Electronic Colloquium on Computational Complexity, Report No. 21 (2021)

Average-Case Perfect Matching Lower Bounds from Hardness of Tseitin Formulas*

Per Austrin and Kilian Risse

KTH Royal Institute of Technology

February 18, 2021

Abstract

We study the complexity of proving that a sparse random regular graph on an odd number of vertices does not have a perfect matching, and related problems involving each vertex being matched some pre-specified number of times. We show that this requires proofs of degree $\Omega(n/\log n)$ in the Polynomial Calculus (over fields of characteristic $\neq 2$) and Sum-of-Squares proof systems, and exponential size in the bounded-depth Frege proof system. This resolves a question by Razborov asking whether the Lovász-Schrijver proof system requires n^δ rounds to refute these formulas for some $\delta > 0$. The results are obtained by a worst-case to average-case reduction of these formulas relying on a topological embedding theorem which may be of independent interest.

*Supported by the Approximability and Proof Complexity project funded by the Knut and Alice Wallenberg Foundation.

ISSN 1433-8092

1 Introduction

Proof complexity is the study of certificates of unsatisfiability, initiated by Cook and Reckhow [CR79] as a program to separate **NP** from **coNP**. The main goal of this program is to prove size lower bounds on proofs of unsatisfiability of logical formulas. This is a daunting job – indeed we are far from proving general size lower bounds on certificates of unsatisfiability. As an intermediate step we study proof systems with restricted deductive power and prove size lower bounds for such restricted certificates of unsatisfiability. The most studied such proof system is resolution [Bla37] which is fairly well understood by now, see e.g., the proof complexity book by Krajíček [Kra19].

But resolution is by far the only proof system. A closely related and quite general proof system is the bounded depth Frege proof system [CR79] which manipulates propositional formulas of bounded depth. While we have some results for the bounded depth Frege proof system, in this introduction we instead focus on two other systems as these were the primary motivation behind our work. These are the two (incomparable) Polynomial Calculus (PC) [CEI96, ABSRW04] and Sum-of-Squares (SoS) [Sho87, Par00, Las01] proof systems. These proof systems do not rely on propositional logic, like resolution or Frege, but rather on algebraic reasoning and are examples of so-called semi-algebraic proof systems (see e.g. [GHP02]).

Both PC and SoS provide refutations of (satisfiability of) a set of polynomial equations $\mathcal{Q} = \{ q_i(x) = 0 \mid i \in [m] \}$ over n variables $x_1, \ldots, x_n$. In the case of PC, these polynomials can be over any field $\mathbb{F}$ (finite or infinite), and in the case of SoS, these polynomials are over $\mathbb{R}$. A key complexity measure of a $\mathrm{PC}_{\mathbb{F}}$ or SoS refutation of $\mathcal{Q}$ is its degree, defined as the maximum degree of any polynomial appearing in the refutation. The degree of refuting $\mathcal{Q}$ in $\mathrm{PC}_{\mathbb{F}}$ or SoS, which we denote by $\mathsf{Deg}(\mathcal{Q} \vdash_{\mathrm{PC}_{\mathbb{F}}} \bot)$ and $\mathsf{Deg}(\mathcal{Q} \vdash_{\mathrm{SoS}} \bot)$ respectively, is the minimum degree of any $\mathrm{PC}_{\mathbb{F}}$ or SoS refutation of $\mathcal{Q}$. For Boolean systems of equations, meaning that $\mathcal{Q}$ contains the equations $x_i^2 - x_i = 0$ for all $i \in [n]$, strong enough degree lower bounds imply size lower bounds in both $\mathrm{PC}_{\mathbb{F}}$ [CEI96, IPS99] and SoS [AH19], where the size of a refutation is the total number of monomials appearing in it.

There is by now a large number of lower bound results for both PC, e.g., [Raz98, IPS99, BGIP01, AR01, GL10, MN15], and SoS, e.g., [Gri01, Sch08, MPW15, BHK+16, KMOW17, AH19, Pot20, AGK20], with SoS in particular having received considerable attention in recent years due to its close connection to the Sum-of-Squares hierarchy of semidefinite programming, a powerful "meta-algorithm" for combinatorial optimization problems [BS14].

In this paper we study the power (or lack thereof) of these proof systems when it comes to refuting the perfect matching formula $\mathsf{PM}(G)$ defined over sparse random graphs $G = (V, E)$ on an odd number of vertices. This formula can be viewed as a system of linear equations over $\mathbb{R}$ on a set of Boolean variables: for each edge $e \in E$ there is a variable $x_e \in \{0, 1\}$ (indicating whether the edge is used in the matching) and for each vertex $v \in V$ there is an equation $\sum_{e \ni v} x_e = 1$. Apart from being a natural well-studied problem on its own, the perfect matching formula is interesting because of its close relation to two other widely studied families of formulas, namely the pigeonhole principle (PHP), and Tseitin formulas.

PHP asserts that m pigeons cannot fit in $n < m$ holes (where each hole can fit at most one pigeon). This can be viewed as a bipartite matching problem on the complete bipartite graph with $m + n$ vertices, where each vertex on the large side (with m vertices) must be matched at least once, and each vertex on the small side (with n vertices) can be matched at most once. There are many variants of PHP (see e.g. the survey [Raz02]), and the one closest to the perfect matching formula is the so-called "onto functional PHP", in which each vertex on both sides must be matched exactly once (rather than at least/at most once). Equivalently, this formula is simply the perfect matching formula on a complete bipartite graph with $n + m$ vertices. While

most variants of PHP are hard for PC [Raz98, MN15], the onto functional PHP variant is in fact easy to refute in PC over any field [Rii93]. In SoS, all variants of PHP are easy to refute [GHP02].

The Tseitin formula over a graph G claims that there is a subgraph of G such that each vertex has odd degree. As the sum of the degrees of a graph is even, this formula is not satisfiable if G has an odd number of vertices. In contrast to the PHP, the Tseitin formula is (almost) always hard: for $\mathrm{PC}_{\mathbb{F}}$ over fields $\mathbb{F}$ of characteristic distinct from 2 [BGIP01, AR01] and SoS [Gri01] these formulas require linear degree if G is a good vertex expander. We cannot hope to prove degree lower bounds over fields of characteristic 2 as the constraints become linear and we can thus refute the Tseitin formula using Gaussian elimination. As the perfect matching formula $\mathsf{PM}(G)$ implies the Tseitin formula, PC over fields of characteristic 2 can also easily refute $\mathsf{PM}(G)$ for G with an odd number of vertices.

In summary, the perfect matching formula lies somewhere in between PHP and Tseitin, of which the former is easy to refute in SoS (and easy to refute in PC in the onto functional variant), and the latter is hard to refute in SoS (as well as in PC with characteristic $\neq 2$). Hence it is natural to wonder whether SoS or PC requires large degree to refute the perfect matching formula over non-bipartite graphs.

The case of perfect matching in the *complete graph* on an odd number of vertices (sometimes called the "MOD 2 principle") is well-understood in both PC [BGIP01] and SoS [Gri01, Pot17], requiring degree $\Omega(n)$ in both proof systems unless the underlying field of PC is of characteristic 2. For sparse graphs, less is known. Buss et al. [BGIP01] obtained worst-case lower bounds in PC showing that there exist bounded degree graphs on n vertices requiring $\Omega(n)$ degree refutations. This is obtained by a reduction from Tseitin formulas and while the work of Buss et al. outdates the current interest in the SoS system, it is not hard to see that the same reduction yields a similar $\Omega(n)$ degree lower bound for SoS (details provided in Appendix A).

However, for random graphs G little is known about the hardness of the perfect matching formula and, e.g., Razborov [Raz17] asked whether it is true that the Lovász-Schrijver hierarchy [LS91] (which is weaker than SoS) requires n^{ε} rounds to refute the perfect matching principle on a random sparse regular graph with high probability.

1.1 Our results

We show that indeed the perfect matching principle requires large size on random d-regular graphs (for some constant d) in the Sum-of-Squares, Polynomial Calculus, and bounded-depth Frege proof systems. Our results apply more generally to Tseitin-like formulas defined by linear equations over the reals induced by some graph, so let us now define these.

For a graph $G = (V, E)$ and integer vector $b \in \mathbb{Z}^V$, consider the system of linear equations over the reals having a variable x_e for each $e \in E$, and the equation $\sum_{e \ni v} x_e = b_v$ for each $v \in V$. Let $\mathrm{Card}(G, b)$ denote this system of linear equations along with the Boolean constraints $x_e \in \{0, 1\}$ (viewed as a quadratic equation $x_e^2 - x_e = 0$) for each edge – in Section 2.2 the encoding is discussed in more detail. Note that $\mathrm{Card}(G, \vec{1})$ corresponds to the perfect matching problem in G and in general $\mathrm{Card}(G, b)$ can be viewed as asserting that G has a "matching" where each vertex is matched exactly b_v times. Note that whenever $\sum_{v \in V} b_v$ is odd, $\mathrm{Card}(G, b)$ is unsatisfiable (since the equations imply $\sum_v b_v = 2 \sum_e x_e$ which is even)[1].

We focus on the special case of $\mathrm{Card}(G, b)$ where G is d-regular and $b = \vec{t} = (t, t, \ldots, t)$ is the all-t vector for some $t \in [d]$. If in this scenario both n and t are odd (implying d is even) then as observed above $\mathrm{Card}(G, \vec{t})$ is unsatisfiable. On the other hand if n is odd and t is even

[1] As pointed out to us by Aleksa Stanković, decidability of $\mathrm{Card}(G, b)$ is in polynomial time as a non-optimal solution always has an augmenting path along which it can be improved.

then $\mathrm{Card}(G, \vec{t})$ is always satisfiable (because such G admits a 2-factorization). The remaining case when n is even may be either satisfiable or unsatisfiable, but for a random d-regular G with $d \geq 3$, $\mathrm{Card}(G, t)$ will be satisfiable with high probability (because such G can be partitioned into perfect matchings with high probability).

If we let $\mathscr{F}_D$ denote a Frege system restricted to depth-D formulas (see Section 2.1), then our main theorem is as follows.

Theorem 1.1. There is a constant d_0 such that for all constants $d \geq d_0$ and $t \in [d]$, the following holds with high probability over a random d-regular graph G on n vertices.

1. $\mathsf{Deg}(\mathrm{Card}(G, \vec{t}) \vdash_{\mathrm{PC}_{\mathbb{F}}} \bot) = \Omega(n/\log n)$ whenever $\mathrm{char}(\mathbb{F}) \neq 2$.
2. $\mathsf{Deg}(\mathrm{Card}(G, \vec{t}) \vdash_{\mathrm{SoS}} \bot) = \Omega(n/\log n)$.
3. There is a $\delta > 0$ such that $\mathsf{Size}(\mathrm{Card}(G, \vec{t}) \vdash_{\mathscr{F}_D} \bot) = \exp\left(\Omega(n^{\delta/D})\right)$, for all $D \leq \frac{\delta \log n}{\log \log n}$.

The interesting case of the above theorem is when both n and t are odd so that $\mathrm{Card}(G, t)$ is unsatisfiable; in the other cases $\mathrm{Card}(G, \vec{t})$ is satisfiability with high probability and the lower bounds are vacuous.

By known size-degree tradeoffs for Polynomial Calculus [IPS99, CEI96] and Sum-of-Squares [AH19] the degree lower bounds in Theorem 1.1 imply near-optimal size lower bounds of $\exp\left(\Omega(n/\log^2 n)\right)$.

Apart from the perfect matching formula, another special case of $\mathrm{Card}(G, \vec{t})$ is the so-called even coloring formula, introduced by Markström [Mar06], which is the case when $t = \deg(v)/2$. An open problem of Buss and Nordström ([BN21], Open Problem 7.7) asks whether these formulas are hard on expanders for Polynomial Calculus over fields of characteristic $\neq 2$. Theorem 1.1 partially resolves this open problem, establishing that it is hard on random graphs (rather than on all expanders).

We will give a more detailed overview of how the results are obtained in Section 1.3 below, but for now let us mention that we obtain them using embedding techniques. In particular for, say, the SoS lower bound, our starting point is the $\Omega(n)$ *worst-case* degree lower bound in sparse graphs, and we then prove that these hard instances can be embedded in a random d-regular graph in such a way that the hardness of refuting the formula is preserved.

To achieve this, one of the components we need is a new graph embedding theorem which may be of independent interest. Very loosely speaking, we show that any bounded-degree graph with $O(n/\log n)$ edges can be embedded as a *topological minor* in any bounded-degree α-expander on n vertices and sufficiently many edges. In addition, for our application to perfect matching (and more generally the $\mathrm{Card}(G, \vec{t})$ formulas), we need to be able to control the parities of the path lengths used in the topological embedding, and we show that as long as large linear sized subgraphs contain odd cycles of length $\Omega(1/\alpha)$, this is indeed possible.

Somewhat informally, we prove the following.

Theorem 1.2 (Informal statement of Theorem 3.3)**.** Let G be a constant degree α-expander on n vertices. If H is a graph with at most $\frac{\varepsilon n}{\log n}$ edges and $\Delta(H) \ll \alpha^2 \cdot d(G)$, then G contains H as a topological minor. Furthermore, if all large vertex induced subgraphs of G contain odd cycles of length $\Omega(1/\alpha)$, then one can choose the parities of the length of all the paths in the embedding of H.

This generalizes various classical results of a similar flavor (e.g. [KR96, KN19, CN19, Kri19]). See the next subsection for a discussion comparing these (and other) existing embedding results to ours.

As a further illustration of the applicability of this theorem we partially resolve a question of Filmus et al. [FLM+13]. They prove that with high probability for random d-regular graphs G, where $d \geq 4$, PC requires *space* $\Omega(\sqrt{n})$ to refute the Tseitin formula, and conjecture that PC in fact requires space $\Omega(n)$. On the other hand, Galesi et al. [GKT19] considered it plausible that the $\Omega(\sqrt{n})$ bound is optimal. We (almost) resolve this question by proving $\Omega(n/\log n)$ space lower bounds for the Tseitin formula defined on vertex expanders, but only of large enough (constant) average degree.

Theorem 1.3. For all $\alpha > 0$ there is a d_0 such that the following holds. Let G be a constant degree α-expander of average degree at least d_0. Then over any field $\mathbb{F}$ it holds that $\mathrm{PC}_{\mathbb{F}}$ requires space $\Omega(n/\log n)$ to refute the Tseitin formula defined on G.

Unlike Theorem 1.1, vertex expansion is sufficient and we require no randomness. This lower bound is obtained by embedding a worst-case instance, due to Filmus et al., into a vertex expander. We provide more details in Section 6.1.

1.2 Related work

Proof Complexity Lower Bounds Using Embedding Techniques To the best of the authors knowledge there are only two other papers that consider embedding arguments in proof complexity, both in connection with Frege systems rather than semi-algebraic systems like PC or SoS. The first one is the paper by Pitassi et al. [PRST16], showing that the Tseitin formula is hard for Frege up to depth $\Theta(\sqrt{\log n})$. This paper was followed up by a paper of Håstad [Hå17] improving the previous lower bounds from super-polynomial to exponential and pusing the depth to $\Theta(\log n/\log\log n)$. This was in turn followed up by Galesi et al. [GIRS19] relating the hardness of Tseitin to treewidth, again using embedding arguments.

Connection to Constraint Satisfaction Problems For a k-ary predicate $P : \{0,1\}^k \to \{0,1\}$, an instance of the $\mathrm{CSP}(P)$ problem consists of a set of constraints over n Boolean variables $x_1, \ldots, x_n$, each constraint being an application of P on a list of k variables. The $\mathrm{Card}(G, \vec{t})$ formulas we study can be viewed as instances of $\mathrm{CSP}(P)$ where each variable appears in exactly 2 constraints and $P : \{0,1\}^d \to \{0,1\}$ is the constraint that exactly t of the d inputs are 1.

CSP problems have been extensively studied throughout the years, and fairly general conditions under which $\mathrm{CSP}(P)$ is hard for PC and SoS are known [AR01, KMOW17]. To be more accurate, these results are for the more general $\mathrm{CSP}(P^{\pm})$ problem in which each constraint is an application of P on k *literals* rather than variables. In particular, Alekhnovich and Razborov [AR01] showed that if P is, say, 8-immune[2] over the underlying field $\mathbb{F}$, then any $\mathrm{PC}_{\mathbb{F}}$ refutation of a random $\mathrm{CSP}(P^{\pm})$ instance with a linear number of constraints requires degree $\tilde{\Omega}(n)$. For SoS, Kothari et al. [KMOW17] showed that, if there exists a pairwise uniform distribution[3] μ over $\{0,1\}^k$ supported on satisfying assignments of P, then with high probability a random $\mathrm{CSP}(P^{\pm})$ instance on $m = \Delta n$ constraints needs degree $\tilde{\Omega}(n/\Delta^2)$ to be refuted by the SoS proof system.

The predicates we study are linear equations over $\mathbb{R}$ and are neither immune nor do they support a pairwise uniform distribution. As such, our results provide CSP lower bounds that fall outside the immunity and pairwise independence frameworks, which are the source of a majority

[2] P is r-immune over $\mathbb{F}$ if there is no degree-r polynomial $q : \{0,1\}^k \to \mathbb{F}$ such that for all satisfying assignments $\alpha \in \{0,1\}^k$ of P it holds that $q(\alpha) = 0$)

[3] A distribution μ over $\{0,1\}^k$ is said to be pairwise uniform if for all $1 \leq i < j \leq k$, the marginal distribution of μ restricted to coordinates i and j is uniform.

of existing CSP lower bounds in PC and SoS. To the authors best knowledge the only other attempt to overcome this framework in the average-case setting is the paper by Deshpande et al. [DMO+19], showing lower bounds for the basic SDP of random regular instances of $\mathrm{CSP}(\mathrm{NAE}_3^{\pm})$, where NAE_3 is the not-all-equal predicate on three bits. In contrast to their work we show (almost) linear degree lower bounds for the stronger Sum-of-Squares hierarchy, but only for a very wide predicate of some large (but constant) arity.

Embedding Theorems There is a rich literature on embeddings of graphs as minors or topological minors into expander graphs. We focus here on the ones most closely related to Theorem 1.2.

The classical result of Kleinberg and Rubinfeld [KR96] shows that a regular expander G on n vertices contains every graph H with $O(n/\mathsf{polylog}(n))$ vertices and edges as a minor. Krivelevich and Nenadov [KN19] simplified and strengthened this by improving the bound on the size of H to $O(n/\log n)$. These results differ from ours in two key ways: (i) we want topological minors, and (ii) we want to be able to control the parities of the path lengths in the embedding. We now discuss these two aspects separately.

Results on topological minors, while somewhat less common, also exist. A result similar to ours is the result of Broder et al. [BFSU96] that with high probability the random graph $\mathcal{G}(n,m)$ on n vertices and $m = \Omega(n \log n)$ edges contains any graph H with $\Delta(H) = O(m/n)$ and at most $O(n/\log n)$ edges (and at most $n/2$ vertices) as a topological minor.

For our second property, the possibility to choose the parities of the paths used in the topological embedding, we are not aware of any previous work studying this question. A related notion are so called *odd minors* which are more general than topological minors with odd length paths. This notion has been considered in connection with a strengthening of Hadwiger's Conjecture, see e.g., the survey by Seymour [Sey16]. This line of research mostly considers complete odd minors, e.g., [GGR+09], and thus is not directly applicable to our situation.

Recently Draganić et al. [DKN20] independently obtained a new embedding theorem similar to ours. They assume the somewhat stronger property that the host graph G is a spectral expander but also obtain a stronger conclusion: each path of the topological embedding is of equal (odd) length and the embedding even works in an adversarial setting. Namely, the adversary is allowed to fix the embedding of the vertices, as long as no neighborhood in G contains too many vertex embeddings.

Extended Formulations There has been a fair amount of work studying the *extension complexity* of the perfect matching polytope [Yan88, Rot17], but these lower bounds do not have any direct implications for the PC and SoS degree of the perfect matching formula. Let us elaborate.

Suppose we have a convex polytope $\mathcal{P}$ consisting of many facets. A natural question is whether there is simpler polytope $\mathcal{Q}$ in a higher dimensional space so that $\mathcal{P}$ is the "shadow" of $\mathcal{Q}$, or a bit more formal that there is a linear projection π such that $\pi(\mathcal{Q}) = \mathcal{P}$. Such a $\mathcal{Q}$ is then called a linear extension of $\mathcal{P}$ and the extension complexity of a polytope $\mathcal{P}$ is the minimum number of facets of any linear extension of $\mathcal{P}$.

Rothvoss [Rot17] proved that the perfect matching polytope of a complete n-node graph has extension compexity $\exp(\Omega(n))$ for n even. This result is incomparable to our lower bounds: as the graphs we consider do not contain a perfect matching, their perfect matching polytope is empty and thus has extension complexity 0. Furthermore, rather than linear programs, i.e., polytopes, we consider semidefinite programs which are more expressive.

1.3 Overview of Proof Techniques

As previously mentioned, our high level approach is to first obtain worst-case perfect matching lower bounds and to then embed these into the $\mathrm{Card}(G, \vec{t})$ formula for G a random regular graph. The worst-case lower bounds are obtained by a gadget reduction from Tseitin to perfect matching, due to Buss et al. [BGIP01]. Using known lower bounds for the Tseitin formula in the corresponding proof systems [BGIP01, Gri01, Hå17] we then obtain the desired worst-case lower bounds for the perfect matching formula.

A naïve attempt to obtain average-case lower bounds from a sparse worst-case instance H on n vertices is to topologically embed the worst-case instance into a random regular graph G on $O(n \log n)$ vertices using Theorem 1.2. One would then like to argue that $\mathsf{PM}(G)$ is hard. But having a worst-case instance as a topological minor is *not* sufficient to conclude that $\mathsf{PM}(G)$ is hard. For instance G may contain an isolated vertex and it is then trivial to refute $\mathsf{PM}(G)$. On the other hand if we could guarantee that there is a perfect matching m in the subgraph of G induced by the vertices *not* used in the embedding of H, we can conclude that $\mathsf{PM}(G)$ is hard: hit the formula with the restriction corresponding to the matching m and we are basically left with the worst-case formula.

Thus if we can ensure that H is a topological minor of G with the two additional properties that (i) every path used in the embedding of H has odd length, and (ii) there exists a perfect matching in the subgraph of G induced by the vertices *not* used in the embedding of H, then we obtain average-case lower bounds for the perfect matching formula $\mathsf{PM}(G) \equiv \mathrm{Card}(G, \vec{1})$. The lower bounds for $\mathrm{Card}(G, \vec{t})$ for $t > 1$ can then be obtained by a reduction to the $t = 1$ case: after fixing the value of the edges in $\lfloor t/2 \rfloor$ cycle covers of G to 1, a restriction of $\mathrm{Card}(G, \vec{t})$ is obtained which behaves like $\mathrm{Card}(G', \vec{1})$ for a somewhat sparser random regular graph G'.

Let us elaborate a bit further on the properties required from the topological minor of H in G. As mentioned previously, our embedding theorem can ensure that all paths are of odd length. To ensure the second property, we in fact do not embed H directly into G but rather into a suitably chosen vertex induced subgraph $G[T]$ with the crucial property that for any set of vertices $U \subseteq T$ of odd cardinality the induced subgraph $G[V \setminus U]$ has a perfect matching. As the embedding of H will consist of an odd number of vertices we then obtain property (ii) above. Since we now want to apply Theorem 1.2 not to G but to $G[T]$, we have to ensure that $G[T]$ satisfies all the conditions of that theorem. We prove what we refer to as the Partition Lemma, which asserts that an induced subgraph $G[T]$ exists that satisfies both the perfect matching property described above, as well as all conditions of Theorem 1.2. The proof of the Partition Lemma relies primarily on the Lovász Local Lemma and spectral bounds to obtain the desired properties.

For the proof of our embedding theorem (Theorem 1.2), we extend an argument due to Krivelevich and Nenadov [KN19] (see also [Kri19]) for ordinary minors (rather than topological minors). In order to obtain a minor embedding of H in G, the idea there is to embed the vertices one by one from H in G while maintaining an "unused" subgraph G' of G which is a slightly worse expander than G is. During this process it may happen that some vertex embedding cannot be connected to a neighbor. If this happens, the embedding of that vertex is removed and it needs to be embedded again.

In order to obtain topological embeddings, we need to adapt this procedure. Since we now want vertex-disjoint paths connecting the embedded vertices, we would ideally like to embed each vertex of H as a large star, and then embed the edges of H as paths connecting different leaves of these stars. In order to make this work out, rather than embedding the vertices as actual stars, we embed them as "star-like" subgraphs of G (more precisely defined in Definition 5.6) that consist of a central vertex connected to many large vertex-disjoint connected subgraphs

of G and show (Lemma 5.7) that we can always embed the vertices of H as such "star-like" subraphs of G.

With this in place, obtaining control of the parities of the path lengths used in the embedding (under the assumption on odd cycles in Theorem 1.2) is relatively straightforward: almost by definition, when embedding an edge of H into a path of G, we can route it via an odd cycle and can then choose which of the two halves of the odd cycles to use, obtaining two possible embeddings with different path length parity, and can choose the one with the appropriate parity.

1.4 Organization

We give some preliminaries in Section 2, formally defining the used proof systems and encodings used, and recalling some general background results. In Section 3 we provide most of the proof of Theorem 1.1 while deferring the proofs of two key results, the aforementioned Partition Lemma and our embedding theorem. The proof of the Partition Lemma is given in Section 4, and the proof of the embedding theorem can be found in Section 5.

In Appendix A we recall the reduction of Buss et al. [BGIP01] from Tseitin to perfect matching and show that it yields lower bounds not only for Polynomial Calculus but also for Sum-of-Squares and bounded depth Frege.

Acknowledgements. The authors are grateful to Susanna de Rezende, Johan Håstad, Jakob Nordström, Dmitry Sokolov and Aleksa Stanković for helpful discussions. In particular we would like to thank Jakob Nordström who suggested to consider the even coloring formula and brought the problem about Polynomial Calculus space to our attention.

2 Preliminaries

Natural logarithms (base e) are denoted by ln, whereas base 2 logarithms are denoted by log. For integers $n \geq 1$ we introduce the shorthand $[n] = \{1, 2, \ldots, n\}$ and sometimes identify singeletons $\{u\}$ by the element u. For a set U we denote the power set of U by 2^U and a transversal A of a family of sets $\mathcal{B} = \{B_1, B_2, \ldots B_n\}$ is a set such that there is a bijective function $f : A \to \mathcal{B}$ satisfying that $a \in f(a)$ for all elements $a \in A$.

2.1 Proof Systems

Let $\mathcal{P} = \{p_1 = 0, \ldots, p_m = 0\}$ be a system of polynomial equations over the set of variables $X = \{x_1, \ldots, x_n, \bar{x}_1, \ldots, \bar{x}_n\}$. Each p_i is called an axiom, and throughout the paper we always assume $\mathcal{P}$ includes all axioms $x_i^2 - x_i$ and $\bar{x}_i^2 - \bar{x}_i$, ensuring that the variables are boolean, as well as the axioms $1 - x_i - \bar{x}_i$, making sure that the "bar" variables are in fact the negation of the "non-bar" variables.

Sum-of-Squares (SoS) is a static semi-algebraic proof system. An SoS proof of $f \geq 0$ from $\mathcal{P}$ is a sequence of polynomials $\pi = (t_1, \ldots, t_m; s_1, \ldots, s_a)$ such that

$$\sum_{i \in [m]} t_i p_i + \sum_{i \in [a]} s_i^2 = f \ . \tag{1}$$

The *degree* of a proof π is

$$\mathsf{Deg}(\pi) = \max\{\max_{i \in [m]} \deg(t_i) + \deg(p_i), \max_{i \in [a]} 2 \deg(s_i)\} \ . \tag{2}$$

An *SoS refutation of* $\mathcal{P}$ is an SoS proof of $-1 \geq 0$ from $\mathcal{P}$, and the SoS degree to refute $\mathcal{P}$ is the minimum degree of any SoS refutation of $\mathcal{P}$: if we let π range over all SoS refutations of $\mathcal{P}$, we can write $\mathsf{Deg}(\mathcal{P} \vdash_{\text{SoS}} \bot) = \min_\pi \mathsf{Deg}(\pi)$.

Definition 2.1 (Pseudoexpectation)**.** A degree d pseudo-expectation for $\mathcal{P}$ is a linear operator $\tilde{\mathbb{E}}$ on the space of real polynomials of degree at most d, such that

(i) $\tilde{\mathbb{E}}[1] = 1$,

(ii) $\tilde{\mathbb{E}}[tp] = 0$ for all polynomials t and $p \in \mathcal{P}$ with $\deg(t) + \deg(p) \leq d$, and

(iii) $\tilde{\mathbb{E}}[s^2] \geq 0$ for all polynomials s of degree $\deg(s) \leq d/2$.

It is easy to check that if there is a degree d pseudo-expectation for $\mathcal{P}$, then there is no SoS refutation of $\mathcal{P}$ of degree at most d: if $\tilde{\mathbb{E}}$ is applied to both sides of (1), where $f = -1$, then the right side is equal to -1 while the left is greater or equal to 0.

The size of an SoS refutation π, $\mathsf{Size}(\pi)$, is the sum of the number of monomials in each polynomial in π and the size of refuting $\mathcal{P}$ is the minimum size over all refutations $\mathsf{Size}(\mathcal{P} \vdash_{\text{SoS}} \bot) = \min_\pi \mathsf{Size}(\pi)$.

Polynomial Calculus is a dynamic proof system operating on polynomial equations over a field $\mathbb{F}$. Let $\mathcal{P}$ be over $\mathbb{F}$. Polynomial Calculus over $\mathbb{F}$ ($\text{PC}_\mathbb{F}$) consists of the derivation rules

- linear combination $\dfrac{p = 0 \qquad q = 0}{\alpha p + \beta q = 0}$, where $p, q \in \mathbb{F}[X]$ and $\alpha, \beta \in \mathbb{F}$, and
- multiplication $\dfrac{p = 0}{xp = 0}$, where $p \in \mathbb{F}[X]$ and $x \in X$.

A PC refutation of $\mathcal{P}$ is a sequence of polynomials $\pi = t_1, \ldots, t_\ell$ such that $t_\ell = 1$ and each polynomial t_i is either in $\mathcal{P}$ or can be derived by one of the derivation rules from earlier polynomials. The degree of a refutation is the maximum degree appearing in the sequence $\mathsf{Deg}(\pi) = \max_{i \in [\ell]} \mathsf{Deg}(t_i)$ and the $\text{PC}_\mathbb{F}$ degree of refuting $\mathcal{P}$ is the minimum degree required of any refutation $\mathsf{Deg}(\mathcal{P} \vdash_{\text{PC}_\mathbb{F}} \bot) = \min_\pi \mathsf{Deg}(\pi)$. Similarly, the size of a refutation π is the sum of the number of monomials in each line of π and the $\text{PC}_\mathbb{F}$ size of refuting $\mathcal{P}$ is the minimum size required of any refutation $\mathsf{Size}(\mathcal{P} \vdash_{\text{PC}_\mathbb{F}} \bot) = \min_\pi \mathsf{Size}(\pi)$.

Frege System Let us describe a Frege system due to Shoenfield, as presented in [UF96]. As Frege systems over the basis $\vee$, $\wedge$ and $\neg$ can polynomially simulate each other [CR79], the details of the system are not essential and hold for any Frege system over the mentioned basis.

Schoenfield's Frege system works over the basis $\vee$ and $\neg$. We treat the conjunction $A \wedge B$ as an abbreviation for the formula $\neg(\neg A \vee \neg B)$ and let 0, 1 denote "false" and "true" respectively. If A is a formula over variables $p_1, \ldots, p_m$, and σ maps the variables $p_1, \ldots, p_m$ to formulas $B_1, \ldots, B_m$, then $\sigma(A)$ is the formula obtained from A by replacing the variable p_i with $B_i = \sigma(p_i)$ for all $i \in [m]$.

A *rule* is a sequence of formulas written as $A_1, \ldots, A_k \vdash A_0$. If every truth assignment satisfying all of $A_1, \ldots, A_k$ also satisfies A_0, then the rule is *sound*. A formula C_0 is inferred from $C_1, \ldots, C_k$ by the rule $A_1, \ldots, A_k \vdash A_0$ if there is a function σ mapping the variables $p_1, \ldots, p_m$, over which $A_0, \ldots, A_k$ are defined, to formulas $B_1, \ldots, B_m$ such that for all $i \in \{0, \ldots, k\}$ it holds that $C_i = f(A_i)$.

The Frege system $\mathscr{F}$ that we consider consists of the following rules:

$$\begin{aligned} \vdash p \vee \neg p & \qquad \text{Excluded Middle,} \\ p \vdash q \vee p & \qquad \text{Expansion rule,} \\ p \vee p \vdash p & \qquad \text{Contraction rule,} \\ p \vee (q \vee r) \vdash (p \vee q) \vee r & \qquad \text{Associative rule,} \\ p \vee q, \neg p \vee r \vdash q \vee r & \qquad \text{Cut rule.} \end{aligned}$$

An *$\mathscr{F}$-refutation* of an unsatisfiable formula $A = C_1 \wedge \ldots \wedge C_m$ is a sequence of formulas $F_1, F_2, \ldots, F_\ell$ such that $F_\ell = 0$ and every formula F_i is either one of $C_1, \ldots, C_m$ or inferred from formulas $F_{j_1}, \ldots, F_{j_k}$ earlier in the sequence by a rule in $\mathscr{F}$. As $\mathscr{F}$ is sound and complete a formula A has a refutation if and only if it is unsatisfiable.

The size of a formula is the number of connectives in the formula and the size of a refutation π, denoted by $\mathsf{Size}(\pi)$, is the sum of the sizes of all formulas in the refutation. The depth of π is the maximum depth of any formula $F \in \pi$. We denote by $\mathscr{F}_d$ the proof system $\mathscr{F}$ restricted to formulas of depth at most d.

2.2 Propositional Formulas

As we are only interested in constant degree graphs all our axioms are of constant size. Hence the precise encoding of the axioms is not significant as we can change the encoding in constant size/degree.

As the encoding is not essential, we view a propositional formula $\mathcal{F}$ over the Boolean variables $x_1, \ldots, x_n$ as a family of functions $\mathcal{F} = \{f_1, \ldots, f_m\}$ where each $f_i : \{0,1\}^n \to \{\text{True}, \text{False}\}$ is a function that depends on a constant number of variables. The formula $\mathcal{F}$ is satisfied by an assignment $\alpha \in \{0,1\}^n$ if under α all functions evaluate to True: $f_i(\alpha) = \text{True}$ for all $i \in [m]$.

For a map $\rho : \{x_1, \ldots, x_n\} \to \{0, 1, x_1, \ldots, x_n, \bar{x}_1, \ldots, \bar{x}_n\}$ and a function $f : \{0,1\}^n \to \{\text{True}, \text{False}\}$, denote by $f\big|_\rho$ the function defined by $f\big|_\rho(x_1, \ldots, x_n) = f(\rho(x_1), \ldots, \rho(x_n))$. We extend this notation to formulas in the obvious way, i.e., $\mathcal{F}\big|_\rho = \{f_1\big|_\rho, f_2\big|_\rho, \ldots, f_m\big|_\rho\}$.

Two formulas $\mathcal{F}$ and $\mathcal{F}'$ are equivalent, denoted by $\mathcal{F} \equiv \mathcal{F}'$ if the formulas are element-wise equivalent, disregarding functions that are constant True. We say that a formula $\mathcal{F}'$ is a *subformula of* $\mathcal{F}$ if there is a map $\rho : \{x_1, \ldots, x_n\} \to \{0, 1, x_1, \ldots, x_n, \bar{x}_1, \ldots, \bar{x}_n\}$ such that $\mathcal{F}' \equiv \mathcal{F}\big|_\rho$. The following lemma states that a formula $\mathcal{F}$ is at least as hard as any of its subformulas.

Lemma 2.2. Let $\mathcal{F}, \mathcal{F}'$ be formulas such that $\mathcal{F}'$ is a subformula of $\mathcal{F}$ and each axiom of $\mathcal{F}$ depends on a constant number of variables. Then,

(i) for any field $\mathbb{F}$ it holds that $\mathsf{Deg}(\mathcal{F} \vdash_{\text{PC}_\mathbb{F}} \bot) \in \Omega\big(\mathsf{Deg}(\mathcal{F}' \vdash_{\text{PC}_\mathbb{F}} \bot)\big)$,

(ii) $\mathsf{Deg}(\mathcal{F} \vdash_{\text{SoS}} \bot) \in \Omega\big(\mathsf{Deg}(\mathcal{F}' \vdash_{\text{SoS}} \bot)\big)$, and

(iii) for all $d \geq 2$ it holds that $\mathsf{Size}(\mathcal{F} \vdash_{\mathscr{F}_d} \bot) \in \Omega\big(\mathsf{Size}(\mathcal{F}' \vdash_{\mathscr{F}_{d+1}} \bot)\big)$.

Proof. Suppose we have a refutation π of $\mathcal{F}$ in one of the mentioned proof systems. We want to show that if we hit the proof with the restriction ρ such that $\mathcal{F}\big|_\rho \equiv \mathcal{F}'$ then we obtain a proof $\pi' = \pi\big|_\rho$ of $\mathcal{F}'$.

First we need to ensure that we can derive all the axioms of $\mathcal{F}'$. These may be encoded in a different manner, but as these proof systems are implicationally complete, and each axiom only depends on a constant number of variables, this can be done in constant degree (constant size).

This shows that the SoS degree of the resulting refutation is at most a constant factor larger. For Polynomial Calculus and Frege the statement is readily verified by an inductive argument over the proof. □

For concreteness let us also define the encoding of the formulas that we are interested in.

Perfect Matching and $\mathrm{Card}(G, \vec{b})$ The Perfect Matching formula $\mathsf{PM}(G)$ encodes the claim that the graph G contains a perfect matching. For every edge $e \in E(G)$ introduce a boolean variable $x_e \in \{0,1\}$ and add for every vertex $v \in V(G)$ an axiom claiming that precisely one incident edge is set to true. As a polynomial over $\mathbb{R}$, we encode this claim as

$$q_v^{\mathsf{PM}} = \sum_{e \ni v} x_e - 1 \ , \tag{3}$$

which is satisfied under an assignment α if $q_v^{\mathsf{PM}}(\alpha) = 0$. Over other fields we encode this as a sum over indicator polynomials (see example for Tseitin below). For the Frege proof system we encode the vertex axiom as the propositional formula

$$q_v^{\mathsf{PM}} = \bigvee_{e \ni v} x_e \wedge \bigwedge_{\substack{e, e' \ni v \\ e \neq e'}} \bar{x}_e \vee \bar{x}_{e'} \ . \tag{4}$$

The formula $\mathrm{Card}(G, \vec{b})$ is encoded in a similar fashion: in the polynomial encoding replace the 1 with b_v, whereas in the propositional encoding we let the later $\wedge$ range over edge-tuples of size $b_v + 1$.

Tseitin Formula The Tseitin formula $\tau(G)$ claims that the edges of the graph G can be labeled by $0, 1$ such that the number of 1-labeled edges incident to any vertex is odd. For every edge $e \in E(G)$ introduce a boolean variable $y_e \in \{0,1\}$, denote the set of variables corresponding to edges incident to v by $Y_v = \{y_e \mid v \in e\}$ and let $A_v \subseteq \{0,1\}^{Y_v}$ contain all assignments to the variables Y_v that set an odd number of variables to 1. We encode the claim that an odd number of edges incident to $v \in V(G)$ are set to 1 as the polynomial

$$q_v^{\tau} = \sum_{\alpha \in A_v} \mathbb{1}_{\{Y_v = \alpha\}} - 1 \ , \tag{5}$$

where $\mathbb{1}_{\{Y_v=\alpha\}} = \prod_{\substack{y \in Y_v \\ \alpha(y)=1}} y \prod_{\substack{y \in Y_v \\ \alpha(y)=0}} \bar{y}$ is the indicator polynomial that is 1 iff the variables in Y_v are set according to α. As before, we also add the boolean axioms to ensure that the variables take values in $\{0,1\}$.

For the Frege system we encode the claim that an odd number of edges incident to $v \in V(G)$ is set to 1 as the propositional formula

$$q_v^{\tau} = \bigvee_{\alpha \in A_v} \mathbb{1}_{\{Y_v = \alpha\}} \ , \tag{6}$$

where the indicator is now encoded as the formula $\mathbb{1}_{\{Y_v=\alpha\}} = \bigwedge_{\substack{y \in Y_v \\ \alpha(y)=1}} y \bigwedge_{\substack{y \in Y_v \\ \alpha(y)=0}} \bar{y}$.

2.3 Graph Theory

This paper only considers simple, undirected graphs: all graphs have no self-loops nor multiple edges. For a graph $G = (V, E)$ the neighborhood of a vertex $u \in V$ is $N(u) = \{ v \in V \mid \{u, v\} \in E \}$, the neighborhood of a set of vertices $U \subseteq V$ is $N(U) = \bigcup_{u \in U} N(u)$ and for sets $U, W \subseteq V(G)$ the neighborhood of U in W is $N(U, W) = N(U) \cap W$. We denote by $\deg(v) = |N(v)|$ the degree of a vertex $v \in V$, by $\Delta(G)$ the maximum degree, $\delta(G)$ the minimum degree and by $d(G)$ the average degree of G. The edges between two vertex sets $U, W \subseteq V$ are denoted by $E(U, W) = \{ \{u, w\} \in E \mid u \in U, w \in W \}$. For a set $U \subseteq V$, we denote by $G[U] = (U, E(U, U))$ the *induced subgraph* of U in G. For a set $T \subseteq V$ we also use $G \setminus T$ as a shorthand for the induced subgraph $G[V \setminus T]$.

A graph G on n vertices is an *α-expander* (has *vertex expansion* α) if for all sets $U \subseteq V(G)$ of size $|U| \leq n/2$ it holds that $|N(U, V \setminus U)| \geq \alpha |U|$. We denote the *uniform distribution over d-regular graphs on n vertices* by $\mathcal{G}(n, d)$ and tacitly assume throughout this paper that nd is even. A graph G contains H as a *topological minor* if there is an injective map $\sigma : V(H) \to V(G)$ and for every $(u, v) \in E(H)$ there is a path $p_{uv} \subseteq G$ from $\sigma(u)$ to $\sigma(v)$ that is pairwise vertex-disjoint with all other paths except in the endpoints.

2.4 Probabilistic Bounds

We use the following version of the multiplicative Chernoff bound.

Theorem 2.3 (Chernoff)**.** Suppose $X_1, \ldots, X_n$ are independent random variables taking values in $\{0, 1\}$. Let X denote their sum and let $\mu = \mathbb{E}[X]$. Then, for and $0 \leq \delta \leq 1$ we have

$$\Pr[|X - \mu| \geq \delta\mu] \leq 2 \exp(-\delta^2 \mu / 3) \ .$$

We also need a similar bound for Poisson random variables.

Theorem 2.4 ([MU05], Theorem 5.4)**.** Let X be a Poisson random variable with parameter μ. If $x > \mu$, then

$$\Pr[X \geq x] \leq e^{-\mu} \left(\frac{e\mu}{x} \right)^x \ .$$

Finally we also need the following form of the Lovász Local Lemma.

Lemma 2.5 (LLL; [AS00], Lemma 5.1.1)**.** Let $A_1, A_2, \ldots, A_n$ be events in an arbitrary probability spacce. A directed graph $D = (V, E)$ on the set of vertices $V = \{1, 2, \ldots n\}$ is called a dependency digraph for the events $A_1, \ldots, A_n$ if for each i, $1 \leq i \leq n$, the event A_i is mutually independent of all the events $\{A_j \mid (i, j) \notin E\}$. Suppose that $D = (V, E)$ is a dependency digraph for the above events and suppose there are real numbers $x_1, \ldots x_n$ such that $0 \leq x_i < 1$ and $\Pr[A_i] \leq x_i \prod_{(i,j) \in E} (1 - x_j)$ for all $1 \leq i \leq n$. Then $\Pr[\wedge_{i=1}^n \bar{A}_i] \geq \prod_{i=1}^n (1 - x_i)$.

3 Average-Case Lower Bounds

In this section we establish our main result Theorem 1.1 giving average-case lower bounds in PC, SoS and bounded depth Frege for the $\mathrm{Card}(G, \vec{t})$ formulas.

3.1 Lower Bounds for Perfect Matching

Recall that we aim to prove that any sparse graph H (in particular a graph where $\mathsf{PM}(H)$ is hard to refute) can be topologically embedded into a random graph such that all paths in the embedding have odd length. In order to do this, we need to assume that the graph is far from bipartite (since otherwise H would need to be bipartite as well, and $\mathsf{PM}(H)$ is easy for bipartite graphs). Furthermore our embedding theorem relies on all large induced subgraphs of G having sufficiently large maximum degree. The two following definitions capture that both properties hold for all large induced subgraphs of G.

Definition 3.1. A graph G on n vertices is (κ, d)-max-degree-robust if for all $U \subseteq V(G)$ of size $|U| \geq \kappa n$ it holds that maximum degree of the induced subgraph $G[U]$ is at least $\Delta(G[U]) \geq d$.

Definition 3.2. A graph G on n vertices is (κ, ℓ)-odd-cycle-robust if for all $U \subseteq V(G)$ of size $|U| \geq \kappa n$ it holds that the induced subgraph $G[U]$ contains an odd cycle C of length at least $|C| \geq \ell$ that is also bounded in length by $|C| \leq 2 \cdot \mathrm{diam}(G[U]) + 1$.

Both properties are clearly monotone in κ: if the properties hold for som $\kappa_0 > 0$, then they also hold for all $\kappa \geq \kappa_0$. With these definitions at hand we can state our embedding theorem.

Theorem 3.3 (Embedding Theorem)**.** For $\alpha > 0$ there are $\epsilon, n_0 > 0$ such that the following holds. Let G be an α-expander on $n > n_0$ vertices, let $k \geq 6$, and H be a graph on at most $\epsilon n / k \log n$ vertices and edges. If G is $(1-4/k, 550\Delta(H)/\alpha^2)$-max-degree-robust, then G contains H as a topological minor. Furthermore, if G is also $(1 - 2/k, 1 + 2/\alpha)$-odd-cycle-robust, then one can choose the parities of the length of all the paths in the embedding of H.

Let us highlight that k may depend on the graph G. The proof of the embedding theorem can be found in Section 5.

As mentioned before we need to ensure that once we obtain an embedding of the worst-case graph H in G, that there is a matching in the graph G with the embedding of H removed. To ensure this we will in fact not embed H directly in G but rather in a subgraph of G: first we identify a set of vertices $T \subseteq V(G)$ such that no matter what set $U \subseteq T$ of odd cardinality is removed from G, the graph $G \setminus U$ still contains a perfect matching. We then proceed to show that the graph $G[T]$ satisfies all the properties required in order to embed H into it. The following lemma captures these properties.

Lemma 3.4 (Partition Lemma)**.** There is a d_0 such that for all $d > d_0$ there is an n_0 such that the following holds. Let $n > n_0$ be odd and $G \sim \mathcal{G}(n, d)$. Then, asymptotically almost surely, there is a set $T \subseteq V(G)$ of size $|T| \geq n/8$ such that $G[T]$ is a 1/3-expander, $(1/2, 7)$-odd-cycle-robust, $(1/3, d/32)$-max-degree-robust and for any set $U \subseteq T$ of odd cardinality it holds that $G \setminus U$ has a perfect matching.

The partition lemma is proved in Section 4. The constants in Lemma 3.4 are rather arbitrarily chosen and their precise values are not significant – the interested reader can find the precise dependencies between them in the proof. With Theorem 3.3 and Lemma 3.4 at hand, we can now easily state and prove our lower bounds for the perfect matching formula (i.e., the special case $t = 1$ of Theorem 1.1).

Theorem 3.5. There is a d_0 and an $\varepsilon > 0$ such that for all $d > d_0$ the following holds. For n and $n' \leq \frac{\varepsilon n}{\log n}$ both odd, let $G \sim \mathcal{G}(n, d)$ and H be any graph on n' vertices of degree $\Delta(H) \leq 5$. Then, asymptotically almost surely, $\mathsf{PM}(H)$ is a subformula of $\mathsf{PM}(G)$.

Using the graphs from Appendix A (i.e., the graphs from Theorems A.1, A.3 and A.4) as our choice of H and combining Theorem 3.5 with Lemma 2.2 finishes the proof of Theorem 1.1 for the perfect matching formula.

Proof of Theorem 3.5. Let $G \sim \mathcal{G}(n, d)$ as in the statement. Apply Lemma 3.4 to G and to obtain a set T with the mentioned properties. Apply Theorem 3.3 to $G[T]$ and the graph H to obtain a topological embedding $B_H \subseteq G[T]$ of H in G, where all paths in B_H are of odd length.

We now construct a restriction ρ to apply Lemma 2.2. As all paths are of odd length, we see that the number of vertices $|V(B_H)|$ is odd. Hence Lemma 3.4 guarantees that there exists a perfect matching m in the graph $G' = G \setminus V(B_H)$. The restriction ρ sets all variables outside of B_H to 0 or 1 depending on whether the edge $e \in m$.

For each path $p_{uv} \in B_H$ connecting two embedded vertices, pick an edge e_{uv} on this path as a "representative". The function ρ maps variables corresponding to edges e on the path p_{uv} to $x_{e_{uv}}$ if there is an odd number of edges between e and e_{uv} on the path p_{uv} and to $\bar{x}_{e_{uv}}$ otherwise. By inspection we see that $\mathsf{PM}(H)$ is a subformula of $\mathsf{PM}(G)$ as claimed. □

3.2 Lower Bounds for $\mathrm{Card}(G, \vec{t})$

In the following we prove the average-case lower bounds on the $\mathrm{Card}(G, \vec{t})$ formulas for $G \sim \mathcal{G}(n, d)$. We consider the special case when n and $t \leq d$ are odd and thus d even. Without loss of generality, assume that $t \leq d/2$: otherwise "flip" the roles of 0 and 1.

The idea is to split the edge set of the graph G into $\lfloor t/2 \rfloor$ 2-regular graphs $G_1, \ldots, G_{\lfloor t/2 \rfloor}$ and one d_0-regular graph G_0, where $d_0 = d - 2\lfloor t/2 \rfloor$. Then we want to set all variables that correspond to an edge in any of the 2-regular graphs $G_1, \ldots, G_{\lfloor t/2 \rfloor}$ to 1 so that we are left with the perfect matching formula $\mathsf{PM}(G_0)$, on which we will embed the worst-case instance of Appendix A.

In order to be able to apply Theorem 3.5 to $\mathsf{PM}(G_0)$, we need to argue that G_0 is a random d_0-regular graph. Also, we need to show that it is in fact possible to decompose a random d-regular into $\lfloor t/2 \rfloor$ 2-regular graphs plus a d_0-regular graph. For this, we use the notion of *contiguity*. Intuitively, two sequences of probability measures are contiguous, if all properties that hold with high probability in one also hold with high probability in the other measure.

Definition 3.6. Let $(P_n)_1^\infty$ and $(Q_n)_1^\infty$ be two sequences of probability measures, such that for each n, P_n and Q_n both are defined on the same measurable space $(\Omega_n, \mathcal{F}_n)$. The two sequences are *contiguous* if for every sequence of sets $(A_n)_1^\infty$, where $A_n \in \mathcal{F}_n$, it holds that

$$\lim_{n\to\infty} P_n(A_n) = 0 \Leftrightarrow \lim_{n\to\infty} Q_n(A_n) = 0 \ .$$

We denote contiguity of two sequences by $P_n \approx Q_n$.

For two random graphs $\mathcal{G}_n$ and $\mathcal{H}_n$ on n vertices, we denote by $\mathcal{G}_n \oplus \mathcal{H}_n$ the union of two independent samples conditioned on the result being simple. If $\mathcal{G}_n = \mathcal{G}(n, d)$ and $\mathcal{H}_n = \mathcal{G}(n, d')$ are uniform distributions over random regular graphs we can think of this as a proccess where we first sample $G \sim \mathcal{G}_n$ and then repeatedly sample $H \sim \mathcal{H}_n$ until the union of G and H is simple.

Theorem 3.7 (Corollary 9.44, [JLR00])**.** For all constants $d \geq 3$, $m \geq 1$ and $d_1, \ldots, d_m \geq 1$ satisfying $d = \sum_{i=1}^m d_i$ it holds that

$$\mathcal{G}(n, d_1) \oplus \cdots \oplus \mathcal{G}(n, d_m) \approx \mathcal{G}(n, d) \ .$$

In other words, if we can show that e.g. SoS requires linear degree for a formula over $G \sim \mathcal{G}(n, d_0) \oplus \bigoplus_{i \in \lfloor t/2 \rfloor} \mathcal{G}(n, 2)$ with high probability, then this also holds for the same formula over graphs $G \sim \mathcal{G}(n, d)$. Implementing our idea in the former probability distribution is straightforward and we have the following theorem.

Theorem 3.8. There is a d_0 and an $\varepsilon > 0$ such that for all $d \geq d_0$ the following holds. Let $n, n' \leq \frac{\varepsilon n}{\log n}$ and $t \in [d]$ all be odd, let $G \sim \mathcal{G}(n, d)$ and H be a graph on n' vertices of degree $\Delta(H) \leq 5$. Then, asymptotically almost surely, $\mathsf{PM}(H)$ is a subformula of $\mathrm{Card}(G, \vec{t})$.

Analogously to how Theorem 3.5 implied the $t = 1$ case of Theorem 1.1, this theorem implies the general case of Theorem 1.1.

Proof of Theorem 3.8. As n is odd d must be even. Note that we may assume that $t \leq d/2$: if $t > d/2$, let us flip the role of 1 and 0 in the formula to obtain $\mathrm{Card}(G, \overrightarrow{d - t})$. Let $d_0 = d - 2\lfloor t/2 \rfloor \geq d/2$ and sample

$$G' = G_0 \cup \bigcup_{1 \leq i \leq \lfloor t/2 \rfloor} G_i \sim \mathcal{G}(n, d_0) \oplus \bigoplus_{1 \leq i \leq \lfloor t/2 \rfloor} \mathcal{G}(n, 2) \ . \tag{7}$$

By Theorem 3.7, if we show the statement for G', then it also holds for $G \sim \mathcal{G}(n, d)$.

Set all variables in $G_1, \ldots G_{\lfloor t/2 \rfloor}$ to 1. When $\mathrm{Card}(G, \vec{t})$ is hit with this restriction we are left with the formula $\mathsf{PM}(G_0)$. As G_0 is distributed according to $\mathcal{G}(n, d_0)$, we may apply Theorem 3.5 to conclude that $\mathsf{PM}(H)$ is a subformula of $\mathrm{Card}(G, \vec{t})$. □

4 Proof of the Partition Lemma

In this section we prove Lemma 3.4, restated here for convenience.

Lemma 3.4 (Partition Lemma)**.** There is a d_0 such that for all $d > d_0$ there is an n_0 such that the following holds. Let $n > n_0$ be odd and $G \sim \mathcal{G}(n, d)$. Then, asymptotically almost surely, there is a set $T \subseteq V(G)$ of size $|T| \geq n/8$ such that $G[T]$ is a 1/3-expander, $(1/2, 7)$-odd-cycle-robust, $(1/3, d/32)$-max-degree-robust and for any set $U \subseteq T$ of odd cardinality it holds that $G \setminus U$ has a perfect matching.

We proceed as follows. First, we partition $V(G) = S \dot{\cup} T$ into two sets such that every vertex $v \in V(G)$ has a good fraction of its neighbors in S.

Definition 4.1. A (c, ε)-degree-balanced cut of a graph G is a partition $S \dot{\cup} T = V(G)$ of the vertices of G such that:

1. $\big||S| - cn\big| \leq \varepsilon n$
2. for every vertex $u \in V$, the fraction of u's neighbors that are in S is at least $c - \varepsilon$ and at most $c + \varepsilon$.

It turns out that in random regular graphs any (c, ε)-degree-balanced cut possesses the properties needed in the Partition Lemma, as summarized in the following lemma.

Lemma 4.2. For all constants $c, \varepsilon, d > 0$, satisfying $c > 1/2 + \varepsilon$ and $d \geq \max\left\{\frac{8}{(c-1/2-\varepsilon)^2}, \frac{4}{\varepsilon^2}\right\}$ the following holds. Let n be odd and $G \sim \mathcal{G}(n, d)$. Then, asymptotically almost surely as $n \to \infty$, for any (c, ε)-degree-balanced cut (S, T) of G it holds that

(i) the graph G is $\left(\kappa, d\big(\kappa - \sqrt{\frac{1-\kappa}{\kappa d}}\big)\right)$-max-degree-robust for all constants $\kappa \in [0, 1]$,

(*ii*) the graph G is $(4/\sqrt{d}, \ell)$-odd-cycle-robust, for any $\ell = O(1)$,

(*iii*) the graph $G[T]$ is an α-expander, where $\alpha = \frac{1-c-2\varepsilon}{2(1-c-\varepsilon)}$, and

(*iv*) the graph $G \setminus U$ has a perfect matching for any $U \subseteq T$ of odd cardinality.

Deferring the proof of this lemma to Section 4.2, let us first show that (c, ϵ)-degree-balanced cuts always exist in regular graphs of large enough degree.

Lemma 4.3. For all $c \in [0,1], \varepsilon > 0$ there is a $d_0 \in O\big(\frac{c}{\varepsilon^2}\log^2(\frac{c}{\varepsilon^2})\big)$ such that the following holds. For every $d > d_0$, every d-regular graph G has a (c, ε)-degree-balanced cut.

Proof. Independently include every vertex $v \in V(G)$ in S with probability c. Let A_u denote the bad event that $\big||N(u,S)| - cd\big| \geq \varepsilon d$. By Theorem 2.3, we have

$$\Pr[A_u] \leq 2\exp(-\varepsilon^2 d/3c) \ . \tag{8}$$

Note that the event A_u depends only on A_v for v within distance 2 of u in G, and there are at most d^2 many such v's. We want to apply Lemma 2.5 to the events $\{A_v \mid v \in V(G)\}$ and $x_v = x$ for some parameter x. The local lemma conditions then require $\Pr[A_u] \leq x(1-x)^{d^2}$ and this right hand side is maximized at $x = \frac{1}{d^2+1}$ where, using the bound $1-x = 1-1/(d^2+1) \geq e^{-1/d^2}$, it becomes

$$x \cdot (1-x)^{d^2} = \frac{1}{d^2+1} \cdot \left(1 - \frac{1}{d^2+1}\right)^{d^2} \geq \frac{1}{d^2+1} \cdot \frac{1}{e}.$$

For large enough $d = \Omega(\frac{c}{\varepsilon^2}\log(\frac{c}{\varepsilon^2}))$, this is much larger than $\Pr[A_u] \leq 2\exp(-\varepsilon^2 d/3c)$ so by Lemma 2.5 we conclude that $\Pr[\wedge_{v\in V(G)}\bar{A}_v] > (1-x)^n \geq \exp(-\frac{2n}{d^2})$. All that remains is to argue that there is a positive probability that both this happens as well as the size of S being close to cn. In particular if $\Pr\big[\big||S| - cn\big| \geq \varepsilon d\big] < \Pr[\wedge_{v\in V(G)}\bar{A}_v]$, the lemma follows.

By Theorem 2.3, the size $|S|$ of S is in $[cn \pm \varepsilon n]$ except with probability at most $2\exp(-\varepsilon^2 n/3c)$. Hence it is sufficient that $2\exp(-\frac{\varepsilon^2 n}{3c}) < \exp(-\frac{2n}{d^2})$, and for $d \gg \sqrt{c}/\varepsilon$ this clearly holds. This concludes the proof. □

With Lemmas 4.2 and 4.3 at hand, proving the Partition Lemma simply boils down to choosing appropriate values for the different constants.

Proof of Lemma 3.4. Fix $c = 3/4$, $\varepsilon = \kappa = 1/16$ and $\ell = 7$. Let (S,T) be the (c,ε)-degree-balanced cut as guaranteed to exist in G by Lemma 4.3. The cut (S,T) satisfies all the properties of Lemma 4.2. Hence all that remains is to verify that the constants were chosen appropriately.

(*i*) $G[T]$ is $(1/3, d/32)$-max-degree-robust: we have that $|T| \geq (1-c-\varepsilon)n = 3n/16$. Thus if the graph G is $(1/16, d/32)$-max-degree-robust, the statement follows. Observe that for our choice of κ and d large enough (e.g. $d \geq 2^{14}$ suffices) it holds that

$$d\left(\kappa - \sqrt{\frac{1-\kappa}{\kappa d}}\right) = d\left(\frac{1}{16} - \sqrt{\frac{15}{d}}\right) \geq d/32.$$

(*ii*) $G[T]$ is $(1/2, 7)$-odd-cycle-robust: as we may assume that $d \geq 64$, this property is satisfied.

(*iii*) $G[T]$ is a $1/3$-expander: the expansion α guaranteed by Lemma 4.2 is

$$\alpha = \frac{1-c-2\varepsilon}{2(1-c-\varepsilon)} = \frac{1/8}{3/8} = 1/3 \ .$$

The statement follows. □

All that remains is to prove Lemma 4.2. In the following section we recall some results from spectral graph theory needed for the proof of Lemma 4.2 which is then given in Section 4.2.

4.1 Spectral Bounds

Let us establish some notation and recall some results from spectral graph theory.

We denote the adjacency matrix of a graph G by A_G and by L_G its Laplacian $L_G = D_G - A_G$ (where D_G is the diagonal matrix containing the degrees of the vertices of G). For a matrix $A \in \mathbb{R}^{n\times n}$, denote by $\lambda_1(A) \leq \lambda_2(A) \leq \cdots \leq \lambda_n(A)$ the eigenvalues of A in non-decreasing order.

The *edge expansion* of a graph G on n vertices is

$$\Phi(G) = \min_{\substack{U\subseteq V(G)\\ |U|\leq n/2}} \frac{|E(U, V(G)\setminus U)|}{|U|} \ . \tag{9}$$

It is well-known that if the second smallest eigenvalue of the laplacian is large, then the graph is a good expander. Note that the following theorem does not require that G is regular.

Theorem 4.4 ([Moh89])**.** For all graphs G it holds that $\frac{\lambda_2(L_G)}{2} \leq \Phi(G)$.

Corollary 4.5. All graphs G have vertex expansion $\frac{\lambda_2(L_G)}{2\Delta(G)}$.

Proof. As every vertex has at most $\Delta(G)$ neighbors, the neighborhood of every set U, satisfying $|U| \leq n/2$, is of size at least $\Phi(G)/\Delta(G)$. The statement follows from Theorem 4.4. □

Recall that regular random graphs are very good spectral expanders. For the sake of conciseness, let $\lambda = \max\{|\lambda_1(A_G)|, |\lambda_{n-1}(A_G)|\}$.

Theorem 4.6 ([Fri08])**.** Let $d \geq 3$ and nd be even. Then, for $G \sim \mathcal{G}(n,d)$ it holds asymptotically almost surely that $\lambda \leq 2\sqrt{d-1} + o(1)$.

Another well-known result from spectral graph theory is that the smallest eigenvalue of the adjacency matrix puts a limit on the maximum size of an independent set.

Theorem 4.7 (Hoffman's bound)**.** Let G be a d-regular graph on n vertices. If $S \subseteq V(G)$ is an independent set of G, then

$$|S| \leq -\frac{n\cdot\lambda_1(A_G)}{d-\lambda_1(A_G)}$$

Corollary 4.8. Let G be a d-regular graph on n vertices. For any set $S \subseteq V(G)$ it holds that if $|S| > -\frac{2\cdot n\cdot\lambda_1(A_G)}{d-\lambda_1(A_G)}$, then $G[S]$ is not bipartite.

Proof. For the sake of contradiction suppose that there is an $S \subseteq V(G)$ such that $G[S]$ is bipartite and $|S| > -\frac{2\cdot n\cdot\lambda_1(A_G)}{d-\lambda_1(A_G)}$. Let us denote the partition by $S = A \dot\cup B$. W.l.o.g., assume that $|A| \geq |S|/2$ and apply Theorem 4.7 to A to conclude that $-\frac{n\cdot\lambda_1(A_G)}{d-\lambda_1(A_G)} < |A| \leq -\frac{n\cdot\lambda_1(A_G)}{d-\lambda_1(A_G)}$. □

Let us recall the mixing lemma; it states that between linearly sized sets of vertices there are about as many edges as expected in a random regular graph.

Lemma 4.9 (Expander Mixing Lemma [HLW06])**.** Let G be a d-regular graph on n vertices. Then for all $S, T \subseteq V(G)$:

$$\left||E(S,T)| - \frac{d|S||T|}{n}\right| \leq \lambda\sqrt{|S||T|} \ .$$

We also rely on the following theorem that relates the spectrum of the Laplacian and the existence of a perfect matching.

Theorem 4.10 ([BH05])**.** Let G be a graph on n vertices. If n is even and $\lambda_n(L_G) \leq 2\lambda_2(L_G)$, then G has a perfect matching.

The following statements consider large induced subgraphs $H \subseteq G$. Proposition 4.13 states that if we have good control of the degees in H, then we have good control of the spectrum of the Laplacian of H in terms of the spectrum of the adjacency matrix of G. The proof uses Weyl's theorem and Cauchy's interlacing theorem, so let us first state these.

Theorem 4.11 (Weyl)**.** Let $A, B \in \mathbb{R}^{n\times n}$ be Hermitian. Then, for all $k \in [n]$,

$$\lambda_k(A) + \lambda_1(B) \leq \lambda_k(A+B) \leq \lambda_k(A) + \lambda_n(B) \ .$$

Theorem 4.12 (Interlacing Theorem)**.** Suppose $A \in \mathbb{R}^{n\times n}$ is symmetric. Let $B \in \mathbb{R}^{m\times m}$, with $m < n$, be a principal submatrix. Then, for all $k \in [m]$,

$$\lambda_k(A) \leq \lambda_k(B) \leq \lambda_{k+n-m}(A) \ .$$

Proposition 4.13. Let G be a graph an n vertices and H be an induced subgraph of G with m vertices. Then, for all $k \in [m]$,

$$\delta(H) - \lambda_{n-k+1}(A_G) \leq \lambda_k(L_H) \leq \Delta(H) - \lambda_{m-k+1}(A_G) \ .$$

Proof. By Theorem 4.12, applied to $-A_G$ and $-A_H$, we see that for all $k \in [m]$

$$\lambda_k(-A_G) \leq \lambda_k(-A_H) \leq \lambda_{k+n-m}(-A_G) \ . \tag{10}$$

Applying Theorem 4.11 to D_H and $-A_H$, we conclude that, for all $k \in [m]$

$$\begin{aligned} \lambda_k(-A_G) + \delta(H) &= \lambda_k(-A_H) + \lambda_1(D_H) && (11)\\ &\leq \lambda_k(D_H - A_H) && (12)\\ &\leq \lambda_k(-A_H) + \lambda_n(D_H) = \lambda_{k+n-m}(-A_G) + \Delta(H) \ . && (13) \end{aligned}$$

As $\lambda_k(-A_G) = -\lambda_{n-k+1}(A_G)$ and $\lambda_{k+n-m}(-A_G) = -\lambda_{m-k+1}(A_G)$, the statement follows. □

Before commencing with the proof of Lemma 4.2, let us state two results that are of non-spectral nature. The following is a theorem by Bollobás which captures the distribution of short cycles in random regular graphs.

Theorem 4.14 ([Bol01], Corollary 2.19)**.** Let $d \geq 2$ and $k \geq 3$ be fixed natural numbers and denote by $Y_i = Y_i(G)$ the number of i-cycles in a graph $G \sim \mathcal{G}(n,d)$. Then $Y_3, Y_4, \dots Y_k$ are asymptotically independent Poisson random variables with means $\lambda_3, \lambda_4, \dots, \lambda_k$, where $\lambda_i = (d-1)^i/(2i)$.

Last, we record a simple observation that establishes a relation between the diameter of a graph and the length of its shortest odd cycle.

Lemma 4.15. Let G be a non-bipartite graph of diameter D. Then G contains an odd cycle of length at most $2D+1$.

Proof. Let G be a non-bipartite graph of diameter D and denote by C a shortest odd cycle in G. For the sake of contradiction, suppose $|C| > 2D+1$. Hence there are vertices $u, v \in C$ at distance at least $D+1$ on C. As the diameter of G is D, there is a path p from u to v of length at most D. Let $q \subseteq p$ be a subpath of p such that

1. q only shares its enpoints w_0, w_1 with C, and
2. the two arcs a_0, a_1 from w_0 to w_1 on C are longer than q.

But note that this gives rise to a shorter odd cycle: either $a_0 \cup q$ or $a_1 \cup q$ is an odd cycle, of length less than C. This is in contradiction to the initial assumption that C is a shortest odd cycle. The statement follows. □

4.2 Proof of Lemma 4.2

Recall that by Theorem 4.6, with high probability all but the largest eigenvalue of the adjacency matrix of G are bounded in magnitude by $2\sqrt{d-1}+o(1)$. Let us argue each property separately.

(i) Apply the mixing lemma (Lemma 4.9) to the graph G to conclude that for any set U of size $|U| \geq \kappa n$ it holds that

$$|E(U, V(G) \setminus U)| \leq \kappa n \cdot d\big((1-\kappa) + \sqrt{\frac{1-\kappa}{\kappa d}}\big) \ .$$

As G is a d-regular graph, we conclude that the average degree in $G[U]$ is at least $d(\kappa - \sqrt{\frac{1-\kappa}{\kappa d}})$, as required.

(ii) Recall that a sum of independent Poisson variables $X_1, \ldots, X_k$ with means $\mu_1, \ldots, \mu_k$ is again a Poisson variable with mean $\sum_{i\in[k]} \mu_i$. Hence the number of cycles in G of length at most ℓ is, according to Theorem 4.14, a Poisson random variable Y with mean

$$\mu = \sum_{i=3}^{\ell} \frac{(d-1)^i}{2i} \leq 2d^{\ell} \ . \tag{14}$$

Theorem 2.4 then tells us that

$$\Pr[Y \geq \log n] \leq e^{-\mu} \left(\frac{e\mu}{\log n}\right)^{\log n} < \frac{1}{n} \ , \tag{15}$$

where the strict inequality holds for n large enough. Hence we may assume that $Y < \log n$, which implies that at most $\ell \cdot \log n$ vertices belong to cycles of length at most ℓ.

We also know that all but the largest eigenvalue of the adjacency matrix of G are bounded in magnitude by $2\sqrt{d-1}+o(1)$. Apply Corollary 4.8 to conclude that no subset $U \subset V(G)$ of size at least $|U| \geq 3\frac{n}{\sqrt{d}}$ induces a bipartite subgraph, in other words any such $G[U]$ contains an odd cycle.

As there are so few vertices in short cycles, for n large enough, we see that all sets of at least a $\frac{4}{\sqrt{d}}$ fraction of the vertices contains an odd cycle C of length at least $|C| \geq \ell$.

Applying Lemma 4.15 to $G[U]$ we see that C can be chosen such that $|C| \leq 2\cdot\mathrm{diam}(G[U])+1$. The claim follows.

(iii) Applying Proposition 4.13 to G and $G[T]$, we see that

$$\lambda_2(L_{G[T]}) \geq \delta(G[T]) - \lambda_{n-1}(A_G) \ . \tag{16}$$

Every vertex $v \in T$ has degree at least $(1-c-\varepsilon)d$ in $G[T]$. Furthermore, as $\lambda_{n-1}(A_G)$ is bounded by $2\sqrt{d-1}+o(1)$ and we assumed that $d \geq 4/\varepsilon^2$, we obtain that $\lambda_2(L_{G[T]}) \geq (1-c-2\varepsilon)d$. Applying Corollary 4.5, we conclude that $G[T]$ has vertex expansion at least $\frac{1-c-2\varepsilon}{2(1-c-\varepsilon)}$.

(iv) Let $U \subseteq T$ of odd cardinality be as in the statement, and denote by m the number of vertices in $G \setminus U$. By Theorem 4.10, it is sufficient to establish the bound $\lambda_m(L_{G\setminus U}) \leq 2\lambda_2(L_{G\setminus U})$ on the eigenvalues of the Laplacian of $G \setminus U$. Applying Proposition 4.13 to

$G \setminus U$, we can bound these eigenvalues in terms of the eigenvalues of the adjacency matrix of G, obtaining

$$\lambda_m(L_{G\setminus U}) \leq d - \lambda_1(A_G) \text{ and} \tag{17}$$

$$\lambda_2(L_{G\setminus U}) \geq (c-\varepsilon)d - \lambda_{n-1}(A_G) \ . \tag{18}$$

As $\lambda_1(A_G)$ and $\lambda_{n-1}(A_G)$ are both bounded in absolute value by $2\sqrt{d-1} + o(1)$ we thus conclude

$$2\lambda_2(L_{G\setminus U}) - \lambda_m(L_{G\setminus U}) \geq (2(c-\varepsilon)-1)d - 2\sqrt{d-1} - o(1).$$

Since $c > 1/2 + \varepsilon$ and we assumed that $d \geq \frac{8}{(c-1/2-\varepsilon)^2}$, we have that $\lambda_m(L_{G\setminus U}) \leq 2\lambda_1(L_{G\setminus U})$ as desired.

5 Embedding Theorem

In this section we prove our embedding theorem (Theorem 3.3). Before starting with the proof, let us establish some notation and recall some facts from graph theory.

5.1 Further Graph Theory Preliminaries

In a graph $G = (V, E)$ on n vertices a vertex set $S \subseteq V$ is a *balanced separator in* G if there is a partition $V = A \dot\cup B \dot\cup S$ of the vertex set of G such that $|A|, |B| \leq 2n/3$, and G has no edges between A and B. For a path p in G we denote by $|p|$ the number of edges and by $V(p) \subseteq V(G)$ the set of vertices of p. For two vertices vertices $u, v \in V(p)$, we let $p[u, v]$ denote the subpath of p between (and including) the vertices u and v. The distance between two vertices $u, v \in V$ is the length of the shortest path from u to v and the distance between two sets $U, W \subset V$ is the minimum distance between any pair of vertices $u \in U$ and $w \in W$. Let $\mathrm{diam}(G)$ denote the diameter of G, that is, the maximum distance between any two vertices in G. For a vertex set $U \subseteq V$, and an integer $r \in \mathbb{N}$, let $B_r^G(U) \subseteq V(G)$ be the *ball around* U *of radius* r *in* G: $B_r^G(U)$ contains all vertices $v \in V$ that are at distance at most r from U.

It is well-known that vertex expanders have small diameter.

Lemma 5.1 ([Kri19])**.** Let G be an α-expander on n vertices. Then the diameter of G is upper bounded by $\lceil \frac{2(\log n - 1)}{\log(1+\alpha)} \rceil + 1 = O_\alpha(\log n)$.

As this constant will show up in a few places, let $D_\alpha^\emptyset = \frac{2}{\log(1+\alpha)} + 3$ and hence $\mathrm{diam}(G) \leq D_\alpha^\emptyset \cdot \log n$, if G is an α-expander.

The following lemma states that even if a small set of vertices is removed from a vertex expander, large sets remain well connected by short paths.

Lemma 5.2. Let G be an α-expander on n vertices. Then for all $r \geq 0$ and all disjoint $S, T \subseteq V(G)$ satisfying $|T| \geq \frac{2}{\alpha}|S|$ it holds that $|B_r^{G\setminus S}(T)| \geq \min\{n/2, (1+\alpha/2)^r|T|\}$.

Proof. Using expansion and $|S| \leq \frac{\alpha}{2}|T| \leq \frac{\alpha}{2}|B_r^{G\setminus S}(T)|$ we have that for all $r \geq 0$

$$|B_{r+1}^{G\setminus S}(T)| \geq (1+\alpha)|B_r^{G\setminus S}(T)| - |S| \geq (1+\alpha/2)|B_r^{G\setminus S}(T)| \ ,$$

unless $B_r^{G\setminus S}(T)$ is already as large as $n/2$. □

A simple consequence of this is that two large sets are connected by short paths even after the removal of a small set of vertices.

Corollary 5.3. Let G be an α-expander on n vertices. Then for all sets $S, T, U \subseteq V(G)$ satisfying $S \cap (T \cup U) = \emptyset$ and $|T|, |U| \geq \frac{2}{\alpha}|S|$ it holds that in $G \setminus S$ the distance between T and U is at most $D^{\emptyset}_{\alpha/2} \log n$.

Proof. Apply Lemma 5.2 to S, T and $r = \lceil \frac{\log n}{\log(1+\alpha/2)} \rceil$ to conclude that at distance r from T there are at least $n/2$ vertices in the graph $G \setminus S$. Applying the same argument to U and S, we see that also from U there are at least $n/2$ vertices reachable by length r paths in $G \setminus S$. But this implies that there is a path of length at most $2r + 1 \leq D^{\emptyset}_{\alpha/2} \log n$ between T and U. □

Large vertex expansion implies that balanced separators are large: the next lemma makes this well-known connection precise.

Lemma 5.4. Let G be an α-expander on n vertices, and let S be a balanced separator in G. Then $|S| \geq \frac{\alpha n}{3(1+\alpha)}$.

Proof. Let S be a balanced separator in G of size $|S| = s$, separating A and B, with $|A| = a$, $|B| = b$. Without loss of generality assume that $a \leq b \leq 2n/3$. Clearly, $a + s \geq n/3$. Further, $N_G(A) \subseteq S$, and since $a \leq n/2$, by expansion, we get that $s \geq \alpha a$. In other words, $s/\alpha \geq a$, which when substituted into $a + s \geq n/3$ yields $s(1 + 1/\alpha) \geq n/3$. □

Finally we have the following lemma on vertex-disjoint paths in expanders.

Lemma 5.5 ([FK19])**.** Let $G = (V, E)$ be an α-expander and let $A, B \subseteq V$ be two vertex sets of sizes $|A|, |B| \geq t$ for some $t > 0$. Then G contains at least $\frac{t\alpha}{1+\alpha}$ vertex-disjoint paths between A and B.

5.2 Proof of Theorem 3.3

We now proceed with the proof of Theorem 3.3, restated here for convenience.

Theorem 3.3 (Embedding Theorem)**.** For $\alpha > 0$ there are $\epsilon, n_0 > 0$ such that the following holds. Let G be an α-expander on $n > n_0$ vertices, let $k \geq 6$, and H be a graph on at most $\epsilon n/k \log n$ vertices and edges. If G is $(1-4/k, 550\Delta(H)/\alpha^2)$-max-degree-robust, then G contains H as a topological minor. Furthermore, if G is also $(1 - 2/k, 1 + 2/\alpha)$-odd-cycle-robust, then one can choose the parities of the length of all the paths in the embedding of H.

When embedding a high degree vertex $x \in V(H)$ into G, we want to find a vertex $v \in V(G)$ of high degree such that many neighbors are connected to large, disjoint sets of vertices. These large sets are very useful as they guarante that there are many vertices to which we can connect a vertex embedding. The following definition makes this intuition precise.

Definition 5.6 (Cross)**.** An (r, s)-*cross* in a graph $G = (V, E)$ is a tuple $(v, \mathcal{U})$, where $v \in V$ is a vertex and $\mathcal{U} \subseteq 2^V$ consists of r pairwise disjoint vertex sets $U \subseteq V \setminus \{v\}$, each of size $|U| = s$, such that $N(v) \cap U \neq \emptyset$ and the graph $G[U]$ is connected. We refer to v as the *center* of the cross and to $\mathcal{U}$ as the *branches* of the cross.

The following lemma shows that crosses always exist in expanders with sufficiently large maximum degree.

Lemma 5.7. For all $\beta > 0$ and $\gamma = \frac{\beta}{3(1+\beta)}$ the following holds. Let G be an β-expander on n vertices that is $(1 - 2/k, (1 + 1/\beta)r)$-max-degree-robust, for some $k \geq 3$ and $r > 0$ such that $r \leq \frac{\gamma^3 n}{k(1+\gamma)}$. Then G contains an (r, s)-cross, for all s that satisfy $r \cdot s \leq \frac{\gamma^2 n}{k(1+\gamma)}$.

The proof is an adaptation of a proof by Krivelevich and Nenadov [KN19] and is deferred to Section 5.3. We also have the following lemma which is what allows us to choose the path length parities in the "furthermore" part of Theorem 3.3. It states that if there is an odd cycle in the graph, then there is an odd and even path between any vertex u and a large enough set A of vertices. Note that this does not necessarily hold if A is too small: the vertex u may have degree 1 and A may be the single neighbor of u. Similarly a lower bound on the length of the odd cycle is needed.

Lemma 5.8. For all $\beta > 0$ the following holds. Let G be an β-expander on n vertices that contains an odd cycle of length $\ell \geq 1 + 2/\beta$. Then, for all $u \in V(G)$ and $A \subseteq V(G)$, of size $|A| \geq (D^{\emptyset}_{\beta} \log n + 1)(1 + 2/\beta)$, there is a vertex $v \in A$ such that u and v are connected by both an odd and an even path, each of length at most $(15D^{\emptyset}_{\beta/2}/\beta) \log n + \ell$.

We defer the proof of Lemma 5.8 to Section 5.4.

We now prove Theorem 3.3 with the assumption of odd-cycle-robustness. Furthermore, the proof makes all paths of odd length, though it is immediate that one can choose the parities. To get the theorem without the assumption of odd-cycle-robustness, one just has to replace the application of Lemma 5.8 by any shortest path (which, by Lemma 5.1 is short).

The main idea is due to Krivelevich and Nenadov [KN19] (see also [Kri19]). In contrast to their work we cannot directly embed the vertices into the graph but rather take a detour by embedding appropriately sized crosses for each vertex and then connect branches of crosses that correspond to embeddings of adjacent vertices. The reason for this difference is that the present theorem deals with topological embeddings rather than plain graph embeddings (the difference is that in topological embeddings vertices are connected by vertex disjoint paths while in graph embeddings subgraphs are connected).

In order for this to work we need to make some further changes to the embedding process used. In their work, three sets of vertices are maintained throughout the process: one set of "discarded" vertices, one set of vertices used in the embedding, and the remaining set of vertices. A key invariant maintained is that the set of discarded vertices expand poorly into the set of remaining vertices, which together with expansion implies that not too many vertices can be discarded. In our case, some of the discarded vertices may in fact have good expansion into C, but we can maintain the property that there are not too many such vertices. The details are worked out in what follows. If the verbal description is ambiguous, there is an algorithmic description in Appendix B (Algorithm 4).

Formally, the algorithm maintains a partition $A \dot\cup A' \dot\cup \dot\bigcup_{B\in\mathcal{B}} B \dot\cup C$ of the vertices of G. When the algorithm terminates, every vertex $v \in V(H)$ (edge $e \in E(H)$, respectively) has a vertex embedding $B_v \in \mathcal{B}$ (an edge embedding $B_e \in \mathcal{B}$) giving a topological embedding of H in G. Initially, all sets except $C = V(G)$ are empty.

Let $\beta = \frac{\alpha}{3(1+\alpha)}$ be the constant from Lemma 5.4 for the lower bound on the size of a balanced separator in an α-expander. At several points in the algorithm we want to ensure that $G[C]$ is a β-expander. This is achieved by removing any subset $U \subseteq C$ of size $|U| \leq |C|/2$ with small neighborhood $|N(U, C \setminus U)| < \beta|U|$ from C and adding it to A (i.e., letting $C \leftarrow C \setminus U$ and $A \leftarrow A \cup U$). Clearly once there are no sets $U \subseteq C$ left as above, $G[C]$ is a β-expander.

Throughout the algorithm the following invariants are maintained:

(*i*) C never increases in size and $|C| \geq n(1 - 2/k)$,

(*ii*) $N(A, C) < \beta|A|$,

(*iii*) $|A'| < \beta|A|/2$, and

(*iv*) $G[C]$ is a β-expander (by restoring expansion as described above whenever needed).

The algorithm maintains the set $I \subseteq V(H)$ to keep track of the vertices already embedded.

Let $r(d) = d(1+4/\beta) - 1 < 25d/\alpha$ and $s = \left(18D^{\emptyset}_{\beta/2}/\beta\right)\log n$. In what follows we assume ε is sufficiently small as a function of β.

Fix a vertex $x \in V(H) \setminus I$ not already embedded and apply Lemma 5.7 to $G[C]$ to obtain a $(r(\deg_H(x)), s)$-cross B_x. Remove B_x from C and add it to $\mathcal{B}$ as the vertex embedding of x (set $C \leftarrow C \setminus B_x$ and $\mathcal{B} \leftarrow \mathcal{B} \cup \{B_x\}$), and restore β-expansion in $G[C]$.

Let us check that all the conditions of Lemma 5.7 are satisfied. First, we need that $G[C]$ is $(1-2/k, (1+3(1+\beta)/\beta)r(\deg_H(x)))$-max-degree-robust. We have $1+3(1+\beta)/\beta \le 22/\alpha$ and thus

$$r(\deg_H(x))(1+3(1+\beta)/\beta) \le r(\Delta(H))\frac{22}{\alpha} < \frac{550}{\alpha^2}\Delta(H).$$

Furthermore since G is $(1-4/k, 550\Delta(H)/\alpha^2)$-max-degree-robust and $|C| \ge (1-2/k)n$, $G[C]$ is $(1-2/k, 550\Delta(H)/\alpha^2)$-max-degree-robust. Second we need to check that

$$r(\deg_H(x)) \le \frac{\gamma^3|C|}{k(1+\gamma)} \qquad \text{and} \qquad r(\deg_H(x))\cdot s \le \frac{\gamma^2|C|}{k(1+\gamma)}\ ,$$

where $\gamma = \frac{\beta}{3(1+\beta)}$. Since $|C| \ge (1-2/k)n$ and $\Delta(H) \le |V(H)| \le \frac{\varepsilon n}{k\log n}$ the first bound clearly holds for n large enough, and provided ε is sufficiently small as a function of α the second bound also holds. Thus we can indeed apply Lemma 5.7 on $G[C]$ with the desired choice of r and s.

After embedding x, we need to connect the embedding B_x to the embeddings of the neighbors $N_H(x) \cap I = \{y_1, \dots, y_\nu\}$ that are already embedded. Suppose, for now, that the vertex embeddings have branches $U_x \in B_x$ and $U_{y_i} \in B_{y_i}$ that are β-expanding into C (i.e. $|N(U_x, C)|, |N(U_{y_i}, C)| \ge \beta s$), and such that neither of the two branches are already used to connect x, resp. y_i, to a neighbor.

By the assumption on odd-cycle-robustness, we see that $G[C]$ is non-bipartite and contains an odd cycle c of length

$$1 + 2/\alpha \le |c| \le 2\cdot \operatorname{diam}(G[C]) + 1 \le 2D^{\emptyset}_{\beta}\log n + 1\ . \tag{19}$$

As each branch is rather large, of size s, we can apply Lemma 5.8 to $G[C]$, $N(U_x, C)$ and $N(U_{y_i}, C)$ to conclude that in $G[C]$ there is an odd path q_i connecting U_x to U_{y_i} of length $\left(18D^{\emptyset}_{\beta/2}/\beta\right)\log n \le s$. Remove q_i from C, add it to $\mathcal{B}$ as the edge embedding $B_{\{x,y_i\}}$ and restore β-expansion in $G[C]$. This process is illustrated in Figure 1 and can be found as pseudo code in Algorithm 4.

If all branches of a vertex embedding B_z have either too few neighbors in C or are already adjacent to an edge embedding (i.e., have already been used to embed some other edge), then we want to remove the embedding of z. This has to be done in a careful manner in order not to break the invariants. First, move all branches that are not used to connect z to a neighbor to A. Note that each such branch U satisfies $|N(U, C)| < \beta|U|$. Next, move the remaining branches along with the adjacent edge embeddings to A'. Last, the center of B_z is moved to A' and z is removed from I. Note that at most $2(\deg_H(z)-1)s$ many vertices are moved to A': at most $\deg_H(z)-1$ many branches of size s and as many edge embeddings, each again of size at most s. On the other hand at least

$$\left(r(\deg_H(z)) - (\deg_H(z)-1)\right)\cdot s = \deg_H(z)\cdot 4s/\beta \tag{20}$$

many vertices are moved to A. Hence the invariant $|A'| < \beta|A|/2$ is maintained.

The algorithm terminates the first time either $I = V(H)$ or $|A| \ge n/k$. This completes the description of the algorithm.

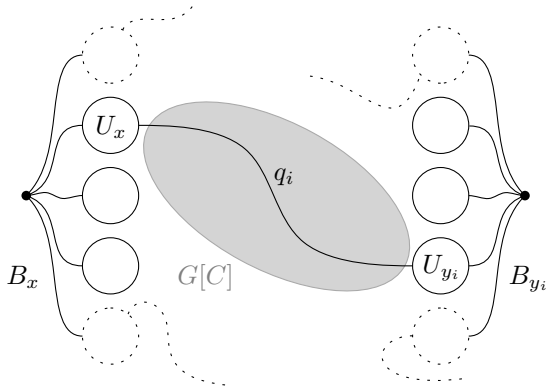


Figure 1: The vertex embedding B_x is connected to B_{y_i} by the path q_i which connects the two branches U_x and U_{y_i}. The dotted branches have an edge embedding adjacent and can thus not be used to connect B_{y_i} to B_x.

It remains to argue that it cannot happen that $|A| \geq n/k$, in other words that when the algorithm terminates, all of H is embedded in G. To this end, observe that the size of $\cup_{B\in\mathcal{B}} B$ is upper bounded by

$$\begin{aligned} s \cdot \left(|E(H)| + \sum_{v \in V(H)} r(\deg_H(v)) \right) &< s \cdot \left(|E(H)| + (4/\beta + 1) \sum_{v \in V(H)} \deg_H(v) \right) \\ &\leq s \cdot |E(H)| \cdot \frac{11}{\beta} \\ &\leq s \cdot \frac{\varepsilon n}{k \log n} \cdot \frac{11}{\beta} \\ &\leq \beta n/2k \ . \end{aligned}$$

Furthermore, while $|A| \leq n/k$ we have that

$$|A'| < \beta|A|/2 \leq \beta n/2k \ .$$

Note that this also holds the first time $|A|$ becomes larger than n/k. This shows, in particular, that the invariant $|C| \geq n(1-2/k)$ is maintained throughout the execution of the algorithm.

For the sake of contradiction, suppose that the algorithm terminates because of $|A| \geq n/k$. Note that $|N(A)| \leq |A'| + |\cup_{B\in\mathcal{B}} B| + |N(A,C)| < \beta(|A| + n/k)$. We do a case distinction, depending on the size of $|A|$. In both cases we derive contradiction and thus show that the algorithm only terminates after having embedded all of H into G.

Case 1: $n/k \leq |A| \leq n/2$. By expansion and using $\beta < \alpha/3$ we have

$$\alpha|A| \leq |N(A)| < \beta(|A| + n/k) < \frac{\alpha}{3}(|A| + n/k) \ ,$$

which together with $|A| \geq n/k$ yields the desired contradiction.

Case 2: $|A| > n/2$. Note that the first time $|A| \geq n/k$, it also holds that $|A| \leq n(1+1/k)/2$ as the sets added to A are of size at most $|C|/2 \leq (n - |A|)/2$. Hence we get that

$$|N(A)| < \beta\big(n/k + |A|\big) < \beta n\big(1/k + (1+1/k)/2\big) \leq \frac{\alpha n}{3(1+\alpha)} \ ,$$

using that $k \geq 3$. Note that $N(A)$ is a balanced separator, separating A from $V(G) \setminus A$. But this is a contradiction, since Lemma 5.4 states that any balanced separator of G has size at least $\frac{\alpha n}{3(1+\alpha)}$.

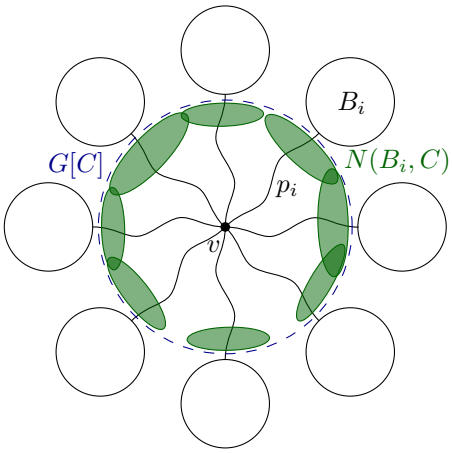


Figure 2: A cross with center v and branches $\{ V(p_i) \cup B_i \mid i \in [r] \}$.

5.3 Crosses in Expanders

Let us now turn to the proof of Lemma 5.7, restated here for convenience.

Lemma 5.7. For all $\beta > 0$ and $\gamma = \frac{\beta}{3(1+\beta)}$ the following holds. Let G be an β-expander on n vertices that is $(1 - 2/k, (1 + 1/\beta)r)$-max-degree-robust, for some $k \geq 3$ and $r > 0$ such that $r \leq \frac{\gamma^3 n}{k(1+\gamma)}$. Then G contains an (r, s)-cross, for all s that satisfy $r \cdot s \leq \frac{\gamma^2 n}{k(1+\gamma)}$.

The proof follows a similar algorithm as the proof of Theorem 3.3. In this case we can in fact more or less use the original argument of Krivelevich and Nenadov [KN19] without any extensions.

Proof. The high-level idea of the proof is as follows. First, using the embedding argument of Krivelevich and Nenadov, we will find some number $r' > r$ pairwise disjoint sets $B_1, \dots, B_{r'}$ of vertices of G and a final set C such that (i) each B_i is a connected subgraph of G on s vertices, (ii) the B_is have many neighbors in C, and (iii) $G[C]$ is expanding. Having these subsets, we can then choose a representative $u_i \in N(B_i, C)$ of each B_i, take a vertex $v \in C$ of high degree (which exists by the max-degree-robustness of G), and apply Lemma 5.5 to find vertex-disjoint paths connecting $N(v)$ to the u_is. This establishes the existence of a cross with v as the center and the B_is together with the respective paths as branches. See Figure 2 for an illustration.

Let us proceed with the details. In case there is some ambiguity in the verbal description there is also a pseudo code description in Appendix B of what follows.

Fix r, set $r' = r(1 + 1/\gamma)$ and choose $s \in \mathbb{N}$ maximal such that $s \leq \frac{\gamma n}{k \cdot r'}$. Note that $s \geq 1/\gamma$ and if the statement holds for this maximal s, then it also holds for smaller values of s, as one can always shrink the branches to the appropriate size.

Let us describe an algorithm to identify the sets $\mathcal{B} = \{ B_i \subseteq V(G) \mid i \in [r'] \}$. The algorithm maintains a partition $A \,\dot\cup\, \dot\bigcup_{B \in \mathcal{B}} B \,\dot\cup\, C$ of the vertices of G. Initially, all sets except $C = V(G)$ are empty. After running the procedure, the set $\mathcal{B}$ contains r' pairwise vertex-disjoint sets such that for each $B_i \in \mathcal{B}$ it holds that $|B_i| = s$ and the induced subgraph $G[B_i]$ is a single connected component. Further, for all sub families $\mathcal{F} \subseteq \mathcal{B}$ it holds that $\left|\bigcup_{F \in \mathcal{F}} N(F, C)\right| \geq \gamma s |\mathcal{F}|$. Throughout the execution of the algorithm the following invariants are maintained

(i) C never increases in size,

(ii) $|C| \geq n(1 - 2/k)$,

(iii) $\left|\dot{\bigcup}_{B\in\mathcal{B}} B\right| \leq r' \cdot s \leq \gamma n/k$, and

(iv) $N(A, C) < \beta|A|$.

The algorithm terminates if $\mathcal{B}$ contains r' vertex sets as described, or if the size of A reaches $|A| \geq n/k$. The latter case can only occur if there is a small balanced separator in G. But G is a β-expander, so we know from Lemma 5.4 that there are no small balanced separators and hence when the algorithm terminates, $\mathcal{B}$ must contain r' sets as described above.

Like in the main algorithm used in the proof of Theorem 3.3, we want to ensure that $G[C]$ is a γ-expander throughout the algorithm, which is achieved by removing any subset $U \subseteq C$ of size $|U| \leq |C|/2$ with small neighborhood $|N(U, C \setminus U)| < \gamma|U|$ from $C = C \setminus U$ and adding it to $A = A \cup U$.

Repeat the following while there are less than r' sets in $\mathcal{B}$. Choose a set of vertices $U \subseteq C$ of size $|U| = s$ such that $G[U]$ is a single connected component. Remove this set from $C = C \setminus U$, add it to $\mathcal{B} = \mathcal{B} \cup \{U\}$ and restore expansion in $G[C]$. After expansion is restored, let $\mathcal{F} \subseteq \mathcal{B}$ be a maximal (possibly empty) family such that $\left|\bigcup_{F\in\mathcal{F}} N(F, C)\right| < \gamma s|\mathcal{F}|$. Remove $\mathcal{F}$ from $\mathcal{B} = \mathcal{B} \setminus \mathcal{F}$, and add these sets to $A = A \cup_{F\in\mathcal{F}} F$.

As mention before, the algorithm terminates once there are either r' sets in $\mathcal{B}$ or the set A is large $|A| \geq n/k$. This completes the description of the algorithm. Let us argue that the latter cannot happen – for the sake of contradiction, suppose the algorithm terminates because $|A| \geq n/k$. Note that we have $|N(A)| \leq |\cup_{B\in\mathcal{B}} B| + |N(A, C)| < \gamma(|A| + n/k)$. We do a case distincion on the size of $|A|$.

Case 1: $n/k \leq |A| \leq n/2$. By expansion, $\beta|A| \leq N(A) < \gamma(|A| + n/k) < \frac{\beta}{3}(|A| + n/k)$ As $|A| \geq n/k$ this is a contradiction.

Case 2: $|A| \geq n/2$. Note that the first time $|A| \geq n/k$, it also holds that $|A| \leq n(1 + 1/k)/2$ as the sets added to A are of size at most $|C|/2 \leq (n - |A|)/2$. Hence we get (using $k \geq 3$) that

$$|N(A)| \leq \gamma(n/k + |A|) \leq \gamma n = \frac{\beta n}{3(1 + \beta)}$$

Note that $N(A)$ is a balanced separator, separating A from $V(G) \setminus A$. But this is a contradiction, since Lemma 5.4 states that any balanced separator of G has size at least $\frac{\beta n}{3(1+\beta)}$.

It remains to obtain an (r, s)-cross from the sets B_i and the remaining part C. Choose a vertex $v \in C$ of degree at least $\deg_{G[C]}(v) \geq r'$. Such a vertex v exists, as $|C| \geq (1 - 2/k)n$ is large (second invariant) and the statement assumes that there is a vertex of degree r' in every induced subgraph of size at least $(1 - 2/k)n$. Apply Lemma 5.5 to $G[C]$, the vertex set $N(v)$, and a transversal of the family $\{ N(B, C) \mid B \in \mathcal{B} \}$. Note that such a transversal exists by Hall's marriage theorem, using that $s \geq 1/\beta$ and that every subset of $\mathcal{B}$ is β-expanding into C. We conclude that there are pairwise vertex-disjoint paths $\{ p_i \mid i \in [r] \}$ each connecting a set $B_i \in \mathcal{B}$ to $N(v)$. The cross has center v and branches $\{ V(p_i) \cup B_i \mid i \in [r] \}$. Each branch is of size at least s. Shrinking the branches to the appropriate size recovers the statement. □

5.4 Odd and Even Paths

In this section we prove Lemma 5.8.

Lemma 5.8. For all $\beta > 0$ the following holds. Let G be an β-expander on n vertices that contains an odd cycle of length $\ell \geq 1 + 2/\beta$. Then, for all $u \in V(G)$ and $A \subseteq V(G)$, of size $|A| \geq (D^{\emptyset}_{\beta} \log n + 1)(1 + 2/\beta)$, there is a vertex $v \in A$ such that u and v are connected by both an odd and an even path, each of length at most $(15D^{\emptyset}_{\beta/2}/\beta) \log n + \ell$.

The lemma is a corollary of a more general statement about short paths in α-expanders. The lemma states that if sets S, T, where $|S| \gtrsim |T|/\alpha$, are connected by $|T|$ many short vertex-disjoint paths, then for any large set U there is again a set of short vertex-disjoint paths that does not only connect every vertex of T to S but also a vertex from U to S.

In order to state the lemma, let us introduce some notation. For a graph G and vertex sets $S, T \subseteq V(G)$, denote by $L^G_{\text{disj}}(T, S)$ the minimum total length of connecting all vertices of T to S by pairwise vertex-disjoint paths;

$$L^G_{\text{disj}}(T, S) = \min_{\{\, p_t \mid t \in T \,\}} \sum_{t \in T} |p_t| \tag{21}$$

where $\{\, p_t \mid t \in T \,\}$ ranges over all sets of pairwise vertex-disjoint paths such that p_t connects t to S. If no such set of paths exists, the value of the minimum is taken to be ∞. If the graph G is clear from context, we omit the superscript.

A similar lemma (though without the essential upper bound on the path lengths) has appeared in e.g. [FK19].

Lemma 5.9. Let G be an β-expander on n vertices and $S, T \subseteq V(G)$ satisfy $|S| \geq |T|(1+2/\beta)$. Then every set $U \subseteq V(G)$, of size $|U| \geq (L_{\text{disj}}(T, S) + |T|)(1 + 2/\beta)$, contains a vertex $u \in U$ such that $L_{\text{disj}}(T \cup \{u\}, S) \leq 7(L_{\text{disj}}(T, S) + |T|)/\beta + 2D^{\emptyset}_{\beta/2} \log n$.

Lemma 5.8 follows by a single application of Lemma 5.9.

Proof of Lemma 5.8. Let C denote an odd cycle of length $\ell \geq 1 + 2/\beta$, as guaranteed to exist, and denote by p a shortest path connecting u to C. By Lemma 5.1, we know that $|p| \leq D^{\emptyset}_{\beta} \log n$. Apply Lemma 5.9 to $S = C$, $T = \{u\}$, $p_u = p$, and $U = A$. We conclude that there is an odd and even length path connecting u to A of length $\ell + (15D^{\emptyset}_{\beta/2}/\beta) \log n$, as required. □

Proof of Lemma 5.9. Denote by $\mathcal{P} = \{\, p_t \mid t \in T \,\}$ a set of pairwise vertex-disjoint paths of smallest total length, where the path p_t connects t to S. Let $V(\mathcal{P}) = \cup_{p \in \mathcal{P}} V(p)$ denote all the vertices in the paths in $\mathcal{P}$. Clearly, $|V(\mathcal{P})| = L_{\text{disj}}(T, S) + |T|$. Set $m = |V(\mathcal{P})|(1 + 2/\beta)$ and $r = \lceil \frac{\log m}{\log(1+\beta/2)} \rceil$. Note that $n \geq |U| \geq m$ and hence $r \leq \frac{1}{2} D^{\emptyset}_{\beta/2} \log n$.

If $|S| \geq m$, apply Corollary 5.3 to $V(\mathcal{P})$, $S \setminus V(\mathcal{P})$ and $U \setminus V(\mathcal{P})$ to conclude that there is a path p of length $D^{\emptyset}_{\beta/2} \log n$ connecting $S \setminus V(\mathcal{P})$ to $U \setminus V(\mathcal{P})$ in $G \setminus V(\mathcal{P})$. The set $\mathcal{P} \cup \{p\}$ clearly satisfies the conclusion of the lemma.

Otherwise, if $|S| < m$, we want to get into a position where we can again apply Corollary 5.3. To this end, we define a sequence of sets of vertices $S = S_0 \subseteq S_1 \subseteq \ldots \subseteq S_\ell \subseteq V(G)$ that are in some sense well-connected to S. We formalize this property after explaining how to obtain these sets.

The set S_{i+1} is defined in terms of S_i using the following process. Let w^t_i be the last vertex on the path p_t (viewed as a path from S to t) that is in S_i and $W_i = \{\, w^i_t \mid t \in T \,\}$. Suppose $|S_i| < m$ and there is a path of length at most r connecting $S_i \setminus W_i$ to $V(\mathcal{P}) \setminus S_i$ in the graph $G \setminus W_i$. Denote by q_i a minimal such path, denote by w the endpoint of q_i in $V(\mathcal{P}) \setminus S_i$, and let

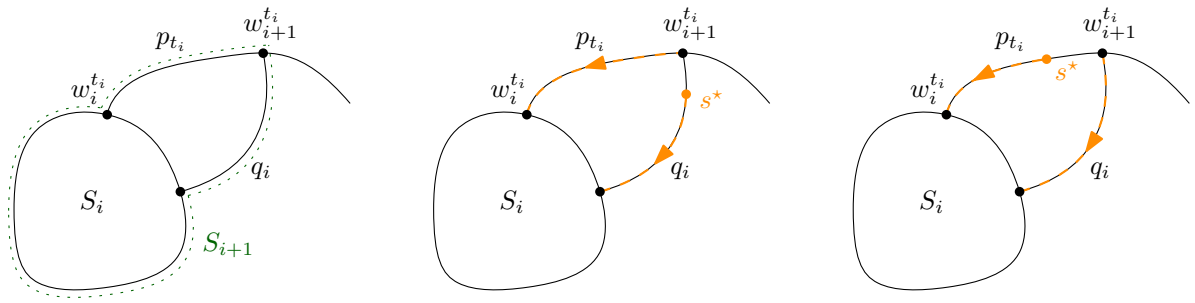


Figure 3: Given the set S_i, the first figure depicts the process of obtaining the set S_{i+1}. The following figures indicate how to route the paths, as in the proof of Claim 5.10, depending on where $s^\star$ is located.

$t_i \in T$ be such that $w \in V(p_{t_i})$. Then, define $S_{i+1} = S_i \cup q_i \cup p_{t_i}[w_i^{t_i}, w]$. Otherwise, if $|S_i| \geq m$ or there is no such q_i, set $\ell = i$ and stop the process. There is an illustration of this process in Figure 3.

The following claim formalizes the well-connectedness property of S_ℓ.

Claim 5.10. For every vertex $s^\star \in S_\ell \setminus W_\ell$ it holds that $L_{\text{disj}}^{G[S_\ell \cup V(\mathcal{P})]}(T \cup \{s^\star\}, S) \leq L_{\text{disj}}^{G}(T, S) + |S_\ell|$ and furthermore the paths achieving this bound are the same as the paths in $\mathcal{P}$ outside $S_\ell \setminus W_\ell$.

Proof. Proof by induction on $i \in \{0, \ldots, \ell\}$. The base case $i = 0$ clearly holds – we have for all $s^\star \in S_0 \setminus W_0 \subseteq S$ that $L_{\text{disj}}^{G[S_0 \cup V(\mathcal{P})]}(T \cup \{s^\star\}, S) = L_{\text{disj}}^{G}(T, S)$.

Suppose the statement is true for some $i \in \{0, \ldots \ell - 1\}$ and let us prove that is then true for $i + 1$ as well. By the inductive hypothesis, $L_{\text{disj}}^{G[S_i \cup V(\mathcal{P})]}(T \cup \{s_i\}, S) \leq L_{\text{disj}}^{G}(T, S) + |S_i|$, and this bound can be achieved by a set of paths $\mathcal{P}'$ which follow $\mathcal{P}$ outside $S_i \setminus W_i$.

Fix an arbitrary $s^\star \in S_{i+1} \setminus W_{i+1}$. By the induction hypothesis the claim holds for $s^\star \in S_i \setminus W_i$, so we may assume [4] that either $s^\star \in q_i$, or $s^\star \in p_{t_i}[w_i^{t_i}, w_{i+1}^{t_i}]$. If $s^\star \in q_i$ (excluding its endpoint $w_{i+1}^{t_i}$) then we simply extend the path in $\mathcal{P}'$ ending in s_i with the subpath of q_i from $s^\star$ to s_i, increasing the total length of $\mathcal{P}'$ by at most $|q_i|$. On the other hand if $s^\star \in p_{t_i}[w_i^{t_i}, w_{i+1}^{t_i}]$ then we reroute the path from t_i in $\mathcal{P}'$ to s_i via q_i and then use the now unused part of p_{t_i} to connect $s^\star$ to S, again increasing the total length of $\mathcal{P}'$ by at most $|q_i|$. There is an illustration of the two cases in Figure 3.

In either case, we can connect T and $s^\star$ to S via vertex-disjoint paths of length at most

$$L_{\text{disj}}^{G[S_{i+1} \cup V(\mathcal{P})]}(T \cup \{s^\star\}, S) \leq L_{\text{disj}}^{G}(T, S) + |S_i| + |q_i| \leq L_{\text{disj}}^{G}(T, S) + |S_{i+1}|,$$

as desired. □

It is easy to see that $|S_\ell| \leq 2m + r$: the number of vertices added by $p_{t_i}[w_i^{t_i}, w_{i+1}^{t_i}]$ is always upper bounded by $|V(\mathcal{P})| \leq m$. Suppose there is a path $p^\star$ of length $|p^\star| \leq D_{\beta/2}^{\emptyset} \log n + r$ connecting some vertex $s^\star \in S_\ell \setminus W_\ell$ to $u^\star \in U \setminus V(\mathcal{P}_\ell)$ in $G \setminus V(\mathcal{P}_\ell)$. We can then "compose"

[4] Here we are using that $W_i = (W_{i-1} \setminus \{w_{i-1}^{t_i}\}) \cup \{w_i^{t_i}\}$.

the paths to conclude that

$$L^G_{\text{disj}}(T \cup \{u^\star\}, S) \leq L^{G[S_\ell]}_{\text{disj}}(W_\ell \cup \{s^\star\}, S) + L^{G\setminus(S_\ell \setminus (W_\ell \cup \{s^\star\}))}_{\text{disj}}(W_\ell \cup \{s^\star\}, T \cup \{u^\star\}) \tag{22}$$

$$\leq |S_\ell| + L^G_{\text{disj}}(T, S) + |p^\star| \tag{23}$$

$$\leq |S_\ell| + L^G_{\text{disj}}(T, S) + D^{\emptyset}_{\beta/2} \log n + r \tag{24}$$

$$\leq 2m + r + L^G_{\text{disj}}(T, S) + D^{\emptyset}_{\beta/2} \log n + r \tag{25}$$

$$\leq 7(L^G_{\text{disj}}(T, S) + |T|)/\beta + 2D^{\emptyset}_{\beta/2} \log n \ , \tag{26}$$

as claimed in the statement.

It remains to establish that such a path $p^\star$ exists. If $|S_\ell| \geq m$, apply Corollary 5.3 to $V(\mathcal{P})$, $S_\ell \setminus W_\ell$ and $U \setminus V(\mathcal{P})$ to conclude that there is a path $p^\star$ of length at most $|p^\star| \leq D^{\emptyset}_{\beta/2} \log n$ that connects $S_\ell \setminus W_\ell$ to $U \setminus V(\mathcal{P})$ in $G \setminus V(\mathcal{P})$.

Otherwise, by construction, $S_\ell \setminus W_\ell$ cannot reach $V(\mathcal{P}) \setminus W_\ell$ within r steps in $G \setminus W_\ell$. Hence, to argue that in $G \setminus V(\mathcal{P})$ the ball of radius r around $S_\ell \setminus W_\ell$ is large, we do not need to apply Lemma 5.2 to $S_\ell \setminus W_\ell$ and $V(\mathcal{P})$ but in fact can apply it to $S_\ell \setminus W_\ell$ and W_ℓ, where we use that $|S_\ell \setminus W_\ell| \geq |S| - |T| \geq 2|T|/\beta$. This enables us to grow $S_\ell \setminus W_\ell$ into a set $S^\star$ of size m. Now we are in a position to apply Corollary 5.3 to $V(\mathcal{P})$, $S^\star$ and $U \setminus V(\mathcal{P})$ to conclude that there is a path of length at most $D^{\emptyset}_{\beta/2} \log n$ that connects $S^\star$ to $U \setminus V(\mathcal{P})$ in $G \setminus V(\mathcal{P})$. Taking an additional r steps in $G[S^\star]$, one can reach $S_\ell \setminus W_\ell$, as required. This concludes the proof of the lemma. □

6 Concluding Remarks

We have established average-case lower bounds for refuting the perfect matching formula and more generally the $\text{Card}(G, \vec{t})$ formula in random d-regular graphs on an odd number of vertices. Let us conclude by discussing some further loose ends and mention some open problems.

6.1 Polynomial Calculus Space Lower Bounds

The space of a PC refutation π is the amount of memory needed to verify π. The PC space of a formula $\mathcal{F}$ is then the minimum space required for any PC refutation π of $\mathcal{F}$. As this is rather tangential to the rest of the paper we refer to [FLM+13] for formal definitions. For convenience, let us restate our result on PC space.

Theorem 1.3. For all $\alpha > 0$ there is a d_0 such that the following holds. Let G be a constant degree α-expander of average degree at least d_0. Then over any field $\mathbb{F}$ it holds that $\text{PC}_{\mathbb{F}}$ requires space $\Omega(n/\log n)$ to refute the Tseitin formula defined on G.

The proof idea is to take the worst-case Tseitin lower bounds from Filmus et al. [FLM+13] for which PC requires $\Omega(n)$ space and embed these into a vertex expander of large enough average degree. The only compication that arises is that these formulas are defined over multigraphs – the multigraph H is obtained from an appropriate[5] constant degree graph G by doubling each edge. An inspection of the proof of Theorem 3.3 reveals that H may be a multigraph and we can thus implement our proof strategy.

Proof Sketch. Consider the worst-case instance H from Filmus et al. [FLM+13] on $\varepsilon n/\log n$ vertices, for some small enough $\varepsilon > 0$. Apply Theorem 3.3 to H and G. This gives a topological

[5] See the proof of Theorem 8 in [FLM+13].

embedding of H in G, with no control of the parities of the length of the paths. Consider a restriction ρ that sets the variables outside the embedding of H such that no axiom is falsified (see, e.g., [PRST16]). By appropriately substituting the variables on each path of the topological embedding we obtain that the worst-case instance $\tau(H)$ is a subformula of $\tau(G)$. As a restriction only reduces the amount of space needed to verify a proof, we see that $\tau(G)$ requires PC space $\Omega(n/\log n)$. □

6.2 Paths in Expanders

The arguments used in the proof of Theorem 3.3 can be adapted to make partial progress on a question by Friedman and Krivelevich [FK19]. They asked, given a positive integer q, whether it is possible to guarantee the existence of a cycle whose length is divisible by q in every α-expander.

We can show that for all primes q satisfying $1/\mathsf{poly}(\alpha) \ll q \ll \sqrt{n/\log n}$, this indeed holds. In fact, for all $a \in \mathbb{Z}_q$, we can show that there is a cycle of length $a \mod q$.

The idea is to embed a cycle C_{q^2} of length q^2 into G such that between any two vertices there are two paths whose length difference is non-zero modulo q. If we can ensure this, as all $0 \neq b \in \mathbb{Z}_q$ are generators, we can choose one path between all embedded vertices such that the length of the cycle is $a \mod q$ for any $a \in \mathbb{Z}_q$.

In order to obtain paths of different length modulo q, let us embed a cycle c_e (of length $\gg 1/\mathsf{poly}(\alpha)$) for each edge $e = \{u, v\}$. We then want to connect the vertex embeddings B_u, B_v to c_e such that the two resulting paths are of different length modulo q. Note that once a vertex is connected to the cycle, there are only about $2/q$ vertices in c_e such that both paths are of equal length modulo q. As q is rather large and thus there are few such "bad" vertices, when an edge embedding has to be moved to the sets A, A', we can ensure that the set A' remains relatively small compared to A.

6.3 Open Problems

The main concrete problem left open is to reduce the degree of the hard graphs: the embedding approach taken in the worst-case to average-case reduction results in very large degree d; while Theorem 1.1 does not give an explicit estimate on d_0 one can trace through the proofs and get an estimate somewhere around 15 000. The main bottleneck that prevents us from reducing this is the Partition Lemma and in particular the dependence of d_0 on c and ϵ in Lemma 4.3. If this part could be significantly improved or circumvented we believe that the degree of the graph could be significantly reduced, although it would still be relatively large (at least a few hundred). It would be interesting to see what happens for very small degrees such as a 4-regular graph (recall that since n is odd, d must be even) – is $\mathsf{PM}(G)$ hard with high probability even for these graphs?

Another interesting question is the proof complexity of perfect matching in Polynomial Calculus over $\mathbb{F}_2$ (or any other field of characteristic 2). While $\mathrm{PC}_{\mathbb{F}_2}$ can refute the perfect matching formula on an odd number of vertices for parity reasons, the situation is less clear when the number of vertices is even. Are there graphs G that do not admit perfect matchings but $\mathrm{PC}_{\mathbb{F}_2}$ requires exponential size refutations?

Theorem 1.1 only gives lower bounds for $\mathrm{Card}(G, b)$ when $b = \vec{t}$ is a constant vector (and G is regular). It would be nice to characterize more generally for which vectors b the formula is hard. In the analogous setting for Tseitin formulas, the precise charges of the vertices do not matter, as long as the sum of charges is odd the formula remains hard to refute on a random graph [BGIP01, Gri01]. In the $\mathrm{Card}(G, b)$ case however this is not the case. For instance, if the vector of target degrees b violates the inequality of the Erdős-Gallai characterization of degree

sequences then SoS can easily refute $\mathrm{Card}(G, b)$. In the case when G is the complete graph this in fact gives a complete characterization of the easy and hard vectors b but for sparse graphs the situation is less clear. Is there a nice characterization of vectors b for which $\mathrm{Card}(G, b)$ is hard for SoS with high probability over a random d-regular G?

More broadly, another open problem is to prove SoS lower bounds for random CSPs that do not support pairwise uniform distributions (c.f. the brief discussion on CSPs in Section 1.2). Viewed this way, our results establish hardness of random monotone 1-in-k-SAT instances with two occurrences per variable, for some large constant k. Reducing k corresponds to the aforementioned problem of reducing the degree, but some other natural questions are to look at other CSPs such as 1-in-k-SAT with negated literals, or to understand the hardness as a function of the density of the instances.

References

[ABSRW04] Michael Alekhnovich, Eli Ben-Sasson, Alexander A. Razborov, and Avi Wigderson. Pseudorandom generators in propositional proof complexity. *SIAM Journal on Computing*, 34(1):67–88, 2004. 1

[AGK20] Jackson Abascal, Venkatesan Guruswami, and Pravesh K. Kothari. Strongly refuting all semi-random boolean csps. *CoRR*, abs/2009.08032, 2020. 1

[AH19] Albert Atserias and Tuomas Hakoniemi. Size-Degree Trade-Offs for Sums-of-Squares and Positivstellensatz Proofs. In Amir Shpilka, editor, *34th Computational Complexity Conference (CCC 2019)*, volume 137 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 24:1–24:20, Dagstuhl, Germany, 2019. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. 1, 3

[AR01] M. Alekhnovich and A. A. Razborov. Lower bounds for polynomial calculus: non-binomial case. In *Proceedings 42nd IEEE Symposium on Foundations of Computer Science*, pages 190–199, 2001. 1, 2, 4

[AS00] Noga Alon and Joel H. Spencer. *The Probabilistic Method*. Wiley Publishing, 2nd edition, 2000. 11

[BFSU96] Andrei Z. Broder, Alan M. Frieze, Stephen Suen, and Eli Upfal. An efficient algorithm for the vertex-disjoint paths problem in random graphs. In *Proceedings of the Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '96, page 261–268, USA, 1996. Society for Industrial and Applied Mathematics. 5

[BGIP01] Samuel R. Buss, Dima Grigoriev, Russell Impagliazzo, and Toniann Pitassi. Linear gaps between degrees for the polynomial calculus modulo distinct primes. *Journal of Computer and System Sciences*, 62(2):267–289, March 2001. Preliminary version in *CCC '99*. 1, 2, 6, 7, 29, 34

[BH05] Andries E. Brouwer and Willem H. Haemers. Eigenvalues and perfect matchings. *Linear Algebra and its Applications*, 395:155 – 162, 2005. 17

[BHK+16] B. Barak, S. B. Hopkins, J. Kelner, P. Kothari, A. Moitra, and A. Potechin. A nearly tight sum-of-squares lower bound for the planted clique problem. In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 428–437, 2016. 1

[Bla37] Archie Blake. *Canonical Expressions in Boolean Algebra*. PhD thesis, University of Chicago, 1937. 1

[BN21] Samuel R. Buss and Jakob Nordström. Proof complexity and SAT solving. In Armin Biere, Marijn J. H. Heule, Hans van Maaren, and Toby Walsh, editors, *Handbook of Satisfiability*, volume 336 of *Frontiers in Artificial Intelligence and Applications*, chapter 7, pages 233–350. IOS Press, 2nd edition, February 2021. 3

[Bol01] Béla Bollobás. *Random Graphs*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 edition, 2001. 17

[BS14] Boaz Barak and David Steurer. Sum-of-squares proofs and the quest toward optimal algorithms. In *Proceedings of the International Congress of Mathematicians (ICM)*, volume IV, pages 509–533, August 2014. Available at http://www.icm2014.org/download/Proceedings_Volume_IV.pdf. 1

[CEI96] Matthew Clegg, Jeffery Edmonds, and Russell Impagliazzo. Using the Groebner basis algorithm to find proofs of unsatisfiability. In *Proceedings of the 28th Annual ACM Symposium on Theory of Computing (STOC '96)*, pages 174–183, May 1996. 1, 3

[CN19] Julia Chuzhoy and Rachit Nimavat. Large minors in expanders, 2019. 3

[CR79] Stephen A. Cook and Robert A. Reckhow. The relative efficiency of propositional proof systems. *The Journal of Symbolic Logic*, 44(1):36–50, 1979. 1, 8

[DKN20] Nemanja Draganić, Michael Krivelevich, and Rajko Nenadov. Rolling backwards can move you forward: on embedding problems in sparse expanders, 2020. 5

[DMO+19] Yash Deshpande, Andrea Montanari, Ryan O'Donnell, Tselil Schramm, and Subhabrata Sen. The threshold for sdp-refutation of random regular nae-3sat. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '19, page 2305–2321, USA, 2019. Society for Industrial and Applied Mathematics. 5

[FK19] Limor Friedman and Michael Krivelevich. Cycle lengths in expanding graphs, 2019. 20, 26, 29

[FLM+13] Yuval Filmus, Massimo Lauria, Mladen Mikša, Jakob Nordström, and Marc Vinyals. Towards an understanding of polynomial calculus: New separations and lower bounds (Extended abstract). In *Proceedings of the 40th International Colloquium on Automata, Languages and Programming (ICALP '13)*, volume 7965 of *Lecture Notes in Computer Science*, pages 437–448. Springer, July 2013. 4, 28

[Fri08] Joel Friedman. *A Proof of Alon's Second Eigenvalue Conjecture and Related Problems*. American Mathematical Society, Providence, R.I., 2008. 16

[GGR+09] Jim Geelen, Bert Gerards, Bruce Reed, Paul Seymour, and Adrian Vetta. On the odd-minor variant of hadwiger's conjecture. *Journal of Combinatorial Theory, Series B*, 99(1):20–29, 2009. 5

[GHP02] Dima Grigoriev, Edward A. Hirsch, and Dmitrii V. Pasechnik. Complexity of semi-algebraic proofs. In *STACS 2002*, pages 419–430, Berlin, Heidelberg, 2002. Springer Berlin Heidelberg. 1, 2

[GIRS19] Nicola Galesi, Dmitry Itsykson, Artur Riazanov, and Anastasia Sofronova. Bounded-Depth Frege Complexity of Tseitin Formulas for All Graphs. In *44th International Symposium on Mathematical Foundations of Computer Science (MFCS 2019)*, pages 49:1–15, 2019. 4

[GKT19] Nicola Galesi, Leszek Kołodziejczyk, and Neil Thapen. Polynomial calculus space and resolution width. In *Proceedings of the 60th Annual IEEE Symposium on Foundations of Computer Science (FOCS '19)*, pages 1325–1337, November 2019. 4

[GL10] Nicola Galesi and Massimo Lauria. Optimality of size-degree trade-offs for polynomial calculus. *ACM Transactions on Computational Logic*, 12(1):4:1–4:22, November 2010. 1

[Gri01] Dima Grigoriev. Linear lower bound on degrees of positivstellensatz calculus proofs for the parity. *Theoretical Computer Science*, 259(1):613 – 622, 2001. 1, 2, 6, 29, 36

[HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *BULL. AMER. MATH. SOC.*, 43(4):439–561, 2006. 16

[Hå17] Johan Håstad. On small-depth frege proofs for tseitin for grids. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 97–108, 2017. 4, 6, 36

[IPS99] Russell Impagliazzo, Pavel Pudlák, and Jiří Sgall. Lower bounds for the polynomial calculus and the Gröbner basis algorithm. *Computational Complexity*, 8(2):127–144, 1999. 1, 3

[JLR00] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński. *Theory of random graphs*. John Wiley & Sons, New York; Chichester, 2000. 13

[KMOW17] Pravesh K. Kothari, Ryuhei Mori, Ryan O'Donnell, and David Witmer. Sum of squares lower bounds for refuting any csp. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2017, page 132–145, New York, NY, USA, 2017. Association for Computing Machinery. 1, 4

[KN19] Michael Krivelevich and Rajko Nenadov. Complete Minors in Graphs Without Sparse Cuts. *International Mathematics Research Notices*, 05 2019. rnz086. 3, 5, 6, 21, 24

[KR96] J. Kleinberg and R. Rubinfeld. Short paths in expander graphs. In *Proceedings of the 37th Annual Symposium on Foundations of Computer Science*, FOCS '96, page 86, USA, 1996. IEEE Computer Society. 3, 5

[Kra19] Jan Krajíček. *Proof Complexity*, volume 170 of *Encyclopedia of Mathematics and Its Applications*. Cambridge University Press, March 2019. 1

[Kri19] Michael Krivelevich. *Expanders – how to find them, and what to find in them*, page 115–142. London Mathematical Society Lecture Note Series. Cambridge University Press, 2019. 3, 6, 19, 21

[Las01] Jean B Lasserre. An explicit exact sdp relaxation for nonlinear 0-1 programs. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 293–303. Springer, 2001. 1

[LS91] László Lovász and Alexander Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM Journal on Optimization*, 1(2):166–190, 1991. 2

[Mar06] Klas Markström. Locality and hard SAT-instances. *Journal on Satisfiability, Boolean Modeling and Computation*, 2(1-4):221–227, 2006. 3

[MN15] Mladen Mikša and Jakob Nordström. A generalized method for proving polynomial calculus degree lower bounds. In *Proceedings of the 30th Annual Computational Complexity Conference (CCC '15)*, volume 33 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 467–487, June 2015. 1, 2

[Moh89] Bojan Mohar. Isoperimetric numbers of graphs. *Journal of Combinatorial Theory, Series B*, 47(3):274 – 291, 1989. 16

[MPW15] Raghu Meka, Aaron Potechin, and Avi Wigderson. Sum-of-squares lower bounds for planted clique. In *Proceedings of the 47th Annual ACM Symposium on Theory of Computing (STOC '15)*, pages 87–96, June 2015. 1

[MU05] Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, USA, 2005. 11

[Par00] Pablo A Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. PhD thesis, California Institute of Technology, 2000. 1

[Pot17] Aaron Potechin. Sum of squares lower bounds from symmetry and a good story. *CoRR*, abs/1711.11469, 2017. 2

[Pot20] Aaron Potechin. Sum of squares bounds for the ordering principle. In Shubhangi Saraf, editor, *35th Computational Complexity Conference, CCC 2020, July 28-31, 2020, Saarbrücken, Germany (Virtual Conference)*, volume 169 of *LIPIcs*, pages 38:1–38:37. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. 1

[PRST16] Toniann Pitassi, Benjamin Rossman, Rocco A. Servedio, and Li-Yang Tan. Polylogarithmic frege depth lower bounds via an expander switching lemma. In *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '16, page 644–657, New York, NY, USA, 2016. Association for Computing Machinery. 4, 29

[Raz98] Alexander A. Razborov. Lower bounds for the polynomial calculus. *Computational Complexity*, 7(4):291–324, December 1998. 1, 2

[Raz02] Alexander A. Razborov. Proof complexity of pigeonhole principles. In *5th International Conference on Developments in Language Theory, (DLT '01), Revised Papers*, volume 2295 of *Lecture Notes in Computer Science*, pages 100–116. Springer, July 2002. 1

[Raz17] Alexander Razborov. On the width of semialgebraic proofs and algorithms. *Math. Oper. Res.*, 42(4):1106–1134, November 2017. 2

[Rii93] Søren Riis. *Independence in Bounded Arithmetic*. PhD thesis, University of Oxford, 1993. 2

[Rot17] Thomas Rothvoss. The matching polytope has exponential extension complexity. *J. ACM*, 64(6), September 2017. 5

[Sch08] Grant Schoenebeck. Linear level Lasserre lower bounds for certain k-CSPs. In *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS '08)*, pages 593–602, October 2008. 1

[Sey16] Paul Seymour. *Hadwiger's Conjecture*, pages 417–437. Springer International Publishing, Cham, 2016. 5

[Sho87] N. Shor. Class of global minimum bounds of polynomial functions. *Cybernetics*, 23:731–734, 1987. 1

[UF96] Alasdair Urquhart and Xudong Fu. Simplified lower bounds for propositional proofs. *Notre Dame Journal of Formal Logic*, 37(4):523–544, 1996. 8

[Yan88] Mihalis Yannakakis. Expressing combinatorial optimization problems by linear programs. In *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing*, STOC '88, page 223–228, New York, NY, USA, 1988. Association for Computing Machinery. 5

A Worst-Case Lower Bounds

In this section we describe a general reduction from the Tseitin formula to the Perfect Matching formula as it appeared in [BGIP01] for Polynomial Calculus. We then observe that this reduction also works for the SoS and bounded depth Frege proof systems.

Starting from a graph G such that the Tseitin formula $\tau(G)$ is hard for a proof system P, we want to craft a graph H so that $\mathsf{PM}(H)$ is hard for P. To simplify the presentation, let us assume that G is d-regular. As we are interested in unsatisfiable instances, i.e., when G has an odd number of vertices, we may assume that d is even.

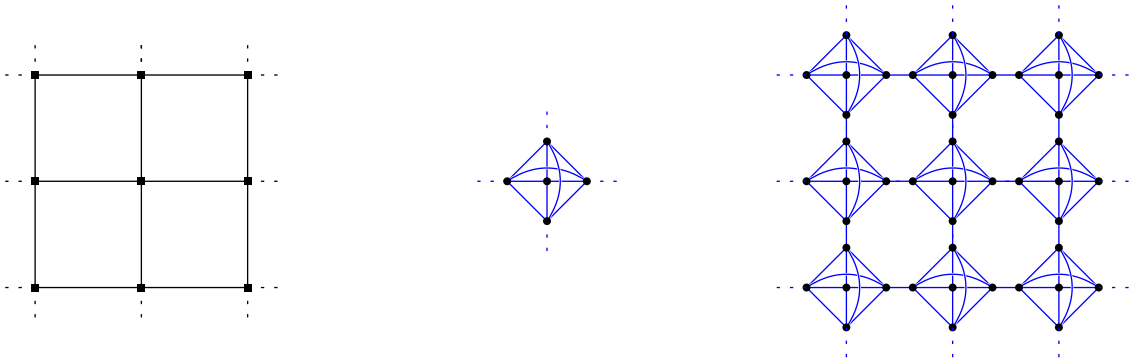
The graph H is a "blow-up" (or "lift") of G: each vertex in $V(G)$ is lifted to a clique of $d+1$ vertices and each lifted edge connects a single pair of vertices from the corresponding cliques. If we denote the lifted vertices of $v \in V(G)$ by $\mathsf{lift}(v) = \{(v,\star),(v,1),\dots,(v,d)\}$, we add for each edge $\{u,v\} \in E(G)$, where v is the ith neighbor of u and u is the jth neighbor of v, an edge $\{(u,i),(v,j)\}$. An illustration of the construction of H can be found in Figure 4.

For intuition, let us describe how we would obtain a satisfying assignment to the Perfect Matching Formula from a hypothetical satisfying assignment to the Tseitin Formula. Set the lifted edges to the same value as they are set to in the Tseitin Formula. Now observe that each charge is odd and hence there is an even number of vertices left that are not matched yet in each $\mathsf{lift}(v)$, for $v \in V(G)$. As the vertices in $\mathsf{lift}(v)$ form a clique, we can select a perfect matching on these unmatched vertices to obtain a satisfying assignment to the Perfect Matching Formula.

Buss et al. [BGIP01] showed that Polynomial Calculus can simulate this reduction.

Theorem A.1 ([BGIP01])**.** There are graphs G on an odd number of vertices n and maximum degree $\Delta(G) = 5$ such that Polynomial Calculus over any field of characteristic different from 2 requires degree $\Theta(n)$ to refute $\mathsf{PM}(G)$.

What remains is to check that this reduction also gives Perfect Matching worst-case lower bounds for Sum-of-Squares and bounded depth Frege. This straightforward verification is carried out in the following two sections.



(a) The graph G. (b) The lift of a single vertex. (c) The lifted graph H.

Figure 4: An illustration of the blow-up construction, starting from a 4-regular graph.

A.1 Sum-of-Squares

A nice property of the Sum-of-Squares system is that if the variables for a formula $\mathcal{Q}$ can be expressed as well-behaved low-degree polynomials in the variables of another formula $\mathcal{P}$ for which a pseudo-expectation exists, then a pseudo-expectation also exists for $\mathcal{Q}$. This property is well-known but let us state and quickly prove the exact version we need.

Claim A.2. Let $\mathcal{P} \subseteq \mathbb{R}[x_1, \ldots, x_n]$ and $\mathcal{Q} \subseteq \mathbb{R}[y_1, \ldots, y_m]$ be two systems of polynomial equations. Let $\tilde{\mathbb{E}}_{\mathcal{Q}}$ be a degree D pseudo-expectation for $\mathcal{Q}$. Suppose there is a function $f : \{x_1, \ldots, x_m\} \to \mathbb{R}[y_1, \ldots y_n]$, mapping the x variables to polynomials in y of degree at most t. Extend f to polynomials by applying the function to each variable individually. If f satisfies that $\tilde{\mathbb{E}}_{\mathcal{Q}}[f(r \cdot p)] = 0$, for all $p \in \mathcal{P}$ and $r \in \mathbb{R}[x_1, \ldots x_m]$ of degree $\deg(r \cdot p) \leq D/t$, then $\tilde{\mathbb{E}}_{\mathcal{Q}} \circ f$ is a degree D/t pseudo-expectation for $\mathcal{P}$.

Proof. As f only maps variables we have that $\tilde{\mathbb{E}}_{\mathcal{Q}}[f(1)] = \tilde{\mathbb{E}}_{\mathcal{Q}}[1] = 1$. Also, we need to check that $\tilde{\mathbb{E}}_{\mathcal{Q}}[f(s^2)] \geq 0$ for $s \in \mathbb{R}[x_1, \ldots, x_m]$ of degree $\deg(s) \leq D/2t$. As we apply f individually to each variable, we can write $\tilde{\mathbb{E}}_{\mathcal{Q}}[f(s^2)] = \tilde{\mathbb{E}}_{\mathcal{Q}}[\big(\sum_{t \in s} f(t)\big)^2] \geq 0$, as $\tilde{\mathbb{E}}_{\mathcal{Q}}$ is a degree D pseudo-expectation. □

In order to apply Claim A.2 with $\mathcal{Q} = \tau(G)$ and $\mathcal{P} = \mathsf{PM}(H)$, we need to express each variable from the Perfect Matching formula as a low degree polynomial in the Tseitin variables.

Let us recall some notation. For a vertex $v \in V(G)$, let Y_v be the set of Tseitin variables corresponding to edges incident to v and denote by A_v all boolean assignments to Y_v that satisfy the vertex axiom of v, i.e., assignments that set an odd number of edges to true. For a Tseitin variable y_e, where $e \in E(G)$, let $\mathsf{lift}(y_e) \in E(H)$ denote the lifted edge variable.

With this notation at hand, let us define the function f to use in Claim A.2. Variables that correspond to lifted edges, $x_e = \mathsf{lift}(y_{e'})$ for some $e' \in E(G)$, are set to 1 if and only if $y_{e'}$ is set to 1 and the variables in Y_v are set according to some assignment in A_v

$$f(x_e) = \sum_{\substack{\alpha \in A_v \\ \alpha(y_{e'})=1}} \mathbb{1}_{\{Y_v = \alpha\}} \ . \tag{27}$$

Note that this is a polynomial of degree $\deg(v) = d$ in the y_e's. For each assignment $\alpha \in A_v$, set the variables in $\mathsf{lift}(Y_v)$ according to α and fix a matching m_α on the vertices in $\mathsf{lift}(v)$ not

matched by α. For any edge $e \subseteq \mathsf{lift}(v)$, let

$$f(x_e) = \sum_{\substack{\alpha \in A_v \\ e \in m_\alpha}} \mathbb{1}_{\{Y_v = \alpha\}} \ . \tag{28}$$

If we apply f individually to each variable, we claim that for $i \in \{\star, 1, \ldots, d\}$ and $v \in V(G)$ the polynomial $f(q^{\mathsf{PM}}_{(v,i)})$ is equal to the Tseitin axiom q^τ_v:

$$\begin{aligned} f(q^{\mathsf{PM}}_{(v,i)}) &= \sum_{e \ni (v,i)} f(x_e) - 1 && (29)\\ &= \sum_{\substack{\alpha \in A_v \\ \alpha(y_{e'})=1}} \mathbb{1}_{\{Y_v=\alpha\}} + \sum_{\substack{\alpha \in A_v \\ \alpha(y_{e'})=0}} \mathbb{1}_{\{Y_v=\alpha\}} - 1 && (30)\\ &= q^\tau_v \ , && (31) \end{aligned}$$

using that the m_α are matchings. As $\tilde{\mathbb{E}}_{\tau(G)}$ maps all axioms multiplied by a low degree polynomial to 0, the same holds for $\tilde{\mathbb{E}}_{\tau(G)} \circ f$ and we can thus apply Claim A.2.

We conclude that if there is a degree D pseudo-expectation $\tilde{\mathbb{E}}_{\tau(G)}$ for the Tseitin Formula $\tau(G)$, then there is a degree D/d pseudo-expectation $\tilde{\mathbb{E}}_{\mathsf{PM}(H)}$ for the Perfect Matching formula over the lifted graph H. Using Grigoriev's Tseitin lower bounds [Gri01] we obtain the following Theorem.

Theorem A.3. There are graphs G on an odd number of vertices n and maximum degree $\Delta(G) = 5$ for which SoS requires degree $\Theta(n)$ to refute $\mathsf{PM}(G)$.

A.2 Bounded Depth Frege

In this section we intend to prove the following theorem.

Theorem A.4. There is a constant $c > 0$ such that the following holds. Suppose $D \leq \frac{c \log n}{\log \log n}$. Then there are graphs G on an odd number of vertices n and maximum degree $\Delta(G) = 5$ such that any depth-D Frege refutation of $\mathsf{PM}(G)$ requires size $\exp(\Omega(n^{c/D}))$.

As in the previous section we use a function f, mapping Perfect Matching variables to low depth formulas in the Tseitin Variables, to argue that we can transform a refutation of $\mathsf{PM}(H)$ into a refutation of the Tseitin formula $\tau(G)$. Assuming that this can be done, we use the following recent result of Håstad about the Tseitin formula over the grid to obtain Theorem A.4.

Theorem A.5 ([Hå17])**.** Suppose that $D \leq \frac{\log n}{59 \log \log n}$, then any depth-$D$ Frege refutation of the Tseitin formula on the $n \times n$ grid requires size $\exp(\Omega(n^{1/58(D+1)}))$.

In the previous section f mapped to polynomials. As we are now working with formulas we need to translate the polynomials to formulas. This is straightforward; reusing notation from the previous section, let

$$f(x_e) = \bigvee_{\substack{\alpha \in A_v \\ \alpha(y_{e'})=1}} \mathbb{1}_{\{Y_v=\alpha\}} \ , \tag{32}$$

if $x_e = \mathsf{lift}(y_{e'})$ is a lifted edge. Else let

$$f(x_e) = \bigvee_{\substack{\alpha \in A_v \\ e \in m_\alpha}} \mathbb{1}_{\{Y_v = \alpha\}} \ . \tag{33}$$

Suppose there is a depth-D Frege refutation π of the Perfect Matching formula $\mathsf{PM}(H)$. Replace each occurrence of a Perfect Matching variable x_e by $f(x_e)$ to obtain a depth-$(D+2)$ refutation π'. We claim that π' is a refutation of the Tseitin formula $\tau(G)$ of size $O_d(\mathsf{Size}(\pi))$.

To this end we need to argue that f maps Perfect Matching axioms to Tseitin Axioms or tautologies that are derivable in small size and depth. Analoguous to SoS observe that

$$f(\bigvee_{e \ni (v,i)} x_e) = \bigvee_{e \ni (v,i)} f(x_e) \tag{34}$$

$$= \bigvee_{\substack{\alpha \in A_v \\ \alpha(y_{e'})=1}} \mathbb{1}_{\{Y_v = \alpha\}} \vee \bigvee_{\substack{\alpha \in A_v \\ \alpha(y_{e'})=0}} \mathbb{1}_{\{Y_v = \alpha\}} \tag{35}$$

$$= q_v^\tau \ , \tag{36}$$

for all $v \in V(G)$ and $i \in \{\star, 1, \ldots, d\}$. Last we need to show that the axioms $\bar{x}_e \vee \bar{x}_{e'}$, for edges $e \neq e' \in E(H)$ satisfying $e \cap e' \neq \emptyset$, are mapped to a tautology derivable in small size and depth. If we let $\{(v,i)\} = e \cap e'$ we can write

$$f(\bar{x}_e \vee \bar{x}_{e'}) = \big(\neg \bigvee_{\beta \in B} \mathbb{1}_{\{Y_v = \beta\}}\big) \vee \big(\neg \bigvee_{\gamma \in C} \mathbb{1}_{\{Y_v = \gamma\}}\big) \ , \tag{37}$$

for disjoint subsets $B, C \subseteq A_v$. Observe that this formula is a tautology and defined on d variables. Thus it is derivable in constant depth and size dependent on d, which is constant in our case.

B Embedding Algorithm

Algorithm 1 Restores β-expansion of $G[C]$.

```
1: procedure FixExpansion(G, C, A, β)
2:     while G[C] is not a β-expander do
3:         U ←_any subset of C such that |U| ≤ |C|/2 and |N(U, C \ U)| < β|U|
4:         C ← C \ U
5:         A ← A ∪ U
```

Algorithm 2 Finds an (r, s)-cross in an β-expander G as in the proof of Lemma 5.7.

Require: Conditions of Lemma 5.7.

procedure EMBEDVERTEX(G, r, s, β, k)
  $\gamma \leftarrow \frac{\beta}{3(1+\beta)}$
  $s \leftarrow \max\{1/\gamma, s\}$
  $r' \leftarrow (1 + 1/\gamma)r$
  $A, \mathcal{B} \leftarrow \emptyset; C \leftarrow V(G)$
  while $|\mathcal{B}| < r'$ **do**
    $U \leftarrow_{\text{any}}$ subset of C such that $|U| = s$ and $G[U]$ is a single connected component
    $\mathcal{B} \leftarrow \mathcal{B} \cup \{U\}$; $C \leftarrow C \setminus U$
    FIXEXPANSION(G, C, A, γ)
    $\mathcal{F} \subseteq \mathcal{B}$ maximal such that $|\cup_{F \in \mathcal{F}} N(F, C)| < \gamma s |\mathcal{F}|$
    $\mathcal{B} \leftarrow \mathcal{B} \setminus \mathcal{F}$; $A \leftarrow A \cup_{F \in \mathcal{F}} F$

  $v \leftarrow_{\text{any}} C$ such that $\deg_{G[C]}(v) \geq r'$
  $F \leftarrow$ a transversal of $\{ N(B, C) \mid B \in \mathcal{B} \}$
  $\{ p_i \mid i \in [r] \} \leftarrow$ from Lemma 5.5 applied to $G[C]$, v and F
  return $\{v\} \cup \{ V(p_i) \cup B_i \mid i \in [r] \}$ ▷ Shrink branches appropriately

Algorithm 3 Remove the embedding of vertex x.

procedure UNEMBEDVERTEX(A, A', B, H, I, x)
  $(v, \mathcal{U}) \leftarrow B_x$ ▷ v is the center and $\mathcal{U}$ are the branches of B_x
  $B \leftarrow B \setminus B_x$; $I \leftarrow I \setminus x$
  $W \leftarrow \emptyset$
  for all $e \in E(H)$ such that $x \in e$ and e is embedded **do**
    let $U \leftarrow \mathcal{U}$ be the branch adjacent to B_e
    $\mathcal{U} \leftarrow \mathcal{U} \setminus U$
    $B \leftarrow B \setminus B_e$
    $W \leftarrow W \cup U \cup B_e$
  $A \leftarrow A \cup \mathcal{U}$ ▷ First add to A, then to A' to maintain the invariant
  $A' \leftarrow A' \cup W \cup \{v\}$

Algorithm 4 Embeds H in an α-expander G as in the proof of Theorem 3.3.

procedure EMBEDGRAPH(H, G, α)
  $\beta \leftarrow \alpha/3(1+\alpha)$
  $A, A', B \leftarrow \emptyset; C \leftarrow V(G)$
  $I \leftarrow \emptyset$
  while $I \neq V(H)$ **do**
    $x \leftarrow_{\text{any}} V(H) \setminus I$
    $B_x \leftarrow$ EMBEDVERTEX($G[C], \deg_H(x), s, \beta, k$)
    $C \leftarrow C \setminus B_x$; $B \leftarrow B \cup B_x$; $I \leftarrow I \cup x$
    FIXEXPANSION(G, C, A, β)

    $\mathcal{U}_{\text{free}}(K) \leftarrow$ branches of the cross K that are *not* used to connect to a neighbor
    for all $\{x, y\} \in E(H)$ such that $y \in I$ **do**
      try
        $U_z \leftarrow_{\text{any}} \mathcal{U}_{\text{free}}(B_z)$ such that $|N(U_z, C)| \geq \beta |U_z|$ for $z \in \{x, y\}$
      catch no such U_z for $z \in \{x, y\}$
        UNEMBEDVERTEX(A, A', B, H, I, z); **continue**
      $B_{xy} \leftarrow$ odd path from Lemma 5.8 applied to $G[C]$, $N(U_x, C)$ and $N(U_y, C)$
      $C \leftarrow C \setminus B_{xy}$; $B \leftarrow B \cup B_{xy}$
      FIXEXPANSION(G, C, A, β)
  return B

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il