Additive Approximation for Edge-Deletion Problems

Noga Alon *　　Asaf Shapira †　　Benny Sudakov ‡

Abstract

A graph property is *monotone* if it is closed under removal of vertices and edges. In this paper we consider the following algorithmic problem, called the edge-deletion problem; given a monotone property $\mathcal{P}$ and a graph G, compute the smallest number of edge deletions that are needed in order to turn G into a graph satisfying $\mathcal{P}$. We denote this quantity by $E'_{\mathcal{P}}(G)$. The first result of this paper states that the edge-deletion problem can be efficiently approximated for any monotone property.

- For any fixed $\epsilon > 0$ and any monotone property $\mathcal{P}$, there is a deterministic algorithm, which given a graph $G = (V, E)$ of size n, approximates $E'_{\mathcal{P}}(G)$ in linear time $O(|V| + |E|)$ to within an additive error of ϵn^2.

Given the above, a natural question is for which monotone properties one can obtain better additive approximations of $E'_{\mathcal{P}}$. Our second main result essentially resolves this problem by giving a precise characterization of the monotone graph properties for which such approximations exist.

(1) If there is a bipartite graph that does not satisfy $\mathcal{P}$, then there is a $\delta > 0$ for which it is possible to approximate $E'_{\mathcal{P}}$ to within an additive error of $n^{2-\delta}$ in polynomial time.

(2) On the other hand, if all bipartite graphs satisfy $\mathcal{P}$, then for any $\delta > 0$ it is NP-hard to approximate $E'_{\mathcal{P}}$ to within an additive error of $n^{2-\delta}$.

While the proof of (1) is relatively simple, the proof of (2) requires several new ideas and involves tools from Extremal Graph Theory together with spectral techniques. Interestingly, prior to this work it was not even known that computing $E'_{\mathcal{P}}$ *precisely* for the properties in (2) is NP-hard. We thus answer (in a strong form) a question of Yannakakis, who asked in 1981 if it is possible to find a large and natural family of graph properties for which computing $E'_{\mathcal{P}}$ is NP-hard.

*Schools of Mathematics and Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel and IAS, Princeton, NJ 08540, USA. Email: nogaa@tau.ac.il. Research supported in part by the Israel Science Foundation, by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University and by the Von Neumann Fund.

†School of Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel. Email: asafico@tau.ac.il. This work forms part of the author's Ph.D. thesis. Research supported in part by a Charles Clore Foundation Fellowship and an IBM Ph.D. Fellowship.

‡Department of Mathematics, Princeton University, Princeton, NJ 08544, USA. E-mail: bsudakov@math.princeton.edu. Research supported in part by NSF grants DMS-0355497, DMS-0106589, and by an Alfred P. Sloan fellowship.

1 Introduction

1.1 Definitions, background and motivation

The topic of this paper is graph modification problems, namely problems of the type: "given a graph G, find the smallest number of modifications that are needed in order to turn G into a graph satisfying property $\mathcal{P}$". The main two types of such problems are the following, in *node modification* problems, one tries to find the smallest set of vertices, whose removal turns G into a graph satisfying $\mathcal{P}$, while in *edge modification* problems, one tries to find the smallest number of edge deletions/additions that turn G into a graph satisfying $\mathcal{P}$. In this paper we will focus on edge-modification problems. Before continuing with the introduction we need to introduce some notations.

For a graph property $\mathcal{P}$, let $\mathcal{P}_n$ denote the set of graphs on n vertices which satisfy $\mathcal{P}$. Given two graphs on n vertices, G and G', we denote by $\Delta(G, G')$ the edit distance between G and G', namely the smallest number of edge additions and/or deletions that are needed in order to turn G into G'. For a given property $\mathcal{P}$, we want to denote how far is a graph G from satisfying $\mathcal{P}$. For notational reasons it will be more convenient to normalize this measure so that it is always in the interval $[0, 1]$ (actually $[0, \frac{1}{2}]$). We thus define

Definition 1.1 **($E_{\mathcal{P}}(G)$)** *For a graph property $\mathcal{P}$ and a graph G on n vertices, let*

$$E_{\mathcal{P}}(G) = \min_{G' \in \mathcal{P}_n} \frac{\Delta(G, G')}{n^2} \ .$$

In words, $E_{\mathcal{P}}(G)$ is the minimum edit distance of G to a graph satisfying $\mathcal{P}$ after normalizing it by a factor of n^2.

Graph modification problems are well studied computational problems. In 1979, Garey and Johnson [28] mentioned 18 types of vertex and edge modification problems. Graph modification problems were extensively studied as these problems have applications in several fields, including Molecular Biology and Numerical Algebra. In these applications a graph is used to model experimental data, where edge modifications correspond to correcting errors in the data: Adding an edge means correcting a false negative, while deleting an edge means correcting a false positive. Computing $E_{\mathcal{P}}(G)$ for appropriately defined properties $\mathcal{P}$ have important applications in physical mapping of DNA (see [17], [29] and [31]). Computing $E_{\mathcal{P}}(G)$ for other properties arises when optimizing the running time of performing Gaussian elimination on a sparse symmetric positive-definite matrix (see [42]). Other modification problems arise as subroutines for heuristic algorithms for computing the largest clique in a graph (see [48]). Some edge modification problems also arise naturally in optimization of circuit design [18]. We briefly mention that there are also many results about *vertex* modification problems, notably that of Lewis and Yannakakis [38], who proved that for any nontrivial hereditary property $\mathcal{P}$, it is NP-hard to compute the smallest number of vertex deletions that turn a graph into one satisfying $\mathcal{P}$. (A graph property is hereditary if it is closed under removal of vertices.)

A graph property is said to be monotone if it is closed under removal of both vertices and edges. Examples of well studied monotone properties are k-colorability, and the property of being H-free for some fixed graph H. (A graph is H-free if it contains no copy of H as a not necessarily induced subgraph.) Note, that when trying to turn a graph into one satisfying a monotone property we will only use edge deletions. Therefore, in these cases the problem is sometimes called *edge-deletion* problem. Our main results, presented in the following subsections, give a nearly complete answer to the hardness of additive approximations of the edge-deletion problem for monotone properties.

1.2 An algorithm for any monotone property

Our first main result in this paper states that for any graph property $\mathcal{P}$ that belongs to the large, natural and well studied family of monotone graph properties, it is possible to derive efficient approximations of $E_{\mathcal{P}}$.

Theorem 1.1 *For any fixed $\epsilon > 0$ and any monotone property $\mathcal{P}$ there is a* **deterministic** *algorithm that given a graph G on n vertices computes in time $O(n^2)$ a real E satisfying $|E - E_{\mathcal{P}}(G)| \leq \epsilon$.*

Note, that the running time of our algorithm is of type $f(\epsilon)n^2$, and can in fact be improved to linear in the size of the input by first counting the number of edges, taking $E = 0$ in case the graph has less than ϵn^2 edges. We note that Theorem 1.1 was not known for many monotone properties. In particular, such an approximation algorithm was not even known for the property of being triangle-free and more generally for the property of being H-free for any non-bipartite H.

Theorem 1.1 is obtained via a novel structural graph theoretic technique. One of the applications of this technique (roughly) yields that every graph G, can be approximated by a small weighted graph W, in such a way that $E_{\mathcal{P}}(G)$ is approximately the optimal solution of a certain related problem (explained precisely in Section 3) that we solve on W. The main usage of this new structural-technique in this paper is in proving Lemmas 3.4 and 3.5 that lie at the core of the proof of Theorem 1.1. This new technique, which may very well have other algorithmic and graph-theoretic applications, applies a result of Alon, Fischer, Krivelevich and Szegedy [4] which is a strengthening of Szemerédi's Regularity Lemma [44]. We then use an efficient algorithmic version of the regularity lemma, which also implies an efficient algorithmic version of the result of [4], in order to transform the existential structural result into the algorithm stated in Theorem 1.1.

We further use our structural result in order to prove the following concentration-type result regarding the edit distance of subgraphs of a graph.

Theorem 1.2 *For every ϵ and any monotone property $\mathcal{P}$ there is a $d = d(\epsilon, \mathcal{P})$ with the following property: Let G be any graph and suppose we randomly pick a subset D, of d vertices from $V(G)$. Denote by G' the graph induced by G on D. Then,*

$$Prob[\ |E_{\mathcal{P}}(G') - E_{\mathcal{P}}(G)| > \epsilon] < \epsilon\ .$$

An immediate implication of the above theorem is the following,

Corollary 1.2 *For every $\epsilon > 0$ and any monotone property $\mathcal{P}$ there is a* **randomized** *algorithm that given a graph G computes in time $O(1)$ a real E satisfying $|E - E_{\mathcal{P}}(G)| \leq \epsilon$ with probability at least $1 - \epsilon$.*

We stress that there are some computational subtleties regrading the implementation of the algorithmic results discussed above. Roughly speaking, one should define how the property $\mathcal{P}$ is "given" to the algorithm and also whether ϵ is a fixed constant or part of the input. These issues are discussed in Section 5.

It is natural to ask if the above results can be extended to the larger family of hereditary properties, namely, properties closed under removal of vertices, but not necessarily under removal of edges. Many natural properties such as being Perfect, Chordal and Interval are hereditary non-monotone properties. By combining the ideas we used in order to prove Theorem 1.1 along with the main ideas of [6] it can be shown that Theorem 1.1 (as well as Theorem 1.2 and Corollary 1.2) also hold for any hereditary graph property.

1.3 On the possibility of better approximations

Theorem 1.1 implies that it is possible to efficiently approximate the distance of an n vertex graph from any monotone graph property $\mathcal{P}$, to within an error of ϵn^2 for any $\epsilon > 0$. A natural question one can ask is for which monotone properties it is possible to improve the additive error to $n^{2-\delta}$ for some fixed $\delta > 0$. In the terminology of Definition 1.1, this means to approximate $E_{\mathcal{P}}$ to within an additive error of $n^{-\delta}$ for some $\delta > 0$. Our second main result in this paper is a precise characterization of the monotone graph properties for which such a $\delta > 0$ exists[1].

Theorem 1.3 *Let $\mathcal{P}$ be a monotone graph property. Then,*

1. *If there is a bipartite graph that does not satisfy $\mathcal{P}$, then there is a fixed $\delta > 0$ for which it is possible to approximate $E_{\mathcal{P}}$ to within an additive error of $n^{-\delta}$ in polynomial time.*

2. *On the other hand, if all bipartite graphs satisfy $\mathcal{P}$, then for any fixed $\delta > 0$ it is NP-hard to approximate $E_{\mathcal{P}}$ to within an additive error of $n^{-\delta}$.*

While the first part of the above theorem follows easily from the known results about the Turán numbers of bipartite graphs (see, e.g., [45]), the proof of the second item involves various combinatorial tools. These include Szemerédi's Regularity Lemma, and a new result in Extremal Graph Theory, which is stated in Theorem 6.1 (see Section 6) that extends the main result of [14] and [15]. We also use the basic approach of [1], which applies spectral techniques to obtain an NP-hardness result by embedding a blow-up of a sparse instance to a problem, in an appropriate dense pseudo-random graph. Theorem 6.1 and the proof technique of Theorem 1.3 may be useful for other applications in graph theory and in proving hardness results. As in the case of Theorem 1.1, the second part of Theorem 1.3 was not known for many specific monotone properties. For example, prior to this paper it was not even known that it is NP-hard to *precisely* compute $E_{\mathcal{P}}$, where $\mathcal{P}$ is the property of being triangle-free. More generally, such a result was not known for the property of being H-free for any non-bipartite H.

1.4 Related work

Our main results form a natural continuation and extension of several research paths that have been extensively studied. Below we survey some of them.

1.4.1 Approximations of graph-modification problems

As we have previously mentioned many practical optimization problems in various research areas can be posed as the problem of computing the edit-distance of a certain graph from satisfying a certain property. Cai [16] has shown that for any hereditary property, which is expressible by a finite number of forbidden induced subgraphs, the problem of computing the edit distance is fixed-parameter tractable. Khot and Raman [33] proved that for some hereditary properties $\mathcal{P}$, finding in a given graph G, a subgraph that satisfies $\mathcal{P}$ is fixed-parameter tractable, while for other properties finding such a subgraph is hard in an appropriate sense (see [33]).

Note that Theorem 1.1 implies that if the edit distance (in our case, number of edge removals) of a graph from a property is $\Omega(n^2)$, then it can be approximated to within any *multiplicative* constant $1 + \epsilon$.

[1]We assume henceforth that $\mathcal{P}$ is not satisfied by all graphs.

1.4.2 Hardness of edge-modification problems

Natanzon, Shamir and Sharan [39] proved that for various hereditary properties, such as being Perfect and Comparability, computing $E_{\mathcal{P}}$ is NP-hard and sometimes even NP-hard to approximate to within some constant. Yannakakis [46] has shown that for several graph properties such as outerplanar, transitively orientable, and line-invertible, computing $E_{\mathcal{P}}$ is NP-hard. Asano [12] and Asano and Hirata [13] have shown that properties expressible in terms of certain families of forbidden minors or topological minors are NP-hard.

The NP-completeness proofs obtained by Yannakakis in [46], were add-hoc arguments that applied only to specific properties. Yannakakis posed in [46] as an open problem, the possibility of proving a general NP-hardness result for computing $E_{\mathcal{P}}$ that will apply to a general family of graph properties. Theorem 1.3 achieves such a result even for the seemingly easier problem of approximating $E_{\mathcal{P}}$.

1.4.3 Approximation schemes for "dense" instances

Fernandez de la Vega [22] and Arora, Karger and Karpinski [11] showed that many of the classical NP-complete problems such as MAX-CUT and MAX-3-CNF have a PTAS when the instance is dense, namely if the graph has $\Omega(n^2)$ edges or the 3-CNF formula has $\Omega(n^3)$ clauses. Approximations for dense instances of Quadratic Assignment Problems, as well as for additional problems, were obtained by Arora, Frieze and Kaplan [10]. Frieze and Kannan [26] obtained approximations schemes for several dense graph theoretic problems via certain matrix approximations. Alon, Fernandez de la Vega, Kannan and Karpinski [3] obtained results analogous to ours for any dense Constraint-Satisfaction-Problem via certain sampling techniques. It should be noted that all the above approximation schemes are obtained in a way similar to ours, that is, by first proving an *additive* approximation, and then arguing that in case the optimal solution is large (that is, $\Omega(n^2)$ in case of graphs, or $\Omega(n^3)$ in case of 3-CNF) the small additive error translates into a small multiplicative error.

All the above approximation results apply to the family of so called Constraint-Satisfaction-Problems. In some sense, these problems can express graph properties for which one imposes restrictions on **pairs** of vertices, such as k-colorability. These techniques thus fall short from applying to properties as simple as Triangle-freeness, where the restriction is on triples of vertices. The techniques we develop in order to obtain Theorem 1.1 enable us to handle restrictions that apply to *arbitrarily* large sets of vertices.

We briefly mention that $E_{\mathcal{P}}$ is related to packing problems of graphs. In [32] and [47] it was shown that by using linear programming one can approximate the packing number of a graph. In Section 9 we explain why this technique does not allow one to approximate $E_{\mathcal{P}}$.

1.4.4 Algorithmic applications of Szemerédi's Regularity Lemma

The authors of [2] gave a polynomial time algorithmic version of Szemerédi's Regularity Lemma. They used it to prove that Theorem 1.1 holds for the k-colorability property. The running time of their algorithm was improved by Kohayakawa, Rödl and Thoma [34]. Frieze and Kannan [25] further used the algorithmic version of the regularity lemma, to obtain approximation schemes for additional graph problems.

Theorem 1.1 is obtained via the algorithmic version of a strengthening of the standard regularity lemma, which was proved in [4], and it seems that these results cannot be obtained using the standard regularity lemma.

1.4.5 Tolerant Property-Testing

In standard Property-Testing (see [23] and [41]) one wants to distinguish between the graphs G that satisfy a certain graph property $\mathcal{P}$, or equivalently those G for which $E_{\mathcal{P}}(G) = 0$, from those that satisfy $E_{\mathcal{P}}(G) > \epsilon$. The main goal in designing property-testers is to reduce their query-complexity, namely, minimize the number of queries of the form "are i and j connected in the input graphs?".

Parnas, Ron and Rubinfeld [40] introduced the notion of Tolerant Property-Testing, where one wants to distinguish between the graphs G that satisfy $E_{\mathcal{P}}(G) < \delta$ from those that satisfy $E_{\mathcal{P}}(G) > \epsilon$, where $0 \leq \delta < \epsilon \leq 1$ are some constants. Recently, there have been several results in this line of work. Specifically, Fischer and Newman [24] have recently shown that if a graph property is testable with number of queries depending on ϵ only, then it is also tolerantly testable for any $0 \leq \delta < \epsilon \leq 1$ and with query complexity depending on $|\epsilon - \delta|$. Combining this with the main result of [7] implies that any monotone property is tolerantly testable for any $0 \leq \delta < \epsilon \leq 1$ and with query complexity depending on $|\epsilon - \delta|$. Note, that Corollary 1.2 implicitly states the same. In fact, the algorithm implied by Corollary 1.2 is the "natural" one, where one picks a random subset of vertices S, and approximates $E_{\mathcal{P}}(G)$ by computing $E_{\mathcal{P}}$ on the graph induced by S. The algorithm of [24] is far more complicated. Furthermore, due to the nature of our algorithm if the input graph satisfies a monotone property $\mathcal{P}$, namely if $E_{\mathcal{P}}(G) = 0$, we will always detect that this is the case. The algorithm of [24] may declare that $E_{\mathcal{P}}(G) > 0$ even if $E_{\mathcal{P}}(G) = 0$.

1.5 Organization

The proofs of the main results of this paper, Theorems 1.1 and 1.3, are independent of each other. Sections 2, 3, 4 and 5 contain the proofs relevant to Theorem 1.1 and Sections 6, 7 and 8 contain the proofs relevant to Theorem 1.3.

In Section 2 we introduce the basic notions of regularity and state the regularity lemmas that we use for proving Theorem 1.1 and some of their standard consequences. In Section 3 we give a high level description of the main ideas behind our algorithms. We also state the main structural graph theoretic lemmas, Lemmas 3.4 and 3.5 that lie at the core of these algorithms. The proofs of these lemmas appear in section 4. In Section 5 we give the proof of Theorems 1.1 and 1.2 as well as a discussion about some subtleties regarding the implementation of these algorithms.

Section 6 contains a high-level description of the proof of Theorem 1.3 as well as a description of the main tools that we apply in this proof. In Section 7 we prove a new Extremal Graph-Theoretic result that lies at the core of the proof of Theorem 1.3. In Section 8 we give the detailed proof of Theorem 1.3.

The final Section 9 contains some concluding remarks and open problems. Throughout the paper, whenever we relate, for example, to a function $f_{3.1}$, we mean the function f defined in Lemma/Claim/Theorem 3.1.

2 Regularity Lemmas and their Algorithmic Versions

In this section we discuss the basic notions of regularity, some of the basic applications of regular partitions and state the regularity lemmas that we use in the proof of Theorems 1.1 and 1.2. See [35] for a comprehensive survey on the regularity-lemma. We start with some basic definitions. For every two nonempty disjoint vertex sets A and B of a graph G, we define $e(A, B)$ to be the number of edges of G between A and B. The *edge density* of the pair is defined by $d(A, B) = e(A, B)/|A||B|$.

Definition 2.1 (γ-regular pair) *A pair (A,B) is* γ-regular*, if for any two subsets $A' \subseteq A$ and $B' \subseteq B$, satisfying $|A'| \geq \gamma|A|$ and $|B'| \geq \gamma|B|$, the inequality $|d(A',B') - d(A,B)| \leq \gamma$ holds.*

Throughout the paper we will make an extensive use of the notion of graph homomorphism which we turn to formally define.

Definition 2.2 (Homomorphism) *A homomorphism from a graph F to a graph K, is a mapping $\varphi : V(F) \mapsto V(K)$ that maps edges to edges, namely $(v,u) \in E(F)$ implies $(\varphi(v), \varphi(u)) \in E(K)$.*

In what follows, $F \mapsto K$ denotes the fact that there is a homomorphism from F to K. We will also say that a graph H is homomorphic to K if $H \mapsto K$. Note, that a graph H is homomorphic to a complete graph of size k if and only if H is k-colorable.

Let F be a graph on f vertices and K a graph on k vertices, and suppose $F \mapsto K$. Let G be a graph obtained by taking a copy of K, replacing every vertex with a sufficiently large independent set, and every edge with a random bipartite graph of edge density d. It is easy to show that with high probability, G contains a copy of F (in fact, many). The following lemma shows that in order to infer that G contains a copy of F, it is enough to replace every edge with a "regular enough" pair. Intuitively, the larger f and k are, and the sparser the regular pairs are, the more regular we need each pair to be, because we need the graph to be "closer" to a random graph. This is formulated in the lemma below. Several versions of this lemma were previously proved in papers using the regularity lemma (see [35]).

Lemma 2.3 *For every real $0 < \eta < 1$, and integers $k, f \geq 1$ there exist $\gamma = \gamma_{2.3}(\eta, k, f)$, and $N = N_{2.3}(\eta, k, f)$ with the following property. Let F be any graph on f vertices, and let $U_1, \ldots, U_k$ be k pairwise disjoint sets of vertices in a graph G, where $|U_1| = \ldots = |U_k| \geq N$. Suppose there is a mapping $\varphi : V(F) \mapsto \{1, \ldots, k\}$ such that the following holds: If (i,j) is an edge of F then $(U_{\varphi(i)}, U_{\varphi(j)})$ is γ-regular with density at least η. Then $U_1, \ldots, U_k$ span a copy of F.*

Comment 2.4 *Observe that the function $\gamma_{2.3}(\eta, k, f)$ may and will be assumed to be monotone non-increasing in k and f and monotone non-decreasing in η. Therefore, it will be convenient to assume that $\gamma_{2.3}(\eta, k, f) \leq \eta^2$. Similarly, we will assume that $N_{2.3}(\eta, k, f)$ is monotone non-decreasing in k and f. Also, for ease of future definitions (in particular those given in (2)) set $\gamma_{2.3}(\eta, k, 0) = N_{2.3}(\eta, k, 0) = 1$ for any $k \geq 1$ and $0 < \eta < 1$.*

A partition $\mathcal{A} = \{V_i \mid 1 \leq i \leq k\}$ of the vertex set of a graph is called an *equipartition* if $|V_i|$ and $|V_j|$ differ by no more than 1 for all $1 \leq i < j \leq k$ (so in particular each V_i has one of two possible sizes). The *order* of an equipartition denotes the number of partition classes (k above). A *refinement* of an equipartition $\mathcal{A}$ is an equipartition of the form $\mathcal{B} = \{V_{i,j} \mid 1 \leq i \leq k,\ 1 \leq j \leq l\}$ such that $V_{i,j}$ is a subset of V_i for every $1 \leq i \leq k$ and $1 \leq j \leq l$.

Definition 2.5 (γ-regular equipartition) *An equipartition $\mathcal{B} = \{V_i \mid 1 \leq i \leq k\}$ of the vertex set of a graph is called γ-regular if all but at most $\gamma\binom{k}{2}$ of the pairs $(V_i, V_{i'})$ are γ-regular.*

The Regularity Lemma of Szemerédi can be formulated as follows.

Lemma 2.6 ([44]) *For every m and $\gamma > 0$ there exists $T = T_{2.6}(m, \gamma)$ with the following property: If G is a graph with $n \geq T$ vertices, and $\mathcal{A}$ is an equipartition of the vertex set of G of order at most m, then there exists a refinement $\mathcal{B}$ of $\mathcal{A}$ of order k, where $m \leq k \leq T$ and $\mathcal{B}$ is γ-regular.*

$T_{2.6}(m,\gamma)$ may and is assumed to be monotone non-decreasing in m and monotone non-increasing in γ. Szemerédi's original proof of Lemma 2.6 was only existential as it supplied no efficient algorithm for obtaining the required equipartition. Alon et. al. [2] were the first to obtain a polynomial time algorithm for finding the equipartition, whose existence is guaranteed by lemma 2.6. The running time of this algorithm was improved by Kohayakawa et. al. [34] who obtained the following result.

Lemma 2.7 ([34]) *For every fixed m and γ there is an $O(n^2)$ time algorithm that given an equipartition $\mathcal{A}$ finds equipartition $\mathcal{B}$ as in Lemma 2.6.*

Our main tool in the proof of Theorem 1.1 is Lemma 2.9 below, proved in [4]. This lemma can be considered a strengthening of Lemma 2.6, as it guarantees the existence of an equipartition and a refinement of this equipartition that poses stronger properties compared to those of the standard γ-regular equipartition. This stronger notion is defined below.

Definition 2.8 ($\mathcal{E}$-regular equipartition) *For a function $\mathcal{E}(r) : \mathbb{N} \mapsto (0,1)$, a pair of equipartitions $\mathcal{A} = \{V_i \mid 1 \leq i \leq k\}$ and its refinement $\mathcal{B} = \{V_{i,j} \mid 1 \leq i \leq k,\ 1 \leq j \leq l\}$, where $V_{i,j} \subset V_i$ for all i, j, are said to be $\mathcal{E}$-regular if*

1. *For all $1 \leq i < i' \leq k$, for all $1 \leq j, j' \leq l$ but at most $\mathcal{E}(k)l^2$ of them, the pair $(V_{i,j}, V_{i',j'})$ is $\mathcal{E}(k)$-regular.*
2. *All $1 \leq i < i' \leq k$ but at most $\mathcal{E}(0)\binom{k}{2}$ of them are such that for all $1 \leq j, j' \leq l$ but at most $\mathcal{E}(0)l^2$ of them $|d(V_i, V_{i'}) - d(V_{i,j}, V_{i',j'})| < \mathcal{E}(0)$ holds.*

It will be very important for what follows to observe that in Definition 2.8 we may use an arbitrary *function* rather than a fixed γ as in Definition 2.5 (such functions will be denoted by $\mathcal{E}$ throughout the paper). The following is one of the main results of [4].

Lemma 2.9 ([4]) *For any integer m and function $\mathcal{E}(r) : \mathbb{N} \mapsto (0,1)$ there is $S = S_{2.9}(m, \mathcal{E})$ such that any graph on at least S vertices has an $\mathcal{E}$-regular equipartition $\mathcal{A}$, $\mathcal{B}$ where $|\mathcal{A}| = k \geq m$ and $|\mathcal{B}| = kl \leq S$.*

In order to make the presentation self contained we briefly review the proof of Lemma 2.9. Fix any m and function $\mathcal{E}$ and put $\zeta = \mathcal{E}(0)$. Partition G into m arbitrary subsets of equal size and denote this equipartition by $\mathcal{A}_0$. Put $M = m$. Iterate the following task: Apply Lemma 2.6 on $\mathcal{A}_{i-1}$ with $m = |\mathcal{A}_{i-1}|$ and $\gamma = \mathcal{E}(M)/M^2$ and let $\mathcal{A}_i$ be the refinement of $\mathcal{A}_{i-1}$ returned by Lemma 2.6. If $\mathcal{A}_{i-1}$ and $\mathcal{A}_i$ form an $\mathcal{E}$-regular equipartition stop, otherwise set $M = |\mathcal{A}_{i-1}|$ and reiterate. It is shown is [4] that after at most $100/\zeta^4$ iterations, for some $1 \leq i \leq 100/\zeta^4$ the partitions $\mathcal{A}_{i-1}$ and $\mathcal{A}_i$ form an $\mathcal{E}$-regular equipartition. Moreover, detecting an i for which this holds is very easy, that is, can be done in time $O(n^2)$ (see the proof in [4]). Note, that one can thus set the integer $S_{2.9}(m, \mathcal{E})$ to be the order of $\mathcal{A}_i$. In particular, the following is an immediate implication of the above discussion.

Proposition 2.10 *If m is bounded by a function of ϵ only, then for any $\mathcal{E}$ the integer $S = S_{2.9}(m, \mathcal{E})$ can be upper bounded by a function of ϵ only.*

The ϵ in the above proposition will be the ϵ from the task of approximating $E_{\mathcal{P}}$ within an error of ϵ in Theorem 1.1. Also, in our application of Lemma 2.9 the function $\mathcal{E}$ will (implicitly) depend on ϵ. For example, it will be convenient to set $\mathcal{E}(0) = \epsilon$. However, it follows from the definition of

$S_{2.9}(m, \mathcal{E})$ given above that even in this case it is possible to upper bound $S_{2.9}(m, \mathcal{E})$ by a function of ϵ only.

In order to design our algorithm we will need to obtain the equipartitions $\mathcal{A}$ and $\mathcal{B}$ that appear in the statement of Lemma 2.9. However, note that by the overview of the proof of Lemma 2.9 given above, in order to obtain this partition one can use Lemma 2.7 as an efficient algorithm for obtaining the regular partitions. Moreover, by Proposition 2.10 whenever we apply either $\mathcal{E}$ or Lemma 2.7 we are guaranteed that m (which in the above overview was M) is upper bounded by some function of ϵ and γ is lower bounded by some function of ϵ. This means that each of the at most $100/\zeta^4$ applications of Lemma 2.10 takes $O(n^2)$ time. We thus get the following:

Proposition 2.11 *If m is bounded by a function of ϵ only, then for any $\mathcal{E}$ there is an $O(n^2)$ algorithm for obtaining the equipartitions $\mathcal{A}$ and $\mathcal{B}$ of Lemma 2.9.*

3 Overview of the Proof of Theorem 1.1

We start with a convenient way of handling a monotone graph property.

Definition 3.1 (Forbidden Subgraphs) *For a monotone graph property $\mathcal{P}$, define $\mathcal{F} = \mathcal{F}_{\mathcal{P}}$ to be the set of graphs which are minimal with respect to not satisfying property $\mathcal{P}$. In other words, a graph F belongs to $\mathcal{F}$ if it does not satisfy $\mathcal{P}$, but any graph obtained from F by removing an edge or a vertex, satisfies $\mathcal{P}$.*

As an example of a family of forbidden subgraphs, consider $\mathcal{P}$ which is the property of being 2-colorable. Then $\mathcal{F}_{\mathcal{P}}$ is the set of all odd-cycles. Clearly, a graph satisfies $\mathcal{P}$ if and only it contains no member of $\mathcal{F}_{\mathcal{P}}$ as a (not necessarily induced) subgraph. We say that a graph is $\mathcal{F}$-free if it contains no (not necessarily induced) subgraph $F \in \mathcal{F}$. Clearly, for any family $\mathcal{F}$, being $\mathcal{F}$-free is a monotone property. Thus, the monotone properties are precisely the graph properties that are equivalent to being $\mathcal{F}$-free for some family $\mathcal{F}$. In order to simplify the notation, it will be simpler to talk about properties of type $\mathcal{F}$-free rather than monotone properties. To avoid confusion we will henceforth denote by $E_{\mathcal{F}}(G)$ the value of $E_{\mathcal{P}}(G)$, where $\mathcal{F} = \mathcal{F}_{\mathcal{P}}$ as above.

The main idea we apply in order to obtain the algorithmic results of this paper is quite simple; given a graph G, a family of forbidden subgraphs $\mathcal{F}$ and $\epsilon > 0$ we use Lemma 2.9 with appropriately defined parameters in order to construct in $O(n^2)$ time a weighted complete graph W, of size depending on ϵ but **independent** of the size of G, such that a solution of a certain "related" problem on W gives a good approximation of $E_{\mathcal{F}}(G)$. As W will be of size independent of the size of G, we may and will use exhaustive search in order to solve the "related" problem on W. In what follows we give further details on how to define W and the "related" problem that we solve on W.

We start with the simplest case, where the property is that of being triangle-free, namely $\mathcal{F} = \{K_3\}$. Let W be some weighted complete graph on k vertices and let $0 \leq w(i,j) \leq 1$ denote the weight of the edge connecting i and j in W. Let $E_{\mathcal{F}}(W)$ be the natural extension of the definition of $E_{\mathcal{F}}(G)$ to weighted graphs, namely, instead of just counting how many edges should be removed in order to turn G into an $\mathcal{F}$-free graph, we ask for the edge set of minimum weight with the above property. Let G be a k-partite graph on n vertices with partition classes $V_1, \ldots, V_k$ of equal size n/k. Suppose for every $i < j$ we have $d(V_i, V_j) = w(i,j)$ (recall that $d(V_i, V_j)$ denotes the edge density between V_i and V_j). In some sense, W can be considered a weighted approximation of G, but to our investigation a more important question is whether W can be used in order to estimate $E_{\mathcal{F}}(G)$? In other words, is it true that $E_{\mathcal{F}}(G) \approx E_{\mathcal{F}}(W)$?

It is easy to see that $E_{\mathcal{F}}(G) \leq E_{\mathcal{F}}(W)$. Indeed, given a set of edges S, whose removal turns W into a triangle free graph, we simply remove all edges connecting V_i and V_j for every $(i, j) \in S$. The main question is whether the other direction is also true. Namely, is it true that if it is possible to remove αn^2 from G and thus make it triangle free, then it is possible to remove from W a set of edges of total weight approximately αk^2 and thus make it triangle-free? If true this will mean that by computing $E_{\mathcal{F}}(W)$ we also approximately compute $E_{\mathcal{F}}(G)$. Unfortunately, this assertion is false in general, as the minimal number of edge modifications that are enough to make G triangle-free, may involve removing *some* and not *all* the edges connecting a pair (V_i, V_j), and in W we can remove only edges and not parts of them. It thus seems natural to ask what kind of restrictions should we impose on G (or more precisely on the pairs (V_i, V_j)) such that the above situation will be impossible, namely, that the optimal way to turn G into a triangle free graph will involve removing either none or all the edges connecting a pair (V_i, V_j) (up to some small error). This will clearly imply that we also have $E_{\mathcal{F}}(G) \approx E_{\mathcal{F}}(W)$.

One natural restriction is that the pairs (V_i, V_j) would be random bipartite graphs. While this restriction indeed works it is of no use for our investigation as we are trying to design an algorithm that can handle arbitrary graphs and not necessarily random graphs. One is thus tempted to replace random bipartite graph with γ-regular pairs for some small enough γ. Unfortunately, we did not manage to prove that there is a small enough $\gamma > 0$ ensuring that even if all pairs (V_i, V_j) are γ-regular then $E_{\mathcal{F}}(G) \approx E_{\mathcal{F}}(W)$. In order to circumvent this difficulty we use the stronger notion of $\mathcal{E}$-regularity defined in Section 2. As it turns out, if one uses an appropriately defined function $\mathcal{E}$, then if all pairs (V_i, V_j) are $\mathcal{E}(k)$-regular, one can infer that $E_{\mathcal{F}}(G) \approx E_{\mathcal{F}}(W)$. This result is (essentially) formulated in Lemma 3.4.

In the above discussion we considered the case $\mathcal{F} = \{K_3\}$. So suppose now that $\mathcal{F}$ is an arbitrary (possibly infinite) family of graph. Suppose we use a weighted complete graph W on k vertices as above in order to approximate some k-partite graph. The question that naturally arises at this stage is what problem should we try to solve on W in order to get an approximation of $E_{\mathcal{F}}(G)$. It is easy to see that G may be very far from being $\mathcal{F}$-free, while at the same time W can be $\mathcal{F}$-free, simply because $\mathcal{F}$ does not contain graphs of size at most k. As an example, consider the case, where the property is that of containing no copy of the complete bipartite graph with two vertices in each side, denoted $K_{2,2}$. Now, if G is the complete bipartite graph $K_{n/2,n/2}$ then it is very far from being $K_{2,2}$-free. However, in this case W is just an edge that spans no copy of $K_{2,2}$.

It thus seems that we must solve a *different* problem on W. To formulate this problem we need the following definitions.

Definition 3.2 (**$\mathcal{F}$-homomorphism-free**) *For a family of graphs $\mathcal{F}$, a graph W is called $\mathcal{F}$-homomorphism-free if $F \not\mapsto W$ for any $F \in \mathcal{F}$.*

We now define a measure analogous to $E_{\mathcal{F}}$ but with respect to making a graph $\mathcal{F}$-homomorphism-free. Note that we focus on weighted graphs.

Definition 3.3 (**$\mathcal{H}_{\mathcal{F}}(W)$**) *For a family of graphs $\mathcal{F}$ and a weighted complete graph W on k vertices, let $\mathcal{H}'_{\mathcal{F}}(W)$ denote the minimum total weight of a set of edges, whose removal from W turns it into an $\mathcal{F}$-homomorphism-free graph. Define, $\mathcal{H}_{\mathcal{F}}(W) = \mathcal{H}'_{\mathcal{F}}(W)/k^2$.*

Note, that in Definition 3.2 the graph W is an unweighed not necessarily complete graph. Also, observe that when $\mathcal{F} = \{K_3\}$ then we have $\mathcal{H}_{\mathcal{F}}(W) = E_{\mathcal{F}}(W)$. As it turns out, the "right" problem to solve on W is to compute $\mathcal{H}_{\mathcal{F}}(W)$. This is formulated in the following key lemma, whose proof appears in Section 4:

Lemma 3.4 (The Key Lemma) *For every family of graphs $\mathcal{F}$, there are functions $N_{3.4}(k,\epsilon)$ and $\gamma_{3.4}(k,\epsilon)$ with the following property*[2]*: Let W be any weighted complete graph on k vertices and let G be any k-partite graph with partition classes $V_1,\ldots,V_k$ of equal size such that*

1. $|V_1| = \ldots = |V_k| \geq N_{3.4}(k,\epsilon)$.
2. *All pairs (V_i,V_j) are $\gamma_{3.4}(k,\epsilon)$-regular.*
3. *For every $1 \leq i < j \leq k$ we have $d(V_i,V_j) = w(i,j)$.*

Then, $E_{\mathcal{F}}(G) \geq \mathcal{H}_{\mathcal{F}}(W) - \epsilon$.

It is easy to argue as we did above and prove that $E_{\mathcal{F}}(G) \leq \mathcal{H}_{\mathcal{F}}(W)$ in Lemma 3.4 (see the proof of Lemma 3.5), however we will not need this (trivial) direction. It is important to note that while Lemma 3.4 is very strong as it allows us to approximate $E_{\mathcal{F}}(G)$ via computing $\mathcal{H}_{\mathcal{F}}(W)$ (recall that W is intended to be very small compared to G) its main weakness is that it requires the regularity between each of the pairs to be a function of k, which denotes the number of partition classes, rather than depending solely on the family of graphs $\mathcal{F}$. We note that even if $\mathcal{F} = \{K_3\}$ as discussed above, we can only prove Lemma 3.4 with a regularity measure that depends on k. This supplies some explanation as to why Lemma 2.6 (the standard regularity lemma) is not sufficient for our purposes; note that the input to Lemma 2.6 is some fixed $\gamma > 0$ and the output is a γ-regular equipartition with number of partition classes that depends on γ (the function $T_{2.6}(m,\gamma)$). Thus, even if all pairs are γ-regular, this γ may be very large when considering the number of partition classes returned by Lemma 2.6 and the regularity measure which Lemma 3.4 requires. Hence, the standard regularity lemma cannot help us with applying Lemma 3.4. In order to overcome this problem we use the notion of $\mathcal{E}$-regular partitions and the stronger regularity-lemma given in Lemma 2.9, which, when appropriately used, allows us to apply Lemma 3.4 in order to obtain Lemma 3.5 below, from which Theorem 1.1 follows quite easily. The proof of this lemma appears in Section 4.

Lemma 3.5 *For any $\epsilon > 0$ and family of graphs $\mathcal{F}$ there are functions $N_{3.5}(r)$ and $\mathcal{E}_{3.5}(r)$ satisfying the following*[3]*: Suppose a graph G has an $\mathcal{E}_{3.5}$-regular equipartition $\mathcal{A} = \{V_i \mid 1 \leq i \leq k\}$, $\mathcal{B} = \{V_{i,j} \mid 1 \leq i \leq k,\ 1 \leq j \leq l\}$, where*

1. $k \geq 1/\epsilon$.
2. $|V_{i,j}| \geq N_{3.5}(k)$ *for every* $1 \leq i \leq k$ *and* $1 \leq j \leq l$.

Let W be a weighted complete graph on k vertices with $w(i,j) = d(V_i,V_j)$. Then,

$$|E_{\mathcal{F}}(G) - \mathcal{H}_{\mathcal{F}}(W)| \leq \epsilon .$$

Using the algorithmic version of Lemma 2.9, which is given in Proposition 2.11, we can rephrase the above lemma in a more algorithmic way, which is more or less the algorithm of Theorem 1.1: Given a graph G we use the $O(n^2)$ time algorithm of Proposition 2.11 in order to obtain the equipartition described in the statement of Lemma 3.5. We then construct the graph W as in Lemma 3.5, and finally use exhaustive search in order to precisely compute $\mathcal{H}_{\mathcal{F}}(W)$. By Lemma 3.5, this gives a good approximation of $E_{\mathcal{F}}(G)$. The proof of Theorem 1.1 appears in Section 5.

[2]The functions $N_{3.4}(k,\epsilon)$ and $\gamma_{3.4}(k,\epsilon)$ will also (implicitly) depend on $\mathcal{F}$.

[3]The functions $N_{3.5}(r)$ and $\mathcal{E}_{3.5}(r)$ will also (implicitly) depend on ϵ and $\mathcal{F}$.

4 Proofs of Lemmas 3.4 and 3.5

In this section we apply our new structural technique in order to prove Lemmas 3.4 and 3.5. Regretfully, it is hard to precisely state what are the ingredients of this technique. Roughly speaking, it uses the notion of $\mathcal{E}$-regularity in order to partition the edges of a graph into a bounded number of edge sets, which have regular-partitions that are almost identical[4] and more importantly, the regularity-measure of each of the bipartite graphs in each of the edge sets can be a function of the number of clusters.

We start this section with some definitions that will be very useful for the proof of Lemma 3.4.

Definition 4.1 *For any (possibly infinite) family of graphs $\mathcal{F}$, and any integer r let $\mathcal{F}_r$ be the following set of graphs: A graph R belongs to $\mathcal{F}_r$ if it has at most r vertices and there is at least one $F \in \mathcal{F}$ such that $F \mapsto R$.*

Definition 4.2 *For any family of graphs $\mathcal{F}$ and integer r for which $\mathcal{F}_r \neq \emptyset$, define*

$$\Psi_{\mathcal{F}}(r) = \max_{R \in \mathcal{F}_r} \min_{\{F \in \mathcal{F} : F \mapsto R\}} |V(F)|. \tag{1}$$

Define $\Psi_{\mathcal{F}}(r) = 0$ if $\mathcal{F}_r = \emptyset$. Therefore, $\Psi_{\mathcal{F}}(r)$ is monotone non-decreasing in r.

Practicing definitions, note that if $\mathcal{F}$ is the family of odd cycles, then $\mathcal{F}_k$ is precisely the family of non-bipartite graphs of size at most k. Also, in this case $\Psi_{\mathcal{F}}(k) = k$ when k is odd, and $\Psi_{\mathcal{F}}(k) = k-1$ when k is even. The "right" way to think of the function $\Psi_{\mathcal{F}}$ is the following: Let R be a graph of size at most k and suppose we are guaranteed that there is a graph $F' \in \mathcal{F}$ such that $F' \mapsto R$ (thus $R \in \mathcal{F}_k$). Then by this information only and *without* having to know the structure of R itself, the definition of $\Psi_{\mathcal{F}}$ implies that there is a graph $F \in \mathcal{F}$ of size at most $\Psi_{\mathcal{F}}(k)$, such that $F \mapsto R$.

The function $\Psi_{\mathcal{F}}$ has a critical role in the proof of Lemma 3.4. While proving this lemma we will use Lemma 2.3 in order to derive that some k sets of vertices, which are regular enough, span some graph $F \in \mathcal{F}$. Roughly speaking, the main difficulty will be that we will not know the size of F, and as a consequence will not know the regularity measure between these sets that is sufficient for applying Lemma 2.3 on these k sets (this quantity is $\gamma_{2.3}(\eta, k, |V(F)|)$). However, we *will* know that there is *some* $F' \in \mathcal{F}$ which is spanned by these sets. The function $\Psi_{\mathcal{F}}(r)$ will thus be very useful as it supplies an upper bound for the size of the smallest $F \in \mathcal{F}$ which is spanned by these sets. See Proposition 4.4, where $\Psi_{\mathcal{F}}(r)$ has a crucial role.

Proof of Lemma 3.4: Given ϵ and k let

$$T = T(k, \epsilon) = T_{2.6}(k, \gamma_{2.3}(\epsilon/2, k, \Psi_{\mathcal{F}}(k))). \tag{2}$$

We prove the lemma with $\gamma_{3.4}(k, \epsilon)$ and $N_{3.4}(k, \epsilon)$ satisfying

$$\gamma_{3.4}(k, \epsilon) = \min(\epsilon/2,\ 1/T), \tag{3}$$

$$N_{3.4}(k, \epsilon) = T \cdot N_{2.3}(\epsilon/2, k, \Psi_{\mathcal{F}}(k)) \tag{4}$$

Suppose G is a graph on n vertices, in which case each set V_i is of size $\frac{n}{k}$. We may thus show that one must remove at least $\mathcal{H}_{\mathcal{F}}(W) \cdot n^2 - \epsilon n^2$ edges from G in order to make it $\mathcal{F}$-free. To this

[4] Two regular partitions $V_1, \ldots, V_k$ and $U_1, \ldots, U_k$ are identical if $d(V_i, V_j) = d(U_i, U_j)$

end, it is enough to show that if there is a graph G' that is obtained from G by removing less than $\mathcal{H}_{\mathcal{F}}(W) \cdot n^2 - \epsilon n^2$ edges and spans no $F \in \mathcal{F}$ then it is possible to remove from W a set of edges of total weight less than $\mathcal{H}_{\mathcal{F}}(W) \cdot k^2$ and obtain a graph W' that is $\mathcal{F}$-homomorphism-free. This will obviously be a contradiction.

Assume such a G' exists and apply Lemma 2.6 on it with $\gamma = \gamma_{2.3}(\frac{1}{2}\epsilon, k, \Psi_{\mathcal{F}}(k))$ and $m = k$ (we use $m = k$ as G is already partitioned into k subsets $V_1, \ldots, V_k$). For the rest of the proof we denote by $V_{i,1}, \ldots, V_{i,l}$ the partition of V_i that Lemma 2.6 returns. Recall that as $|V_1| = \ldots = |V_k|$ and Lemma 2.6 partitions a graph into subsets of equal size, then all the sets V_i are partitioned into the same number l of subsets. Note also that by Lemma 2.6 and the definition of T in (2) we have $l < T$. Observe, that T is in fact an upper bound for the *total* number of partition classes $V_{i,j}$).

By Lemma 2.6 (recall that by Comment 2.4 we may assume $\gamma_{2.3}(\frac{1}{2}\epsilon, k, \Psi_{\mathcal{F}}(k)) \leq \frac{1}{2}\epsilon$), we are guaranteed that out of the lk sets $V_{i,j}$ at most $\frac{\epsilon}{2}\binom{lk}{2}$ pairs are not $\gamma_{2.3}(\frac{1}{2}\epsilon, k, \Psi_{\mathcal{F}}(k))$-regular. We define a graph G'', which is obtained from G' by removing all the edges connecting pairs $(V_{i,i'}, V_{j,j'})$ that are not $\gamma_{2.3}(\frac{1}{2}\epsilon, k, \Psi_{\mathcal{F}}(k))$-regular, and all edges connecting pairs $(V_{i,i'}, V_{j,j'})$ for which their edge density in G' is smaller than $\frac{1}{2}\epsilon$.

Proposition 4.3 *There are k sets $V_{1,t_1}, \ldots, V_{k,t_k}$ such that the graphs induced by G and G'' on these k sets differ by less than $\mathcal{H}_{\mathcal{F}}(W) \cdot \frac{n^2}{l^2} - \frac{\epsilon n^2}{2l^2}$ edges.*

Proof: We first claim that G'' is obtained from G' by removing less than $\frac{\epsilon}{2}n^2$ edges. To see this note that the number of edges connecting a pair $(V_{i,i'}, V_{j,j'})$ is at most $(n/kl)^2$. As there are at most $\frac{\epsilon}{2}\binom{lk}{2}$ pairs that are not $\gamma_{2.3}(\frac{1}{2}\epsilon, k, \Psi_{\mathcal{F}}(k))$-regular, we remove at most $\frac{\epsilon}{4}n^2$ edges due to such pairs. Finally, as due to pairs, whose edge density is at most $\frac{1}{2}\epsilon$, we remove at most $\binom{kl}{2}\frac{\epsilon}{2}(n/kl)^2 \leq \frac{\epsilon}{4}n^2$ edges, the total number of edges removed is at most $\frac{\epsilon}{2}n^2$, as needed.

As we assume that G' is obtained from G by removing less than $\mathcal{H}_{\mathcal{F}}(W) \cdot n^2 - \epsilon n^2$ edges, we get from the previous paragraph that G'' is obtained from G be removing less than $\mathcal{H}_{\mathcal{F}}(W) \cdot n^2 - \frac{\epsilon}{2}n^2$ edges. Suppose for every $1 \leq i \leq k$ we randomly and uniformly pick one of the sets $V_{i,1}, \ldots, V_{i,l}$. The probability that an edge, which belongs to G and not to G'', is spanned by these k sets is l^{-2}. As G and G'' differ by less than $\mathcal{H}_{\mathcal{F}}(W) \cdot n^2 - \frac{\epsilon}{2}n^2$ edges, we get that the expected number of such edges is less than $\mathcal{H}_{\mathcal{F}}(W) \cdot \frac{n^2}{l^2} - \frac{\epsilon n^2}{2l^2}$ and therefore there must be a choice of k sets that span less than this number of such edges. ■

We are now ready to arrive at a contradiction by showing that if it is possible to remove less than $\mathcal{H}_{\mathcal{F}}(W) \cdot n^2 - \epsilon n^2$ edges from G and thus turn it into an $\mathcal{F}$-free graph G', then we can remove from W a set of edges of total weight less than $\mathcal{H}_{\mathcal{F}}(W) \cdot k^2$ and thus turn it into an $\mathcal{F}$-homomorphism-free graph W'. Let $V_{1,i_1}, \ldots, V_{k,i_k}$ be the k sets satisfying the condition of Proposition 4.3 and obtain from W a graph W' by removing from W edge (i,j) if and only if the density of (V_{i,t_i}, V_{j,t_j}) in G'' is 0.

Proposition 4.4 *W' is $\mathcal{F}$-homomorphism-free.*

Proof: Assume $F' \mapsto W'$ for some $F' \in \mathcal{F}$. As W' is a graph of size k this means (recall Definition 4.2) that there is $F \in \mathcal{F}$ of size at most $\Psi_{\mathcal{F}}(k)$ such that $F \mapsto W'$. Let φ be a homomorphism from F to W'. By definition of φ, for any $(u,v) \in E(F)$ we have $(\varphi(u), \varphi(v))$ is an edge of W'. Recall that by definition of G'' either the density of a pair $(V_{i,i'}, V_{j,j'})$ in G'' is zero, or this density is at least $\frac{1}{2}\epsilon$ and the pair is $\gamma_{2.3}(\frac{1}{2}\epsilon, k, \Psi_{\mathcal{F}}(k))$-regular. By definition of W', this means that for every $(u,v) \in E(F)$ the pair $(V_{\varphi(u),t_{\varphi(u)}}, V_{\varphi(v),t_{\varphi(v)}})$ has density at least $\frac{\epsilon}{2}$ in G'' and is $\gamma_{2.3}(\frac{1}{2}\epsilon, k, \Psi_{\mathcal{F}}(k))$-regular. By item

1 of the lemma we have for all $1 \le i \le k$ that $|V_i| \ge N_{3.4}(k, \epsilon)$. By our choice in (4) and the fact that $l \le T$, the sets V_{i,t_i} must therefore be of size at least

$$|N_{3.4}(k, \epsilon)|/l \ge |N_{3.4}(k, \epsilon)|/T = N_{2.3}(\frac{1}{2}\epsilon, k, \Psi_{\mathcal{F}}(k)).$$

Hence, the sets $V_{1,t_1}, \dots, V_{k,t_k}$ satisfy all the necessary requirements needed in order to apply Lemma 2.3 on them in order to deduce that they span a copy of F in G'' (recall, that we have already argued that $|V(F)| \le \Psi_{\mathcal{F}}(k)$). This, however, is impossible, as we assumed that G' was already $\mathcal{F}$-free and G'' is a subgraph of G'. ■

Proposition 4.5 *For any $i < j$ the edge densities of (V_i, V_j) and (V_{i,t_i}, V_{j,t_j}) satisfy in G*

$$|d(V_i, V_j) - d(V_{i,t_i}, V_{j,t_j})| \le \frac{1}{2}\epsilon.$$

Proof: Recall that $1/l > 1/T$ and by (3) we have $1/T > \gamma_{3.4}(k, \epsilon)$. We infer that $|V_{i,t_i}| = |V_i|/l \ge \gamma_{3.4}(k, \epsilon)|V_i|$. By item 2 of the lemma, each pair (V_i, V_j) is $\gamma_{3.4}(k, \epsilon)$-regular in G. Hence, by definition of a regular pair, we must have $|d(V_i, V_j) - d(V_{i,t_i}, V_{j,t_j})| \le \gamma_{3.4}(k, \epsilon) \le \frac{1}{2}\epsilon$. ■

Proposition 4.6 *W' is obtained from W by removing a set of edges of weight less than $\mathcal{H}_{\mathcal{F}}(W) \cdot k^2$.*

Proof: Let S be the set of edges removed from W and denote by $w(S)$ the total weight of edges in S. Let $e(V_{i,t_i}, V_{j,t_j})$ denote the number of edges connecting the pair (V_{i,t_i}, V_{j,t_j}) in G. We claim that the following series of inequalities, which imply that $w(S) < \mathcal{H}_{\mathcal{F}}(W) \cdot k^2$, hold:

$$\begin{aligned}
\mathcal{H}_{\mathcal{F}}(W) \cdot \frac{n^2}{l^2} - \frac{\epsilon n^2}{2l^2} &> \sum_{(i,j)\in S} e(V_{i,t_i}, V_{j,t_j}) \\
&\ge \sum_{(i,j)\in S} (w(i,j) - \frac{\epsilon}{2})\frac{n^2}{l^2k^2} \\
&\ge \sum_{(i,j)\in S} w(i,j)\frac{n^2}{l^2k^2} - \frac{\epsilon n^2}{2l^2} \\
&= w(S)\frac{n^2}{l^2k^2} - \frac{\epsilon n^2}{2l^2} \ .
\end{aligned}$$

Indeed, recall that by the definition of W', we have $(i, j) \in S$ if and only if the density of the pair $(V_{i,i'}, V_{j,j'})$ in G'' is 0, which means that all the edges connecting this pair were removed in G''. As by Proposition 4.3 the total difference between G and G'' is less than $\mathcal{H}_{\mathcal{F}}(W) \cdot \frac{n^2}{l^2} - \frac{\epsilon n^2}{2l^2}$ we infer that the first (strict) inequality is valid. The second inequality follows from Proposition 4.5 together with the fact that by the condition of the lemma we have $d(V_i, V_j) = w(i, j)$. The third inequality is due to the fact that W has k vertices and thus $|S| \le k^2$. ■

The sought after contradiction now follows immediately from Propositions 4.4 and 4.6. This completes the proof of the lemma. ■

We continue with the proof of Lemma 3.5.

Proof of Lemma 3.5: We prove the lemma with:

$$\mathcal{E}_{3.5}(r) = \begin{cases} \frac{1}{16}\epsilon^2, & r = 0 \\ \min(\frac{1}{8}\epsilon r^{-2},\ \frac{1}{8}\epsilon^2,\ \gamma_{3.4}(r, \frac{1}{8}\epsilon)), & r \geq 1 \end{cases} \tag{5}$$

and

$$N_{3.5}(r) = N_{3.4}(r, \frac{1}{8}\epsilon)\ .$$

We start with showing that $E_{\mathcal{F}}(G) \leq \mathcal{H}_{\mathcal{F}}(W) + \epsilon$. Suppose G is a graph of n vertices, in which case the number of edges connecting V_i and V_j is $w(i,j) \cdot \frac{n^2}{k^2}$. We first remove all the edges within the sets $V_1, \ldots, V_k$. As $k \geq 1/\epsilon$ the total number of edges removed in this step is at most $k\binom{n/k}{2} \leq \epsilon n^2$.

Let S be the set of minimal weight whose removal turns W into an $\mathcal{F}$-homomorphism-free graph W'. We claim that if for every $(i,j) \in S$ we remove all the edges connecting V_i and V_j the resulting graph G' spans no copy of a graph $F \in \mathcal{F}$. Suppose to the contrary that G' spans a copy of $F \in \mathcal{F}$, and consider the mapping $\varphi : V(F) \mapsto \{1, \ldots, k\}$ that maps every vertex of F that belongs to V_j to j. As we have removed all the edges within the sets $V_1, \ldots, V_k$ and all edges between V_i and V_j for any $(i,j) \in S$ we get that φ is a homomorphism from F to W' contradicting our choice of S. Finally, note that the number of edges removed in the second step is

$$\sum_{(i,j)\in S} w(i,j) \cdot \frac{n^2}{k^2} = n^2 \cdot \mathcal{H}_{\mathcal{F}}(W)\ .$$

Combined with the first step the total number of edges removed is at most $n^2 \cdot \mathcal{H}_{\mathcal{F}}(W) + \epsilon n^2$, as needed.

For the rest of the proof we focus on proving $\mathcal{H}_{\mathcal{F}}(W) \leq E_{\mathcal{F}}(G) + \epsilon$. Let $\mathcal{A}$ and $\mathcal{B}$ be the two equipartitions from the statement of the lemma. Suppose for every $1 \leq i \leq k$ we randomly, uniformly and independently pick a set V_{i,t_i} out of the sets $V_{i,1}, \ldots, V_{i,l}$. Let P denote the event that (i) All the pairs $(V_{i,t_i}, V_{i',t_{i'}})$ are $\mathcal{E}(k)$-regular. (ii) All but at most $\frac{1}{2}\epsilon\binom{k}{2}$ of the pairs $(V_{i,t_i}, V_{i',t_{i'}})$ satisfy $|d(V_{i,t_i}, V_{i',t_{i'}}) - d(V_i, V_{i'})| \leq \mathcal{E}(0)$. We need the following observations:

Proposition 4.7 *P holds with probability at least $1 - \frac{1}{2}\epsilon$.*

Proof: Fix any $i < i'$. By definition of $\mathcal{E}_{3.5}$ we have $\mathcal{E}(k) \leq \frac{1}{8}\epsilon k^{-2}$, thus by item 1 of Definition 2.8, the probability that $(V_{i,t_i}, V_{i',t_{i'}})$ is not $\mathcal{E}(k)$-regular is at most $\frac{1}{8}\epsilon k^{-2}$. By the union bound, the probability that one of the pairs is not $\mathcal{E}(k)$-regular is at most $\binom{k}{2}\frac{1}{8}\epsilon k^{-2} \leq \frac{1}{4}\epsilon$.

Item 2 of Definition 2.8 can be rephrased as stating that there are at most $\mathcal{E}(0)\binom{k}{2} = \frac{1}{16}\epsilon^2\binom{k}{2}$ choices of $i < i'$ for which the probability that $|d(V_{i,t_i}, V_{i',t_{i'}}) - d(V_i, V_{i'})| > \mathcal{E}(0) = \frac{1}{16}\epsilon^2$ is larger than $\mathcal{E}(0) = \frac{1}{16}\epsilon^2$. Thus, the expected number of $i < i'$ for which $|d(V_{i,t_i}, V_{i',t_{i'}}) - d(V_i, V_{i'})| > \mathcal{E}(0)$ is at most $\frac{1}{16}\epsilon^2\binom{k}{2} \cdot 1 + \binom{k}{2} \cdot \frac{1}{16}\epsilon^2 \leq \frac{1}{8}\epsilon^2\binom{k}{2}$. By Markov's inequality, the probability that more than $\frac{1}{2}\epsilon\binom{k}{2}$ of $i < i'$ violate the above inequality is at most $\frac{\epsilon}{4}$.

As properties (i) and (ii) of event P each hold with probability at least $1 - \frac{1}{4}\epsilon$, we get that P holds with probability at least $1 - \frac{1}{2}\epsilon$. ■

Proposition 4.8 *Assume event P holds and denote by G' the subgraph of G that is spanned by the sets $V_{1,t_1}, \ldots, V_{k,t_k}$. Then, $E_{\mathcal{F}}(G') \geq \mathcal{H}_{\mathcal{F}}(W) - \frac{1}{2}\epsilon$.*

Proof: Let W' be a weighted complete graph on k vertices satisfying $w(i,i') = d(V_{i,t_i}, V_{i',t_{i'}})$. Event P assumes that all the pairs $(V_{i,t_i}, V_{i',t_{i'}})$ are $\mathcal{E}(k)$-regular. As $\mathcal{E}(k) \leq \gamma_{3.4}(k, \frac{1}{8}\epsilon)$ and the lemma assumes that $|V_{i,j}| \geq N_{3.5}(k) = N_{3.4}(k, \frac{1}{8}\epsilon)$ we may deduce from Lemma 3.4 that

$$E_{\mathcal{F}}(G') \geq \mathcal{H}_{\mathcal{F}}(W') - \frac{\epsilon}{8}. \tag{6}$$

Now, event P also assumes that all but at most $\frac{\epsilon}{2}\binom{k}{2}$ of the pairs $i < i'$ are such that $|d(V_i, V_{i'}) - d(V_{i,t_i}, V_{i',t_{i'}})| \leq \mathcal{E}(0) < \frac{\epsilon}{8}$. This means that the sum of edge weights of W' differs from the sum of edge weights of W by at most $\frac{\epsilon}{2}\binom{k}{2}$ due to pairs that violate the above inequality and by at most $\binom{k}{2}\frac{\epsilon}{8}$ due to the other pairs. This means that the sum of edge weights of W' differs from that of W by at most $\frac{\epsilon}{4}k^2 + \frac{\epsilon}{16}k^2 \leq \frac{3\epsilon}{8}k^2$. This clearly implies that

$$\mathcal{H}_{\mathcal{F}}(W') \geq \mathcal{H}_{\mathcal{F}}(W) - \frac{3\epsilon}{8}\ . \tag{7}$$

The proof now follows by combining (6) and (7). ■

Let R be an arbitrary set of edges whose removal from G turns it into an $\mathcal{F}$-free graph. Randomly and uniformly select a set V_{i,t_i} from each of the sets $V_{i,1}, \ldots, V_{i,l}$, and let R' denote the set of edges of R that are spanned by these k sets. We claim that the following upper and lower bound on the expected size of R' hold:

$$\begin{aligned}
\frac{1}{l^2} \cdot |R| &= \mathbb{E}[|R'|] \\
&\geq \mathbb{E}[|R'| \mid P] \cdot Prob[P] \\
&\geq (1 - \frac{\epsilon}{2}) \cdot \mathbb{E}[|R'| \mid P] \\
&\geq (1 - \frac{\epsilon}{2}) \cdot (\mathcal{H}_{\mathcal{F}}(W) - \frac{\epsilon}{2}) \cdot k^2 \frac{n^2}{(kl)^2} \\
&\geq (\mathcal{H}_{\mathcal{F}}(W) - \epsilon) \cdot \frac{n^2}{l^2}\ .
\end{aligned}$$

Indeed, the equality is due to the fact than an edge of R has probability precisely $1/l^2$ to be in R'. The second inequality is due to Proposition 4.7, the third is due to Proposition 4.8 and the last is valid because $\mathcal{H}_{\mathcal{F}}(W) \leq 1$. As we thus infer that $|R| \geq \mathcal{H}_{\mathcal{F}}(W) \cdot n^2 - \epsilon n^2$ for *arbitrary* R, we get that $E_{\mathcal{F}}(G) \geq \mathcal{H}_{\mathcal{F}}(W) - \epsilon$, thus completing the proof. ■

5 Proofs of Algorithmic Results

The technical lemmas proved in the previous sections enabled us to infer that certain $\mathcal{E}$-regular partitions may be very useful for approximating $E_{\mathcal{P}}$. In this section we apply Proposition 2.11 in order to efficiently obtain these partitions. We first prove Theorem 1.1, while overlooking some subtle issues. We then discuss them is detail.

Proof of Theorem 1.1: Fix any $\epsilon > 0$ and monotone graph property $\mathcal{P}$. Let $\mathcal{F} = \mathcal{F}_{\mathcal{P}}$ be the family of forbidden subgraphs of $\mathcal{P}$ as in Definition 3.1. As satisfying $\mathcal{P}$ is equivalent to being $\mathcal{F}$-free, we focus on approximating $E_{\mathcal{F}}(G)$. Let $\mathcal{E}_{3.5}(r)$ and $N_{3.5}(r)$ be the appropriate function with respect to $\mathcal{F}$ and ϵ. Put $S(\epsilon) = S_{2.9}(1/\epsilon, \mathcal{E}_{3.5})$ and recall that by Proposition 2.10 the integer S can indeed be upper bounded by a function of ϵ.

If an input graph has less than $S(\epsilon) \cdot N_{3.5}(S(\epsilon))$ vertices we use exhaustive search in order to precisely compute $E_{\mathcal{F}}(G)$. Assume then that G has more than $S(\epsilon) \cdot N_{3.5}(S(\epsilon))$ vertices, and use Proposition 2.11 with $m = 1/\epsilon$ and $\mathcal{E}_{3.5}(r)$ as above in order to compute the equipartition $\mathcal{A} = \{V_i \mid 1 \leq i \leq k\}$ and its refinement $\mathcal{B} = \{V_{i,j} \mid 1 \leq i \leq k,\ 1 \leq j \leq l\}$ satisfying the conditions of Lemma 2.9. As m is bounded by a function of ϵ we get from Proposition 2.11 that this step takes time $O(n^2)$. Also, by Lemma 2.9 we have $kl \leq S$, therefore, as G has at least $S(\epsilon) \cdot N_{3.5}(S(\epsilon))$ vertices each of the sets $V_{i,j}$ is of size at least $N_{3.5}(S(\epsilon)) \geq N_{3.5}(k)$. Let W be a weighted complete graph of size k where $w(i, j) = d(V_i, V_j)$. Using exhaustive search, we can now precisely compute the value of $\mathcal{H}_{\mathcal{F}}(W)$. By Lemma 3.5 we may infer that $|E_{\mathcal{F}}(G) - \mathcal{H}_{\mathcal{F}}(W)| \leq \epsilon$. ■

As we have mentioned in the introduction, one should specify how the property $\mathcal{P}$ is given to the algorithm. For example, $\mathcal{P}$ may be an undecidable property, in which case we cannot do anything. We thus focus on decidable graph properties. However, even in this case we may face some unexpected problems. Note, that for a general infinite family of graphs $\mathcal{F}$ it is not clear how to compute $\mathcal{H}_{\mathcal{F}}$ in finite time. Also, returning to the overview of the proof of Lemma 2.9 given in Section 2, note that we have implicitly assumed that one can compute the function $\mathcal{E}$, as this is needed in order to compute the parameters with which one applies Lemma 2.10. A close inspection of the proofs of Lemmas 3.4 and 3.5 reveals that computing $\mathcal{E}$ involves computing the function $\Psi_{\mathcal{F}}$ (see (2), (3) and (5)). One of the main results of [5] asserts that somewhat surprisingly, there is a family of graph properties $\mathcal{F}$, for which the property of being $\mathcal{F}$-free is decidable (in fact, in $coNP$) but at the same time $\Psi_{\mathcal{F}}$ is not computable. Therefore, even if we confine ourselves to decidable graph properties we still run into trouble.

Suppose first that ϵ is not part of the input to the algorithm. As we have discussed in Section 2, in this case all the applications of $\mathcal{E}_{3.5}$ are on inputs of size depending on ϵ only, thus the algorithm may "keep" the answers to these (finitely many) applications of $\mathcal{E}_{3.5}$ as part of its description. Similarly, in this case we may need to compute $\mathcal{H}_{\mathcal{F}}$ on graphs of size depending on ϵ only[5], thus the algorithm may "keep" the answers to these (finitely many) applications of $\mathcal{H}_{\mathcal{F}}$ as part of its description. Observe, that we don't need to keep the answer of $\mathcal{H}_{\mathcal{F}}$ for all the (infinite) range of edge weights. Rather, as we only need to approximate $E_{\mathcal{F}}$ within an additive error of ϵ, it is enough to consider edge weights $\{0, \epsilon, 2\epsilon, 3\epsilon, \ldots, 1\}$.

If we want the algorithm to be able to accept ϵ as part of the input, then we must confine ourselves to properties for which $\Psi_{\mathcal{F}}$ is computable. However, as for any reasonable graph property this function is computable, this is not a real constraint. For example, as we have mentioned in Section 4, if $\mathcal{P}$ is the property of being bipartite, then $\Psi_{\mathcal{F}}(k)$ is either k or $k - 1$. Another natural family of properties for which $\Psi_{\mathcal{F}}(k)$ is computable is that of being H-free for a fixed graph H, as in this case $\Psi_{\mathcal{F}}(k) \leq |V(H)|$. By the definition of the function $\mathcal{E}_{3.5}$ we get that if $\Psi_{\mathcal{F}}$ is computable then so is $\mathcal{E}_{3.5}$. It is also not difficult to see that if $\Psi_{\mathcal{F}}$ is computable then so is $\mathcal{H}_{\mathcal{F}}$. Therefore, in case $\Psi_{\mathcal{F}}$ is computable, there is no problem with accepting ϵ as part of the input.

We now turn to prove Theorem 1.2. We note that the above difficulties are also relevant for Corollary 1.2, which applies Theorem 1.2, but we refrain from discussing them again.

Proof of Theorem 1.2: (sketch) As in the previous proof, we focus on the property of being $\mathcal{F}$-

[5]Recall that the size of the graph on which we compute $\mathcal{H}_{\mathcal{F}}$ is the number of partition classes of the $\mathcal{E}$-regular partition, and this number is at most $S_{2.9}(m, \mathcal{E})$, which is bounded by a function of ϵ.

free, where $\mathcal{F}$ is the family of forbidden subgraphs of $\mathcal{P}$. Suppose, as in the previous proof, that G is a large enough graph (in terms of ϵ) as otherwise we can take D to be the entire vertex set of G. Assume, we *implicitly* apply Lemma 2.9 on G and let $\mathcal{A} = \{V_i \mid 1 \leq i \leq k\}$, $\mathcal{B} = \{V_{i,j} \mid 1 \leq i \leq k,\ 1 \leq j \leq l\}$ be the equipartitions returned by the lemma. Let W be a weighted complete graph on k vertices, where $w(i,j) = d(V_i, V_j)$. By Lemma 3.5 we have

$$|E_{\mathcal{F}}(G) - \mathcal{H}_{\mathcal{F}}(W)| \leq \epsilon \ . \tag{8}$$

Let D be a random set of vertices and for $1 \leq i \leq k$ let U_i denote the vertices of D that belong to V_i, and for $1 \leq i \leq k, 1 \leq j \leq l$ let $U_{i,j}$ denote the vertices of D that belong to $V_{i,j}$. Recall that k and l are bounded by functions of ϵ. Using standard Chernoff Bounds (see, e.g., [8]), it is easy to see that if we use a large enough sample of vertices D (but only large enough in terms of ϵ), then with high probability (whp) we will have $|d(V_i, V_{i'}) - d(U_i, U_{i'})| \leq \epsilon$ for any $i < i'$ and $|d(V_{i,j}, V_{i',j'}) - d(U_{i,j}, U_{i',j'})| \leq \epsilon$ for any $i < i'$ and $j \neq j'$. Therefore, if W' is a weighted complete graph on k vertices, where $w(i,j) = d(U_i, U_j)$ then

$$|\mathcal{H}_{\mathcal{F}}(W) - \mathcal{H}_{\mathcal{F}}(W')| \leq \epsilon \ . \tag{9}$$

Furthermore, using Chernoff bounds again, one can show that whp all the pairs $(U_i, U_{i'})$ and $(U_{i,j}, U_{i',j'})$ are as regular as $(V_i, V_{i'})$ and $(V_{i,j}, V_{i',j'})$ (up to ϵ). Therefore, the graph induced by D, denoted G', will have equipartitions $\mathcal{A}', \mathcal{B}'$ satisfying the requirements of Lemma 2.9. This means that

$$|E_{\mathcal{F}}(G') - \mathcal{H}_{\mathcal{F}}(W')| \leq \epsilon \ . \tag{10}$$

As (8), (9) and (10) all hold with high probability for any $\epsilon > 0$, we can thus make sure that with probability at least $1 - \epsilon$, we will have $|E_{\mathcal{F}}(G') - E_{\mathcal{F}}(G)| \leq \epsilon$. This completes the proof. ■

6 Overview of the Proof of Theorem 1.3

For the proof of Theorem 1.3 it will be more convenient to denote by $E'_{\mathcal{P}}(G)$ the number of edge removals needed to make G satisfy $\mathcal{P}$, in other words $E'_{\mathcal{P}}(G) = n^2 \cdot E_{\mathcal{P}}(G)$. In particular, $E'_H(G)$ denotes the number of edge removals needed to turn G into an H-free graph. We will also denote by $E'_r(G)$ the number of edge removals needed to turn G into an r-partite graph (or equivalently r-colorable graph). Note, that approximating $E'_{\mathcal{P}}(G)$ within $n^{2-\delta}$ is equivalent to approximating $E_{\mathcal{P}}(G)$ within $n^{-\delta}$.

The main technical result we need in order to obtain Theorem 1.3 is an extension of some classical results in Extremal Graph Theory. Recall, that Turán's Theorem (see [45]) states that the largest K_{r+1}-free graph on n vertices (K_{r+1} = complete graph on $r+1$ vertices) is precisely the largest r-partite graph on n vertices. Another classical result is the Erdős-Stone-Simonovits Theorem (see [45]), which states that for any graph H of chromatic number $r+1$, the largest H-free graph on n vertices has at most $o(n^2)$ more edges than the largest r-partite graph on n vertices. As any r-partite graph does not contain a copy of a graph of chromatic number $r+1$, the above results can thus be restated as saying that when $H = K_{r+1}$ we have $E'_H(K_n) = E'_r(K_n)$ and that for any H of chromatic number $r+1$ we have $E'_r(K_n) - o(n^2) \leq E'_H(K_n) \leq E'_r(K_n)$.

The main extremal graph-theoretic tool that we use in order to obtain Theorem 1.3 is the following result, which greatly extends one of the main results of [14]. Note, that this result also extends Turán's

Theorem and the Erdős-Stone-Simonovits Theorem as it states that $E'_H(G)$ and $E'_r(G)$ are very close not only when G is K_n but already when G has a sufficiently large minimal degree.

Theorem 6.1 *Let H be a graph of chromatic number $r+1 \geq 3$.*

(i) If there is an edge of H whose removal reduces its chromatic number, then there is constant $\mu = \mu(H) > 0$ such that if $G = (V, E)$ is a graph on n vertices of minimum degree at least $(1-\mu)n$, then $E'_H(G) = E'_r(G)$.

(ii) Otherwise, there are constants $\gamma = \gamma(H) > 0$ and $\mu = \mu(H) > 0$ such that if $G = (V, E)$ is a graph on n vertices of minimum degree at least $(1-\mu)n$, then

$$E'_r(G) - O(n^{2-\gamma}) \leq E'_H(G) \leq E'_r(G).$$

The assertion of this theorem for the special case of H being a triangle is proved in [14] and in a stronger form in [15]. We note that the $n^{2-\gamma}$ term in the second item of the theorem cannot be avoided. Note, that the error term we obtain in the second part of the theorem is better than the error term of the classical Erdős-Stone-Simonovits Theorem. Such improvement of the error term was previously known (see, e.g., [20] and [43]) but only for the case of G being K_n and not for G of sufficiently high minimal degree. The proof of Theorem 6.1 appears in Section 7.

Our second tool in the proof of Theorem 1.3 is certain pseudo-random graphs. An (n, d, λ)-*graph* is a d-regular graph on n vertices all of whose eigenvalues, except the first one, are at most λ in their absolute values. This notation was introduced by the first author in the 80s, motivated by the fact that if λ is much smaller than d, then such graphs have strong pseudo-random properties. In particular, (see, e.g., [8], Chapter 9), in this case the number of edges between any two sets of vertices U and W of G is roughly its expected value, which is $|U||W|d/n$, (see Section 8 for the precise statement). There are many known explicit constructions of (n, d, λ)-graphs that suffice for our purpose here. Specifically, we can use, for example, the graph constructed by Delsarte and Goethals and by Turyn (see [37]). In this graph the vertex set $V(G)$ consist of all elements of the two dimensional vector space over $GF(q)$ (q is any prime power), so G has $n = q^2$ vertices. To define the edges of G we fix a set L of k lines through the origin. Two vertices x and y of the graph G are adjacent if $x - y$ is parallel to a line in L. It is easy to check that this graph is $d = k(q-1)$-regular. Moreover, because it is a strongly regular graph, one can compute its eigenvalues precisely and show that besides the first one they all are either $-k$ or $q - k$. Therefore, by choosing $k = (1-\mu)\frac{q^2}{q-1}$ we obtain an (n, d, λ)-graph with $d = (1-\mu)n$ and $\lambda \leq \sqrt{n}$ (μ will be chosen as the constant from Theorem 6.1).

Given a graph F let F_b denote the b-blowup of F, that is, the graph obtained from F by replacing every vertex $v \in V(F)$ with an independent set I_v, of size b, and by replacing every edge $(u, v) \in E(F)$, with a complete bipartite graph, whose partition classes are the independent sets I_u and I_v. It is not difficult to show (see Claim 8.2) that for any integer r, we have $E'_r(F_b) = b^2 E'_r(F)$. The final piece of notation we need is the Boolean Or, denoted by $G_1 \cup G_2$ of two graphs G_1 and G_2 on the same set of vertices V. Its set of vertices is V, and its set of edges contains all edges of G_1 and all edges of G_2.

Armed with these preparations, we can now outline the proof of Theorem 1.3. Its first part is an easy application of Turán's Theorem for bipartite graphs. The proof of the second part is more interesting. Suppose all bipartite graphs satisfy $\mathcal{P}$, and let $r+1$ (≥ 3) be the minimum chromatic number of a graph that does not satisfy this property. Fix a graph H of chromatic number $r+1$ that does not satisfy $\mathcal{P}$ and let μ be the constant of Theorem 6.1. Consider, first, the case $r \geq 3$.

In this case we show that any efficient algorithm that approximates $E'_{\mathcal{P}}(G)$ up to $n^{2-\delta}$ will enable us to decide efficiently if a given input graph $F = (V(F), E(F))$ is r-colorable. Indeed, given such an F on m vertices, let $b = m^c$ where c is large constant, to be chosen appropriately. Let F_b be the b-blowup of F, and let F' be the vertex disjoint union of r copies of F_b. Let G' be the (n, d, λ)-graph with $d = (1 - \mu)n$ and $\lambda \leq \sqrt{n}$, whose number of vertices n, is at least the number of vertices of F', and not more than four times of that, and identify the vertices of F' with some of those of G'. Let $G = G' \cup F'$ be the Boolean Or of these two graphs. If F is r-colorable, then so is its blowup F_b, and hence in this case F' has a proper r-coloring in which all color classes have the same size. This can be extended to a partition of the vertices of G to r nearly equal color classes, providing an r-colorable subgraph of G (which satisfies $\mathcal{P}$ by our choice of r) that contains all edges of F', and some edges of G' that do not belong to F'. The pseudo-random properties of G' enable us to approximate this number well.

On the other hand, if F is not r-colorable, then any r-colorable subgraph of G misses at least $b^2 r$ edges of F', and, by the pseudo-random properties of G' cannot contain too many edges of this graph that do not belong to F'. With the right choice of c, this will ensure that if we can approximate the number of edges in a maximum r-colorable subgraph of G up to an $n^{2-\delta}$-additive error, this will enable us to know for sure whether F is r-colorable or not. However, by Theorem 6.1, and as the minimum degree of our graph is at least $(1 - \mu)n$, the maximum size of an H-free subgraph of G is very close to the maximum size of an r-colorable subgraph of it, which is therefore also very close to the maximum number of edges in a subgraph of G satisfying $\mathcal{P}$. This implies that approximating well this last quantity is NP-hard. The case $r = 2$ is similar, but here we have to use that the MAX-CUT problem is NP-hard. The full details appear in Section 8.

7 Proof of Theorem 6.1

Throughout this section we will assume that the number of vertices n in our graph is sufficiently large. We first prove the first part of Theorem 6.1, which is an extension of Turán's theorem. To this end, we need a result proved for K_{r+1}-free graphs by Andrásfai, Erdős and Sós [9] and in a more general form by Erdős and Simonovits [21].

Theorem 7.1 ([9],[21]) *Let H be a fixed graph with chromatic number $r+1 \geq 3$ which contains an edge e such that $\chi(H - e) = r$. If G is an H-free graph of order n with minimal degree $\delta(G) > \frac{3r-4}{3r-1}n$ then G is r-colorable.*

We will also need the following simple lemma.

Lemma 7.1 *Let $r \geq 2$ be an integer and suppose G' is an r-partite subgraph of a graph G (which may be empty) such that there are m edges incident to the vertices in $V(G) \backslash V(G')$. Then G has an r-partite subgraph of size at least $e(G') + \frac{r-1}{r}m$.*

Proof: Let $(A'_1, \ldots, A'_r)$ be the partition of G'. Consider an r-partite subgraph Γ of G with parts $(A_1, \ldots, A_r)$ such that $A'_i \subset A_i$ for every i, where we place each vertex $v \in V(G) \backslash V(G')$ in A_i randomly and independently with probability $1/r$. All edges of G' are edges of Γ, and each edge incident to a vertex in $V(G) \backslash V(G')$ appears in Γ with probability $\frac{r-1}{r}$. By linearity of expectation $\mathbb{E}[e(\Gamma)] = e(G') + \frac{r-1}{r}m$, so some r-partite subgraph of G has at least this many edges. ■

In particular, by taking G' to be the empty graph we obtain that every G contains an r-partite subgraph of size at least $\frac{r-1}{r}e(G)$.

Proof of Theorem 6.1 part (i): We prove that $E'_H(G) = E'_r(G)$ for all graphs G on n vertices with minimum degree

$$\delta(G) \geq \left(1 - \frac{3}{4(r-1)(3r-1)}\right) n + 1.$$

Let Γ be the largest (in terms of number of edges) r-partite subgraph of G and let F be the largest H-free subgraph of G. To prove the first part of the theorem one needs to show that $e(F) = e(\Gamma)$. As H is not r-colorable we trivially have $e(F) \geq e(\Gamma)$. In the rest of the proof we establish that $e(\Gamma) \geq e(F)$. First, note that by Lemma 7.1 we have

$$\begin{aligned} e(\Gamma) &\geq \frac{r-1}{r} e(G) \\ &= \frac{r-1}{r}\left(\left(1 - \frac{3}{4(r-1)(3r-1)}\right)n + 1\right) n/2 \\ &= \frac{12r^2 - 16r + 1}{8r(3r-1)} n^2 + \frac{r-1}{2r} n. \end{aligned}$$

If F has a vertex of degree at most $\frac{3r-4}{3r-1}n$ we delete it and continue. We construct a sequence of graphs $F = F_n, F_{n-1}, ...$, where if F_k has a vertex of degree $\leq \frac{3r-4}{3r-1}k$ we delete that vertex to obtain F_{k-1}. Let F' be the final graph of this sequence which has s vertices and minimal degree greater than $\frac{3r-4}{3r-1}s$. Since F' is H-free, by Theorem 7.1, it is r-partite. Therefore we have that

$$\begin{aligned} \frac{r-1}{2r} s^2 &\geq e(F') \geq e(F) - \frac{3r-4}{3r-1}\left(\binom{n+1}{2} - \binom{s+1}{2}\right) \\ &\geq e(\Gamma) - \frac{3r-4}{2(3r-1)}(n^2 - s^2) - \frac{3r-4}{2(3r-1)} n \\ &\geq \frac{12r^2 - 16r + 1}{8r(3r-1)} n^2 - \frac{3r-4}{2(3r-1)}(n^2 - s^2). \end{aligned}$$

This implies that $\frac{s^2}{2r(3r-1)} \geq \frac{n^2}{8r(3r-1)}$ and so $s \geq n/2$.

Let X be the set of $n - s$ vertices which we deleted, i.e., $X = V(G) - V(F')$. By the minimal degree assumption there are at least

$$m \geq \delta(G)|X| - \binom{|X|}{2} \geq \frac{12r^2 - 16r + 1}{4(r-1)(3r-1)} n(n-s) + (n-s) - \frac{(n-s)^2}{2}$$

edges incident with vertices in X. Thus, by Lemma 7.1, the size of the largest r-partite subgraph of G is at least

$$\begin{aligned} e(\Gamma) &\geq e(F') + \frac{r-1}{r} m \geq e(F) - \frac{3r-4}{3r-1}\left(\binom{n+1}{2} - \binom{s+1}{2}\right) + \frac{r-1}{r} m \\ &= e(F) - \frac{3r-4}{2(3r-1)}(n^2 - s^2) - \frac{3r-4}{2(3r-1)}(n-s) + \frac{r-1}{r} m \\ &\geq e(F) - \frac{3r-4}{2(3r-1)}(n^2 - s^2) + \frac{r-1}{r}\left(\frac{12r^2 - 16r + 1}{4(r-1)(3r-1)} n(n-s) - \frac{(n-s)^2}{2}\right) \\ &= e(F) + \frac{(n-s)(2s-n)}{4r(3r-1)} \geq e(F). \end{aligned}$$

This implies that $e(\Gamma) \geq e(F)$ and completes the proof. ■

We turn to prove Theorem 6.1 part (ii). To this end, we first prove the main technical result of this section, Theorem 7.2 below, which is a version of Theorem 7.1 that applies to arbitrary graphs H. We then apply this theorem in order to prove Theorem 6.1 part (ii). The reader may want to note that this application of Theorem 7.2 is similar to the way we applied Theorem 7.1 in order to prove the first part of Theorem 6.1.

Theorem 7.2 *Let H be a fixed graph on h vertices with chromatic number $r+1 \geq 3$ and let G be an H-free graph of order n with minimum degree $\delta(G) \geq \left(\frac{r-1}{r} - \frac{1}{3hr^2}\right) n$. Then one can delete at most $O\left(n^{2-(r+1)/h}\right)$ edges to make G r-colorable.*

Proof: First we need the following weaker bound on $E'_r(G)$.

Claim 7.2 *G can be made r-partite by deleting $o(n^2)$ edges.*

Proof: We use the Regularity Lemma given in Lemma 2.6. For every constant $0 < \eta < \frac{1}{12hr^2}$ let $\gamma = \gamma_{2.3}(\eta, r+1, h) < \eta^2$ be sufficiently small to guarantee that the assertion of Lemma 2.3 holds[6]. Consider a γ-regular partition $(U_1, U_2, \ldots U_k)$ of G. Let G' be a new graph on the vertices $1 \leq i \leq k$ in which (i,j) is an edge iff (U_i, U_j) is a γ-regular pair with density at least η. Since G is an H-free graph and H is homomorphic to K_{r+1} (as $\chi(H) = r+1$), by Lemma 2.3, G' contains no clique of size $r+1$. Call a vertex of G' *good* if there are at most ηk other vertices j such that the pair (U_i, U_j) is not γ-regular, otherwise call it *bad*. Since the number of non-regular pairs is at most $\gamma\binom{k}{2} \leq \eta^2 k^2/2$ we have that all but at most ηk vertices are good. By definition, the degree of each good vertex in G' is at least $\left(\frac{r-1}{r} - \frac{1}{3hr^2}\right) k - 2\eta k - 1$, since deletion of the edges from non-regular pairs and sparse pairs can decrease the degree by at most ηk each and the deletion of edges inside the sets U_i can decrease it by 1. By deleting all bad vertices we obtain a K_{r+1}-free graph on at most k vertices with minimal degree at least

$$\begin{aligned}\left(\frac{r-1}{r} - \frac{1}{3hr^2}\right) k - 3\eta k - 1 &\geq \left(\frac{r-1}{r} - \frac{2}{3hr^2}\right) k \\ &\geq \left(\frac{r-1}{r} - \frac{1}{3r^2}\right) k \\ &> \frac{3r-4}{3r-1} k.\end{aligned}$$

Therefore, by Theorem 7.1, this graph is r-partite. This implies that to make G r-partite we can delete at most $\gamma n^2 + \eta n^2 + (\eta n) \cdot n + k \cdot (n/k)^2 \leq 3\eta n^2 + n^2/k = o(n^2)$ edges. ∎

Consider a partition $(V_1, \ldots, V_r)$ of the vertices of G into r parts which maximizes the number of crossing edges between the parts. Then for every $x \in V_i$ and $j \neq i$ the number of neighbors of x in V_i is at most the number of its neighbors in V_j, as otherwise by shifting x to V_j we increase the number of crossing edges. By Claim 7.2, we have that this partition satisfies that $\sum_i e(V_i) = o(n^2)$. Call a vertex x of G *typical* if $x \in V_i$ and has at most $n/(10hr^2)$ neighbors in V_i. Note that there are at most $o(n)$ non-typical vertices in G and, in particular, every part V_i contains a typical vertex. By definition, the degree of this vertex outside V_i is at least $\left(\frac{r-1}{r} - \frac{1}{3hr^2}\right) n - \frac{n}{10hr^2} > \left(\frac{r-1}{r} - \frac{1}{2hr^2}\right) n$ and at most $n - |V_i|$. Therefore $|V_i| \leq (\frac{1}{r} + \frac{1}{2hr^2})n$. Also note that the number of neighbors in V_i of

[6] Recall that by Comment 2.4 we may assume that $\gamma_{2.3}(\eta, r+1, h) < \eta^2$.

every typical vertex $x \in V_j, j \neq i$ is at least

$$\begin{aligned} d_{V_i}(x) &\geq d(x) - d_{V_j}(x) - (r-2)\max_k |V_k| \\ &\geq \left(\frac{r-1}{r} - \frac{1}{3hr^2}\right)n - \frac{n}{10hr^2} - (r-2)\left(\frac{1}{r} + \frac{1}{2hr^2}\right)n \\ &> \left(\frac{1}{r} - \frac{r-1}{2hr^2}\right)n. \end{aligned} \tag{11}$$

The next claim is an immediate corollary of the above observation.

Claim 7.3 *Let U be a subset of V_j of size at least $(\frac{1}{2r} - \frac{1}{4hr})n$ and let $y_1, \ldots, y_k$ be an arbitrary set of $k \leq r-1$ typical vertices outside V_j. Then, there are at least $\frac{n}{2r(r+1)}$ vertices in U, which are adjacent to all vertices y_i.*

Proof: By definition, there are at most $|V_j| - d_{V_j}(y_i)$ non-neighbors of y_i in V_j and thus there are at most that many vertices in U not adjacent to y_i. Delete from U any vertex, which is not a neighbor of either $y_1, y_2, \ldots, y_k$. The remaining set is adjacent to every vertex y_i and has size at least

$$|U| - \sum_i \big(|V_j| - d_{V_j}(y_i)\big).$$

Since by (11) the degree in V_j of every typical vertex $y_i \notin V_j$ is at least $d_{V_j}(y_i) \geq (\frac{1}{r} - \frac{r-1}{2hr^2})n$, we obtain that the number of common neighbors of $y_1, \ldots, y_k$ in U is at least

$$\begin{aligned} |U| - \sum_i \big(|V_j| - d_{V_j}(y_i)\big) &\geq k\left(\frac{1}{r} - \frac{r-1}{2hr^2}\right)n - k|V_j| + |U| \\ &\geq k\left(\frac{1}{r} - \frac{r-1}{2hr^2}\right)n - k\left(\frac{1}{r} + \frac{1}{2hr^2}\right)n + |U| \\ &\geq |U| - \frac{k}{2hr}n \geq \left(\frac{1}{2r} - \frac{1}{4hr}\right)n - \frac{k}{2hr}n \\ &\geq \left(\frac{1}{2r} - \frac{k+1}{2hr}\right)n \geq \frac{n}{2r} - \frac{n}{2h} \geq \frac{n}{2r(r+1)}. \end{aligned}$$

Here we used that $k+1 \leq r$ and $h \geq r+1$. ■

Claim 7.4 *For every non-typical vertex $x \in V_i$ there are at least $\frac{n^r}{5h(3r^2)^r}$ r-cliques $y_1, \ldots, y_r$ such that $y_j \in V_j$ for all $1 \leq j \leq r$ and all vertices y_j are adjacent to x.*

Proof: Without loss of generality let $i = 1$ and let $x \in V_1$ be a non-typical vertex. Since for every $j \neq 1$ the number of neighbors of x in V_j is at least as large as the number of its neighbors in V_1 we have that

$$\begin{aligned} d_{V_j}(x) &\geq \frac{d_{V_j}(x) + d_{V_1}(x)}{2} \geq \frac{1}{2}\left(\Big(\frac{r-1}{r} - \frac{1}{3hr^2}\Big)n - (r-2)\max_i |V_i|\right) \\ &\geq \frac{1}{2}\left(\Big(\frac{r-1}{r} - \frac{1}{3hr^2}\Big)n - (r-2)\Big(\frac{1}{r} + \frac{1}{2hr^2}\Big)n\right) \\ &\geq \left(\frac{1}{2r} - \frac{1}{4hr}\right)n. \end{aligned} \tag{12}$$

To construct the r-cliques satisfying the assertion of the claim, first observe, that since x is non-typical it has at least $n/(10hr^2)$ neighbors in V_1 and at least $n/(10hr^2) - o(n) > n/(15hr^2)$ of these neighbors are typical. Choose y_1 to be an arbitrary typical neighbor of x in V_1 and continue. Suppose at step $1 \leq k \leq r-1$ we already have a k-clique $y_1, \ldots, y_k$ such that $y_i \in V_i$ for all i and all vertices y_i are adjacent to x. Let U_{k+1} be the set of neighbors of x in V_{k+1}. Then, by (12) we have that $|U_{k+1}| = d_{V_{k+1}}(x) \geq (\frac{1}{2r} - \frac{1}{4hr})n$ and therefore by Claim 7.3 there are at least $\frac{n}{2r(r+1)}$ common neighbors of the vertices y_i in U_{k+1}. Moreover, at least $\frac{n}{2r(r+1)} - o(n) > \frac{n}{3r^2}$ of them are typical and we can choose y_{k+1} to be any of them. Therefore at the end of the process we indeed obtained at least $\frac{n}{15hr^2}(\frac{n}{3r^2})^{r-1} = \frac{n^r}{5h(3r^2)^r}$ r-cliques with the desired property. ■

Claim 7.5 *Each V_i contains at most $O(1)$ non-typical vertices.*

Proof: Suppose that the number of non-typical vertices in V_i is at least $5h^2(3r^2)^r$. Consider an auxiliary bipartite graph F with parts W_1, W_2, where W_1 is the set of some $t = 5h^2(3r^2)^r$ non-typical vertices in V_i, W_2 is the family of all n^r r-element multi-sets of $V(G)$ such that $x \in W_1$ is adjacent to multi-set Y from W_2 iff Y is an r-clique in G with exactly one vertex in every V_j and all vertices of Y are adjacent to x. By the previous claim, F has at least $e(F) \geq t\frac{n^r}{5h(3r^2)^r} = hn^r$ edges and therefore the average degree of a vertex in W_2 is at least $d_{av} = e(F)/|W_2| = e(F)/n^r \geq h$. By the convexity of the function $f(z) = \binom{z}{h}$, we can find h vertices $x_1, \ldots, x_h$ in W_1 such that the number of their common neighbors in W_2 is at least

$$m \geq \frac{\sum_{Y \in W_2} \binom{d(Y)}{h}}{\binom{t}{h}} \geq n^r \frac{\binom{d_{av}}{h}}{t^h} = \Omega\big(n^r\big).$$

Thus we proved that G contains h vertices $X = \{x_1, \ldots, x_h\}$ and a family of r-cliques $\mathcal{C}$ of size $m = \Omega\big(n^r\big)$ such that every clique in $\mathcal{C}$ is adjacent to all vertices in X. Next we need the following well-known lemma which appears first implicitly in Erdős [19] (see also, e.g., [27]). It states that if an r-uniform hypergraph on n vertices has $m = \Omega\big(n^r\big)$ edges, then it contains a complete r-partite r-uniform hypergraph with parts of size h. By applying this statement to $\mathcal{C}$, we conclude that there are r disjoint set of vertices $A_1, \ldots, A_r$ each of size h such that every r-tuple $a_1, \ldots, a_r$ with $a_i \in A_i$ forms a clique which is adjacent to all vertices in X. The restriction of G to $X, A_1, \ldots, A_r$ forms a complete $(r+1)$-partite graph with parts of size h each, which clearly contains H. This contradiction shows that there are less than $5h^2(3r^2)^r = O(1)$ non-typical vertices in V_i and completes the proof of the claim. ■

Having finished all the necessary preparations, we are now ready to complete the proof of Theorem 7.2. Let $h_1 \leq h_2 \leq \ldots \leq h_{r+1}$ be the sizes of the color-classes in an $r+1$ coloring of H. Clearly $h_1 \leq h/(r+1)$. Without loss of generality, suppose that V_1 spans at least $2hn^{2-(r+1)/h}$ edges. By the previous claim, only at most $O(n)$ of these edges are incident to non-typical vertices. Therefore the set of typical vertices in V_1 spans at least $hn^{2-(r+1)/h}$ edges. Then, by the well known result of Kövari, Sós and Turán [36] about Turán numbers of bipartite graphs, V_1 contains a complete bipartite graph $H_1 = K_{h_1,h_2}$ all of whose vertices are typical. If there are at least h_3 typical vertices in V_2 which are adjacent to all vertices of H_1 then we add them to H_1 to form a complete 3-partite graph H_2 with parts of sizes h_1, h_2 and h_3 and continue. We claim that if at step $1 \leq k \leq r-1$ there is a $k+1$-partite graph $H_k \subset \cup_{i=1}^k V_i$ with parts of sizes $h_1, \ldots, h_{k+1}$ all of whose vertices are typical, then we can extend it to the complete $k+2$-partite graph H_{k+1} by adding h_{k+2} typical vertices from V_{k+1} which are adjacent to all vertices of H_k. Indeed, recall that by (11) the number of neighbors in V_{k+1} of every typical vertex $x \in V_i, i \neq k+1$ is at least $d_{V_{k+1}}(x) \geq (\frac{1}{r} - \frac{r-1}{2hr^2})n$. Let $t \leq h$ be the order of H_k. Then, as in Claim 7.3 the number of vertices in V_{k+1} which are adjacent to all vertices of H_k is at least

$$\begin{aligned}
|V_{k+1}| - t\left(|V_{k+1}| - \left(\frac{1}{r} - \frac{r-1}{2hr^2}\right)n\right) &\geq t\left(\frac{1}{r} - \frac{r-1}{2hr^2}\right)n - (t-1)\left(\frac{1}{r} + \frac{1}{2hr^2}\right)n \\
&= \frac{n}{r} - \frac{t(r-1)+t-1}{2hr^2}n \\
&\geq \frac{n}{r} - \frac{t}{2hr}n \geq \frac{n}{r} - \frac{n}{2r} = \frac{n}{2r}
\end{aligned}$$

and thus at least $n/(2r) - O(1) > h_{k+2}$ of these vertices are typical. Continuing the above process $r-1$ steps we obtain a complete $(r+1)$-partite graph with parts of sizes $h_1, \ldots, h_{r+1}$, which clearly contains H. This contradicts our assumption that G is H-free and shows that every V_i spans at most $O\big(n^{2-(r+1)/h}\big)$ edges. Therefore the number of edges we need to delete to make G r-partite is bounded by $\sum_i e(V_i) \leq O\big(n^{2-(r+1)/h}\big)$. This completes the proof. ■

Proof of Theorem 6.1 part (ii): Let H be a fixed graph on h vertices with chromatic number $r+1 \geq 3$. We show that the constants $\gamma(H)$ and $\mu(H)$ in the assertion of the theorem can be chosen to be $(r+1)/h$ and $1/(4hr^2)$ respectively. Let G be an H-free graph of order n with minimal degree $\delta(G) \geq (1 - \frac{1}{4hr^2})n$ and let Γ be the largest r-partite subgraph of G and F be a largest H-free subgraph of G. To prove the second item of the theorem it is enough to show that $e(\Gamma) \leq e(F) \leq e(\Gamma) + O(n^{2-(r+1)/h})$. As H is not r-colorable we trivially have $e(\Gamma) \leq e(F)$. In the rest of the proof we establish that $e(F) \leq e(\Gamma) + O(n^{2-(r+1)/h})$. By Lemma 7.1 we have that

$$e(\Gamma) \geq \frac{r-1}{r}e(G) = \frac{r-1}{r}\left(1 - \frac{1}{4hr^2}\right)n^2/2 = \left(\frac{r-1}{2r} - \frac{r-1}{8hr^3}\right)n^2.$$

If F has a vertex of degree at most $(\frac{r-1}{r} - \frac{1}{3hr^2})n$ we delete it and continue. We construct a sequence of graphs $F = F_n, F_{n-1}, \ldots$, where if F_k has a vertex of degree $\leq (\frac{r-1}{r} - \frac{1}{3hr^2})k$ we delete that vertex to obtain F_{k-1}. Let F' be the final graph of this sequence which has s vertices and minimal degree greater than $(\frac{r-1}{r} - \frac{1}{3hr^2})s$ and let Γ' be the largest r-partite subgraph of F'. Since F' is H-free, Theorem 7.2 implies $e(F') \leq e(\Gamma') + O(n^{2-(r+1)/h})$. Therefore we have that

$$\begin{aligned}
\frac{r-1}{2r}s^2 + o(n^2) &\geq e(F') \geq e(F) - \left(\frac{r-1}{r} - \frac{1}{3hr^2}\right)\left(\binom{n+1}{2} - \binom{s+1}{2}\right) \\
&\geq e(\Gamma) - \left(\frac{r-1}{2r} - \frac{1}{6hr^2}\right)(n^2 - s^2) - O(n) \\
&\geq \left(\frac{r-1}{2r} - \frac{r-1}{8hr^3}\right)n^2 - \left(\frac{r-1}{2r} - \frac{1}{6hr^2}\right)(n^2 - s^2) - o(n^2).
\end{aligned}$$

This implies that

$$\frac{s^2}{6hr^2} \geq \left(\frac{1}{6hr^2} - \frac{r-1}{8hr^3}\right)n^2 - o(n^2) > \left(\frac{1}{6hr^2} - \frac{1}{8hr^2}\right)n^2 = \frac{n^2}{24hr^2}$$

and so $s \geq n/2$.

Let X be the set of $n-s$ vertices which we deleted, i.e., $X = V(G) - V(F')$. By the minimal degree assumption there are at least

$$\begin{aligned} m \geq \delta(G)|X| - \binom{|X|}{2} &\geq \left(1 - \frac{1}{4hr^2}\right) n(n-s) - \frac{(n-s)^2}{2} \\ &= (n-s)\left(\left(\frac{1}{2} - \frac{1}{4hr^2}\right)n + \frac{s}{2}\right) \end{aligned}$$

edges incident with vertices in X. Thus, by Lemma 7.1, the size of the largest r-partite subgraph of G is at least

$$\begin{aligned} e(\Gamma) &\geq e(\Gamma') + \frac{r-1}{r}m \geq e(F') - O\big(n^{2-(r+1)/h}\big) + \frac{r-1}{r}m \\ &\geq e(F) - \left(\frac{r-1}{r} - \frac{1}{3hr^2}\right)\left(\binom{n+1}{2} - \binom{s+1}{2}\right) + \frac{r-1}{r}m - O\big(n^{2-(r+1)/h}\big) \\ &\geq e(F) - \left(\frac{r-1}{2r} - \frac{1}{6hr^2}\right)(n^2 - s^2) + \frac{r-1}{r}m - O\big(n^{2-(r+1)/h}\big) \\ &\geq e(F) - \left(\frac{r-1}{2r} - \frac{1}{6hr^2}\right)(n^2 - s^2) + (n-s)\left(\left(\frac{r-1}{2r} - \frac{r-1}{4hr^3}\right)n + \frac{(r-1)s}{2r}\right) - O\big(n^{2-\frac{r+1}{h}}\big) \\ &= e(F) + \frac{(n-s)(2s - \frac{r-3}{r}n)}{12hr^2} - O\big(n^{2-(r+1)/h}\big) \geq e(F) - O\big(n^{2-(r+1)/h}\big). \qquad \blacksquare \end{aligned}$$

8 Proof of Theorem 1.3

We start with the proof of the first part of Theorem 1.3. If there is a bipartite graph H that does not satisfy $\mathcal{P}$, then, by the known results about the Turán numbers of bipartite graphs proved in [36], there exists a positive $\delta > 0$ such that for any large n, any graph with n vertices and at least $n^{2-\delta}$ edges contains a copy of H. Thus, given a graph G on n vertices, one must delete all its edges besides, possibly, $n^{2-\delta}$ of them, to obtain a subgraph satisfying $\mathcal{P}$. As certainly the edgeless graph satisfies $\mathcal{P}$, this provides the required approximation in this case.

The proof of the second part is more complicated, and requires all the preparations obtained in the previous section. Suppose all bipartite graphs satisfy $\mathcal{P}$, and let $r+1 \geq 3$ be the minimum chromatic number of a graph that does not satisfy this property. Fix a graph H of chromatic number $r+1$ that does not satisfy $\mathcal{P}$. We will show that any efficient algorithm that approximates $E'_{\mathcal{P}}(G)$ up to $n^{2-\delta}$ will enable us to decide efficiently how many edges we need to delete from a given input graph $F = (V(F), E(F))$ to make it r-partite. For $r \geq 3$ this problem contains the r-colorability problem, and for $r = 2$ it is the MAX-CUT problem and therefore it is NP-hard for every $r \geq 2$.

Given a graph F on m vertices such that we need to delete ℓ edges to make it r-partite, let $b = m^c$ where c is a large constant, to be chosen later. Let F_b be the b-blowup of F, and let F' be the vertex disjoint union of r copies of F_b. Let $\mu = \mu(H)$ be the constant from Theorem 6.1 and let G' be the (n, d, λ)-graph with $d = (1-\mu)n$ and $\lambda \leq \sqrt{n}$, described in Section 6. As the integer q in the construction discussed in Section 6 can be a prime power, we can always choose the number of vertices of G', which is q^2, to be at least the number of vertices of F', and not more than 4 times of that. In particular, we have $n = \Theta(rmb) = \Theta\big(m^{c+1}\big)$. Identify the vertices of F' with some of those of G'. Let $G = G' \cup F'$ be the Boolean Or of these two graphs.

Suppose, that instead of adding to F' a pseudo-random graph G', we would put any non-edge of F' in G with probability $1-\mu$. It is easy to see that in this case the expected number of edges, which would be spanned by a set of a vertices that span t edges in F', would be $(1-\mu)\binom{a}{2}+\mu t$. The following claim establishes that this is approximately what we find when we add to F' a pseudo-random graph. We then use this claim to show that we can also estimate $E'_r(G)$ as a function of $\ell = E'_r(F)$.

Claim 8.1 *Let A be a subset of the vertices of G of size a which contains precisely t edges of F'. Then the number of edges of G in A satisfies*

$$(1-\mu)\frac{a^2}{2}+\mu t - O\big(m^2 n^{3/2}\big) \le e_G(A) \le (1-\mu)\frac{a^2}{2}+\mu t + O\big(m^2 n^{3/2}\big).$$

Proof: By construction, the edges of the subgraph of F' induced on the set A form an edge disjoint union of complete bipartite graphs we denote by $\Gamma_i = (U_i, W_i), 1 \le i \le k$. Thus $\sum_i |U_i||W_i| = t$ and the fact that F' is a blowup of r disjoint copies of F, which altogether have rm vertices and at most $r\binom{m}{2}$ edges, implies that $k \le r\binom{m}{2} < rm^2$. The number of edges of G spanned on A is the number of edges of G' inside A, minus the number of edges of G' spanned by the pairs (U_i, W_i), plus the number of edges of F' inside A. To estimate this quantity, we need the well-known fact (see, e.g, Chapter 9 of [8]), that the number of edges between two subsets X, Y of an (n, d, λ)-graph G' satisfies

$$\left| e(X,Y) - \frac{|X||Y|d}{n} \right| \le \lambda\sqrt{|X||Y|}$$

and the fact that in such a graph $\left|e(X) - \frac{d|X|^2}{2n}\right| \le \lambda|X|$. Therefore we obtain that

$$\begin{aligned}
e_G(A) &= e_{G'}(A) - \sum_{i=1}^{k} e_{G'}(U_i, W_i) + t = e_{G'}(A) + \sum_{i=1}^{k}\Big(|U_i|W_i| - e_{G'}(U_i, W_i)\Big) \\
&\ge \frac{d|A|^2}{2n} - \lambda|A| + \sum_{i=1}^{k}\left(|U_i|W_i| - \frac{d}{n}|U_i|W_i| - \lambda\sqrt{|U_i|W_i|}\right) \\
&\ge \frac{d|A|^2}{2n} - \lambda n + \sum_{i=1}^{k}\big(\mu|U_i|W_i| - \lambda n\big) \\
&= (1-\mu)\frac{a^2}{2} + \mu\sum_{i=1}^{k}|U_i|W_i| - (k+1)\lambda n \\
&= (1-\mu)\frac{a^2}{2} + \mu t - O\big(m^2 n^{3/2}\big).
\end{aligned}$$

The upper bound $e_G(A) \le (1-\mu)\frac{a^2}{2} + \mu t + O\big(m^2 n^{3/2}\big)$ can be obtained similarly. ■

Recall that the b-blowup F_b of a graph F, defined in Section 6, is the graph obtained from F by replacing every vertex $v \in V(F)$ with an independent set I_v, of size b, and by replacing every edge $(u,v) \in E(F)$, with a complete bipartite graph, whose partition classes are the independent sets I_u and I_v.

Claim 8.2 *For any graph F and any integer b, we have $E'_r(F_b) = b^2 E'_r(F)$.*

Proof: We start by showing that $E'_r(F_b) \leq b^2 E'_r(F)$. Suppose S is a set of $E'_r(F)$ edges whose removal turns F into an r-colorable graph F'. Suppose we remove from F_b all the edges connecting I_u and I_v for any $(u,v) \in S$. Note, that we thus remove $b^2 E'_r(F)$ edges from F_b. We claim that the resulting graph F'_b is r-colorable. Indeed, let $c : V(F) \mapsto \{1,\ldots,r\}$ be a r-coloring of F' and note that by definition of F'_b, if we color all the vertices of I_v with the color $c(v)$, we get a legal r-coloring of F'. Therefore $E'_r(F_b) \leq b^2 E'_r(F)$.

To see that $E'_r(F_b) \geq b^2 E'_r(F)$, let S be a set of edges whose removal turns F_b into an r-colorable graph, and suppose for every $v \in V(F)$ we randomly pick a single vertex from each of the sets I_v. For every edge of S, the probability that we picked both of its endpoints is b^{-2}, therefore the expected number of edges spanned by these vertices is $|S|/b^2$. As the removal of the edges of S makes F_b r-colorable, this in particular applies to all of its subgraphs. Note, that for any choice of a single vertex from each of the independent sets I_v, the graph they span is isomorphic to F. Thus, any such choice spans at least $E'_r(F)$ of the edges of S. It thus must be the case that $|S|/b^2 \geq E'_r(F_b)$, and the proof is complete. ■

Claim 8.3 *The graph G satisfies*

$$\left| E'_r(G) - \left((1-\mu)\frac{n^2}{2r} + \mu r\ell b^2\right) \right| \leq O(m^2 n^3). \tag{13}$$

Proof: Fix a partition of F into r parts which misses exactly ℓ edges and consider r disjoint copies of F. By taking appropriately different parts in every copy of F we can partition this new graph into r equal parts such that exactly $r\ell$ edges are non-crossing. Since F' is a b-blowup of r disjoint copies of F, this gives a partition of F' into equal parts which misses $r\ell b^2$ edges. We can extend this to a partition of G into r nearly equal sets $V(G) = V_1 \cup \ldots \cup V_r$ which misses exactly $r\ell b^2$ edges of F'. Let t_i be the number of edges of F' inside V_i, then $\sum_i t_i = r\ell b^2$. This, together with Claim 8.1, implies that it is enough to delete at most

$$\begin{aligned}
\sum_{i=1}^{r} e_G(V_i) &\leq \sum_{i=1}^{r} \left((1-\mu)\frac{|V_i|^2}{2} + \mu t_i + O\big(m^2 n^{3/2}\big) \right) \\
&\leq (1-\mu) r \frac{(n/r+1)^2}{2} + \mu \sum_{i=1}^{r} t_i + O\big(m^2 n^{3/2}\big) \\
&= (1-\mu)\frac{n^2}{2r} + \mu r\ell b^2 + O\big(m^2 n^{3/2}\big).
\end{aligned}$$

edges to make G r-partite and hence to satisfy property $\mathcal{P}$.

On the other hand, by Claim 8.2, any partition of F', which is b-blowup of r disjoint copies of F, into r parts misses at least $r\ell b^2$ edges. Therefore for every partition of the vertices of G into r sets there are at least $r\ell b^2$ edges of F' which are non-crossing. Let $V_1 \cup \ldots \cup V_r$ be a partition of $V(G)$ that maximizes the number of crossing edges and let again t_i be the number of edges of F' inside V_i (note that in this case the sets V_i are not necessarily of the same size). Using Claim 8.1, together with the fact that $\sum_i t_i \geq r\ell b^2$ and the Cauchy-Schwartz inequality, we conclude that

$$\begin{aligned}
\sum_{i=1}^{r} e_G(V_i) &\geq \sum_{i=1}^{r} \left((1-\mu)\frac{|V_i|^2}{2} + \mu t_i - O\big(m^2 n^{3/2}\big) \right) \\
&\geq \frac{1-\mu}{2} r \left(\frac{\sum_i |V_i|}{r} \right)^2 + \mu r \ell b^2 - O\big(m^2 n^{3/2}\big) \\
&= (1-\mu)\frac{n^2}{2r} + \mu r \ell b^2 - O\big(m^2 n^{3/2}\big).
\end{aligned}$$

This completes the proof of the claim. ■

We are now ready to complete the proof of Theorem 1.3. Choose the constant c to be sufficiently large so that $2/(c+1) < \min(\delta, \gamma, 1/4)$. Recall, that as we chose $b = m^c$ and $n = \Theta(m^{c+1})$, we have

$$n^{2-\delta} = o(b^2), \qquad n^{2-\gamma} = o(b^2), \qquad m^2 n^{3/2} = o(b^2). \tag{14}$$

Also, as G has minimum degree $(1-\mu)n$ we get from Theorem 6.1, that

$$E'_H(G) \geq E'_r(G) - O(n^{2-\gamma}). \tag{15}$$

As H does not satisfy $\mathcal{P}$ we clearly have $E'_{\mathcal{P}}(G) \geq E'_H(G)$. Combining this with (13), (14) and (15) we get

$$\begin{aligned}
E'_{\mathcal{P}}(G) \geq E'_H(G) &\geq E'_r(G) - O(n^{2-\gamma}) \\
&\geq (1-\mu)\frac{n^2}{2r} + \mu r \ell b^2 - O\big(m^2 n^{3/2}\big) - O\big(n^{2-\gamma}\big) \\
&\geq (1-\mu)\frac{n^2}{2r} + \mu r \ell b^2 - o\big(b^2\big)\ .
\end{aligned}$$

Furthermore, by our choice of r, we get that any r-colorable graph satisfies $\mathcal{P}$, hence we infer from (13) and (14) that

$$\begin{aligned}
E'_{\mathcal{P}}(G) \leq E'_r(G) &\leq (1-\mu)\frac{n^2}{2r} + \mu r \ell b^2 + O\big(m^2 n^{3/2}\big) \\
&\leq (1-\mu)\frac{n^2}{2r} + \mu r \ell b^2 + o\big(b^2\big)\ .
\end{aligned}$$

We thus conclude that $|E'_{\mathcal{P}}(G) - ((1-\mu)\frac{n^2}{2r} + \mu r \ell b^2)| \leq o(b^2)$. Therefore, if one can approximate $E'_{\mathcal{P}}(G)$ in time polynomial in n (and hence also in m) within an additive error of $n^{2-\delta} = o(b^2)$ then one thus efficiently computes an integer L, which is within an additive error of $o(b^2)$ from $(1-\mu)\frac{n^2}{2r} + \mu r \ell b^2$. But as in this case ℓ is precisely the nearest integer to $(L - (1-\mu)\frac{n^2}{2r})/\mu r b^2$, this implies that we can *precisely* compute the number of edge removals, needed in order to turn the input graph F into an r-partite graph. This implies that the problem of approximating $E'_{\mathcal{P}}(G)$ within $n^{2-\delta}$ is NP-hard, and completes the proof of Theorem 1.3.

9 Concluding Remarks and Open Problems

- We have shown that for any monotone graph property $\mathcal{P}$ and any $\epsilon > 0$ one can approximate efficiently the minimum number of edges that have to be deleted from an n-vertex input graph to get a graph that satisfies $\mathcal{P}$, up to an additive error of ϵn^2. Moreover, for any *dense* monotone property, that is, a property for which there are graphs on n vertices with $\Omega(n^2)$ edges that satisfy it, it is NP-hard to approximate this minimum up to an additive error of $n^{2-\delta}$. It will be interesting to obtain similar sharp results for the case of sparse monotone properties. In some of these cases (like the property of containing no cycle, or the property of containing no vertex of degree at least 2) the above minimum can be computed precisely in polynomial time, and in some other cases, a few of which are treated in [12], [13], [46], a precise computation is known to be hard. Obtaining sharp estimates for the best approximation achievable efficiently seems difficult.

- As we have mentioned in Section 1, a special case of Theorem 1.3 implies that for any non-bipartite H, computing the smallest number of edge removals that are needed to make a graph H-free is NP-hard. This is clearly not the case for some bipartite graphs such as a single edge or any star. It will be interesting to classify the bipartite graphs for which this problem is NP-hard.

- It seems interesting to decide if one can obtain a result analogous to Theorem 1.3 for the family of hereditary properties.

- A weaker version of Theorem 1.1 can be derived by combining the results of [7] and [24]. However, this only enables one to approximate $E_{\mathcal{P}}(G)$ within an additive error ϵ in time $n^{f(\epsilon)}$, while the running time of our algorithm is of type $f(\epsilon)n^2$.

- Recall that $E'_{\mathcal{F}}(G)$ denotes the smallest number of edge deletions that are needed in order to make G $\mathcal{F}$-free. For a family of graphs $\mathcal{F}$, let $\nu_{\mathcal{F}}(G)$ denote the $\mathcal{F}$-packing number of G, which is the size of the largest family of edge-disjoint copies of members of $\mathcal{F}$, which is spanned by G. Let $\nu^*_{\mathcal{F}}(G)$ denote the natural Linear Programming relaxation of $\nu_{\mathcal{F}}(G)$. Haxell and Rödl [32] and Yuster [47] have shown that $\nu_{\mathcal{F}}(G) \leq \nu^*_{\mathcal{F}}(G) \leq \nu_{\mathcal{F}}(G) + \epsilon n^2$ for any $\mathcal{F}$ and any $\epsilon > 0$, implying that for any finite $\mathcal{F}$, $\nu_{\mathcal{F}}(G)$ can be approximated within any additive error of ϵn^2 by solving the Linear Program for computing $\nu^*_{\mathcal{F}}(G)$. One may wonder whether it is possible to obtain Theorem 1.1 by solving the natural Linear Programming relaxation of $E'_{\mathcal{F}}(G)$, which we denote by $E^*_{\mathcal{F}}(G)$. Regretfully, this is not the case. Linear Programming duality implies that $E^*_{\mathcal{F}}(G) = \nu^*_{\mathcal{F}}(G)$ and by the results of [32] and [47] we thus have

$$\nu_{\mathcal{F}}(G) \leq E^*_{\mathcal{F}}(G) \leq \nu_{\mathcal{F}}(G) + \epsilon n^2 \ . \tag{16}$$

  Consider now any $\mathcal{F}$, which does not contain the single edge graph and note that we trivially have $\nu_{\mathcal{F}}(K_n) \leq \frac{1}{2}\binom{n}{2} \leq \frac{1}{4}n^2$ (we denote by K_n the n-vertex complete graph). If $\mathcal{F}$ contains a bipartite graph then by the theorem of Kövari, Sós and Turán (see Section 6) we have $E'_{\mathcal{F}}(K_n) > \binom{n}{2} - n^{2-\delta} \geq (\frac{1}{2} - o(1))n^2$. If on the other hand all the graphs in $\mathcal{F}$ are of chromatic number $r \geq 3$ then clearly they all must contain at least $\binom{r}{2}$ edges, and therefore we must have $\nu_{\mathcal{F}}(K_n) \leq \binom{n}{2}/\binom{r}{2} \leq \frac{n^2}{r(r-1)}$. On the other hand, by the theorem of Erdős-Stone-Simonovits (see Section 6) $E'_{\mathcal{F}}(K_n) > \frac{n^2}{2(r-1)} - o(n^2)$. In any case, we have that $\nu_{\mathcal{F}}(K_n) + \delta n^2 \leq E'_{\mathcal{F}}(K_n)$ for some fixed $\delta = \delta(\mathcal{F}) > 0$. Combined with (16) we get that for any $\mathcal{F}$ not containing the single edge graph $E^*_{\mathcal{F}}(K_n) + \delta n^2 < E'_{\mathcal{F}}(K_n)$. Thus, the (trivial) case in which $\mathcal{F}$ contains a single edge is the only one for which computing $E^*_{\mathcal{F}}(G)$ is guaranteed to approximate $E'_{\mathcal{F}}(G)$ within ϵn^2 for any $\epsilon > 0$. In fact, in this degenerate case we actually have $E^*_{\mathcal{F}}(G) = E'_{\mathcal{F}}(G)$.

References

[1] N. Alon, Ranking tournaments, SIAM J. Discrete Math. 20 (2006), 137-142.

[2] N. Alon, R. A. Duke, H. Lefmann, V. Rödl and R. Yuster, The algorithmic aspects of the Regularity Lemma, Proc. 33^{rd} IEEE FOCS, Pittsburgh, IEEE (1992), 473-481. Also: J. of Algorithms 16 (1994), 80-109.

[3] N. Alon, W. Fernandez de la Vega, R. Kannan and M. Karpinski, Random Sampling and Approximation of MAX-CSP Problems, Proc. of 34^{th} ACM STOC, ACM Press (2002), 232-239.

[4] N. Alon, E. Fischer, M. Krivelevich and M. Szegedy, Efficient testing of large graphs, Proc. of 40^{th} FOCS, New York, NY, IEEE (1999), 656–666. Also: Combinatorica 20 (2000), 451-476.

[5] N. Alon and A. Shapira, A separation theorem in property-testing, manuscript.

[6] N. Alon and A. Shapira, A characterization of the (natural) graph properties testable with one-sided error, Proc. of FOCS 2005, 429-438.

[7] N. Alon and A. Shapira, Every monotone graph property is testable, Proc. of 37^{th} STOC 2005, 128-137. Also, SIAM J. on Computing, to appear.

[8] N. Alon and J. H. Spencer, **The Probabilistic Method**, Second Edition, Wiley, New York, 2000

[9] B. Andrásfai, P. Erdős and V. Sós, On the connection between chromatic number, maximal clique and minimal degree of a graph, Discrete Math. 8 (1974), 205-218.

[10] S. Arora, A. Frieze and H. Kaplan, A new rounding procedure for the assignment problem with applications to dense graph arrangement problems, Proc. of 36^{th} FOCS (1996), 21-30. Also, Mathematical Programming 92:1 (2002), 1-36.

[11] S. Arora, D. Karger and M. Karpinski, Polynomial time approximation schemes for dense instances of graph problems, Proc. of 28^{th} STOC (1995). Also, JCSS 58 (1999), 193-210.

[12] T. Asano, An application of duality to edge-deletion problems, SIAM J. on Computing, 16 (1987), 312-331.

[13] T. Asano and T. Hirata, Edge-deletion and edge-contraction problems, Proc. of STOC (1982), 245-254.

[14] J. Bondy, J. Shen, S. Thomassé and C. Thomassen, Density conditions for triangles in multipartite graphs, *Combinatorica*, to appear.

[15] J. Balogh, P. Keevash and B. Sudakov, On the minimal degree implying equality of the largest triangle-free and bipartite subgraphs, submitted.

[16] L. Cai, Fixed-parameter tractability of graph modification problems for hereditary properties, Information Processing Letters, 58 (1996), 171-176.

[17] K. Cirino, S. Muthukrishnan, N. Narayanaswamy and H. Ramesh, graph editing to bipartite interval graphs: exact and asymptotic bounds, Proc. of 17^{th} FSTTCS (1997), 37-53.

[18] E. S. El-Mallah and C. J. Colbourn, The complexity of some edge-deletion problems, IEEE transactions on circuits and systems, 35 (1988), 354-362.

[19] P. Erdős, On extremal problems of graphs and generalized graphs, *Israel J. Math.* 2 (1964), 183–190.

[20] P. Erdős, On some new inequalities concerning extremal properties of graphs, Theory of Graphs (Proc. Colloq., Tihany, 1966), Academic Press, New York, 1968, 77–81.

[21] P. Erdős and M. Simonovits, On a valence problem in extremal graph theory, Discrete Math. 5 (1973), 323-334.

[22] W. Fernandez de la Vega, Max-Cut has a randomized approximation scheme in dense graphs, Random Structures and Algorithms, 8(3) 1996, 187-198.

[23] E. Fischer, The art of uninformed decisions: A primer to property testing, The Computational Complexity Column of The Bulletin of the European Association for Theoretical Computer Science 75 (2001), 97-126.

[24] E. Fischer and I. Newman, Testing versus estimation of graph properties, Proc. of 37^{th} STOC 2005, 138-146.

[25] A. Frieze and R. Kannan, The regularity lemma and approximation schemes for dense problems, Proc. of 37^{th} FOCS, 1996, 12-20.

[26] A. Frieze and R. Kannan, Quick approximation to matrices and applications, Combinatorica, 19(2), 1999, 175-220.

[27] Z. Füredi, Turán type problems, in: *Surveys in combinatorics*, London Math. Soc. Lecture Note Ser. 166, Cambridge Univ. Press, Cambridge, 1991, 253–300

[28] M.R. Garey and D.S. Johnson, Computers and Intractability: A guide to the Theory of NP-Completeness, W.H. Freeman and Co., San Francisco, 1979.

[29] P. W. Goldberg, M. C. Golumbic, H. Kaplan and R. Shamir, Four strikes against physical mapping of DNA, Journal of Computational Biology 2 (1995), 139–152.

[30] O. Goldreich, Combinatorial property testing - a survey, In: Randomization Methods in Algorithm Design (P. Pardalos, S. Rajasekaran and J. Rolim eds.), AMS-DIMACS (1998), 45-60.

[31] M. C. Golumbic, H. Kaplan and R. Shamir, On the complexity of DNA physical mapping, Advances in Applied Mathematics, 15 (1994), 251-261.

[32] P. E. Haxell and V. Rödl, Integer and fractional packings in dense graphs, Combinatorica 21 (2001), 13-38.

[33] S. Khot and V. Raman, Parameterized complexity of finding subgraphs with hereditary properties, COCOON 2000, 137-147.

[34] Y. Kohayakawa, V. Rödl and L. Thoma, An optimal algorithm for checking regularity, SIAM J. on Computing 32 (2003), no. 5, 1210-1235.

[35] J. Komlós and M. Simonovits, Szemerédi's Regularity Lemma and its applications in graph theory. In: *Combinatorics, Paul Erdös is Eighty*, Vol II (D. Miklós, V. T. Sós, T. Szönyi eds.), János Bolyai Math. Soc., Budapest (1996), 295–352.

[36] T. Kövari, V.T. Sós and P. Turán, On a problem of K. Zarankiewicz,*Colloquium Math.* 3 (1954), 50-57.

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

[38] J. Lewis and M. Yannakakis, The node deletion problem for hereditary properties is NP-complete, JCSS 20 (1980), 219-230.

[39] A. Natanzon, R. Shamir and R. Sharan, Complexity classification of some edge modification problems, Discrete Applied Mathematics 113 (2001), 109–128.

[40] M. Parnas, D. Ron and R. Rubinfeld, Tolerant property testing and distance approximation, manuscript, 2004.

[41] D. Ron, Property testing, in: P. M. Pardalos, S. Rajasekaran, J. Reif and J. D. P. Rolim, editors, *Handbook of Randomized Computing*, Vol. II, Kluwer Academic Publishers, 2001, 597–649.

[42] J. D. Rose, A graph-theoretic study of the numerical solution of sparse positive-definite systems of linear equations, Graph Theory and Computing, R.C. Reed, ed., Academic Press, N.Y., 1972, 183-217.

[43] M. Simonovits, A method for solving extremal problems in graph theory, stability problems, Theory of Graphs (Proc. Colloq., Tihany, 1966), Academic Press, New York, 1968, 279–319.

[44] E. Szemerédi, Regular partitions of graphs, In: *Proc. Colloque Inter. CNRS* (J. C. Bermond, J. C. Fournier, M. Las Vergnas and D. Sotteau, eds.), 1978, 399–401.

[45] D. B. West, **Introduction to Graph Theory**, Prentice Hall, 2001.

[46] M. Yannakakis, Edge-deletion problems, SIAM J. Comput. 10 (1981), 297-309.

[47] R. Yuster, Integer and fractional packing of families of graphs, Random Structures and Algorithms 26 (2005), 110-118.

[48] J. Xue, Edge-maximal triangulated subgraphs and heuristics for the maximum clique problem. Networks 24 (1994), 109-120