From small eigenvalues to large cuts, and Chowla's cosine problem

Zhihan Jin* Aleksa Milojević* István Tomon† Shengtong Zhang‡

Abstract

We show that there exists an absolute constant $\gamma > 0$ such that for every $A \subseteq \mathbb{Z}_{>0}$ we have

$$\min_{x\in[0,2\pi]} \sum_{a\in A} \cos(ax) \le -\Omega(|A|^\gamma).$$

This gives the first polynomial bound for Chowla's cosine problem from 1965. To show this, we prove structural statements about graphs whose smallest eigenvalue is small in absolute value. As another application, we show that any graph G with m edges and no clique of size $m^{1/2-\delta}$ has a cut of size at least $m/2 + m^{1/2+\varepsilon}$ for some $\varepsilon = \varepsilon(\delta) > 0$. This proves a weak version of a celebrated conjecture of Alon, Bollobás, Krivelevich, and Sudakov. Our proofs are based on novel spectral and linear algebraic techniques, involving subspace compressions and Hadamard products of matrices.

1 Introduction

The central question in spectral graph theory is how the structural properties of a graph influence its spectrum and, conversely, what information about the graph the spectrum encodes. Here and throughout, the spectrum of the graph refers to the set of eigenvalues of its adjacency matrix. Some of the most prominent relations between eigenvalues and structural properties of a graph are the Expander Mixing Lemma [4] (which connects pseudorandomness properties of the graph to its eigenvalues) and the Hoffman bound [43] (which bounds the size of the largest independent set based on the smallest eigenvalue). The question which concerns us in this paper is the following: what can be said about the graphs whose last eigenvalue is small in absolute value?

The prime examples of such graphs are cliques, since all of their eigenvalues are at least -1. As taking disjoint unions of graphs corresponds to taking a union of their spectra, and changing a small number of edges does not affect the spectrum significantly, it is not hard to see that graphs which are close to disjoint unions of cliques also have small $|\lambda_n|$, where λ_n denotes the smallest eigenvalue of the graph. The main technical result of our paper is a converse to this statement: we prove that already a mild restriction $|\lambda_n| \le n^{1/4-o(1)}$ on the smallest eigenvalue λ_n of an n-vertex graph forces the graph to be close to a disjoint union of cliques, and the exponent $1/4$ is best possible.

We also obtain a variant of this result which applies to sparse graph. More precisely, we show that graphs of average degree d and smallest eigenvalue $|\lambda_n| \le d^\gamma$ contain cliques of size $d^{1-O(\gamma)}$. Perhaps

*Department of Mathematics, ETH Zürich, Switzerland. Email: `{zhihan.jin, aleksa.milojevic}@math.ethz.ch`.
Research supported in part by SNSF grant 200021-228014.

†Umeå University, *e-mail*: `istvantomon@gmail.com`, Research supported in part by the Swedish Research Council grant VR 2023-03375.

‡Stanford University, *e-mail*: `stzh1555@stanford.edu`. This work was partially supported by the National Science Foundation under Grant No. DMS-1928930, while the author was in residence at the Simons Laufer Mathematical Sciences Institute in Berkeley, California, during the Spring 2025 semester. This work was also partially supported by NSF Award DMS-2154129.

surprisingly, this simple graph-theoretic statement has a powerful application to Chowla's cosine problem [21] from 1965. In fact, it allows us to prove that if A is a finite set of positive integers, then the function

$$f(x) = \sum_{a \in A} \cos(ax)$$

takes values as small as $-|A|^{\Omega(1)}$. For a more detailed overview of Chowla's cosine problem and related work, consult Section 1.1.

Our methods can also be used to study graphs with small maximum cut, allowing us to make substantial progress on a celebrated conjecture of Alon, Bollobás, Krivelevich, and Sudakov [3]. This is a central problem in the study of the maximum cut in H-free graphs, initiated by Erdős and Lovász [35] in the 1970's. If H is a fixed graph, this conjecture states that any graph G with m edges and no subgraph isomorphic to H has a cut of size at least $m/2 + m^{1/2+\varepsilon_H}$ for some $\varepsilon_H > 1/4$. Here, a *cut* is a partition of a vertex set of G into two parts, and its size is the number of edges crossing the partition. We prove that forbidding a clique of size as large as $m^{1/2-\delta}$ in G already allows to show that G has a cut with at least $m/2 + m^{1/2+\varepsilon}$ edges, for some $\varepsilon = \varepsilon(\delta) > 0$. The problem of estimating the size of the maximum cut is discussed in further detail in Section 1.2.

Finally, our results have further implications about graphs of small second eigenvalue. A classical result of Alon and Boppana [63] gives a bound on the second-largest eigenvalue λ_2 of d-regular n-vertex graphs, and it states that $\lambda_2 \geq 2\sqrt{d-1}(1 - \frac{1}{\lfloor D/2 \rfloor})$, where D is the diameter of the graph. Note that this bound is trivial for graphs of diameter at most 3, which could happen already when $d \geq n^{1/3}$. In this paper, we extend Alon–Boppana theorem to dense graphs, and we show that if an n-vertex regular graph is far from a Turán graph, then its second eigenvalue satisfies $\lambda_2 \geq n^{1/4-o(1)}$, with the exponent $1/4$ being optimal. Finally, we prove similar structural results about graphs of large bisection width. Our proofs introduce novel spectral and linear-algebraic techniques based on subspace compressions of matrices and the use of Hadamard products, which may be of independent interest.

1.1 Chowla's cosine problem

In 1948, in the study of certain Dedekind zeta functions, Ankeny and Chowla came across the the following question, see [20] — is it true that for every $K > 0$ and sufficiently large $n > 0$, if $a_1, \ldots, a_n$ are distinct integers, then the minimum of the function $f(x) = \cos(a_1 x) + \cdots + \cos(a_n x)$ is less than $-K$? This was proved by Uchiyama and Uchiyama [74], albeit with poor quantitative dependencies. Soon after this work, in 1965, Chowla [21] revisited the problem and made a more precise conjecture, today known as Chowla's cosine problem. He asked to show that $\min_{x \in [0,2\pi]} f(x) = \min_{x \in [0,2\pi]} \sum_{i=1}^{n} \cos(a_i x) \leq -\Omega(\sqrt{n})$. The bound $-\Omega(\sqrt{n})$ comes from the fact that in case $A = \{a_1, \ldots, a_n\}$ can be written as $A = B - B$, where B is a Sidon set, one has $\min_{x \in [0,2\pi]} f(x) = -\Theta(\sqrt{n})$, see [62] for a detailed proof. The bounds of Uchiyama and Uchiyama were later improved by Roth [67], who showed a lower bound of $\min_x f(x) \leq -\Omega(\sqrt{\log n / \log \log n})$.

An important observation in the early study of Chowla's cosine problem was its connection to Littlewood's L_1-problem, which asks to show that for each n-element set $A \subseteq \mathbb{Z}$, the L_1-norm of the Fourier transform of $\mathbb{1}_A$ is bounded below by $\log n$, i.e.

$$\|\widehat{\mathbb{1}_A}\|_1 = \int_0^1 \Big| \sum_{a \in A} e^{2\pi i a x} \Big| \, d\, x = \Omega(\log n).$$

Indeed, any lower bound on Littlewood's L_1-problem gives a comparable upper bound for the cosine problem (see [67] for a detailed derivation). Thus, the resolution of the Littlewood L_1-problem by Konyagin [53] and McGehee, Pigno and Smith [61] immediately implies that $\min_x f(x) \leq -\Omega(\log n)$.

It was Bourgain [14] who first broke this logarithmic barrier, and then his method was further refined by Ruzsa [68] to give the previously best known bound $\min_x f(x) = -\exp\left(\Omega(\sqrt{\log n})\right)$. Chowla's cosine problem also appears on Green's 100 problems list [41] as problem number 81. Here, we give the first polynomial bound.

Theorem 1.1. *There exists an absolute constant $\gamma > 0$ such that for any A of positive integers, there exists $x \in [0, 2\pi]$ such that*

$$\sum_{a \in A} \cos(ax) \leq -\Omega(|A|^\gamma).$$

Following our proof, one can take $\gamma = 0.01$, so it remains an interesting open problem to decide whether $\gamma = 1/2$ is the best exponent. Our graph theoretic machinery has a hard theoretical barrier at the exponent $1/4$, and possibly substantial new ideas are needed to move beyond this point (however, we do not claim that $\gamma = 1/4$ is reachable either with our methods). We now say a couple of words about the proof of Theorem 1.1. The key ingredient of the proof is the following purely graph-theoretic result.

Theorem 1.2. *For every $\gamma > 0$, the following holds for every sufficiently large d. Let G be a graph of average degree d and assume that $|\lambda_n| \leq d^\gamma$. Then G contains a clique of size at least $d^{1-O(\gamma)}$.*

The first idea is to embed A into the group $\mathbb{Z}/n\mathbb{Z}$ for a sufficiently large prime n and to consider the Cayley graph $\Gamma = \mathrm{Cay}(\mathbb{Z}/n\mathbb{Z}, A \cup -A)$. It is well-known that the eigenvalues of Cayley graphs correspond to the Fourier coefficients of the generating set, and thus the smallest eigenvalue λ_n satisfies $\lambda_n = \sum_{a \in A \cup -A} \exp(2\pi i a k/n) = 2\sum_{a \in A} \cos(2\pi a k/n)$ for some $k \in [n]$. Hence, $\frac{1}{2}\lambda_n \geq \min_x f(x)$. In other words, unless the minimum of $f(x)$ is very negative, $|\lambda_n|$ is small. But then we can use Theorem 1.2 to conclude that G contains very large cliques. However, the transitive symmetry of the Cayley graph does not allow such large cliques, without violating the property of having small $|\lambda_n|$.

We conclude the discussion of Chowla's cosine problem by remarking some other interesting questions about cosine polynomials. For example, the problem of estimating the number of zeros of such polynomials have also attracted lots of attention recently. Namely, an old problem of Littlewood [59] asks study the minimum number of zeros of the function $f(x) = \sum_{i=1}^n \cos(a_i x)$ in the interval $[0, 2\pi]$. Although Littlewood guessed that this number should always be linear in n, Borwein, Erdélyi, Ferguson and Lockhart [15] showed that the there are integers $a_1, \dots, a_n$ such that $f(x)$ has at most $n^{5/6}$ zeros. This result was later improved to $(n \log n)^{2/3}$ by Juškevičius and Sahasrabudhe [49] and, independently, by Konyagin [54]. A complementary bound has been proven by Sahasrabudhe [69] and Erdélyi [33, 34], who showed that $f(x)$ always has at least $(\log\log\log n)^{1/2-\varepsilon}$ roots, and this was later improved to $(\log\log n)^{1-o(1)}$ by Bedert [13].

In the Appendix, we discuss further extensions of Chowla's problem in arbitrary finite groups. The proof of Theorem 1.1 is presented in Section 3, and the proof of Theorem 1.2 is presented in Section 12.

1.2 Maximum Cut

Given a graph G, a *cut* in G is a partition (U, V) of the vertex set together with all the edges having exactly one endpoint in both parts. The *size* of the cut is the number of its edges. The *MaxCut* of G is the maximum size of a cut, denoted by $\mathrm{mc}(G)$. The MaxCut is among one the most studied graph parameters, lying at the intersection of theoretical computer science [32, 40, 51], extremal combinatorics [2, 3, 30, 35] and probabilistic graph theory [23, 22, 27]. In theoretical computer science, one is usually interested in approximating the size of the MaxCut efficiently, and in extremal combinatorics the goal is often to establish good bounds on the MaxCut depending on various graph parameters, such as the number of edges or vertices of the graph.

A simple probabilistic argument shows that every graph with m edges has a cut of size at least $m/2$. Indeed, a random cut, chosen from the uniform distribution on all cuts, has size $m/2$ in expectation. The constant $1/2$ cannot be improved in general, and therefore when measuring the size of the MaxCut in G, it is often more natural to talk about the *surplus* of G, which is defined as $\mathrm{surp}(G) = \mathrm{mc}(G) - m/2$. The trivial bound can be improved, a fundamental result of Edwards [30, 31] shows that any graph G with m edges has $\mathrm{mc}(G) \geq \frac{m}{2} + \frac{\sqrt{8m+1}-1}{8}$ or, equivalently, that $\mathrm{surp}(G) \geq \frac{\sqrt{8m+1}-1}{8}$, which is sharp when G is a clique on an odd number of vertices.

In general, if G is a disjoint union of constantly many cliques, then the MaxCut of G is of size $m/2 + O(\sqrt{m})$. This raises the following natural question.

Can this bound be improved if G is far from a disjoint union of cliques?

One way to ensure that a graph is far from a disjoint union of cliques is to assume that it does not contain some fixed graph H as a subgraph. The study of the size of the MaxCut, and in turn the surplus, in such graphs was initiated by Erdős and Lovász in the 1970's (see [35]). One of the first major results in the area is due to Alon [2], who proved that if a graph G has m edges and no triangles, then $\mathrm{surp}(G) = \Omega(m^{4/5})$, and this bound is tight. There are two natural ways to generalize this result - one is to study graphs without short cycles and the other is to study graphs avoiding K_r, the complete graph on r vertices.

The surplus in graphs without short cycles, studied in [3, 6, 10, 39], proved to be the easier of these two problems, and [10, 39] achieve tight bounds for this problem. On the other hand, finding the size of the minimum surplus in K_r-free graphs seems to be much more difficult. Alon, Bollobás, Krivelevich, and Sudakov [6] proved that for every r, there exists $\varepsilon_r > 0$ such that every K_r-free graph has surplus at least $m^{1/2+\varepsilon_r}$. This was improved by Carlson, Kolla, Li, Mani, Sudakov, and Trevisan [18], and then Glock, Janzer, and Sudakov [39] established $\mathrm{surp}(G) \geq m^{\frac{1}{2}+\frac{3}{4r-2}}$.

However, these results seem to be far from the truth, since Alon, Bollobás, Krivelevich and Sudakov conjectured in [3] that the answer should be $\mathrm{surp}(G) \geq m^{3/4+\varepsilon_r}$ for some $\varepsilon_r > 0$. This conjecture is still wide open, and, in fact, for a long time, it was a tantalizing open problem to find any absolute constant $\varepsilon > 0$ (independent of r), such that every K_r-free graph has surplus $\Omega_r(m^{1/2+\varepsilon})$. Glock, Janzer and Sudakov [39] write *"Arguably, the main open problem is to decide whether there exists a positive absolute constant ε such that any K_r-free graph with m edges has surplus $\Omega_r(m^{1/2+\varepsilon})$."* Our next main result not only proves this, but shows that we can achieve such a large surplus by forbidding extremely large cliques as well.

Theorem 1.3. *For every $\delta > 0$ there exists $\varepsilon > 0$ such that the following holds for every sufficiently large m. Let G be a graph with m edges such that G contains no clique of size $m^{1/2-\delta}$. Then G has a cut of size at least $\frac{m}{2} + m^{1/2+\varepsilon}$.*

In the very extreme case, Balla, Hambardzumyan, and Tomon [9] recently showed that graphs with clique number $o(\sqrt{m})$ already have surplus $\omega(m^{1/2})$. Despite the similarity between this result and the previous theorem, there is no implication between the two due to the hidden dependencies. The methods achieving these results are also very different, despite both being algebraic in nature.

Coming back to the motivating question of whether graphs with small surplus must necessarily look like unions of cliques, we also prove the following stability result.

Theorem 1.4. *There exists $\varepsilon > 0$ such that the following holds for every sufficiently large n. If G is an n-vertex m-edge graph with no cut of size larger than $\frac{m}{2} + n^{1+\varepsilon}$, then G is $n^{-\varepsilon}$-close to the disjoint union of cliques.*

There is a close relationship between the MaxCut of a graph G and its smallest eigenvalue. It is well known that $\mathrm{surp}(G) \leq |\lambda_n|n$ (see e.g. Claim 5.1 for a short proof). For many algebraically defined graph families, we also have $\mathrm{surp}(G) = \Theta(|\lambda_n|n)$, but in general these quantities can be far apart. A good way to think about the surplus as a robust version of the smallest eigenvalue: in many natural cases $\mathrm{surp}(G) = \Theta(|\lambda_n|n)$, but $\mathrm{surp}(G)$ is much less sensitive to local modifications. Therefore, it is not unexpected that we obtain similar results as in the smallest eigenvalue case, as we present in the next section.

The proof of Theorem 1.3 is presented in Section 12, and the proof of Theorem 1.4 in Section 13.2.

1.3 Smallest eigenvalue

A central topic of spectral graph theory is understanding the structure of graphs, whose adjacency matrix has large smallest eigenvalue. Let G be an n-vertex graph and let λ_n denote the smallest eigenvalue of its adjacency matrix. A simple consequence of the Cauchy-interlacing theorem is that if G is non-empty, then $\lambda_n \leq -1$ with equality if and only if G is the disjoint union of cliques. In the 1970's, Cameron, Goethels, Seidel, and Shult [17] gave a complete characterization of graphs satisfying $|\lambda_n| \leq 2$, which are exactly generalized line graphs and some sporadic examples with at most 36 vertices. More recently, Koolen, Yang and Yang [58] obtained a partial characterization in the case $|\lambda_n| \leq 3$ by integral lattices. Beyond these specific values, much less is known. Kim, Koolen, and Yang [52] proved the following structure theorem for regular graphs satisfying $|\lambda_n| \leq \lambda$. One can find dense induced subgraphs $Q_1, \dots, Q_c$ in G such that each vertex lies in at most λ of $Q_1, \dots, Q_c$, and almost all edges are covered by the union of $Q_1, \dots, Q_c$. However, the proof of this is based on certain forbidden subgraph characterizations and Ramsey theoretic arguments, and the results are no longer meaningful if λ grows faster than polylogarithmic in n. For highly structured graphs, such as strongly regular graphs (SRG), it is known [57] that if $|\lambda_n|$ is at most a small polynomial of the average degree, then the graph belongs to one of two special families. However, these results rely on the highly structured nature of SRGs. We refer the interested reader to the survey of Koolen, Cao, and Yang [56] for a general overview of the topic.

The smallest eigenvalue of the adjacency matrix also has theoretical importance. The celebrated Hoffman bound (see e.g. [43]) states that it controls the independence number of the graph. In particular, if G is an n-vertex d-regular graph, then $\alpha(G) \leq \frac{n|\lambda_n|}{|\lambda_n|+d}$. Furthermore, the maximum of $|\lambda_n|$ and the second largest eigenvalue λ_2 determines the expansion and mixing properties of the graph [4], and as we will discuss later, λ_n controls the maximum cut.

Many of these results show that the property of having small $|\lambda_n|$ and the existence of large trivial substructures, such as cliques, are interconnected. However, such results were previously only known when $|\lambda_n|$ is bounded by a constant, or growing very slowly with n. We prove that this phenomenon already starts to appear when $|\lambda_n| < n^{1/4-o(1)}$, and we show that graphs with smallest eigenvalue below this threshold converge to a trivial structure: a disjoint union of cliques. The exponent $1/4$ is also sharp, a celebrated construction of de Caen [26] related to equiangular lines provides a graph with smallest eigenvalue $|\lambda_n| = \Theta(n^{1/4})$ which is far from the disjoint union of cliques. We say that an n-vertex graph G is *μ-close* to some family of graphs $\mathcal{F}$ if the edit distance of G to some member of $\mathcal{F}$ is at most μn^2.

Theorem 1.5. *Let $\varepsilon, \delta > 0$, then the following holds for every sufficiently large n. Let G be an n-vertex graph such that $|\lambda_n| \leq n^{1/4-\varepsilon}$. Then G is δ-close to the disjoint union of cliques.*

The previous theorem only ensures $o(1)$-closeness in case $|\lambda_n| \leq n^{1/4-o(1)}$. However, by requiring a slightly stronger upper bound on $|\lambda_n|$, we can also establish polynomial proximity to a disjoint union of cliques.

Theorem 1.6. *Let $\varepsilon > 0$, then there exists $\alpha > 0$ such that the following holds for every sufficiently large n. Let G be an n-vertex graph such that $|\lambda_n| \leq n^{1/6-\varepsilon}$. Then G is $n^{-\alpha}$-close to the disjoint union of cliques.*

While these results give strong structural results about somewhat dense graphs with small $|\lambda_n|$, they are no longer meaningful for sparse graphs G. On the other hand, it is not possible to formulate any reasonable extension of the previous theorems for sparse graphs, as the following example shows. The line graph of a graph always satisfies that $|\lambda_n| \leq 2$, but the line graph of the complete graph K_s has $m = \Theta(s^3)$ edges and it is not possible to add/remove $o(m)$ edges to get a disjoint union of cliques.

Despite this, we recall that Theorem 1.2 shows that large cliques, with size comparable to the average degree, do emerge in graphs of any sparsity and small $|\lambda_n|$. This suggests that such graphs might be close to the blow-up of much smaller graphs, and shows that trivial structures start to appear at any sparsity, assuming $|\lambda_n|$ is sufficiently small.

We prove Theorem 1.5 in Section 9, Theorem 1.6 in Section 13.1.

1.4 Alon–Boppana theorem

The Alon–Boppana theorem [63] is a cornerstone result of spectral graph theory. It states that if G is an n-vertex d-regular graph, then the second largest eigenvalue λ_2 of the adjacency matrix is at least

$$\lambda_2 \geq 2\sqrt{d-1} - o_n(1).$$

This result is often misquoted, with the $o_n(1)$ term forgotten or not understood properly. In its precise formulation, the Alon–Boppana theorem states that if D is the diameter of G, then $\lambda_2 \geq 2\sqrt{d-1} - \frac{2\sqrt{d-1}}{\lfloor D/2 \rfloor}$. In particular, if $D \to \infty$, which is satisfied in case $d = n^{o(1)}$, one gets the former lower bound. For fixed d, families of graphs satisfying $\max\{|\lambda_n|, \lambda_2\} \leq 2\sqrt{d-1}$ are called Ramanujan graphs, and their existence is known for many different values of d [47]. A breakthrough of Friedman [38] shows that random d-regular graphs are close to being Ramanujan. Since the spectral gap $d - \lambda_2$ controls the expansion properties of the graphs, the Ramanujan graphs are optimal expanders. For this reason, such graphs are of great interest in the design of resilient networks, with countless further applications in theoretical computer science and extremal combinatorics.

In the case where the diameter D is at most three, which can already happen if $d \approx n^{1/3}$, the Alon–Boppana bound is no longer meaningful. Also, one cannot hope for the bound $\lambda_2 = \Omega(\sqrt{d})$ to hold unconditionally; for example the complete bipartite graph has $\lambda_2 = 0$. Recently, a number of authors [8, 11, 48, 64] studied the second eigenvalue in the case of denser graphs, and uncovered some highly unexpected behavior of its extremal value. In particular, [64] (see also [11] for a short note) proved that $\lambda_2 = \Omega(d^{1/2})$ continues to hold for $d \leq n^{2/3}$, however, $\lambda_2 = \Omega(n/d)$ for $d \in [n^{2/3}, n^{3/4}]$, and this is (essentially) sharp by an old strongly regular graph construction of Metz (see [75]). Moreover, for $d \in [n^{3/4}, (1/2 - \varepsilon)n]$, we have $\lambda_2 \geq \Omega_\varepsilon(d^{1/3})$, which is also sharp for $d = \Omega(n)$ by recent constructions of Davis, Huczynska, Johnson, and Polhill [25]. As we observed earlier, if $d = n/2$, we might have $\lambda_2 = 0$ by the complete bipartite graph. In general, when $d = (1 - 1/r)n$ for some positive integer r, then the *Turán graph* $T_r(n)$, the complete r-partite graph with parts of size n/r, is d-regular and satisfies $\lambda_2 = 0$.

However, what happens when d is not of the form $(1 - 1/r)n$, or G is far from a Turán graph? The methods of [64] and related papers no longer apply when $d > n/2$, and there are no obvious further obstructions for having large second eigenvalue. In [64], it was conjectured that the answer to the second question is $\Omega(n^{1/4})$, which is then sharp by the equiangular lines construction of de Caen [26]. Considering complements, Theorem 1.5 immediately implies an almost complete solution of this conjecture. If G is a regular graph with second eigenvalue λ_2, then the complement of G has smallest eigenvalue $-\lambda_2 - 1$.

Theorem 1.7. *Let $\varepsilon, \delta > 0$, then the following holds for every sufficiently large n. Let G be an n-vertex d-regular graph such that $\lambda_2 \leq n^{1/4-\varepsilon}$. Then G is δ-close to a Turán graph. Thus, if $\lambda_2 < n^{1/4-\varepsilon}$, then*

$$\frac{d}{n} \in \left\{1 - \frac{1}{r} : r \in \mathbb{Z}^+\right\} + [-\delta, \delta].$$

1.5 Bisection width

The *bisection width* of a graph is defined as the minimum number of edges crossing a balanced partition of the vertex set, and it is denoted by $\mathrm{bw}(G)$. As a natural dual to the maximum cut, this parameter is also of central interest in theoretical computer science [45, 46, 50], probabilistic [12, 28, 27, 29] and extremal graph theory [1, 64, 66].

It is convenient to measure the bisection width via the *deficit*, which is defined as

$$\mathrm{dfc}(G) = e(G)\left(\frac{1}{2} + \frac{1}{2n-2}\right) - \mathrm{bw}(G).$$

By the uniform random balanced cut, the deficit is always non-negative, and if G is a regular graph that is neither empty nor complete, then $\mathrm{dfc}(G) = \Omega(n)$, see e.g. [64]. This is optimal if G is a Turán graph.

A classic result of Alon [1] states that if G is d-regular, and $d = O(n^{1/9})$, then $\mathrm{dfc}(G) = \Omega(\sqrt{d}n)$, which is optimal for random d-regular graphs. Recently, Räty, Sudakov, and Tomon [64] greatly extended this bound by showing that

$$\mathrm{dfc}(G) = \begin{cases} \Omega(\sqrt{d}n) & \text{if } d \leq n^{2/3}, \\ \Omega(n^2/d) & \text{if } d \in [n^{2/3}, n^{4/5}], \\ \tilde{\Omega}(d^{1/4}n) & \text{if } d \in [n^{4/5}, (1/2-\varepsilon)n]. \end{cases}$$

These results are sharp for $d \in [1, n^{3/4}]$, and there are d-regular graphs for $d \approx n/3$ with deficit $O(n^{4/3})$. For $d = n(1-1/r)$, where r is a positive integer, the Turán graph $T_r(n)$ shows that we cannot hope for a bound better than $\Omega(n)$. Räty, Sudakov, and Tomon [64] conjectured that Turán graphs are the only obstruction to large deficit. Using the terminology of *positive discrepancy*, they conjectured that if $\mathrm{dfc}(G) = o(n^{5/4})$, then G is $o(1)$-close to a Turán graph. We prove that this conjecture holds qualitatively, by establishing the bisection width analogue of our MaxCut result (Theorem 1.4).

Theorem 1.8. *There exists $\varepsilon > 0$ such that the following holds for every sufficiently large n. Let G be an n-vertex d-regular graph. If the bisection width of G is more than $\frac{dn}{4} - n^{1+\varepsilon}$, then G is $n^{-\varepsilon}$-close to a Turán graph. Thus, if $\mathrm{dfc}(G) \leq n^{1+\varepsilon}$, then*

$$\frac{d}{n} \in \left\{1 - \frac{1}{r} : r \in \mathbb{Z}^+\right\} + [-n^{-\varepsilon}, n^{\varepsilon}].$$

The proof of Theorem 1.8 is presented in Section 13.2.

2 Proof overview

First, we outline the proof of Theorem 1.5, that is, that if a graph G has smallest eigenvalue $|\lambda_n| \leq n^{1/4-\varepsilon}$, then G is δ-close to a disjoint union of cliques. Let A be the adjacency matrix of G. We study the identity $A = A \circ A$ from a spectral perspective, where $\circ$ denotes the Hadamard product (see Section 4 for formal definitions). Writing $A = \sum_{i=1}^n \lambda_i v_i v_i^T$ for the spectral decomposition, we get that

$$\sum_{i=1}^n \lambda_i v_i v_i^T = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j (v_i \circ v_j)(v_i \circ v_j)^T. \tag{1}$$

But how to use this identity? An instructive case is when G is a Cayley graph on some finite abelian group $(\Gamma, +)$. In this case, (1) reduces to a clean convolution relation: $\lambda_a = \frac{1}{n}\sum_{b+c=a} \lambda_b \lambda_c$ for all $a \in \Gamma$. Here, the eigenvalues are re-indexed by the group elements, and the identity follows from special properties of the characters of the group, which are also the eigenvectors of G. See Brouwer and Haemers [16] for further details. We can almost ignore the negative terms in the sum since the smallest eigenvalue is bounded. Hence, this convolution relation basically tells that large eigenvalues reinforce each other, i.e. if $\lambda_b, \lambda_c \geq T$, then $\lambda_{b+c} \gtrsim T^2/n$. This motivates the definition $S_T = \sum_{\lambda_i \geq T} \lambda_i$, the spectral weight above threshold T. Summing over all $\lambda_b \geq T$ and $\lambda_c \geq T$, the above observation essentially gives that

$$S_{T^2/n} = \sum_{\lambda_a \geq T^2/n} \lambda_a \gtrsim \frac{1}{n} \sum_{\lambda_b, \lambda_c \geq T} \lambda_b \lambda_c = \frac{1}{n} S_T^2.$$

Surprisingly, this does generalize to arbitrary graphs to give the curious recursive inequality on the sum of large eigenvalues: for all $T \geq 2|\lambda_n|\sqrt{n}$,

$$4n S_{\frac{T^2}{2n}} \geq S_T^2. \tag{2}$$

We achieve this by compressing both sides of (1) to the subspace W spanned by the vectors $v_i \circ v_j$ where $\lambda_i, \lambda_j \geq T$; see Section 6.1 for a detailed argument. In terms of the quadratic sum of the eigenvalues, this implies that most contribution comes from large eigenvalues; we show this in Section 6.3. But this means that A is close to a low-rank positive semidefinite matrix in the Frobenius norm, which is only possible if G is close to a disjoint union of cliques. We prove this in Section 9.

Now we discuss the proof of Theorem 1.6, which requires a lot more work. Recall that this theorem states that if a graph G has smallest eigenvalue $|\lambda_n| \leq n^{1/6-\varepsilon}$, then G is $n^{-\alpha}$-close to a disjoint union of cliques. The bottleneck in the previous argument is the last part, it requires that the rank of the low-rank approximation of A is at most constant, in which case we can establish $o(1)$-closeness to a disjoint union of cliques. In order to overcome this, we first aim to show that either G is already sparse (in which case G is $n^{-\alpha}$-close to the empty graph), or G contains very large cliques. If we are able to do this, then we repeatedly pull out large cliques, which gives enough structure to easily conclude the desired result. In order to find our large cliques, we go through 4 phases of densification, i.e. we find denser and denser subgraphs of G. We keep in mind that the Cauchy interlacing theorem ensures that the induced subgraphs of G also have large smallest eigenvalue. In what follows, we use c to denote a small positive constant depending only on ε and α, and different occurrences of c might denote different quantities.

Phase 1. We show that G contains an unusually high number of triangles. We count triangles by the cubic sum of eigenvalues, and we argue that this sum is large because most of the mass of the quadratic sum of eigenvalues is concentrated on the few largest eigenvalues. Having many triangles means that we can find a vertex whose neighbourhood is much denser than G. We repeat this process until we find a subgraph G_1 on n^{1-c} vertices of positive constant density. This phase of the argument requires $|\lambda_n| \leq n^{1/6-\varepsilon}$. This can be found in Section 8.

Phase 2. In Theorem 1.5, we already established that G_1 is $o(1)$-close to a disjoint union of cliques. Therefore, if G_1 has positive constant edge density, this implies that G_1 contains a linear size subgraph G_2 of edge density $1 - o(1)$. This can be found in Section 9.

Phase 3. For very dense graph, we employ a new method, inspired by the works of [64]. Let G be an n-vertex graph of density $1 - p$, with $10^{-5} > p > n^{-1/2}$, and $|\lambda_n| \leq n^{1/6-o(1)}$. We consider the matrix

$$B = A - \lambda_1 v_1 v_1^T + |\lambda_n| I,$$

and we study the triple Hadamard product $B \circ B \circ B$. The matrix B is positive semidefinite, so the Schur product theorem ensures that $B \circ B \circ B$ is also positive semidefinite. But then $x^T(B \circ B \circ B)x \geq 0$ for every vector x. We choose x to be the characteristic vector of a carefully chosen, linear sized set $U \subset V(G)$, and argue that the previous inequality can be only satisfied if the density of $G[U]$ is at least $1 - O(p^3)$. Therefore, repeating this argument on G_2, we get a rapid density increment until we reach a graph G_3 of edge density $1 - n^{-1/2}$. This can be found in Section 10.

Phase 4. Above edge density $1-n^{-1/2}$, we may consider simple identities involving the eigenvalues. Let G be a graph of edge density at least $1 - n^{-1/2}$ and $|\lambda_n| \leq n^{1/6}$, and consider the four identities

$$\Delta(G) \geq \lambda_1, \quad \sum_{i=1}^{n} \lambda_i = 0, \quad \sum_{i=1}^{n} \lambda_i^2 = 2e(G), \quad \sum_{i=1}^{n} \lambda_i^3 \geq 0.$$

We find an almost regular subgraph G_4 of G_3, and use simple algebraic manipulations to argue that the expressions above can be only satisfied if G_4 has density at least $1 - n^{-1+c}$. At this point, we simply apply Turán's theorem to find a clique of size n^{1-c}. This can be found in Sections 11 and 12.

Next, we discuss Theorem 1.2, which states that any graph with average degree d and smallest eigenvalue $|\lambda_n| \leq d^{\gamma}$ contains a clique of size $d^{1-O(\gamma)}$. In order to prove this, we introduce a phase 0 of the densification process, which immediately lets us move to density at least n^{-c}, assuming $|\lambda_n|$ is sufficiently small with respect to the average degree. Then, we apply the previous four densification steps to conclude the proof.

Phase 0. Let G be a graph of average degree d, then we show that G contains a subgraph on d vertices of edge density $\Omega(1/|\lambda_n|)$. This follows by picking a vertex x with a set of d neighbours S, and then analyzing $v^T A v$ for an appropriately chosen v with support $\{x\} \cup S$. This can be found in Section 7.

In order to prove our results concerning graphs with small maximum cut, that is, Theorems 1.3 and 1.4, we follow the same steps. In Section 5, we present a toolkit that gives various lower bounds on the MaxCut based on the negative eigenvalues of the graph. With the help of these, instead of having a bound on $|\lambda_n|$, we can bound the sum, quadratic sum, and cubic sum of the negative eigenvalues. This allows us to transfer most of the machinery developed for graphs with bounded smallest eigenvalue to graphs with bounded MaxCut, but with the cost of incurring some losses quantitatively.

3 Chowla's cosine problem

In this section, we give a short self–contained proof of Theorem 1.1, assuming Theorem 1.2. We begin the section by recalling some standard notation. Let G be a finite group, $A \subset G$ be a symmetric subset (i.e. a set satisfying $A = A^{-1}$), and let $\Gamma = \mathrm{Cay}(G, A)$. Recall that $\mathrm{Cay}(G, A)$ is the *Cayley graph* on G generated by A, that is, the graph on vertex set G in which $x, y \in G$ are joined by an edge if $xy^{-1} \in A$. In case G is abelian, we use $+$ to denote the group operation. In this case, it is well known that the eigenvalues of G are the values of the discrete Fourier transform $\widehat{\mathbb{1}_A}$. In the special case $G = \mathbb{Z}/n\mathbb{Z}$, this gives that the eigenvalues of the Cayley graph are

$$\sum_{a \in A} e^{\frac{2\pi i}{n} \cdot a\xi} = \sum_{a \in A} \cos\left(\frac{2\pi a \xi}{n}\right)$$

for $\xi \in \mathbb{Z}/n\mathbb{Z}$. We restate Theorem 1.1 for the reader's convenience.

Theorem 3.1. *There exists an absolute constant $\gamma > 0$ such that for any set A of positive integers, there exists $x \in [0, 2\pi]$ such that*

$$\sum_{a \in A} \cos(ax) \leq -\Omega(|A|^{\gamma}).$$

Proof. Without loss of generality, we may assume $|A|$ is sufficiently large. Let $n > 4|A|$ be a prime larger than all elements of A, and let $\Gamma = \mathrm{Cay}(\mathbb{Z}/n\mathbb{Z}, A \cup -A)$ be the Cayley graph with the generating set $A \cup -A$ (where $-A = \{-a : a \in A\}$). Then Γ is an n-vertex d-regular graph with $d = 2|A|$. Each $\xi \in \mathbb{Z}/n\mathbb{Z}$ corresponds to an eigenvalue of Γ given by

$$\lambda_\xi = \sum_{a \in A \cup -A} e^{2\pi i a \xi / n} = 2 \sum_{a \in A} \cos\left(\frac{2\pi a \xi}{n}\right)$$

for some $\xi \in \mathbb{Z}/n\mathbb{Z}$. Hence, if we denote by λ_n the smallest eigenvalue of Γ, there exists $x = \frac{2\pi\xi}{n}$ such that

$$\sum_{a \in A} \cos(ax) = \frac{1}{2}\lambda_n.$$

Thus, our aim is to show that $|\lambda_n| \geq d^\gamma$, for an absolute constant $\gamma > 0$, since this implies that $\sum_{a\in A}\cos(ax) \leq -d^\gamma \leq -\Omega(|A|^\gamma)$, as needed. Assume, for the sake of contradiction, that $|\lambda_n| < d^\gamma$. We use Theorem 1.2 to find a clique S of size $|S| = d^{1-c\gamma}$ for some absolute constant c, assuming d is sufficiently large with respect to γ. We now argue that Γ cannot contain such large cliques.

Claim 3.2. *There exists a non-zero $t \in \mathbb{Z}/n\mathbb{Z}$ such that $|(t+S)\cap S| \geq |S|(|S|-1)/d$.*

Proof. As S is a clique in Γ, we have $S - S \subset A \cup -A \cup \{0\}$. By a simple averaging argument, there exists $t \in A \cup -A$ such that $s - s' = t$ for at least $\frac{|S|(|S|-1)}{2|A|} = \frac{|S|(|S|-1)}{d}$ pairs $(s,s') \in S \times S$. Hence, for at least $|S|(|S|-1)/d$ values of $s \in S$ we have $s + t \in S$, and therefore $|(t+S)\cap S| \geq |S|(|S|-1)/d$. □

For a positive integer k, let H_k is the graph that is formed by a clique of size $2k$, and an additional vertex connected to half of the vertices of the clique.

Claim 3.3. *The smallest eigenvalue μ of H_k satisfies $\mu < -\sqrt{k/2}$.*

Proof. Let $X \cup Y$ be the partition of C, where X is the neighbourhood of x_0. Let B be the adjacency matrix of H_k and let $v \in \mathbb{R}^{V(H_k)}$ be the vector defined as $v(x_0) = 1/\sqrt{2}$, $v(y) = -\frac{1}{2\sqrt{k}}$ if $y \in X$, and $v(y) = \frac{1}{2\sqrt{k}}$ if $y \in Y$. Then $||v||_2 = 1$ and

$$v^T Bv = 2\sum_{ab\in E(H_k)} v(a)v(b) = 2k\cdot\frac{1}{\sqrt{2}}\cdot\big(-\frac{1}{2\sqrt{k}}\big) + 4\cdot\binom{k}{2}\cdot\frac{1}{4k} - 2\cdot k^2\cdot\frac{1}{4k} = -\sqrt{\frac{k}{2}} - \frac{1}{2} < -\sqrt{\frac{k}{2}},$$

where the first term is the contribution of the edges x_0y, $y \in X$, the second term is the contribution of edges yy' with $y, y' \in X$ or $y, y' \in Y$, and the third term is the contribution of edges yy' with $y \in X$ and $y' \in Y$. □

By Claim 3.3, the smallest eigenvalue μ of H_k satisfies $\mu < -\frac{1}{2}\sqrt{k}$. Therefore, by the Cauchy interlacing theorem, Γ does not contain H_k as an induced subgraph for $k = 4d^{2\gamma}$. As $|S| > 2k$, each vertex of Γ sends either at most k edges to S, or at least $|S| - k$ edges. We prove that every vertex in Γ must send at least $|S| - k$ edges to S. This easily leads to a contradiction for n sufficiently large: this implies that there are at least $(n-|S|)(|S|-k) \geq \frac{n}{2}\cdot\frac{|S|}{2} > d|S|$ edges with an endpoint in S, contradicting that Γ is d-regular.

Claim 3.4. *Every $v \in V(\Gamma)$ sends at least $|S| - k$ edges to S.*

Proof. We prove by induction on ℓ that every vertex of $\ell t + S$ sends at least $|S| - k$ edges to S. As every vertex $v \in V(\Gamma)$ is contained in some $\ell t + S$, this finishes the proof. The base case $\ell = 0$ is trivial, so let $\ell \geq 1$. By our induction hypothesis and translation invariance, every vertex $v \in \ell t + S$ sends at least $|S| - k$ edges to $t + S$. But then v sends at least

$$|S \cap (t+S)| - k = \Omega(|S|^2/d) - k = \Omega(d^{1-2c\gamma}) - 4d^{2\gamma} > 4d^{2\gamma} = k$$

edges to $S \cap (t+S)$, and in particular, more than k edges to S. Here the last inequality holds if we take $\gamma = \frac{1}{3+2c}$ and $|A|$ sufficiently large. Therefore, as Γ contains no induced copy of H_k, v must send at least $|S| - k$ edges to S, and we are done. □

□

4 Preliminaries

We recall some basic facts and standard notation from linear algebra and graph theory. The *edge density* of a graph G is $m/\binom{n}{2}$, where $m = e(G)$ is the number of edges. Given a subset U of the vertices, $G[U]$ denotes the subgraph of G induced on vertex set U. Also, if $V \subset V(G)$ is disjoint form U, then $G[U, V]$ is the bipartite subgraph of $V(G)$ induced between U and V. The *complement* of G is denoted by $\overline{G}$. The *maximum degree* of G is denoted by $\Delta(G)$, and the average degree by $d(G)$. If G has n vertices, we will often identify the set of vertices of G with $[n]$.

The *MaxCut* of G, denoted by $\mathrm{mc}(G)$, is the maximum size of a cut, where a *cut* is a partition (U, V) of the vertices into two parts, with all the edges having exactly one endpoint in both parts. The size of a cut is the number of its edges. The *surplus* of G is defined as $\mathrm{surp}(G) = \mathrm{mc}(G) - m/2$, where m is the number of edges of G. Note that $\mathrm{surp}(G)$ is always nonnegative. A useful property of the surplus is that if G_0 is an induced subgraph of G, then $\mathrm{surp}(G_0) \le \mathrm{surp}(G)$, see e.g. [39].

Given an $n \times n$ real symmetric matrix M, we denote by $\lambda_1(M) \ge \dots \ge \lambda_n(M)$ the eigenvalues of M with multiplicity. As we will see, the surplus of G is controlled by the negative eigenvalues of the graph. If G is an n-vertex graph whose adjacency matrix is A, then we denote by $\lambda_i = \lambda_i(A)$ the eigenvalues of A, sometimes also calling them the eigenvalues of G. We also denote by $v_1, \dots, v_n$ a corresponding orthonormal basis of eigenvectors. By the Perron–Frobenius theorem, we may take v_1 to be a vector with non-negative entries, which we call the *principal eigenvector* of A. Furthermore, the corresponding eigenvalue satisfies $\lambda_1 \ge d(G)$. See the survey [24] as a general reference on the principal eigenvector.

Given two $n \times n$ matrices A and B, their scalar product is defined as

$$\langle A, B\rangle = \mathrm{tr}(AB^T) = \sum_{1 \le i,j \le n} A_{i,j} B_{i,j}.$$

The *Frobenius-norm* of an $n \times n$ matrix A is

$$\|A\|_F^2 = \langle A, A\rangle = \sum_{i,j=1}^{n} A_{i,j}^2.$$

If A is symmetric with eigenvalues $\lambda_1, \dots, \lambda_n$, then we also have

$$\|A\|_F^2 = \langle A, A\rangle = \mathrm{tr}(A^2) = \sum_{i=1}^{n} \lambda_i^2.$$

The *Hadamard product* (also known as entry-wise product) of A and B is the $n \times n$ matrix $A \circ B$ defined as $(A \circ B)_{i,j} = A_{i,j} B_{i,j}$. We denote the k-term Hadamard product $A \circ \dots \circ A$ by $A^{\circ k}$. A useful feature of the Hadamard product, which is a key component of our arguments, is that it preserves positive semidefiniteness.

Theorem 4.1 (Schur product theorem)**.** *If A and B are positive semidefinite matrices, then $A \circ B$ is also positive semidefinite.*

We also exploit the simple observation that if A is an adjacency matrix, then $A = A \circ A$. Another useful identity involving the Hadamard product is that if x, y, u, v are vectors, then

$$(xy^T) \circ (uv^T) = (x \circ u)(y \circ v)^T.$$

We also use the Hadamard product for vectors: for $u, v \in \mathbb{R}^n$, their Hadamard product vector $u \circ v \in \mathbb{R}^n$ is defined by $(u \circ v)(i) := u(i)v(i)$ for all $i \in \{1, \dots, n\}$.

5 Spectral lower bounds for the surplus

In this section, we present bounds on the MaxCut of a graph in terms of its spectrum. These inequalities will be crucial in transferring our results for the smallest eigenvalue to the MaxCut setting.

Claim 5.1. *For an n-vertex graph G with the smallest eigenvalue λ_n, we have* $\mathrm{surp}(G) \le |\lambda_n| n/4$.

Proof. Let A be the adjacency matrix of G. We can assign a vector with entries ± 1 to each cut $V(G) = X \cup Y$, by setting $x_u = 1$ if $u \in X$ and $x_u = -1$ otherwise. Then, the surplus of this cut equals $\frac{1}{2}\big(e(X,Y) - e(X) - e(Y)\big) = -\frac{1}{2}\sum_{\{u,v\}\subseteq V(G)} x_u A_{uv} x_v = -\frac{1}{4}\sum_{u,v\in V(G)} x_u A_{uv} x_v$. Hence, we have

$$\mathrm{surp}(G) = \frac{1}{4}\max_{x\in\{-1,1\}^n} -x^T A x = \frac{1}{4}\max_{x\in[-1,1]^n} -x^T A x.$$

But, we have $-x^T A x \le |\lambda_n| \|x\|_2^2$ for every vector $x \in \mathbb{R}^n$, and so $\mathrm{surp}(G) \le \frac{1}{4}|\lambda_n|\sqrt{n}^2 = |\lambda_n| n/4$. □

The key ingredient of the above proof is the relation $\mathrm{surp}(G) = \frac{1}{4}\max_{x\in[-1,1]^n} -x^T A x$, which can also be written as $\mathrm{surp}(G) = \frac{1}{4}\max_{x\in[-1,1]^n}\langle -A, xx^T\rangle$, where we observe that xx^T is a positive-semidefinite matrix with diagonal entries bounded by 1. As we will see, it will be very convenient to define the semidefinite relaxation of the surplus as follows. Given an n-vertex graph G with adjacency matrix A, define

$$\mathrm{surp}^*(G) = \max_X -\langle A, X\rangle,$$

where the maximum is taken over all $n \times n$ positive semidefinite matrices X such that $X_{i,i} \le 1$ for every $i \in [n]$. The following inequality between $\mathrm{surp}(G)$ and $\mathrm{surp}^*(G)$ can be found in [65], and it is a simple application of the graph Grothendieck inequality of Charikar and Wirth [19].

Claim 5.2 ([65])**.** *For every graph G, we have* $\mathrm{surp}^*(G) \ge \mathrm{surp}(G) \ge \Omega\Big(\frac{\mathrm{surp}^*(G)}{\log n}\Big)$.

The semidefinite relaxation $\mathrm{surp}^*(G)$ allows us to obtain lower bounds on the surplus using the negative eigenvalues of a graph G. Parts of the following lemma and similar bounds can be also found in [64, 65]. Given a graph G, let

$$\Delta^*(G) := \min\{\Delta(G), \Delta(\overline{G})\}.$$

Lemma 5.3. *There exists an absolute constant $c > 0$ such that the following holds. Let G be a graph on n vertices with eigenvalues $\lambda_i = \lambda_i(G)$, and let $\Delta^* = \Delta^*(G)$. Then*

(i) $\mathrm{surp}^*(G) \ge \sum_{\lambda_i<0} |\lambda_i|$

(ii) $\mathrm{surp}^*(G) \ge \frac{c}{\sqrt{\Delta^*+1}} \sum_{\lambda_i<0} \lambda_i^2$

(iii) $\mathrm{surp}^*(G) \ge \frac{c}{\Delta^*+1} \sum_{\lambda_i<0} |\lambda_i|^3$.

Before we prove Lemma 5.3, we briefly discuss two preliminary results. First, we show that the entries of eigenvectors corresponding to large eigenvalues are smoothly distributed. Then, we show that the entries of the principal eigenvector are especially well behaved.

Lemma 5.4. *Let G be an n-vertex graph, and let λ be an eigenvalue with normalized eigenvector v. Then*

$$||v||_\infty \le \frac{\sqrt{n}}{|\lambda|}.$$

Proof. For every $b \in [n]$, we have $\lambda v(b) = \sum_{b\sim i} v(i)$, where we use $x \sim y$ to denote that x is connected to y by an edge in G. By the inequality between the arithmetic and quadratic mean,

$$\frac{1}{n}\left|\sum_{b\sim i} v(i)\right| \leq \frac{1}{n}\sum_{i=1}^{n}|v(i)| \leq \sqrt{\frac{\sum_i v(i)^2}{n}} = \frac{1}{\sqrt{n}},$$

where we used that $\sum_{i=1}^{n} v_1(i)^2 = 1$. Hence $|\lambda||v(b)| \leq \sqrt{n}$. □

Lemma 5.5. *Let G be an n-vertex graph, whose complement has edge density $p \leq 1/10$ and maximum degree $\bar{\Delta} = \Delta(\overline{G})$. If v_1 is the principal eigenvector of G, then for each $i \in [n]$ we have*

$$\frac{1 - 2\bar{\Delta}/n}{\sqrt{n}} \leq v_1(i) \leq \frac{1 + 2p + 2/n}{\sqrt{n}}.$$

Proof. Let $d = d(G) = (1-p)(n-1)$ be the average degree of G, and recall that $\lambda_1 \geq d$. By Lemma 5.4,

$$v_1(b) \leq \frac{\sqrt{n}}{\lambda_1} \leq \frac{\sqrt{n}}{d} = \frac{\sqrt{n}}{(1-p)(n-1)} \leq \frac{1 + 2p + 2/n}{\sqrt{n}}.$$

In the last inequality, we used that $p < 1/10$. To prove the lower bound, we first observe that

$$1 = \sum_{i=1}^{n} v_1(i)^2 \leq \|v_1\|_\infty \sum_{i=1}^{n} v_1(i),$$

which implies that $\sum_{i=1}^{n} v_1(i) \geq \frac{\lambda_1}{\sqrt{n}}$. But then using the identity $Av_1 = \lambda_1 v_1$,

$$\lambda_1 v_1(b) = \sum_{i\sim b} v_1(i) \geq \sum_{i=1}^{n} v_1(i) - \bar{\Delta}\|v_1\|_\infty \geq \frac{\lambda_1}{\sqrt{n}} - \bar{\Delta}\frac{\sqrt{n}}{\lambda_1} = \frac{\lambda_1}{\sqrt{n}}\Big(1 - \bar{\Delta}\frac{n}{\lambda_1^2}\Big) \geq \frac{\lambda_1}{\sqrt{n}}\Big(1 - \frac{2\bar{\Delta}}{n}\Big),$$

where we used that $\lambda_1^2 \geq d^2 \geq n^2/2$ in the last inequality. Canceling λ_1 gives $v_1(b) \geq \frac{1-2\bar{\Delta}/n}{\sqrt{n}}$. □

Proof of Lemma 5.3. We begin by showing the inequalities *(i)* and *(iii)*, which we then combine to derive *(ii)*. Let $v_1, \dots, v_n$ be an orthonormal basis of eigenvectors corresponding to the eigenvalues $\lambda_1, \dots, \lambda_n$; so $A = \sum_{i=1}^{n} \lambda_i v_i v_i^T$. The inequalities *(i)* and *(iii)* will be shown by plugging in the appropriate test matrix X in the formula $\mathrm{surp}^*(G) = \max_X -\langle A, X\rangle$. Observe that, if we choose $X = \sum_{i=1}^{n} \alpha_i v_i v_i^T$ for some real numbers $\alpha_1, \dots, \alpha_n$, then

$$\langle A, X\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i\lambda_j \langle v_i v_i^T, v_j v_j^T\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i\lambda_j \langle v_i, v_j\rangle^2 = \sum_{i=1}^{n} \alpha_i\lambda_i.$$

(i) Let $X = \sum_{\lambda_i<0} v_i v_i^T$. Then X is positive semidefinite, and as $v_1, \dots, v_n$ is an orthonormal basis, we have

$$X_{j,j} = \sum_{\lambda_i<0} v_i(j)^2 \leq \sum_{i=1}^{n} v_i(j)^2 = \sum_{i=1}^{n}\langle v_i, \mathbf{e}_j\rangle^2 = \|\mathbf{e}_j\|^2 = 1.$$

Therefore,

$$\mathrm{surp}^*(G) \geq -\langle A, X\rangle = \sum_{\lambda_i<0} |\lambda_i|.$$

(iii) Let $\beta = \frac{1}{100(\Delta^*+1)}$, and $X = \beta\sum_{\lambda_i<0} \lambda_i^2 v_i v_i^T$. Then X is positive semidefinite. It is enough to prove that the diagonal entries of X are bounded by 1, as then $\mathrm{surp}^*(G) \geq -\langle A, X\rangle = \beta\sum_{\lambda_i<0}|\lambda_i|^3$.

First, we consider the (easier) case $\Delta^* = \Delta(G)$. We observe that

$$\beta A^2 - X = \beta \sum_{\lambda_i > 0} \lambda_i^2 v_i v_i^T$$

is positive semidefinite, so the diagonal entries of βA^2 dominate the diagonal entries of X. But $(A^2)_{i,i}$ is the degree of vertex i, so $(A^2)_{i,i} \leq \Delta \leq 1/\beta$, and the claim follows.

Next, we consider the case $\Delta^* = \Delta(\overline{G})$. We may assume that the edge density of $\overline{G}$ is less than $1/10$, otherwise $\Delta^* = \Omega(n)$, and the previous case implies $\text{surp}^*(G) = \Omega(\frac{1}{n}) \sum_{\lambda_i < 0} |\lambda_i|^3$. To show that $X_{i,i} \leq 1$, we analyze the matrix $B = A - \lambda_1 v_1 v_1^T$. Since

$$\beta B^2 - X = \beta \sum_{i \neq 0, \lambda_i > 0} \lambda_i^2 v_i v_i^T,$$

we have that $\beta B^2 - X$ is positive semidefinite. Hence, $(\beta B^2)_{i,i} \geq X_{i,i}$ for every $i \in [n]$. Therefore, it is enough to show that $(B^2)_{i,i} \leq 1/\beta = 100(\Delta^* + 1)$.

To show this, we first bound the entries of B. We denote by p the density of $\overline{G}$, and observe that $p \leq \frac{\Delta^* n/2}{\binom{n}{2}} = \frac{\Delta^*}{n-1}$. Then, Lemma 5.5 implies that for any $i, j \in [n]$ we have

$$1 - \frac{5(\Delta^* + 1)}{n} \leq (n-1)(1-p)\left(\frac{1 - 2\Delta^*/n}{\sqrt{n}}\right)^2 \leq \lambda_1 v_1(i) v_1(j) \leq n\left(\frac{1 + 2p + 2/n}{\sqrt{n}}\right)^2 \leq 1 + \frac{5(\Delta^* + 1)}{n}.$$

Therefore, for every $i, j \in [n]$, if $ij \in E(G)$ and $A_{i,j} = 1$, then $|B_{i,j}| \leq \frac{5(\Delta^*+1)}{n}$. Otherwise, we have $|B_{i,j}| \leq 1 + 5(\Delta^* + 1)/n \leq 6$. From this, we have

$$(B^2)_{i,i} = \sum_{j=1}^{n} (B_{i,j})^2 \leq 36\Delta^* + n\frac{25(\Delta^* + 1)^2}{n^2} \leq 100(\Delta^* + 1).$$

(ii) We show that *(i)* and *(iii)* can be combined to give the desired lower bound on $\text{surp}^*(G)$. Namely, we have

$$\text{surp}^*(G)^2 \geq \beta \left(\sum_{\lambda_i < 0} |\lambda_i|^3\right)\left(\sum_{\lambda_i < 0} |\lambda_i|\right) \geq \beta \left(\sum_{\lambda_i < 0} \lambda_i^2\right)^2.$$

Note that the first inequality is the combination of *(i)* and *(iii)*, while the second one is simply the Cauchy-Schwartz inequality applied to the sequences $(|\lambda_i|^3)_{\lambda_i < 0}$ and $(|\lambda_i|)_{\lambda_i < 0}$. Taking square roots then proves *(ii)*. □

Finally, we remark two simple, but important properties of $\text{surp}^*(\cdot)$, that will be used repeatedly.

Claim 5.6. *If G' is an induced subgraph of G, then* $\text{surp}^*(G') \leq \text{surp}^*(G)$.

Proof. Let A' be the adjacency matrix of G' and let $X' \in \mathbb{R}^{V(G') \times V(G')}$ be a matrix such that X' is positive semidefinite, $X_{i,i} \leq 1$ for every $i \in V(G')$, and $\text{surp}^*(G') = -\langle A', X' \rangle$. Let $X \in \mathbb{R}^{V(G) \times V(G)}$ be the matrix that agrees with X' on every entry $(x, y) \in V(G') \times V(G')$, and zero everywhere else. Then

$$\text{surp}^*(G) \geq -\langle A, X \rangle = -\langle A', X' \rangle = \text{surp}^*(G').$$

□

Claim 5.7. *If G is an n-vertex graph with smallest eigenvalue λ_n, then* $\text{surp}^*(G) \leq |\lambda_n| n$.

Proof. Let $X \in \mathbb{R}^{n\times n}$ be a positive semidefinite matrix such that $X_{i,i} \leq 1$ for every $i \in [n]$. Let $A = \sum_{i=1}^n \lambda_i v_i v_i^T$ be the spectral decomposition of A, then

$$-\langle A, X\rangle = -\sum_{i=1}^n \lambda_i \langle v_i v_i^T, X\rangle \leq \sum_{i=1}^n |\lambda_n| \langle v_i v_i^T, X\rangle = |\lambda_n| \langle I, X\rangle \leq |\lambda_n| n.$$

In the first inequality, we used the fact that the scalar product of positive semidefinite matrices is nonnegative (which follows easily by considering spectral decompositions). □

6 Main lemmas

An important component of the proofs of our main results is the notion and properties of the *subspace compression* of matrices. This is a special instance of the compression of linear operators, see the book of Halmos [44] as a general reference.

W-compression and W-trace. Let $W < \mathbb{R}^n$ be a subspace. We denote by Π_W the orthogonal projection matrix onto W. Given an $n \times n$ symmetric matrix M, the *W-compression* of M is the symmetric matrix

$$M_W := \Pi_W M \Pi_W.$$

Furthermore, the W-trace of M is

$$\mathrm{tr}_W(M) := \mathrm{tr}(M_W).$$

Clearly, tr_W is a linear functional. Observe that if $M = uu^T$, then $M_W = (\Pi_W u)(\Pi_W u)^T$ and thus

$$\mathrm{tr}_W(uu^T) = ||\Pi_W u||_2^2.$$

Finally, given an orthonormal basis $w_1, \dots, w_d$ of W, the W-trace can be calculated as

$$\mathrm{tr}_W(M) = \sum_{i=1}^d w_i^T M w_i.$$

From this equality, it also follows that $\mathrm{tr}_W(I) = \dim(W)$. We present an upper bound on the W-trace that will be used repeatedly in our proofs.

Lemma 6.1. $|\mathrm{tr}_W(M)| \leq \dim(W)^{1/2} ||M||_F$.

Proof. Let $M = \sum_{i=1}^n \mu_i v_i v_i^T$ be the spectral decomposition of M. Then

$$\begin{aligned} |\mathrm{tr}_W(M)| &= \left| \sum_{i=1}^n \mu_i \,\mathrm{tr}_W(v_i v_i^T) \right| = \left| \sum_{i=1}^n \mu_i ||\Pi_W v_i||_2^2 \right| \leq \sum_{i=1}^n |\mu_i| \cdot ||\Pi_W v_i||_2 \\ &\leq \left(\sum_{i=1}^n \mu_i^2 \right)^{1/2} \cdot \left(\sum_{i=1}^n ||\Pi_W v_i||_2^2 \right)^{1/2} = ||M||_F \dim(W)^{1/2}. \end{aligned}$$

Here, the first inequality uses that $||\Pi_W v_i|| \leq 1$ for every $i \in [n]$, and the second inequality is due to the Cauchy-Schwartz inequality. □

The importance of the W-compression and W-trace is that it allows us to analyze the contribution of the top eigenvalues of a matrix, by choosing an appropriate subspace W. Given a graph G with adjacency matrix A, eigenvalues $\lambda_1 \geq \dots \geq \lambda_n$ and a real number T, we write

$$S_T(G) = \sum_{i:\lambda_i \geq T} \lambda_i.$$

If the graph G is clear from the context, we simply write S_T instead of $S_T(G)$. Furthermore, let $N_T = N_T(G)$ denote the number of eigenvalues at least T. We will use repeatedly that

$$N_T \leq \frac{S_T}{T}.$$

The next lemma gives a simple upper bound on the trace of the W-compression of A.

Lemma 6.2. *Let G be an n-vertex graph with adjacency matrix A and let $W < \mathbb{R}^n$. Then for every $K > 0$,*

$$\mathrm{tr}_W(A) \leq S_K + K \dim(W).$$

Proof. We have

$$\mathrm{tr}_W(A) = \sum_{i=1}^{n} \lambda_i ||\Pi_W v_i||_2^2 \leq \sum_{\lambda_i \geq K} \lambda_i + K \sum_{i=1}^{n} ||\Pi_W v_i||_2^2 = S_K + K \dim(W).$$

□

6.1 Main lemma - least eigenvalue version

The following lemma is the heart of our argument. It shows a curious recursive relation between the sums of the largest eigenvalues. Later, we show how to use this relation to conclude that the quadratic sum of all but the top eigenvalues are negligible. Considering the Frobenius norm, this is equivalent to saying that the adjacency matrix is well approximated by the top part of the spectral decomposition.

Lemma 6.3. *Let G be an n-vertex graph. If $T \geq 2|\lambda_n|\sqrt{n}$, then*

$$4nS_{\frac{T^2}{2n}} \geq S_T^2.$$

Proof. The main idea is to analyze the identity

$$A = A \circ A.$$

Here, we further rewrite $A \circ A$ as

$$A = (A + |\lambda_n| I) \circ (A + |\lambda_n| I) - \lambda_n^2 I, \tag{3}$$

where the advantage is that $A + |\lambda_n| I$ is a positive semidefinite matrix. Let $v_1, \dots, v_n$ be an orthonormal basis of eigenvectors corresponding to the eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$. By looking at the spectral decompositions $A = \sum_{i=1}^n \lambda_i v_i v_i^T$ and $A + |\lambda_n| I = \sum_{i=1}^n (\lambda_i + |\lambda_n|) v_i v_i^T$, the identity (3) is equivalent to

$$\sum_{i=1}^{n} \lambda_i v_i v_i^T = \left(\sum_{i=1}^{n} (\lambda_i + |\lambda_n|) v_i v_i^T \right)^{\circ 2} - \lambda_n^2 I = \sum_{i,j} (\lambda_i + |\lambda_n|)(\lambda_j + |\lambda_n|)(v_i \circ v_j)(v_i \circ v_j)^T - \lambda_n^2 I.$$

We show that the desired inequality can be deduced by considering the W-traces of both sides of this identity for an appropriately chosen W. Define W to be the subspace generated by those vectors $v_i \circ v_j$, where λ_i and λ_j are both at least T. Formally,

$$W = \langle v_i \circ v_j : \lambda_i, \lambda_j \geq T \rangle.$$

Note that

$$\dim(W) \leq N_T^2 \leq \frac{S_T^2}{T^2}.$$

Let $K = T^2/2n$. Then by Lemma 6.2, we have

$$\mathrm{tr}_W(A) \le S_K + K\dim(W) \le S_K + K\frac{S_T^2}{T^2} = S_{\frac{T^2}{2n}} + \frac{S_T^2}{2n}.$$

On the other hand

$$\begin{aligned}\mathrm{tr}_W(A \circ A) &= \mathrm{tr}_W((A + |\lambda_n|I) \circ (A + |\lambda_n|I) - \lambda_n^2 I) \\ &= \sum_{i,j}(\lambda_i + |\lambda_n|)(\lambda_j + |\lambda_n|)||\Pi_W v_i \circ v_j||_2^2 - \lambda_n^2 \dim(W).\end{aligned} \tag{4}$$

Note that if $\lambda_i, \lambda_j \ge T$, then $v_i \circ v_j \in W$, so $\Pi_W v_i \circ v_j = v_i \circ v_j$ and

$$(\lambda_i + |\lambda_n|)(\lambda_j + |\lambda_n|)||\Pi_W v_i \circ v_j||_2^2 \ge \lambda_i\lambda_j||v_i \circ v_j||_2^2.$$

Also, each term in the sum is nonnegative, so if $\lambda_i < T$ or $\lambda_j < T$, we simply lower bound the contribution of $(\lambda_i + |\lambda_n|)(\lambda_j + |\lambda_n|)||\Pi_W v_i \circ v_j||_2^2$ by 0. Hence the right-hand-side of (4) can be lower bounded as

$$\sum_{\lambda_i,\lambda_j \ge T} \lambda_i\lambda_j||v_i \circ v_j||_2^2 - \lambda_n^2 \dim(W).$$

We further lower bound this new sum as follows.

Claim 6.4.

$$\sum_{\lambda_i,\lambda_j \ge T} \lambda_i\lambda_j||v_i \circ v_j||_2^2 \ge \frac{S_T^2}{n}.$$

Proof. We have

$$\sum_{\lambda_i,\lambda_j > T} \lambda_i\lambda_j||v_i \circ v_j||_2^2 = \sum_{\lambda_i,\lambda_j \ge T} \lambda_i\lambda_j\langle v_i \circ v_i, v_j \circ v_j\rangle = \left\|\sum_{\lambda_i \ge T}\lambda_i v_i \circ v_i\right\|_2^2.$$

Here, the right-hand-side is

$$\sum_{k=1}^n\left(\sum_{\lambda_i \ge T}\lambda_i v_i(k)^2\right)^2 \ge \frac{1}{n}\left(\sum_{\lambda_i \ge T}\sum_{k=1}^n \lambda_i v_i(k)^2\right)^2 = \frac{S_T^2}{n},$$

where the first inequality is due to the inequality between the quadratic and arithmetic mean. □

Using this claim, we thus proved that

$$\mathrm{tr}_W(A \circ A) \ge \frac{S_T^2}{n} - \lambda_n^2 \dim(W).$$

Recalling that $\dim(W) \le \frac{S_T^2}{T^2}$, and that the conditions of the lemma imply $\lambda_n^2 \le \frac{T^2}{16n}$, we can further write

$$\frac{S_T^2}{n} - \lambda_n^2 \dim(W) \ge \frac{S_T^2}{n} - \frac{S_T^2}{4n} = \frac{3S_T^2}{4n}.$$

In conclusion, we proved that

$$S_{\frac{T^2}{2n}} + \frac{S_T^2}{2n} \ge \mathrm{tr}_W(A) = \mathrm{tr}_W(A \circ A) \ge \frac{3S_T^2}{4n}.$$

From this, the desired inequality follows. □

6.2 Main lemma - MaxCut version

In this section, we present a variant of the previous lemma for graphs with small MaxCut. In order to prove this lemma, we employ a similar strategy as in the proof of Lemma 6.3. However, instead of writing $A = (A + |\lambda_n| I)^{\circ 2} - |\lambda_n| I$, we write $A = (A+E) \circ (A+E) - 2A \circ E - E \circ E$, where E is the contribution of the negative eigenvalues. Then, most of the proof comes down to showing that if the surplus is small, then $\mathrm{tr}_W(A \circ E)$ and $\mathrm{tr}_W(E \circ E)$ are also small for an appropriately chosen W.

Lemma 6.5. *Let G be an n-vertex graph such that $\mathrm{surp}^*(G) \leq n^{1+s}$. Let $T \geq Cn^{1-\frac{1}{24}+\frac{s}{4}}$ for some sufficiently large absolute constant $C > 0$. Then*

$$4nS_{\frac{T^2}{2n}} \geq S_T^2.$$

Proof. Let $Q = \mathrm{surp}^*(G)$. We again analyze the identity

$$A = A \circ A.$$

Let $v_1, \ldots, v_n$ be an orthonormal basis of eigenvectors corresponding to the eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$. Set E to be the "negative part" of A, that is,

$$E := \sum_{\lambda_i < 0} |\lambda_i| v_i v_i^T.$$

Then we can rewrite the previous equality as

$$A = (A+E) \circ (A+E) - 2A \circ E - E \circ E. \tag{5}$$

Using the spectral decomposition of A, this is equivalent to

$$\sum_{i=1}^n \lambda_i v_i v_i^T = \left(\sum_{\lambda_i > 0} \lambda_i v_i v_i^T\right) \circ \left(\sum_{\lambda_i > 0} \lambda_i v_i v_i^T\right) - 2A \circ E - E \circ E,$$

which can be further written as

$$\sum_{i=1}^n \lambda_i v_i v_i^T = \sum_{\lambda_i, \lambda_j > 0} \lambda_i \lambda_j (v_i \circ v_j)(v_i \circ v_j)^T - 2A \circ E - E \circ E.$$

We bound the W-traces of both sides with separate methods for an appropriate subspace W. The terms $2A \circ E$ and $E \circ E$ constitute as error terms, for which we wish to show that their contribution to the W-trace is not too large.

Let W_0 be the subspace generated by those vectors $v_i \circ v_j$, where λ_i and λ_j are both at least T. Formally,

$$W_0 = \langle v_i \circ v_j : \lambda_i, \lambda_j \geq T \rangle.$$

The subspace W_0 is almost what we want. However, when we try to bound the contribution of the error term $\mathrm{tr}_{W_0}(E \circ E)$, large entries of E may cause trouble. In order to overcome this, we introduce a cutoff

$$\beta := \frac{Q^{1/4} n^{7/8}}{T} > 1.$$

Let $J \subset [n]$ be the set of indices i such that $E_{i,i} > \beta$. Note that as E is positive semidefinite, we have $\max_{i,j} |E_{i,j}| = \max_{i,i} E_{i,i}$, so $|E_{i,j}| \leq \beta$ for every $i, j \in [n] \setminus J$. We can bound the size of J as follows.

Claim 6.6. $|J| \leq Q/\beta$.

Proof. By Lemma 5.3, (i), we have

$$\sum_{i=1}^{n} E_{i,i} = \mathrm{tr}(E) = \sum_{\lambda_i<0} |\lambda_i| \le Q.$$

From this, the claim follows immediately. □

Let $Y < \mathbb{R}^n$ be the subspace of vectors that vanish on J, that is,

$$Y := \{y \in \mathbb{R}^n : \forall i \in J, y(i) = 0\}.$$

Finally, define

$$W := \Pi_Y(W_0).$$

Note that

$$\dim(W) \le \dim(W_0) \le N_T^2 \le \frac{S_T^2}{T^2}.$$

Consider the trace of the W-compressions of both sides of (5). Let $K = \frac{T^2}{2n}$, then using Lemma 6.2 we can upper bound the left hand side as

$$\mathrm{tr}_W(A) \le S_K + K\dim(W) \le S_K + K\frac{S_T^2}{T^2} = S_{\frac{T^2}{2n}} + \frac{S_T^2}{2n}.$$

Next, we consider the first term of the right hand side of (5), and write

$$\mathrm{tr}_W((A+E)\circ(A+E)) = \sum_{\lambda_i,\lambda_j>0} \lambda_i\lambda_j ||\Pi_W v_i \circ v_j||_2^2 \ge \sum_{\lambda_i,\lambda_j\ge T} \lambda_i\lambda_j ||\Pi_W v_i \circ v_j||_2^2.$$

Here, by definition, we have $v_i \circ v_j \in W_0$, so $\Pi_W(v_i\circ v_j) = \Pi_Y(v_i \circ v_j)$. Thus,

$$||\Pi_W(v_i\circ v_j)||_2^2 = ||\Pi_Y(v_i\circ v_j)||_2^2 = ||v_i\circ v_j||_2^2 - \sum_{k\in J}(v_i(k)v_j(k))^2.$$

Using Lemma 5.4, the entries of v_i and v_j are bounded as $|v_i(k)|, |v_j(k)| \le \frac{\sqrt{n}}{T}$, so we get

$$||\Pi_W v_i\circ v_j||_2^2 = ||v_i\circ v_j||_2^2 - \sum_{k\in J}(v_i(k)v_j(k))^2 \ge ||v_i\circ v_j||_2^2 - \frac{|J|n^2}{T^4}.$$

With this bound, we get

$$\begin{aligned}\mathrm{tr}_W((A+E)\circ(A+E)) &\ge \sum_{\lambda_i,\lambda_j\ge T}\lambda_i\lambda_j\left(||v_i\circ v_j||_2^2 - \frac{|J|n^2}{T^4}\right)\\ &\ge \left(\sum_{\lambda_i,\lambda_j\ge T}\lambda_i\lambda_j||v_i\circ v_j||_2^2\right) - S_T^2\frac{|J|n^2}{T^4} \ge \frac{S_T^2}{n} - \frac{QS_T^2n^2}{\beta T^4}.\end{aligned}$$

Here, the second inequality follows by writing $\sum_{\lambda_i,\lambda_j\ge T}\lambda_i\lambda_j = S_T^2$, and the last inequality follows by Claim 6.4, and writing $|J| \le Q/\beta$.

Finally, we bound $\mathrm{tr}_W(E\circ A)$ and $\mathrm{tr}_W(E\circ E)$. First, we have

$$||E\circ A||_F \le ||E||_F = \left(\sum_{\lambda_i<0}\lambda_i^2\right)^{1/2} = O(n^{1/4}Q^{1/2}),$$

where the last equality follows by Lemma 5.3, (ii). Therefore, by Lemma 6.1, we have

$$\mathrm{tr}_W(A \circ E) = \dim(W)^{1/2} ||E \circ A||_F \leq O(\dim(W)^{1/2} n^{1/4} Q^{1/2}).$$

Now consider $\mathrm{tr}_W(E \circ E)$. Let $E' = E_Y$, then $E'_{i,j} = E_{i,j}$ if $i, j \notin J$, and $E'_{i,j} = 0$ otherwise. Therefore,

$$\begin{aligned} ||E \circ E||_F = ||E' \circ E'||_F &= \left(\sum_{i,j \notin J} E_{i,j}^4 \right)^{1/2} \\ &\leq \beta \left(\sum_{i,j \notin J} E_{i,j}^2 \right)^{1/2} \leq \beta ||E||_F = O(\beta n^{1/4} Q^{1/2}). \end{aligned}$$

From this, using Lemma 6.1,

$$\mathrm{tr}_W(E \circ E) \leq \dim(W)^{1/2} ||E \circ E||_F \leq O(\dim(W)^{1/2} \beta n^{1/4} Q^{1/2}).$$

Hence, the total contribution from the error terms can be bounded as

$$\begin{aligned} 2\,\mathrm{tr}_W(A \circ E) + \mathrm{tr}_W(E \circ E) &\leq O(\dim(W)^{1/2} \beta n^{1/4} Q^{1/2}) \\ &\leq O(\dim(W) \beta n^{1/4} Q^{1/2}) \leq O\left(\frac{S_T^2 \beta n^{1/4} Q^{1/2}}{T^2} \right). \end{aligned}$$

In the second inequality, we upper bounded $\dim(W)^{1/2}$ by $\dim(W)$, which is quite wasteful, but it simplifies upcoming calculations. Putting everything together, we proved that

$$\begin{aligned} \mathrm{tr}_W(A \circ A) &= \mathrm{tr}_W((A + E) \circ (A + E)) - 2\,\mathrm{tr}_W(A \circ E) - \mathrm{tr}_W(E \circ E) \\ &\geq \frac{S_T^2}{n} - \frac{Q S_T^2 n^2}{\beta T^4} - O\left(\frac{S_T^2 \beta n^{1/4} Q^{1/2}}{T^2} \right) = S_T^2 \left(\frac{1}{n} - \frac{Q n^2}{\beta T^4} - O\left(\frac{\beta n^{1/4} Q^{1/2}}{T^2} \right) \right). \end{aligned}$$

The parameter β was chosen such that the two negative terms have the same order of magnitude. After substituting $\beta = \frac{Q^{1/4} n^{7/8}}{T}$, we get

$$S_T^2 \left(\frac{1}{n} - O\left(\frac{Q^{3/4} n^{9/8}}{T^3} \right) \right) \geq \frac{3 S_T^2}{4}.$$

Here, the last inequality holds by our assumptions that $Q \leq n^{1+s}$ and $T \geq C n^{1 - \frac{1}{24} + \frac{s}{4}}$. In conclusion, comparing the left-hand-side and right-hand-side of (5), we get

$$S_{\frac{T^2}{2n}} + \frac{S_T^2}{2n} \geq \mathrm{tr}_W(A) = \mathrm{tr}_W(A \circ A) \geq \frac{3 S_T^2}{4n}.$$

From this, the desired inequality follows. □

6.3 Recursion

In this section, we show what information can be extracted from the recursive relationship between the sum of top eigenvalues, that is Lemmas 6.3 and 6.5. The proof is just a simple analysis on sequences.

Lemma 6.7. *Let n, k be integers, $0 < s < q$ real numbers, and let $\lambda_1 \geq \cdots \geq \lambda_k \geq 0$. For every real number T, define $S_T = \sum_{\lambda_i \geq T} \lambda_i$. Assume that $\lambda_1 + \cdots + \lambda_k \leq n^{1+s}$, and for every $T \geq n^{1-q}$,*

$$S_T^2 \leq 4nS_{\frac{T^2}{2n}}.$$

Then for every $H \leq n$,

$$\sum_{\lambda_i \leq H} \lambda_i^2 \leq \frac{16q}{q-s} n^{1+s/q} H^{1-s/q}.$$

Proof. Let $T_0 = n^{1-2q}/2$, and recursively define $T_{i+1} = (2nT_{i-1})^{1/2}$ for $i = 0, 1, \ldots$. Observe that

$$\frac{2n}{T_{i+1}} = \left(\frac{2n}{T_i}\right)^{1/2}.$$

The sequence $T_0, T_1, \ldots$ is monotone increasing, and we have $T_1 = (2nT_0)^{1/2} \geq n^{1-q}$, so we get that

$$S_{T_{i+1}}^2 \leq 4nS_{T_i}$$

for all $i \geq 0$. We prove by induction on i that

$$S_{T_i} \leq 4n \left(\frac{2n}{T_{i+1}}\right)^{s/q}.$$

For the base case $i = 0$, this is true as

$$S_{T_0} \leq \sum_{i=1}^{k} \lambda_i \leq n^{1+s} = n\left(\frac{n}{n^{1-q}}\right)^{s/q} \leq 4n\left(\frac{2n}{T_1}\right)^{s/q}.$$

Now let $i > 0$, then

$$S_{T_{i+1}} \leq \sqrt{4nS_{T_i}} \leq \sqrt{16n^2\left(\frac{2n}{T_{i+1}}\right)^{s/q}} = 4n\left(\frac{2n}{T_{i+2}}\right)^{s/q}.$$

It is easy to see that $\lim_{i\to\infty} T_i = 2n$, so for every $T \in [T_0, n]$, there exists an index i such that $T_i \leq T < T_{i+1}$. As S_T is decreasing in T, we have

$$S_T \leq S_{T_i} \leq 4n\left(\frac{2n}{T_{i+1}}\right)^{s/q} \leq 4n\left(\frac{2n}{T}\right)^{s/q}.$$

This inequality also holds for $T \leq T_0 = n^{1-2p}/2$ as

$$S_T \leq S_{T_0} \leq n^{1+s} \leq 4n\left(\frac{2n}{T}\right)^{s/q}.$$

Hence, we can write

$$\sum_{\lambda_i \leq H} \lambda_i^2 \leq \int_0^H 2S_t dt \leq \int_0^H 8n\left(\frac{2n}{t}\right)^{s/q} dt = \frac{8 \cdot 2^{s/q}}{1 - s/q} n^{1+s/p} H^{1-s/q}.$$

In the last equality, we use that $s/q < 1$. Moreover, the first inequality follows as

$$\sum_{\lambda_i \leq H} \lambda_i^2 = \sum_{\lambda_i} \int_0^{\lambda_i} 2t dt = \int_0^H 2N_t t dt \leq \int_0^H 2S_t dt,$$

where $N_t = \#\{i : \lambda_i \geq t\}$. ∎

7 Densification — Phase 0

In this section, we prove that every graph of average degree d and smallest eigenvalue λ_n contains a subgraph on d vertices of density $\Omega(1/|\lambda_n|)$.

Lemma 7.1. *Let G be an n-vertex graph of average degree d with smallest eigenvalue λ_n satisfying $|\lambda_n| \leq \frac{1}{2}d^{1/2}$. Then G contains an induced subgraph on d vertices with edge density $\Omega(1/|\lambda_n|)$.*

Proof. Let x be a vertex with at least d neighbours, and let S be a set of exactly d neighbours of x. Define the vector $v \in \mathbb{R}^n$ such that $v(x) = 1$, $v(y) = \frac{\lambda_n}{d}$ if $y \in Y$, and $v(z) = 0$ if $z \notin S \cup \{x\}$. Then

$$v^T A v = 2\lambda_n + 2\frac{\lambda_n^2}{d^2} e(G[S]).$$

On the other hand

$$v^T v = 1 + \frac{\lambda_n^2}{d} \leq \frac{3}{2}.$$

Therefore, we obtain that

$$2\lambda_n + 2\frac{\lambda_n^2}{d^2} e(G[S]) = v^T A v \geq \lambda_n v^T v \geq \frac{3}{2}\lambda_n.$$

From this,

$$e(G[S]) \geq \frac{d^2}{4|\lambda_n|},$$

showing that $G[S]$ is the desired subgraph. $\square$

8 Densification — Phase 1

Lemma 6.7, combined with earlier results, has a number of powerful consequences. First, we use it to show that a graph with large smallest eigenvalue (or small surplus) contains a subgraph of positive density. We prove this via the following density increment argument. We apply Lemma 6.7 to show that the cubic sum of eigenvalues is large, which in turn coincides with six times the number of triangles. But if a graph has too many triangles, it means that there is a vertex, whose neighborhood is dense, so we pass to this neighborhood, and repeat. The main step of this argument is presented in the next lemma.

Lemma 8.1. *Let $s \in (0, 1/6)$, $C > 2$, and let G be a n-vertex graph with edge density $p > n^{-1/2}$, $\Delta(G) \leq Cpn$, and smallest eigenvalue λ_n satisfying $|\lambda_n| \leq n^s$. Then G has a subgraph on at least pn vertices of edge density at least $c_0 p^{6s}$ for some $c_0 = c_0(s, C) > 0$.*

Proof. Let $m = p\binom{n}{2}$ denote the number of edges. Let $\lambda_1 \geq \cdots \geq \lambda_n$ be the eigenvalues of G, then

$$\sum_{\lambda_i > 0} \lambda_i = \sum_{0 < \lambda_i} |\lambda_i| \leq n|\lambda_n| \leq n^{1+s}.$$

By Lemma 6.3, we also have

$$4nS_{\frac{T^2}{2n}} \geq S_T^2$$

for every $T \geq 2n^{1/2+s}$. So applying Lemma 6.7 with $q = 1/3 > s$ and the sequence of positive eigenvalues, we get that for every $H \leq n$,

$$\sum_{0 < \lambda_i \leq H} \lambda_i^2 \leq cn^{1+3s} H^{1-3s}$$

for some $c > 1$ depending only on s. Write $u := 3s < 1/2$, and set $H := (8c)^{-1/(1-u)} p^{1/(1-u)} n$, then $H \leq n$ and the right-hand-side equals $pn^2/8$. On the other hand, we have

$$\sum_{\lambda_i < 0} \lambda_i^2 \leq n\lambda_n^2 \leq n^{1+2s} < pn^2/8,$$

so

$$\sum_{\lambda_i \leq H} \lambda_i^2 \leq pn^2/4.$$

But $\sum_{i=1}^n \lambda_i^2 = ||A||_F^2 = 2m$, so we conclude that

$$\sum_{\lambda_i > H} \lambda_i^2 \geq 2m - \sum_{\lambda_i \leq H} \lambda_i^2 \geq pn^2/2.$$

Writing N for the number of triangles, we have

$$6N = \sum_{i=1}^n \lambda_i^3 \geq H \sum_{\lambda_i \geq H} \lambda_i^2 - \sum_{\lambda_i < 0} |\lambda_i|^3 \geq Hpn^2/2 - n|\lambda_n|^3 \geq c' p^{\frac{2-u}{1-u}} n^3 - n^{1+3s}$$

for some constant $c' > 0$ depending only on s. Here, using that $s < 1/6$, $u < 1/2$ and $p > n^{-1/2}$, we have $n^{1+3s} < \frac{c'}{2} p^{\frac{2-u}{1-u}} n^3$ for sufficiently large n. Hence, we get

$$N \geq \frac{c'}{12} p^{\frac{2-u}{1-u}} n^3.$$

Counting triangles by vertices, we observe that there is a vertex $v \in V(G)$ whose neighborhood X contains at least $\frac{3N}{n}$ edges. Here, note that $|X| \leq \Delta(G) \leq Cpn$. Let X' be an arbitrary $\min\{Cpn, n\}$ element subset of $V(G)$ containing X, then the edge density of $G[X']$ is

$$\frac{3N/n}{\binom{|X'|}{2}} \geq \frac{c' p^{\frac{2-u}{1-u}} n^2}{2C^2 p^2 n^2} = \frac{c'}{2C^2} p^{\frac{u}{1-u}}.$$

As $u/(1-u) = 3s/(1-3s) < 6s$, this finishes the proof. □

In the next lemma, we show how to handle the case when G has some vertices of too large degrees.

Lemma 8.2. *Let $C > 2$, let G be an n-vertex graph of average degree d such that* $\mathrm{surp}(G) \leq 0.01dn$. *Then either*

(i) G contains a subgraph on at least $n/2$ vertices with average degree at least $d/4$, and maximum degree at most Cd.

(ii) G contains a subgraph on at least n/C vertices of average degree at least $0.2Cd$.

Proof. Let $X \subset V(G)$ be the set of vertices of degree more than Cd, then $|X| \leq n/C$. Let $Y = V(G) \setminus X$. The maximum degree of $G[Y]$ is at most Cd, so if $G[Y]$ has average degree at least $d/4$, then (i) holds. Otherwise, there are at most $nd/8$ edges in $G[X]$. As $\mathrm{surp}(G) \leq 0.1dn$, we have $e(G[X,Y]) \leq \frac{dn}{4} + \mathrm{surp}(G) \leq 0.26dn$. Hence, we must have that $G[Y]$ has at least $dn/2 - dn/8 - 0.26dn > 0.1dn$ edges. Let X_0 be an arbitrary n/C sized subset of $V(G)$ containing X, then X_0 has average degree at least $0.2Cd$, so $G[X_0]$ satisfies the requirements. □

Lemma 8.3. *Let $s \in (0, 1/6)$, $\varepsilon \in (0, 1/2)$, then there exists $\rho = \rho(s, \varepsilon) > 0$ and $c_1 = c_1(s, \varepsilon) > 0$ such that the following holds for every $n > n_0(s, \varepsilon)$. Let G be an n-vertex graph with edge density p and smallest eigenvalue λ_n such that $p > n^{-\rho}$ and $|\lambda_n| \leq n^s$. Then G has a subgraph on at least $n^{1-\varepsilon}$ vertices with edge density at least c_1.*

We may choose $\rho = \alpha\varepsilon$ with some decreasing function $\alpha = \alpha(s) > 0$.

Proof. Let $\alpha = \frac{1}{3} - 2s > 0$, then we show that $\rho = \alpha\varepsilon$ suffices. Set $C := 5$, $s' = \frac{s}{1-\rho}$, then $s' < 1/6$ and $6s' < 1 - 2\alpha$. Let $c_0 = c_0(s', 8C)$ be the constant guaranteed by Lemma 8.1. Then there exists some $c_1 = c_1(s', \varepsilon) > 0$ such that for every $q \in (0, c_1)$, we have $c_0 q^{6s'} > q^{1-2\alpha}$.

We define the sequence of progressively denser induced subgraphs $G = G_0 \supset G_1 \supset \dots$ as follows. Assume that G_k is already defined on n_k vertices with edge density p_k. If $p_k > c_1$ or $n_k < n^{1-\varepsilon}$, we stop, otherwise we define G_{k+1} as follows. By Lemma 8.2 (noting that the condition $\mathrm{surp}(G) \leq |\lambda_n| n < 0.01pn^2$ is trivially satisfied), either (i) G_k contains a subgraph G' on at least $n_k/2$ vertices with average degree at least $p_k n_k/4$ and $\Delta(G') \leq Cp_k n_k$, or (ii) G contains a subgraph G'' on at least n_k/C vertices of average degree at least $0.2Cp_k n_k$.

First, consider the case if outcome (ii) happens. Then set $G_{k+1} := G''$, and note the $n_{k+1} \geq n_k/C$ and $p_{k+1} \geq 0.2C^2 p_k > Cp_k$.

Now consider the case if outcome (i) happens. Let $n' = v(G')$ and let p' be the edge density of G', then $n' \geq n_k/2$, $p' \geq p_k/4$ and $\Delta(G') \leq Cp_k n_k \leq 8Cp'n'$. As $n_k \geq n^{1-\varepsilon}$, we have $p' \geq p/4 \geq n^{-\rho}/4 \geq (n')^{-\rho/(1-\varepsilon)}/8 > (n')^{-1/2}$. Also, if λ' is the smallest eigenvalue of G', then

$$|\lambda'| \leq |\lambda_n| \leq n^s \leq (n'/2)^{s/(1-\rho)} \leq (n')^{s'}.$$

Hence, we can apply Lemma 8.1 to conclude that G' contains a subgraph on at least $p'n'$ vertices of edge density at least $c_0(p')^{6s'}$. Let G_{k+1} be this subgraph. Then $n_{k+1} \geq p'n' \geq p_k n_k/8$ and

$$p_{k+1} \geq c_0(p')^{6s'} \geq (p_k/8)^{1-2\alpha}.$$

Here, the last inequality holds by our assumption that $p_k \leq c_1$.

We can unite outcomes (i) and (ii) by observing that in both cases there is some $q_k > 1$ such that $n_{k+1} > n_k/q_k$ and $p_{k+1} > p_k q_k^{2\alpha}$. Indeed, in case (i), we can choose $q_k = C$, and in case (ii), we choose $q_k = p_k$. From this, it is easy to see that the sequence $n_k^{2\alpha} p_k$ is monotone increasing, so we have $n_k^{2\alpha} p_k \geq n^{2\alpha} p$. Hence, as long as $p_k \leq c_1$, we have $n_k \geq n(p/c_1)^{1/2\alpha} > n^{1-\rho/2\alpha}/c_1^{1/2\alpha} = n^{1-\varepsilon/2}/c_1^{1/2\alpha} > n^{1-\varepsilon}$. Let K be the last index k for which G_k is defined, so either $p_K > c_1$ or $n_K < n^{1-\varepsilon}$. Then the previous argument shows that we must have $p_K > c_1$ and $n_K > n^{1-\varepsilon}$, so the graph G_K suffices. $\square$

Next, we prove a version of the previous lemmas for small surplus. As the proofs are more or less the same, with only some parameters changed, we only highlight the key differences.

Lemma 8.4. *Let $s \in (0, 1/60)$, $C > 2$, and let G be a n-vertex graph with edge density $p > n^{-1/3}$, $\Delta(G) \leq Cpn$, and $\mathrm{surp}^*(G) \leq n^{1+s}$. Then G has a subgraph on at least pn vertices of edge density at least $c_0 p^{4/5}$ for some $c_0 = c_0(s, C) > 0$.*

Proof. By Lemma 5.3, (i), we have

$$\sum_{\lambda_i > 0} \lambda_i = \sum_{0 < \lambda_i} |\lambda_i| \leq \mathrm{surp}^*(G) \leq n^{1+s}.$$

By Lemma 6.5, we also have

$$4nS_{\frac{T^2}{2n}} \geq S_T^2$$

for every $T > C_0 n^{1-\frac{1}{24}+\frac{s}{4}}$. So we can apply Lemma 6.7 with $q = 3/80 > s$ and the sequence of positive eigenvalues to get that for every $H \leq n$,

$$\sum_{0<\lambda_i \leq H} \lambda_i \leq c n^{1+s/q} H^{1-s/q}.$$

Furthermore, by Lemma 5.3, (ii), we have

$$\sum_{\lambda_i<0} \lambda_i^2 \leq O(n^{1/2}\operatorname{surp}^*(G)) = O(n^{3/2+s}).$$

Write $u := s/q < 4/9$, then setting $H := (8c)^{-1/(1-u)} p^{1/(1-u)} n$, we get $\sum_{\lambda_i<H} \lambda_i^2 \leq pn^2/4$. Thus, if N is the number of triangles, then

$$6N = \sum_{i=1}^{n} \lambda_i^3 \geq H \sum_{\lambda_i \geq H} \lambda_i^2 - \sum_{\lambda_i<0} |\lambda_i|^3 \geq Hpn^2/2 - O(n \operatorname{surp}^*(G)) \geq c' p^{\frac{2-u}{1-u}} n^3 - n^{2+s}.$$

Here, in the second inequality, we used Lemma 5.3, (iii). The rest of the proof is identical to the proof of Lemma 8.1. □

Lemma 8.5. *Let $s \in (0, 1/60)$, $\varepsilon \in (0, 1/2)$, then there exists $\rho = \rho(s, \varepsilon) > 0$ and $c_1 = c_1(s, \varepsilon) > 0$ such that the following holds for every $n > n_0(s, \varepsilon)$. Let G be an n-vertex graph with edge density p such that $\operatorname{surp}^*(G) \leq n^{1+s}$ and $p > n^{-\rho}$. Then G has a subgraph on at least $n^{1-\varepsilon}$ vertices with edge density at least c_1.*

Proof. The proof of this is almost identical to the proof of Lemma 8.3, but we use Lemma 8.4 instead of Lemma 8.1. We omit further details. □

9 Densification — Phase 2

In this section, we prove that graphs of positive constant density and large smallest eigenvalue (or small surplus) are $o(1)$-close to the disjoint union of cliques. In particular, this implies that such graphs must contain subgraphs of density $1 - o(1)$. Beyond the use of the main lemmas, Lemma 6.3 and 6.5, another key component is showing that if the adjacency matrix of a graph is close to a small rank matrix (in Frobenius norm), then G admits an *ultra-strong regularity partition*. These kind of partitions are closely related to Szemerédi's regularity lemma, but they provide substantially stronger quantitative bounds. Ultra-strong regularity lemmas first appeared in relation to graphs of bounded VC-dimension, see the seminal work of Lovász and Szegedy [60].

A partition $V_1, \dots, V_K$ of a set of size n is an *equipartiton*, if $|V_i| \in \{\lfloor n/K \rfloor, \lceil n/K \rceil\}$ for $i \in [K]$. Given a graph G, $\delta \in (0, 1)$, and two disjoint sets $X, Y \subset V(G)$, the pair (X, Y) is δ-*empty* if there are at most $\delta|X||Y|$ edges between X and Y. Also, (X, Y) is δ-*full* if there are at least $(1 - \delta)|X||Y|$ edges between X and Y. Then (X, Y) is δ-*homogeneous* if it is either δ-empty or δ-full.

A δ-*regular partition* of G is an equipartition $V_1, \dots, V_K$ of the vertex set such that all but at most δK^2 of the pairs (V_i, V_j) for $1 \leq i < j \leq K$ are δ-homogeneous.

Lemma 9.1. *For every $\delta > 0$, there exists $\varepsilon > 0$ such that the following holds for every positive integer r, and every n that is sufficiently large with respect to δ, r. Let G be a graph with adjacency matrix A. Assume that there exists an $n \times n$ symmetric matrix B of rank r such that $||A - B||_F^2 \leq \varepsilon n^2$. Then G has a δ-regular partition into K parts, where $1/\delta < K < O_{r,\delta}(1)$.*

Proof. We show that $\varepsilon = \delta^2/100$ suffices. Let $B = \sum_{i=1}^r \mu_i w_i w_i^T$ be the spectral decomposition of B. Then

$$\left(\sum_{i=1}^r \mu_i^2\right)^{1/2} = ||B||_F \leq ||A||_F + ||B - A||_F < 2n,$$

which shows that $|\mu_i| \leq 2n$ for all $i \in [r]$. Next, we group the coordinates of the vectors $w_1, \dots, w_r$ with respect to how close they are, which then we use to form a partition of B into submatrices that are close to constant matrices.

Let $\beta := 10^{-3}\delta^{1/2}r^{-3/2}$, and for $\ell \in \mathbb{Z}$, let

$$X_{i,\ell} = \left\{ j \in [n] : \frac{\beta}{\sqrt{n}}\ell \leq w_i(j) < \frac{\beta}{\sqrt{n}}(\ell+1) \right\}.$$

That is, for fixed $i \in [n]$, the sets $X_{i,\ell}$ form a partition of the coordinates of w_i into chunks that are close to constant. Next, we show that most coordinates of w_i are covered by $O_{r,\delta}(1)$ of these sets. Set $h := 10^4 r^2/\delta$. As $\sum_{j=1}^n w_i(j)^2 = 1$, the number of $j \in [n]$ not contained in $\bigcup_{\ell=-h}^h X_{i,\ell}$ is at most $n/(h^2\beta^2) < \frac{\delta n}{8r}$.

Let $I = \{-h, \dots, h\}^r$. For every $\overline{\ell} \in I$, let $X_{\overline{\ell}} = \bigcap_{i \in [r]} X_{i,\overline{\ell}(i)}$. Then

$$\bigcup_{\overline{\ell} \in I} X_{\overline{\ell}} \geq n\left(1 - \frac{\delta}{8r}\right)^r \geq n(1 - \delta/8). \tag{6}$$

Thus, the sets $X_{\overline{\ell}}$ form a disjoint covering of all but at most $\delta n/8$ of the indices. Next, our goal is to show that if $\overline{\ell}_1, \overline{\ell}_2 \in I$, then the submatrix of B induced on $X_{\overline{\ell}_1} \times X_{\overline{\ell}_2}$ is close to a constant matrix. We refer to the rectangles $X_{\overline{\ell}_1} \times X_{\overline{\ell}_2}$ as *blocks*. Let

$$\gamma = \gamma_{\overline{\ell}_1, \overline{\ell}_2} = \sum_{i=1}^r \mu_i \cdot \frac{\beta^2}{n} \cdot \overline{\ell}_1(i)\overline{\ell}_2(i).$$

Note that for every $(j_1, j_2) \in X_{\overline{\ell}_1} \times X_{\overline{\ell}_2}$, we have

$$\left| w_i(j_1)w_i(j_2) - \frac{\beta^2}{n}\overline{\ell_1}(i)\overline{\ell_2}(i) \right| \leq 4h,$$

which we get from the general inequality $|ab - cd| \leq |a||b-d| + |d||a-c|$. Therefore, we have

$$|B_{j_1,j_2} - \gamma| \leq \sum_{i=1}^r |\mu_i| \cdot \left| w_i(j_1)w_i(j_2) - \frac{\beta^2}{n}\overline{\ell_1}(i)\overline{\ell_2}(i) \right| \leq \sum_{i=1}^r |\mu_i| \cdot \frac{\beta^2}{n} \cdot 4h \leq 8r\beta^2 h < \frac{1}{3}. \tag{7}$$

Observe that if $X \subset X_{\overline{\ell}_1}$ and $Y \subset X_{\overline{\ell}_2}$ are such that (X, Y) is not δ-homogeneous, then

$$||A[X \times Y] - B[X \times Y]||_F^2 \geq \frac{\delta}{36}|X||Y|.$$

Indeed, if $\gamma_{\overline{\ell}_1, \overline{\ell}_2} < 1/2$, then $A_{j_1,j_2} - B_{j_1,j_2} \geq 1/6$ for every one entry A_{j_1,j_2}, otherwise $|A_{j_1,j_2} - B_{j_1,j_2}| \geq 1/6$ for every zero entry A_{j_1,j_2}.

Now let $K = |I|/(8\delta)$, and let $V_1, \dots, V_K$ be an equipartition of $V(G)$ as follows. Let V^* be the set of elements not covered by any of the $X_{\overline{\ell}}$ for $\overline{\ell} \in I$. For each $\overline{\ell} \in I$, choose as many of the V_i to be completely contained in $X_{\overline{\ell}}$ as possible, and then move the not covered elements of $X_{\overline{\ell}}$ to V^*. Then finally partition V^*. Each $X_{\overline{\ell}}$ contributes at most n/K elements to V^*, so in the end we have $|V^*| \leq \delta n/8 + |I| \cdot (n/K) \leq \delta n/4$. Therefore, at most $\delta K/4$ sets V_i are contained in V^*. We show that $V_1, \dots, V_K$ is a δ-regular partition.

Assume that (V_i, V_j) is not δ-homogeneous. There are at most $\delta K^2/2$ such pairs where either $V_i \subset V^*$ or $V_j \subset V^*$. On the other hand, if $V_i, V_j \not\in V^*$, then $||A[V_i \times V_j] - B[V_i \times V_j]||_F^2 \geq \frac{\delta}{36}|V_i||V_j|$. As $||A - B||_F^2 \leq \varepsilon n^2$, this means that the number of such pairs is at most $36\varepsilon/\delta K^2 \leq \delta K^2/2$. Hence, the total number of pairs that are not δ-homogeneous is at most δK^2, as desired. $\square$

An important feature of Lemma 9.1 that ε only depends on δ, and not on r. To continue from this point, we observe that if X, Y, Z are sets of size $\Omega(n)$ such (X, Y) and (Y, Z) are δ-full, then (X, Z) cannot be δ-empty, assuming $\text{surp}^*(G)$ is small.

Lemma 9.2. *Let G be a graph on n vertices. Let $X, Y, Z \subset V(G)$ be disjoint sets such that $|X| = |Y| = |Z|$ and (X, Y) and (Y, Z) are δ-full, but (Y, Z) is δ-empty. Then* $\text{surp}(G) \geq (1/4 - 3\delta)|X|$.

Proof. Let $G' = G[X \cup Y \cup Z]$, and consider the cut $(Y, X \cup Z)$ in G'. This cut has at least $|X|^2(2 - 2\delta)$ edges. On the other hand, $e(G') \leq \frac{3}{2}|X|^2 + 2|X|^2 + \delta|X|^2 \leq (\frac{7}{2} + \delta)|X|^2$. Therefore,

$$\text{surp}(G) \geq \text{surp}(G') = \text{mc}(G') - \frac{e(G')}{2} \geq |X|^2(2 - 2\delta) - \left(\frac{7}{4} + \frac{\delta}{2}\right)|X|^2 = |X|^2\left(\frac{1}{4} - \frac{5}{2}\delta\right).$$

$\square$

A graph is the disjoint union of cliques if and only if it does not contain an induced *cherry*, that is, the path of length 2. Therefore, by the *Induced graph removal lemma* [5], being close to the disjoint union of cliques is equivalent to having few cherries. For the special case of cherries, one does not need the full power of this lemma, and a simple proof of the following quantitatively stronger bound is given by Alon and Shapira [7].

Lemma 9.3. *Let G be an n-vertex graph containing at most εn^3 cherries. Then G is ε^c-close to the disjoint union of cliques for some absolute constant $c > 0$.*

Furthermore, if G is δ-close to the union of cliques, then G contains at most $3\delta n^3$ cherries.

Proof. The first part follows from Alon and Shapira [7], so we only prove the second part. Let $\tilde{G}$ be the disjoint union of cliques that is δ-close to G. Then each cherry of G contains at least one edge or non-edge from $\tilde{G}\Delta G$, so we are done. $\square$

Now we are ready to prove Theorem 1.5, which we restate here for convenience.

Theorem 9.4. *Let $s \in (0, 1/4)$ and $\delta > 0$, then the following holds for every sufficiently large n. Let G be an n-vertex graph such that $|\lambda_n| \leq n^s$. Then G is δ-close to the union of cliques.*

Proof. Let $\delta_0 > 0$ be specified later, depending only on δ. Let ε be the constant guaranteed by Lemma 9.1 with respect to δ_0. We have

$$\sum_{\lambda_i > 0} \lambda_i = \sum_{\lambda_i < 0} |\lambda_i| \leq |\lambda_n| n \leq n^{1+s}.$$

Also, by Lemma 6.3, we have $4nS_{\frac{T^2}{2n}} \geq S_T^2$ for every $T \geq n^{1/2+s} \geq 2|\lambda_n|\sqrt{n}$. Hence, we can apply Lemma 6.7 to the sequence of positive eigenvalues with $q := 1/4 > s$ to conclude that for every $H \leq n$, we have

$$\sum_{0 < \lambda_i < H} \lambda_i^2 \leq O(n^{1+4s} H^{1-4s}).$$

Furthermore, $\sum_{\lambda_i < 0} \lambda_i^2 \leq n|\lambda_n|^2 < n^{3/2}$, so

$$\sum_{\lambda_i < H} \lambda_i^2 \leq O_s(n^{1+4s} H^{1-4s}) + n^{3/2}.$$

Set $H = \varepsilon_0 n$, where $\varepsilon_0 > 0$ is specified later. Then $n^{1+4s}H^{1-4s} = n^2\varepsilon_0^{1-4s}$. Hence, we can choose ε_0 (depending only on s and ε) such that

$$\sum_{\lambda_i < H} \lambda_i^2 \leq \varepsilon n.$$

Let A be the adjacency matrix of G, then A has $pn(n-1) > pn^2/2$ one entries, and we note that $p/2 \geq 4\delta_0$ holds. Let

$$B = \sum_{\lambda_i \geq H} \lambda_i v_i v_i^T,$$

then

$$||A - B||_F^2 = \sum_{\lambda_i < H} \lambda_i^2 \leq \varepsilon n^2.$$

Furthermore, the rank of B is $r := N_H \leq \frac{n^2}{H^2} \leq \varepsilon_0^{-2}$, where in the first inequality we used that $n^2 \geq \sum_{\lambda_i \geq H} \lambda_i^2 \geq H^2 N_H$. Hence, we can apply Lemma 9.1 to conclude that there is a δ_0-regular partition $V_1, \dots, V_K$ for some K with $1/\delta_0 < K < O_{r,\delta_0}(1) = O_{s,\delta}(1)$.

In order to finish the proof, we count cherries. Let x, y, z be the vertices of a cherry with $xy, yz \in E(G), xz \notin E(G)$, and let $x \in V_i, y \in V_j, z \in V_k$. We put this cherry into one of the following categories:

(i) i, j, k are not all distinct,

(ii) (V_i, V_j) or (V_j, V_k) or (V_i, V_k) is not δ_0-homogeneous,

(iii) (V_i, V_j) or (V_j, V_k) is δ_0-empty,

(iv) (V_i, V_k) is δ_0-full.

By Lemma 9.2, we cannot have that (V_i, V_j) and (V_j, V_k) are δ_0-full, but (V_i, V_k) is δ_0-empty. Therefore, each cherry belongs to one of the four categories. We observe that the number of cherries belonging to each category is at most $O(\delta_0 n^3)$. Indeed, for (i), there are $O(K^2)$ choices for the set $\{i, j, k\}$, and then there are at most $(n/K)^3$ choices for x, y, z, so in total $O(K^2(n/K)^3) = O(n^3/K) = O(\delta_0 n^3)$. For (ii), we use the fact that there are at most $\delta_0 K^2$ non-δ_0-homogeneous pairs (V_i, V_j), so the number of choices for (V_i, V_j, V_k) is at most $O(\delta_0 K^3)$. For (iii) and (iv), we observe that if we fixed (V_i, V_j, V_k), then there are at most $\delta_0(n/K)^3$ choices for x, y, z. Indeed, if say (V_i, V_j) is δ_0-empty, the pair (x, y) can be chosen from only the $\delta_0(n/K)^2$ edges between V_i and V_j.

In conclusion, the number of cherries in G is $O(\delta_0 n^3)$. But then by Lemma 9.3, G is $O(\delta_0)^c$-close to a disjoint union of cliques for some absolute constant $c > 0$. We are done by setting $\delta_0 > 0$ sufficiently small with respect to δ.

□

The following immediate corollary of this lemma will be used later.

Corollary 9.5. *Let $s \in (0, 1/4)$, $p > 0$ and $\delta > 0$, then the following holds for every sufficiently large n. Let G be an n-vertex graph of edge density p such that $|\lambda_n| \leq n^s$. Then G contains a subgraph on at least $pn/2$ vertices of edge density at least $1 - \delta$.*

Proof. Let $\delta_0 = \delta p^2/2$. By Theorem 9.4, G is δ_0-close to some graph H that is the disjoint union of cliques. Let $C_1, \dots, C_k$ be the vertex sets of the cliques forming H, then

$$e(H) = \sum_{i=1}^{k} \binom{|C_i|}{2} \leq \sum_{i=1}^{k} |C_i|^2 \leq n \max_{i \in [k]} |C_i|.$$

As $e(H) \geq e(G) - \delta_0 n^2 \geq pn^2/2$, this shows that at least one of the C_i's has size at least $pn/2$. Without loss of generality, $|C_1| \geq pn/2$. But then $G[C_1]$ has at least $\binom{|C_1|}{2} - \delta_0 n^2$ edges, so $G[C_1]$ has edge density at least $1 - 2\delta_0/p^2 = 1 - \delta$. □

Next, we present the MaxCut version of the previous lemma, whose proof is almost identical. We only highlight the key differences.

Theorem 9.6. *Let $s \in (0, 1/30)$, $\delta > 0$, then the following holds for every sufficiently large n. Let G be an n-vertex graph such that $\text{surp}^*(G) \leq n^{1+s}$. Then G is δ-close to a disjoint union of cliques.*

Proof. One of the key differences compared to the proof of Theorem 9.4 is that we use Lemma 6.5 to have $4nS_{\frac{T^2}{2n}} \geq S_T^2$ satisfied for every $T \geq n^{1-\frac{1}{24}+\frac{s}{4}}$. Then setting $q = 1/30$, the condition $s \in (0, 1/30)$ ensures that $s < q$ and $1 - q \geq 1 - \frac{1}{24} + \frac{s}{4}$, so we can apply Lemma 6.7. Another difference is that we bound $\sum_{0<\lambda_i} \lambda_i^2$ using Lemma 5.3, (ii), which gives $\sum_{0<\lambda_i} \lambda_i^2 \leq O(\sqrt{n}\,\text{surp}^*(G)) \leq O(n^{3/2+s}) = o(n^2)$. □

Finally, we deduce the immediate corollary of this theorem about finding dense subgraphs. The proof of this is identical to the proof of Corollary 9.5, so we omit it.

Corollary 9.7. *Let $s \in (0, 1/30)$, $p > 0$ and $\delta > 0$, then the following holds for every sufficiently large n. Let G be an n-vertex graph of edge density p such that $\text{surp}^*(G) \leq n^{1+s}$. Then G contains a subgraph on at least $pn/2$ vertices of edge density at least $1 - \delta$.*

10 Densification — Phase 3

In this section, we prove that graphs of edge density $1-\varepsilon$ and large smallest eigenvalue (or small surplus) must contain a subgraph of edge density $1 - n^{-\varepsilon'}$ on almost the same number of vertices. The proof of this proceeds via a density increment argument, as summarized in the following lemma.

Lemma 10.1. *Let $s \in (0, 1/3)$, and let n be sufficiently large with respect to s. Let G be an n-vertex graph with edge density $1 - p$, where $p < 10^{-5}$. Assume that $|\lambda_n| < n^s$. Then G contains an $n/2$-vertex induced subgraph of edge density at least $1 - 10^7 p^3 - O(n^{3s-1})$.*

Proof. Let A be the adjacency matrix of G with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$ and corresponding orthonormal basis of eigenvectors $v_1, \ldots, v_n$. Define $B = A - \lambda_1 v_1 v_1^T$. The key idea of the proof is to consider the following triple Hadamard product:

$$D = (B + |\lambda_n| I)^{\circ 3} = B^{\circ 3} + 3|\lambda_n| B \circ B \circ I + 3|\lambda_n|^2 B \circ I + |\lambda_n|^3 I.$$

As $B + |\lambda_n| I = |\lambda_n| v_1 v_1^T + \sum_{i=2}^n (\lambda_i + |\lambda_n|) v_i v_i^T$, we have that $B + |\lambda_n| I$ is positive semidefinite. Therefore, D is also positive semidefinite by the Schur product theorem. Then, using the principal eigenvector v_1, we identify a set of well-behaved vertices U, and evaluate the product $0 \leq \mathbb{1}_U^T D \mathbb{1}_U$.

We now give the details. Let U be the set of vertices $i \in [n]$ that satisfy $v_1(i) \geq (1 - 8p)/\sqrt{n}$.

Claim 10.2. $|U| \geq n/2$.

Proof. By Lemma 5.5, we have $|v_1(i)| \leq (1 + 2p + 2/n)/\sqrt{n}$ for every $i \in [n]$, thus

$$\begin{aligned} 1 = \sum_{i=1}^n v_1(i)^2 &\leq |U| \frac{(1 + 2p + 2/n)^2}{n} + (n - |U|) \frac{(1 - 8p)^2}{n} \\ &\leq \frac{1}{n} \left(|U|(1 + 8p) + (n - |U|)(1 - 8p) \right) = (1 - 8p) + \frac{16p|U|}{n}. \end{aligned}$$

From this, we get $|U| \geq n/2$. □

We now evaluate the terms of $\mathbb{1}_U^T D \mathbb{1}_U$, starting with the main term $\mathbb{1}_U^T B^{\circ 3} \mathbb{1}_U$.

Claim 10.3. $\mathbb{1}_U^T B^{\circ 3} \mathbb{1}_U \leq 10^6 |U|^2 p^3 - e(\overline{G}[U])/4$.

Proof. We have

$$B_{i,j} = \begin{cases} 1 - \lambda_1 v_1(i) v_1(j) & \text{if } ij \in E(G), \\ -\lambda_1 v_1(i) v_1(j) & \text{if } ij \notin E(G). \end{cases}$$

If $i, j \in I$, then

$$1 - \lambda_1 v_1(i) v_1(j) \leq 1 - (1-p)(n-1) \cdot \left(\frac{1-8p}{\sqrt{n}} \right)^2 \leq 100p,$$

and thus $\lambda_1 v_1(i) v_1(j) > \frac{1}{2}$. Therefore,

$$\mathbb{1}_U^T B^{\circ 3} \mathbb{1}_U = \sum_{i,j \in I, i \sim j} (1 - \lambda_1 v_1(i) v_1(j))^3 - \sum_{i,j \in U, i \not\sim j} (\lambda_1 v_1(i) v_1(j))^3 \leq 100^3 |U|^2 p^3 - 2e(\overline{G}[U])/8.$$

□

Claim 10.4. *We have*

$$\mathbb{1}_U^T (3|\lambda_n| B \circ B \circ I + 3|\lambda_n| B \circ I + |\lambda_n|^3 I) \mathbb{1}_U \leq O(n^{1+3s}).$$

Proof. We have $B_{i,i} = -\lambda_1 v_1(i)^2$ for every $i \in [n]$. By Lemma 5.5, we have $|B_{i,i}| \leq 2$. Then the claim follows immediately as the expression above evaluates to

$$\sum_{i \in U} (3|\lambda_n| B_{i,i}^2 + 3|\lambda_n|^2 B_{i,i} + |\lambda_n|^3) \leq O(n|\lambda_n|^3).$$

□

In conclusion, we proved that

$$0 \leq \mathbb{1}_U^T D \mathbb{1}_U \leq 10^6 |U|^2 p^3 - e(\overline{G}[U])/4 + O(n^{1+3s}).$$

Hence, we must have $e(\overline{G}[U]) \leq 10^7 |U|^2 p^3 / 2 + O(n^{1+3s})$, which shows that the edge density of $G[U]$ is at least $1 - 10^7 p^3 - O(n^{3s-1})$. □

Lemma 10.5. *Let $0 < s < 1/3$, and let n be sufficiently large with respect to s. Let G be an n-vertex graph with edge density at least $1 - 10^{-5}$. If $|\lambda_n| < n^s$, then G contains an induced subgraph of size $n^{1-o(1)}$ with edge density at least $1 - n^{3s-1+o(1)}$.*

Proof. Let s_0 be such that $s < s_0 < 1/3$, and let $1-p$ be the edge density of G, $p < 10^{-5}$. Let $G_0 = G$, and define the sequence of induced subgraphs $G_0 \supset G_1 \supset \ldots$ with increasing density as follows. If G_i is already defined with n_i vertices and edge density $1 - p_i$, then stop if either $p_i < n_i^{3s_0-1}$, or $n_i < n^{s/s_0}$. Otherwise, as $n_i > n^{s/s_0}$ and the smallest eigenvalue of G_i is at least $\lambda_n > -n^s > -n_i^{s_0}$, we can apply Lemma 10.1 to find an induced subgraph G_{i+1} of G_i on at least $n_i/2$ vertices of edge density at least $1 - 10^7 p_i^3 - O(n_i^{3s_0-1}) \geq 1 - \max\{10^8 p_i^3, O(n_i^{3s_0-1})\}$.

Let L be the last index i for which G_i is defined. Then $n_L \geq n 2^{-L}$. On the other hand $p_{L-1} > n_{L-1}^{3s_0-1}$, and so for every $i \leq L-2$, we have $p_{i+1} \leq 10^8 p_i^3$. This show that

$$p_{L-1} < (10^8)^{3^{L-2}+3^{L-3}+\cdots+1} p^{3^{L-1}} < (10^4 p)^{3^{L-1}} < 10^{-3^{L-1}}.$$

Hence, we must have stopped because $p_L < n_i^{3s_0-1}$, which happens for some $L = O(\log\log n)$. But then $n_L \geq n 4^{-O(\log\log n)} = n^{1-o(1)}$. We thus get that G_L is a subgraph of G of size $n^{1-o(1)}$ with edge density at least $1 - n^{3s_0-1+o(1)}$. As we may choose s_0 arbitrarily close to s, this finishes the proof. □

Now we establish the MaxCut version of Lemma 10.1. While the core idea remains the same, the proof becomes more involved. Therefore, we give a detailed proof.

Lemma 10.6. *Let $0 < s < 1/4$ and $0 < \alpha < \min\{\frac{1}{12} - \frac{s}{6}, \frac{1}{6} - \frac{2s}{3}\}$, and let n be sufficiently large with respect to s, α. Let G be an n-vertex graph with edge density $1 - p$, where $n^{-\alpha} \leq p < 10^{-5}$. Assume that* $\mathrm{surp}^*(G) < n^{1+s}$. *Then G contains an $n/4$-vertex induced subgraph of edge density at least $1 - 10^8 p^3$.*

Proof. Let A be the adjacency matrix of G with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$ and corresponding orthonormal basis of eigenvectors $v_1, \ldots, v_n$. Define $B = A - \lambda_1 v_1 v_1^T$ and $E = \sum_{\lambda_i < 0} |\lambda_i| v_i v_i^T$. Then, the matrices E and $B + E = \sum_{\lambda_i > 0, i \neq 1} \lambda_i v_i v_i^T$ are positive semidefinite.

The key idea of the proof is to consider the following triple Hadamard product:

$$D = (B + E)^{\circ 3} = B^{\circ 3} + 3B \circ B \circ E + 3B \circ E \circ E + E^{\circ 3}.$$

As $B + E$ is positive semidefinite, so is D by the Schur product theorem. Then, using the matrix E and the principal eigenvector v_1, we identify a set of well-behaved vertices U, and evaluate the product

$$0 \leq \mathbb{1}_U^T D \mathbb{1}_U = \mathbb{1}_U^T B^{\circ 3} \mathbb{1}_U + \mathbb{1}_U^T \big(3B \circ B \circ E + 3B \circ E \circ E + E^{\circ 3}\big) \mathbb{1}_U.$$

By carefully analyzing the terms of this product, we will conclude that the graph $\overline{G}[U]$ must be *much* sparser than G.

We now give the details. First, observe that

$$\mathrm{tr}(E) = \sum_{\lambda_i < 0} |\lambda_i| \leq \mathrm{surp}^*(G) \leq n^{1+s}$$

by *(i)* in Lemma 5.3, and

$$\|E\|_F^2 = \sum_{\lambda_i < 0} \lambda_i^2 \leq O(\sqrt{\Delta + 1})\, \mathrm{surp}^*(G) \leq O(\sqrt{\Delta + 1}) n^{1+s} \leq O(n^{3/2+s})$$

by *(ii)* in Lemma 5.3. Let U be the set of vertices $i \in [n]$ that satisfy $E_{i,i} \leq 4n^s$ and $v_1(i) \geq (1 - 8p)/\sqrt{n}$.

Claim 10.7. $|U| \geq n/4$.

Proof. Let U_0 be the set of vertices $i \in [n]$ such that $v_1(i) \geq (1 - 8p)/\sqrt{n}$. Then using Lemma 5.5,

$$\begin{aligned} 1 = \sum_{i=1}^n v_1(i)^2 &\leq |U_0| \frac{(1 + 2p + 2/n)^2}{n} + (n - |U_0|) \frac{(1 - 8p)^2}{n} \\ &\leq \frac{1}{n} \left(|U_0|(1 + 8p) + (n - |U_0|)(1 - 8p)\right) = (1 - 8p) + \frac{16p|U_0|}{n}. \end{aligned}$$

From this, we get $|U_0| \geq n/2$. Now U is those set of vertices $i \in U_0$ that satisfy $E_{i,i} \leq 4n^s$. As $\mathrm{tr}(E_{i,i}) \leq n^{1+s}$, the number of vertices such that $E_{i,i} > 4n^s$ is at most $n/4$, giving the desired bound $|U| \geq n/4$. □

We now evaluate the terms of $\mathbb{1}_U^T D \mathbb{1}_U$. For the main term$\mathbb{1}_U^T B^{\circ 3} \mathbb{1}_U$, we already proved in Claim 10.3 that

$$\mathbb{1}_U^T B^{\circ 3} \mathbb{1}_U \leq 10^6 |U|^2 p^3 - e(\overline{G}[U])/4.$$

So we consider the rest of the terms.

Claim 10.8. *We have*

$$\mathbb{1}_U^T (3B \circ B \circ E + 3B \circ E \circ E + E^{\circ 3}) \mathbb{1}_U \leq O(n^{7/4+s/2} + n^{3/2+2s}).$$

Proof. We bound each summand of the error term independently. Firstly, note that every entry of B is between 1 and -2, as $0 \le \lambda_1 v_1(i) v_2(j) \le n\left(\frac{1+2p+2/n}{\sqrt{n}}\right)^2 \le 2$. Therefore, we have

$$\mathbb{1}_U^T(3B \circ B \circ E)\mathbb{1}_U \le 12 \sum_{i,j\in U} |E_{i,j}| \le 12|U|\sqrt{\sum_{i,j\in U} E_{i,j}^2} \le 12n\|E\|_F \le O(n^{7/4+s/2}).$$

Here, the second inequality holds by the inequality between the arithmetic and quadratic mean.

To bound the second summand, we again use that entries of B are bounded by 2 in absolute value, and so

$$\mathbb{1}_U^T(3B \circ E \circ E)\mathbb{1}_U \le 6 \sum_{i,j\in U} E_{i,j}^2 \le 6\|E\|_F^2 \le O(n^{3/2+s}).$$

Finally, in bounding the last term we use that $E_{i,i} \le n^s$ for all $i \in U$. In particular, since E is a positive definite matrix, this implies that $|E_{i,j}| \le n^s$ for all $i, j \in U$. So,

$$\mathbb{1}_U^T E^{\circ 3}\mathbb{1}_U \le \sum_{i,j\in U} |E_{i,j}|^3 \le \max_{i,j\in U} |E_{i,j}| \cdot \sum_{i,j\in U} |E_{i,j}|^2 \le n^s \cdot \|E\|_F^2 \le O(n^{3/2+2s}).$$

The conclusion now follows by summing up the bounds obtained on each of the error terms above. $\square$

In conclusion, we proved that

$$0 \le \mathbb{1}_U^T D\mathbb{1}_U \le 10^6|U|^2p^3 - e(\overline{G}[U])/4 + O(n^{7/4+s/2} + n^{3/2+2s}) \le 10^7|U|^2p^3 - e(\overline{G}[U])/4,$$

where in the last inequality we used our bounds on p and s, and that n is sufficiently large. Hence, we must have $e(\overline{G}[U]) \le 10^8|U|^2p^3/2$, which shows that the edge density of $G[U]$ is at least $(1 - 10^8p^3)$. $\square$

Lemma 10.9. *Let $0 < s < 1/4$ and $0 < \alpha < \min\{\frac{1}{12} - \frac{s}{6}, \frac{1}{6} - \frac{2s}{3}\}$, and let n be sufficiently large with respect to s, α. Let G be an n-vertex graph with edge density at least $1 - 10^{-5}$. If $\mathrm{surp}^*(G) < n^{1+s}$, then G contains an induced subgraph of size $n^{1-o(1)}$ with edge density at least $1 - n^{-\alpha}$.*

Proof. The proof of this is almost identical to the proof of Lemma 10.5, but instead on using Lemma 10.1, we use Lemma 10.6. Then, we just need to adjust the parameters s and α slightly. We skip further details. $\square$

11 Densification — Phase 4

In this section, we show that graphs with large smallest eigenvalue (or small surplus) of edge density $1 - n^{-\varepsilon}$ contain subgraphs of edge density $1 - n^{-1+\varepsilon}$ on almost the same number of vertices.

In order to prove this, it is more convenient to work with the complement $\overline{G}$ of G. Let $\lambda_1 \ge \cdots \ge \lambda_n$ be the eigenvalues of G, and let $\mu_1 \ge \cdots \ge \mu_n$ be the eigenvalues of $\overline{G}$. Unfortunately, as G is not necessarily regular, there is no simple formula to express μ_i in terms of $\lambda_1, \dots, \lambda_n$. However, we can use Weyl's inequality to establish the following interlacing property.

Lemma 11.1. *Let G be an n vertex graph with eigenvalues $\lambda_1 \ge \cdots \ge \lambda_n$, and let $\mu_1 \ge \cdots \ge \mu_n$ be the eigenvalues of the complement of G. For each $i = 1, 2 \dots, n-1$, we have*

$$1 + \mu_{i+1} \le -\lambda_{n+1-i}.$$

Proof. Weyl's inequality states that if X and Y are $n \times n$ symmetric matrices, and $1 \le i, j \le n$ and $i + j \le n + 1$, then

$$\lambda_{i+j-1}(X + Y) \le \lambda_i(X) + \lambda_j(Y),$$

where $\lambda_1(X) \geq \cdots \geq \lambda_n(X)$ denote the eigenvalues of a matrix X. Let A be the adjacency matrix of G, and let B be the adjacency matrix of $\overline{G}$. Then $B = J - I - A$. Let $X = -A$ and $Y = J - I$, then $\lambda_i(X) = -\lambda_{n+1-i}$, $\lambda_1(Y) = n-1$, $\lambda_i(X) = -1$ for $i = 2, \dots, n$, and $\lambda_i(X+Y) = \mu_i$. Hence, applying the above inequality with $j = 2$, we get

$$\mu_{i+1} \leq -\lambda_{n+1-i} - 1.$$

□

The next lemma provides a bound on the surplus of very dense graphs. The result and its proof are similar to the proof of Lemma 5.9 of Räty, Sudakov and Tomon [64] for the complementary quantity called as *positive discrepancy*. Say that a graph G is C-*balanced* if $\Delta(G) \leq Cd(G)$.

Lemma 11.2. *Let G be an n-vertex graph of density $(1-p)$ such that the complement of G is C-balanced, and $p < 0.001C^{-2}$. Then*

$$\mathrm{surp}^*(G) \geq \Omega\left(\min\left\{\frac{n}{C^3 p}, C^{-1}p^{1/2}n^{3/2}\right\}\right).$$

Proof. Let A be the adjacency matrix of G with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$, and let B be the adjacency matrix of $\overline{G}$ with eigenvalues $\mu_1 \geq \cdots \geq \mu_n$.

Let $\bar{\Delta}$ be the maximum degree of $\overline{G}$, so $\mu_1 \leq \bar{\Delta} \leq Cpn$. We may assume that $p > 0$ and thus $\bar{\Delta} \geq 1$, otherwise the statement is trivial. For $k = 1, 2, 3$, set

$$P_k = \sum_{i \neq 1, \mu_i > 0} \mu_i^k \quad \text{and} \quad N_k = \sum_{\mu_i < 0} |\mu_i|^k.$$

Lemma 11.1 applied with $i \geq 2$ shows that whenever $\mu_i \geq 0$ we also have $\lambda_{n+1-i} \leq -\mu_i - 1 < 0$. Combined with Lemma 5.3, this shows that

$$\mathrm{surp}^*(G) \geq \sum_{\lambda_i < 0} |\lambda_i| \geq \sum_{i \neq 1, \mu_i > 0} \mu_i = P_1,$$

$$\mathrm{surp}^*(G) \geq \Omega\left(\bar{\Delta}^{-1/2} \sum_{\lambda_i < 0} |\lambda_i|^2\right) \geq \Omega\left(\bar{\Delta}^{-1/2} \sum_{i \neq 1, \mu_i > 0} \mu_i^2\right) = \Omega\left(\bar{\Delta}^{-1/2} P_2\right),$$

$$\mathrm{surp}^*(G) \geq \Omega\left(\frac{1}{\bar{\Delta}} \sum_{\lambda_i < 0} |\lambda_i|^3\right) \geq \Omega\left(\frac{1}{\bar{\Delta}} \sum_{i \neq 1, \mu_i > 0} \mu_i^3\right) = \Omega\left(\frac{1}{\bar{\Delta}} P_3\right).$$

We show that these three inequalities together with some simple identities suffice to prove the lemma.

First, assume that $N_2 \leq \frac{1}{8}pn^2$. Note that $\mu_1^2 + P_2 + N_2 = \|B\|_F^2$ is twice the number of edges of $\overline{G}$, so $\mu_1^2 + P_2 + N_2 = 2p\binom{n}{2}$, from which

$$P_2 \geq pn^2/2 - \mu_1^2 - N_2 \geq pn^2/2 - C^2p^2n^2 - pn^2/8 \geq pn^2/4,$$

where we have used that $pC^2 \leq 10^{-3}$ in the last inequality. But then $\mathrm{surp}^*(G) = \Omega(p^{1/2}n^{3/2})$ by the second highlighted inequality, and we are done.

Hence, in the rest of the proof, we may assume that $N_2 \geq \frac{1}{8}pn^2$. By the inequality between the quadratic and cubic mean, we have

$$\left(\frac{N_2}{n}\right)^{1/2} \leq \left(\frac{N_3}{n}\right)^{1/3}$$

which gives $N_3 \geq N_2^{3/2} n^{-1/2} \geq p^{3/2}n^{5/2}/64$.

Next, consider the quantity $T = N_3 - P_3$. Observe that $\mu_1^3 - T = \sum_{i=1}^n \mu_i^3$ is six-times the number of triangles of $\overline{G}$. In particular, $\mu_1^3 - T$ it is nonnegative, showing that $T \leq \mu_1^3 \leq \bar{\Delta}^3$. Assume that $N_3 \geq 2T$, then $P_3 \geq N_3/2$. By the third highlighted inequality, we then have

$$\mathrm{surp}^*(G) \geq \Omega\left(\frac{P_3}{\bar{\Delta}}\right) \geq \Omega\left(\frac{N_3}{\bar{\Delta}}\right) \geq \Omega(C^{-1}p^{1/2}n^{3/2}).$$

Hence, we are done in this case as well.

Finally, assume that $N_3 \leq 2T$, then $\bar{\Delta}^3 \geq T \geq N_3/2$. By the Cauchy-Schwartz inequality applied to the sequences $(|\mu_i|^3)_{\mu_i<0}$ and $(|\mu_i|)_{\mu_i<0}$, we have the inequality $N_1N_3 \geq N_2^2$, which gives

$$N_1 \geq \frac{N_2^2}{N_3} \geq \frac{p^2n^4}{128\bar{\Delta}^3} \geq \frac{n}{128C^3p}.$$

But $0 = \mathrm{tr}(B) = \mu_1 + P_1 - N_1$, from which $P_1 = N_1 - \mu_1 \geq \frac{n}{128C^3p} - \bar{\Delta} \geq \frac{n}{500C^3p}$. Hence, as $\mathrm{surp}^*(G) \geq P_1$, we are done. ☐

Next, we present a simple technical lemma which shows that every graph contains a large induced $O(\log n)$-balanced graph.

Lemma 11.3. *Let G be an n-vertex graph of edge density p, and let $C \geq 4\log_2 n$. Then G contains a C-balanced induced subgraph on at least $(1 - 2\log_2 n/C)n$ vertices of edge density at most p.*

Proof. Let $d = d(G)$ be the average degree of G. We perform an inductive process, where in each step we delete the vertices of high degree. More precisely, let $G_0 = G$, and define the sequence of induced subgraphs $G_0 \supset G_1 \supset \dots$ as follows. Having already defined G_i, we denote its number of vertices by n_i and its average degree by d_i. To define G_{i+1}, remove from G_i all vertices of degree at least $Cd_i/2$, if any such vertices exist. The process halts once either all vertices of G_i have degree less than $Cd_i/2$, or the average degree of G_i is at least $d_{i-1}/2$. Let I be the last index i for which G_i is defined.

For each i, it is not hard to verify that the density of G_i is smaller than the density of G_{i-1}. Moreover, since the average degree of G_i is d_i, there are at most $2n_i/C$ vertices of degree larger than $Cd_i/2$ in G_i, so $v(G_i) \geq v(G_{i-1}) - 2n_i/C \geq v(G_{i-1}) - 2n/C$. Thus, we have $v(G_i) \geq n(1 - 2i/C)$ for all $i = 1, \dots, I$. Furthermore, if the process did not halt at index i, we have $d_i \leq d_{i-1}/2$, and so $d_i \leq d2^{-i}$. This shows that $I \leq \log_2 n$ and $n_I \geq n(1 - 2\log_2 n/C)$.

Finally, note that G_I has no vertices of degree more than $Cd(G_I)$. Indeed, if the process has halted because G_I contains no vertices of $Cd_I/2$, this is immediate, and if the process has halted because $d_I \geq d_{I-1}/2$, then G_I contains no vertex of degree more than $Cd_{I-1}/2 \leq Cd_I$. Hence, G_I is a C-balanced induced subgraph of G on at least $(1 - 2\log_2/C)n$ vertices. ☐

Now we are ready to prove the main result of this section. Conveniently, we do not have separate versions for the minimum eigenvalue and the MaxCut, as our result for the latter is already optimal.

Lemma 11.4. *Let $0 < \varepsilon < \alpha$, then the following holds if n is sufficiently large. Let G be an n-vertex graph of edge density at least $1 - n^{-\alpha}$. If $\mathrm{surp}^*(G) \leq n^{1+\varepsilon}$, then G contains an induced subgraph of density at least $(1 - O((\log n)^2 n^{2\varepsilon-1}))$ on at least $n/2$ vertices.*

Proof. Let $1 - p$ the edge density of G, and let $C = 4\log_2 n$. Applying Lemma 11.3 to the complement of G, we find an induced subgraph $G_0 \subseteq G$ on at least $n/2$ vertices with edge density $1 - p_0 \geq 1 - p$ such that the complement of G_0 is C-balanced. As $p_0 \leq p \leq 0.001C^{-3}$, we can apply Lemma 11.2 to conclude that

$$\mathrm{surp}^*(G_0) = \Omega\left(\min\left\{\frac{n}{C^3p_I}, C^{-1}p_I^{1/2}n^{3/2}\right\}\right).$$

However, as $p_0 < n^{-\alpha}$ and $\varepsilon < \alpha$, the inequality $\mathrm{surp}^*(G_0) \leq n^{1+\varepsilon}$ is only possible if

$$p_0 \leq O(C^2n^{2\varepsilon-1}) = O(n^{2\varepsilon-1}(\log n)^2).$$

☐

12 Finding large cliques

In this section, we combine our densification steps to prove that graphs with large smallest eigenvalue (or small surplus) contain large cliques.

Theorem 12.1. *Let $s \in (0, 1/6)$ and $\delta > 0$, then there exist $\rho > 0$ such that the following holds for every sufficiently large n. Let G be an n-vertex graph of edge density at least $n^{-\rho}$ such that $|\lambda_n| \leq n^s$. Then G contains a clique of size at least $n^{1-2s-\delta}$.*

Furthermore, we may choose $\rho = \alpha\delta$ for some decreasing $\alpha = \alpha(s) > 0$.

Proof. Let $\varepsilon = \min\{\delta/20, (1/6 - s)/5\}$. Applying the phase 1 densification step, that is, Lemma 8.3, there exists $\rho = \rho(s, \varepsilon) > 0$ and $c_1 = c_1(s, \varepsilon)$ such that if G has edge density at least $n^{-\rho}$ and the absolute value of the smallest eigenvalue $|\lambda_n| \leq n^s$, then G contains a subgraph G_1 on at least $n_1 \geq n^{1-\varepsilon}$ vertices with edge density at least c_1. Note that G_1 satisfies that if its smallest eigenvalue is $\lambda^{(1)}$, then $|\lambda^{(1)}| \leq |\lambda_n| \leq n^{1+s} \leq n_1^{(1+s)/(1-\varepsilon)} \leq n_1^{1+s_1}$ for $s_1 = s + 2\varepsilon \in (0, 1/6)$. Furthermore, we may choose $\rho = \alpha_0\varepsilon$ for some $\alpha_0 = \alpha_0(s)$, which shows that we may choose $\alpha = \alpha(s) > 0$ such that $\rho = \alpha\delta$.

Next, applying the phase 2 densification to G_1, that is, Corollary 9.5, we get that G_1 contains a subgraph G_2 on $n_2 \geq c_1 n_1/2 \geq n^{1-2\varepsilon}$ vertices with edge density at least $1 - 10^{-5}$. The smallest eigenvalue $\lambda^{(2)}$ of G_2 still satisfies that $|\lambda^{(2)}| \leq n_2^{1+s_2}$ for $s_2 = s_1 + \varepsilon = s + 3\varepsilon \in (0, 1/6)$.

Apply the phase 3 densification step to G_2, that is, Lemma 10.5. Then G_2 contains a subgraph G_3 of edge density $1 - n_2^{3s_2-1+o(1)} > 1 - n_2^{-1/2}$ on $n_3 \geq n_2^{1-o(1)} > n^{1-3\varepsilon}$ vertices. The smallest eigenvalue $\lambda^{(3)}$ of G_3 satisfies that $|\lambda^{(3)}| \leq n_3^{1+s_3}$ for $s_3 = s_2 + \varepsilon = s + 4\varepsilon \in (0, 1/6)$.

Finally, apply the phase 4 densification step to G_3, that is, Lemma 11.4. As $\text{surp}^*(G_3) \leq |\lambda^{(3)}| n_3 \leq n_3^{1+s_3}$ and $s_3 < 1/6 < 1/2$, the lemma is indeed applicable. Therefore, we get a graph G_4 of density at least $1 - O((\log n_3)^2 n_3^{2s_3-1}) > 1 - n_3^{2s_3+\varepsilon-1}$ on at least $n_3/2 > n^{1-4\varepsilon}$ vertices. But then by Turán's theorem, G_4 contains a clique of size $n_3^{1-2s_3-\varepsilon}/4 \geq n^{(1-4\varepsilon)(1-2s-9\varepsilon)} > n^{1-2s-13\varepsilon} > n^{1-2s-\delta}$, finishing the proof. □

Here, the bound $n^{1-2s-\delta}$ is optimal up to the δ error term. Indeed, the Erdős-Rényi random graph with edge probability $p = 1 - n^{2s-1}$ has no clique of size larger that $n^{1-2s+o(1)}$, and its smallest eigenvalue satisfies $|\lambda_n| = O(n^s)$. Next, we use the previous theorem and Densification step 0 to prove Theorem 1.2, which we restate here.

Theorem 12.2. *There exists $c > 0$ such that the following holds for every $s > 0$, and every d sufficiently large with respect to s. Let G be a graph of average degree d and assume that $|\lambda_n| \leq d^s$. Then G contains a clique of size at least d^{1-cs}.*

Proof. Let $s_0 = 1/12$ and $\alpha = \alpha(s_0)$ be the constant given by Theorem 12.1. Then $\alpha(s) \geq \alpha$ for every $s \leq s_0$. By densification step 0, that is Lemma 7.1, G contains a subgraph G_0 on d vertices of density at least $\Omega(1/|\lambda_n|) \geq \Omega(d^{-s})$. Set $\delta = 2s/\alpha$, then the density of G_0 is at least $d^{-\alpha\delta}$, so we can apply Theorem 12.1 to G_0 to get a clique of size at least $d^{1-2s-\delta} = d^{1-s(2+2/\alpha)}$, so $c = \min\{12, 2 + 2/\alpha\}$ suffices. □

Now we prove the MaxCut version of Theorem 12.1, and use it to deduce Theorem 1.3.

Theorem 12.3. *Let $s \in (0, 1/60)$ and $\delta > 0$, then there exist $\rho > 0$ such that the following holds for every sufficiently large n. Let G be an n-vertex graph of edge density at least $n^{-\rho}$ such that $\text{surp}(G) \leq n^{1+s}$. Then G contains a clique of size at least $n^{1-2s-\delta}$.*

Proof. The proof of this is essentially identical to the proof of Theorem 12.1. The only difference is that we cite the MaxCut versions of our main densification results: Lemma 8.5, Corollary 9.7, Lemma 10.9 and Lemma 11.4. We also use that $\text{surp}^*(G) \leq O(\text{surp}(G) \log n)$ by Lemma 5.2. □

Proof of Theorem 1.3. Let $s = \delta_0 = \min\{1/100, \delta/10\}$, and let ρ be the constant guaranteed by Theorem 12.3 with δ_0 instead of δ. We show that $\varepsilon = \min\{\rho/3, s/2\}$ suffices. Let G be a graph with $\mathrm{surp}(G) \leq m^{1/2+\varepsilon}$, and let n be the number of vertices of G. We may assume that G contains no isolated vertices. Then, a result of Erdős, Gyárfás, and Kohayakawa [36] implies that $\mathrm{surp}(G) \geq n/6$. If $m \leq n^{2-3\varepsilon}$, then $\mathrm{surp}(G) \geq n/6 \geq \frac{1}{6}m^{1/(2-3\varepsilon)} \geq m^{1/2+\varepsilon}$, contradiction. Hence, $m \geq n^{2-3\varepsilon}$, so the edge density of G is at least $n^{-3\varepsilon} > n^{-\rho}$. As $\mathrm{surp}(G) \leq m^{1/2+\varepsilon} \leq n^{1+2\varepsilon} \leq n^{1+s}$, G contains a clique of size

$$n^{1-2s-\delta_0} = n^{1-3s} \geq m^{(1-3s)/2} \geq m^{1/2-\delta}$$

by Theorem 12.3. □

13 Edit distance from the union of cliques

The limitations of Theorems 9.4 and 9.6 are that they are only meaningful if the graph has large density. Indeed, these theorems imply that a graph of large minimum eigenvalue (or small surplus) is $o(1)$-close to a disjoint union of cliques, where a careful inspection of the parameters show that the $o(1)$ term is polylogarithmic in n. Hence, if the graph G has density at most n^{-c} for any small constant $c > 0$, they become meaningless as G is already $o(1)$-close to the empty graph. However, in this section we prove that polynomial proximity can also be establish, albeit under somewhat stronger condition on the smallest eigenvalue or the surplus.

We start with the following simple lemma, which will be used to argue that a dense graph with small surplus cannot induce sparse subgraphs.

Lemma 13.1. *Let G be a graph on n vertices. Let $X \cup Y$ be a partition of $V(G)$, and let $b = e(G[X, Y])$ and $c = e(G[Y])$. Then* $\mathrm{surp}(G) \geq \frac{b^2}{4n^2} - c$.

Proof. If $a = e(G[X])$ satisfies $a \leq b/2$, then $\mathrm{surp}(G)$ is at least

$$e(G[X,Y]) - \frac{e(G)}{2} = b - \frac{a+b+c}{2} = \frac{b-a-c}{2} \geq \frac{b}{4} - \frac{c}{2} \geq \frac{b^2}{4n^2} - c,$$

as desired.

Otherwise, we have $b < 2a$ and we can take $p = b/(4a) \in [0, 1/2)$. Let U be a random subset of X, where each vertex is included independently with probability $1/2+p$, and consider the cut $(U, (X \setminus U) \cup Y)$. Each edge in $G[X]$ has probability $1/2 - 2p^2$ of being cut, and each edge between X and Y is cut with probability $1/2 + p$. Therefore, the expected size of this cut is $a(1/2 - 2p^2) + b(1/2 + p)$, showing that the expected surplus is

$$a\Big(\frac{1}{2} - 2p^2\Big) + b\Big(\frac{1}{2} + p\Big) - \frac{a+b+c}{2} = bp - 2ap^2 - \frac{c}{2} = \frac{b^2}{8a} - \frac{c}{2} \geq \frac{b^2}{4n^2} - c,$$

where we have used that $a = e(G[X]) \leq n^2/2$ in the last step. □

13.1 Least eigenvalue version

Next, we show that a graph with large smallest eigenvalue contains a collection of large cliques such that almost all edges are contained in the subgraph induced by the union of these cliques.

Lemma 13.2. *Let $s \in (0, 1/6)$ and $\delta > 2s$, then there exists $\varepsilon > 0$ such that the following holds for sufficiently large n. Let G be a graph on n vertices such that $|\lambda_n| \leq n^s$. Then there exists $X \subset V(G)$ such that the number of edges not in $G[X]$ is at most $n^{2-\varepsilon}$, and $G[X]$ can be partitioned into cliques of size $n^{1-\delta}$.*

Proof. Choose any constants $s_0, \delta_0 > 0$ such that $s < s_0$ and $2s_0 + \delta_0 < \delta$. By Theorem 12.1, there exists $\rho = \rho(s_0, \delta_0) > 0$ such that every n_0-vertex graph of edge density at least $n_0^{-\rho}$ and smallest eigenvalue at least $-n_0^{s_0}$ contains a clique of size $n_0^{1-2s_0-\delta_0}$ as long as n_0 is sufficiently large.

We show that $\varepsilon = \min\{1/16, \rho/5, (1 - s/s_0)/2, (\delta - \delta_0 - 2s_0)/2\}$ suffices. Delete all vertices of G of degree less than $d = n^{1-2\varepsilon}$, and let G_0 be the resulting graph. Note that we removed at most dn edges. Repeat the following procedure. If G_i is defined and G_i contains a clique of size $n^{1-\delta}$, then let C_{i+1} be such a clique, and set $G_{i+1} = G_i \setminus C_{i+1}$. Otherwise, stop, and let I be the last index i for which G_i is defined. We show that $X = C_1 \cup \dots \cup C_I$ is the required set. It is clear that X can be partitioned into cliques of size $n^{1-\delta}$, so it remains to show that the number of edges not in X is at most $n^{2-\varepsilon}$.

Let $Y = V(G_I)$, $b = e(G_0[X, Y])$ and $c = e(G_0[Y])$.

Claim 13.3. *If $|Y| \geq n^{3/4}$, then $c \geq d^2|Y|^2/(20n^2)$.*

Proof. As G has minimum degree d, we have $b + 2c \geq d|Y|$. Assume that $c < d^2|Y|^2/(20n^2) < d|Y|/4$, then $b \geq d|Y|/2$. By Lemma 13.1, we have

$$\mathrm{surp}^*(G) \geq \mathrm{surp}^*(G_0) \geq \frac{b^2}{4n^2} - c.$$

Using that $\mathrm{surp}^*(G) \leq |\lambda_n| n \leq n^{1+s}$, we have

$$c \geq \frac{b^2}{4n^2} - n^{1+s} \geq \frac{d^2|Y|^2}{16n^2} - n^{1+s} \geq \frac{d^2|Y|^2}{20n^2}.$$

In the last inequality we used that $d = n^{1-2\varepsilon} \geq n^{7/8}$ and $|Y| \geq n^{3/4}$. $\square$

Now we show that $|Y| \leq n^{1-2\varepsilon}$. Indeed, suppose $|Y| > n^{1-2\varepsilon}$. By the above claim, G_I has at least $d^2|Y|^2/(20n^2) > |Y|^2 n^{-4\varepsilon}/20 \geq |Y|^{2-5\varepsilon}$ edges, which shows that G_I has edge density at least $|Y|^{-5\varepsilon} \geq |Y|^{-\rho}$. Moreover, by Cauchy's interlacing theorem, the smallest eigenvalue of G_I is at least that of G, which is at least $-n^s \geq -|Y|^{s/(1-2\varepsilon)} \geq -|Y|^{s_0}$.

But then, as discussed above, Theorem 12.1 guarantees that Y contains a clique of size $|Y|^{1-2s_0-\delta_0} > n^{1-2s_0-\delta_0-2\varepsilon} \geq n^{1-\delta}$, contradicting that G_I contains no clique of size $n^{1-\delta}$. Therefore, we must have $|Y| \leq n^{1-2\varepsilon}$.

From this, the number of edges of G not in $G[X]$ is at most

$$dn + e(G[X, Y]) + e(G[Y]) \leq dn + |Y|n \leq n^{1-\varepsilon}.$$

This finishes the proof. $\square$

Next, we show that the graph between two cliques must be either very dense or very sparse. In order to prove this, we use that graphs of small $|\lambda_n|$ avoid the following simple graph as an induced subgraph.

Lemma 13.4. *Let G be an n-vertex graph with the smallest eigenvalue λ_n and let $X, Y \subset V(G)$ be disjoint cliques of the same size. Then $G[X, Y]$ has either at most $O(|\lambda_n|^2|X|)$ edges, or at least $|X|^2 - O(|\lambda_n|^2|X|)$ edges.*

Proof. By the Cauchy interlacing theorem and Claim 3.3, G does not contain H_k as an induced subgraph for $k = 2|\lambda_n|^2$. We may assume that $|X| = |Y| \geq 4k$, otherwise the statement is trivial. Then each vertex in X has either at most k neighbours or at most k non-neighbours in Y. Moreover, each vertex in Y has at most k neighbours or at most k non-neighbours in X. Let $X_0 \subset X$ be the set of vertices with at most k neighbours, and let $X_1 = X \setminus X_0$, and define $Y_0, Y_1 \subset Y$ analogously. Suppose that X_0 and Y_1 both have size at least $2c|\lambda_n|^2$. If the number of edges between X_0 and Y_1 is at least $|X_0||Y_1|/2$, then there is a vertex in X_0 with at least $|Y_1|/2 \geq k$ neighbours in Y_1, contradiction. On the other hand, if the number of edges between X_0 and Y_1 is at most $|X_0||Y_1|/2$, then there is a vertex in Y_0 with at least $|X_0|/2 \geq k$

non-neighbours, contradiction. Therefore, it must hold that at least one of X_0 or Y_1 has size at most $2k$. If $|X_0| \leq 2k$, then $G[X,Y]$ has at least $|X_1|(|X|-k) \geq |X|^2 - 3k|X| = |X|^2 - O(|\lambda_n|^2|X|)$ edges. Otherwise, if $|Y_1| \leq 2k$, then $G[X,Y]$ has at most $|Y_0|k + |Y_1||X| \leq 3k|X| = O(|\lambda_n|^2|X|)$ edges. $\square$

We are ready to prove Theorem 1.6, which we restate here for convenience.

Theorem 13.5. *Let $s \in (0, 1/6)$, then there exists $\varepsilon > 0$ such that the following holds. Let G be an n-vertex graph such that $|\lambda_n| \leq n^s$. Then G is $n^{-\varepsilon}$-close to the disjoint union of cliques.*

Proof. Let $\delta = 1/2 > 2s$, and let $\varepsilon_0 = \varepsilon_0(s, \delta) > 0$ be the constant guaranteed by Lemma 13.2. We show that $\varepsilon = \min\{1/7, \varepsilon_0/2\}$ suffices.

By Lemma 13.2, there exists a set $X \subset V(G)$ such that X can be partitioned into the union of cliques of size $n_0 = n^{1-\delta}$, and G has at most $n^{2-\varepsilon_0}$ edges not in $G[X]$. Let $C_1, \dots, C_I$ be the cliques of size n_0 partitioning X, then $I = |X|/n_0$. Lemma 13.4 implies that the bipartite graph between C_i and C_j has either at most $O(n_0|\lambda_n|^2)$, or at least $n_0^2 - O(n_0|\lambda_n|^2)$ edges. In other words, (C_i, C_j) is a $O(|\lambda_n|^2/n_0)$-homogeneous pair, whose definition appears in Section 9. Here, $O(|\lambda_n|^2/n_0) < n^{-1/6}$. Define the auxiliary graph Γ on vertex set $\{1, \dots, I\}$, where we connect i and j if (C_i, C_j) is $n^{-1/6}$-full.

Claim 13.6. *Γ contains no cherry.*

Proof. If Γ contains a cherry, it means that there is a triple (C_i, C_j, C_k) such that (C_i, C_j) and (C_j, C_k) are $n^{-1/6}$-full, but (C_i, C_k) is $n^{-1/6}$-empty. By Lemma 9.2, then $G[C_i \cup C_j \cup C_k]$ has surplus at least $(1/4 - 3n^{-1/6})|C_i|^2 \geq |C_i|n^{1/2}/8$, which shows that the smallest eigenvalue of $G[C_i \cup C_j \cup C_k]$ is at most $-n^{1/2}/24$, contradiction. $\square$

Recall that graphs containing no cherry are the disjoint union of cliques. Therefore, we can partition $V(\Gamma)$ into sets $I_1, \dots, I_s$ such that $\Gamma[I_i]$ is a clique and there are no edges between I_i and I_j in Γ. But this gives a partition of X into sets $Y_1, \dots, Y_s$ by setting $Y_a = \bigcup_{i \in I_a} C_i$. Define $\widetilde{G}$ to be the graph on vertex set $V(G)$, where $Y_1, \dots, Y_s$ are cliques, and all edges of $\widetilde{G}$ are contained in one of these cliques.

We prove that $\widetilde{G}$ is $n^{-\varepsilon}$-close to G. For $1 \leq i < j \leq I$, $G[C_i, C_j]$ and $\widetilde{G}[C_i, C_j]$ differ by at most $n_0^2 \cdot n^{-1/6}$ edges. Therefore, $G[X]$ and $\widetilde{G}[X]$ differ by at most

$$\binom{|X|/n_0}{2} n_0^2 \cdot n^{-1/6} \leq n^{2-1/6}$$

edges. Furthermore, there are at most $n^{2-\varepsilon_0}$ edges of G not in $G[X]$, so G and $\widetilde{G}$ differ by at most $n^{2-\varepsilon_0} + n^{2-1/6} \leq n^{2-\varepsilon}$ edges. This finishes the proof. $\square$

13.2 MaxCut version

Finally, we prove the MaxCut version of Theorem 13.5. Most of the proof of this theorem is identical to the proof of Theorem 13.5, but a substantial difficulty arises when one tries the adapt the proof of Lemma 13.4. Before, we start with a variant of Lemma 13.2.

Lemma 13.7. *Let $s \in (0, 1/60)$ and $\delta > 2s$, then there exists $\varepsilon > 0$ such that the following holds. Let G be a graph on n vertices such that $\mathrm{surp}^*(G) \leq n^{1+s}$. Then there exists $X \subset V(G)$ such that the number of edges not in $G[X]$ is at most $n^{2-\varepsilon}$, and $G[X]$ can be partitioned into cliques of size $n^{1-\delta}$.*

Proof. The proof this is almost identical to the proof of Lemma 13.2. The only difference is that we cite Theorem 12.3 instead of Theorem 12.1. We omit further details. $\square$

Recall that Lemma 13.4 states that a pair of cliques in a graph with large smallest eigenvalue must form a δ-homogeneous pair for some very small δ. The proof of this relied on finding a simple forbidden subgraph H_k. Unfortunately, for graphs of small surplus, we do not have such a simple forbidden structure anymore. Instead, we show that if G contains two large cliques, then the bi-adjacency matrix of the bipartite graph between them is close to a rank-1 matrix. Then we show that this implies that the bipartite graph is close to a graph of rank one, which is just a complete bipartite graph. We prepare the proof of this with a couple of lemmas.

A *Boolean matrix* is a matrix with only zero and one entries. First, we show that if a Boolean matrix is approximated by a rank one matrix, then it is also approximated by a rank one Boolean matrix, or equivalently, a combinatorial rectangle.

Lemma 13.8. *Let A be an $n \times n$ Boolean matrix, and let $\delta \geq 0$. If there exist $u, v \in \mathbb{R}^n$ such that $\|A - uv^T\|_F^2 \leq \delta n^2$, then there exist $x, y \in \{0,1\}^n$ such that $\|A - xy^T\|_F^2 \leq O(\delta^{1/3} n^2)$.*

Proof. Without loss of generality, we may assume that $\delta \leq 1$. Furthermore, we may assume that u and v has nonnegative entries, as replacing every entry with the absolute value does not increase $\|A - uv^T\|_F^2$. Observe that $\|u\|_2^2 \|v\|_2^2 = \|uv^T\|_F^2$, which shows that

$$\|u\|_2 \|v\|_2 \leq \|A\|_F + \sqrt{\delta} n \leq 2n.$$

We may rescale u and v such that $\|u\|_2 = \|v\|_2 \leq \sqrt{2n}$. Let $\alpha = \delta^{1/6}$, and define $x, y \in \{0,1\}^n$ such that

$$x(i) = \begin{cases} 1 & \text{if } u(i) \geq \alpha \\ 0 & \text{otherwise,} \end{cases}$$

and similarly

$$y(i) = \begin{cases} 1 & \text{if } v(i) \geq \alpha \\ 0 & \text{otherwise.} \end{cases}$$

We show that xy^T is a good approximation of A. Note that $\|A - xy^T\|_F^2$ is the number of pairs (i,j) such that $A_{i,j} \neq x_i y_j$. We count these pairs in three cases.

Case 1. $A_{i,j} = 1$ and $x_i = 0$.

In this case, we have $u_i < \alpha$. If $v_j \leq 1/(2\alpha)$, then $(A_{i,j} - u_i v_j)^2 > 1/4$, so there are at most $4\delta n^2$ such pairs (i,j). On the other hand, the number of j such that $v_j \geq 1/(2\alpha)$ is at most $4\alpha^2 n$, as $\|v\|_2^2 = \sum_{j=1}^n v_j^2 \leq 2n$. Therefore, the number of (i,j) such that $A_{i,j} = 1$ and $x_i = 0$ is at most

$$4\delta n^2 + 4\alpha^2 n^2 = 4\delta n^2 + 4\delta^{1/3} n^2 = O(\delta^{1/3} n^2).$$

Case 2. $A_{i,j} = 1$ and $y_j = 0$.

This is symmetric to the previous case, so the number of such pairs is also at most $O(\delta^{1/3} n^2)$.

Case 3. $A_{i,j} = 0$ and $x_i = y_j = 1$.

In this case, $u_i \geq \alpha$ and $v_j \geq \alpha$, so $(A_{i,j} - u_i v_j)^2 \geq \alpha^4$. Thus, the total number of pairs (i,j) in this case is at most $\delta n^2 / \alpha^4 = \delta^{1/3} n^2$. $\square$

Next, we prove a simple technical lemma which shows that the union of two cliques has large surplus as long as the two cliques are not too disjoint, and do not overlap too much.

Lemma 13.9. *Let G be a graph such that $V(G) = C_1 \cup C_2$ and $E(G) = \binom{C_1}{2} \cup \binom{C_2}{2}$. Let $|C_1 \setminus C_2| = a$, $|C_2 \setminus C_1| = b$ and $|C_1 \cap C_2| = c$. Then*

$$\operatorname{surp}(G) \geq \frac{1}{4} \min\{a^2, b^2, c^2\}.$$

Proof. Let $A = C_1 \setminus C_2$, $B = C_2 \setminus C_1$, and $C = C_1 \cap C_2$. We may assume that $a = b$. Otherwise, if, say $a \leq b$, we remove vertices of $B \setminus C$ until its size is exactly a. Then it is enough to show that the resulting graph has surplus at least $\frac{1}{4}\min\{a^2, c^2\}$.

The number of edges of G is

$$2\binom{a+c}{2} - \binom{c}{2} < a^2 + 2ac + \frac{c^2}{2}.$$

If $c \leq a$, then define the cut (U, V) such that U is some $(a+c)/2$ element subset of A together with some $(a+c)/2$ element subset of B. The number of edges in this cut is $(a+c)^2/2$. Hence, the surplus of G is at least $\frac{1}{2}(a+c)^2 - \frac{1}{2}e(G) = c^2/4$.

If $c \geq a$, then define the cut (U, V) such that U is some $(a+c)/2$ element subset of C. Then the number of edges in this cut is $\frac{a+c}{2} \cdot \frac{c-a}{2} + 2\frac{a+c}{2} \cdot a = \frac{c^2}{4} + \frac{3}{4}a^2 + ac$. Therefore, the surplus is at least $a^2/4$. □

Now we are ready to prove our lemma about the surplus of the complement of bipartite graphs.

Lemma 13.10. *Let s and $\delta > 0$ such that $s + 6\delta < 1/2$. Let H be a bipartite graph with vertex classes of size n, and let $G = \overline{H}$. If $\operatorname{surp}^*(G) \leq n^{1+s}$, then either $e(H) \leq n^{2-\delta}$ or $e(H) \geq n^2 - n^{2-\delta}$.*

Proof. Let A be the adjacency matrix of G with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_{2n}$. Furthermore, let M be the adjacency matrix of H, and let $\mu_1 \geq \cdots \geq \mu_{2n}$ be the eigenvalues of M. As H is bipartite, we have $\mu_i = -\mu_{2n+1-i}$ for $i \in [2n]$. By Lemma 5.3 *(ii)* and Lemma 11.1, we have

$$\operatorname{surp}^*(G) \geq \Omega\left(\frac{1}{\sqrt{n}} \sum_{\lambda_i < 0} \lambda_i^2\right) \geq \Omega\left(\frac{1}{\sqrt{n}} \sum_{i \neq 1, \mu_i > 0} \mu_i^2\right) = \Omega\left(\frac{1}{\sqrt{n}} \sum_{i \neq 1, 2n} \mu_i^2\right).$$

Hence, if $\operatorname{surp}^*(G) \leq n^{1+s}$, then we have $\sum_{i \neq 1,2n} \mu_i^2 = O(n^{3/2+s})$.

On the other hand, we can express $\sum_{i \neq 1,2n} \mu_i^2$ as follows. The matrix M has the form $M = \begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix}$ with an appropriate $n \times n$ matrix B. Let v_1 be the principal eigenvector of M, then we can write $v_1 = (u, v)$, where $u, v \in \mathbb{R}^n$ correspond to the two vertex classes of H. Then the eigenvector corresponding to the smallest eigenvalue $\lambda_{2n} = -\lambda_1$ is $v_{2n} = (u, -v)$, and we have

$$\begin{aligned}\sum_{i \neq 1,2n} \mu_i^2 &= \|M - \lambda_1 v_1 v_1^T - \lambda_{2n} v_{2n} v_{2n}^T\|_F^2 \\ &= \left\| \begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix} - \lambda_1 \begin{pmatrix} uu^T & uv^T \\ vu^T & vv^T \end{pmatrix} + \lambda_1 \begin{pmatrix} uu^T & -uv^T \\ -vu^T & vv^T \end{pmatrix} \right\|_F^2 = 2\|B - 2\lambda_1 uv^T\|_F^2.\end{aligned}$$

Therefore, $\|B - 2\lambda_1 vu^T\|_F = O(n^{3/2+s})$. But then by Lemma 13.8, there exist $x, y \in \{0,1\}^n$ such that $\|B - xy^T\|_F^2 = O(n^{5/6+s/3})$. The matrix xy^T corresponds to a complete bipartite graph between the vertex classes of H, let $\widetilde{H}$ denote this complete bipartite graph, and let X_0 and Y_0 denote its vertex classes. Note that $e(\widetilde{H}) = \|xy^T\|_F^2 = |X_0||Y_0|$, and $\|B - xy^T\|_F^2$ is the number of edges $\widetilde{H}$ differs from H. Therefore, if $e(\widetilde{H}) \leq n^{2-\delta}/2$, then $e(H) \leq e(\widetilde{H}) + \|B - xy^T\|_F \leq n^{2-\delta}$, so we are done. We can proceed similarly if $e(\widetilde{H}) \geq n^2 - n^{2-\delta}/2$. Hence, we may assume that $n^{2-\delta}/2 \leq e(\widetilde{H}) \leq n^2 - n^{2-\delta}/2$. We show that this is impossible, by deriving that the surplus of G is too large in this case.

Let $\widetilde{G}$ be the complement of $\widetilde{H}$. Then $\widetilde{G}$ and G differ by at most $O(n^{5/6+s/3})$ edges. On the other hand, $\widetilde{G}$ is the union of two cliques, having vertex sets C_1 and C_2, where $X_0 = C_1 \setminus C_2$, $Y_0 = C_2 \setminus C_1$, and $C_1 \cap C_2 = V(G) \setminus (X_0 \cup Y_0)$. As $n^{2-\delta}/2 \leq e(\widetilde{H}) = |X_0||Y_0|$, we have $|X_0|, |Y_0| \geq n^{1-\delta}/2$. Also, as $e(\widetilde{H}) \leq n - n^{2-\delta}/2$, we have $|C_1 \cap C_2| = |V(G) \setminus (X_0 \cup Y_0)| \geq n^{2-\delta}/2$. Hence, by applying Lemma

13.9, we get that $\mathrm{surp}^*(\widetilde{G}) \geq \Omega(n^{2-2\delta})$. But as G and $\widetilde{G}$ differ by less than $O(n^{5/6+s/3})$ edges, and $5/6+s/3 < 2-2\delta$, this gives

$$\mathrm{surp}^*(G) \geq \mathrm{surp}^*(\widetilde{G}) - O(n^{5/6+s/3}) \geq \Omega(n^{2-2\delta}) > n^{1+s}$$

as well, contradiction. $\square$

Now we are ready to prove the main theorem of this section, that is, Theorem 1.4, which we restate here for the reader's convenience. The proof is essentially the same as the proof of Theorem 13.5.

Theorem 13.11. *Let $s \in (0, 1/60)$, then there exists $\varepsilon > 0$ such that the following holds. Let G be an n-vertex graph such that $\mathrm{surp}^*(G) \leq n^{1+s}$. Then G is $n^{-\varepsilon}$-close to the disjoint union of cliques.*

Proof. The proof is almost identical to the proof of Theorem 13.5, but we use Lemmas 13.7 and 13.10 instead of Lemmas 13.2 and 13.4. We omit further details. $\square$

Finally, we show that the previous theorem indeed implies Theorem 1.8 after complementation.

Proof of Theorem 1.8. Define the *positive discrepancy* of a graph G of edge density p as

$$\mathrm{disc}^+(G) = \max_{U \subset V(G)} e(G[U]) - p\binom{|U|}{2},$$

and define the negative discrepancy as

$$\mathrm{disc}^-(G) = \max_{U \subset V(G)} p\binom{|U|}{2} - e(G[U]).$$

It was proved in [64], Lemma 2.6, that if G is regular, then $\mathrm{surp}(G) = \Theta(\mathrm{disc}^-(G))$ and $\mathrm{bw}(G) = e(G)/2 - \Theta(\mathrm{disc}^+(G))$. Moreover, $\mathrm{disc}^+(G) = \mathrm{disc}^-(\overline{G})$. Therefore, the theorem follows from Theorem 1.4 after taking complements, and noting that if a *regular graph* G is close to a complement of a disjoint union of cliques, then G is close to a Turán graph. $\square$

14 Concluding remarks — Chowla's cosine problem in finite groups

Littlewood's L_1-problem and Chowla's cosine problem have also been studied in the setting of groups. Green and Konyagin [42] proposed to study the smallest L_1-norm of a dense set $A \subseteq \mathbb{Z}/p\mathbb{Z}$. If $\widehat{\mathbb{1}_A}$ denotes the Fourier transform of the indicator function of A over $\mathbb{Z}/p\mathbb{Z}$, they showed that $\sum_r |\widehat{\mathbb{1}_A}(r)| \geq (\log p)^{1/3-o(1)}$, which was later improved by Sanders [70] to $(\log p)^{1/2-o(1)}$. For sparser sets $A \subseteq \mathbb{Z}/p\mathbb{Z}$, a similar question has been studied by Schoen [72] and Konyagin and Shkredov [55].

Paralleling the extensions of the Littlewood L_1-problem, Sanders [71] extended Chowla's problem to finite abelian groups G as follows. For a symmetric subset $A \subset G$, one can define

$$M_G(A) = \sup_{y \in G} -\widehat{\mathbb{1}_A}(y),$$

where $\widehat{f} : G \to \mathbb{C}$ is the discrete Fourier transform of a function $f : G \to \mathbb{C}$, and $\mathbb{1}_A$ is the indicator function of A. Since A is a symmetric set, $\widehat{\mathbb{1}_A}$ is a real function. The natural analogue of Chowla's problem is to estimate the minimum of $M_G(A)$ over all centrally symmetric subsets A of size n. One can quickly observe that $M_G(A)$ need not go to infinity with the size of A. Indeed, if A is a subgroup of G, then $M_G(A) = 0$. On the other hand, Sanders ([71, Theorem 1.3]) proved that if A is far from a subgroup of G, then $M_G(A)$ is necessarily large. Formally, he proved that for every $\delta > 0$ there exists $c(\delta) > 0$ such that if $M_G(A) \leq |G|^{c(\gamma)}$, then there is some subgroup $H < G$ satisfying $|H \triangle A| \leq \delta|G|$. Noting that the image of the Fourier transform $\widehat{\mathbb{1}_A}$ is the spectrum of the Cayley graph $\mathrm{Cay}(G, A)$ generated by A, we can use our main results to immediately improve these bounds, and to extend them to non-abelian groups as well.

If G is a finite group and $A \subset G$ such that $A = A^{-1}$, define $M_G(A) = \max -\lambda$, where the maximum is taken among all eigenvalues of the Cayley graph $\mathrm{Cay}(G, A)$, which then coincides with the earlier definition for finite abelian groups. We show that $M_G(A)$ is small if and only if A is close to a subgroup of G.

Theorem 14.1. *Let $\delta, \varepsilon > 0$, then the following holds for every sufficiently large finite group G. Let $A \subset G$ such that $A = A^{-1}$. If $M_G(A) \leq |G|^{1/4-\varepsilon}$, then there exists a subgroup $H < G$ such that*

$$|H \triangle A| \leq \delta |G|.$$

Moreover, if $\alpha > 0$ is sufficiently small with respect to ε and $M_G(A) < |G|^{1/6-\varepsilon}$, then there exists a subgroup $H < G$ such that

$$|H \triangle A| \leq |G|^{1-\alpha}.$$

The main idea of the proof is to show that $\mathrm{Cay}(G, A)$ is close to the disjoint union of cliques if and only if A is close to a subgroup of G. This is proved in the following lemma.

Lemma 14.2. *Let G be a group and let $A \subset G$, $A = A^{-1}$, such that the number of pairs $(x, y) \in A \times A$ such that $xy^{-1} \notin A$ is at most $\varepsilon |A|^2$. Then there exists a subgroup $H < G$ such that $|H \triangle A| \leq O(\varepsilon^{1/2}|A|)$.*

In the proof, we use an old theorem of Freiman [37] on sets of very small doubling, sometimes referred to as Freiman's 3/2-theorem. See also the blog of Tao [73] for a short proof. Given subsets A, B of a group G, we write $AB := \{xy : x \in A, y \in B\}$.

Lemma 14.3 (Freiman's 3/2-theorem)**.** *Let G be a group and let $A \subset G$ such that $|AA^{-1}| \leq \frac{3}{2}|A|$. Then AA^{-1} and $A^{-1}A$ are both subgroups of G.*

Proof of Lemma 14.2. We may assume that $\varepsilon < 1/1000$, otherwise the statement is trivial. Also, we can assume that the identity $1_G \in A$, as adding 1_G does not change the number of pairs $(x, y) \in A \times A$ with $xy^{-1} \notin A$, and it only changes the size of A by 1.

Let N be the number of pairs $(x, y) \in A \times A$ such that $xy^{-1} \notin A$. Also, for every $x \in A$, let

$$N(x) = |(xA) \triangle A|.$$

Then using $A = A^{-1}$, we have

$$N = \frac{1}{2} \sum_{x \in A} N(x).$$

Therefore, $\frac{1}{|A|} \sum_{x \in A} N(x) \leq 2\varepsilon |A|$. Let $\delta = (2\varepsilon)^{1/2}$, and define

$$B = \{x \in A : N(x) \leq \delta |A|\}.$$

Then by simple averaging, we have $|B| \geq (1 - 2\varepsilon/\delta)|A| = (1 - \delta)|A|$. We also note that $B = B^{-1}$ as $N(x) = N(x^{-1})$, and $1_G \in B$. Observe that for every $x_1, x_2 \in B$, we can use the triangle inequality to write

$$|(x_1 x_2 A) \triangle A| \leq |(x_1 A) \triangle A| + |(x_1 x_2 A) \triangle (x_1 A)| \leq 2\delta |A|.$$

In particular, for every $x \in B \cdot B$, we have $|(xA) \triangle A| \leq 2\delta |A|$. Therefore,

$$\sum_{x \in B \cdot B} |(xA) \triangle A| \leq 2\delta |A||B \cdot B|.$$

On the other hand, $\sum_{x\in B\cdot B}|(xA)\triangle A|$ counts the number of pairs $(x,y)\in(B\cdot B)\times A$ such that $xy\notin A$ or $y\notin xA$. For every fixed y, the number of such pairs is clearly lower bounded by $|B\cdot B|-|A|$. Therefore, we can also write

$$\sum_{x\in B\cdot B}|(xA)\triangle A|\geq|A|(|B\cdot B|-|A|).$$

Comparing the lower and upper bounds on $\sum_{x\in B\cdot B}|(xA)\triangle A|$, we get

$$2\delta|A||B\cdot B|\geq|A|(|B\cdot B|-|A|),$$

from which

$$|B\cdot B|\leq\frac{1}{1-2\delta}|A|<(1+4\delta)|A|.$$

Therefore,

$$(1-\delta)|A|\leq|B|\leq|B\cdot B|\leq(1+4\delta)|A|,$$

which also implies that $|B\cdot B|\leq 3|B|/2$. Hence, we can apply Lemma 14.3 to conclude that $B\cdot B$ is a subgroup of G. Furthermore, as $1_G\in B$, we have $B\subset B\cdot B$, so $|A\cap B\cdot B|\geq|B|$. In conclusion

$$|A\triangle(B\cdot B)|\leq|A|+|B\cdot B|-2|B|\leq 6\delta|A|,$$

showing that $H=B\cdot B$ suffices. □

With Lemma 14.2 in our hands, Theorem 14.1 follows almost immediately from Theorems 9.4 and 13.5.

Proof of Theorem 14.1. We start with the first part of the theorem. We may assume that $|A|\geq\delta|G|$, otherwise the statement is trivial by choosing $H=\{1_G\}$. Let $\Gamma=\mathrm{Cay}(G,A)$, and let $\lambda_n=-M_G(A)$ be the smallest eigenvalue of Γ, $n=|G|$. We may assume that $1_G\notin A$, by noting that removing 1_G shifts the eigenvalues by -1. Therefore, Γ is a simple graph with no self loop. By Theorem 9.4, the inequality $|\lambda_n|\leq n^{1/4-\varepsilon}$ implies that Γ is β-close to a disjoint union of cliques for any $\beta>0$, given $n>n_0(\beta,\varepsilon)$. But then Γ contains at most $3\beta n^3$ cherries by Lemma 9.3. This implies that there are at most $6\beta n^2\leq 6(\beta/\delta^2)|A|^2$ pairs $(x,y)\in A\times A$ such that $xy^{-1}\notin A$. By Lemma 14.2, then $|A\triangle H|\leq O(\beta^{1/2}/\delta|A|)\leq\delta n$ for some subgroup $H<G$, assuming $\beta\ll\delta^4$.

The second part of the theorem follows essentially in the same manner, but we cite Theorem 13.5 instead of Theorem 9.4. We omit the details. □

Acknowledgements

We are grateful to Ilya Shkredov for bringing Chowla's problem to our attention, and pointing out connections to our earlier results. Also, we would like to thank Igor Balla, Clive Elphick, Jacob Fox, Lianna Hambardzumyan, Jack Koolean, Anqi Li, Nitya Mani, Benny Sudakov, and Quanyu Tang for many valuable discussions on various parts of this project.

References

[1] N. Alon. *On the edge-expansion of graphs.* Combinatorics, Probability and Computing 11 (1993): 1–10. 1.5

[2] N. Alon. *Bipartite subgraphs.* Combinatorica 16 (1996): 301–311. 1.2

[3] N. Alon, B. Bollobás, M. Krivelevich, and B. Sudakov. *Maximum cuts and judicious partitions in graphs without short cycles.* J. Combin. Theory Ser. B, 88(2) (2003): 329–346. 1, 1.2

[4] N. Alon, F.R.K. Chung. Explicit construction of linear sized tolerant networks. Discrete Mathematics 72 (1988): 15–19. 1, 1.3

[5] N. Alon, E. Fischer, M. Krivelevich, and M. Szegedy. *Efficient testing of large graphs.* Combinatorica 20(4) (2000): 451–476. 9

[6] N. Alon, M. Krivelevich, and B. Sudakov. *MaxCut in H-free graphs.* Combin. Probab. Comput. 14 (2005): 629–647. 1.2

[7] N. Alon, and A. Shapira. *A Characterization of Easily Testable Induced Subgraphs.* Comb. Prob. Comp. 15(6) (2006): 791–805. 9, 9

[8] I. Balla. *Equiangular lines via matrix projection.* preprint, arXiv:2110.15842 (2021). 1.4

[9] I. Balla, L. Hambardzumyan, and I. Tomon. *Factorization norms and an inverse theorem for MaxCut.* preprint, arxiv:2506.23989 (2025). 1.2

[10] I. Balla, O. Janzer, and B. Sudakov. *On MaxCut and the Lovász theta function.* Proc. Amer. Math. Soc. 152 (2024): 1871–1879. 1.2

[11] I. Balla, E. Räty, B. Sudakov, and I. Tomon. *Note on the second eigenvalue of regular graphs.* preprint, arXiv:2311.07629, 2023. 1.4

[12] P. Bärnkopf, Z. L. Nagy, and Z. Paulovics. *A Note on Internal Partitions: The 5-Regular Case and Beyond.* Graphs and Combinatorics 40 (2024): paper 36. 1.5

[13] B. Bedert. *An improved lower bound for a problem of Littlewood on the zeros of cosine polynomials* preprint, arXiv:2407.16075, 2024. 1.1

[14] J. Bourgain. *Sur le minimum d'une somme de cosinus.* Acta Arithmetica 45 (1986): 381–389. 1.1

[15] P. Borwein, T. Erdélyi, R. Ferguson, and R. Lockhart. *On the zeros of cosine polynomials: solution to a problem of Littlewood.* Ann. of Math. 167 (2008) 1109–1117. 1.1

[16] A. E. Brouwer and W. H. Haemers. *Spectra of Graphs.* Springer Heidelberg, 2012. 2

[17] P. J. Cameron, J. M. Goethals, J. J. Seidel, and E. E. Shult. *Line graphs, root systems, and elliptic geometry.* Journal of Algebra, 43(1) (1976): 305–327. 1.3

[18] C. Carlson, A. Kolla, R. Li, N. Mani, B. Sudakov, and L. Trevisan. *Lower bounds for max-cut in H-free graphs via semidefinite programming.* SIAM J. Discrete Math., 35(3) (2021): 1557–1568. 1.2

[19] M. Charikar, and A. Wirth. *Maximizing quadratic programs: extending Grothendieck's Inequality.* FOCS (2004): 54–60. 5

[20] S. Chowla. *The Riemann zeta and allied functions.* Bull. Amer. Math. Soc. 58 (1952) 287–305. 1.1

[21] S. Chowla. *Some applications of a method of A. Selberg.* Journal für die Reine und Angewandte Mathematik 217 (1965): 128–132. 1, 1.1

[22] A. Coja-Oghlan, P. Loick, B. F. Mezei, and G. B. Sorkin. *The Ising Antiferromagnet and Max Cut on Random Regular Graphs.* SIAM Journal on Discrete Mathematics 36 (2) (2022): 1306–1342. 1.2

[23] D. Coppersmith, D. Gamarnik, M. Hajiaghayi, G. B. and Sorkin. *Random MAX SAT, random MAX CUT, and their phase transitions.* Random Struct. Alg. 24 (2004): 502–545. 1.2

[24] D. Cvetković, and P. Rowlinson. *The largest eigenvalue of a graph: A survey.* Linear and multilinear algebra 28 (1-2) (1990): 3–33. 4

[25] J. A. Davis, S. Huczynska, L. Johnson, and J. Polhill. *Denniston partial difference sets exist in the odd prime case.* Finite Fields and Their Applications 99 (2024): Article 102499. 1.4

[26] D. de Caen. *Large equiangular sets of lines in Euclidean space.* Electronic Journal of Combinatorics 7 (2000): #R55. 1.3, 1.4

[27] A. Dembo, A. Montanari, and S. Sen. *Extremal cuts of sparse random graphs.* The Annals of Probability, Ann. Probab. 45(2) (2017): 1190–1217. 1.2, 1.5

[28] J. Díaz, N. Do, M. Serna, and N. Wormald. *Bounds on the max and min bisection of random cubic and random 4-regular graphs.* Theoretical Computer Science 307 (2003): 531–548. 1.5

[29] J. Díaz, M. Serna, and N. Wormald. *Bounds on the bisection width for random d-regular graphs.* Theoretical Computer Science 382 (2007): 120–130. 1.5

[30] C. S. Edwards. *Some extremal properties of bipartite subgraphs.* Canadian J. Math. 25 (1973): 475–485. 1.2

[31] C. S. Edwards. *An improved lower bound for the number of edges in a largest bipartite subgraph.* Recent advances in graph theory (Proc. Second Czechoslovak Sympos., Prague, 1974) (1975): 167–181. 1.2

[32] A. El Alaoui, A. Montanari, and M. Sellke. *Local algorithms for maximum cut and minimum bisection on locally treelike regular graphs of large degree.* Random Struct. Algorithms. 63 (2023): 689–715. 1.2

[33] T. Erdélyi. *The number of unimodular zeros of self-reciprocal polynomials with coefficients in a finite set.* Acta Arith., 176 (2016): 177–200. 1.1

[34] T. Erdélyi. *Improved lower bound for the number of unimodular zeros of self-reciprocal polynomials with coefficients in a finite set.* Acta Arith., 192 (2020): 189–210. 1.1

[35] P. Erdős. *Problems and results in graph theory and combinatorial analysis.* Graph theory and related topics (Proc. Conf., Univ. Waterloo, Waterloo, 1977), Academic Press, (1979) 153–163. 1, 1.2

[36] P. Erdős, A. Gyárfás, and Y. Kohayakawa. *The size of the largest bipartite subgraphs.* Disc. Math. 177 (1997): 267– 271. 12

[37] G. A. Freiman (Editor). *Number-Theoretic Studies in Markov Spectrum and in the structural theory of set addition.* Kalinin Gos. Univ. Moscow 1973 [Russian]. 14

[38] J. Friedman. *A proof of Alon's second eigenvalue conjecture and related problems.* Mem. Amer. Math. Soc., 195(910) (2008):viii+100. 1.4

[39] S. Glock, O. Janzer, and B. Sudakov. *New results for MaxCut in H-free graphs.* J. London Math. Soc. 108 (2023): 441–481. 1.2, 4

[40] M. X. Goemans, and D. P. Williamson. *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming.* Journal of the ACM 42 (6) (1995): 1115–1145. 1.2

[41] B. Green. *100 open problems.* https://people.maths.ox.ac.uk/greenbj/papers/open-problems.pdf 1.1

[42] B. Green and S. Konyagin. *On the Littlewood Problem Modulo a Prime.* Canadian Journal of Mathematics. 61 (2009): 141–164. 14

[43] W. H. Haemers. *Hoffman's ratio bound.* Lin. Alg. Appl. 617 (2021): 215–219. 1, 1.3

[44] P. Halmos. *A Hilbert Space Problem Book.* Second Edition, Springer-Verlag, 1982. 6

[45] M. Henzinger, J. Li, S. Rao and D. Wang. *Deterministic Near-Linear Time Minimum Cut in Weighted Graphs.* In: Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA) (2024): 3089–3139. 1.5

[46] M. Henzinger, S. Rao, and D. Wang. *Local flow partitioning for faster edge connectivity.* SIAM J.Comput. 49(1) (2020): 1–36. 1.5

[47] S. Hoory, N. Linial, and A. Wigderson. *Expander graphs and their applications.* Bull. Am. Math. Soc., 43 (4) (2006): 439–561. 1.4

[48] F. Ihringer. *Approximately Strongly Regular Graphs.* Discrete Math. 346 (3) (2023): 113299. 1.4

[49] T. Juškevičius and J. Sahasrabudhe. *Cosine polynomials with few zeros.* Bull. Lond. Math. Soc. 53 (2021): 877–892. 1.1

[50] D. Karger. *Minimum cuts in near-linear time.* In: Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing, STOC '96 (1996): 56–63. 1.5

[51] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. *Optimal Inapproximability Results for MAX-CUT and Other 2-Variable CSPs?* SIAM Journal on Computing (37) 1 (2007): 319–357. 1.2

[52] H. K. Kim, J. H. Koolen, and J. J. Yang. *A structure theory for graphs with fixed smallest eigenvalue.* Lin. Alg. Appl. 504 (2016): 1–13. 1.3

[53] S. V. Konyagin. *On the Littlewood problem.* Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya 45 (1981): 243–265, 463. 1.1

[54] S. V. Konyagin. *On zeros of sums of cosines.* Mat. Zametki, 108 (2020) 547–551. 1.1

[55] S. V. Konyagin and I. D. Shkredov. *On Wiener norm of subsets of Zp of medium size.* Journal of Mathematical Sciences, 218 (2016), 599–608. 14

[56] J. H. Koolen, M. Y. Cao, and Q. Yang. *Recent progress on graphs with fixed smallest adjacency eigenvalue: A survey.* Graphs and Combinatorics 37(4) (2021): 1139–1178. 1.3

[57] J. H. Koolen, C. Lv, G. Markowsky, and J. Park. *An improved bound for strongly regular graphs with smallest eigenvalue -m.* preprint, arXiv:2506.04964, 2025. 1.3

[58] J. H. Koolen, J. Y. Yang, and Q. Yang. *On graphs with smallest eigenvalue at least -3 and their lattices.* Adv. Math. 338 (2018): 847–864. 1.3

[59] J. E. Littlewood. *Some problems in real and complex analysis.* D. C. Heath and Company Raytheon Education Company, Lexington, MA, 1968. 1.1

[60] L. Lovász and B. Szegedy. *Regularity partitions and the topology of graphons.* An Irregular Mind, Imre Bárány, József Solymosi, and Gábor Sági editors, Bolyai Society Mathematical Studies 21 (2010): 415–446. 9

[61] O. C. McGehee, L. Pigno and B. Smith. *Hardy's inequality and the L_1 norm of exponential sums.* Annals of Mathematics. Second Series 113 (1981): 613–618. 1.1

[62] I. Mercer. *Finite searches, Chowla's cosine problem and large Newman polynomials.* Integers 19 (2019). 1.1

[63] A. Nilli. *On the second eigenvalue of a graph.* Discrete Mathematics, 91 (2) (1991): 207–210. 1, 1.4

[64] E. Räty, B. Sudakov, and I. Tomon. *Positive discrepancy, MaxCut, and eigenvalues of graphs.* to appear in Trans. AMS. 1.4, 1.5, 2, 5, 11, 13.2

[65] E. Räty, and I. Tomon. *Large Cuts in Hypergraphs via Energy.* Math. Proc. Camb. Soc. 179 (1) (2025): 45–61. 5, 5.2, 5

[66] E. Räty, and I. Tomon. *Bisection Width, Discrepancy, and Eigenvalues of Hypergraphs* preprint, arxiv:2409.15140, 2024. 1.5

[67] K. F. Roth. *On cosine polynomials corresponding to sets of integers* Acta Arithmetica 24 (1973): 87–98. 1.1

[68] I. Z. Ruzsa. *Negative values of cosine sums.* Acta Arithmetica 111 (2004): 179–186. 1.1

[69] J. Sahasrabudhe. *Counting zeros of cosine polynomials: On a problem of Littlewood.* Adv. Math. 343 (2019): 495–521. 1.1

[70] T. Sanders. *The Littlewood-Gowers problem.* J Anal Math 101 (2007): 123–162. 14

[71] T. Sanders. *Chowla's cosine problem.* Israel J. Math. 179 (2010): 1–28. 14

[72] T. Schoen. *On the Littlewood conjecture in $\mathbb{Z}/p\mathbb{Z}$.* Mosc. J. Comb. Number Theory 7 (2017) 66–72. 14

[73] T. Tao. *An elementary non-commutative Freiman theorem.* blog post `https://terrytao.wordpress.com/2009/11/10/an-elementary-non-commutative-freiman-theorem/` 14

[74] M. Uchiyama and S. Uchiyama. *On the cosine problem.* Proc. Japan Acad. 36 (1960): 475–479. 1.1

[75] J. H. van Lint, and A. E. Brouwer. *Strongly regular graphs and partial geometries.* Enumeration and design. Academic Press Inc. (1984): 85–122. 1.4