Near-perfect clique-factors in sparse pseudorandom graphs [1]

Jie Han [a,2,5] Yoshiharu Kohayakawa [a,3,6] Yury Person [b,4,7]

[a] *Instituto de Matemática e Estatística, Universidade de São Paulo, Brazil*

[b] *Institut für Mathematik, Technische Universität Ilmenau, 98684 Ilmenau, Germany*

Abstract

We prove that, for any $t \geq 3$, there exists a constant $c = c(t) > 0$ such that any d-regular n-vertex graph with the second largest eigenvalue in absolute value λ satisfying $\lambda \leq cd^{t-1}/n^{t-2}$ contains $(1 - o(1))n/t$ vertex-disjoint copies of K_t. This provides further support for the conjecture of Krivelevich, Sudakov and Szábo [*Triangle factors in sparse pseudo-random graphs*, Combinatorica **24** (2004), pp. 403–426] that (n, d, λ)-graphs with $n \in 3\mathbb{N}$ and $\lambda \leq cd^2/n$ for a suitably small absolute constant $c > 0$ contain triangle-factors.

Keywords: clique-factors, (n, d, λ)-graphs, pseudorandomness

---

[1] The cooperation of the authors is supported by a joint CAPES-DAAD PROBRAL project (Proj. no. 430/15, 57350402, 57391197)

[2] Email: `jhan@ime.usp.br`

[3] Email: `yoshi@ime.usp.br`

[4] Email: `yury.person@tu-ilmenau.de`

[5] Supported by FAPESP (2014/18641-5, 2013/03447-6)

[6] Partially supported by CNPq (310974/2013-5)

[7] Supported by DFG grant PE 2299/1-1

1 Introduction

The study of conditions under which certain spanning or almost spanning structures are forced in random or pseudorandom graphs is one of the central topics in extremal graph theory and in random graphs.

An (n, d, λ)-graph is an n-vertex d-regular graph whose second largest eigenvalue in absolute value is at most λ. Graphs with $\lambda \ll d$ are considered to be *pseudorandom*, i.e., they behave in certain respects as random graphs do; for example, the edge count between 'not too small' vertex subsets is close to what one sees in random graphs of the same density. As usual, let $e(A, B) = e_G(A, B)$ denote the number of pairs $(a, b) \in A \times B$ so that ab is an edge of G (note that edges in $A \cap B$ are counted twice). The following result makes what we discussed above precise.

Theorem 1.1 (Expander mixing lemma, see e.g. [6]) *If G is an (n, d, λ)-graph and A, $B \subseteq V(G)$, then*

$$\left| e(A, B) - \frac{d}{n}|A||B| \right| < \lambda\sqrt{|A||B|}. \tag{1}$$

As starting points to the extensive literature on pseudorandom graphs, the reader is refereed to, e.g., [6, Chapter 9] and [11].

It is an interesting problem to understand optimal or asymptotically optimal conditions on the parameter λ in terms of d and n that force an (n, d, λ)-graph to possess a desired property. To demonstrate the optimality of a condition, one needs to show the existence of an (n, d, λ)-graph that certifies that the condition is indeed optimal.

Unfortunately, there are very few examples certifying optimality. A celebrated example is due to Alon, who showed [3] that there are (n, d, λ)-graphs that are K_3-free and yet satisfy $\lambda = cd^2/n$ for some absolute constant $c > 0$. This is in contrast with the fact that, as it follows easily from the expander mixing lemma above, for, say, $\lambda \leq 0.1d^2/n$, any (n, d, λ)-graph contains a triangle (in fact, every vertex lies in a triangle). It turns out that (n, d, λ)-graphs with $\lambda = \Theta(d^2/n)$ must satisfy $d = \Omega(n^{2/3})$. The construction of Alon [3] provides an example of the essentially sparsest possible K_3-free (n, d, λ)-graph with $d = \Theta(n^{2/3})$ and $\lambda = \Theta(n^{1/3})$. The other known example is a generalization of this construction by Alon and Kahale [5] (see also [11, Section 3]) to graphs without odd cycles of length at most $2\ell + 1$.

Our focus here is on spanning or almost spanning structures in (n, d, λ)-graphs. One of the simplest spanning structures is that of a perfect matching. Alon, Krivelevich and Sudakov (see [11]) proved that (n, d, λ)-graphs with

$\lambda \le d-2$ and n even contain perfect matchings. *Factors* generalize perfect matchings: for a graph F, an *F-factor* in a graph G is a collection of vertex-disjoint copies of F in G whose vertex sets cover $V(G)$ (this requires that $v(G) := |V(G)|$ should be divisible by $v(F)$). Motivated by the study of spanning structures in graphs, Krivelevich, Sudakov and Szabó [12] proved that (n,d,λ)-graphs with $\lambda = o\left(d^3/(n^2 \log n)\right)$ contain a triangle-factor if $3 \mid n$.

A *fractional* triangle-factor in a graph $G=(V,E)$ is a non-negative weight function f on the set $\mathcal{K}_3(G)$ of all triangles T of G, such that, for every $v \in V$, we have $\sum_{T:v\in V(T)} f(T) = 1$. Krivelevich, Sudakov and Szabó further proved [12] that (n,d,λ)-graphs with $\lambda \le 0.1d^2/n$ admit a fractional triangle-factor. Moreover, they conjectured the following.

Conjecture 1.2 (Conjecture 7.1 in [12]) *There exists an absolute constant $c>0$ such that if $\lambda \le cd^2/n$, then every (n,d,λ)-graph G on $n \in 3\mathbb{N}$ vertices has a triangle-factor.*

The t^{th} *power* H^t of a graph H is the graph on the vertex set $V(H)$ where uv ($u \ne v$) is an edge if there is a u-v-path of length at most t in H. Since the $(t-1)^{\text{st}}$ power of a Hamilton cycle contains a K_t-factor if $t \mid n$, powers of Hamilton cycles are also of interest when investigating clique-factors.

Allen, Böttcher, Hàn and two of the authors [2] proved that, if $\lambda = o(d^{3t/2}n^{1-3t/2})$ and $t \ge 3$, then any (n,d,λ)-graph contains the t^{th} power of the Hamilton cycle (and thus a K_{t+1}-factor if $(t+1) \mid n$).

In the case $t=2$, it was further proved in [2] that the condition $\lambda = o\left(d^{5/2}/n^{3/2}\right)$ suffices to guarantee squares of Hamilton cycles, and thus K_3-factors, improving over the aforementioned result of Krivelevich, Sudakov and Szabó. Very recently, it was proved in [8] that $\lambda = o(d^t/n^{t-1})$ guarantees a K_t-factor, which further improved [2] for $t \ge 4$.

The construction of Alon of K_3-free (n,d,λ)-graphs shows that the condition on λ in Conjecture 1.2 cannot be weakened. The result from [12] on the existence of fractional triangle-factors supports Conjecture 1.2. As a further evidence in support of that conjecture we prove here the following result.

Theorem 1.3 *For any $t \ge 3$, there is $n_0 > 0$ such that the following holds. Every (n,d,λ)-graph G with $n \ge n_0$ and $\lambda \le (1/(50t4^{t-2}))d^{t-1}/n^{t-2}$ contains vertex-disjoint copies of K_t covering all but at most $n^{1-1/(8t^4)}$ vertices of G.*

We remark that, under the condition $\lambda \le cd^{t-1}/n^{t-2}$ for some appropriate $c = c(t) > 0$, Krivelevich, Sudakov and Szabó [12] proved that any (n,d,λ)-graph contains a fractional K_t-factor.

A naïve approach to proving Theorem 1.3 is to pick cliques K_t one after another, each vertex-disjoint from the previous ones, by appealing to the pseudorandomness of G via the expander mixing lemma. However, even for triangles, if $\lambda = cd^2/n$, then all what one gets this way is that G has $(1-c)n/3$ vertex-disjoint triangles: one can see that a set of cn vertices in G induces a graph of average degree roughly cd, but the condition on λ and the expander mixing lemma do not guarantee that sets of size roughly cd contain an edge, and hence we do not know whether cn vertices necessarily span a triangle. Thus our naïve greedy approach will get stuck leaving cn vertices uncovered. What our result establishes is that, even for some absolute constant $c > 0$, we can cover all but $o(n)$ vertices of G by vertex-disjoint copies of K_3. Moreover, $o(n)$ can be taken to be of the form $n^{1-\varepsilon}$ for some $\varepsilon > 0$. We have restricted ourselves to triangles in this paragraph, but a similar reasoning applies to general cliques K_t as well.

Now let $p = d/n$ and suppose $G = (V, E)$ is an (n, d, λ)-graph with $\lambda \leq cd^2/n$. Inequality (1) implies that

$$\left| \frac{e(A,B)}{|A||B|} - p \right| < \frac{cp^2 n}{\sqrt{|A||B|}} \leq c^{1/2} \tag{2}$$

for all A, $B \subseteq V$ with $|A|$, $|B| \geq c^{1/2}n$. Let us now focus on the case in which d is linear in n, that is, $p = d/n$ is a constant independent of n. The powerful blow-up lemma of Komlós, Sárközy and Szemerédi [9] implies that, if c is small enough in comparison with p and $1/t$, then *any* graph $G = (V, E)$ on n vertices with minimum degree at least pn that satisfies (2) contains a K_t-factor as long as $t \mid n$. Thus, Conjecture 1.2 holds for dense graphs.

We remark that the blow-up lemmas for sparse graphs developed recently by Allen, Böttcher, Hàn and two of the authors [1] provide bounds on λ to establish the existence of K_t-factors, but those bounds are worse than those from [2] discussed above.

2 A proof outline

In the following we provide a proof overview in the case of triangles, since the general case is similar. Our arguments combine tools from linear programming with probabilistic techniques. In fact, they can be seen as a synthesis of some methods in Alon, Frankl, Huang, Rödl, Ruciński and Sudakov [4] and in Krivelevich, Sudakov and Szabó [12].

Let an (n, d, λ)-graph G with $\lambda \leq cd^2/n$ be given. From the expander mixing lemma, Theorem 1.1, it follows that every vertex of G lies in $\frac{1}{2}\left(d^3/n \pm \lambda d\right) = \left(d^3/2n\right)(1 \pm c)$ triangles. The naïve greedy approach above does not guarantee a collection of $(1-o(1))n/3$ vertex-disjoint triangles. Another attempt would be to apply some theorem that would tell us that the 3-uniform hypergraph $\mathcal{K}_3(G)$ of the triangles in G contains an almost perfect matching. A theorem of Pippenger (see [7]) would do if we knew that $\mathcal{K}_3(G)$ is pseudorandom enough (roughly speaking, one needs that $\mathcal{K}_3(G)$ should be approximately ℓ-regular for some $\ell \to \infty$ and that pairs of vertices of $\mathcal{K}_3(G)$ should be contained in $o(\ell)$ triples of $\mathcal{K}_3(G)$ (i.e., the 'codegrees' should be small)). However, for c an absolute constant, this property of $\mathcal{K}_3(G)$ cannot be deduced.

We circumvent the fact that $\mathcal{K}_3(G)$ is not necessarily pseudorandom enough by finding a subhypergraph H of $\mathcal{K}_3(G)$ in which the 'deviation' of the number of triangles at any vertex is 'smoothed out' (thus H will be almost ℓ-regular). This can be done if G has $\ell = n^{\Theta(1)}$ fractional K_3-factors $f_1, \dots, f_\ell$ such that $\sum_{i=1}^{\ell} f_i(T) \le 1$ for each $T \in \mathcal{K}_3(G)$ and, for any edge $e \in E(G)$, the sum of the weights on the triangles containing e across $f_1, \dots, f_\ell$ is at most $\ell^{1-\gamma}$ for some $\gamma \in (0,1)$. This latter condition helps us force small codegrees.

Indeed, with these fractional K_3-factors, we can select $H \subseteq \mathcal{K}_3(G)$ at random, by including each $T \in \mathcal{K}_3(G)$ in H independently with probability $\sum_{i=1}^{\ell} f_i(T)$. Then Chernoff's inequality guarantees that H satisfies, with high probability, the assumptions of a packing result in [10], a strengthening of Pippenger's result. Such a 'randomization' strategy has previously been successfully employed in [4] in the context of perfect matchings in hypergraphs.

Thus, it suffices to find such K_3-factors $f_1, \dots, f_\ell$. In fact, we find such f_i with the property that, for any $e \in E(G)$, we have $\sum_{e \in E(T)} \sum_{i=1}^{\ell} f_i(T) \le 1$ (hence $\sum_{i=1}^{\ell} f_i(T) \le 1$ for each $T \in \mathcal{K}_3(G)$ is automatically true).

Theorem 1.3 is vacuously true for $d = o(n^{2/3})$ when $t = 3$. We thus suppose $d = \Omega(n^{2/3})$. We consider two cases. We pick any $\beta \in (0, 1/3)$ independent of n. Our first approach works as long as d is not too small, say, $d \ge n^{2/3+\beta}$. In contrast, the second approach works as long as d is not too large, say, $d \le n^{1-\beta}$.

In the first approach, we consider edge-weighted graphs and we repeatedly 'remove' fractional K_3-factors from G (removing from edges e the weights of the triangles T with $e \subseteq V(T)$). We show that we can repeat this process n^β times. To establish this, we use linear programming techniques to find fractional triangle-factors in weighted graphs. In doing so, we generalize the linear programming arguments from [12] used to study fractional triangle-factors in (n, d, λ)-graphs.

When d is close to $n^{2/3}$, our approach above fails because we cannot execute it sufficiently many times. To circumvent this, we randomly split $E(G)$

into $\ell = n^{\Omega(1)}$ sets $E_1, \ldots, E_\ell$, with each subgraph $G_i := (V, E_i)$ distributed as a random subgraph G_p of G, where each edge is included in G_p with probability $p = 1/\ell$, independently of all the other edges. Then we find in each G_i a fractional K_3-factor f_i with high probability, again by linear programming arguments. This second approach works only for $d \leq n^{1-o(1)}$, which makes both approaches necessary.

References

[1] Allen, P., J. Böttcher, H. Hàn, Y. Kohayakawa and Y. Person, *Blow-up lemmas for sparse graphs*, arXiv:1612.00622, submitted.

[2] Allen, P., J. Böttcher, H. Hàn, Y. Kohayakawa and Y. Person, *Powers of Hamilton cycles in pseudorandom graphs*, Combinatorica **37** (2017), pp. 573–616.

[3] Alon, N., *Explicit Ramsey graphs and orthonormal labelings*, Electron. J. Combin. **1** (1994), pp. Research Paper 12, approx. 8.

[4] Alon, N., P. Frankl, H. Huang, V. Rödl, A. Ruciński and B. Sudakov, *Large matchings in uniform hypergraphs and the conjectures of Erdős and Samuels*, J. Combin. Theory Ser. A **119** (2012), pp. 1200–1215.

[5] Alon, N. and N. Kahale, *Approximating the independence number via the ϑ-function*, Mathematical Programming **80** (1998), pp. 253–264.

[6] Alon, N. and J. H. Spencer, "The probabilistic method," Hoboken, NJ: John Wiley & Sons, 2016, 4th edition edition, 384 pp.

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

[8] Han, J., Y. Kohayakawa, P. Morris and Y. Person, *Clique-factors in sparse pseudorandom graphs*, manuscript.

[9] Komlós, J., G. N. Sárközy and E. Szemerédi, *Blow-up lemma*, Combinatorica **17** (1997), pp. 109–123.

[10] Kostochka, A. V. and V. Rödl, *Partial Steiner systems and matchings in hypergraphs*, Random Structures Algorithms **13** (1998), pp. 335–347.

[11] Krivelevich, M. and B. Sudakov, *Pseudo-random graphs*, in: *More sets, graphs and numbers*, Bolyai Soc. Math. Stud. **15**, Springer, Berlin, 2006 pp. 199–262.

[12] Krivelevich, M., B. Sudakov and T. Szabó, *Triangle factors in sparse pseudo-random graphs*, Combinatorica **24** (2004), pp. 403–426.