problems for further study.

2 Preliminaries

Given a graph $G = (V, E)$ and two vertex sets $X, Y \subset V(G)$, we denote by $E(X, Y)$ the set of edges which have one end point in X and the other in Y. Also we write $e(X, Y) = |E(X, Y)|$ to indicate the number of edges and $d(X, Y) = \frac{e(X,Y)}{|X||Y|}$ for the density. For a cut $\mathbf{X} = (X_1, \cdots, X_r)$ of the vertex set, we say that a triangle with vertices u, v, w *crosses the cut* $\mathbf{X}$ if it contains at most one vertex from each set, and denote it by $(u, v, w) \pitchfork \mathbf{X}$. We use $Tr(\mathbf{X})$ for the number of triangles with vertices $(u, v, w) \pitchfork \mathbf{X}$. For a k-uniform hypergraph and a partition $V_1, \ldots, V_t$ of its vertex set V into t parts, we define its *density vector* as the vector in $\mathbb{R}^{\binom{t}{k}}$ indexed by the k-subsets of $[t]$ whose $\{i_1, \cdots, i_k\}$-entry is the density of hyperedges which have exactly one vertex in each of the sets $V_{i_1}, \cdots, V_{i_k}$. Throughout the paper, we always use subscripts such as $\delta_{2.6}$ to indicate that the parameter δ comes from Theorem 2.6.

To state asymptotic results, we utilize the following standard notations. For two functions $f(n)$ and $g(n)$, write $f(n) = \Omega(g(n))$ if there exists a positive constant c such that $\liminf_{n\to\infty} f(n)/g(n) \geq c$, $f(n) = o(g(n))$ if $\limsup_{n\to\infty} f(n)/g(n) = 0$. Also, $f(n) = O(g(n))$ if there exists a positive constant $C > 0$ such that $\limsup_{n\to\infty} f(n)/g(n) \leq C$.

To isolate the unnecessary complication arising from the error terms, we will use the notation $x =_\varepsilon y$ if $|x - y| = O(\varepsilon)$ and say that x, y are ε-equal. For two vectors, we define $\vec{x} =_\varepsilon \vec{y}$ if $\|\vec{x} - \vec{y}\|_\infty = O(\varepsilon)$. We omit the proof of the following simple properties (we implicitly assume that the following operations are performed a constant number of times in total). Let C and c be positive constants.

(1a) (Finite transitivity) If $x =_\varepsilon y$ and $y =_\varepsilon z$, then $x =_\varepsilon z$.
(1b) (Complete transitivity) For a finite set of numbers $\{x_1, \cdots, x_n\}$. If $x_i =_\varepsilon x_j$ for every i, j, then there exists x such that $x_i =_\varepsilon x$ for all i.
(2) (Additivity) If $x =_\varepsilon z$ and $y =_\varepsilon w$, then $x + y =_\varepsilon z + w$.
(3) (Scalar product) If $x =_\varepsilon y$ and $0 < c \le a \le C$, then $ax =_\varepsilon ay$ and $x/a =_\varepsilon y/a$.
(4) (Product) If x, y, z, w are bounded above by C, then $x =_\varepsilon y$ and $z =_\varepsilon w$ implies that $xz =_\varepsilon yw$.
(5) (Square root) If both x and y are greater than c, then $x^2 =_\varepsilon y^2$ implies that $x =_\varepsilon y$.
(6) For the linear equation $A\vec{x} =_\varepsilon \vec{y}$, if all the entries of an invertible matrix A are bounded by C, and the determinant of A is bounded from below by c, then $\vec{x} =_\varepsilon A^{-1}\vec{y}$.
(7) If $xy =_\varepsilon 0$, then either $x =_{\sqrt{\varepsilon}} 0$ or $y =_{\sqrt{\varepsilon}} 0$.

2.1 Extremal Graph Theory

To prove the main theorem, we use the regularity lemma developed by Szemerèdi [16]. Let $G = (V, E)$ be a graph and $\varepsilon > 0$ be fixed. A disjoint pair of sets $X, Y \subset V$ is called an *ε-regular pair* if $\forall A \subset X, B \subset Y$ such that $|A| \ge \varepsilon|X|, |B| \ge \varepsilon|X|$ satisfies $|d(X, Y) - d(A, B)| \le \varepsilon$. A vertex partition $\{V_i\}_{i=1}^t$ is called an *ε-regular partition* if (i) V_i have equal size for all i, and (ii) (V_i, V_j) is ε-regular for all but at most εt^2 pairs $1 \le i < j \le n$. The regularity lemma states that every large enough graph admits a regular partition. In our proof, we use a slightly different form which can be found in [11]:

Theorem 2.1 (Regularity Lemma) *For every real $\varepsilon > 0$ and positive integers m, r there exists constants $T(\varepsilon, m)$ and $N(\varepsilon, m)$ such that given any $n \ge N(\varepsilon, m)$, the vertex set of any n-vertex graph G can be partitioned into t sets $V_1, \cdots, V_t$ for some t divisible by r and satisfying $m \le t \le T(\varepsilon, m)$, so that*

- *$|V_i| < \lceil \varepsilon n \rceil$ for every i.*
- *$||V_i| - |V_j|| \le 1$ for all i, j.*
- *Construct a reduced graph H on t vertices such that $i \sim j$ in H if and only if (V_i, V_j) is ε-regular in G. Then the reduced graph has minimum degree at least $(1 - \varepsilon)t$.*

As one can see in the following lemma, regular pairs are useful in counting small subgraphs of a graph (this lemma can easily be generalized to other subgraphs).

Lemma 2.2 *Let V_1, V_2, V_3 be subsets of vertices. If the pair (V_i, V_j) is ε-regular with density d_{ij} for every distinct i, j, then the number of triangles $Tr(V_1, V_2, V_3)$ is*

$$Tr(V_1, V_2, V_3) = (d_{12}d_{23}d_{31} + O(\varepsilon))|V_1||V_2||V_3|.$$

Proof. If a vertex $v \in V_1$ has degree $(1 + O(\varepsilon))d_{12}|V_2|$ in V_2 and $(1 + O(\varepsilon))d_{13}|V_3|$ in V_3, then by the regularity of the pair (V_2, V_3), there will be $(1 + O(\varepsilon))|V_2||V_3|d_{12}d_{23}d_{31}$ triangles which contain the vertex v. By the regularity of the pairs (V_1, V_2) and (V_1, V_3), there are at least $(1 - 2\varepsilon)|V_1|$ such vertices. Moreover, there are at most $2\varepsilon|V_1||V_2||V_3|$ triangles which does not contain such vertex from V_1. Therefore we have,

$$Tr(V_1, V_2, V_3) = (1 + O(\varepsilon))|V_1||V_2||V_3|d_{12}d_{23}d_{31} + 2\varepsilon|V_1||V_2||V_3| = (d_{12}d_{23}d_{31} + O(\varepsilon))|V_1||V_2||V_3|.$$

□

For a fixed graph H, a perfect H-factor of a large graph G is a collection of vertex disjoint copies of H that cover all the vertices of G. The next theorem is a classical theorem proved by

Hajnal and Szemerédi [8] which establishes a sufficient minimum degree condition for the existence of a perfect clique factor.

Theorem 2.3 *Let k be a fixed positive integer and n be divisible by k. If G is a graph on n vertices with minimum degree at least $(1-1/k)n$, then G contains a perfect K_k-factor.*

2.2 Concentration

The following concentration result of Hoeffding [9] and Azuma [1] will be used several times during the proof (see also [13, Theorem 3.10]).

Theorem 2.4 (Hoeffding-Azuma Inequality) *Let $c_1, \ldots, c_n$ be constants, and let $X_1, \ldots, X_n$ be a martingale sequence with $|X_k| \le c_k$ for each k. Then for any $t \ge 0$,*

$$Pr\left(\left|\sum_{i=1}^{n} X_i\right| \ge t\right) \le 2\exp\left(-\frac{t^2}{2\sum_{i=1}^{n} c_i^2}\right)$$

The next lemma is a corollary of Hoeffding-Azuma's inequality.

Lemma 2.5 *Let $G=(V,E)$ be a graph with $|V|=n$ and $|E| = d\binom{n}{2}$ for some real d. Let U be a random subset of V constructed by selecting every vertex independently with probability α. Then $e(U) = \alpha^2 d\binom{n}{2} + o(n^2)$ with probability at least $1-e^{-O(n^{1/2})}$.*

Proof. Arbitrarily label the veritces by $1, \ldots, n$ and consider the vertex exposure martingale. More precisely, let X_k be the number of edges within U incident to k among the vertices $1, \ldots, k-1$ ($X_k = 0$ if $k \notin U$). Then $e(U) = X_1 + \cdots + X_n$ and $\big(X_1 + \cdots + X_k - \mathbb{E}[X_1 + \cdots + X_k]\big)_k$ forms a martingale such that $|X_k - \mathbb{E}[X_k]| \le n$ for all k. Thus by Hoeffding-Azuma's inequality (Theorem 2.4),

$$Pr(|e(U) - \mathbb{E}[e(U)]| \ge C) \le 2e^{-2C^2/n^3}.$$

Since $\mathbb{E}[e(U)] = \alpha^2 d\binom{n}{2}$, by selecting $C = n^{7/4}$, we obtain $e(U) = \alpha^2 d\binom{n}{2} + o(n^2)$ with probability at least $1-e^{-O(n^{1/2})}$. $\square$

2.3 Quasi-randomness of hypergraph cut properties

Recall the cut property $\mathcal{P}_C(\vec{\alpha})$ defined in the introduction, and the fact that it is closely related to the clique cut property $\mathcal{C}_k(\vec{\alpha})$. While proving Theorem 1.3, Shapira and Yuster also characterized the structure of hypergraphs which satisfies the balanced cut property $\mathcal{P}_C(1/r, \cdots, 1/r)$. Let $p \in (0,1)$ be fixed and t be an integer. In order to classify the k-uniform hypergraphs satisfying the balanced cut property, we first look at certain edge-weighted hypergraphs. Fix a set $I \subset [t]$ of size $|I| = t/2$, and consider the weighted hypergraph on the vertex set $[t]$ such that the hyperedge e has density $2p|e \cap I|/k$ for all e. Let $\mathbf{u}_{t,p,I}$ be the vector representing this weighted hypergraph, and let $W_{t,p}$ be the affine subspace of $\mathbb{R}^{\binom{t}{k}}$ spanned by the vectors $\mathbf{u}_{t,p,I}$ for all possible sets I of size $|I| = t/2$. In [14], the authors proved that the structure of a (non-weighted) hypergraph which is δ-close to satisfying the balanced cut property $\mathcal{P}_C(1/r, \cdots, 1/r)$ can be described by the vector space $W_{t,p}$ (note that the vector which has constant weight lies in this space).

Theorem 2.6 *Let $p \in (0,1)$ be fixed. There exists a real t_0 such that For every $\varepsilon > 0$, and for every $t \ge t_0$ divisible by $2r$* [1]*, there exists $\delta = \delta(t, \varepsilon) > 0$ so that the following holds. If G is a k-uniform*

[1] The authors omiitted the divisibility condition in their paper [14].

hypergraph with density p which is δ-close to satisfying the balanced cut property $\mathcal{P}_C(1/r, \cdots, 1/r)$, then for any partition of $V(G)$ into t equal parts, the density vector $\mathbf{d}$ of this partition satisfies $\|\mathbf{d} - \mathbf{y}\|_\infty \leq \varepsilon$ for some vector $\mathbf{y} \in W_{t,p}$.

A part of the proof of Shapira and Yuster's theorem relies on showing that certain matrices have full rank, and they establish this result by using the following famous result from algebraic combinatorics proved by Gottlieb [7]. For a finite set T and integers h and k satisfying $|T| > h \geq k \geq 2$, denote by $B(T, h, k)$ the h versus k inclusion matrix of T which is the $\binom{|T|}{h} \times \binom{|T|}{k}$ 0-1 matrix whose rows are indexed by the h-element subsets of T, columns are indexed by the k-elements subsets of T, and entry (I, J) is 1 if and only if $J \subset I$.

Theorem 2.7 *$rank(B(T, h, k)) = \binom{|T|}{k}$ for all $|T| \geq h + k$.*

3 Base case - Triangle Balanced Cut

In this section we prove a special case, triangle balanced cut property, of the main theorem. Our proof consists of several steps. Let G be a graph which satisfies the triangle balanced cut property. First we apply the regularity lemma to describe the structure of G by an ε-regular partition $\{V_i\}_{i=1}^t$. This step allows us to count the edges or triangles effectively using regularity of the pairs. From this point on, we focus only on the cuts whose parts consist of a union of the sets V_i. In the next step, we swap some vertices of V_i and V_j. By the triangle cut property, we can obtain an algebraic relation of the densities inside V_i and between V_i and V_j. After doing this, the problem is transformed into solving a system of nonlinear equations, which basically implies that inside any clique of the reduced graph, most of the densities are very close to each other. Finally resorting to results from extremal graph theory, we can conclude that almost all the densities are equal and thus prove the quasi-randomness of triangle balanced cut property.

Theorem 3.1 *Fix a real $p \in (0, 1)$ and an integer $r \geq 3$. For every $\varepsilon > 0$, there exists a positive real δ such that the following is true. If G is a graph which has density p and is δ-close to satisfying the triangle balanced cut property $\mathcal{C}_3(1/r, \cdots, 1/r)$, then G is ε-close to being p-quasi-random.*

Let G be a graph δ-close to satisfying $\mathcal{C}_3(1/r, \cdots, 1/r)$. By applying the regularity lemma, Theorem 2.1, to G, we get an ε-regular equipartition $\pi = \{V_i\}_{i=1}^t$. We can assume that $|V_1| = \cdots = |V_t|$ by deleting at most t vertices. The reason this can be done is that later when we use the triangle cut property to count the number of triangles, the error term that this deletion creates is at most tn^2 which is negligible comparing to δn^3 when n is sufficiently large. Also in the definition of quasi-randomness, the error term from counting edges is at most tn, which is also $o(n^2)$.

Now denote the edge density within V_i by x_i, the edge density between V_i and V_j by d_{ij}, and the density of triangles in the tripartite graph formed by (V_i, V_j, V_k) by d_{ijk}. Call a triple (V_i, V_j, V_k) *regular* if each of the three pairs is regular.

Consider a family $\{\pi_\alpha\}_{\alpha \in [0,1]}$ of partitions of G given as follows:

$$\pi_\alpha = ((1 - \alpha)V_1 + \alpha V_2, \alpha V_1 + (1 - \alpha)V_2, V_3, \cdots V_t).$$

In other words, we pick U_1 and U_2 both containing α-proportion of vertices in V_1 and V_2 uniformly at random and exchange them to form a new equipartition π_α. To be precise, for fixed α, the partition π_α is not a well-defined partition. However, for convenience we assume that π_α is a partition constructed as above which satisfies some explicit properties that we soon mention which a.a.s. holds for random partitions. Denote the new triangle density vector of π_α by $\mathbf{d}^\alpha = (d^\alpha_{ijk})$.

We know that every $(1/r,\cdots,1/r)$-cut $\mathbf{X}=(X_1,\cdots,X_r)$ of the vertex set $[t]$ also gives a $(1/r,\cdots,1/r)$-cut of $V(G)$. By the triangle balanced cut property, for every $\alpha\in[0,1]$,

$$(p^3\pm\delta)\left(\frac{n}{r}\right)^3\cdot\binom{r}{3}=\sum_{(i,j,k)\pitchfork\mathbf{X}}Tr(X_i,X_j,X_k)=\sum_{(i,j,k)\pitchfork\mathbf{X}}d^{\alpha}_{ijk}\left(\frac{n}{t}\right)^3.$$

So $\sum_{(i,j,k)\pitchfork\mathbf{X}}d^{\alpha}_{ijk}=(p^3\pm\delta)\binom{r}{3}\left(\frac{t}{r}\right)^3$. Let M be the $\binom{t}{t/r,\cdots,t/r}\times\binom{t}{3}$ $0-1$ matrix whose rows are indexed by the $(1/r,\cdots,1/r)$-cuts of the vertex set $[t]$ and columns are indexed by the triples $\binom{[t]}{3}$. Where the $(\mathbf{X},(i,j,k))$-entry of M is 1 if and only if $(i,j,k)\pitchfork\mathbf{X}$. The observation above implies $M\mathbf{d}^{\alpha}=(p^3\pm\delta)\binom{r}{3}\left(\frac{t}{r}\right)^3\cdot\mathbf{1}$ where $\mathbf{1}$ is the all-one vector. Thus if we let $\mathbf{d'}=\mathbf{d}^{1/2}-\frac{1}{2}\mathbf{d}^0-\frac{1}{2}\mathbf{d}^1$, then $M\mathbf{d'}=_{\delta t^3}\mathbf{0}$. From this equation we hope to get useful information about the densities x_i and d_{ij}. With the help of the following lemma, we can compute the new densities d^{α}_{ijk}, and thus the modified density vector $\mathbf{d'}$, in terms of the densities x_i and d_{ij}.

Lemma 3.2 *Let ε satisfy $2\varepsilon<d_{ij}$ for every i,j and assume that the graph G is large enough. Then for all $\alpha\in(\varepsilon,1-\varepsilon)$, there exists a choice of sets U_1,U_2 such that the following holds.*

(1)

$$d^{\alpha}_{ijk}=\begin{cases}d_{ijk} & \text{if } \{i,j,k\}\cap\{1,2\}=\emptyset\\ (1-\alpha)d_{1jk}+\alpha d_{2jk}+o(1) & \text{if } i=1 \text{ and } 2\notin\{j,k\}\\ \alpha d_{1jk}+(1-\alpha)d_{2jk}+o(1) & \text{if } i=2 \text{ and } 1\notin\{j,k\}\\ \text{see (2)} & \text{if } i=1 \text{ and } j=2\end{cases}.$$

(2) *If (V_1,V_2,V_k) is a regular triple, then*

$$d^{\alpha}_{12k}=((1-\alpha)^2+\alpha^2)d_{12}d_{1k}d_{2k}+\alpha(1-\alpha)(x_1d^2_{1k}+x_2d^2_{2k})+O(\varepsilon).$$

(3) *Let $d'_{ijk}=d^{\alpha}_{ijk}-(1-\alpha)d^0_{ijk}-\alpha d^1_{ijk}$. Then*

$$d'_{ijk}=\begin{cases}0 & \text{if } \{i,j,k\}\cap\{1,2\}=\emptyset\\ o(1) & \text{if } i=1 \text{ and } 2\notin\{j,k\}\\ o(1) & \text{if } i=2 \text{ and } 1\notin\{j,k\}\end{cases}.$$

Moreover, for the case $i=1$ and $j=2$, if (V_1,V_2,V_k) is a regular triple, then

$$d'_{12k}=\alpha(1-\alpha)(x_1d^2_{1k}+x_2d^2_{2k}-2d_{12}d_{1k}d_{2k})+O(\varepsilon).$$

Proof. (1) The claim clearly holds for the cases $\alpha=0$ and $\alpha=1$.

For $\alpha\in(0,1)$, if $\{i,j,k\}\cap\{1,2\}=\emptyset$, then the density d^{α}_{ijk} is not affected by the swap of vertices in V_1 and V_2 so it remains the same with d_{ijk}. In the case that $\{i,j,k\}\cap\{1,2\}=\{1\}$, without loss of generality we assume $i=1$ and $j,k\neq2$. We also assume that there are S_x triangles with a fixed vertex $x\in V_1\cup V_2$ and two other vertices belonging to V_j and V_k respectively (note that $S_x\leq m^2$). After swapping subset $U_1\subset V_1$ with $U_2\subset V_2$ such that $|U_1|=|U_2|=\alpha|V_i|$, we know that the number of triangles in triple $(((V_1\cup U_2)\backslash U_1),V_j,V_k)$ changes by $\sum_{u\in U_2}S_u-\sum_{u\in U_1}S_u$.

Assume $|V_i|=m$, instead of taking αm vertices uniformly at random, take every vertex in V_1 (or V_2) independently with probability α. This gives random variables X_i for $1\leq i\leq m$ having Bernoulli distribution with parameter α. Let $R=\sum_{i=1}^m X_i$ and $S=\sum_{i=1}^m X_iS_i$. These random variables represent the number of vertices chosen for U_1, and the number of triangles in the triple that contain these chosen vertices, respectively. It is easy to see

$$Pr(R = \alpha m) = \alpha^{\alpha m}(1-\alpha)^{(1-\alpha)m}\binom{m}{\alpha m} \sim \Omega\left(\frac{1}{\sqrt{\alpha(1-\alpha)}}m^{-1/2}\right),$$

and by Hoeffding-Azuma's inequality (Theorem 2.4)

$$Pr(|S - \mathbb{E}S| \geq C) \leq 2\exp\left(-\frac{C^2}{2\sum_{i=1}^m S_i^2}\right) \leq 2\exp\left(-\frac{C^2}{2m^3}\right).$$

Let $C = m^2$, and the second probability decreases much faster than the first probability, thus we know that conditioned on the event $R = \alpha m$, S is also concentrated at its expectation $\mathbb{E}S = \sum_{i=1}^m \alpha S_x = \alpha d_{1jk} m^3$. From here we know the number of triangles changes by

$$\sum_{u\in U_2} S_u - \sum_{u\in U_1} S_u = \alpha d_{2jk} m^3 - \alpha d_{1jk} m^3 + o(m^3)$$

Therefore the new density is

$$d_{1jk}^{\alpha} = d_{1jk} + (\sum_{u\in U_2} S_u - \sum_{u\in U_1} S_u)/m^3 = (1-\alpha)d_{1jk} + \alpha d_{2jk} + o(1)$$

We can use similar method to compute d_{2jk}^{α}.

(2) Let $U_1 \subset V_1$, $U_2 \subset V_2$ be as in (1), and let $V_1' = (V_1 \setminus U_1) \cup U_2$, $V_2' = (V_2 \setminus U_2) \cup U_1$. Then we have the identity

$$Tr(V_1', V_2', V_k) = Tr(U_1, U_2, V_k) + Tr(U_1, V_1 \setminus U_1, V_k) + Tr(V_2 \setminus U_2, U_2, V_k) + Tr(V_2 \setminus U_2, V_1 \setminus U_1, V_k).$$

Since $\alpha \in (\varepsilon, 1-\varepsilon)$, the triples (U_1, U_2, V_k) and $(V_1 \setminus U_1, V_2 \setminus U_2, V_k)$ are regular. Thus by Lemma 2.2,

$$Tr(U_1, U_2, V_k) = (d_{12}d_{1k}d_{2k} + O(\varepsilon))|U_1||U_2||V_k| = (d_{12}d_{1k}d_{2k} + O(\varepsilon))\alpha^2 m^3 \qquad \text{and,}$$
$$Tr(V_1 \setminus U_1, V_2 \setminus U_2, V_k) = (d_{12}d_{1k}d_{2k} + O(\varepsilon))(1-\alpha)^2 m^3.$$

To compute $Tr(U_1, V_1 \setminus U_1, V_k)$, let $E_1^k \subset E(V_1)$ be the collection of edges which have $(d_{1k} \pm \varepsilon)^2 m$ common neighbors in V_k. By the regularity of the pair (V_1, V_k), there are at most $2\varepsilon m$ vertices in V_1 which does not have $(d_{1k} \pm \varepsilon)m$ neighbors in V_k. If v is not such a vertex, then by $\varepsilon < d_{1k} - \varepsilon$ and regularity, there are at most $2\varepsilon m$ other vertices in V_1 which do not have $(d_{1k} \pm \varepsilon)^2 m$ common neighbors with v. Consequently there are at most $4\varepsilon m^2$ edges inside V_1 which does not have $(d_{1k} \pm \varepsilon)^2 m$ common neighbors inside V_k, thus $|E_1^k| \geq x_1\binom{m}{2} - 4\varepsilon m^2$. By Lemma 2.5 and the calculation from part (1) there exists a choice of U_1 of size αm such that,

$$\begin{aligned}|E_1^k(U_1, V_1 \setminus U_1)| &= |E_1^k| - |E_1^k(U_1)| - |E_1^k(V_1 \setminus U_1)| \\ &= (1 - \alpha^2 - (1-\alpha)^2 + o(1))|E_1^k| = \alpha(1-\alpha)x_1 m^2 + O(\varepsilon)m^2.\end{aligned}$$

for all k. Note that the number of triangles $Tr(U_1, V_1 \setminus U_1, V_k)$ can be computed by adding the number of triangles containing the edges E_1^k and then the number of triangles containing at most $O(\varepsilon)m^2$ of the "exceptional" edges. The latter can be crudely bounded by $O(\varepsilon)m^2 \cdot m \leq O(\varepsilon)m^3$. Since each edge in E_1^k is contained in $(d_{1k} \pm \varepsilon)^2 m$ triangles (within the triple (V_1, V_2, V_k)),

$$Tr(U_1, V_1 \setminus U_1, V_k) = |E_1^k(U_1, V_1 \setminus U_1)| \cdot (d_{1k}^2 + O(\varepsilon))m + O(\varepsilon)m^3 = \alpha(1-\alpha)x_1 d_{1k}^2 m^3 + O(\varepsilon)m^3.$$

Similarly we can show that there exists a choice of U_2 of size αm that gives

$$Tr(U_2, V_2 \setminus U_2, V_k) = \alpha(1-\alpha)x_2 d_{2k}^2 m^3 + O(\varepsilon)m^3.$$

for all k. Combining all the results together, we can conclude the existence of sets U_1, U_2 such that

$$d_{12k}^{\alpha} = ((1-\alpha)^2 + \alpha^2)d_{12}d_{1k}d_{2k} + \alpha(1-\alpha)(x_1 d_{1k}^2 + x_2 d_{2k}^2) + O(\varepsilon).$$

Part (3) is just a straightforward computation from the definition of d'_{ijk} and (2). $\square$

Lemma 3.3 *Let ε satisfy $2\varepsilon < d_{ij}$ for every i, j and assume that the graph G is large enough. If (V_i, V_j, V_k) is a regular triple, then $x_i d_{ik}^2 + x_j d_{jk}^2 - 2d_{ij}d_{ik}d_{jk} =_{\delta t^3 + \varepsilon} 0$.*

Proof. As mentioned before Lemma 3.2, the vector $\mathbf{d}' = \mathbf{d}^{1/2} - \frac{1}{2}\mathbf{d}^0 - \frac{1}{2}\mathbf{d}^1$ satisfies $M\mathbf{d}' =_{\delta t^3} \mathbf{0}$. For an index $k \neq 1, 2$, consider a balanced partition $\mathbf{X}$ of the vertex set $[t]$ such that 1 and 2 lies in different parts, and let Y be the union of the parts which contains neither 1 nor 2. Then by Lemma 3.2 (3),

$$0 =_{\delta t^3} \sum_{(i,j,k) \pitchfork \mathbf{X}} d'_{ijk} = t^3 \cdot o(1) + \sum_{k \in Y} d'_{12k},$$

where $o(1)$ goes to 0 as the number of vertices in the graph G grows. Since Y can be an arbitrary set of size $(r-2)t/r$ not containing 1 and 2, this immediately implies that $d'_{12k} =_{\delta t^3} 0$ for all k. Thus if (V_1, V_2, V_k) is a regular triple, then by Lemma 3.2 (3), $x_1 d_{1k}^2 + x_2 d_{2k}^2 - 2d_{12}d_{1k}d_{2k} =_{\delta t^3 + \varepsilon} 0$. By symmetry, we can replace 1 and 2 by arbitrary indices i, j. $\square$

Using Theorem 2.6 which characterizes the non quasi-random hypergraphs satisfying the balanced cut property, we can prove the following lemma which allows us to bound the densities from below.

Lemma 3.4 *There exists t_0 such that for fixed $p \in (0,1)$ and every $t \geq t_0$ which is divisible by $2r$, there exist $c = c(p)$ and $\delta_0 = \delta_0(t,p) > 0$ so that the following holds for every $\delta \leq \delta_0$. If G is a graph with density p which is δ-close to satisfying the triangle balanced cut property, then for any partition π of $V(G)$ into t equal parts, the density vector $\mathbf{d} = (d_{ij})_{i,j}$ satisfies $d_{ij} \geq c$ for all distinct $i, j \in [t]$.*

Proof. Let $t_0 = t_{2.6}$, $\eta = p^3/6$, and for a given $t \geq t_0$ divisibly by $2r$, let $\delta_0 = \min\{\delta_{2.6}(t,\eta), p^3/10\}$. Let $V = V(G)$, and let G' be the hypergraph over the vertex set V such that $\{i,j,k\} \in E(G')$ if and only if i, j, k forms a triangle in the graph G. Let π be an arbitrary partition of V into t equal parts $V_1, \ldots, V_t$, and let $(d_{ij})_{i,j}$ be the density vector of the graph G, and $(d_{ijk})_{i,j,k}$ be the density vector of the hypergraph G' with respect to π. It suffices to show the bound $d_{ij} \geq p^3/10$ for every distinct $i, j \in [t]$. For simplicity we will only verify it for d_{12}. Note that number of triangles which cross V_1, V_2, V_k is at most $e(V_1, V_2) \cdot |V_k| = (|V_1||V_2|d_{12}) \cdot |V_k|$, and thus $d_{12k} \leq d_{12}$ for all $k \geq 3$. Consequently, we can obtain the following inequality which would be crucial in our argument:

$$\sum_{k=3}^{t} d_{12k} \leq (t-2) \cdot d_{12}. \tag{1}$$

Since G is δ-close to satisfying the triangle balanced cut property, we know that the density q of triangles is at least $q \geq p^3 - \delta$. By Theorem 2.6, $(d_{ijk})_{i,j,k}$ is η-equal to some vector in $W_{t,q}$. Recall that the vectors in $W_{t,q}$ can be expressed as an affine combination of the vectors $\mathbf{u}_{t,q,I} = (u_{ijk}^I)_{i,j,k}$

for sets $I \subset [t]$ of size $|I| = t/2$, and note that the following is true no matter how we choose the set I:

$$\sum_{k=3}^{t} u^I_{12k} \geq \sum_{k\in I} \frac{2q}{3} \geq \left(\frac{t}{2} - 2\right) \frac{2q}{3}.$$

Since $(d_{ijk})_{i,j,k}$ is ε-equal to an affine combination of these vectors, for large enough t we have

$$\sum_{k=3}^{t} d_{12k} \geq \left(\frac{t}{2} - 2\right) \frac{2q}{3} - t\varepsilon \geq \frac{tq}{3} - \frac{4q}{3} - \frac{tq}{6} \geq \frac{tq}{8}. \tag{2}$$

By combining (1) and (2), we obtain $d_{12} \geq q/8 \geq (p^3 - \delta)/8 \geq p^3/10$. Similarly we can deduce $d_{ij} \geq p^3/10$ for all distinct $i, j \in [t]$. $\square$

Since Lemma 3.4 asserts that all the pairwise densities d_{ij} are bounded from below by some constant, we are allowed to divide each side of an ε-equality by d_{ij}. This turns out to be a crucial ingredient in solving the equations given by Lemma 3.3.

Lemma 3.5 *Given a positive real c and an integer $n \geq 4$, if for every distinct $i, j \in [n]$, $d_{ij} \geq c$ and for every distinct $i, j, k \in [n]$, $x_i d^2_{ik} + x_j d^2_{jk} - 2d_{ij}d_{ik}d_{jk} =_\varepsilon 0$, then there exists $s \in [n]$, $x, y > 0$ such that for any distinct $i, j \neq s$, we have $d_{ij} =_\varepsilon \sqrt{x}$. For any $i \neq s$, $d_{is} =_\varepsilon \sqrt{y}$. Moreover, $x_i =_\varepsilon \sqrt{x}$ if $i \neq s$ and $x_s =_\varepsilon \frac{\sqrt{x}}{y}(2y - x)$ (see, figure 3.1).*

Proof. Throughout the proof, we heavily rely on properties of ε-equality given in Section 2.

First consider the case $n = 4$. By taking $(i, j, k) = (1, 2, 3), (2, 3, 1), (3, 1, 2)$ respectively, we get the following system of equations:

$$\begin{cases} d^2_{13}x_1 + d^2_{23}x_2 =_\varepsilon 2d_{12}d_{13}d_{23} \\ d^2_{12}x_1 + d^2_{23}x_3 =_\varepsilon 2d_{12}d_{13}d_{23} \\ d^2_{12}x_2 + d^2_{13}x_3 =_\varepsilon 2d_{12}d_{13}d_{23} \end{cases} \tag{3}$$

Considering this as a system of linear equations with unknowns x_1, x_2, x_3, the determinant of the coefficient matrix becomes $2d^2_{12}d^2_{13}d^2_{23} \geq 2c^6$. Moreover, the coefficients in the matrix are bounded from above by 1. Therefore we can solve the linear system by appealing to property (6) of ε-equality and get

$$\begin{cases} x_1 =_\varepsilon \dfrac{d_{23}}{d_{12}d_{13}}(d^2_{12} + d^2_{13} - d^2_{23}) \\ x_2 =_\varepsilon \dfrac{d_{13}}{d_{12}d_{23}}(d^2_{12} + d^2_{23} - d^2_{13}) \\ x_3 =_\varepsilon \dfrac{d_{12}}{d_{13}d_{23}}(d^2_{13} + d^2_{23} - d^2_{12}) \end{cases}. \tag{4}$$

Then

$$\begin{aligned} x_1x_2 &=_\varepsilon \frac{d_{23}}{d_{12}d_{13}}(d^2_{12} + d^2_{13} - d^2_{23}) \cdot \frac{d_{13}}{d_{12}d_{23}}(d^2_{12} + d^2_{23} - d^2_{13}) \\ &=_\varepsilon \frac{1}{d^2_{12}}[d^4_{12} - (d^2_{13} - d^2_{23})^2] \\ &=_\varepsilon \frac{1}{d^2_{12}}[d^4_{12} - (d^2_{14} - d^2_{24})^2]. \end{aligned} \tag{5}$$

The last equation comes from repeating the same step for the system of equations for indices 1, 2, and 4. Equation (5) implies $d^2_{13} - d^2_{23} =_\varepsilon \pm(d^2_{14} - d^2_{24})$, and $d^2_{ik} - d^2_{jk} =_\varepsilon \pm(d^2_{il} - d^2_{jl})$ for all distinct i, j, k, l in general. Assume that there exists an assignment $\{i, j, k, l\} = \{1, 2, 3, 4\}$ such that $d_{ik}^2 - d_{jk}^2 =_\varepsilon -(d_{il}^2 - d_{jl}^2) \neq_\varepsilon 0$, (we call such case as a "flip"). Without loss of generality let $d_{13}^2 - d_{23}^2 =_\varepsilon -(d_{14}^2 - d_{24}^2)$. By equation (3), we know that $x_1 d_{12}^2 + x_3 d_{23}^2 =_\varepsilon 2d_{12}d_{13}d_{23} =_\varepsilon x_2 d_{12}^2 + x_3 d_{13}^2$, from which we get

$$d_{12}^2(x_1 - x_2) =_\varepsilon (d_{13}^2 - d_{23}^2)x_3.$$

Replace the index 3 by 4 and we get

$$d_{12}^2(x_1 - x_2) =_\varepsilon (d_{14}^2 - d_{24}^2)x_4.$$

By the assumption on a "flip", by subtracting the two equalities we get $x_3 + x_4 =_\varepsilon 0$, thus $x_3 =_\varepsilon 0$ and $x_4 =_\varepsilon 0$. This is impossible from the equation $x_3 d_{13}^2 + x_4 d_{14}^2 =_\varepsilon 2d_{13}d_{14}d_{34}$ and the fact $d_{ij} \geq c$. Therefore no flip exists and we have

$$d_{ik}^2 - d_{il}^2 =_\varepsilon d_{jk}^2 - d_{jl}^2 \qquad \forall \{i, j, k, l\} = \{1, 2, 3, 4\}. \tag{6}$$

Since $d_{ij} \geq c$, the sum of $d_{12}^2 + d_{13}^2 - d_{23}^2$, $d_{12}^2 + d_{23}^2 - d_{13}^2$ and $d_{13}^2 + d_{23}^2 - d_{12}^2$ is equal to $d_{12}^2 + d_{13}^2 + d_{23}^2 \geq 3c^2$. So at least one of the terms is greater than c^2, without loss of generality we can assume $d_{12}^2 + d_{13}^2 - d_{23}^2 \geq c^2$. Recall that $x_1 =_\varepsilon \dfrac{d_{23}}{d_{13}d_{12}}(d_{12}^2 + d_{13}^2 - d_{23}^2)$. By equation (6), we also have $x_1 =_\varepsilon \dfrac{d_{24}}{d_{14}d_{12}}(d_{12}^2 + d_{14}^2 - d_{24}^2) =_\varepsilon \dfrac{d_{24}}{d_{14}d_{12}}(d_{12}^2 + d_{13}^2 - d_{23}^2)$. Therefore

$$\frac{d_{23}}{d_{13}} =_\varepsilon \frac{d_{24}}{d_{14}}.$$

By appealing to the bound $d_{ij} \geq c$, we get $d_{23}d_{14} =_\varepsilon d_{24}d_{13}$ and $d_{23}^2 d_{14}^2 =_\varepsilon d_{24}^2 d_{13}^2$, which implies

$$\begin{aligned}
&(d_{13}^2 - d_{23}^2)(d_{13}^2 - d_{14}^2) \\
&= d_{13}^2(d_{13}^2 - d_{23}^2) - d_{13}^2 d_{14}^2 + d_{23}^2 d_{14}^2 \\
&=_\varepsilon d_{13}^2(d_{14}^2 - d_{24}^2) - d_{13}^2 d_{14}^2 + d_{13}^2 d_{24}^2 = 0
\end{aligned}$$

So either $d_{13}^2 =_{\sqrt{\varepsilon}} d_{14}^2$ or $d_{13}^2 =_{\sqrt{\varepsilon}} d_{23}^2$. Thus at this point we may assume the existence of indices i, j, k satisfying $d_{ik}^2 =_{\sqrt{\varepsilon}} d_{jk}^2$. Assume that $d_{13}^2 =_{\sqrt{\varepsilon}} d_{14}^2$ as the other case can be handled identically.

So $d_{13}^2 =_{\sqrt{\varepsilon}} x$, $d_{14}^2 =_{\sqrt{\varepsilon}} x$ for some x and by equation (6) we have $d_{23}^2 =_{\sqrt{\varepsilon}} y$, $d_{24}^2 =_{\sqrt{\varepsilon}} y$ for some y. We let $d_{34}^2 = z$, and the equation $d_{14}^2 - d_{34}^2 = d_{12}^2 - d_{32}^2$ given by (6) translates to $d_{12}^2 =_\varepsilon x + y - z$. Moreover, from equation (4) for indices $\{1, 3, 4\}$ and $\{1, 2, 4\}$ we know that

$$x_1 =_\varepsilon \frac{d_{34}}{d_{14}d_{13}}(d_{14}^2 + d_{13}^2 - d_{34}^2) =_\varepsilon \frac{d_{24}}{d_{14}d_{12}}(d_{14}^2 + d_{12}^2 - d_{24}^2)$$

If we plug all equalities for d_{ij} into this identity, we get

$$(2x - z)\frac{\sqrt{z}}{x} =_{\sqrt{\varepsilon}} (2x - z)\frac{\sqrt{y}}{\sqrt{x(x + y - z)}}$$

So either $\dfrac{\sqrt{z}}{x} =_{\varepsilon^{1/4}} \dfrac{\sqrt{y}}{\sqrt{x(x + y - z)}}$ or $z =_{\varepsilon^{1/4}} 2x$. In the first case, by solving this equation we get either $z =_{\varepsilon^{1/8}} x$ or $z =_{\varepsilon^{1/8}} y$ (before multiplying each side of the equation by its denominators, one must establish the fact that $x + y - z$ is bounded away from 0. This can be done by first figuring out the equation $\sqrt{\frac{z}{xy}} =_{\varepsilon^{1/4}} \sqrt{\frac{1}{x+y-z}}$, and then realizing that the left hand side is bounded from above). Both of the above solutions gives us a graph as claimed (see figure 3.1, for the case $z =_{\varepsilon^{1/8}} x$).

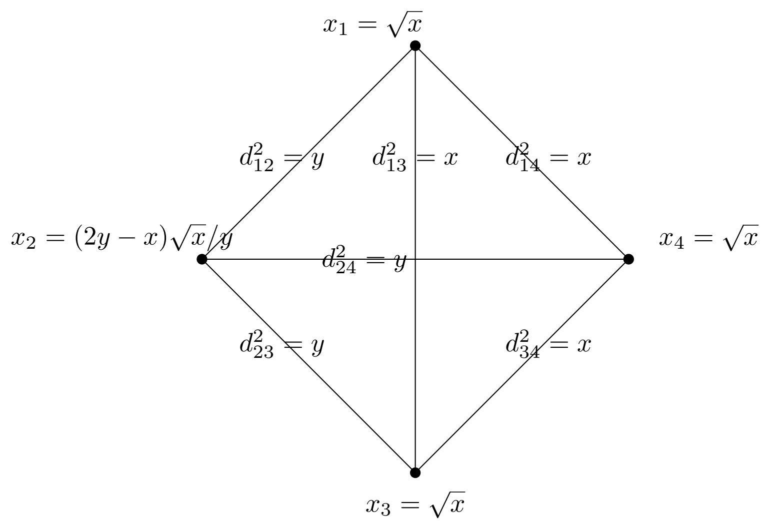

Figure 3.1: The structure of solution for $n = 4$.

In the second case $z =_{\varepsilon^{1/4}} 2x$, we consider the equation

$$x_2 =_\varepsilon \frac{d_{34}}{d_{24}d_{23}}(d_{24}^2 + d_{23}^2 - d_{34}^2) =_\varepsilon \frac{d_{14}}{d_{24}d_{12}}(d_{24}^2 + d_{12}^2 - d_{14}^2)$$

to get

$$(2y - z)\frac{\sqrt{z}}{x} =_{\sqrt{\varepsilon}} (2y - z)\frac{\sqrt{y}}{\sqrt{x(x + y - z)}}.$$

By the previous analysis we may assume $z =_{\sqrt{\varepsilon}} 2y$, which implies $x =_{\varepsilon^{1/4}} y$ and $d_{12}^2 =_{\varepsilon^{1/4}} 0$. This is impossible by the fact $d_{12} \geq c$.

For $n = 5$, suppose not all the edge densities are ε-equal to the same value. In this case, there must be four vertices such that not all the densities between them are equal. Without loss of generality, we can assume $d_{12} =_\varepsilon d_{13} =_\varepsilon d_{14} =_\varepsilon x$, $d_{23} =_\varepsilon d_{24} =_\varepsilon d_{34} =_\varepsilon y$ and $x \neq y$. Now let us consider the collections of vertices $\{v_1, v_2, v_3, v_5\}$, $\{v_1, v_2, v_4, v_5\}$, $\{v_1, v_3, v_4, v_5\}$. From the case $n = 4$, we know that $d_{15} =_\varepsilon x$, $d_{25} =_\varepsilon d_{35} =_\varepsilon d_{45} =_\varepsilon y$. We can generalize it to arbitrary $n \geq 5$. $\square$

Note that if in the regular partition, every pair of sets were regular, then Lemma 3.5 itself forces the graph to be quasi-random, as apart from one part (which is negligible), all the densities are equal. However, the regularity lemma inevitably produces a partition which contains some irregular pairs, and in the remaining of the proof of Theorem 3.1 we will show how to handle this subtlety. The main idea is that since there are only a small number of irregular pairs, the reduced graph will contain many cliques, and thus that we can use Lemma 3.5 to study its structure.

From now on in the reduced graph, when a clique of size at least 4 is given, we will call the exceptional vertex s to be "bad" and all others to be "good" vertices. We also call the densities d_s and d_{is} for any $i \neq s$ "bad" and d_{ij} "good" for $i, j \neq s$. However as it will later turn out, most cliques of size 4 has $x = y$, and in this case we call every vertex and edge to be "good".

Now we can combine Lemmas 3.3, 3.4, and 3.5 above to prove the main theorem which says that triangle balanced cut property is quasi-random.

Proof of Theorem 3.1 (triangle case). Let $c = c_{3.4}(p)$. We may assume that $\varepsilon < \min\{c/2, 1/4\}$. Let $t_0 = t_{3.4}$ and $T = T_{2.1}(\varepsilon, t_0)$. Let $\delta = \min_{t_0 \leq t \leq T}\{\varepsilon/t^3, \delta_{3.4}(t, p)\}$.

Let G be a graph which is δ-close to satisfying $\mathcal{C}_3(1/r, \cdots, 1/r)$. Consider the ε-regular equipartition π of G: $V(G) = V_1 \cup \cdots \cup V_t$ we mentioned before. This gives a reduced graph H on t vertices of minimum degree at least $(1-\varepsilon)t$ (we may assume that t is divisible by $4r$). Every edge ij corresponds to an ε-regular pair (V_i, V_j). We mark on each edge of H a weight d_{ij} which is the density of edges in (V_i, V_j), and also the density x_i inside V_i on the vertices. Parameters are chosen so that Lemma 3.3 and Lemma 3.4 holds. Moreover, by the fact $\delta t^3 \leq \varepsilon$, we have $x_i d_{ik}^2 + x_j d_{jk}^2 - 2d_{ij}d_{ik}d_{jk} =_\varepsilon 0$ for every regular triple (V_i, V_j, V_k). Thus whenever there is a clique of size at least 4 in H, by Lemma 3.5 we know that all the densities are ε-equal to each other, except for at most one "bad" vertex. Since $\varepsilon < 1/4$ and $4|t$, we can apply Hajnal-Szemerédi theorem (Theorem 2.3) to the reduced graph H and get an equitable partition of the vertices of H into vertex disjoint 4-cliques $C_1, \cdots, C_{t/4}$.

For every 4-clique C_i, from Lemma 3.5 we know that there is at most one "bad" vertex. For two 4-cliques C_i and C_j, we can consider the bipartite graph $\mathcal{B}(C_i, C_j)$ between them which is induced from H. If $\mathcal{B}(C_i, C_j) = K_{4,4}$, then it contains a subgraph isomorphic to $K_{2,2}$ where all the vertices are "good" (two vertices are good in C_i and other two in C_j). If we apply the structural lemma, Lemma 3.5, to this new 4-clique (together with two edges coming from the two known cliques), we get that the "good" densities of C_i and C_j are ε-equal to each other.

Now consider the reduced graph H' whose vertices correspond to the 4-cliques C_i, and C_i and C_j are adjacent in H' if and only if there is a complete bipartite graph between them. It is easy to see that $\delta(H') \geq (1-4\varepsilon)|H'|$. Take any two vertex $u', v' \in V(H')$, since $d(u') + d(v') > |H'|$ they have a common neighbor w', and thus by the discussion above, the "good" density in $C_{u'}$ or $C_{v'}$ are ε-equal to the "good" density in $C_{w'}$. So all the "good" densities are ε-equal to each other. Thus by the total transitivity of ε-equality (see, Section 2), all the "good" densities are ε-equal to p' for some p'.

We would like to show that $d_{ij} =_\varepsilon p'$ for all but at most $O(\varepsilon)t^2$ edges of the reduced graph H. We already verified this for "good" edges $\{i, j\}$ belonging to the cliques $C_1, \ldots, C_{t/4}$. If C_i is adjacent to C_j in H' then they actually form a clique of size 8 in H, and by Lemma 3.5 there is at most one "bad" vertex there. Consequently the total number of cliques that contain at least one "bad" vertex cannot exceed the independence number of H', which is at most $|H'| - \delta(H') \leq \varepsilon t$. Thus among the cliques $C_1, \ldots, C_{t/4}$ there are at most εt cliques which contain at least one "bad" vertex. Moreover, the density of an edge in H which is part of a $K_{4,4}$ connecting two "good" cliques C_i and C_j are ε-equal to p' again by Lemma 3.5. Among the remaining edges, all but εt^2 are such edges as otherwise $e(H) < \binom{t}{2} - \varepsilon t^2$ which is a contradiction. Therefore all but at most $O(\varepsilon)t^2$ edges of H have density ε-equal to p'. This in turn implies that the density of G is equal to $p' + O(\varepsilon)$. On the other hand we know that the density is p, thus $p' =_\varepsilon p$.

Now by verifying that G satisfies $\mathcal{P}_2(1/2)$ (see Theorem 1.1), we will show that G is quasi-random. For an arbitrary subset $U \subset V(G)$ of size $n/2$, let us compute the number of edges in $e(U)$ and estimate its difference with the number of edges of a subset of size $n/2$ in $G(n, p)$.

$$\begin{aligned}
&\left| e(U) - \binom{n/2}{2} p \right| \\
\leq &\sum_{i=1}^{t} \left| e(U \cap V_i) - \binom{|U \cap V_i|}{2} p \right| + \sum_{i,j} |\, e(U \cap V_i, U \cap V_j) - |U \cap V_i||U \cap V_j| p \,| \\
\leq &\sum_{i=1}^{t} |V_i|^2 + \left(\sum_{i,j} |U \cap V_i||U \cap V_j| \right) |p' - p + O(\varepsilon)| \\
\leq &n^2/t + O(\varepsilon |U|^2/2) = O(\varepsilon n^2).
\end{aligned} \tag{7}$$

In the last equation, we took sufficiently large t depending on ε and p. Therefore by the quasi-randomness of $\mathcal{P}_2(1/2)$, we can conclude that G is a quasi-random graph. $\square$

4 General Cliques

Throughout the section, k and r are fixed integers satisfying $r \geq k \geq 4$. Let a r-balanced cut be a $(1/r, \cdots, 1/r)$-cut. In this section, we will prove the remaining cases of the main theorem, quasi-randomness of general k-clique r-balanced cut properties.

Theorem 4.1 *Fix a real $p \in (0,1)$ and positive integers r, k such that $r \geq k \geq 4$. For every $\varepsilon > 0$, there exists a positive real δ such that the following is true. If G is a graph which has density p and is δ-close to satisfying the K_k balanced cut property $\mathcal{C}_k(1/r, \cdots, 1/r)$, then G is ε-close to being p-quasi-random.*

Let G be a graph which is δ-close to satisfying the k-clique r-balanced cut property. Apply the regularity lemma (Theorem 2.1) to this graph to obtain an ε-regular partition $(V_i)_{i=1}^t$ of the vertex set. For $i \in [t]$, let x_i be the density of the edges within V_i, and for distinct $i, j \in [t]$, let d_{ij} be the density of the pair (V_i, V_j). For $k \geq 2$, a k-tuple $J = (i_1, \ldots, i_k)$ is a multiset of k-indices (not necessarily distinct). Let d_J be the density of k-cliques which have exactly one vertices in each of the V_{i_a} for $a = 1, \ldots, k$. A k-tuple J is called *regular* if (V_{i_a}, V_{i_b}) forms an ε-regular pair for all $a, b \in [k]$. For a k-tuple J and a cut $\mathbf{X} = (X_1, \ldots, X_r)$, we say that J crosses the cut $\mathbf{X}$ if $|J \cap X_i| \leq 1$ for all i, and denote it by $J \pitchfork \mathbf{X}$.

The proof of the k-clique r-balanced cut case follows the same line of the proof of the triangle case. First we develop two lemmas which correspond to Lemma 3.3 and Lemma 3.4.

Lemma 4.2 *Let ε be small enough depending on the densities d_{ij} for all $i, j \in [t]$. There exists a function $f : \mathbb{R} \to \mathbb{R}$ such that the following holds. Let J be a regular k-tuple, $J' \subset J$ be such that $|J'| = k - 2$, and $\{j_1, j_2\} = J \setminus J'$. Then,*

$$x_{j_1}\left(\prod_{a\in J'} d_{aj_1}\right)^2\left(\prod_{a,b\in J',a<b} d_{ab}\right)+x_{j_2}\left(\prod_{a\in J'} d_{aj_2}\right)^2\left(\prod_{a,b\in J',a<b} d_{ab}\right)-2\left(\prod_{a,b\in J,a<b} d_{ab}\right) =_{\varepsilon+\delta\cdot f(t)} 0.$$

Proof. For the sake of clarity, without loss of generality we consider the case $\{1, 2\} \subset J$ and $j_1 = 1, j_2 = 2$. As in the triangle case, by considering the family of t-partitions

$$\pi_\alpha = ((1-\alpha)V_1 + \alpha V_2, \alpha V_1 + (1-\alpha)V_2, V_3, \ldots, V_t),$$

and the density vector $(d_J^\alpha)_{J\in\binom{[t]}{r}}$ which arise from these partitions, we can define $\mathbf{d}' = (d'_J)_{J\in\binom{[t]}{r}}$ as $d'_J = d_J^{1/2} - \frac{1}{2}d_J^0 - \frac{1}{2}d_J^1$. The same proof as in Lemma 3.2 gives us,

$$d'_J = \begin{cases} 0 & \text{if } J \cap \{1,2\} = \emptyset \\ o(1) & \text{if } J \cap \{1,2\} = \{1\} \\ o(1) & \text{if } J \cap \{1,2\} = \{2\} \end{cases}, \tag{8}$$

and if $\{1, 2\} \subset J$ (let $J' = J \setminus \{1, 2\}$)) and J is a regular r-tuple, we have

$$d'_J = \alpha(1-\alpha)\left(\sum_{i=1}^{2} x_i\left(\prod_{a\in J'} d_{ai}\right)^2\left(\prod_{a,b\in J',a<b} d_{ab}\right) - 2\left(\prod_{a,b\in J,a<b} d_{ab}\right)\right) + O(\varepsilon). \tag{9}$$

If $\{1,2\} \subset J$ and J is not a regular k-tuple, then we do not have any control on d'_J.

Let M be the $\binom{t}{t/r,\ldots,t/r} \times \binom{t}{k}$ 0-1 matrix whose rows are indexed by r-balanced cuts of the vertex set $[t]$ and columns are indexed by the k-tuples $\binom{[t]}{k}$. Where the $((X_1, X_2, \ldots, X_r), J)$-entry of M is 1 if and only if $J \pitchfork (X_1, X_2, \ldots, X_r)$. We know that $M\mathbf{d}' =_{\delta t^k} \mathbf{0}$.

Consider the submatrix N of M formed by the rows of partitions which have 1 and 2 in different parts, and columns of r-tuples which include both 1 and 2. Let $\mathbf{d}''$ be the projection of $\mathbf{d}'$ onto the coordinates corresponding to the r-tuples which contain both 1 and 2. By equation (8) and $M\mathbf{d}' =_{\delta t^k} \mathbf{0}$, we can conclude that $N\mathbf{d}'' =_{\delta t^k} \mathbf{0}$ given that the graph is large enough. Thus if we can show that N has full rank, then this implies that $d'_J =_{\delta \cdot f(t)} 0$ for all $\{1,2\} \subset J$ (appeal to property (6) of ε-equality - the entries of N are bounded and the size of N depends on t).

Observe that the matrix N can be considered as a 0-1 matrix whose rows are indexed by the collection of subsets $\mathbf{Y} = (Y_1, \ldots, Y_{r-2})$ of the set $T = \{3, \ldots, t\}$ where each part has size t/r, and columns are indexed by the $(k-2)$-tuples $J' \in \binom{T}{k-2}$, where the entry $(\mathbf{Y}, J')$ is 1 if and only if $J' \pitchfork \mathbf{Y}$. By fixing a subset of T of size $(r-2)t/r$ and considering all possible $\mathbf{Y}$ arising within this set, one can see that the row-space of N generates the row-space of the $(r-2)t/r$ verses $k-2$ inclusion matrix of T, which we know by Gottlieb's theorem, Theorem 2.7, has full rank. This implies that N has full rank as well. $\square$

Even though the equation which we obtained in Lemma 4.2 looks a lot more complicated than the triangle case, as it turns out, it is possible to make a substitution of variables so that the equations above become exactly the same as the equations in the triangle case. For a regular $(k-3)$-tuple I, an index $j \notin I$, and distinct $j_1, j_2 \notin I$, define

$$d^I_{j_1 j_2} := d_{j_1 j_2} \left(\prod_{a \in I} d_{a j_1} \right)^{1/2} \left(\prod_{a \in I} d_{a j_2} \right)^{1/2} \left(\prod_{a,b \in I, a<b} d_{ab} \right)^{1/3} \quad \text{and}$$

$$x^I_j := x_j \left(\prod_{a \in I} d_{aj} \right) \left(\prod_{a,b \in I, a<b} d_{ab} \right)^{1/3}.$$

Claim 4.3 *Let J be a regular k-tuple, $I \subset J$ be of size $|I| = k-3$, $\{j_1, j_2, j_3\} = J \setminus I$, and $J' = I \cup \{j_3\}$. Then*

$$d^I_{j_1 j_2} d^I_{j_2 j_3} d^I_{j_3 j_1} = \prod_{a,b \in J, a<b} d_{ab}, \quad \textit{and} \quad x^I_{j_1} (d^I_{j_1 j_3})^2 = x_{j_1} \left(\prod_{a \in J'} d_{a j_1} \right)^2 \left(\prod_{a,b \in J', a<b} d_{ab} \right).$$

Proof. The claim follows from a direct calculation. $\square$

In other words, Claim 4.3 transforms the computation of the density of K_r in the graph into the computation of the density of triangles in another graph. This observation will greatly simplify the equations obatained from Lemma 4.2.

Lemma 4.4 *Let ε be small enough depending on the densities d_{ij} for all $i, j \in [t]$. There exists a function $f : \mathbb{R} \to \mathbb{R}$ such that the following holds. Let J be a regular k-tuple and $I \subset J$ be of size $|I| = k-3$. For $\{j_1, j_2, j_3\} = J \setminus I$, we have $x^I_{j_1} (d^I_{j_1 j_3})^2 + x^I_{j_2} (d^I_{j_2 j_3})^2 - 2 d^I_{j_1 j_2} d^I_{j_2 j_3} d^I_{j_3 j_1} =_{\varepsilon + \delta \cdot f(t)} 0$.*

Proof. This is an immediate corollary of Lemma 4.2 and Claim 4.3. $\square$

Next lemma corresponds to Lemma 3.4 and establishes a lower bound on the densities. We omit the proof which is a straightforward generalization of the proof of Lemma 3.4.

Lemma 4.5 *There exists t_0 such that for fixed $p \in (0,1)$ and every $t \geq t_0$ divisible by $2r$, there exist $c = c(k,p)$ and $\delta_0 = \delta_0(t,p) > 0$ so that the following holds for every $\delta \leq \delta_0$. If G is a graph with density p which is δ-close to satisfying the k-clique balanced cut property, then for any partition π of $V(G)$ into t equal parts, the density vector $\mathbf{d} = (d_{ij})_{i,j}$ satisfies $d_{ij} \geq c$ for all distinct $i, j \in [t]$.*

For every fixed regular $(k-3)$-tuple I, the set of equations that Lemma 4.4 gives is exactly the same as the set of equations obtained from Lemma 3.3. Consequently, by using Lemma 4.5, we can solve these equations for every fixed I just as in the triangle case.

Thus as promised, we can reduce the case of general cliques to the case of triangles. However, this observation does not immediately imply that $d_{j_1 j_2} =_\varepsilon p$ for most of the pairs j_1, j_2, since the only straightforward conclusion that we can draw is that for every regular $(k-3)$-tuple I, there exists a constant p_I such that $d^I_{j_1 j_2} =_\varepsilon p_I$ for most of the pairs j_1, j_2. In order to prove the quasi-randomness of balanced cut properties, we will need some control on the relation between different p_I. Call a k-tuple J *excellent* if it is regular, and for every $(k-3)$-tuple $I \subset J$, we have $d^I_{j_1 j_2} =_\varepsilon p_I$ for all distinct $j_1, j_2 \in J \setminus I$.

Lemma 4.6 *Let J be an excellent k-tuple. Then the density of every pairs in J are ε-equal to each other.*

Proof. For the sake of clarity, assume that $J = (1, 2, \ldots, k)$. First, consider $I = (4, \ldots, k)$. Then by the assumption, we have $d^I_{13} =_\varepsilon d^I_{23}$, which by definition gives,

$$d_{13} \left(\prod_{a \in I} d_{a1} \right)^{1/2} \left(\prod_{a \in I} d_{a3} \right)^{1/2} \left(\prod_{a,b \in I, a<b} d_{ab} \right)^{1/3} =_\varepsilon d_{23} \left(\prod_{a \in I} d_{a2} \right)^{1/2} \left(\prod_{a \in I} d_{a3} \right)^{1/2} \left(\prod_{a,b \in I, a<b} d_{ab} \right)^{1/3}.$$

After cancelation of the same terms, we can rewrite this as,

$$d_{13} \left(\prod_{a=4}^{k} d_{a1} \right)^{1/2} =_\varepsilon d_{23} \left(\prod_{a=4}^{k} d_{a2} \right)^{1/2} \quad \Leftrightarrow \quad d_{13}^{1/2} \left(\prod_{a=3}^{k} d_{a1} \right)^{1/2} =_\varepsilon d_{23}^{1/2} \left(\prod_{a=3}^{k} d_{a2} \right)^{1/2}. \tag{10}$$

We can replace 3 by 4 up to k and multiply each side of all these equations to obtain,

$$\prod_{i=3}^{k} \left(d_{1i}^{1/2} \left(\prod_{a=3}^{k} d_{a1} \right)^{1/2} \right) =_\varepsilon \prod_{i=3}^{k} \left(d_{2i}^{1/2} \left(\prod_{a=3}^{k} d_{a2} \right)^{1/2} \right),$$

which is equivalent to

$$\left(\prod_{i=3}^{k} d_{1i} \right)^{(k-1)/2} =_\varepsilon \left(\prod_{i=3}^{k} d_{2i} \right)^{(k-1)/2}.$$

If we plug this back into equation (10), we get $d_{13} =_\varepsilon d_{23}$. By repeating this process for other choice of indices, we can conclude that the density of every pairs are ε-equal to each other. $\square$

We now combine all these observations to show that $d_e =_\varepsilon p$ for most of the edges e of the reduced graph, which will in turn imply the quasi-randomness.

Proof of Theorem 4.1. Choose ε_0 small enough depending on the constant $c = c_{4.5}(p)$ so that the condition of Lemma 4.4 holds, and let f be the function from Lemma 4.4. Let $\varepsilon \leq \min\{\varepsilon_0, 1/4\}$, $t_0 = t_{4.5}$, and let $T = T_{2.1}(\varepsilon, t_0)$. Let $\delta = \min_{t_0 \leq t \leq T}\{\varepsilon/f(t), \delta_{4.5}(t,p)\}$.

Let G be a graph which is δ-close to satisfying the k-clique r-balanced cut property. Apply the regularity lemma (Theorem 2.1) to this graph to obtain an ε-regular partition $\{V_i\}_{i=1}^t$ of the vertex set where t is divible by $2r$. For distinct $i, j \in [t]$, let d_{ij} be the density of the pair (V_i, V_j). Note that the parameters are chosen so that Lemma 4.4 and Lemma 4.5 holds.

For every regular $(k-3)$-tuple I, define a graph H_I as following. The vertex set of H_I is the collection of elements of $[t]\backslash I$ which form a regular $(k-2)$-tuple together with I. And $j_1, j_2 \in V(H_I)$ forms an edge if and only if the $(k-1)$-tuple $I \cup (j_1, j_2)$ is regular. Since each part of the regular partition forms a regular pair with at least $(1-\varepsilon)t$ of the other parts, we know that the graph H_I has at least $(1-k\varepsilon)t$ vertices and minimum degree at least $(1-2k\varepsilon)t$. Thus by Lemma 4.4, Lemma 4.5, Lemma 3.5 and the proof of Theorem 3.1, we know that there exists a p_I such that at least $(1-O(\varepsilon))$-proportion of the edges of H_I have density ε-equal to p_I.

Select k indices $j_1, \ldots, j_k$ out of $[t]$ independently and uniformly at random. With probability at least $1-O(\varepsilon)$, the k-tuple is regular. Moreover, with probability at least $1-O(\varepsilon)$, $d_{j_1 j_2}^{(j_4,\ldots,j_k)} =_\varepsilon p_{(j_4,\ldots,j_k)}$ and the same is true for other choices of indices as well. Therefore by the union bound, the k-tuple $(j_1, \ldots, j_k)$ is excellent with probability at least $1-O(\varepsilon)$. Equivalently, the number of excellent k-tuples is at least $(1-O(\varepsilon))\binom{t}{k}$.

Call a pair of indices in $[t]$ *excellent* if it is contained in at least $\frac{2}{3}\binom{t}{k-2}$ excellent k-tuples. Assume that there are ηt^2 non-excellent edges. Then the number of non-excellent k-tuples are at least

$$\eta t^2 \times \frac{1}{3}\binom{t}{k-2} / \binom{k}{2} = \Omega(\eta)\binom{t}{k}.$$

Therefore, $\eta = O(\varepsilon)$ and there are at most $O(\varepsilon)t^2$ non-excellent edges. We claim that all the excellent edges are ε-equal to each other. Take two excellent edges e, f. Since each of these edges form an excellent k-tuple with more than $\frac{2}{3}\binom{t}{k-2}$ of the $(k-2)$-tuples, there exists an $(k-2)$-tuple which forms an excellent k-tuple with both of these edges. Thus by Lemma 4.6 applied to each of these k-tuples separately, we can conclude that $d_e =_\varepsilon d_f$.

Consequently, by the total transitivity of ε-equality (see, Section 2), we can conclude that $d_e =_\varepsilon p'$ for some p' for every excellent edge e. Then apply the same reasoning as in the triangle case to show that $p' =_\varepsilon p$ and $G \in \mathcal{P}_2(1/2)$. This proves the quasi-randomness of the graph G. □

5 Concluding Remarks

In this paper, we proved the quasi-randomness of k-clique balanced cut properties for $k \geq 3$ and thus answered an open problem raised by both Shapira-Yuster [14] and Janson [10]. The most important base case was $k = 3$ where we solved a system of equations given by Lemma 3.3. The existence of "bad" vertex in Lemma 3.5 complicated the proof of the main theorem. It is hard to believe that the case can be significantly simplified since even if we assume that all the pairs are regular in the regular partition, there is an assignment of variables x_i and d_{ij} which is not all constant but forms a solution of the system.

We conclude this paper with an open problem for further study.

Question 5.1 *Let k, r be positive integers satisfying $r \geq k \geq 3$. Let H be a nonempty graph on k vertices, and assume that every $(1/r, \cdots, 1/r)$-cut of a graph G has the "correct" number of copies of H such that every vertex of H is in a different part of the cut. Does this condition force G to be quasi-random?*

Acknowledgement. We would like to thank Asaf Shapira, and our advisor Benny Sudakov for the kindness and advice they provided that greatly helped us in doing this work. We also want to thank Svante Janson for valuable corrections.

References

[1] K. Azuma, Weighted sums of certain dependent random variables, *Tôkuku Math. J.* **19** (1967), 357–367.

[2] F.R.K. Chung, R.L. Graham, and R.M. Wilson, Quasi-random graphs, *Combinatorica* **9**(4) (1989), 345–362.

[3] F.R.K. Chung, R.L. Graham, and R.M. Wilson, Quasi-random set systems, *Journal of the AMS* **4** (1991), 151–196.

[4] F.R.K. Chung and R.L. Graham, Quasi-random tournaments, *Journal of Graph Theory* **15** (1991), 173–198.

[5] F.R.K. Chung and R.L. Graham, Quasi-random hypergraphs, *Random Structures and Algorithms* **1** (1990), 105–124.

[6] F.R.K. Chung, R.L. Graham, and R.M. Wilson, Maximum cuts and quasi-random graphs, *Random Graphs*, (Poznan Conf., 1989) Wiley-Intersci, Publ. vol 2, 151–196.

[7] D.H. Gottlieb, A class of incidence matrices, *Proc. Amer. Math. Soc.* **17** (1966), 1233–1237.

[8] A. Hajnal and E. Szemerédi, Proof of a conjecture of P. Erdős, In *Combinatorial theory and its applications, II (Proc. Colloq., Balatonfüred, 1969)*, North-Holland, Amsterdam (1970), 601–623.

[9] W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Ameri. Statist. Assoc.*, **58** (1963), 13–30.

[10] S. Janson, Quasi-random graphs and graph limits, arXiv:0905.3241 [math.CO].

[11] J. Komlós and M. Simonovits, Szemerédi's regularity lemma and its applications in graph theory, In *Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993)*, volume 2 of *Bolyai Soc. Math. Stud.*, pages 295–352. János Bolyai Math. Soc., Budapest, 1996.

[12] M. Krivelevich and B. Sudakov, Pseudo-random graphs, More sets, graphs and numbers, E. Györi, G. O. H. Katona and L. Lovász, Eds., Bolyai Society Mathematical Studies Vol. 15, 199–262.

[13] C. McDiarmid, Concentration, **Probabilistic Methods for Algorithmic Discrete Mathematics** (1998), 1–46.

[14] A. Shapira and R. Yuster, The quasi-randomness of hypergraph cut properties, arXiv:1002.0149v1 [math.CO].

[15] M. Simonovits and V. T. Sós, Hereditarily extended properties, quasi-random graphs and not necessarily induced subgraphs, *Combinatorica*, **17** (1997), 577–596.

[16] E. Szemerédi, Regular partitions of graphs, In *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, volume 260 of *Colloq. Internat. CNRS*, pages 399–401. CNRS, Paris, 1978.

[17] A. Thomason, *Pseudo-random graphs*, in: Proceedings of Random Graphs, Poznań 1985, M. Karoński, ed., Annals of Discrete Math. **33** (North Holland 1987), 307–331.

[18] A. Thomason, *Random graphs, strongly regular graphs and pseudo-random graphs*, Surveys in Combinatorics, 1987, C. Whitehead, ed., LMS Lecture Note Series **123** (1987), 173–195.