# Refuting Perfect Matchings in Spectral Expanders is Hard

Ari Biswas, *University Of Warwick*, tcs@randomwalks.xyz
Rajko Nenadov, *University Of Auckland*, rajko.nenadov@auckland.ac.nz

**ABSTRACT**

This work studies the complexity of refuting the existence of a perfect matching in spectral expanders with an odd number of vertices, in the Polynomial Calculus (PC) and Sum of Squares (SoS) proof system. Austrin and Risse [SODA, 2021] showed that refuting perfect matchings in sparse $d$-regular *random* graphs, requires with high probability, proofs with degree $\Omega\left(\frac{n}{\log n}\right)$ in the above proof systems. We extend their result by showing the same lower bound holds for **all** $d$-regular graphs with a mild spectral gap. As a direct consequence, we also positively resolve the open problem posed by Buss and Nordström, which asks, *"Are even colouring formulas over expander graphs hard for polynomial calculus over fields of characteristic distinct from 2 ?"*

# 1 Introduction

Perhaps the most fundamental problem in computation is to provide an answer to the question asked by Stephen A. Cook and Robert A. Reckhow. in their seminal paper [10] - "*Given a true statement $A$, is there a short proof of the claim that $A$ is true*". In trying to answer this question, we must first describe what constitutes a valid proof. That is, we must describe the language in which the proof is written (axioms), and the rules for checking it (the verifier). Each set of rules for writing and checking a proof defines a proof system. Therefore, a precise restatement of the question above is the following: "*Given a true statement $A$ and a proof system $S$, what is the length of a shortest proof $\pi \in S$ that proves $A$*?" If we could show that there exists a proof system $S$, such that for *any* true statement $A$, the length of the shortest proof in $S$ is upper bounded by some polynomial in the length of $A$, it would imply that CoNP = NP and consequently the polynomial hierarchy collapes to NP. Conversely, if we could show large proof size lower bounds for some true statement $A$ in *all* proof systems, it would lead to a formal proof of the widely believed conjecture that P $\neq$ NP. Unfortunately, such lower bounds for *arbitrary* proof systems are out of reach. As an intermediate step, the research community has invested a significant effort in proving lower bounds for increasingly expressive proof systems (e.g., see [1, 2, 4, 8, 9, 17, 20, 26, 28, 29, 32, 33]).

In this work, we focus on the algebraic and semi-algebraic proof systems[1] of *polynomial calculus* (PC) and *sum of squares* (SoS). In algebraic proof systems, we are given a set $\mathcal{Q} = \{q_i(\vec{x}) \mid i \in [m]\}$ of $m$ polynomial equations[2] over $n$ variables $\vec{x} = \{x_1, ..., x_n\}$. In PC, the equations can be

[1]Similar to the work of Per Austrin and Kilian Risse. [5], our lower bounds also extend to bounded depth Frege proof systems. However, the key technical component in this work is the graph theoretic techniques proposed for embedding carefully chosen hard instances into the host graph. As this applies broadly across proof systems, we restrict our preliminaries to PC and SoS for brevity.

[2]Semi-algebraic proof systems also allow for inequalities but we will not deal with inequality constraints in this paper.

over an arbitrary, but fixed field $\mathbb{F}$, and in the SoS the coefficients are over the reals. We say a proof $\pi$ is a refutation of $\mathcal{Q}$, if it is a proof of the claim (in the specified language) that there exists no assignment of $\vec{x} \in \mathbb{F}^n$ that satisfies *all* the polynomial equations in $\mathcal{Q}$. In PC and SoS, the proof $\pi$ is itself expressed as a sequence of polynomials. Two common measures of the complexity of a semi-algebraic proof are size (the number of monomials appearing in the proof) and the degree (the largest degree of the proof polynomials that refute $\mathcal{Q}$, see **Definition 2.1.2**). Trade-offs between the two are well known; in particular, any degree $d$ proof has size at most $n^{O(d)}$. Therefore, in this work we focus on the former[3]. We denote the smallest maximum degree over all proofs that refute $\mathcal{Q}$ in PC and SoS, with $\mathrm{Deg}\left(\mathcal{Q} \underset{\mathrm{PC}(\mathbb{F})}{\vdash} \bot\right)$ and $\mathrm{Deg}\left(\mathcal{Q} \underset{\mathrm{SOS}}{\vdash} \bot\right)$, respectively. One motivation for proving lower bounds for algebraic proof systems, as opposed to propositional proof systems, is that often they imply lower bounds for a broad family of related algorithms for solving combinatorial optimisation problems. Similarly, upper bounding the proof length has led to the fruitful discovery of many efficient algorithms. The SoS proof system is of particular interest because of its close connection to the sum-of-squares hierarchy of semi-definite programming. We refer the reader to the survey by Noah Fleming, Pravesh Kothari, Toniann Pitassi, and others. [13] for more details about the connections between the semi-algebraic proof systems and combinatorial optimisation. In this work, we study the complexity of refuting *perfect matchings* in PC and SoS. Apart from being a natural problem in its own right, perfect matchings are also related to the pigeon hole principle [6, 22, 27, 30, 31] and Tseitin formula [12, 14–16], two well studied formulae in proof complexity. Assuming at most one pigeon fits in a single hole, the pigeon hole principle says $m$ pigeons cannot fit in $n < m$ holes. If we construct the complete bipartite graph with the left vertices as $m$ pigeons and the right vertices as $n < m$ holes, proving the pigeon hole principle amounts to proving that such a bipartite graph does not have a perfect matching. There are other formulations of the pigeon hole principle (see the survey by Alexander A. Razborov. [30]), and almost all of them have short proofs in the sum of squares proof system. In contrast, Tseitin formulae are known to require long proofs. The Tseitin formula over a graph claims that there is a spanning subgraph in which every vertex has odd degree. If a graph has a perfect matching, then the subgraph described by the matching ensures that every vertex has odd degree. However, formally refuting Tseitin formulae for expander graphs with an odd number of vertices, in the SoS proof system, requires degree linear in the number of vertices in the graph [16]. Given its close connections to the pigeon-hole and Tseitin, and the different behaviour of the two formulae, it is natural to determine the complexity of refuting perfect matchings for non-bipartite graphs.

To refute perfect matchings in an algebraic proof system, we first need to specify combinatorial constraints as algebraic equalities. Given an undirected graph $G = (V, E)$, $V = \{1, \ldots, n\}$, and a vector $\vec{b} = (b_1, \ldots, b_n) \in \mathbb{F}^n$, we define $\mathrm{Card}(G, \vec{b})$ as the following set of polynomial constraints over variables $x_e$ for $e \in E$:

[3] Note that exponential size lower bounds only follow from degree lower bounds $d \gg \sqrt{n}$, as $d = O(\sqrt{n})$ yields only subexponential size bounds $n^{O(\sqrt{n})} = 2^{O(\sqrt{n} \log n)}$.

$$\mathrm{Card}(G, \vec{b}) := \begin{cases} x_e(1-x_e) = 0 & \text{for every } e \in E \\ \sum_{e \sim v} x_e \quad = b_v & \text{for every } v \in V \end{cases}$$

where we use notation $e \sim v$ to denote the set of edges $e$ incident on node $v$. For every $e \in E$, the equation $x_e(1-x_e) = 0$ restricts the domain of the above variables to bits. In plain words, $\mathrm{Card}(G, \vec{b})$ denotes the claim that there exists a spanning subgraph $G' \subseteq G$ such that a vertex $v \in V(G)$ has $b_v$ edges incident to it in $G'$. Note if there was an assignment of variables $(x_e)_{e \in E}$ that satisfies all equations in $\mathrm{Card}(G, \vec{1})$, where $\vec{1} = (1, \ldots, 1) \in \mathbb{F}^n$, it would imply that the graph $G$ has a perfect matching (given by the edges corresponding to variables with assignment 1). Therefore, we define $\mathrm{PM}(G) := \mathrm{Card}(G, \vec{1})$. When $|V|$ is odd, $G$ trivially does not contain a perfect matching. How difficult is it to refute $\mathrm{PM}(G)$ in this case? In recent work, Per Austrin and Kilian Risse. [5] showed that refuting $\mathrm{PM}(G)$, in the Polynomial Calculus and Sum-of-Squares system, in the case $G$ is a *random $d$-regular graphs* with an odd number of vertices typically requires proofs with degree $\Omega\left(\frac{n}{\log n}\right)$. They conjecture that the hardness results should also apply to general expander graphs but leave showing so as an open problem [5: see Section 6]. In this work, we verify this by extending their result to all $d$-regular spectral expanders, that is, $d$-regular graphs with a mild condition on the spectral gap. In fact, similar to Austrin and Risse, we reduce the hardness of refuting $\mathrm{Card}(G, \vec{t})$, where $\vec{t} = (t, \ldots, t)$, for any odd value $t$, to the hardness of refuting $\mathrm{Card}(G, \vec{1})$, where $\vec{1} = (1, \ldots, 1)$. As another special case, this answers the *even-colouring* case when $t = \frac{d}{2}$ is odd, a problem posed by Buss and Nordström [7: see Open Problem 7.7], which asks, *"Are even colouring formulas over expander graphs hard for polynomial calculus over fields of characteristic distinct from 2 ?"* Formally, we prove the following (for the definition of $(n, d, \lambda)$-graphs see **Section 2.2**).

**Theorem 1.1 (Hardness Result For $\mathrm{Card}(G, \vec{t})$)**

There exist universal constants $\varepsilon, n_0, d_0 \in \mathbb{N}$ such that for any **odd** $n \geq n_0$ and **even** $d \in [d_0, n]$, the following holds for **any** $(n, d, \lambda)$-graph $G$ with $\lambda < \varepsilon d$, and for any **odd** $1 \leq t \leq d$:

$$\mathrm{Deg}\left(\mathrm{Card}(G, \vec{t}) \underset{\mathrm{PC}(\mathbb{F})}{\vdash} \bot =\right) = \Omega\left(\frac{n}{\log n}\right)$$

$$\mathrm{Deg}\left(\mathrm{Card}(G, \vec{t}) \underset{\mathrm{SOS}}{\vdash} \bot =\right) = \Omega\left(\frac{n}{\log n}\right)$$

We follow the overall approach of Per Austrin and Kilian Risse. [5]. Very briefly, the strategy is to obtain an affine restriction (see **Definition 2.1.6**) $\mathrm{Card}(G, \vec{t})|_\rho \equiv \mathrm{PM}(H)$ where $H$ is some graph for which refuting $\mathrm{PM}(H)$ requires large degree. An example of such $H$ is given by Sam Buss, Dima Grigoriev, Russell Impagliazzo, and Toniann Pitassi. [8]. We now describe how to find such a restriction in more details: Using a result of Nemanja Draganić, Michael Krivelevich, and Rajko Nenadov. [11], we show that $H$ topologically embeds into a given expander graph $G$ with

$\lambda < \varepsilon d$ for some universal small constant $\varepsilon \in (0,1)$, such that all paths corresponding to the embedding have odd length. The main technical ingredient of Austrin and Risse is also a similar embedding theorem, albeit a significantly more complicated one. Moreover, we show that one can find such an embedding so that the subgraph of $G$ induced by vertices which are not part of the embedding has a perfect matching. This allows us to use the restriction argument to transfer the hardness of $\mathrm{PM}(H)$ into the hardness of $\mathrm{PM}(G)$. To extend this to hardness of $\mathrm{Card}(G, \vec{t})$ for an odd $3 \leq t \leq d$, it suffices to show that the graph $G'$ obtained from $G$ by removing all edges that participate in the embedding and the matching contains a $(t-1)$-regular spanning subgraph. Austrin and Risse achieve this using the contiguity property of random regular graphs (and hence their hardness result for $\mathrm{Card}(G, \vec{t})$ critically relies on randomness). Instead, we provide a significantly simpler and shorter argument based on Tutte's criterion. As a random $d$ -regular graph is with high probability an $(n, d, \lambda)$-graph with $\lambda = \Theta\left(\sqrt{d}\right)$ ( [34: see Theorem A], our embedding theorem readily applies in the context of [5].

The rest of the document is structured as follows. In **Section 2**, we describe the requisite background from graph theory and proof complexity. In **Section 3**, we describe the machinery for finding a desired topological embedding. In **Section 4**, we prove conditions under which the residual graph has a perfect matching or, more generally, a $(t-1)$-regular spaning subgraph. In **Section 5**, we use the tools from the previous sections to prove **Theorem 1.1**. In Section **Section 6** we briefly discuss a few other lower bounds using embeddings in proof complexity, and conclude with some future directions.

# 2 Preliminaries

## 2.1 Proof Complexity Preliminaries

Let $\mathcal{Q} = \{p_1 = 0, \ldots, p_m = 0\}$ be a set of polynomial equations[4], which we refer to as axioms, over variables $\vec{X} = \{x_1, \ldots, x_n, \overline{x}_1, \ldots, \overline{x}_n\}$.

[4]The sum of squares proof system is a semi-algebraic proof system where $\mathcal{Q}$ may also contain inequalities of the form $p_i(x) \geq 0$. However, we only need equality constraints to express the existence of Perfect Matchings over graphs. Therefore to simplify our exposition, we write all our definitions using equality constraints only.

**Definition 2.1.1 (Sum Of Squares Refutations)**

Given a set of $m$ polynomial equality constraints $\mathcal{Q}$ over the reals, a Sum of Squares (SoS) refutation is a sequence of polynomials $\pi = (t_1, ..., t_m; s_1, ..., s_a)$ such that

$$h := \sum_{i\in[m]} t_i p_i + \sum_{i\in[a]} s_i^2 = -1$$

The degree of a proof $\pi$ is

$$\mathrm{Deg}(\pi) := \max\left\{\max_{i\in[m]} \mathrm{Deg}(t_i) + \mathrm{Deg}(p_i), \max_{i\in[a]} 2\ \mathrm{Deg}(s_i)\right\}$$

Note that $s_i^2(x) \geq 0$ for any $x$ by definition. Therefore, if there were to exist some $x^*$ such that $p_i(x^*) = 0$ for all $p_i \in \mathcal{Q}$, then $\sum_{i\in[m]} t_i(x^*)p_i(x^*) = 0$. This would imply that $h \geq 0$, but if proof $\pi$ shows that $h = -1$, then by the contrapositive, no such $x^*$ can exist. Therefore, the existence of the sequence of polynomials $\pi$ act as a formal proof of the claim that the set of polynomial equations in $\mathcal{Q}$ is unsatisfiable.

**Definition 2.1.2 (Complexity Of SoS Refutation)**

If we let $\Pi$ denote the set of all valid SoS refutations for $\mathcal{Q}$, then the complexity of refuting $\mathcal{Q}$ in the SoS proof system is given by

$$\mathrm{Deg}\left(\mathcal{Q} \underset{\text{SOS}}{\vdash} \perp\right) := \min_{\pi\in\Pi} \mathrm{Deg}(\pi)$$

Polynomial Calculus (PC) is a dynamic version of the static Nullstellensatz proof system [13: see Section 1.3 for the defintion of Nullstellensatz proof systems] operating over an arbitrary but fixed field, based on the following inference rules.

1. From polynomial equations $f = 0$ and $g = 0$ where $f, g \in \mathbb{F}\left[\vec{X}\right]$ we can derive $\alpha f + \beta g = 0$ for $\alpha, \beta \in \mathbb{F}$.
2. From polynomial $f = 0$ where $f \in \mathbb{F}\left[\vec{X}\right]$, we can derive $xf = 0$ where $x \in \vec{X}$.

**Definition 2.1.3 (Polynomial Calculus Refutations)**

A Polynomial Calculus (PC) refutation of $\mathcal{Q}$ over $\mathbb{F}$ is a sequence of polynomials $\pi = (t_1, ..., t_\ell)$ such that $t_\ell = 1$, and for each $i \neq l$, either (1) $t_i \in \mathcal{Q}$, or (2) $t_i$ is derived from $\left(t_j\right)_{j<i}$ using the above rules. The degree of the proof is given by $\mathrm{Deg}(\pi) = \max_{i\in l} \mathrm{Deg}(t_i)$. If we let $\Pi$ denote the set of all PC refutations of $\mathcal{Q}$, then

$$\mathrm{Deg}\left(\mathcal{Q} \underset{\text{PC}(\mathbb{F})}{\vdash} \perp\right) := \min_{\pi\in\Pi} \mathrm{Deg}(\pi)$$

To ensure Boolean variables, we assume the axioms $\mathcal{Q}$ always contain the equations $x_i^2 - x_i = 0$ and $\overline{x}_i^2 - \overline{x}_i = 0$ for all $i \in [n]$. Equivalently, we can also just work in the ring $\mathbb{F}[x_1, ..., \overline{x}_n]/(x_1^2 - x_1, ..., \overline{x}_n^2 - \overline{x}_n)$ of multi-linear polynomials. Multi-linearity implies that the degree of any proof can be at most $n$ i.e. a proof of degree $\Omega(n)$ is the largest lower bound one can hope to achieve. Additionally, we will also assume that $1 - x_i - \overline{x}_i = 0$ is also included in $\mathcal{Q}$, for all $i \in [n]$, which ensures that the bar elements are bit complements of the non-bar elements. The following lemma is by Sam Buss, Dima Grigoriev, Russell Impagliazzo, and Toniann Pitassi. [8] and gives an instance where perfect matching is hard to refute in the worst case.

**Lemma 2.1.4 (Worst Case Hard Instance For PC)**

Given any odd $n \in \mathbb{N}$, there exists a graph $H$ with $n$ vertices and maximum degree $\Delta_H = 5$ such that Polynomial Calculus over any field of characteristic different from 2 requires degree $\Theta(n)$ to refute $\text{Card}(H, \vec{1})$.

A description of the worst case hard instance for SoS can be found in [5:Theorem A.3] which is also derived from [8].

**Lemma 2.1.5 (Worst Case Hard Instance For SOS)**

Given any odd $n \in \mathbb{N}$, there exists a graph $H$ with $n$ vertices and maximum degree $\Delta_H = 5$ such that SoS refutations requires degree $\Theta(n)$ to refute $\text{Card}(H, \vec{1})$.

An important lemma we will need is that given a set of axioms $\mathcal{Q}$ over the ring $\mathbb{F}[x_1, ..., x_n]$, a partial assignment of variables can only make refuting $\mathcal{Q}$ easier. Given a set of $m$ polynomial equality constraints $\mathcal{Q}$ over boolean variables $\{x_1, ..., x_n\}$, let the family of functions $\{f_i : \{0,1\}^n \to \{\text{True}, \text{False}\}\}_{i \in [m]}$, denote predicates for satisfiability for each constraint. For example, given $\alpha \in \{0,1\}^n$, $f_i(\alpha) = \text{True}$ if the $i$'th polynomial constraint $q_i \in \mathcal{Q}$ is satisfied i.e $q_i(\alpha) = 0$. We say $\mathcal{Q}$ is satisfied if there exists $\alpha \in \{0,1\}^n$ such that $f_i(\alpha) = \text{True} \Leftrightarrow q_i(\alpha) = 0$ for all $i \in [m]$. Given a map $\rho : \{x_1, ..., x_n\} \to \{x_1, ..., x_n, \overline{x}_1, ..., \overline{x}_n, 1, 0\}$, the restriction of a function $f : \{0,1\}^n \to \{0,1\}$, denoted by $f|_\rho$, is defined as $f|_\rho\ (x_1, ..., x_n) = f(\rho(x_1), ..., \rho(x_n))$. Similarly, the restriction of formula $\mathcal{Q}$ is defined as $\mathcal{Q}|_\rho = \left\{f_{\{1\}}|_\rho, ..., f_{\{m\}}|_\rho\right\}$. Two formula $\mathcal{Q}$ and $\mathcal{Q}'$ are equivalent if they are element-wise equal, ignoring any functions that are constantly True. For example, $\mathcal{Q} = \{f_a, f_b, \text{True}\}$ and $\mathcal{Q}' = \{f_a, f_b\}$ are equivalent, denoted as $\mathcal{Q} \equiv \mathcal{Q}'$.

**Definition 2.1.6 (Affine Restriction)**

We say that an axiom $\mathcal{Q}'$ is an affine restriction of $\mathcal{Q}$ if there is a map $\rho : \{x_1, ..., x_n\} \to \{x_1, ..., x_n, \overline{x}_1, ..., \overline{x}_n, 1, 0\}$ such that $\mathcal{Q} \equiv \mathcal{Q}|_\rho$.

**Lemma 2.1.7**

Let $\mathcal{Q}, \mathcal{Q}'$ be axioms such that $\mathcal{Q}'$ is an affine restriction of $\mathcal{Q}$, and each axiom of $\mathcal{Q}$ depends on a constant number of variables, then

- For any arbitrary but fixed $\mathbb{F}$ it holds that $\mathrm{Deg}\left(\mathcal{Q} \underset{\mathrm{PC}(\mathbb{F})}{\vdash} \bot\right) \in \Omega\left(\mathrm{Deg}\left(\mathcal{Q}' \underset{\mathrm{PC}(\mathbb{F})}{\vdash} \bot\right)\right)$
- $\mathrm{Deg}\left(\mathcal{Q} \underset{\mathrm{SOS}}{\vdash} \bot\right) \in \Omega\left(\mathrm{Deg}\left(\mathcal{Q}' \underset{\mathrm{SOS}}{\vdash} \bot\right)\right)$

The proof for Lemma 2.1.7 can be found in [5: see Lemma 2.2]. What the above lemma says is that if we have a graph $G$ with odd vertices with constant degree, that has a perfect matching on a subset of even vertices on the graph, then the size of the proof to refute $\mathrm{PM}(G)$ is at least as large as refuting a perfect matching in $G$ with the even vertices removed.

## 2.2 Graph Theory Preliminaries

We use standard graph theoretic notation. For a graph $G$, we use $V(G)$ and $E(G)$ to denote the vertices and edges of $G$. For a vertex $v \in V(G)$, we use $\Gamma_G(v) = \{u \in V(G) : (u, v) \in E'\}$ to denote the neighbourhood of $v$ in $G$, and $\deg_G(v) := |\Gamma_G(v)|$. Given two sets $S, T \subseteq V(G)\}$, we use $e_G(S, T)$ to denote the number of edges in $G$ with one endpoint in $S$ and one endpoint in $T$. Note that we do not require $S$ and $T$ to be disjoint; in case they are not disjoint, every edge with both endpoints in $S \cap T$ is counted twice in $e_G(S, T)$. If the graph $G$ is clear from the context, we omit the subscript. Given two vertices $u$ and $v$, we use $u \rightsquigarrow v$ to denote the sequence of edges in the path from $u$ and $v$. Given $W \subseteq V(G)$, we denote with $G[W]$ the subgraph of $G$ induced by $W$. We say that a subgraph $G' \subseteq G$ is *spanning* if $V(G') = V(G)$. Next, we give a definition of pseudorandom graphs.

**Definition 2.2.1 ($(n, d, \lambda)$-graphs)**

Let $G$ be a $d$-regular graph on $n$ vertices, and, let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ denote eigenvalues of the adjacency matrix of $G$. We say $G$ is an $(n, d, \lambda)$-graph if $\lambda(G) := \max_{2,\ldots,n} \max|\lambda_i| \leq \lambda$.

The following is a well known result of N. Alon and F. R. K. Chung. [3].

**Lemma 2.2.2 (Expander Mixing Lemma)**

Given an $(n, d, \lambda)$-graph $G$, for any $S, T \subseteq V(G)$ we have

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

We make use of the following two well known criteria of Tutte [35, 36]. Note that both of the

lemmata ask for properties which are stronger than what Tutte criteria requires[5] however, they are easier to state and verify in our application. We denote with $q(G)$ the number of connected components in a graph $G$.

**Lemma 2.2.3 (Tutte's Criterion)**

If a graph $G$ has *even* number of vertices and for every subset $S \subseteq V(G)$ we have $q(G \setminus S) \leq |S|$, then $G$ contains a perfect matching.

**Lemma 2.2.4 (Tutte's Generalised Criterion)**

Let $f \in \mathbb{N}$ be even. Suppose $G$ is a graph such that for every pair of disjoint sets $S, T \subseteq V(G)$ the following holds:

$$q(G \setminus (S \cup T)) \leq |S|f - \sum_{w \in T}(f - |\Gamma_G(w) \setminus S|)$$

Then $G$ contains a spanning subgraph $G' \subseteq G$ which if $f$-regular.

## 2.3 Probabilistic Tools

Next we introduce standard tools for randomised algorithms. A dependency graph for a set of events $E_1, ..., E_n$ is a graph $G = (V, E)$ such that $V = \{1..., n\}$ and, for $i = 1, ..., n$, event $E_i$ is mutually independent of the events $\{E_j \mid (i, j) \notin E\}$. The degree of the dependency graph is the maximum degree of any vertex in the graph.

**Lemma 2.3.1 (Lovász Local Lemma)**

Let $E_1, ..., E_n$ be a set of events over some probability space with probability $\mathcal{D}$, and assume that for some $\beta \in (0, 1)$ the following hold:

- The degree of the dependency graph given by $(E_1, ..., E_n)$ is bounded by $d$.
- For all $i \in [n]$, $\mathsf{Pr}_{E_i \leftarrow \$\mathcal{D}}[E_i] \leq \beta$.
- $\beta \leq \frac{1}{4d}$.

Then $\mathsf{Pr}[\cap_{i=1}^{n} E_i] > 0$

[5] More specifically, in the Tutte criterion $q(G)$ denotes the number of *odd* sized connected components.

**Lemma 2.3.2 (Multiplicative Chernoff bound)**

Suppose $X_1, \ldots, X_n$ are identical independent random variables taking values in $\{0,1\}$. Let $X$ denote their sum and let $\mu = n\mathbb{E}[X_1]$ denote the sum's expected value. Then, for any $0 < \delta < 1$, we have

$$\Pr[|X - \mu| \geq \delta\mu] \leq 2e^{-(\delta^2\mu)/3}$$

The proof of Lemma 2.3.2 and Lemma 2.3.1 can be found in any textbook on randomised algorithms (for example, see [23: see Ch. 1 and 7]).

Lemma 2.3.3 is originally by [5: see Lemma 4.3], re-derived here for completeness.

**Lemma 2.3.3 (Partition Theoem)**

For every $0 < c < 1$ and $\gamma > 0$, there exists $d_0$ such that the following holds. If $G$ is a $d$-regular graph, for some $d \geq d_0$, then there exists a subset $A \subseteq V(G)$ such that

$$cd - \gamma d \leq |\Gamma_G(v) \cap A| \leq cd + \gamma d \tag{1}$$

for every $v \in V(G)$.

*Proof.* We prove the existence of such a partition $A \subseteq V(G)$ using the probabilistic method. For each $v \in V(G)$, we toss an independent coin $X_i$ with bias $c$. We include $v$ in $A$ if and only if $X_i = 1$. Thus, $\vec{X} := (X_1, \ldots, X_n) \in \{0,1\}^n$ is a random variable that describes how we choose $A$. For any $v \in V$, let $Y_v := |\Gamma_G(v) \cap A|$ denote the random variable that counts the number of neighbours of $v$ in $A$. Define $\delta := \frac{\gamma}{c}$, and for every $v \in V$ let $E_v = \mathbb{1}(|Y_v - dc| \geq \delta cd)$ denote the bad event that $v$ has too many or too few neighbours in $A$. Observe that the dependency graph of events $\{E_v\}_{\{v \in V\}}$ has maximum degree at most $d^2$ (only vertices at most two hops away from $v$ affect how many of $v$'s neighbours are in $A$; there at most $d^2$ such vertices). As $G$ is $d$-regular, $\mathbb{E}_{\vec{X}}[Y_v] = cd$. By the Lemma 2.3.2, for any $v \in V$ we have $\Pr_{\vec{X}}[E_v] \leq 2e^{-\delta^2 cd/3} =: \beta$. For $d$ sufficiently large we have $\beta \leq \frac{1}{4d^2}$, and so $\beta d^2 \leq \frac{1}{4}$. All the conditions of Lemma 2.3.1 are satisfied, from which we conclude that, with positive probability, none of the bad events happen. This implies the desired $A \subseteq V(G)$ exists. □

# 3 Topological embedding

In this section we describe the topological embedding result of Nemanja Draganić, Michael Krivelevich, and Rajko Nenadov. [11]. We start with a necessary definition.

**Definition 3.1 (Sub-divisions)**

Given a graph $H$ and a function $\sigma : E(H) \to \mathbb{N}$, the $\sigma$-subdivision of $H$, denoted by $H^\sigma$, is the graph obtained by replacing each edge in $E(H)$ with a path of length $\sigma(e)$ joining the end points of $e$ such that all these paths are mutually vertex disjoint, except at the end points.

If a graph $G$ contains $H^\sigma$ for some $\sigma : E(H) \to \mathbb{N}$, then we say $G$ contains $H$ as a *topological minor*. In our application, it will be important that we can control the parity of $\sigma(e)$. The following result follows directly from [11: see Theorem 1].

**Theorem 3.2 (Embedding Theorem)**

For every $D \in \mathbb{N}$ there exist $\alpha, \xi, C > 0$, such that the following holds. Suppose $G$ is a graph with $n$ vertices and $m \geq Cn$ edges such that for every pair of disjoint subsets $S, T \subseteq V(G)$ of size $|S|, |T| \geq \xi n$, we have

$$\left| e_{\{G\}}(S,T) - |S\|T|p \right| \leq \xi \ |S\|T| \ p$$

where $p = m/\binom{n}{2}$. Then $G$ contains $H^\sigma$, where $H$ is any graph with maximum degree at most $D$, $H^\sigma$ has at most $\alpha n$ vertices, and $\sigma(e) \geq \log n$ for every $e \in E(H)$.

When $G$ is an $(n, d, \lambda)$ graph, we will we make use of Theorem 3.2 to show that $G$ satisfies the required properties, thereby contains $H$ as a topological minor. This gives us the following corollary.

**Corollary 3.2.1**

For every $D \in \mathbb{N}\}$ there exist $d_0, n_0 \in \mathbb{N}, \varepsilon, \alpha \in (0,1)$, such that the following holds. Suppose $G$ is an $(n, d, \lambda)$-graph where $d \geq d_0$, and $\lambda < \varepsilon d$, and $n \geq n_0$. Let $B \subseteq V(G)$ be a subset of size $|B| \geq \frac{n}{20}$, and $H$ is any graph with maximum degree at most $D$ and at most $\alpha \frac{n}{\log n}$ vertices. Then the induced sub-graph $G[B]$ contains $H^\sigma$ such that $\sigma(e)$ is odd for every $e \in E(H)$.

*Proof.* Let $m$ denote the number of edges in the induced subgraph $G[B]$, which gives us $2m = e_G(B,B)$. Denote $b := |B|$ and define $p = m/\binom{b}{2}$. By the Lemma 2.2.2, we have

$$\left| 2m - \frac{d}{n} b^2 \right| \leq \lambda b \tag{2.1}$$

Dividing both sides with $b(b-1)$, and observing that $\frac{db^2}{n(b-1)b} = \left(\frac{d}{n}\right)\left(\frac{b}{b-1}\right) = \frac{d}{n}\left(1 + \frac{1}{b-1}\right)$, we further get

$$\left|\frac{2m}{b(b-1)} - \frac{d}{n} - \frac{d}{n(b-1)}\right| \leq \frac{\lambda}{b-1} \tag{3.1}$$

From this we have

$$\left|p - \frac{d}{n}\right| = \left|p - \frac{d}{n} - \frac{d}{n(b-1)} + \frac{d}{n(b-1)}\right| \tag{4.1}$$

$$\leq \left|p - \frac{d}{n} - \frac{d}{n(b-1)}\right| + \frac{1}{(b-1)} \tag{4.2}$$

$$\leq \frac{\lambda}{(b-1)} + \frac{1}{b-1} \tag{4.3}$$

$$\leq \frac{2\lambda}{b} \tag{4.4}$$

Let us briefly justify each step: Equation 4.2 comes from the triangle inquality and $\frac{d}{n} \leq 1$; Equation Equation 4.3 comes from Equation Equation 3.1; the last inequality comes from the assumption that $b \in \Omega(n)$, so the inequality holds for $n$ large enough.
Let $\xi$ be as given by the Theorem 3.2. Using the bound on the difference between $p$ and $\frac{d}{n}$, for every disjoint subsets $S, T \subseteq B$ of size $|S|, |T| \geq \xi n$, for $\lambda < \varepsilon d$ where $\varepsilon$ is sufficiently small, we have

$$|e_G(S,T) - p|S\|T|| \leq \left|e_G(S,T) - \frac{d}{n}|S\|T|\right| + \left|\frac{d}{n}|S\|T| - p|S\|T|\right| \tag{5.1}$$

$$\leq \lambda\sqrt{|S\|T|} + \frac{2\lambda}{b}\ |S\|T| \leq \xi\ |S\|T|p \tag{5.2}$$

With the lower bounds on $S, T$ and $B$, we can make $\varepsilon$ sufficiently small with respect to $\xi$ to get the upper bound in the last step. Let $\sigma : E(H) \to \mathbb{N}$ be the constant function where $\sigma(e)$ is the smallest odd integer larger than $\log n$. As $G[B]$ has at least $Cn$ edges (by the Lemma 2.2.2), $\sigma(e) \leq 2 + \log n$, and $H$ has at most $\alpha\frac{n}{\log n}$ vertices, we can invoke the Theorem 3.2 to conclude that $G[B]$ contains $H^\sigma$.

□

# 4 Perfect matching and regular subgraphs

As described earlier, the second ingredient in our hardness proof is showing that a certain residual graph contains a perfect matching or a spanning $(t-1)$-regular subgraph. In this section we state and prove these ingredients.

**Lemma 4.1 (Perfect Matching Lemma)**

Let $G$ be an $(n, d, \lambda)$-graph with $\lambda < d/50$, and suppose $G' \subseteq G$ satisfies $\delta(G') \geq 0.9d$. Then for all $S \subseteq V(G')$, $G' \setminus S$ has at most $|S|$ connected components, that is, $q(G' \setminus S) \leq |S|$. Therefore, if $G'$ has an even number of vertices then it contains a perfect matching.

*Proof.* Let $U = V(G')$. We aim to show that the graph $G' \setminus S$ has at most $|S|$ connected components. If $|S| \geq |U|/2$ then $G' \setminus S$ has at most $|S|$ vertices, so the the upper bound on connected components trivially holds. For the remainder of the proof we can assume $|S| < |U|/2$. We claim the following:

> **Claim:** For every partition $X \cup Y = U \setminus S$, with $|X|, |Y| \geq \frac{|S|}{3}$, we have
>
> $$e_{G'}(X, Y) \geq 1 \Rightarrow q(G' \setminus S) \leq |S|$$

> *Proof Of Claim.* To see why, assume towards a contradiction that there exists an edge in $G'$ between *every* partition $X \cup Y = U \setminus S$, where $|X|, |Y| \geq |S|/3$, *and* $G' \setminus S$ has more than $|S|$ connected components. Denote the vertex sets of these components by $C_1, ..., C_k$, for some $k > |S|$. Let $X^* := C_1 \cup ... \cup C_s$ and $Y^* := C_{s+1} \cup ... \cup C_k$, where $s = \lfloor |S|/2 \rfloor \geq |S|/3$. By construction, even if each component $C_i$ is a singleton set, we get that $|X^*|, |Y^*| \geq |S|/3$. Now as all $C_i$'s are disjoint connected components, there can be no edge between $X^*$ and $Y^*$. Therefore, we have found a partition $X^* \cup Y^* = U \setminus S$ with $|X^*|, |Y^*| \geq |S|/3$ without an edge between them, which contradicts our assumption that all appropriately sized partitions have at least one edge between them. □

To complete our main proof, it suffices to show $e_{G'}(X, Y) \geq 1$ for every partition $X \cup Y = U \setminus S$ with $|X|, |Y| \geq |S|/3$.

Consider some arbitrary partition $X \cup Y$ of $U \setminus S$, with $|X|, |Y| \geq |S|/3$, and without loss of generality assume $|X| \leq |Y|$. Then by a simple counting argument we get

$$|X| \leq \frac{|U| - |S|}{2} \leq \frac{n - |S|}{2} \tag{6.1}$$

We have:

$$\begin{aligned} e_{G'}(X, X) + e_{G'}(X, S) &\leq e_G(X, X) + e_G(X, S) && (7.1)\\ &\leq \frac{d}{n} |X|^2 + \lambda |X| + \frac{d}{n} |X||S| + \lambda\sqrt{|X||S|} && (7.2)\\ &\leq \frac{d}{n} |X| \frac{(n - |S|)}{2} + \lambda |X| + \frac{d}{n} |X||S| + \lambda\sqrt{3} |X| && (7.3)\\ &< \frac{d|X|}{2} + \frac{d|X||S|}{2n} + 3\lambda |X| && (7.4) \end{aligned}$$

$$< \frac{d}{2} |X| + \frac{d}{4} |X| + 3\lambda |X| \tag{7.5}$$

$$< \frac{9d}{10} |X| \tag{7.6}$$

These steps are justified as follows: The first equation follows from the Lemma 2.2.2; Equation 7.3 comes from Equation 6.1 and $|X| \geq |S|/3$; Equation 7.5 comes from $|S| < \frac{n}{2}$; and Equation 7.6 comes the assumption $\lambda < \frac{d}{50}$. By the assumption $\delta(G') > 0.9d$ we conclude that there is an edge in $G'$ with one vertex in $X$ and the other in $V(G') \setminus (X \cup S) = Y$.

□

The next lemma shows that subgraphs of $(n, d, \lambda)$-graphs with large minimum degree contain regular spanning subgraphs.

**Lemma 4.2 (Regular Subgraph Lemma)**

For every $C > 1$ there exists $d_0 = d_0(C)$ such that the following holds. Suppose $G$ is an $(n, d, \lambda)$ graph with $\lambda < \varepsilon d$ and $d \geq d_0$, where $\varepsilon < 1/\left(100C^{\frac{3}{2}}\right)$. If $G' \subseteq G$ has minimum degree $\delta_{G'} \geq d - C$, then $G'$ contains a spanning $f$-regular subgraph for any even $2 \leq f \leq d/2$.

*Proof.* We prove this lemma using Lemma 2.2.4. We need to show that for any pair of disjoint sets $S, T \subseteq V(G')$, we have

$$q(G' \setminus (S \cup T)) \leq |S| f - \sum_{w \in T} (f - |\Gamma_{G'}(w) \setminus S|) \tag{8.1}$$

As $\varepsilon < 1/\left(100C^{\frac{3}{2}}\right)$ and $C > 1$, we have that $\varepsilon < 1/100$. This implies that $G$ is an $(n, d, \lambda)$ graph with $\lambda < \frac{d}{100}$. We set $d_0 := d_0(C)$ large enough such that for all $d \geq d_0$, even after deleting at most $C$ edges incident on each vertex of the $d$-regular graph $G$ to get $G'$, we have the minimum degree of $G'$ to be $\delta_{G'} \geq d - C > 9d/10$. Therefore the conditions of Lemma 4.1 are satisfied, thus

$$q(G' \setminus (S \cup T)) \leq |S \cup T| = |S| + |T| \tag{9}$$

To prove Equation 8.1, it suffices to show

$$|S| + |T| \leq |S|\ f - |T| f + |T|(d - C) - e_{G'}(S, T) \tag{10.1}$$

$$\leq |S|\ f - \sum_{w \in T} (f - |\Gamma_{G'}(w) \setminus S|) \tag{10.2}$$

We distinguish a few cases.

**Case 1**: Suppose $|S| \leq |T|$. As $f \leq d/2$, we have

$$|S|f + |T|(d - C - f) \geq (|S| + |T|)\left(\frac{d}{2} - C\right) \tag{11.1}$$

The condition described by the inequality Equation 10.1 is satisfied via the following analysis.

$$|S| + |T| + e_{G'}(S,T) \leq |S| + |T| + |S||T|\ \frac{d}{n} + \varepsilon d\sqrt{|S||T|} \tag{12.1}$$

$$\leq |S| + |T| + \frac{d}{4}(|S| + |T|) + \varepsilon d\frac{1}{2}(|S| + |T|) \tag{12.2}$$

$$\leq (|S| + |T|)\left(\frac{d}{4} + 1 + \frac{\varepsilon d}{2}\right) \tag{12.3}$$

$$< (|S| + |T|)\left(\frac{d}{2} - C\right) \tag{12.4}$$

$$\leq |S|f - |T|(d - C - f) \tag{12.5}$$

Equation 12.1 comes from the Lemma 2.2.2 and $\lambda < \varepsilon d$, together with an obvious upper bound $e_{G'}(S,T) \leq e_{G(S,T)}$. Equation Equation 12.2 comes from the fact that $|S||T| \leq \left(\frac{|S|+|T|}{2}\right)^2 \leq n\frac{(|S|+|T|)}{4}$. Equation 12.4 comes from $\varepsilon < 1/\left(100C^{\frac{3}{2}}\right)$, $C > 1$ and $d_0$ being sufficiently large. Equation 12.5 follows from Equation 11.1 which gives us what we want.

**Case 2**: Suppose $|S| > |T|$. As $f \geq 2$, we have

$$|S|f + |T|(d - C - f) \geq 2|S| + |T|(d - C - 2) \tag{13.1}$$

To show Equation 10.1, it suffices to show that

$$e_{G'}(S,T) \leq |S| + |T|(d - C - 3) \tag{14.1}$$

Now we distinguish between two subcases.

- If $|T| \leq \frac{|S|}{C+3}$, then Equation 14.1 follows from a trivial bound $e_{G'}(S,T) \leq |T|d$.
- $\frac{|S|}{C+3} < |T| < |S|$. As $|S| + |T| < n$ we have $|S| < n - \frac{|S|}{C+3}$, thus $|S| < n \leq \frac{C+3}{C+4}$.

Using the Lemma 2.2.2, we have

$$e_G(S,T) \leq e_{G'}(S,T) \leq \frac{d}{n}|S||T| + \varepsilon d\sqrt{|S||T|} \tag{15.1}$$

$$< d\frac{(C+3)}{(C+4)}|T| + \varepsilon d|T|\sqrt{(C+3)} \tag{15.2}$$

$$= |T|d \leq \left(\frac{C+3}{(C+4\}}\right) + \varepsilon\sqrt{C+3}) \tag{15.3}$$

$$< |T|d \leq \left(\frac{C+3}{C+4} + \frac{1}{2(C+4)}\right) \tag{15.4}$$

$$= |T|d \leq \left(1 - \frac{1}{2(C+4)}\right) \tag{15.5}$$

where the penultimate inequality follows from the upper bound on $\varepsilon$. For $d$ sufficiently large in terms of $C$ we obviously have

$$|T|d \leq \left(1 - \frac{1}{2(C+4)}\right) < |T|d - |T|(C+3) < |S| + |T|(d-C-3) \tag{16}$$

hence Equation 14.1 is satisfied.

□

# 5 Proof of Theorem Theorem 1.1

In this section we prove Theorem 1.1. As $\mathrm{Card}(G, \vec{t}) \equiv \mathrm{Card}(G, \vec{d} - \vec{t})$, without loss of generality we only prove the theorem for $t \leq d/2$. Let $G = (V, E)$ be an $(n, d, \lambda)$-graph on an odd number of vertices with $\lambda < \varepsilon d$, where $\varepsilon < 1/\left(100C^{\frac{3}{2}}\right)$ and $C = 6$. For sufficiently small constant $\alpha \in (0, 1)$, let $H$ denote the graph on $h = \alpha\frac{n}{\log n}$ vertices as given by Lemma 2.1.5 (to show lower bounds for PC, we use $H$ from Lemma 2.1.4). Recall that any SoS proof which refutes $\mathrm{PM}(H)$ has degree $\Omega(h)$. We now make use of $H$ to show the hardness of refuting $\mathrm{Card}(G, \vec{t})$. The idea is to find a restriction $\rho$ such that $\mathrm{Card}(G, \vec{t})|_\rho \equiv \mathrm{PM}(H)$. We achieve this through the following steps.

- Invoke Lemma 2.3.3 with parameters $c = 0.925$ and $\gamma = 0.025$ to get subsets $A \subseteq V(G)$ and $B = V(G) \setminus A$, such that for every $u \in V(G)$ we have

$$0.9d \leq |\Gamma_G(u) \cap A| \leq 0.95d \tag{17.1}$$

$$0.05d \leq |\Gamma_G(u) \cap B| \leq 0.1d \tag{17.2}$$

- From equations Equation 17.1 and Equation 17.2, $|B| > \frac{n}{20}$, with room to spare and $|E(G[B])| \geq \frac{nd}{20}$. By Corollary Corollary 3.2.1, $G[B]$ contains $H^\sigma$ such that each $\sigma(e)$ is odd. Let us denote a subgraph of $G[B]$ corresponding to $H^\sigma$ by $G_\psi$. We can describe $G_\psi$ as a function $\psi : V(H) \to B$ together with a collection of pairwise vertex-disjoint (other than at the endpoints) paths $\psi(u) \rightsquigarrow \psi(v)$ in $G[B]$, for $(u, v) \in E(H)$. Observe that it is at least as hard to refute[6] $\mathrm{PM}(G_\psi)$ as it is to refute $\mathrm{PM}(H)$. To see why, let $y_1, \ldots, y_{E(H)}$ denote the variables for the $\mathrm{PM}(H)$ formulae for each edge of $H$. We use as shorthand $\mathcal{Y} = (y_e)_{e \in E(H)}$ and $\overline{\mathcal{Y}} = (\overline{y}_e)_{e \in E(H)}$. Define a mapping $\rho' : E(G_\psi) \to \{0, 1, \mathcal{Y}, \overline{\mathcal{Y}}\}$ as follows. For each $(u, v) \in E(H)$, let $\rho'(x_e) = y_{(u,v)}$ where $e$ is the first edge on the path $\psi(u) \rightsquigarrow \psi(v)$. Subsequently, map each

[6]Note that this by itself does not guarantee it is hard to refute $\mathrm{PM}(G)$. We need item (3) and this to show hardness of refuting $\mathrm{PM}(G)$.

variable $x_e$ for $e \in \psi(u) \rightsquigarrow \psi(v)$ alternately to $y_{u,v}$ or $\overline{y}_{u,v}$, such that the edges of the path adjacent to $\psi(u)$ and $\psi(v)$ are set to $y_{(u,v)}$. This is always possible as $u \rightsquigarrow v$ has odd length. Observe that $\mathrm{PM}(G_\psi)|_{\rho'} \equiv \mathrm{PM}(H)$.

- As $n$ is odd and $\left|V\left(G_\psi\right)\right|$ is odd, we have that $U = V(G) \setminus V\left(G_\psi\right)$ has even size. From Equation 17.1 we have that $G[U]$ has minimum degree at least $\frac{9d}{10}$. As $\lambda < \frac{d}{50}$ (with room to spare), we can invoke Lemma 4.1 to conclude $G[U]$ has a perfect matching $M$.
- Consider the subgraph $G' \subseteq G$ obtained by deleting all edges $e \in E\left(G_\psi\right) \cup M$, where $M$ is the perfect matching from the step above. As $\Delta_H \leq 5$, every vertex $u \in G$ loses at most 5 edges in this process. Thus, we have $\delta_{G'} \geq d - 5$. As $\lambda < \varepsilon d$, by the Lemma 4.2 we have that $G'$ contains a $(t-1)$-regular spanning subgraph $G''$.

We finally define $\rho$ as follows:

$$\rho(x_e) := \begin{cases} \rho'(e) & \text{if } e \in E\left(G_\psi\right) \\ 1 & \text{if } e \in M \cup E(G'') \\ 0 & \text{otherwise} \end{cases}$$

Then $\mathrm{Card}(G, \vec{t})|_\rho \equiv \mathrm{PM}(H)$, thus our theorem follows from Lemma 2.1.7.

# 6 Related Work

In propositional[7] proof complexity, there a few prior examples of the strategy of embedding a worst case instance into a host graph to show lower bounds for a larger class of objects [18, 25]. Toniann Pitassi, Benjamin Rossman, Rocco A. Servedio, and Li-Yang Tan. [25] show Tseitin lower bounds for Frege proof systems by relying on the embedding result by J. Kleinberg and R. Rubinfeld. [19], which allows one to embed any bounded degree graph $H$ of size $O(n/(\mathrm{poly}(\log n)))$ into an expander graph on $n$ vertices as a minor (not necessarily a topological one). Michael Krivelevich and Rajko Nenadov. [21] simplify and improve the above embedding theorem to allow for embedding any graph $H$ with size $O(n/\log n)$ as an ordinary minor . However, embedding a hard instance $H$ into $G$ as an ordinary minor does not guarantee that the hardness of $H$ is preserved in the setting considered in this paper. In particular, it is entirely possible that one of the edge contractions to obtain the minor results in $H$ now being easy to refute. Consequently, these embedding theorems cannot be directly applied show hardness of refuting perfect matchings in our setting Instead, as described in section **Section 5**, one way to preserve hardness is to use embedding theorems that allow for topological embeddings that allow for edge sub-divisions of odd size [11, 24]. In order to get a topological embedding, Austrin and Risse modify the ordinary embedding theorem in [21] but critically rely on the host graph being random. In this work, we use the embedding theorem by Nemanja Draganić, Michael Krivelevich, and Rajko Nenadov. [11], which greatly simplifies the argument. Moreover, we avoid the use of

[7] As opposed to algebraic proof complexity

the contiguity argument present in [5] by directly utilising Tutte's criterion and the Expander Mixing Lemma.

In summary, we show degree lower bounds for refuting $\mathrm{Card}(G, \vec{t})$ for odd $t$ in $(n, d, \lambda)$ graphs in the SoS and PC proof systems. There is still a $\log n$ gap between the largest possible proof in such systems, and our lower bounds (similar to [5]). It is not inherently clear that such a gap should exist. The gap is an artefact of $d$ being constant, which makes the graphs sparse i.e we need $\Theta(\log n)$ edges to form a path between any two nodes. This implies, that $\Omega(n/\log n)$ is the largest hard instance we can topologically embed in any graph. Thus, if the worst case lower for refuting perfect matchings was indeed $\Omega(n)$, we would need a more direct proof of the statement without using a smaller hard instance. We leave the issue of resolving the tightness of our lower bound as an open problem for future work.

# References


[1] Jackson Abascal, Venkatesan Guruswami, and Pravesh K. Kothari. 2021. Strongly refuting all semi-random Boolean CSPs. *Proceedings of the 32nd annual ACM-SIAM symposium on discrete algorithms, SODA 2021, Alexandria, VA, USA, virtual, January 10–13, 2021*, 454–472. https://doi.org/10.5555/3458064.3458092

[2] M. Alekhnovich and A. Razborov. 2001. Lower Bounds for Polynomial Calculus: Non-Binomial Case. In *(FOCS '01)*, 2001. IEEE Computer Society, USA, 190.

[3] N. Alon and F. R. K. Chung. 1988. Explicit construction of linear sized tolerant networks. *Discrete Math.* 72, 1–3 (1988), 15–19. https://doi.org/10.1016/0012-365X(88)90189-6

[4] Albert Atserias and Tuomas Hakoniemi. 2020. Size-degree trade-offs for sums-of-squares and positivstellensatz proofs. In *Proceedings of the 34th Computational Complexity Conference (CCC '19)*, 2020. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, New Brunswick, New Jersey. https://doi.org/10.4230/LIPIcs.CCC.2019.24

[5] Per Austrin and Kilian Risse. 2022. Perfect Matching in Random Graphs is as Hard as Tseitin. *TheoretiCS* (December 2022), 979–1012. https://doi.org/10.46298/theoretics.22.2

[6] Paul Beame, Russell Impagliazzo, Jan Krajčíček, Toniann Pitassi, Pavel Pudlák, and Alan Woods. 1992. Exponential Lower Bounds for the Pigeonhole Principle. In *Proceedings of the Twenty-Fourth Annual ACM Symposium on Theory of Computing (STOC '92)*, 1992. Association for Computing Machinery, Victoria, British Columbia, Canada, 200–220. https://doi.org/10.1145/129712.129733

[7] Sam Buss and Jakob Nordström. 2021. Proof complexity and SAT solving. *Handbook of Satisfiability*, 233–350.

[8] Sam Buss, Dima Grigoriev, Russell Impagliazzo, and Toniann Pitassi. 1999. Linear Gaps Between Degrees for the Polynomial Calculus Modulo Distinct Primes. In *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing (STOC '99)*, 1999. Associ-

ation for Computing Machinery, Atlanta, Georgia, USA, 547–556. https://doi.org/10.1145/301250.301399

[9] Jonas Conneryd, Susanna F. De Rezende, Jakob Nordström, Shuo Pang, and Kilian Risse. 2023. Graph Colouring Is Hard on Average for Polynomial Calculus and Nullstellensatz. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, 2023. 1–11. https://doi.org/10.1109/FOCS57990.2023.00007

[10] Stephen A. Cook and Robert A. Reckhow. 2023. The Relative Efficiency of Propositional Proof Systems. In *Logic, Automata, and Computational Complexity: The Works of Stephen A. Cook* (1st ed.). Association for Computing Machinery, New York, NY, USA, 173–192. Retrieved from https://doi.org/10.1145/3588287.3588299

[11] Nemanja Draganić, Michael Krivelevich, and Rajko Nenadov. 2022. Rolling backwards can move you forward: on embedding problems in sparse expanders. *Trans. Am. Math. Soc.* 375, 7 (2022), 5195–5216. https://doi.org/10.1090/tran/8660

[12] Yuval Filmus, Massimo Lauria, Mladen Mikša, Jakob Nordström, and Marc Vinyals. 2013. Towards an understanding of polynomial calculus: new separations and lower bounds (extended abstract). *Automata, languages, and programming. 40th international colloquium, ICALP 2013, Riga, Latvia, July 8–12, 2013, Proceedings, Part I*, 437–448. https://doi.org/10.1007/978-3-642-39206-1_37

[13] Noah Fleming, Pravesh Kothari, Toniann Pitassi, and others. 2019. Semialgebraic proofs and efficient algorithm design. *Foundations and Trends® in Theoretical Computer Science* 14, 1–2 (2019), 1–221.

[14] Nicola Galesi, Dmitry Itsykson, Artur Riazanov, and Anastasia Sofronova. 2023. Bounded-depth Frege complexity of Tseitin formulas for all graphs. *Ann. Pure Appl. Logic* 174, 1 (2023), 23. https://doi.org/10.1016/j.apal.2022.103166

[15] Ludmila Glinskih and Dmitry Itsykson. 2017. Satisfiable Tseitin formulas are hard for nondeterministic read-once branching programs. *42nd international symposium on mathematical foundations of computer science, MFCS 2017, August 21–25, 2017, Aalborg, Denmark*, 12. https://doi.org/10.4230/LIPIcs.MFCS.2017.26

[16] Dima Grigoriev. 2001. Linear lower bound on degrees of Positivstellensatz calculus proofs for the parity. *Theoretical Computer Science* 259, 1–2 (2001), 613–622.

[17] Russell Impagliazzo, Pavel Pudlák, and Jiří Sgall. 1999. Lower Bounds for the Polynomial Calculus and the Gröbner Basis Algorithm. *Comput. Complex.* 8, 2 (November 1999), 127–144. https://doi.org/10.1007/s000370050024

[18] Dmitry Itsykson, Artur Riazanov, Danil Sagunov, and Petr Smirnov. 2021. Near-optimal lower bounds on regular resolution refutations of tseitin formulas for all constant-degree graphs. *computational complexity* 30, 2 (2021), 13.

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

[20] Pravesh K. Kothari, Ryuhei Mori, Ryan O'Donnell, and David Witmer. 2017. Sum of squares lower bounds for refuting any CSP. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing (STOC 2017)*, 2017. Association for Computing Machinery, Montreal, Canada, 132–145. https://doi.org/10.1145/3055399.3055485

[21] Michael Krivelevich and Rajko Nenadov. 2021. Complete minors in graphs without sparse cuts. *Int. Math. Res. Not.* 2021, 12 (2021), 8996–9015. https://doi.org/10.1093/imrn/rnz086

[22] Alexis Maciel, Toniann Pitassi, and Alan R. Woods. 2000. A new proof of the weak pigeonhole principle. In *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing (STOC '00)*, 2000. Association for Computing Machinery, Portland, Oregon, USA, 368–377. https://doi.org/10.1145/335305.335348

[23] Michael Mitzenmacher and Eli Upfal. 2017. *Probability and computing: Randomization and probabilistic techniques in algorithms and data analysis*. Cambridge university press.

[24] Rajko Nenadov. 2023. Routing permutations on spectral expanders via matchings. *Combinatorica* 43, 4 (2023), 737–742.

[25] Toniann Pitassi, Benjamin Rossman, Rocco A. Servedio, and Li-Yang Tan. 2016. 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)*, 2016. Association for Computing Machinery, Cambridge, MA, USA, 644–657. https://doi.org/10.1145/2897518.2897637

[26] Aaron Potechin. 2020. Sum of Squares Bounds for the Ordering Principle. In *35th Computational Complexity Conference (CCC 2020) (Leibniz International Proceedings in Informatics (LIPIcs)*, 2020. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 38:1–38:37. https://doi.org/10.4230/LIPIcs.CCC.2020.38

[27] Ran Raz. 2004. Resolution lower bounds for the weak pigeonhole principle. *J. ACM* 51, 2 (March 2004), 115–138. https://doi.org/10.1145/972639.972640

[28] Ran Raz. 2008. Elusive functions and lower bounds for arithmetic circuits. In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing (STOC '08)*, 2008. Association for Computing Machinery, Victoria, British Columbia, Canada, 711–720. https://doi.org/10.1145/1374376.1374479

[29] Alexander A. Razborov. 1998. Lower bounds for the polynomial calculus. *Comput. Complex.* 7, 4 (December 1998), 291–324. https://doi.org/10.1007/s000370050013

[30] Alexander A. Razborov. 2002. Proof Complexity of Pigeonhole Principles. In *Developments in Language Theory*, 2002. Springer Berlin Heidelberg, Berlin, Heidelberg, 100–116.

[31] Alexander A. Razborov. 2003. Resolution lower bounds for the weak functional pigeonhole principle. *Theor. Comput. Sci.* 303, 1 (2003), 233–243. https://doi.org/10.1016/S0304-3975(02)00453-X

[32] Susanna F. De Rezende, Aaron Potechin, and Kilian Risse. 2023. Clique Is Hard on Average for Unary Sherali-Adams. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, 2023. 12–25. https://doi.org/10.1109/FOCS57990.2023.00008

[33] Grant Schoenebeck. 2008. Linear Level Lasserre Lower Bounds for Certain k-CSPs. In *Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS '08)*, 2008. IEEE Computer Society, USA, 593–602. https://doi.org/10.1109/FOCS.2008.74

[34] Konstantin Tikhomirov and Pierre Youssef. 2016. The spectral gap of dense random regular graphs. Retrieved from https://arxiv.org/abs/1610.01765

[35] William T Tutte. 1947. The factorization of linear graphs. *Journal of the London Mathematical Society* 1, 2 (1947), 107–111.

[36] William Thomas Tutte. 1952. The factors of graphs. *Canadian Journal of Mathematics* 4, (1952), 314–328.