A Ihara-Bass Formula for Non-Boolean Matrices and Strong Refutations of Random CSPs

Tommaso d'Orsi
Department of Computer Science, ETH Zürich, Switzerland

Luca Trevisan
Department of Computing Sciences, Bocconi University, Milano, Italy

Abstract

We define a novel notion of "non-backtracking" matrix associated to any symmetric matrix, and we prove a "Ihara-Bass" type formula for it.

We use this theory to prove new results on polynomial-time strong refutations of random constraint satisfaction problems with k variables per constraints (k-CSPs). For a random k-CSP instance constructed out of a constraint that is satisfied by a p fraction of assignments, if the instance contains n variables and $n^{k/2}/\epsilon^2$ constraints, we can efficiently compute a certificate that the optimum satisfies at most a $p + O_k(\epsilon)$ fraction of constraints.

Previously, this was known for even k, but for odd k one needed $n^{k/2}(\log n)^{O(1)}/\epsilon^2$ random constraints to achieve the same conclusion.

Although the improvement is only polylogarithmic, it overcomes a significant barrier to these types of results. Strong refutation results based on current approaches construct a certificate that a certain matrix associated to the k-CSP instance is quasirandom. Such certificate can come from a Feige-Ofek type argument, from an application of Grothendieck's inequality, or from a spectral bound obtained with a trace argument. The first two approaches require a union bound that cannot work when the number of constraints is $o(n^{\lceil k/2 \rceil})$ and the third one cannot work when the number of constraints is $o(n^{k/2}\sqrt{\log n})$.

We further apply our techniques to obtain a new PTAS finding assignments for k-CSP instances with $n^{k/2}/\epsilon^2$ constraints in the semi-random settings where the constraints are random, but the sign patterns are adversarial.

2012 ACM Subject Classification Theory of computation; Mathematics of computing → Probability and statistics

Keywords and phrases CSP, k-XOR, strong refutation, sum-of-squares, tensor, graph, hypergraph, non-backtracking walk

Digital Object Identifier 10.4230/LIPIcs.CCC.2023.27

Related Version *Full Version*: arXiv:2204.10881

Funding This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreements No 815464 and 834861).

Acknowledgements Thanks to Pravesh Kothari for useful discussions about semi-random CSPs.

1 Introduction

If we take a random instance of 3SAT with n variables and $m \geq cn$ clauses where c is a sufficiently large constant, then almost surely the instance is not satisfiable. Indeed, an instance of random 3SAT with n variables and n/ϵ^2 clauses is almost surely such that at most a $7/8 + O(\epsilon)$ fraction of clauses can be simultanously satisfied by the best assignment. Finding a *certificate* that a specific random formula exhibits such behaviour is, however, believed to be quite hard.

© Tommaso d'Orsi and Luca Trevisan;
licensed under Creative Commons License CC-BY 4.0

38th Computational Complexity Conference (CCC 2023).
Editor: Amnon Ta-Shma; Article No. 27; pp. 27:1–27:16
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

In 2002, Feige [8] formulated the hypothesis that it is computationally intractable to find *strong refutations* of random 3-SAT formulas when the number of clauses is slightly superlinear in the number of variables. A *strong refutation* of a 3-SAT formula is a certificate, verifiable in polynomial time, that every assignment fails to satisfy a constant fraction of the clauses. Feige proved that his hypothesis has several consequences for the hardness of approximation of various problems.

Because of its centrality to the theories of proof complexity and of average-case complexity, and its connection to other questions in cryptography, computational complexity, and statistical physics, the complexity of strong refutations for random 3SAT and other random constraint satisfaction problems has been extensively studied since the 1980s.

Among several important algorithmic milestones, we mention the idea of using spectral techniques to find refutations and strong refutations (introduced in [13, 12] and then refined in subsequent work) and a reduction from the problem of finding strong refutations for random 3SAT to the problem of finding strong refutations for random 3XOR (introduced in [8] and then refined in subsequent work).

The state of the art concerning polynomial-time computable strong refutations of random constraint satisfaction problems is a 2015 paper by Allen, O'Donnell and Witmer [2]. We refer the reader to the introduction of [2] for an extended survey of algorithmic ideas and results related to refutations of random constraint satisfaction problems. Allen, O'Donnell and Witmer [2] show how to obtain strong refutations for random k-XOR constraint satisfaction problems on n variables and $n^{k/2}(\log n)^{O(1)}$ constraints. When k is even, $O(n^{k/2})$ constraints suffice. Thanks to a reduction from arbitrary constraint satisfaction to k-XOR (of which we provide a self-contained simpler proof in the full version of the paper), similar bounds hold for any constraint satisfaction problem over k variables.

To illustrate the difference between odd k and even k, we briefly discuss how a strong refutation for random 4-XOR and random 3-XOR instances is constructed.

In general, if we have an instance of k-XOR with m constraints and n variables, a strong refutation is a certificate that

$$\max_{x\in\{-1,1\}^n} \sum_{i_1,\ldots,i_k} T_{i_1,\ldots,i_k} x_{i_1} \cdots x_{i_k} \leq \epsilon m$$

where T is a symmetric tensor of order k such that $T_{i_1,\ldots,i_k} = 0$ if there is no constraint on the k-tuple of variables $x_{i_1}, \ldots, x_{i_k}$, and otherwise $T_{i_1,\ldots,i_k} = \pm 1$ depending on the right-hand-side of the constraint.

When $k = 4$, we can flatten the tensor to an $n^2 \times n^2$ symmetric matrix M (where $M_{(a,b),(c,d)} = T_{a,b,c,d}$) and we have

$$\max_{x\in\{-1,1\}^n} \sum_{i_1,\ldots,i_4} T_{i_1,\ldots,i_4} x_{i_1} \cdots x_{i_4} = \max_{x\in\{-1,1\}^n} (x^{\otimes 2})^{\mathsf{T}} M x^{\otimes 2}$$

Now we can relax the right-hand side to a maximization over arbitrary n^2-dimensional Boolean vectors and further relax to the ∞-to-1 norm:

$$\max_{x\in\{-1,1\}^n} (x^{\otimes 2})^{\mathsf{T}} M x^{\otimes 2} \leq \max_{y\in\{-1,1\}^{n^2}} y^{\mathsf{T}} M y \leq \max_{y,z\in\{-1,1\}^{n^2}} y^{\mathsf{T}} M z = ||M||_{\infty\to 1}$$

Finally, the last expression above can be upper bounded by ϵm, by using Chernoff bounds and a union bound over all the 2^{2n^2} possible choices for y and z, which is possible if m is a sufficiently large constant times n^2/ϵ^2. Finally, we can use Grothendieck's inequality to get us a certified upper bound of the $\infty \to 1$ norm in polynomial time up to a constant factor.

For 3-XOR, the idea is to apply a Cauchy-Schwarz step to reduce the problem of bounding a degree-4 problem, and then to flatten the resulting 4-tensor to an $n^2 \times n^2$ matrix M such that

$$\max_{x \in \{-1,1\}^n} \sum_{i_1,i_2,i_4} T_{i_1,i_2,i_3} x_{i_1} x_{i_2} x_{i_3} \leq \sqrt{n} \cdot \sqrt{\max_{x \in \{\pm 1\}^n} \quad (x^{\otimes 2})^{\mathsf{T}} M x^{\otimes 2}} \leq \sqrt{n} \cdot \sqrt{\max_{y,z \in \{\pm 1\}^{n^2}} \quad y^{\mathsf{T}} M z}$$

Unfortunately, now it is not possible any more to bound the maximum on the above right-hand via a union bound over 2^{2n^2} cases. Indeed, for this to be possible, we would need our distribution to have at least order of n^2 bits of entropy, and so we would need to have order of n^2 constraints.

The alternative is to obtain a bound in terms of the spectral norm of M, using the fact that

$$\max_{y,z \in \{\pm 1\}^{n^2}} \quad y^{\mathsf{T}} M z \leq n^2 \cdot ||M|| \,.$$

But for a sparse matrix to have a non-trivial bound on its spectral norm, we have to have at least $\operatorname{poly}\log n$ non-zero entries per row on average[1], and for this to happen the number of constraints has to be at least of the order of $n^{1.5} \operatorname{poly}\log n$. In the regime of $n^{1.5} \operatorname{poly}\log n$ random 3-XOR constraints, a spectral norm bound on M can be established via trace methods, and this is how the results of [2] are proved in the case of odd k.

Semi-random CSPs

The complementary question to that of certifying strong refutations, concerns the design of algorithms that satisfy as-many-as-possible clauses in the given CSP instance. As for refutations, complexity theory paints a grim picture for (approximately) solving worst case instances [18, 6, 11]. But, in the average case, polynomial time approximation schemes are known [3, 1] when the number of clauses is of the order $n^{k/2}(\log n)^{O(1)}$.

The algorithmic techniques behind these PTAS are closely related to those used for refutations and, in particular, again boils down to studying the spectrum of the flattened tensor representing the instance.

Remarkably, groundbreaking work [15], showed that a similar picture holds in the significantly more general settings of *smoothed* CSPs: where both the literal negation patterns and clauses are chosen arbitrarily, but then signs are randomly flipped with a small, yet constant, probability.[2].

1.1 Our Results

Strong refutations

Our first result breaks the $n^{k/2} \operatorname{poly}\log n$ barrier for strong refutations of random k-XOR instances, with odd k.

[1] This is similar to the phenomenon that the quasirandomness of a $G_{n,p}$ random graphs can be certified in terms of the non-trivial eigenvalues of the adjacency matrix only if the average degree is at least logarithmic. We will return to the graph analogy shortly.

[2] Smoothed CSPs were first introduced in [9]

► **Theorem 1** (Strong refutations of random k-XOR)**.** *There exists an efficient algorithm that, given an instance $\boldsymbol{I}$ of random k-XOR with $n^{k/2}/\epsilon^2$ constraints, with probability at least* 0.99*, finds strong refutation of $\boldsymbol{I}$, that is, a certificate that*

$$\mathrm{Opt}_{\boldsymbol{I}} \leq \frac{1}{2} + O(\epsilon)\,.$$

Using the known reduction of general k-CSP to k-XOR, of which we provide a simple self-contained proof, we have the following consequence.

► **Theorem 2** (Strong refutations of random CSPs)**.** *Let $P : \{-1, +1\}^k \longrightarrow \{0, 1\}$ be a Boolean k-ary predicate, and call $\mathbb{E}\,P$ the probability that P is satisfied by a random assignment. There exists a polynomial time algorithm that given a random instance $CSP(P)$ instances $\boldsymbol{I}$, over n variables, with at least $n^{k/2}/\epsilon^2$ constraints, with probability at least* 0.99*, finds a strong refutation of $\boldsymbol{I}$, that is, a certificate that*

$$\mathrm{Opt}_{\boldsymbol{I}} \leq \mathbb{E}\,P + O(\epsilon)\,.$$

Robust approximation algorithms against adversarial sign patterns

Our techniques can be further applied to design efficient algorithms finding an assignment with value Opt $- O(\epsilon)$ beyond the $n^{k/2}$ polylog n barrier. Our sharp results not only works for random instances, but also in the semi-random settings where: *first*, clauses are sampled randomly, and *second*, given the instance, the sign pattern of *each* clause is adversarially perturbed. Such perturbations are not captured by the smooth models of [9, 15] and hence require different algorithmic challenges. In the special case of even k, [17] provided a PTAS whenever $p \geq n^{k/2}$ polylog n .

► **Theorem 3** (Algorithm for k-XOR with adversarial patterns)**.** *Let n , k be positive integers, $\epsilon > 0$, n and $n^{-k/2}/\epsilon^2 < 1$. Let $\mathcal{I}$ be a k-XOR instance constructed through the following process:*

- *Sample a random k-XOR instance $\boldsymbol{I}'$ with at least $n^{k/2}/\epsilon^2$ constraints.*
- *Given $\boldsymbol{I}'$, arbitrarily (possibly adversarially) replace the sign of each clause in $\boldsymbol{I}'$.*

There exists a randomized algorithm, running in time $n^{O(k/\epsilon^2)}$, that returns an assignment $\hat{\mathbf{x}}$ with value

$$\mathrm{Val}_{\mathcal{I}}\,(\hat{\mathbf{x}}) \geq \mathrm{Opt}_{\mathcal{I}} - O(\epsilon)\,,$$

with probability at least 0.99*.*

As in the case of strong refutations, Theorem 3 can be extended to k-CSPs.

► **Theorem 4** (Algorithm for semi-random k-CSPs)**.** *Let n , k be positive integers, $\epsilon > 0$, n and $n^{-k/2}/\epsilon^2 < 1$. Let $P : \{-1, +1\}^k \longrightarrow \{0, 1\}$ be a Boolean k-ary predicate. Let $\mathcal{I}$ be a $CSP(P)$ instance constructed through the following process:*

- *Sample a random $CSP(P)$ instance $\boldsymbol{I}'$ with at least $n^{k/2}/\epsilon^2$ constraints.*
- *Given $\boldsymbol{I}'$, for each clause in $\boldsymbol{I}'$, replace the sign pattern with an arbitrary (possibly adversarial) sign pattern.*

There exists a randomized algorithm, running in time $n^{O(k/\epsilon^2)}$, that returns an assignment $\hat{\mathbf{x}}$ with value

$$\mathrm{Val}_{\mathcal{I}}\,(\hat{\mathbf{x}}) \geq \mathrm{Opt}_{\mathcal{I}} - O(\epsilon)\,,$$

with probability at least 0.99*.*

1.2 Our Techniques

We develop new techniques to bound[3]

$$\max_{x \in \{\pm 1\}^N} x^{\mathsf{T}} \mathbf{M} x \tag{1}$$

when $\mathbf{M}$ is a random $N \times N$ matrix with only a constant expected number of non-zero entries per row and per column, and in which such entries are not independent.

A toy problem

Before we explain our ideas, consider the following question, which models some of the difficulties that we encounter: suppose that we are given a random graph on N vertices, and such that every edge exists with probability d/N, where d is a constant, but the edges are only known to be $poly \log N$-wise independent, and not fully independent. Can we certify that the graph has interesting quasirandom properties, for example can we certify that the Max Cut optimum is at most a $1/2 + O(1/\sqrt{d})$ fraction of edges?

One approach could be to bound $||\mathbf{A} - \mathbb{E}\,\mathbf{A}||_{\infty \to 1}$ where $\mathbf{A}$ is the adjacency matrix of the graph. If the graph has mutually independent random edges, that is, if it is sampled from an Erdős-Reniy distribution $G_{N,\frac{d}{N}}$, then we can use a union bound over 2^{2N} cases to argue that with high probability

$$||\mathbf{A} - \mathbb{E}\,\mathbf{A}||_{\infty \to 1} \leq O(\sqrt{d}N)$$

which is certifiable in polynomial time, up to a constant factor loss, using Grothendieck's inequality and which certifies that the Max Cut optimum is at most $1/2 + O(1/\sqrt{d})$. Unfortunately, if the edges are only polylog N-wise independent, then it is not possible to take such union bound.

Another option in the fully independent case is to use the results of Feige and Ofek [10], which show that, after removing nodes of degree larger than, say, $10d$, the adjacency matrix of the residual graph has second eigenvalue at most $O(\sqrt{d})$ with high probability. Unfortunately the proof of Feige and Ofek also relies on a union bound over $2^{O(N)}$ cases, and so it cannot work in the polylog N-wise independent case.

A trace argument can be used to prove that, with high probability, we have

$$||\mathbf{A} - \mathbb{E}\,\mathbf{A}|| \leq O(\sqrt{d \log N})$$

which provides a polynomial time certificate that the Max Cut optimum is at most $1/2 + O(\sqrt{\log N}/\sqrt{d})$, and the trace calculation only requires $O(\log N)$-wise independence. It does, however, introduce an extra logarithmic factor, which is unavoidable because the spectral norm of $||\mathbf{A} - \mathbb{E}\,\mathbf{A}||$ is $\tilde{\Omega}(\sqrt{\log N})$ when d is constant.

It is conceivable that one could prove the result of Feige and Ofek (that the adjacency matrix has second largest eigenvalue $O(\sqrt{d})$ after the removal of high-degree vertices) through a trace bound on the adjacency matrix of the truncated graph, although it seems very difficult to deal with the conditional distribution of edges given that the edges survive the truncation.

[3] We use boldface to denote random variables.

A solution to the toy problem

Although all standard techniques fail, there is a way to combine certain recent results to solve our toy problem. The starting point is the fact that, given an undirected graph $G = (V, E)$, we can define the "non-backtracking" $2|E| \times 2|E|$ matrix B of G, and that this matrix satisfies the Ihara-Bass equation

$$\det(\mathrm{Id} - xB) = (1 - x^2)^{|E|-|V|} \cdot \det(\mathrm{Id} - xA + x^2(D - \mathrm{Id}))$$

where A is the adjacency matrix of the graph, D is the diagonal matrix of degrees, and the above equation holds as an identity of polynomials of degree $2|E|$ in x. See the survey of Horton [16] for an exposition of these definitions and results.

Fan and Montanari [7] show that bounds on the spectral radius of B imply useful PSD inequalities on A. In particular, if $\lambda_{\min}$ is the smallest real eigenvalue of B, then we have

$$A \succeq -|\lambda_{\min}| \cdot \mathrm{Id} - \frac{1}{|\lambda_{\min}|} \cdot (D - \mathrm{Id})$$

In the context of their work on the Stochastic Block Model, Bordenave, Lalarge and Massoulié [5] use a trace argument to prove a result that implies that $\lambda_{\min} \geq -(1 + o(1)) \cdot \sqrt{d}$ in $G_{N,\frac{d}{N}}$ random graphs, and so all these results together imply that the Max Cut of a $G_{N,\frac{d}{N}}$ random graph is with high probability at most $1/2 + (1 + o(1))/\sqrt{d}$, and that this upper bound is efficiently certifiable, for example by the dual of the Goemans-Williamson relaxation.

The key point is that there was never a union bound over $2^{O(N)}$ cases in the above argument and that, in fact, everything works assuming polylog N-wise independence of the edges.[4]

From unweighted graphs to general symmetric matrices

Our goal is to develop an analog of this argument where we work with the $n^2 \times n^2$ matrix M that comes up in the analysis of 3-XOR (or, in general, with the $n^{\lceil k/2 \rceil} \times n^{\lceil k/2 \rceil}$ matrix that comes up in the analysis of k-XOR when k is odd) instead of the adjacency matrix A of the pseudorandom graph analysed above.

The first challenge in carrying out this program is that the original notion of non-backtracking matrix is defined only with respect to 0/1 Boolean symmetric matrices, while we want to study matrices with positive and negative entries that can be arbitrary integers.

A certain generalization of non-backtracking matrices was already introduced in [20, 7], however for technical reasons *we* were not able to use it to carry out our program. We thus introduce a novel theory of "non-backtracking" matrices associated to any given symmetric matrix. In Section 3, given a symmetric $N \times N$ matrix M with Nz non-zero entries, we define an $Nz \times Nz$ "non-backtracking" matrix B_M associated to M, and we prove (see Theorem 7) an Ihara-Bass-type identity

$$\det(\mathrm{Id} - xB_M + x(L_M - J_M)) = (1 - x^2)^{Nz/2-N} \cdot \det(\mathrm{Id} - xM + x^2(D_M - \mathrm{Id}))$$

[4] Incidentally, this combination of Fan-Montanari ideas and Bordenave-Lalarge-Massouli's bounds, also implies that if A' is the adjacency matrix of a graph G sampled from a distribution in which edges have probability d/N and are polylog N wise independent, and then truncated by removing all vertices of degree more than, say, $10d$, then we have with high probability $A' \succeq -O(\sqrt{d}) \cdot I$, proving a one-sided version of the result of Feige and Ofek.

where D_M, L_M and J_M are certain matrices that are associated to M. When M is Boolean, $L_M = J_M$ and D_M is the diagonal matrix such that $(D_M)_{i,i} = \sum_j M_{i,j}$, so our equation becomes the standard Ihara-Bass equation in the case of Boolean M. Conveniently, closed non-backtracking walks W arising from the definition of B_M take value in $\{\pm \prod_{(i,j)\in W} M_{ij}\}$, allowing one to easily mimic arguments used for standard non-backtracking matrices.

Now, given a bound on the spectral radius of $B_M - L_M + J_M$, it is possible, with an argument in the style of Fan and Montanari, to deduce a certifiable bound on the ∞-to-1 norm of M.

Bounding the spectral radius via weighted hyper-walks

Studying the spectral radius of $B_{\mathbf{M}} - L_{\mathbf{M}} + J_{\mathbf{M}}$ –matrices associated to the matrix $\mathbf{M}$ coming from random k-XOR instances– is the main technical challenge of this work.

Our bound relies on a trace argument of $B_{\mathbf{M}}$. However, compared to Bordenave, Lalarge and Massoulié [5] our setup presents a number of new technical challenges.

One challenge comes from the extra terms that we have in the non-Boolean case. In particular, our non-backtracking matrix $B_{\mathbf{M}}$ has entries that are the absolute values of certain entries of $\mathbf{M}$. To compute an expectation of the trace of the symmetrization of a power of $B_{\mathbf{M}}$, we replace absolute values with squares, and bound the error that we incur because of this.

Perhaps the most important challenge comes from the fact that the trace bound ultimately boils down to a weighted count of certain closed "hypergraph walks" performed on the hypergraph corresponding to constraints of the k-XOR instance. These objects arise from our notion of non-backtracking walks on the symmetric matrix $\mathbf{M}$ obtained from the instance. This count is performed by showing that such walks can be encoded with a small number of bits. It is enough to count walks in which every hyperedge is repeated at least twice, and the crux of the argument is that the second time we see a hyperedge we can encode that hyperedge in a compact way. A naive way of doing that would point back to the previous step in the walk in which that hyperedge appeared, and this costs $\log \ell$ bits where ℓ is the length of the walk. To obtain the right result, however, repeated hyperedges have to be represented with an amortized constant number of bits per occurrence. The argument of Bordenave, Lalarge and Massoulié [5] relies on the assumption, which is true with high probability, that the graph is "tangle-free," meaning that small subgraphs have at most one cycle. We have to work with a looser notion of tangle-free hypergraph in order to prove that it holds with high probability, but we are still able to obtain the desired bound.

From spectral bounds to algorithms

It is clear that an algorithm certifying tight bounds on Equation (1) for the matrix M obtained from k-XOR instances can be used for strong refutations. Instead, to obtain Theorem 3 additional ideas are needed.

Our starting point is the local-to-global rounding paradigm of [3]. As it is often the case, the odd settings are significantly more challenging than the regimes with k even. Hence consider first a 2-XOR random instance $\boldsymbol{\mathcal{I}}$. Up to the signs of the clauses, this may be represented as a graph $\mathbf{G}$ over n vertices. Now, for a distribution ν over assignments, one may define the local and global correlations as

$$\begin{aligned}\mathrm{LC}_{\mathbf{G}}(\nu) &= \mathop{\mathbb{E}}_{(\mathbf{a},\mathbf{b})\sim E(\mathbf{G})} \left|\mathrm{Cov}_\nu\left(\mathbf{x}^{\mathbf{a}}, \mathbf{x}^{\mathbf{b}}\right)\right| \\ \mathrm{GC}(\nu) &= \mathop{\mathbb{E}}_{(\mathbf{a},\mathbf{b})\sim[n]\times[n]} \left|\mathrm{Cov}_\nu\left(\mathbf{x}^{\mathbf{a}}, \mathbf{x}^{\mathbf{b}}\right)\right| .\end{aligned}$$

If the local correlation is bounded by ϵ, it is possible to obtain an assignment with value $\mathrm{Opt}_{\mathbf{I}} - O(\epsilon)$ simply looking at the *first moment* of ν. Moreover, one can always transform ν into a distribution with small global correlation in polynomial time.

With these observations, the argument of [3] comes down to: *(i)* bounding the difference between local and global correlation in terms of the spectral radius $\rho_{\mathbf{G}}$ of the centered adjacency matrix of the graph $\mathbf{G}$, *(ii)* showing that one can always find, in time $n^{O(1/\epsilon^2)}$, a degree $O(1)$ pseudo-distribution over the hypercube with global correlation at most ϵ. As we only required low-degree moments to obtain the desired assignment, the argument goes through in this case as well.

To combine this approach with the bounds previously illustrated and extend the argument to random k-XOR instances with $m \geq \Omega(n^{k/2}/\epsilon^2)$ clauses, we need to introduce two novel ingredients. *First*, we need new notions of local and global correlations which difference can be bounded studying the matrix $\mathbf{M}$ arising from the instance. *Second*, we need to bound this difference not in term of the eigenvalues of $\mathbf{M}$ but rather in terms of Equation (1).

A careful Cauchy-Schwarz application allows us to formulate notions of local and global correlations in terms of $\mathbf{M}$. Its squaring step, further allows us to get rid of absolute values, thus providing an avenue to bound the difference between local and global correlation in terms of $\max_{x\in\{\pm1\}^n} x^{\mathsf{T}}\mathbf{M}x$.

Finally, since the adversarial perturbations in Theorem 3 cannot alter the "hypergraph walks" required to prove our bound, we are able to generalize our result to these settings.

1.3 Perspective

Several results on the average-case complexity of Sum-of-Square relaxations rely on proving that sparse matrices with non-independent entries are "quasirandom" in an appropriate sense. We have introduced a new approach to prove results of this form, which applies to very sparse matrices that have only a constant expected number of non-zero entries per row and per column. We hope that our ideas will find further application, for example to the context of semi-random instances of constraint satisfaction problems [14] or of higher-degree Sum-of-Square relaxations of random constraint satisfaction problems [19, 21].

Our theory could also be useful to study problems on random weighted graphs.

Our certificates prove certain PSD inequalities, and can be seen as Semidefinite Duals of certain Sum-of-Squares relaxations, but the computation of the certificate only requires an eigenvalue computation of a certain matrix, and does not require the solution of an SDP. There might be other ways to apply our theory so that one uses SDP relaxations only in the analysis, but the algorithm itself is purely spectral.

1.4 Organization

In the rest of the paper we first introduce preliminary notions, including those of CSPs and strong refutation, then present a proof of our generalized Ihara-Bass formula. We show the proofs of our main Theorems in the full version of the paper.

2 Preliminaries

We introduce some notation, useful facts and needed preliminary notions. We denote random variables in **bold**. We use lower case letters $a, b, c, d, \ldots$ to denote indices or scalars (the use willl be clear from context). We use the greek letters α, β, η to denote multi-indices. The cardinality of a multi-index α is $|\alpha|$. The i-th index in α is $\alpha(i)$. We may thus write a monomial (with coefficient c) in indeterminates $x_1, \ldots, x_n$ as $c \cdot x^\alpha$. For two multi indices $\alpha, \beta \in [n]^k$ we denote by (α, β) the multi-index obtained concatenating α and β. Multi-indices $\alpha, \beta \in [n]^k$ satisfy $\alpha = \beta$ if at each position the corresponding indices are identical. We use $S(\alpha)$ to denote the unordered multi-set of indices in α. We use n to denote our ambient dimension. For functions $f, g : \mathbb{R} \to \mathbb{R}$ we write $f = o(g)$ and $g = \omega(f)$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$.

Matrices

For a matrix $M \in \mathbb{R}^{n\times n}$, we denote by $\lambda_1(M) \geq \ldots \geq \lambda_n(M)$ its eigenvalues. Then $\rho(M) := \max_i |\lambda_i(M)|$ is the spectral radius of M. W When the context is clear we simply write $\lambda_1, \ldots, \lambda_n$. The spectral radius of a matrix satisfies the following inequality.

▶ **Lemma 5** (Gelfand's Formula). *Let $M \in \mathbb{R}^{n\times n}$ and let $\|\cdot\|_*$ be a norm. Then for any positive integer z*

$$\rho(M) \leq \|M^z\|_*^{1/z} .$$

We write $\|M\|$ for the spectral norm of a matrix M, $\|M\|_{\mathrm{F}}$ for its Frobenius norm $\|M\|_{\infty\to 1} := \max_{x,y\in\{\pm 1\}^n} \langle M, xy^\mathsf{T}\rangle$ and $\|M\|_{\max} := \max_{ij} |M_{ij}|$. Furthermore, we let

$$\|M\|_{\mathrm{Gr}} = \max \{\langle M, X\rangle \mid X \succeq 0, X_{ii} \leq 1 , \forall i \in [n]\} .$$

We denote by $|M|$ the matrix with entries $(|M|)_{ij} := |M_{ij}|$. We write Id_t for the t-by-t identity matrix, $\mathbf{0}$ for the zero matrix and J for the all-ones matrix.

Graphs

For a graph G, $V(G)$ and $E(G)$ denotes respectively its set of vertices and edges. $\vec{E}(G) := \{(u, v) : u \neq v \in V(G), uv \in E(G)\}$ is the set of all its ordered pairs such that $\{u, v\} \in E(G)$. For $e \in \vec{E}(G)$, $s(e)$ and $t(e)$ are respectively the source and target of the oriented edge. We write e^{-1} for its inverse. We also write K_n for the complete graph over n vertices. For a graph G with n vertices, we write $A(G) \in \mathbb{R}^{n\times n}$ for its adjacency matrix. For a vertex $v \in V(G)$, we denote by $\deg_G(v)$ its degree. We denote by $N_{G,t}(v)$ the set of vertices in G at distance t from v. We and drop the subscript G when the context is clear. If the graph G is weighted with weights given by $w : V(G) \times V(G) \to \mathbb{R}$, then $A_{uv} = w(\{uv\})$. If $e \neq E(G)$, then we assume $w(e) = 0$. A walk W in a graph G is a sequence of vertices $(v_1, \ldots, v_{z+1})$. A walk $v_1, \ldots, v_{z+1}$ is said to be non-backtracking if for any $i \leq z - 1$, $v_i \neq v_{i+2}$.

2.1 CSPs, k-XOR and strong refutations

k-XOR

A random k-XOR instance $\boldsymbol{I}$ with n variables and $p\binom{n}{k}(1 \pm o(1))$ clauses can be generated by picking a random symmetric tensor $\mathbf{T}$, with independent entries, such that $\mathbf{T}_\alpha = 0$ if the indices in the multi-index $\alpha \in [n]^k$ are not distinct and otherwise:

$$\mathbf{T}_\alpha = \begin{cases} 0 & \text{with probability } 1-p\,, \\ +1 & \text{with probability } p/2\,, \\ -1 & \text{with probability } p/2\,. \end{cases}$$

We denote by m the exact number of clauses in the instance. Then $\mathcal{I}$ consists of the m k-XOR predicates $k\text{-XOR}(\alpha) = \frac{1-x^\alpha(-\mathbf{T})_\alpha}{2}$ where $\mathbf{T}_\alpha$ is non-zero. We use $\mathcal{F}_{k\text{-XOR}(n,p)}$ to denote such distribution and $\mathcal{I} \sim \mathcal{F}_{k\text{-XOR}(n,p)}$ to denote a random instance. We let $\mathrm{Val}_{\mathcal{I}}(x)$ be the fraction of constrained satisfied by the assignment $x \in \{\pm 1\}^n$ and $\mathrm{Opt}_{\mathcal{I}} := \max_{x\in\{\pm1\}^n} \mathrm{Val}_{\mathcal{I}}(x)$. For any assignment $x \in \{\pm 1\}^n$ we have

$$\mathrm{Val}_{\mathcal{I}}(x) = \frac{1}{2} + \frac{1}{m(\mathcal{I})} \sum_{\alpha\in[n]^k} \frac{x^\alpha \mathbf{T}_\alpha}{2}\,.$$

Notice that since m will be $(1 \pm o(1))p\binom{n}{k}$ with overwhelming probability, we blur the distinction between these parameters. Then the max k-XOR problem is that of finding an assignment with value

$$\max_{x\in\{\pm1\}^n} \sum_{\alpha\in[n]^k} \mathbf{T}_\alpha x^\alpha\,. \tag{2}$$

This is captured by the following proposition.

▶ **Proposition 6.** *Let $\mathcal{I} \sim \mathcal{F}_{k\text{-}XOR(n,p)}$ and let $\mathbf{T}$ be the associated k-th order tensor. Then with overwhelming probability*

$$\mathrm{Opt}_{\mathcal{I}} \leq \frac{1}{2} + (1+o(1))\left(\binom{n}{k}\cdot p\right)^{-1} \cdot \sum_{\alpha\in[n]^k} \mathbf{T}_\alpha x^\alpha\,.$$

Throughout the paper we assume k to be an *odd* integer as for the even case sharp refutation algorithms are already known (e.g see [2]).

A random k-XOR instance $\mathcal{I}$ with n variables and exactly m clauses can be generated by picking m times a clause at random out of the $\binom{n}{k}$ possible k-XOR-clause. It is possible to show that a refutation algorithm for $\mathcal{I} \sim \mathcal{F}_{k\text{-XOR}(n,p)}$ can also be used for refutation of k-XOR instances sampled through this second process. For this reason, we blur the distinction between these two processes. We direct the reader interested in a formal reduction to [2] (Appendix D).

CSPs

Given a predicate $P : \{-1,1\}^k \to \{0,1\}$, an instance $\mathcal{I}$ of the CSP(P) problem over variables $x_1,\ldots,x_n$ is a multi-set of pairs (c,α) representing constraints of the form $P(c\circ x^\alpha) := P(c_1 x^{\alpha(1)},\ldots,c_k x^{\alpha(k)}) = 1$ where $\alpha\in[n]^k$ is the scope and $c\in\{\pm1\}^k$ is the negation pattern. We can represent the predicate P as a multi-linear polynomial of degree k in indeterminates $c_1x^{\alpha(1)},\ldots,c_kx^{\alpha(k)}$,

$$P(c\circ x^\alpha) = \sum_{d\leq k} P_d(c\circ x^\alpha)\,,$$

where P_d denotes the degree d part of the predicate. In particular $P_0 := P_0(c\circ x^\alpha)$ denotes the constant part of the polynomial, which does not depend on the assignment.

The fraction of all possible assignments that satisfy P is given by $\mathbb{E}_{\mathbf{z} \overset{u.a.r}{\sim} \{\pm 1\}^k}[P(\mathbf{z})]$. For any assignment $x \in \{\pm 1\}^n$ and an instance $\mathcal{I}$ over m constraints we have

$$\mathrm{Val}_{\mathcal{I}}(x) = \frac{1}{m} \sum_{(c,\alpha) \in \mathcal{I}} P(c \circ x^{\alpha})$$

$$\text{and} \quad \mathrm{Opt}_{\mathcal{I}} = \max_{x \in \{\pm 1\}^n} \mathrm{Val}_{\mathcal{I}}(x)\,.$$

A *random* CSP(P) instance $\boldsymbol{\mathcal{I}}$ with $(1+o(1))m = p \cdot 2^k \cdot n^k$ constraints can be generated as follows:

(i) Pick independently with probability p each pair $(\mathbf{c}, \boldsymbol{\alpha})$ where $\mathbf{c}$ is a random negation pattern from $\{-1,+1\}^k$ and α is a multi-index from $[n]^k$,

(ii) For each such pair $(\mathbf{c}, \boldsymbol{\alpha})$ add the constraint $P(\mathbf{c} \circ x^{\boldsymbol{\alpha}}) = 1$ to $\boldsymbol{\mathcal{I}}$.

Notice that we do not rule out predicates with same multi-index but different negation pattern as multi-indices in which an index appears multple time. We also do not assume P to be symmetric. We denote such distribution by $\mathcal{F}_{\mathrm{CSP(P)}}(n,p)$.

As in the case of k-XOR a random CSP(P) instance $\boldsymbol{\mathcal{I}}$ with n variables and exactly m clauses can be generated by picking m times a clause and a negation pattern at random. Again it is possible to show that a refutation algorithm for $\boldsymbol{\mathcal{I}} \sim \mathcal{F}_{\mathrm{CSP(P)}}(n,p)$ can also be used for refutation of instances sampled through this second process (see Appendix D in [2]).

Refutation and certification

We say that $\mathcal{A}$ is a δ-*refutation algorithm* for random CSP(P) if $\mathcal{A}$ has the following properties:

(i) on all instances $\mathcal{I}$ the output of $\mathcal{A}$ si either $\mathrm{Opt}_{\mathcal{I}} \leq 1-\delta$ or "fail",

(ii) if $\mathrm{Opt}_{\mathcal{I}} > 1-\delta$ then $\mathcal{A}$ *never* outputs $\mathrm{Opt}_{\mathcal{I}} \leq 1-\delta$.

More generally, for an set of possible inputs $\mathcal{S}$ and a property p over instances in $\mathcal{S}$, we say that an algorithm $\mathcal{A}$ *certifies* p if:

(i) on all inputs $\mathcal{I} \in \mathcal{S}$ the output of $\mathcal{A}$ is either "$\mathcal{I}$ satisfies p" or "fail",

(ii) if $\mathcal{I} \in \mathcal{S}$ does not satisfy p then $\mathcal{A}$ *never* outputs "$\mathcal{I}$ satisfies p".

In the context of random CSP(P) (and hence k-XOR), a ***strong refutation*** is a δ-refutation for $1-\delta \leq \mathbb{E}_{\mathbf{x} \overset{u.a.r}{\sim} \{\pm 1\}^k}[P(\mathbf{x})] + o(1)$.

3 A generalized Ihara-Bass formula

In this section we present an extension of the Ihara-Bass theorem (see [16] and references therein) to arbitrary real symmetric matrices. We remark that our extension differs from the one in [7].

Throughout the section we assume to be given a *symmetric* matrix $A \in \mathbb{R}^{n \times n}$ with $2m$ non-zero entries and zeroed diagonal. We use the following notation. We will use letters u, v to denote indices in $[n]$ and e, f for indices in $[2m]$. We conveniently think of A as the adjacency matrix of a weighted undirected graph G with n vertices and $2m$ oriented edges. Then $uv \in E(G)$ if $A_{uv} \neq 0$, moreover then the inverse edge vu is also in $E(G)$ since $A_{uv} = A_{vu}$ by definition. Recall for an edge $e \in E(G)$ we write e^{-1} for its inverse and for a vertex $v \in V(G)$ we write $N^+(v)$ (respectively $N^-(v)$) for its set of outgoing (resp. incoming) oriented edges in G. We write $\sigma_{uv} = \mathrm{sign}(A_{uv})$. To reason about the spectrum of A, we introduce several matrices: the diagonal matrices

$$D(A) \in \mathbb{R}^{n\times n}, \quad \text{with } D_{uv}(A) = \begin{cases} \sum_w |A_{uw}| & u = v \\ 0 & \text{otherwise.} \end{cases}$$

$$Q(A) \in \mathbb{R}^{m\times m}, \quad \text{with } Q_{ef}(A) = \begin{cases} |A_e| & e = f \\ 0 & \text{otherwise.} \end{cases}$$

the block matrices

$$J(A) = \begin{pmatrix} 0 & \mathrm{Id}_m \\ \mathrm{Id}_m & 0 \end{pmatrix} \in \mathbb{R}^{2m\times 2m}$$

$$L(A) = \begin{pmatrix} 0 & Q(A) \\ Q(A) & 0 \end{pmatrix} \in \mathbb{R}^{2m\times 2m}$$

and the source, target and non-backtracking matrices

$$S(A) \in \mathbb{R}^{n\times 2m}, \quad \text{with } S_{ue}(A) = \begin{cases} \sigma_{uv}\sqrt{|A_{uv}|} & \text{if } u \text{ is the source of } e = uv \text{ and } u < v \\ \sqrt{|A_{uv}|} & \text{if } u \text{ is the source of } e = uv \text{ and } u > v \\ 0 & \text{otherwise.} \end{cases}$$

$$T(A) \in \mathbb{R}^{n\times 2m}, \quad \text{with } T_{ue}(A) = \begin{cases} \sigma_{uv}\sqrt{|A_{uv}|} & \text{if } u \text{ is the target of } e = vu \text{ and } u < v \\ \sqrt{|A_{uv}|} & \text{if } u \text{ is the target of } e = vu \text{ and } u > v \\ 0 & \text{otherwise.} \end{cases}$$

$$B(A) \in \mathbb{R}^{2m\times 2m}, \quad \text{with } B_{ef}(A) = \begin{cases} \sigma_e\sigma_f\sqrt{|A_eA_f|} & \text{if } ef \text{ is a non-backtracking walk} \\ & e = uv,\ \text{f=}vw \text{ and } v < u, w \\ \sigma_e\sqrt{|A_eA_f|} & \text{if } ef \text{ is a non-backtracking walk} \\ & e = uv,\ \text{f=}vw \text{ and } w < v < u \\ \sigma_f\sqrt{|A_eA_f|} & \text{if } ef \text{ is a non-backtracking walk} \\ & e = uv,\ \text{f=}vw \text{ and } u < v < w \\ \sqrt{|A_eA_f|} & \\ & \text{if } ef \text{ is a non-backtracking walk} \\ & e = uv,\ \text{f=}vw \text{ and } v > u, w \\ 0 & \text{otherwise.} \end{cases}$$

When the context is clear we simply write B for $B(A)$ (analogously for the other matrices). To gain intuition on these linear maps, it is instructive to consider the case when A is the adjacency matrix of an unweighted graph G. Then D is the degree diagonal matrix with $D_{uu} = deg_G(u)$, $L = J$ and B corresponds to the non-backtracking matrix of G.

Throughout the other sections of the paper, for a given non-backtracking matrix $B \in \mathbb{R}^{2m\times 2m}$, we will consider the related extension matrix $B^* \in \mathbb{R}^{2n^2\times 2n^2}$ with entries

$$B^*_{ef} = \begin{cases} B_{ef} & \text{if } e, f \in E(G) \\ 0 & \text{otherwise.} \end{cases}$$

For simplicity of the notation, we will often denote B^* simply by B. The context will always be clarified by the ambient dimension. We can now state the main result of the section.

▶ **Theorem 7** (Generalized Ihara-Bass Theorem). *Let n, m be integers and let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix with m non-zero entries, all off-diagonal. Let B, L, J, D defined as above. Then, for any $u \in \mathbb{R}$,*

$$\det\left(\mathrm{Id}_{2m} - u(B+L-J)\right) = (1-u^2)^{m-n}\left(\mathrm{Id}_n - uA + u^2 D - u^2\mathrm{Id}_n\right) .$$

Our proof of Theorem 7 closely resembles the proof of Bass [4]. We first observe that the matrices above satisfy several useful identities, than tackle the theorem.

▶ **Lemma 8.** *Using the definitions above:*

i) $SJ = T$ *and* $TJ = S$ *,*
ii) $A = ST^{\mathsf{T}}$ *,*
iii) $D = SS^{\mathsf{T}} = TT^{\mathsf{T}}$ *,*
iv) $B + L = T^{\mathsf{T}}S$ *.*

Proof. For i), notice that $SJ \in \mathbb{R}^{n\times 2m}$ and $SJ_{ue} = \langle S_{u,-}, J_{-,e}\rangle = S_{ue^{-1}} = T_{ue}$, where in the third step we used symmetry of A. A similar argument can be made to show $TJ = S$. For ii) observe that

$$A_{uv} = \langle S_{u,-}, T_{v,-}\rangle = \sum_e S_{ue}T_{ve}$$

which is nonzero only when $e = uv$. In that case, by definition $A_{uv} = \sigma_{uv}\,|A_{uv}| = S_{ue}T_{ve}$ since either $u < v$ or $u > v$. Consider now SS^{T}, the matrix is diagonal since each edge has at most one source vertex, then

$$(SS^{\mathsf{T}})_{uu} = \sum_e S^2_{ue} = \sum_{v\in N^+(u)} |A_{uv}| = D_{uu} .$$

A symmetric derivation shows $D_{uu} = (TT^{\mathsf{T}})_{uu}$. It remains to prove iv). It is trivial to check that

$$(T^{\mathsf{T}}S)_{ee} = \langle T_{-,e}, S_{-,e}\rangle = \sum_u T_{ue}S_{ue} = 0 ,$$

since there are no self-loops in the graph. For distinct $e, f \in [2m]$

$$(T^{\mathsf{T}}S)_{ef} = \sum_u T_{ue}S_{uf} .$$

There is at most one non-zero element in the sum, corresponding to the case when u is the target vertex of e and the source of f, which means ef is a walk of length 2 in G. If ef is a non-backtracking walk (that is, $e \neq f^{-1}$) then $B_{ef} = (T^{\mathsf{T}}S)_{ef}$ and $L_{ef} = 0$. Conversely, if $e = f^{-1}$ then $B_{ef} = 0$ and $L_{ef} = (T^{\mathsf{T}}S)_{ef}$. Finally, signs can be checked case by case. ◀

We are now ready to prove Theorem 7.

Proof of Theorem 7. In the following identities all matrices are $(n+2m)\times(n+2m)$ block matrices where the first block has size $n \times n$. Let $u \in \mathbb{R}$,

$$\begin{aligned}
&\begin{pmatrix}\mathrm{Id}_n & 0\\ T^{\mathsf{T}} & \mathrm{Id}_{2m}\end{pmatrix}\begin{pmatrix}\mathrm{Id}_n(1-u^2) & Su\\ 0 & \mathrm{Id}_{2m} - (B+L-J)u\end{pmatrix} \qquad (3)\\
&= \begin{pmatrix}\mathrm{Id}(1-u^2) & Su\\ T^{\mathsf{T}}(1-u^2) & T^{\mathsf{T}}Su + \mathrm{Id}_{2m} - (B+L-J)u\end{pmatrix}\\
&= \begin{pmatrix}\mathrm{Id}(1-u^2) & Su\\ T^{\mathsf{T}}(1-u^2) & \mathrm{Id}_{2m} + Ju\end{pmatrix} .
\end{aligned}$$

On the other hand

$$\begin{aligned}
&\begin{pmatrix}\mathrm{Id}_n(1-u^2)-Au+Du^2 & Su\\ 0 & \mathrm{Id}_{2m}+Ju\end{pmatrix}\begin{pmatrix}\mathrm{Id}_n & 0\\ T^\intercal - S^\intercal u & \mathrm{Id}_{2m}\end{pmatrix}\\
&=\begin{pmatrix}\mathrm{Id}_n(1-u^2)-Au+Du^2+ST^\intercal u - SS^\intercal u^2 & Su\\ T^\intercal - S^\intercal u + JT^\intercal u - JS^\intercal u^2 & \mathrm{Id}_{2m}+Ju\end{pmatrix}\\
&=\begin{pmatrix}\mathrm{Id}_n(1-u^2) & Su\\ T^\intercal(1-u^2) & \mathrm{Id}_{2m}+Ju\end{pmatrix}.
\end{aligned} \tag{4}$$

Putting Equation (3) and Equation (4) together and taking determinants we get

$$(1-u^2)^n \det\left(\mathrm{Id}_{2m}-(B+L-J)u\right)=\det\left(\mathrm{Id}_n(1-u^2)-Au+Du^2\right)\det\left(\mathrm{Id}_{2m}+Ju\right).$$

Now notice that

$$\mathrm{Id}_{2m}+Ju=\begin{pmatrix}\mathrm{Id}_m & \mathrm{Id}_m u\\ \mathrm{Id}_m u & \mathrm{Id}_m\end{pmatrix}$$

and thus $\det\left(\mathrm{Id}_{2m}+Ju\right)=(1-u^2)^m$. Rearranging, the result follows. ◀

3.1 Norm bounds via the Ihara-Bass formula

In this section we show how Theorem 7 can be used to study the spectrum of a real symmetric matrix A via the spectrum of related matrices. The central tool is the theorem below.

▶ **Theorem 9.** *Let $A \in \mathbb{R}^{n\times n}$ a symmetric matrix with zero diagonal. Let B, L, J, D be as defined in Section 3. Let $\lambda_{\min}$ be the smallest eigenvalue of the matrix $B+L-J \in \mathbb{R}^{2m\times 2m}$. Then for any $\lambda \le \lambda_{\min}$*

$$A \succeq -|\lambda|\,\mathrm{Id}_n - |\lambda|^{-1}\left(D-\mathrm{Id}_n\right).$$

Proof. Let $\lambda_{\min}$ be the smallest real eigenvalue of $B+L-J$. By Theorem 7 we know -1 is a real eigenvalue of $B+L-J$ and thus $\lambda_{\min} \le -1$. Moreover, for every $\lambda < \lambda_{\min}$ we have $\det\left(\mathrm{Id}_{2m}-\lambda^{-1}B+\lambda^{-1}L-\lambda^{-1}J\right)\neq 0$ otherwise λ would be an eigenvalue smaller than $\lambda_{\min}$. Define the matrix

$$M_\lambda := \mathrm{Id}_n - \lambda^{-1}A + \lambda^{-2}(D-\mathrm{Id}_n).$$

By the same reasoning as in Theorem 7, $\det(M_\lambda)\neq 0$ as long as $\lambda<\lambda_{\min}$. We make the stronger claim

$$\forall\lambda,\ \lambda_{\min}\ :\ M_\lambda \succ \mathbf{0}.$$

To prove the above claim, suppose toward a contradiction that $\lambda' < \lambda_{\min}$ is such that $M_{\lambda'}$ has a negative eigenvalue. Since M_λ tends to Id_n when $\lambda \to -\infty$, there is a value $\lambda_{PD} < \lambda'$ such that $M_{\lambda_{PD}}$ is strictly positive definite. Consider now the smallest eigenvalue of M_λ for values of λ in the range (λ_{PD}, λ'). The smallest eigenvalue of M_λ varies continuously with λ, it is positive for $\lambda=\lambda_{PD}$ and it is negative for $\lambda=\lambda'$, so it must be equal to zero for some $\lambda^* \le \lambda' < \lambda_{\min}$. But this means that $\det(M_{\lambda^*})=0$ and so λ^* is an eigenvalue of $B+L-J$, which contradicts the definition of $\lambda_{\min}$. We have thus established our claim. Rearranging the result follows. ◀

A crucial consequence of Theorem 9 is that, exploiting the diagonal structure of the matrices $D\,,\mathrm{Id}_n$ one can bound the norm $\|A\|_{\infty\to 1}$ as a function of the smallest eigenvalue of the associated non-backtracking matrix.

▶ **Corollary 10.** *Let $A \in \mathbb{R}^{n\times n}$ a symmetric matrix with zero diagonal. Let $\lambda_{\min}$ and $\lambda'_{\min}$ be respectively the smallest eigenvalue of the matrix $B(A) + L(A) - J(A) \in \mathbb{R}^{2m\times 2m}$ and $B(-A) + L(A) - J(A) \in \mathbb{R}^{2m\times 2m}$, for $B\,,L\,,J\,,D$ as defined in Section 3. Then, for any $\lambda \geq \max\left\{|\lambda_{\min}|\,,|\lambda'_{\min}|\right\}$,*

$$\|A\|_{\infty\to 1} \leq 2\,\mathrm{Tr}\left|\left(\lambda\mathrm{Id}_n + \lambda^{-1}(D(A) - \mathrm{Id}_n)\right)\right|$$

Proof. Define

$$R := \left|\lambda\mathrm{Id}_n + \lambda^{-1}(D(A) - \mathrm{Id}_n)\right|\,.$$

By Theorem 9 for any $x \in \{\pm 1\}^n$ we have $\left|x^\top A x\right| \leq \left|x^\top R x\right|$. For any $y \in \{\pm 1\}^n$ we can write

$$\begin{aligned} 2\left|x^\top A y\right| &\leq \left|(x+y)^\top A(x+y) - x^\top A x - y^\top A y\right| \\ &\leq \left|(x+y)^\top A(x+y)\right| + \left|x^\top A x\right| + \left|y^\top A y\right|\,. \end{aligned}$$

Now $x + y \in \{-2, 0, +2\}^n$ and thus

$$\left|(x+y)^\top A(x+y)\right| \leq 4 \max_{z\in\{\pm 1\}^n} z^\top R z\,,$$

the result follows by definition of R. ◀

References

1 Vedat Levi Alev, Fernando Granha Jeronimo, and Madhur Tulsiani. Approximating constraint satisfaction problems on high-dimensional expanders. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 180–201. IEEE, 2019.

2 Sarah R. Allen, Ryan O'Donnell, and David Witmer. How to refute a random CSP. In *IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS 2015, Berkeley, CA, USA, 17-20 October, 2015*, pages 689–708, 2015. doi:10.1109/FOCS.2015.48.

3 Boaz Barak, Prasad Raghavendra, and David Steurer. Rounding semidefinite programming hierarchies via global correlation. In *2011 ieee 52nd annual symposium on foundations of computer science*, pages 472–481. IEEE, 2011.

4 Hyman Bass. The Ihara-Selberg zeta function of a tree lattice. *International Journal of Mathematics*, 3(06):717–797, 1992.

5 Charles Bordenave, Marc Lelarge, and Laurent Massoulié. Non-backtracking spectrum of random graphs: Community detection and non-regular ramanujan graphs. In *IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS 2015, Berkeley, CA, USA, 17-20 October, 2015*, pages 1347–1357, 2015. doi:10.1109/FOCS.2015.86.

6 Siu On Chan. Approximation resistance from pairwise-independent subgroups. *Journal of the ACM (JACM)*, 63(3):1–32, 2016.

7 Zhou Fan and Andrea Montanari. How well do local algorithms solve semidefinite programs? In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 604–614. ACM, 2017.

8 Uriel Feige. Relations between average case complexity and approximation complexity. In *STOC 2002*, pages 534–543, 2002.

9 Uriel Feige. Refuting smoothed 3cnf formulas. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)*, pages 407–417. IEEE, 2007.

10 Uriel Feige and Eran Ofek. Spectral techniques applied to sparse random graphs. *Random Struct. Algorithms*, 27(2):251–275, 2005.

11 Dimitris Fotakis, Michael Lampis, and Vangelis Th Paschos. Sub-exponential approximation schemes for csps: From dense to almost sparse. *arXiv preprint*, 2015. `arXiv:1507.04391`.

12 Joel Friedman and Andreas Goerdt. Recognizing more unsatisfiable random 3-SAT instances efficiently. In *Automata, Languages and Programming, 28th International Colloquium, ICALP 2001, Crete, Greece, July 8-12, 2001, Proceedings*, volume 2076 of *Lecture Notes in Computer Science*, pages 310–321. Springer, 2001.

13 Andreas Goerdt and Michael Krivelevich. Efficient recognition of random unsatisfiable k-SAT instances by spectral methods. In *STACS 2001, 18th Annual Symposium on Theoretical Aspects of Computer Science, Dresden, Germany, February 15-17, 2001, Proceedings*, volume 2010 of *Lecture Notes in Computer Science*, pages 294–304. Springer, 2001.

14 Venkatesan Guruswami, Pravesh K. Kothari, and Peter Manohar. Algorithms and certificates for boolean CSP refutation: "smoothed is no harder than random". *arXiv*, 2109.04415, 2021. `arXiv:2109.04415`.

15 Venkatesan Guruswami, Pravesh K Kothari, and Peter Manohar. Algorithms and certificates for boolean csp refutation: smoothed is no harder than random. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 678–689, 2022.

16 Matthew D Horton, HM Stark, and Audrey A Terras. What are zeta functions of graphs and what are they good for? *Contemporary Mathematics*, 415:173–190, 2006.

17 Pravesh K. Kothari. Personal communication, 2022.

18 Dana Moshkovitz and Ran Raz. Two-query pcp with subconstant error. *Journal of the ACM (JACM)*, 57(5):1–29, 2008.

19 Prasad Raghavendra, Satish Rao, and Tselil Schramm. Strongly refuting random CSPs below the spectral threshold. In Hamed Hatami, Pierre McKenzie, and Valerie King, editors, *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 121–131. ACM, 2017.

20 Yusuke Watanabe and Kenji Fukumizu. Graph zeta function in the bethe free energy and loopy belief propagation. *Advances in Neural Information Processing Systems*, 22, 2009.

21 Alexander S. Wein, Ahmed El Alaoui, and Cristopher Moore. The Kikuchi hierarchy and tensor PCA. In *60th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2019, Baltimore, Maryland, USA, November 9-12, 2019*, pages 1446–1468, 2019.