Triangle factors in sparse pseudo-random graphs

Michael Krivelevich [*] Benny Sudakov [†] Tibor Szabó [‡]

Abstract

The goal of the paper is to initiate research towards a general, Blow-up Lemma type embedding statement for pseudo-random graphs with *sublinear* degrees. In particular, we show that if the second eigenvalue λ of a d-regular graph G on $3n$ vertices is at most $cd^3/n^2 \log n$, for some sufficiently small constant $c > 0$, then G contains a triangle factor. We also show that a fractional triangle factor already exists if $\lambda < 0.1d^2/n$. The latter result is seen to be best possible up to a constant factor, for various values of the degree $d = d(n)$.

1 Introduction

Let H be a fixed graph on n vertices. We say that a graph G on n vertices has an H*-factor* if G contains n/h vertex disjoint copies of H. For example, if H is just an edge $H = K_2$, then an H-factor is a perfect matching in G. Of course, a trivial necessary condition for the existence of an H-factor in G is that h divides n.

Providing sufficient conditions for the existence of an H-factor is one of the most important lines of research in Extremal Graph Theory. Such renowned results as matching theorems of Hall and Tutte can certainly be classified into this category. Usually we assume that H is fixed while the order n of the underlying graph G grows. In many cases sufficient conditions are formulated in terms of the minimum degree of G. For example, a classical result of Corrádi and Hajnal [11] asserts that every graph on n vertices with minimum degree at least $2n/3$ contains a triangle factor (assuming again 3 divides n). This was generalized later by Hajnal and Szemerédi [16] who proved that if the minimum degree $\delta(G)$ satisfies $\delta(G) \geq (1 - \frac{1}{r})n$, then G contains $\lfloor n/r \rfloor$ vertex disjoint copies of K_r. The statement of this theorem is easily seen to be tight. Therefore for sparse graphs we need more subtle conditions to

[*]Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel. Email: krivelev@math.tau.ac.il. Supported by a USA-Israeli BSF grant, by a grant from the Israel Science Foundation and by a Bergmann Memorial Award.

[†]Department of Mathematics, Princeton University, Princeton, NJ 08540, USA and Institute for Advanced Study, Princeton, NJ 08540, USA. Email address: bsudakov@math.princeton.edu. Research supported in part by NSF grants DMS-0106589, CCR-9987845 and by the State of New Jersey.

[‡]Department of Computer Science, ETH Zürich, 8092 Switzerland. Email address: szabo@inf.ethz.ch. Research supported in part by NSF grant DMS 99-70270 and by the joint Berlin/Zurich graduate program Combinatorics, Geometry, Computation, financed by the German Science Foundation (DFG) and ETH Zürich.

guarantee the existence of an H-factor. It turns out that *pseudo-randomness* allows in many cases to obtain results which fail to hold for general graphs of the same edge density.

Pseudo-random graphs can be informally described as graphs whose edge distribution resembles closely that of a truly random graph $G(n,p)$ of the same edge density. They have been a subject of intensive study during the last two decades and have seen numerous applications both in combinatorics and theoretical computer science (see, e.g. [23], [24], [10], [22], [4] and the recent survey [19]).

Among several possible ways, we choose the one through eigenvalues to formulate precisely the concept of "pseudo-random graph". This will enable us to use the powerful and well developed machinery of spectral graph theory (see [9]) to connect between the eigenvalues of a graph and its edge distribution. For simplicity, we restrict our attention to regular pseudo-random graphs (our methods should apply also to almost regular graphs, but exact statements are harder to formulate). Some definitions are in place here. Let $G = (V, E)$ be a graph with vertex set $V = \{1, \ldots, n\}$. The adjacency matrix $A = A(G)$ is an n-by-n $0,1$-matrix whose entry a_{ij} is 1 whenever $(i, j) \in E(G)$, and is 0 otherwise. As A is a real symmetric matrix, all its eigenvalues are real. We thus denote the eigenvalues of A, usually also called the eigenvalues of the graph G itself, by $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. In case G is a d-regular graph, it follows from the Perron-Frobenius Theorem that $\lambda_1 = d$ and $|\lambda_i| \leq d$ for all $2 \leq i \leq n$. Let now $\lambda = \lambda(G) = \max\{|\lambda_i(G)| : i = 2, 3, \ldots, n\}$. The parameter λ is usually called *the second eigenvalue of* G.

It is well known that the larger the so called spectral gap (i.e. the difference between the order of magnitude of d and λ) is, the more closely the edge distribution of G approximates that of a random graph $G(n, d/n)$. We will cite relevant quantitative results in the next section, for now we just state that the value of λ will serve us as the measure of randomness.

The subject of this paper is to obtain a sparse version of the result of Corrádi and Hajnal. We show that under certain conditions pseudo-randomness ensures the existence of a triangle factor. First we discuss the case of constant edge density p. In this case the celebrated Blow-up Lemma of Komlós, Sárközy and Szemerédi [17] can be used to show the existence of H-factors.

In order to formulate the Blow-up Lemma we need to introduce the notion of a super-regular pair. Given $\epsilon > 0$ and $0 < p < 1$, a bipartite graph G with bipartition (V_1, V_2), $|V_1| = |V_2| = n$, is called *super* (p, ϵ)*-regular* if for all vertices $v \in V(G)$,

$$(p - \epsilon)n \leq d(v) \leq (p + \epsilon)n \ ,$$

and for every pair of sets (U, W), $U \subset V_1$, $W \subset V_2$, $|U|, |W| \geq \epsilon n$,

$$\left| \frac{e(U, W)}{|U||W|} - \frac{|E(G)|}{n^2} \right| \leq \epsilon \ .$$

The Blow-up Lemma roughly states that any r-partite graph H with parts of order $n > n_0(\epsilon, r, p, \Delta)$, and of maximum degree Δ is a subgraph of any r-partite graph G with parts $V_1, \ldots, V_r$ of order also n, for which all $\binom{r}{2}$ bipartite subgraphs $G[V_i, V_j]$ are super (p, ϵ)-regular. In particular such G contains an H-factor for every fixed graph with chromatic number r. This version of the lemma, due to Rödl and Ruciński [21], is somewhat different from and yet equivalent to the original formulation of Komlós et al. We use it here as it is closer in spirit to the notion of pseudo-randomness.

The Blow-up Lemma is a very powerful embedding tool. Using it together with the properties of pseudo-random graphs (see the next section) one can easily prove that if the vertex degree d is linear in the number of vertices n, then already a very weak condition on the spectral gap $\lambda = o(d)$ guarantees a triangle factor. As usual, the case of a vanishing edge density $p = o(1)$ is significantly more complicated. Here a sufficient condition for the existence of an H-factor should depend heavily on the graph H, as there may exist quite dense pseudo-random graphs without a single copy of H. We discuss such examples in Section 6. The case when $H = K_2$ is relatively simple. Indeed, already an extremely weak pseudo-randomness condition, $\lambda \leq d - 2$, suffices to guarantee a perfect matching (see [19]).

In this paper we consider the first substantial case, when $H = K_3$. We prove the following theorem, that provides a sufficient condition for a triangle factor in a pseudo-random regular graph G and does not require its degree d to be linear in $|V(G)|$. This result can also be interpreted as a first attempt to extend the Blow-up Lemma to sparse graphs.

Theorem 1.1 *Let G be a d-regular graph on n vertices, such that n is divisible by 3. If*

$$\lambda(G) = o\left(\frac{d^3}{n^2 \log n}\right),$$

then G has a triangle factor.

Our proof uses properties of random matchings in bipartite graphs similar to that in [21]. But applying directly the approach of [21], one can only infer the existence of a triangle factor in pseudo-random graphs, which are quite dense. So to obtain the bound in Theorem 1.1, new ideas are needed.

Since for graphs with sublinear degrees $\lambda(G) = \Omega(\sqrt{d})$ (see Proposition 2.3 below), Theorem 1.1 starts working for degrees $d = d(n)$ asymptotically larger than $n^{4/5} \log^{2/5} n$. The next statement, whose proof is based on a construction of Alon [1], shows that $4/5$ cannot be replaced by a power smaller than $2/3$.

Theorem 1.2 *For every function $d(n)$ such that $\Omega(n^{2/3}) \leq d \leq n$ there exists a triangle-free graph G that satisfies:*

1. *$|V(G)| = n_1 = \Theta(n)$;*
2. *G is d_1-regular with $d_1 = \Theta(d)$;*
3. *$\lambda(G) = O({d_1}^2/n_1)$.*

We tend to think that the condition of Theorem 1.1 can be further improved, maybe even down to the bound of Theorem 1.2. To substantiate this conjecture we discuss the fractional variant of the problem. Let us introduce the relevant terminology. Let $G = (V, E)$ be a graph. Denote by $T = T(G)$ the set of all triangles of G. A function $f : T \to \mathbb{R}_+$ is called a *fractional triangle factor* if for every $v \in V(G)$ one has $\sum_{t \in T, v \in t} f(t) = 1$. If G contains a triangle factor T_0, then assigning values $f(t) = 1$ for all $t \in T_0$, and $f(t) = 0$ for all other $t \in T$ produces a fractional triangle factor. This simple argument shows that the existence of a triangle factor in G implies the existence of a fractional triangle factor. The converse statement is easily seen to be invalid in general.

It turns out that this fractional relaxation of the integer problem is much more tractable, and quite precise result can be obtained about it.

Theorem 1.3 *Let $G = (V, E)$ be a d-regular graph on n vertices. If*

$$\lambda(G) \leq \frac{0.1d^2}{n},$$

then G has a fractional triangle factor.

The lack of triangles of course prevents the existence of a fractional triangle factor. So Theorem 1.2 shows that for any sensible degree d the assertion of Theorem 1.3 is best possible up to a constant factor.

The rest of the paper is organized as follows. The next section collects useful properties of pseudo-random graphs together with some simple facts about eigenvalues. In Section 3 we provide estimates for the number of perfect matchings in bipartite graphs, to be used as technical tools in the proof of our main result. Section 4 is devoted to the proof of Theorem 1.1, the sufficient condition for a triangle factor. We prove Theorem 1.3 in Section 5. In Section 6 we present a construction of pseudo-random triangle-free graphs of various vertex degrees which will show the optimality of the fractional result of Section 5. Section 7, the last section of the paper, is devoted to concluding remarks and discussion of relevant open problems.

For a vertex v, $N(v)$ denotes the neighborhood of v, and $d(v)$ denotes its degree. For a subset X of the vertex set, $N_X(v) = N(v) \cap X$, $d_X(v) = |N_X(v)|$. The number of common neighbors of vertices u and v is denoted by $d(u, v)$. Given subsets of vertices $B, C \subseteq V(G)$, $e(B, C)$ denotes the number of ordered pairs (u, v), where $u \in B, v \in C$ and $(u, v) \in E(G)$; of course when B and C are disjoint, $e(B, C)$ is just the number of edges of G connecting B and C. Finally a graph $G = (V, E)$ is called an (n, d, λ)*-graph* if it is d-regular, has n vertices and the second eigenvalue of G is at most λ. All logarithms in this paper are natural. We will also routinely omit floor and ceiling signs in order to simplify the presentation.

2 Properties of (n, d, λ)-graphs

This section contains description of basic facts about (n, d, λ)-graphs. For a more detailed information about pseudo-random graphs we refer the interested reader to the recent survey of the first two authors [19].

The following two results, quite well known by now, provide a quantitative support to the thesis that spectral gap in regular graphs implies pseudo-randomness. This is especially visible in Theorem 2.2, where it is stated explicitly that the number of edges between any pair of sufficiently large sets

B, C in an (n, d, λ)-graph is close to its expected value $|B||C|d/n$ in the random graph $G(n, d/n)$ if the second eigenvalue λ, governing the error term, is relatively small.

Theorem 2.1 *([6, page122, Theorem 2.4]) Let G be an (n, d, λ)-graph. For any $B \subseteq V = V(G)$,*

$$\sum_{v \in V} (N_B(v) - |B|d/n)^2 \leq \lambda^2 |B|(n - |B|)/n.$$

Theorem 2.2 *([6, page 122, Corollary 2.5]) Let G be an (n, d, λ)-graph. For any $B \subseteq V(G), C \subseteq V(G)$,*

$$|e(B, C) - |B||C|d/n| \leq \lambda \sqrt{|B||C|}.$$

Theorems 2.1, 2.2 indicate that the smaller the second eigenvalue λ of a d-regular graph G on n vertices is, the closer its edge distribution approximates that of the truly random graph $G(n, d/n)$. A legitimate question is thus how small can λ be for given n and d? As the following easy proposition shows, $\lambda = \Omega(\sqrt{d})$ unless the degree d is very close to n.

Proposition 2.3 *Let G be an (n, d, λ)-graph. Then $\lambda \geq \sqrt{(nd - d^2)/(n-1)}$. In particular for $d \leq n/2$, $\lambda \geq \sqrt{d}/\sqrt{2}$.*

Proof. Let A be the adjacency matrix of G. The trace of A^2 is easily seen to be equal to the number of ones in A, which is exactly $2|E(G)| = nd$. We thus obtain:

$$nd = Tr(A^2) = \sum_{i=1}^{n} \lambda_i^2 \leq d^2 + (n-1)\lambda^2.$$

Solving the above inequality for λ establishes the claim of the proposition. □

Theorem 2.4 *Let G be an (n, d, λ)-graph. For an integer $r \geq 2$, define:*

$$s_r = \frac{\lambda n}{d} \left(1 + \frac{n}{d} + \ldots + \left(\frac{n}{d}\right)^{r-2} \right).$$

Then every set of more than s_r vertices of G spans a copy of K_r.

This theorem can be proved by induction on r using repeatedly Theorem 2.2. A proof can be found in [3].

Given a graph $G = (V, E)$, a *k-blow-up* of G, denoted by $G(k)$, is obtained be replacing each vertex of G by an independent set of size k and connecting two vertices of $G(k)$ by an edge if and only if the corresponding vertices of G are connected by an edge.

Proposition 2.5 *Let G be an (n, d, λ)-graph. Then the blow-up $G(k)$ is a $(kn, kd, k\lambda)$-graph.*

Proof. It is obvious that $G(k)$ has kn vertices and is kd-regular. In order to estimate the eigenvalues of $G(k)$, recall that the *Kronecker product* of an $m \times n$ dimensional matrix $A = (a_{ij})$ and an $s \times t$-dimensional matrix $B = (b_{kl})$ is an $ms \times nt$-dimensional matrix $A \otimes B$, whose entry labeled $((i, k)(j, l))$

is $a_{ij}b_{kl}$. In case A and B are symmetric matrices with spectra $\{\lambda_1, \dots, \lambda_n\}$, $\{\mu_1, \dots, \mu_t\}$ respectively, it is a simple consequence of the definition that the spectrum of $A \otimes B$ is $\{\lambda_i \mu_k : i = 1, \dots, n, k = 1, \dots, t\}$ (see, e.g. [12]). Observe that the adjacency matrix of $G(k)$ is formed by taking the Kronecker product of the adjacency matrix of G with the all-1 matrix J_k of order k-by-k. Therefore the eigenvalues of $G(k)$ are obtained by taking all possible products of an eigenvalue of G and that of J_k. Since the distinct eigenvalues of J_k are k and 0, the eigenvalues of $G(k)$ are obtained by multiplying the eigenvalues of G by k, and adding 0 to this list, in case it is not there already. This immediately implies that $\lambda(G(k)) = k\lambda(G)$. $\square$

3 Perfect matchings in pseudo-random bipartite graphs

In this, mostly technical, section we obtain estimates on the number of perfect matchings in bipartite graphs with some pseudo-random properties, to be used later in the proof of Theorem 1.1. Our approach is based on known estimates for permanents of matrices and is similar to the one used in [5]. First we need a lemma about the existence of factors of large degree in pseudo-random bipartite graphs.

Lemma 3.1 *Let H' be a bipartite graph with bipartition $V(H') = U \cup V$, $|U| = |V| = n$, such that $|d(v) - d| < r$ for every vertex $v \in V(H')$. Suppose furthermore that there exists a λ such that for any $A \subseteq U, B \subseteq V$*

$$\left| e(A,B) - \frac{|A||B|d}{n} \right| < \lambda \sqrt{|A||B|}.$$

Let $t \geq 0$ be an integer and let $k = d - 2r - 2t - \lambda$. Let H be a subgraph of H' obtained by deleting some edges of H' so that at any vertex not more than t incident edges are deleted.

If $\lambda < \frac{d^2}{20n}$, and $t + r \leq \frac{d^2}{40n}$, then H has a k-factor.

Proof. The existence of a k-factor [20, page 70, Theorem 2.4.2], is equivalent to the property that for every two subsets $X \subseteq U$, $Y \subseteq V$

$$e_H(U - X, V - Y) \geq k(n - |X| - |Y|).$$

This condition can be reformulated in the following more convenient form: For every subsets $X \subseteq U$, $Y \subseteq V$ of cardinalities $|X| = x$, $|Y| = y$, respectively,

$$e_H(X, Y) \geq k(x + y - n)\ . \tag{1}$$

Checking the above condition appears to be a matter of routine, yet tedious calculations and case analysis. Following the referee's advice, we defer the technical details to the appendix. $\square$

From this lemma one can easily deduce the following estimate on the number of perfect matchings in H.

Corollary 3.2 *Let H be a bipartite graph as defined in Lemma 3.1. Then H has at least $(k/n)^n n!$ perfect matchings.*

Proof. To obtain this bound, observe that the van der Waerden conjecture on the permanent of doubly stochastic matrices, proved by Egorychev [13] and Falikman [14], implies that the number of matchings in a k-regular bipartite graph is at least $(k/n)^n n!$. $\square$

We finish this section with an upper bound on the number of perfect matchings in bipartite graphs with a given degree sequence.

Lemma 3.3 *Let H be a bipartite graph with bipartition $V(H) = U \cup V$, $|U| = |V| = n$, such that the maximum degree of a vertex from U is at most $d_2 > 6$. Suppose that there are x vertices from U having degree at most $d_1 > 6$. Then the number of perfect matchings in H is at most*

$$e^{-n} d_1^{x+\frac{x}{d_1}} d_2^{n-x+\frac{n-x}{d_2}}.$$

Proof. Let us denote by d_v the degree of vertex $v \in U$. By Brégman's Theorem [8] (see also [6]) the number of perfect matchings between U and V is at most

$$\prod_{v \in U} (d_v!)^{\frac{1}{d_v}}.$$

As the function $f(t) = t!^{1/t}$ is increasing and $t! < t(t/e)^t$ for $t > 6$, the above estimate is at most

$$(d_1!)^{\frac{x}{d_1}} (d_2!)^{\frac{n-x}{d_2}} < \left(d_1 \left(\frac{d_1}{e}\right)^{d_1}\right)^{\frac{x}{d_1}} \left(d_2 \left(\frac{d_2}{e}\right)^{d_2}\right)^{\frac{n-x}{d_2}} = e^{-n} d_1^{x+\frac{x}{d_1}} d_2^{n-x+\frac{n-x}{d_2}}. \quad \square$$

4 Triangle factors

In this section we prove our main result, Theorem 1.1. Since it is known that $\lambda(G) \geq \sqrt{d}/\sqrt{2}$ for $d \leq n/2$ (see Proposition 2.3), the assumptions of the theorem imply that $d > n^{4/5}$ and $\lambda(G) = O(d/\log n)$. We will assume whenever it is needed that n is sufficiently large. For the sake of simplicity we will also write λ instead of $\lambda(G)$. Our proof borrows some ideas from the argument of [21].

Let $d_1 = d/3$, $n_1 = n/3$ and $r = 2\sqrt{d \log n}$. Clearly $d_1/n_1 = d/n$. Consider a random partition of $V(G)$ into three equal parts V_1, V_2, V_3. A standard application of the Chernoff-type bounds gives that with positive probability,

$$|d_{V_i}(v) - d_1| \leq r \text{ for every } v \in V \text{ and } i = 1, 2, 3. \tag{2}$$

Let us fix a particular partition $V_1 \cup V_2 \cup V_3 = V$, $|V_1| = |V_2| = V_3| = n/3 = n_1$, which has property (2) and denote by H' the bipartite subgraph of G, which contains all the edges between V_1 and V_2. By definition and by Theorem 2.2, H' satisfies the conditions of Lemma 3.1.

We call a pair of vertices $u, v \in V$ *bad* if for some $i = 1, 2$ or 3,

$$\left| d_{V_i}(u, v) - \frac{d_1^2}{n_1} \right| > \frac{d_1^2}{3n_1}.$$

Otherwise the pair is called *good*. First we need to estimate from above the number of bad pairs containing a particular vertex of G.

Lemma 4.1 *The number of vertices forming a bad pair with a fixed vertex $u \in V(G)$ is at most $\lambda + r$.*

Proof. Let us first estimate the number s of vertices, which form a bad pair with u because of the size of their common neighborhood in V_1. Let $B = N_{V_1}(u)$. By (2), $\big||B| - d_1\big| \le r$. A vertex $v \in V(G)$ forms a bad pair with u if and only if $\Big||N_B(v)| - d_1^2/n_1\Big| > d_1^2/(3n_1)$. Then by applying Theorem 2.1 we obtain

$$s \cdot \frac{d_1^4}{9n_1^2} \le \sum_{v \in V(H)} \Big(|N_B(v)| - \frac{d_1^2}{n_1}\Big)^2 \le 2 \sum_{v \in V(H)} \left(\Big(|N_B(v)| - \frac{|B|}{n} d\Big)^2 + \Big(\frac{|B|}{n} d - \frac{d_1^2}{n_1}\Big)^2\right)$$

$$= 2 \sum_{v \in V(H)} \Big(|N_B(v)| - \frac{|B|}{n} d\Big)^2 + 2 \sum_{v \in V(H)} \Big(\frac{d_1}{n_1}(|B| - d_1)\Big)^2 \le 2\lambda^2 |B| \Big(1 - \frac{|B|}{n}\Big) + 2n \frac{r^2 d_1^2}{n_1^2}$$

$$\le 2\lambda^2(d_1 + r) + 6\frac{r^2 d_1^2}{n_1} \le 3\lambda^2 d_1 + 6\frac{r^2 d_1^2}{n_1}.$$

Therefore

$$s \le \frac{27\lambda^2 n_1^2}{d_1^3} + \frac{54 r^2 n_1}{d_1^2} = \lambda \frac{27\lambda n_1^2}{d_1^3} + r\frac{54 r n_1}{d_1^2} \le \frac{1}{3}(\lambda + r).$$

The last inequality follows from the facts that $27\lambda n_1^2/d_1^3 \le O(\log^{-1} n_1)$, $d_1 > n_1^{4/5}$ and n_1 is large enough.

To obtain the assertion of the lemma, note that the number of vertices which form a bad pair with u because of the size of their common neighborhood in V_2, V_3 can be bound similarly. The sum of these three bounds is $\lambda + r$. This completes the proof. $\square$

An edge $e = (u, v)$, $u \in V_1, v \in V_2$, is called *bad* if its endpoints form a bad pair. Let H be the bipartite subgraph of G containing all edges between V_1 and V_2, except the bad ones. By Lemma 4.1, H is a subgraph of H' obtained by deleting some edges of H' so that at any vertex not more than $\lambda + r$ incident edges are deleted. Thus H satisfies the condition of Lemma 3.1 with parameter $t = \lambda + r$ and therefore, by Corollary 3.2, has a lot of perfect matchings. Given a perfect matching M in H, define an auxiliary bipartite graph $H(M)$ on the vertex set $V(H(M)) = V_3 \cup M$ by connecting a vertex $v \in V_3$ to an edge $e = \{x_1, x_2\} \in M$, if v, x_1 and x_2 form a triangle in G.

Note that by definition, a perfect matching in $H(M)$ corresponds to a triangle factor in G. Hence our ultimate goal is to prove that there exists a perfect matching M of H, for which $H(M)$ satisfies Hall's Condition. First we bound from below the degree of a vertex from M in $H(M)$. Since by our construction every $e \in M$ is a good edge, we get that the degree of e in $H(M)$ is at least $2d_1^2/(3n_1)$ for any perfect matching M. Using this fact we can easily deduce the following statement.

Lemma 4.2 *Let M be an arbitrary perfect matching of H and let X be a subset of V_3, such that its neighborhood in $H(M)$ has size less than $|X|$. Then the size of X is less than $n_1 - d_1/2$.*

Proof. Let $(v_i, u_i), v_i \in V_1, u_i \in V_2, i = 1, \dots, s$, be the set of edges of the matching M which are not adjacent in the graph $H(M)$ to any vertex in X, and let $Y = V_3 - X$. If the neighborhood of X has size less than $|X|$, then $|Y| < s$. By the definition of $H(M)$, all (v_i, u_i) are good edges and therefore in graph G the vertices v_i, u_i have at least $2d_1^2/(3n_1)$ common neighbors, all of them belonging to the set Y. This implies that the number of edges of G between the sets Y and $\{v_1, \ldots, v_s\}$ is at least $(2d_1^2/(3n_1))s$. On the other hand, by Theorem 2.2, this number is at most $|Y|sd_1/n_1 + \lambda\sqrt{|Y|s}$. Using that $|Y| < s$ and $\lambda = o(d_1^2/n_1)$ we obtain that

$$\frac{2d_1^2}{3n_1}s \leq |Y|s\frac{d_1}{n_1} + \lambda\sqrt{|Y|s} < |Y|s\frac{d_1}{n_1} + \lambda s < |Y|s\frac{d_1}{n_1} + \frac{d_1^2}{6n_1}s.$$

Therefore $|Y| > d_1/2$ and thus $|X| = n_1 - |Y| < n_1 - d_1/2$. This completes the proof. $\square$

In order to verify Hall's condition in $H(M)$ to subsets $X \subseteq V_3$ of size $|X| \leq n_1 - d_1/2$, just an arbitrary perfect matching of H will not do. Thus, we choose a perfect matching M in H *uniformly at random* and prove that Hall's condition fails with probability $o(1)$.

We first estimate degrees of vertices from V_3 in the graph $H(M)$. This gives us a lower bound on the size of a set $X \subset V_3$, that can violate Hall's condition.

Lemma 4.3 *The probability that there exists a vertex $v \in V_3$, whose degree in the bipartite graph $H(M)$ is less than $d_1^2/(n_1 \log n_1)$, is $o(1)$.*

Proof. We interpret a perfect matching M of H as a bijection σ_M from V_1 to V_2. Let v be a vertex of V_3 and let X and Y be the sets of its neighbors in V_1, V_2 respectively. Thus the degree of v in $H(M)$ is $|\sigma_M(X) \cap Y|$. To prove the lemma it is enough to show, that for every vertex v, $|\sigma_M(X) \cap Y| < d_1^2/(n_1 \log n_1)$ happens with probability less than $o(1/n_1)$.

Let $X_1 \subseteq X$ be the set of those vertices, which (in graph H) have *at least* $d_1^2/2n_1$ neighbors in Y. If $u \notin X_1$, then u forms a bad pair with v, as in graph G it has at most $d_1^2/2n_1 + \lambda + r < 2d_1^2/3n_1$ neighbors in Y. Lemma 4.1 then implies that $|X - X_1| \leq \lambda + r$. By (2), $d_1 + r > |X| \geq |X_1| \geq |X| - (\lambda + r) > d_1 - \lambda - 2r$, so we have $|X_1| = (1 - o(1))d_1$.

As $X_1 \subseteq X$, it is sufficient to prove $Pr(|\sigma_M(X_1) \cap Y| < d_1^2/(n_1 \log n_1)) = o(1/n_1)$ for every vertex $v \in V_3$. If $|\sigma_M(X_1) \cap Y| < d_1^2/(n_1 \log n_1)$, then there is a subset $S \subseteq X_1$ of size $|S| = |X_1| - d_1^2/(n_1 \log n_1)$, such that $|\sigma_M(S) \cap Y| = 0$.

Denote by z_1 the number of perfect matchings in the bipartite graph H. Let $S \subseteq X_1$ be a set of size $|S| = |X_1| - d_1^2/n_1 \log n_1$ and denote by $z_2(S)$ the number of perfect matchings M of H which have the property $|\sigma_M(S) \cap Y| = 0$. The maximum of the $z_2(S)$, taken over all subsets $S \subseteq X_1$ of size

$|S| = |X_1| - d_1^2/n_1 \log n_1$, is denoted by z_2^{max}. Then it is easy to see that

$$\begin{aligned}
Pr\left(|\sigma_M(X_1) \cap Y| < \frac{d_1^2}{n_1 \log n_1}\right) &\le \sum_{\substack{S \subseteq X_1 \\ |S| = |X_1| - d_1^2/n_1 \log n_1}} Pr\left(|\sigma_M(S) \cap Y| = 0\right) \\
&= \sum_{\substack{S \subseteq X_1 \\ |S| = |X_1| - d_1^2/n_1 \log n_1}} \frac{z_2(S)}{z_1} \\
&\le \binom{|X_1|}{\frac{d_1^2}{n_1 \log n_1}} \frac{z_2^{max}}{z_1} \le \left(\frac{e|X_1| n_1 \log n_1}{d_1^2}\right)^{\frac{d_1^2}{n_1 \log n_1}} \frac{z_2^{max}}{z_1} \\
&\le \left(\frac{3n_1 \log n_1}{d_1}\right)^{\frac{d_1^2}{n_1 \log n_1}} \frac{z_2^{max}}{z_1} \le e^{\frac{d_1^2}{4n_1}} \frac{z_2^{max}}{z_1}.
\end{aligned}$$

In the last inequality we use that $3n_1 \log n_1/d_1 < n_1^{1/4}$.

As we already mentioned, the bipartite graph H satisfies the condition of Lemma 3.1 with $t = r + \lambda$. Therefore by Corollary 3.2 the number of perfect matchings z_1 in H is at least $(k/n_1)^{n_1} n_1! > (k/e)^{n_1}$, where $k = d_1 - 2r - \lambda - 2t = d_1 - 4r - 3\lambda > d_1 - \lambda \log n_1$. On the other hand, $z_2(S)$ is the number of perfect matchings in the bipartite graph H_S obtained from H by deleting all the edges between S and Y. We proved that this graph has $|S| = s = |X_1| - d_1^2/(n_1 \log n_1) = (1 - o(1))d_1$ vertices of degree at most $d_1 + r - d_1^2/(2n_1)$ while the rest of the vertices have degree at most $d_1 + r$. Thus by Lemma 3.3

$$\begin{aligned}
z_2(S) &< e^{-n_1} \left(d_1 + r - \frac{d_1^2}{2n_1}\right)^{s + \frac{s}{d_1 + r - d_1^2/(2n_1)}} (d_1 + r)^{n_1 - s + \frac{n_1 - s}{d_1 + r}} \\
&\le d_1^{\frac{3n_1}{d_1}} e^{-n_1} \left(d_1 + r - \frac{d_1^2}{2n_1}\right)^s (d_1 + r)^{n_1 - s} .
\end{aligned}$$

Finally, we obtain that

$$\begin{aligned}
Pr\left(|\sigma_M(X_1) \cap Y| < \frac{d_1^2}{n_1 \log n_1}\right) &\le e^{\frac{d_1^2}{4n_1}} \frac{z_2^{max}}{z_1} \le e^{\frac{d_1^2}{4n_1}} \frac{d_1^{\frac{3n_1}{d_1}} e^{-n_1} \left(d_1 + r - \frac{d_1^2}{2n_1}\right)^s (d_1 + r)^{n_1 - s}}{\left((d_1 - \lambda \log n_1)/e\right)^{n_1}} \\
&\le e^{\frac{d_1^2}{4n_1}} d_1^{\frac{3n_1}{d_1}} \left(\frac{d_1 + r - \frac{d_1^2}{2n_1}}{d_1 - \lambda \log n_1}\right)^{(1 - o(1))d_1} \left(\frac{d_1 + r}{d_1 - \lambda \log n_1}\right)^{n_1} \\
&\le e^{\frac{d_1^2}{4n_1}} d_1^{\frac{3n_1}{d_1}} \left(1 - \frac{d_1}{3n_1}\right)^{(1 - o(1))d_1} \left(1 + \frac{3\lambda \log n_1}{d_1}\right)^{n_1} \\
&\le e^{\frac{d_1^2}{4n_1} + \frac{3n_1 \log n_1}{d_1} + \frac{3n_1 \lambda \log n_1}{d_1} - \frac{(1 - o(1))d_1^2}{3n_1}} \le e^{-(1 - o(1)) \frac{d_1^2}{12n_1}} = o(n_1^{-1}).
\end{aligned}$$

Here we use the facts that $\lambda n_1 \log n_1/d_1 = o(d_1^2/n_1)$ and $r < \lambda \log n_1 < 3n_1 \lambda \log n_1/d_1$. This completes the proof of the lemma. $\square$

Finally, it remains to verify Hall's condition for the subsets of V_3 of size between $d_1^2/(n_1 \log n_1)$ and $n_1 - d_1/2$. This is done in the following lemma.

Lemma 4.4 *With probability $1-o(1)$, every subset $X \subset V_3$ of size $d_1^2/(n_1 \log n_1) \le |X| \le n_1 - d_1/2$ has at least $|X|$ neighbors in the bipartite graph $H(M)$.*

Proof. Let $Y = Y(X)$ be the subset of V_1 composed of all vertices with at most $|X|d_1/(2n_1)$ neighbors in X and let x, y be the sizes of X, Y, respectively. Obviously, $e(X,Y) \le d_1 xy/(2n_1)$. On the other hand, by Theorem 2.2 the number of edges between X and Y satisfies

$$e(X,Y) \ge \frac{d_1 xy}{n_1} - \lambda\sqrt{xy}.$$

Comparing the above estimates for $e(X,Y)$ we derive that $d_1 xy/(2n_1) \ge d_1 xy/n_1 - \lambda\sqrt{xy}$, and therefore $y \le 4n_1^2\lambda^2/(xd_1^2)$.

Let $Y_1 = Y_1(X) = V_1 - Y$. Call an edge $(v,u) \in E_H(Y_1, V_2)$ *useless* (*for* X) if there are no edges from u to $N_X(v)$, otherwise call it *useful* (*for* X). We also say that u is a *useless neighbor* (*useful neighbor*, resp.) of v. Denote by W_v the set of all useless neighbors of v in V_2. Obviously, $e(W_v, N_X(v)) = 0$. On the other hand, it follows from Theorem 2.2 that

$$e(W_v, N_X(v)) \ge \frac{|W_v||N_X(v)|d_1}{n_1} - \lambda\sqrt{|W_v||N_X(v)|}\ .$$

Recalling that $|N_X(v)| \ge d_1 x/(2n_1)$, it follows that

$$|W_v| \le \frac{\lambda^2 n_1^2}{d_1^2 |N_X(v)|} \le \frac{2\lambda^2 n_1^3}{d_1^3 x}\ .$$

Note that if $e = (v,u) \in E_H(Y_1, V_2)$ is an edge of the perfect matching M of H, then e belongs to the neighborhood of X in the graph $H(M)$ if and only if e is a useful edge. Clearly, the probability $\mathbf{P_X}$ of X violating Hall's condition is at most the probability that *there exists* a set $S \subset Y_1$ of size $s = |Y_1| - x$, incident to *no* useful edge from M. Denote again by z_1 the number of perfect matchings in the bipartite graph H. Let $S \subseteq Y_1$ be a subset of size $s = |Y_1| - x$, and denote by $z_2(S)$ the number of perfect matchings M of H, which have only useless edges incident to the vertices of S. Then it is easy to see that

$$\mathbf{P_X} \le \sum_{\substack{S \subseteq Y_1 \\ |S| = |Y_1| - x}} \frac{z_2(S)}{z_1} \le \binom{|Y_1|}{s} \frac{\max_S z_2(S)}{z_1}.$$

Thus to finish the proof of the lemma it is enough to show that

$$\sum_{\substack{X \subseteq V_3 \\ |X| = x}} \mathbf{P_X} = o(1/n_1)$$

for all $d_1^2/n_1 \log n_1 \le x \le n_1 - d_1/2$.

First we consider the case when $x \le n_1/2$. Since $x \ge d_1^2/(n_1 \log n_1)$, then for every $v \in Y_1$ we have that

$$|W_v| \le \frac{2n_1^3\lambda^2}{d_1^3 x} \le \frac{2n_1^4\lambda^2 \log n_1}{d_1^5} = 2\left(\frac{n_1^2\lambda \log n_1}{d_1^3}\right)^2 \frac{d_1}{\log n_1} = o\left(\frac{d_1}{\log n_1}\right),$$

and

$$y \leq \frac{4n_1^2\lambda^2}{xd_1^2} \leq \frac{4n_1^3\lambda^2 \log n_1}{d_1^4} = 4 \ \frac{n_1^2 \lambda \log n_1}{d_1^3} \ \frac{\lambda}{d_1} \ n_1 = o(n_1).$$

As we already mentioned in the proof of Lemma 4.3, z_1 is at least $(d_1 - \lambda \log n_1)^{n_1} e^{-n_1}$. For any S, $z_2(S)$ is equal to the number of matchings in a bipartite subgraph of H with $s = |S|$ vertices of degree $|W_v| \leq d_1/\log n_1$. The rest of the vertices are of degree at most $d_1 + r$, therefore via Lemma 3.3 we obtain

$$\begin{aligned} z_2(S) \ &\leq \ e^{-n_1} \left(\frac{d_1}{\log n_1} \right)^{s + \frac{s}{d_1/\log n_1}} (d_1 + r)^{n_1 - s + \frac{n_1 - s}{d_1 + r}} \\ &\leq \ d_1^{\frac{2n_1 \log n_1}{d_1}} e^{-n_1} \left(\frac{d_1}{\log n_1} \right)^s \left(d_1 + r \right)^{n_1 - s}. \end{aligned}$$

Now, taking into account that $s = |Y_1| - x = n_1 - y - x \geq n_1 - o(n_1) - n_1/2 > n_1/3$, we can deduce that

$$\begin{aligned} \sum_{\substack{X \subseteq V_3 \\ |X| = x}} \mathbf{P_X} \ &\leq \ \sum_{\substack{X \subseteq V_3 \\ |X| = x}} \ \sum_{\substack{S \subseteq Y_1 \\ |S| = |Y_1| - x}} \frac{z_2(S)}{z_1} \leq \sum_{\substack{X \subseteq V_3 \\ |X| = x}} \ \sum_{\substack{S \subseteq Y_1 \\ |S| = |Y_1| - x}} \frac{d_1^{\frac{2n_1 \log n_1}{d_1}} e^{-n_1} (d_1/\log n_1)^s (d_1 + r)^{n_1 - s}}{(d_1 - \lambda \log n_1)^{n_1} e^{-n_1}} \\ &\leq \binom{n_1}{x} \binom{|Y_1|}{s} e^{\frac{2n_1 \log^2 n_1}{d_1}} \left(\frac{2}{\log n_1} \right)^{n_1/3} \left(\frac{d_1 + r}{d_1 - \lambda \log n_1} \right)^{n_1} \leq 2^{3n_1} \left(\frac{2}{\log n_1} \right)^{n_1/3} 2^{n_1} = o(n_1^{-1}). \end{aligned}$$

Here we use that $r < \lambda \log n_1 = o(d_1)$ and $2n_1 \log^2 n_1/d_1 = o(n_1)$.

Next assume that $n_1/2 \leq x \leq n_1 - d_1/2$. Then

$$y \leq \frac{4n_1^2\lambda^2}{xd_1^2} \leq \frac{8n_1\lambda^2}{d_1^2} = 8\frac{\lambda n_1^2 \log n_1}{d_1^3} \ \frac{\lambda}{n_1} \ \frac{d_1}{\log n_1} = o\left(\frac{d_1}{\log n_1} \right)$$

and hence $s = |Y_1| - x = n_1 - y - x = (n_1 - x) - o(d_1/\log n_1)$. Also in this case we have for every $v \in Y_1$ that

$$|W_v| \leq \frac{2n_1^3\lambda^2}{d_1^3 x} \leq \frac{4n_1^3\lambda^2}{d_1^3 n_1} = 4\frac{n_1^2 \lambda \log n_1}{d_1^3} \ \frac{\lambda}{\log n_1} = o\left(\frac{\lambda}{\log n_1} \right).$$

Similarly as before, $z_2(S)$ is equal to the number of matchings in a bipartite subgraph of H with s vertices of degree less than $|W_v| < \lambda/\log n_1$. The maximum degree is again at most $d_1 + r$. Therefore by Lemma 3.3 we obtain

$$\begin{aligned} z_2(S) \ &< \ e^{-n_1} \left(\frac{\lambda}{\log n_1} \right)^{s + \frac{s}{\lambda/\log n_1}} (d_1 + r)^{n_1 - s + \frac{n_1 - s}{d_1 + r}} \\ &\leq \ d_1^{\frac{2n_1 \log n_1}{\lambda}} e^{-n_1} \left(\frac{\lambda}{\log n_1} \right)^s (d_1 + r)^{n_1 - s} \ . \end{aligned}$$

We finish the proof of this case by the following estimate.

$$
\begin{aligned}
\sum_{\substack{X\subseteq V_3\\|X|=x}} \mathbf{P_X} &\le \sum_{\substack{X\subseteq V_3\\|X|=x}} \sum_{\substack{S\subseteq Y_1\\|S|=|Y_1|-x}} \frac{z_2(S)}{z_1} \le \sum_{\substack{X\subseteq V_3\\|X|=x}} \sum_{\substack{S\subseteq Y_1\\|S|=|Y_1|-x}} \frac{d_1^{\frac{2n_1\log n_1}{\lambda}} e^{-n_1}(\lambda/\log n_1)^s(d_1+r)^{n_1-s}}{(d_1-\lambda\log n_1)^{n_1}e^{-n_1}} \\
&\le \binom{n_1}{n_1-x}\binom{|Y_1|}{s}\frac{d_1^{\frac{2n_1\log n_1}{\lambda}}(\lambda/\log n_1)^s(d_1+r)^{n_1-s}}{(d_1-\lambda\log n_1)^{n_1}} \\
&\le \left(\frac{en_1}{n_1-x}\right)^{n_1-x}\left(\frac{e|Y_1|}{s}\right)^s\frac{d_1^{\frac{2n_1\log n_1}{\lambda}}(\lambda/\log n_1)^s(d_1+r)^{n_1-s}}{(d_1-\lambda\log n_1)^{n_1}} \\
&= d_1^{\frac{2n_1\log n_1}{\lambda}}\left(\frac{en_1}{n_1-x}\right)^{n_1-x-s}\left(\frac{e^2n_1|Y_1|\lambda}{(n_1-x)s(d_1+r)\log n_1}\right)^s\left(\frac{d_1+r}{d_1-\lambda\log n_1}\right)^{n_1} \\
&\le e^{\frac{2n_1\log^2 n_1}{\lambda}}\, e^{y\log n_1}\left(\frac{4e^2n_1^2\lambda}{(1+o(1))d_1^3\log n_1}\right)^s\left(1+\frac{3\lambda\log n_1}{d_1}\right)^{n_1} \\
&\le e^{d_1}\left(\frac{1}{\log n_1}\right)^{(1+o(1))d_1/2} e^{\frac{3n_1\lambda\log n_1}{d_1}} = o(1/n_1).
\end{aligned}
$$

In these computations we used that $n_1-x\ge d_1/2$, $y=n_1-x-s=o(d_1/\log d_1)$, $s\ge(1+o(1))d_1/2$, $n_1^2\lambda/d_1^3=o(1/\log n_1)$, $2n_1\log^2 n_1/\lambda=o(d_1)$, $n_1\lambda\log n_1/d_1=o(d_1)$ and $r<\lambda\log n_1=o(d_1)$. This completes the proof of the lemma. □

Finally, to complete the proof of Theorem 1.1 note that by Lemmas 4.3 and 4.4, with probability $1-o(1)$ there exists a matching M such that the bipartite graph $H(M)$ satisfies Hall's condition. Pick such a matching M. Then $H(M)$ contains a perfect matching which supplies a triangle factor of the graph G. □

5 Fractional triangle factors

In this section we prove Theorem 1.3. The constant 0.1 in this theorem is certainly not best possible, but we will make no serious attempt to optimize it.

The fact that a fractional triangle factor f can take non-integer values, as opposed to the characteristic vector of a "usual" (i.e. integer) triangle factor, enables us to invoke the powerful machinery of Linear Programming to prove the above result. Below we provide relevant necessary background, needed to apply Linear Programming Duality to the problem of showing the existence of a fractional factor.

Let $H=(V,E)$ be a hypergraph. Recall that a non-negative real-valued function $f:E\to\mathbb{R}_+$ is called a *fractional matching* with *value* $|f|=\sum_{e\in E}f(e)$ if $\sum_{v\in e}f(e)\le 1$ for every $v\in V$. The maximum of $|f|$ over all fractional matchings of H is the *fractional matching number* of H, denoted by $\nu^*(H)$. Similarly, a *fractional cover* of H is a non-negative real-valued function $g:V\to\mathbb{R}_+$ such that $\sum_{v\in e}g(v)\ge 1$ for every $e\in E(H)$. The *value* of g is $|g|=\sum_{v\in V}g(v)$. The minimum of $|g|$ over all fractional covers of H is the *fractional covering number* of H, denoted by $\tau^*(H)$.

It is easy to see that the above two definitions of $\nu^*(H)$ and $\tau^*(H)$ can be represented as optimal solutions of a pair of dual linear programming problems. The Duality Theorem of Linear Programming thus implies that $\tau^*(H) = \nu^*(H)$. It can also be used to get the following:

Proposition 5.1 *For every r-uniform hypergraph $H = (V, E)$ one has :*

1. *$\nu^*(H) \geq \nu(H)$, where $\nu(H)$ is a size of a maximum matching in H;*
2. *$\nu^*(H) \leq |V|/r$;*
3. *If $g : V \to \mathbb{R}_+$ is a fractional cover of H, then for every subset $U \subseteq V$ the function $g' : U \to \mathbb{R}_+$ defined by $g'(v) = g(v)$ for every $v \in U$ (that is, g' is the restriction of g on U) is a fractional cover of the induced hypergraph $H[U]$;*
4. *Let $g : V \to \mathbb{R}_+$ be an optimal fractional cover of H and denote $V_1 = \{v \in V : g(v) > 0\}$, then $\nu^*(H) \geq |V_1|/r$.*

The proof of the above proposition is quite standard and can be found, e.g., in [18] (see Proposition 2). The reader is also referred to [15] for additional information about integer and fractional matchings and covers in hypergraphs.

Now let us return to the problem of a fractional triangle factor and introduce relevant notation. Let $G = (V, E)$ satisfy the conditions of Theorem 1.3. We denote by $H = H(G)$ the hypergraph of the triangles of G defined as follows: the vertex set of H coincides with V; for every triangle (u, v, w) in G, there is an edge $(u, v, w) \in E(H)$. Thus H is a 3-uniform hypergraph on n vertices. For a subset $U \subseteq V$, we let $H[U]$ to be the subhypergraph of H spanned by U. In other words, $H[U]$ is formed by all the triangles of G spanned by U. With some abuse of notation we set $\nu^*(U) = \nu^*(H[U])$.

Notice that due to Proposition 5.1.2 we have $\nu^*(H) \leq n/3$. Therefore, the value of any fractional matching in H is at most $n/3$. Thus, proving that G has a fractional triangle factor is equivalent to proving that an optimal fractional matching in H has the maximal possible value, $n/3$.

The proof of Theorem 1.3 will proceed as follows. We first derive some properties of the graph G, based on the information about its eigenvalues. Then we show that every graph having these properties has a fractional triangle factor. Our approach is similar to the one used in [18] with some additional ideas. From now on we assume that the graph G satisfies the conditions of Theorem 1.3.

First we apply Theorem 2.4 to graph G with $r = 1$ and 2 to establish the following properties.

Fact 5.2 **(a)** *Every $0.1d$ vertices of G span an edge;* **(b)** *Every $0.2n$ vertices of G span a triangle.*

Proposition 5.3 *Let $v \in V(G)$. Then there exists a family $T(v)$ of $d/4$ triangles in G, so that for all $t \neq t' \in T(v)$, one has $t \cap t' = \{v\}$.*

Proof. Consider a maximum matching M_0 formed by the edges of G in the neighborhood $N(v)$. If M_0 has less than $d/4$ edges, then by Fact 5.2 (a), the vertices of $N(v)$ not covered by M_0 span an edge $e_0 \in E(G)$. Adding e_0 to M_0 produces a larger matching, a contradiction to the maximality of M_0. We conclude that M_0 has at least $d/4$ edges. Each edge of M_0 together with v forms a triangle in G. Choosing $d/4$ such triangles, we get a desired collection of triangles. $\square$

Proposition 5.4 *For every subset $U \subset V$ of size $|U| \geq n - d/2$ and for every subset $U_0 \subset U$ of size $|U_0| = d/4$, there exists a family T_0 of at least $0.2n - d/4$ triangles in G so that $t \subset U$ and $|t \cap U_0| = 1$ for every $t \in T_0$, and $t \cap t' \subset U_0$ for every pair $t \neq t' \in T_0$.*

Proof. We may assume that $0.2n \geq d/4$, otherwise there is nothing to prove. Let T_0 be a maximum family of triangles in G, satisfying the requirements of the proposition. We will show that T_0 contains at least $0.2n - d/4$ triangles.

Assume to the contrary that $|T_0| < 0.2n - d/4$. Let $W = \bigcup_{t \in T_0} t \setminus U_0$. Then $|W| < 0.4n - d/2$. Due to the maximality of T_0, the graph G does not contain a triangle with one vertex in U_0 and the other two in $U \setminus (U_0 \cup W)$. Notice that $|U \setminus (U_0 \cup W)| > n - d/2 - d/4 - (0.4n - d/2) = 0.6n - d/4$. Then it follows from Theorem 2.2 that

$$\begin{aligned} e(U_0, U \setminus (U_0 \cup W)) &\geq \frac{\frac{d}{4}\left(0.6n - \frac{d}{4}\right)d}{n} - \lambda\sqrt{\frac{d}{4}\left(0.6n - \frac{d}{4}\right)} \\ &> \frac{d^2}{4n}0.3n - \frac{0.1d^2}{n}\sqrt{\frac{dn}{4}} > 0.075d^2 - 0.05d^2 = 0.025d^2 \, . \end{aligned}$$

Therefore there is a vertex $u_0 \in U_0$ for which $e(\{u_0\}, U \setminus (U_0 \cup W)) > 0.025d^2/|U_0| = 0.1d$. Then by Fact 5.2 (a), the set of neighbors of u_0 in $U \setminus (U_0 \cup W)$ contains an edge $e = (w_1, w_2) \in E(G)$ – a contradiction. □

To complete the proof of Theorem 1.3 we show that every graph G having the properties stated in Fact 5.2 (b) and Propositions 5.3, 5.4 satisfies $\nu^*(H(G)) = n/3$.

Lemma 5.5 *If $G = (V, E)$ has the property stated in Fact 5.2 (b). Then for every subset $U \subseteq V$, one has: $\nu^*(U) \geq \frac{|U| - 0.2n}{3}$.*

Proof. Let T_0 be a maximum family of pairwise disjoint triangles of G spanned by U. Recall that by Fact 5.2 (b), every set of more than $0.2n$ vertices of U spans a triangle in G. Therefore the family T_0 leaves uncovered at most $0.2n$ vertices of U, implying $|T_0| \geq (|U| - 0.2n)/3$. Then by Proposition 5.1.1 we have $\nu^*(U) \geq (|U| - 0.2n)/3$. □

Lemma 5.6 *If $G = (V, E)$ has the properties stated in Fact 5.2 (b) and Proposition 5.4, then for every $U \subseteq V$ of size $|U| \geq n - d/2$, one has: $\nu^*(U) \geq \frac{|U| - \frac{d}{4}}{3}$.*

Proof. Consider the hypergraph of triangles $H[U]$. Let $g : U \to \mathbb{R}_+$ be an optimal fractional cover of $H[U]$. Denote

$$U' = \{u \in U : g(u) = 0\} \, .$$

Consider two possible cases.

Case 1: $|U'| \leq \frac{d}{4}$. In this case Proposition 5.1.4 implies $\nu^*(U) \geq \frac{|U \setminus U'|}{3} \geq \frac{|U| - \frac{d}{4}}{3}$.

Case 2: $|U'| \geq \frac{d}{4}$.

Choose a subset $U_0 \subset U'$ of size $|U_0| = d/4$. By Proposition 5.4, G contains a family T_0 of $0.2n - d/4$

triangles so that $|t \cap U_0| = 1$ for every $t \in T_0$ and $t \cap t' \subset U_0$ for every $t \neq t' \in T_0$. Let $W = \bigcup_{t \in T_0} t \setminus U_0$. Then $|W| = 2|T_0| = 0.4n - d/2$. Since g is a fractional cover of the triangles of G spanned by U, g has total weight at least 1 on every triangle from T_0. As all vertices from U_0 have zero weight in g, and the triangles in T_0 are disjoint outside U_0, we have: $\sum_{w \in W} g(w) \geq |T_0| = |W|/2$.

Define a function $g_1 : U \setminus W \to \mathbb{R}_+$ by setting $g_1(u) = g(u)$ for all $u \in U \setminus W$ (i.e. g_1 is the restriction of g on $U \setminus W$). Then by Proposition 5.1.3, g_1 is a fractional cover of the hypergraph $H[U \setminus W]$. Therefore from Lemma 5.5 it follows that $|g_1| \geq \nu^*(U \setminus W) \geq \frac{|U \setminus W| - 0.2n}{3}$. This completes the proof, since

$$\begin{aligned}
\nu^*(U) &= |g| = \sum_{w \in W} g(w) + |g_1| \geq \frac{|W|}{2} + \frac{|U \setminus W| - 0.2n}{3} \\
&= \frac{|U| - 0.2n + \frac{|W|}{2}}{3} = \frac{|U| - \frac{d}{4}}{3}. \qquad \square
\end{aligned}$$

Proof of Theorem 1.3. Consider the hypergraph $H = H(G)$ of triangles of G. Let $g : V \to \mathbb{R}_+$ be an optimal fractional cover of H of weight $|g| = \nu^*(H)$. We treat two cases again.
Case 1: $g(v) > 0$ for all $v \in V$. Then by Proposition 5.1.4 $\nu^*(H) \geq |V|/3$.
Case 2: There exists a vertex $v_0 \in V$ such that $g(v_0) = 0$.
According to Proposition 5.3, G contains a family T_0 of $d/4$ triangles whose pairwise intersection is v_0. Let $W = \bigcup_{t \in T_0} t \setminus \{v_0\}$. Then $|W| = 2|T_0| = d/2$. Also, $\sum_{v \in W} g(v) \geq |T_0| = |W|/2$. Define a function $g_1 : V \setminus W \to \mathbb{R}_+$ by setting $g_1(v) = g(v)$ for all $v \in V \setminus W$. Then by Proposition 5.1.3, g_1 is a fractional cover of the hypergraph $H[V \setminus W]$, and thus $|g_1| \geq \nu^*(V \setminus W)$. As $|V \setminus W| = n - d/2$, it follows from Lemma 5.6 that $\nu^*(V \setminus W) \geq (|V \setminus W| - d/4)/3$. Then

$$\begin{aligned}
\nu^*(H) &= |g| = \sum_{w \in W} g(w) + |g_1| \geq \frac{|W|}{2} + \frac{|V \setminus W| - \frac{d}{4}}{3} \\
&= \frac{|V| + \frac{|W|}{2} - \frac{d}{4}}{3} = \frac{|V|}{3}.
\end{aligned}$$

Thus in both cases we have proven $\nu^*(H) \geq |V|/3$. As the opposite inequality $\nu^*(H) \leq |V|/3$ is valid by Proposition 5.1.2, the theorem follows. $\square$

6 Pseudo-random graphs without triangles

The purpose of this section is to show the existence of a family of dense pseudo-random graphs without triangles and to prove Theorem 1.2.

In [1] Alon constructs triangle-free graphs with small independence number to devise a constructive lower bound for the Ramsey number $R(3, t)$. For an infinite sequence of graph sizes n, he defines a triangle-free d-regular graph G on n vertices with second eigenvalue $\lambda = \lambda(G)$, where $d = \Theta(n^{2/3})$ and $\lambda = \Theta(n^{1/3})$. Specifically, he proves the following result.

Theorem 6.1 *Let $k > 0$ be an integer not divisible by 3. Set $n = 2^{3k}$. Then there exists a graph G on n vertices with the following properties:*

1. *G is $d_n = 2^{k-1}(2^{k-1} - 1)$-regular;*
2. *$\lambda(G) \leq 9 \cdot 2^k + 3 \cdot 2^{k/2} + \frac{1}{4}$;*
3. *G is triangle-free.*

Now we use Alon's construction (Theorem 6.1) and apply the blow-up operation for various values of the blow-up parameter k. Notice that the graph G described in Theorem 6.1 satisfies $\lambda(G) = O(d^2/n)$, where d, n are the degree and the number of vertices of G, respectively. Observe also that according to Proposition 2.5, the k-blow-up of G preserves the above relation between the parameters n, d, λ. Finally, it is easy to see that the blow-up graph $G(k)$ is triangle-free as well. Therefore Theorem 1.2 follows. □

7 Concluding remarks

Our main result, Theorem 1.1 provides a sufficient condition for the existence of a triangle factor in a pseudo-random graph. Note that while our fractional result, Theorem 1.3 is asymptotically optimal due to Theorem 1.2, we cannot say the same about Theorem 1.1. In fact, we conjecture that the optimal condition in the integer case should be essentially identical to that of the fractional case:

Conjecture 7.1 *There exists an absolute constant $c > 0$ so that every d-regular graph G on $3n$ vertices, satisfying $\lambda(G) \leq cd^2/n$, has a triangle factor.*

Given graphs G, H on n, h vertices respectively, an H-factor in G is a subgraph of G consisting of n/h vertex disjoint copies of H. Thus a triangle factor corresponds to the case $H = K_3$. A fractional H-factor is defined in an obvious way. A natural possible extension of our result would be to find sufficient conditions for the existence of an H-factor in a pseudo-random graph G, for various choices of the fixed graph H. While the techniques of the present paper certainly have a potential to be adopted for other instances of H, we are unable at this stage to predict the optimal relation between parameters n, d, λ for most of the choices of H, for example, for the case $H = K_4$. This is due to the absence of known optimal constructions of pseudo-random graphs without a single copy of H (see [3] for partial results in the case of H being a fixed clique). We conjecture however that conditions on n, d, λ, sufficient to guarantee a copy of H in a pseudo-random graph G, will be already sufficient to guarantee the existence of a fractional or even an integer H-factor. Finally it is worth stating here that using the technique from Section 5 we can show the following statement:

Theorem 7.2 *For every $r \geq 3$ there exists a constant $c = c(r) > 0$ such that every d-regular graph G on n vertices, satisfying $\lambda(G) \leq cd^{r-1}/n^{r-2}$, has a fractional K_r factor.*

It is instructive to observe that the appearance of triangles in pseudo-random graphs experiences a kind of a threshold behavior – if $\lambda = cd^2/n$ for large enough constant $c > 0$, then there are graphs with parameters (n, d, λ) and without a single triangle; changing the value of c guarantees not only the existence of a triangle, but already the existence of a fractional triangle factor. It would be interesting to understand this phenomenon more deeply, possibly comparing it with the corresponding phenomenon in the theory of random graphs (see, e.g. [7]), and to figure out the scope of its applicability.

In our opinion, an ultimate goal of the line of research suggested by this paper is to prove a generalization of the Blow-Up Lemma to pseudo-random subgraphs with a subquadratic number of edges. Specifically, the following question appears to be both very interesting and challenging: given an integer Δ, what are the optimal conditions on the relation between parameters n, d, λ that guarantee that every d-regular graph G on n vertices with a second eigenvalue λ contains a copy of every n-vertex graph H with maximum degree at most Δ? Our main result, Theorem 1.1, can be viewed as the first step in settling this problem.

Acknowledgement We thank the referees, whose thorough comments improved significantly the presentation of this paper.

References

[1] N. Alon, Explicit Ramsey graphs and orthonormal labelings, *Electronic Journal of Combinatorics* **1** (1994), R12.

[2] N. Alon and F. R. K. Chung, Explicit construction of linear sized tolerant networks, *Discrete Math.* **72** (1988), 15–19.

[3] N. Alon and M. Krivelevich, Constructive bounds for a Ramsey-type problem, *Graphs and Combinatorics* **13** (1997), 217–225.

[4] N. Alon, M. Krivelevich and B. Sudakov, List coloring of random and pseudo-random graphs, *Combinatorica* **19** (1999), 453–472.

[5] N. Alon, V. Rödl and A. Ruciński, Perfect matchings in ϵ-regular graphs, *Electronic J. Combinatorics* **5** (1998), R13.

[6] N. Alon and J. H. Spencer, **The probabilistic method**, 2^{nd} ed., Wiley, New York, 2000.

[7] B. Bollobás and A. Thomason, Threshold functions, *Combinatorica* **7** (1987), 35–38.

[8] L.M. Brégman, Some properties of nonnegative matrices and their permanents, *Soviet Math. Dokl.* **14** (1973), 945-949 [*Dokl. Akad. Nauk SSSR*, **211** (1973), 27-30].

[9] F. R. K. Chung, **Spectral graph theory**, Regional Conference Series in Mathematics 92, Amer. Math. Soc., Providence, 1997.

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

[11] K. Corrádi and A. Hajnal, On the maximal number of independent circuits in a graph, *Acta Math. Acad. Sci. Hungar.* **14** (1963), 423–439.

[12] D. Cvetković, M. Doob and H. Sachs, **Spectra of graphs. Theory and applications**, 3^{rd} ed., Johann Ambrosius Barth, Heidelberg, 1995.

[13] G. Egorychev, Proof of the van der Waerden conjecture for permanents, *Sibirsk. Mat. Zh.* **22** (1981), 65–71.

[14] D.I. Falikman, A proof of van der Waerden's conjecture on the permanent of a doubly stochastic matrix, *Mat. Zametki* **29** (1981), 931-938.

[15] Z. Füredi, Matchings and covers in hypergraphs, *Graphs and Combinatorics* **4** (1988), 115–206.

[16] A. Hajnal and E. Szemerédi, Proof of a conjecture of Erdős, in: *Combinatorial Theory and its applications*, Vol. II, P. Erdős, A. Rényi and V. T. Sós, Eds., Colloq. Math. Soc. J. Bolyai 4, North Holland, Amsterdam, 1970, pp. 601–623.

[17] J. Komlós, G. N. Sárközy and E. Szemerédi, The Blow-up Lemma, *Combinatorica* **17** (1997), 109-123.

[18] M. Krivelevich, Perfect fractional matchings in random hypergraphs, *Random Structures and Algorithms* **9** (1996), 317–334.

[19] M. Krivelevich and B. Sudakov, Pseudo-random graphs, to appear.

[20] L. Lovász, M.D. Plummer, **Matching Theory**, Akadémia Kiadó, Budapest, 1986.

[21] V. Rödl and A. Ruciński, Perfect matchings in ϵ-regular graphs and the Blow-Up Lemma, *Combinatorica*, **19** (1999), 437-452.

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

[23] A. Thomason, Pseudorandom graphs, In: *Random graphs'85 (Poznań 85)*, North-Holland Math. Stud. 144, North Holland, Amsterdam, 1987, pp. 307–331.

[24] A. Thomason, Random graphs, strongly regular graphs and pseudorandom graphs, In: *Surveys in Combinatorics 1987*, London Math. Soc. Lecture Note Ser. 123, Cambridge Univ. Press, Cambridge, 1987, pp. 173–195.

8 Appendix: Proof of Lemma 3.1

Obviously, it is enough to check condition (1) for $x+y \geq n$. Due to symmetry, we may assume $x \leq y$.

Case 1. $x \leq k$.

Every $v \in X$ has at least $deg_{H'}(v) - t - (n-y) \geq d - r - t + y - n$ neighbors in Y. Hence

$$\begin{aligned} e_H(X,Y) &\geq x(d-t-r+y-n) \geq x(k+\lambda+y-n) \\ &= kx + xy - nx + \lambda x = kx + ky - nx - (k-x)y + \lambda x \\ &\geq kx + ky - nx - (k-x)n + \lambda x \qquad \text{(we use here } x \leq k,\ y \leq n) \\ &= kx + ky - kn + \lambda x \geq k(x+y-n)\ , \end{aligned}$$

as required.

Case 2. $x \geq k$. From the assumptions of the lemma we obtain

$$\begin{aligned} e_H(X,Y) &\geq e_H(X,V) - e_{H'}(X, V-Y) \\ &\geq (d-t-r)x - \frac{x(n-y)d}{n} - \lambda\sqrt{x(n-y)} \\ &= \frac{xyd}{n} - (t+r)x - \lambda\sqrt{x(n-y)}\ . \end{aligned}$$

Therefore it is enough to show

$$xyd - (t+r)nx - \lambda n\sqrt{x(n-y)} \geq n(d-\lambda-2t-2r)(x+y-n)\,,$$

or

$$d(n-x)(n-y) \geq \lambda n\Big(\sqrt{x(n-y)} - (x+y-n)\Big) + (t+r)n\Big(x - 2(x+y-n)\Big)\ . \tag{3}$$

As both $x, y \leq n$, the LHS of (3) is non-negative. First we present two cases in which the RHS is negative.

Consider first the case when $y \geq n - k/3$. Then, recalling that $x \geq k$, we get

$$\sqrt{x(n-y)} - (x+y-n) \leq \sqrt{xk/3} - x + \frac{k}{3} \leq \frac{x}{\sqrt{3}} - x + \frac{x}{3} < 0\ .$$

Also in this case,

$$x - 2(x+y-n) \leq x - 2\Big(x - \frac{k}{3}\Big) \leq -\frac{x}{3} < 0\ .$$

These imply that the RHS of (3) is negative and establish the validity of (3) when $y \geq n - k/3$.

Now assume that $x \geq 3n/4$. In this case we have:

$$\begin{aligned} \sqrt{x(n-y)} - (x+y-n) &\leq \sqrt{x(n-x)} - 2x + n \leq \sqrt{\frac{3n}{4}\cdot\frac{n}{4}} - 2\cdot\frac{3n}{4} + n \\ &= n\left(\frac{\sqrt{3}}{4} - \frac{3}{2} + 1\right) < 0\ . \end{aligned}$$

Also,

$$x - 2(x+y-n) \leq n - 2\Big(\frac{3n}{4} + \frac{3n}{4} - n\Big) = 0\ .$$

Therefore the RHS of (3) is negative again, implying the validity of (3).

Finally we dispose of the case $y < n - k/3$, $x < 3n/4$. Notice that $k = d - 2(r+t) - \lambda \geq d - 2d^2/(40n) - d^2/(20n) > d/2$. Therefore, for the LHS of (3) we have:

$$d(n-y)(n-x) \geq \frac{dk}{3}(n-x) \geq \frac{d^2}{6}(n-x) \ .$$

The RHS of (3) can be estimated from above by:

$$\begin{aligned}
\lambda n\Big(\sqrt{x(n-y)} - (x+y-n)\Big) + (t+r)n\Big(x - 2(x+y-n)\Big) &\leq \lambda n\sqrt{x(n-y)} + (t+r)nx \\
&\leq \frac{d^2}{20}\sqrt{x(n-x)} + (t+r)nx \\
&\leq \frac{d^2}{20}\sqrt{3(n-x)^2} + (t+r)3n(n-x) \\
&\leq \frac{\sqrt{3}}{20}(n-x)d^2 + \frac{3}{40}(n-x)d^2 \\
&< \frac{(n-x)d^2}{6} \ .
\end{aligned}$$

This proves (3) in this case and thus completes the proof of the lemma. $\square$