The SDP value of random 2CSPs

Amulya Musipatla

CMU-CS-21-108
May 2021

Computer Science Department
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee:
Ryan O'Donnell, Chair
Pravesh Kothari

*Submitted in partial fulfillment of the requirements
for the degree of Master of Science.*

Keywords: constraint satisfaction problems, semidefinite programming, SDP bound, eigenvalue bound, positive semidefinite, matrix weighted polynomials, matrix weighted graphs, lifts, infinite lifts, duality

Abstract

We consider a very wide class of models for sparse random Boolean 2CSPs; equivalently, degree-2 optimization problems over $\{\pm 1\}^n$. Specifically, we look at models $\mathcal{M}$ which can be represented by matrix weighted polynomials over indeterminates taking the values of either matching matrices or permutation matrices. We interpret these polynomials also as models of graphs with matrix weighted edges. For each model $\mathcal{M}$, we identify the "high probability value" $s^*_{\mathcal{M}}$ of the natural SDP relaxation (equivalently, the quantum value). That is, for all $\epsilon > 0$ we show that the SDP optimum of a random n-variable instance is (when normalized by n) in the range $(s^*_{\mathcal{M}} - \epsilon, s^*_{\mathcal{M}} + \epsilon)$ with high probability.

This thesis contains joint work with Ryan O'Donnell, Tselil Schramm, and Xinyu Wu.

Acknowledgements

I'd like to thank Ryan O'Donnell for being an amazing advisor this past year. Thanks for all the meetings, for always going through ideas with me and providing me with useful feedback on my work, and especially for being invested in my future in research. Even though we weren't able to meet in person this past year, I've learned a lot through the many Zoom calls and appreciate all the time you spent working with me.

I'd also like to thank Xinyu Wu for getting me onboarded with all of the concepts from her previous work, and answering every single question I had while working on this project. Thank you for always checking in with me to see how things were going, for helping me out with my presentation, and for always being willing to meet to talk about things- you were seriously a great help during this year.

A combined thanks to Ryan, Xinyu, and Tselil Schramm for working with me on this project. I've learned so much from each of you and really enjoyed the process of doing research together. Also, special thanks to Pravesh Kothari for being on my thesis committee.

Finally, I'd like to thank my friends and family for all the support I've received during this really weird year. Thanks for listening to my ramblings and always hyping me up. :)

Contents

1 Introduction

A large number of important algorithmic tasks can be construed as constraint satisfaction problems (CSPs): finding an assignment to Boolean variables to optimize the number of satisfied constraints. Almost every form of constraint optimization is NP-complete; thus one is led to questions of efficiently finding near-optimal solutions, or understanding the complexity of average-case rather than worst-case instances. Indeed, understanding the complexity of random sparse CSPs is of major importance not just in traditional algorithms theory, but also in, e.g., cryptography [JLS20], statistical physics [MM09], and learning theory [DLS14].

Suppose we fix the model $\mathcal{M}$ for a random sparse CSP on n variables (e.g., random k-SAT with a certain clause density). Then it is widely believed that there should be a constant $c^*_{\mathcal{M}}$ such that the optimal value of a random instance is $c^*_{\mathcal{M}} \pm o_{n\to\infty}(1)$ with high probability (whp). (Here we define the optimal value to mean the maximum number of simultaneously satisfiable constraints, divided by the number of variables.) Unfortunately, it is extremely difficult to prove this sort of result; indeed, it was considered a major breakthrough when Bayati, Gamarnik, and Tetali [BGT13] established it for one of the simplest possible cases: Max-Cut on random d-regular graphs (which we will denote by $\mathcal{MC}_d$). Actually "identifying" the value $c^*_{\mathcal{M}}$ (beyond just proving its existence) is even more challenging. Generally there is a conjectured value coming from statistical physics, but proving it is beyond the reach of current methods. Taking again the example of Max-Cut on d-regular random graphs, it was only recently [DMS⁺17] that the value $c^*_{\mathcal{MC}_d}$ was determined up to a factor of $1 \pm o_{d\to\infty}(1)$. The value for any particular d, e.g. $c^*_{\mathcal{MC}_3}$, has yet to be established.

Returning to algorithmic questions, we can ask about the *computational feasibility* of optimally solving sparse random CSPs. There are two complementary questions to ask: given a random instance from model $\mathcal{M}$ (with presumed optimal value $c^*_{\mathcal{M}} \pm o_{n\to\infty}(1)$), can one efficiently *find* a solution achieving value $\gtrsim c^*_{\mathcal{M}}$, and can one efficiently *certify* that every solution achieves value $\lesssim c^*_{\mathcal{M}}$? The former question is seemingly a bit more tractable; for example, a very recent breakthrough of Montanari [Mon21] gives an efficient algorithm for (whp) finding a cut in a random graph $\mathcal{G}(n,p)$ graph of value at least $(1-\epsilon)c^*_{\mathcal{G}(n,p)}$.On the other hand, we do not know any algorithm for efficiently certifying (whp) that a random instance has value at most $(1+\epsilon)c^*_{\mathcal{G}(n,p)}$. Indeed, it reasonable to conjecture that no such algorithm exists, leading to an example of a so-called "information-computation gap".

To bring evidence for this we can consider *semidefinite programming* (SDP), which provides efficient algorithms for certifying an upper bound on the optimal value of a CSP [FL92]. Indeed, it is known [Rag09] that, under the Unique Games Conjecture, the basic SDP relaxation provides essentially optimal certificates for CSPs in the *worst case*. In this paper we in particular consider Boolean 2CSPs — more generally, optimizing a homogeneous degree-2 polynomial over the hypercube — as this is the setting where semidefinite programming is most natural. Again, for a fixed model $\mathcal{M}$ of random sparse Boolean 2CSPs, one expects there should exist a constant $s^*_{\mathcal{M}}$ such that the optimal SDP-value of an instance from $\mathcal{M}$ is whp $s^*_{\mathcal{M}} \pm o_{n\to\infty}(1)$. Philosophically, since semidefinite programming is in P, one may be more optimistic proving this and explicitly identifying $s^*_{\mathcal{M}}$. Indeed, some results in this direction have recently been established.

1.1 Prior work on identifying high-probability SDP values

Let us consider the most basic case: $\mathcal{MC}_d$, Max-Cut on random d-regular graphs. For ease of notation, we will consider the equivalent problem of maximizing $\frac{1}{n}x^\mathsf{T}(-\boldsymbol{A})x$ over $x \in \{\pm 1\}^n$, where $\boldsymbol{A}$ is the adjacency matrix of a random n-vertex d-regular graph.[1] Although $s^*_{\mathcal{MC}_d}$, the high-

[1]Throughout this work, **boldface** is used to denote random variables.

probability SDP relaxation value, was pursued as early as 1987 [Bop87] (see also [DH73]), it was not until 2015 that Montanari and Sen [MS16] established the precise result $s^*_{\mathcal{MC}_d} = 2\sqrt{d-1}$. That is, in a random d-regular graph, whp the basic SDP relaxation value [Bop87, RW95, DP93, PR95] for the size of the maximum cut is $(\frac{d}{4} + \sqrt{d-1} \pm o_{n\to\infty}(1))n$. Here the special number $2\sqrt{d-1}$ is the maximum eigenvalue of the d-regular infinite tree.

The proof of this result has two components: showing $\textsc{Sdp}(-\boldsymbol{A}) \geq 2\sqrt{d-1} - \epsilon$ whp, and showing $\textsc{Sdp-Dual}(-\boldsymbol{A}) \leq 2\sqrt{d-1} + \epsilon$ whp. Here $\textsc{Sdp}(A) = \max\{\langle \rho, A\rangle : \rho \geq 0,\ \rho_{ii} = \frac{1}{n}\ \forall i\}$ denotes the "primal" SDP value on matrix A (commonly associated with the Goemans–Williamson rounding algorithm [GW95]), and $\textsc{Sdp-Dual}(A) = \min\{\lambda_{\max}(A + \mathrm{diag}(u)) : \mathrm{avg}_i(u_i) = 0\}$ denotes the (equal) "dual" SDP value on A. To show the latter bound, it is sufficient to observe that $\textsc{Sdp-Dual}(-\boldsymbol{A}) \leq \lambda_{\max}(-\boldsymbol{A})$, the "eigenvalue bound", and $\lambda_{\max}(-\boldsymbol{A}) \leq 2\sqrt{d-1} + o_n(1)$ whp by Friedman's Theorem [Fri08]. As for lower-bounding $\textsc{Sdp}(-\boldsymbol{A})$, Montanari and Sen used the "Gaussian wave" method [Elo09, CGHV15, HV15] to construct primal SDP solutions achieving at least $2\sqrt{d-1} - \epsilon$ (whp). The idea here is essentially to build the SDP solutions using an approximate eigenvector (of finite support) of the infinite d-regular tree achieving eigenvalue $2\sqrt{d-1} - \epsilon$; the fact that SDP constraint "$\rho_{ii} = \frac{1}{n}\ \forall i$" can be satisfied relies heavily on the regularity of the graph.

Remark 1.1. The Montanari–Sen result in passing establishes that the (high-probability) eigenvalue and SDP bounds coincide for random regular graphs. This is consistent with a known theme, that the two bounds tend to be the same (or nearly so) for graphs where "every vertex looks similar" (in particular, for regular graphs). This theme dates back to Delorme and Poljak [DP93], who showed that $\textsc{Sdp-Dual}(-A) = \lambda_{\max}(-A)$ whenever A is the adjacency matrix of a vertex-transitive graph.

Subsequently, the high-probability SDP value $s^*_{\mathcal{M}}$ was established for a few other models of random regular 2CSPs. Deshpande, Montanari, O'Donnell, Schramm, and Sen [DMO+19] showed that for $\mathcal{M} = \mathcal{NAE}3_c$ — meaning random regular instances of NAE-3SAT (not-all-equals 3Sat) with each variable participating in c clauses — we have $s^*_{\mathcal{M}} = \frac{9}{8} - \frac{3}{8} \cdot \frac{\sqrt{c-1}-\sqrt{2}}{c}$. We remark that NAE-3SAT is effectively a 2CSP, as the predicate $\mathrm{NAE}_3 : \{\pm 1\}^3 \to \{0,1\}$ may be expressed as $\frac{3}{4} - \frac{1}{4}(xy+yz+zx)$, supported on the "triangle" formed by vertices x, y, z. The analysis in this paper is somewhat similar to that in [MS16], but with the infinite graph $\mathfrak{X} = K_3 \star K_3 \star \cdots \star K_3$ (c times) replacing the d-regular infinite tree. This $\mathfrak{X}$ is the $2c$-regular infinite "tree of triangles" depicted (partly, in the case $c = 3$) in Figure 1. More generally, [DMO+19] established the high-probability SDP value for large random (edge-signed) graphs that locally resemble $K_r \star K_r \star \cdots \star K_r$, the $(r-1)c$-regular infinite "tree of cliques K_r". (The $r = 2$ case essentially generalizes [MS16].) As in [MS16], $s^*_{\mathcal{M}}$ coincides with the (high-probability) eigenvalue bound. The upper bound on $s^*_{\mathcal{M}}$ is shown by using Bordenave's proof [Bor20] of Friedman's Theorem for random (c, r)-biregular graphs. The lower bound on $s^*_{\mathcal{M}}$ is shown using the Gaussian wave technique, relying on the distance-regularity of the graphs $K_r \star K_r \star \cdots \star K_r$ (indeed, it is known that every infinite distance-regular graph is of this form).

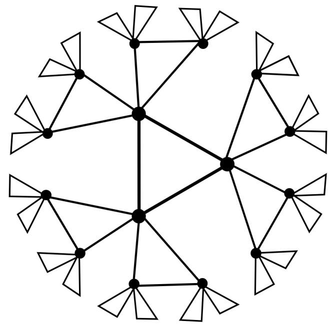

Figure 1: The 6-regular infinite graph $K_3 \star K_3 \star K_3$, modeling random 3-regular NAE3-SAT.

Mohanty, O'Donnell, and Paredes [MOP20] generalized the preceding two results to the case of "two-eigenvalue" 2CSPs. Roughly speaking, these are 2CSPs formed by placing copies of a small weighted "constraint graph" $\mathcal{H}$ — required to have just two distinct eigenvalues — in a random regular fashion onto n vertices/variables. (This is indeed a generalization [DMO+19], as cliques have just two distinct eigenvalues.) As two-eigenvalue examples, [MOP20] considered CSPs with the "CHSH constraint" — and its generalizations, the "Forrelation$_k$" constraints — which are important in quantum information theory [CHSH69, AA18]. Here the SDP value of an instance is particularly relevant as it is precisely the optimal "quantum entangled value" of the 2CSP [CHTW04]. Once again, it is shown in [MOP20] that the high-probability SDP and eigenvalue bounds coincide for these types of CSPs. The two-eigenvalue condition is used at a technical level in both the variant of Bordenave's theorem proven for the eigenvalue upper bound, and in the Gaussian wave construction in the SDP lower bound.

Most recently, O'Donnell and Wu [OW20] used the results of Bordenave and Collins [BC19] (a substantial generalization of [Bor20]) to establish the high-probability *eigenvalue relaxation bound* for very wide range of random 2CSPs, encompassing all those previously mentioned: namely, any quadratic optimization problem defined by random "matrix polynomial lifts" over literals.

1.2 Our work

In this work, we establish the high-probability SDP value $s^*_{\mathcal{M}}$ for random instances of any 2CSP model $\mathcal{M}$ arising from lifting matrix polynomials (as in [OW20]). This generalizes all previously described work on SDP values, and covers many more cases, including random lifts of any base 2CSP and random graphs modeled on any free/additive/amalgamated product. See Section 2 for more details and definitions, and see [OW20] for a thorough description of the kinds of random graph/2CSP models that can arise from matrix polynomials.

Very briefly, a matrix polynomial p is a small, explicit "recipe" for producing random n-vertex edge-weighted graphs, each of which "locally resembles" an associated *infinite* graph $\mathfrak{X}_p$. For example, $p_3(Y_1, Y_2, Y_3) \coloneqq Y_1 + Y_2 + Y_3$ is a recipe for random (edge-signed) 3-regular n-vertex graphs, and here $\mathfrak{X}_{p_3}$ is the infinite 3-regular tree. As another example, if $p_{333}(Z_{1,1}, \dots, Z_{3,3})$ denotes the following matrix polynomial —

$$\begin{pmatrix} 0 & Z_{1,1}Z^*_{1,2} + Z_{2,1}Z^*_{2,2} + Z_{3,1}Z^*_{3,2} & Z_{1,1}Z^*_{1,3} + Z_{2,1}Z^*_{2,3} + Z_{3,1}Z^*_{3,3} \\ Z_{1,2}Z^*_{1,1} + Z_{2,2}Z^*_{2,1} + Z_{3,2}Z^*_{3,1} & 0 & Z_{1,2}Z^*_{1,3} + Z_{2,2}Z^*_{2,3} + Z_{3,2}Z^*_{3,3} \\ Z_{1,3}Z^*_{1,1} + Z_{2,3}Z^*_{2,1} + Z_{3,3}Z^*_{3,1} & Z_{1,3}Z^*_{1,2} + Z_{2,3}Z^*_{2,2} + Z_{3,3}Z^*_{3,2} & 0 \end{pmatrix}$$

— then p_{333} is a recipe for random (edge-signed) 6-regular n-vertex graphs where every vertex participates in 3 triangles. In this case, $\mathfrak{X}_{p_{333}}$ is the infinite graph (partly) depicted in Figure 1. The

Bordenave–Collins theorem [BC19] shows that if $\boldsymbol{A}$ is the adjacency matrix of a random *unsigned* n-vertex graph produced from a matrix polynomial p, then whp the "nontrivial" spectrum of $\boldsymbol{A}$ will be within ϵ (in Hausdorff distance) of the spectrum of $\mathfrak{X}_p$. In the course of derandomizing this theorem, O'Donnell and Wu [OW20] established that for random *edge-signed* graphs, the modifier "nontrivial" should be dropped. As a consequence, in the signed case one gets $\lambda_{\max}(\boldsymbol{A}) \approx \lambda_{\max}(\mathfrak{X}_p)$ up to an additive ϵ, whp; i.e., the high-probability eigenvalue bound for CSPs derived from p is precisely $\lambda_{\max}(\mathfrak{X}_p)$. We remark that for simple enough p there are formulas for $\lambda_{\max}(\mathfrak{X}_p)$; regarding our two example above, it is $2\sqrt{2}$ for $p = p_3$, and it is 5 for $p = p_{333}$. In particular, if p is a linear matrix polynomial, $\lambda_{\max}(\mathfrak{X}_p)$ may be determined numerically with the assistance of a formula of Lehner [Leh99] (see also [GVK21] for the case of standard random lifts of a fixed base graph).

In this paper we investigate the high-probability *SDP* value — denote it s_p^* — of a large random 2CSP (Boolean quadratic optimization problem) produced by a matrix polynomial p. Critically, our level of generality lets us consider *non-regular* random graph models, in contrast to all previous work. Because of this, see we cases in contrast to Remark 1.1, where (whp) the SDP value is strictly smaller than the eigenvalue relaxation bound. As a simple example, for random edge-signed $(2,3)$-biregular graphs, the high-probability eigenvalue bound is $\sqrt{2-1} + \sqrt{3-1} = 1 + \sqrt{2} \approx 2.414$, but our work establishes that the high-probability SDP value is $\sqrt{\frac{13}{4} + 2\sqrt{2}} - \frac{1}{10} \approx 2.365$.

An essential part of our work is establishing the appropriate notion of the "SDP value" of an infinite graph $\mathfrak{X}_p$, with adjacency operator A_∞. While the eigenvalue bound $\lambda_{\max}(A_\infty)$ makes sense for the infinite-dimensional operator A_∞, the SDP relaxation does not. The definition $\textsc{Sdp}(A_\infty) = \max\{\langle \rho, A_\infty \rangle : \rho \geq 0,\ \rho_{ii} = \frac{1}{n}\ \forall i\}$ does not make sense, since "n" is ∞. Alternatively, if one tries the normalization $\rho_{ii} = 1$, any such ρ will have infinite trace, and hence $\langle \rho, A_\infty \rangle$ may be ∞. Indeed, since the only control we have on A_∞'s "size" will be an operator norm ("∞-norm") bound, the expression $\langle \rho, A_\infty \rangle$ is only guaranteed to make sense if ρ is restricted to be trace-class (i.e., have finite "1-norm").

On the other hand, we know that the eigenvalue bound $\lambda_{\max}(A_\infty)$ is too weak, intuitively because it does not properly "rebalance" graphs $\mathfrak{X}_p$ that are not regular/vertex-transitive. The key to obtaining the correct bound is introducing a new notion, intermediate between the eigenvalue and SDP bounds, that is appropriate for graphs $\mathfrak{X}_p$ arising from matrix polynomial recipes. Although these graphs may be irregular, their definition also allows them to be viewed as *vertex-transitive* infinite graphs with $r \times r$ *matrix* edge-weights. In light of their vertex-transitivity, Remark 1.1 suggests that a "maximum eigenvalue"-type quantity — suitably defined for matrix-edge-weighted graphs — might serve as the sharp replacement for SDP value. We introduce such a quantity, calling it (for lack of a better term) the *partitioned SDP* bound. Let G be an n-vertex graph with $r \times r$ Hermitian matrices as edge weights, and let A be its adjacency matrix, thought of as an $n \times n$ matrix whose entries are $r \times r$ matrices. We will define

$$\textsc{PartSdp}(A) = \sup\{\langle \rho, A_\infty \rangle : \rho \geq 0,\ \mathrm{tr}(\rho)_{ii} = \tfrac{1}{r}\}, \tag{1}$$

where here $\mathrm{tr}(\rho)$ refers to the $r \times r$ matrix obtained by summing the entries on A's main diagonal (themselves $r \times r$ matrices), and $\mathrm{tr}(\rho)_{ii}$ denotes the scalar in the (i,i)-position of $\mathrm{tr}(\rho)$. This partitioned SDP bound may indeed be regarded as intermediate between the maximum eigenvalue and the SDP value. On one hand, given an scalar-edge-weighted n-vertex graph with adjacency matrix A, we may take $r = 1$ and then it is easily seen that $\textsc{PartSdp}(A)$ coincides with $\lambda_{\max}(A)$. On the other hand, if we regard A as a 1×1 matrix and take $r = n$ (so that we a have single vertex with a self-loop weighted by all of A), then $\textsc{PartSdp}(A) = \textsc{Sdp}(A)$.

As we will see in Section 3 when we make these definitions precise, $\textsc{PartSdp}(A)$ remains well-defined even for bounded-degree infinite graphs with $r \times r$ edge-weights. Furthermore, it has the following SDP dual:

$$\textsc{PartSdp-Dual}(A) = \inf\{\lambda_{\max}(A + \mathbb{1}_{n\times n} \otimes \operatorname{diag}(u)) : \operatorname{avg}(u_1, \dots, u_r) = 0\}.$$

In the technical Section 3.1, we show that there is no SDP duality gap between $\textsc{PartSdp}(A)$ and $\textsc{PartSdp-Dual}(A)$, even in the case of infinite graphs. It is precisely the common value of $\textsc{PartSdp}(\mathfrak{X}_p)$ and $\textsc{PartSdp-Dual}(\mathfrak{X}_p)$ that is the high-probability SDP value of large random 2CSPs produced from p; our main theorem is the following:

Theorem 1.2. *Let p be a matrix polynomial with $r \times r$ coefficients. Let A_∞ be the adjacency operator (with $r \times r$ entries) of the associated infinite lift $\mathfrak{X}_p$, and write $s_p^* = \textsc{PartSdp}(A_\infty) = \textsc{PartSdp-Dual}(A_\infty)$. Then for any $\epsilon > 0$, if $\boldsymbol{A}_n$ is the adjacency matrix of a random edge-signed n-lift of p, it holds that $s_p^* - \epsilon \le \textsc{Sdp}(\boldsymbol{A}_n) \le s_p^* + \epsilon$ except with probability at most $o_{n\to\infty}(1)$.*

Note that $\textsc{PartSdp}(A_\infty)$ is a fixed value only dependent on the matrix polynomial p, a finitary object.

The upper bound $\textsc{Sdp}(\boldsymbol{A}_n) \le \textsc{PartSdp-Dual}(A_\infty) + \epsilon$ in this theorem follows straightforwardly from the results of [BC19, OW20]. Our main work is to prove the lower bound $\textsc{Sdp}(\boldsymbol{A}_n) \ge \textsc{PartSdp}(A_\infty) - \epsilon$. For this, we first show that there is a ρ_0 with only finitely many nonzero entries that achieves the sup in Equation (1) up to ϵ. This is effectively a finite $r \times r$ matrix-edge-weighted graph. We then show that this ρ_0 can whp be "pasted" almost everywhere into the graph $\boldsymbol{G}_n$ defined by $\boldsymbol{A}_n$ — somewhat similar to the Gaussian wave constructions — and this gives an SDP solution of nearly the same value. The fact that $\mathfrak{X}_p$ and $\boldsymbol{G}_n$ are regarded as regular tree-like graphs with matrix edge-weights (rather than as irregular graphs with scalar edges-weights) is crucially used to show that the "pasted solution" satisfies the finite SDP's constraints "$\rho_{ii} = \frac{1}{n}\ \forall i$".

2 Preliminaries

To preface the following definitions and concepts, we note that our "real" interest is in scalar-weighted graphs, and the following terminology with matrix weights helps us define interesting scalar-weighted graphs via matrix polynomial lifts (our recipes for producing random CSP instances), and facilitates the definition of $\textsc{PartSdp}()$, which we use to bound the $\textsc{Sdp}$ value of the CSPs of interest.

2.1 Matrix-weighted graphs

In the following definitions, we'll be focusing on graphs with at-most-countable vertex sets and bounded degrees. We also often use bra-ket notation, where $(|v\rangle)_{v\in V}$ denotes the standard orthonormal basis for the complex vector space $\ell_2(V)$.

Definition 2.1 (Matrix-weighted graph)**.** Let $G = (V, E)$ be a directed multigraph, in which we explicitly allow self-loops. We say that G is *matrix-weighted* if, for some $r \in \mathbb{N}^+$, each directed edge $e \in E$ is given an associated nonzero "weight" $a_e \in \mathbb{C}^{r\times r}$. We say that G is an *undirected* matrix-weighted graph if (with a minor exception) its directed edges are partitioned into pairs e and e^*, where e^* is the reverse of e and where $a_{e^*} = a_e^*$. We call each such pair an undirected edge. The minor exception is that we allow any subset of the self-loops in E to be unpaired, provided each unpaired self-loop e has a self-adjoint weight, $a_e^* = a_e$. The adjacency operator A for G, acting on $\ell_2(V) \otimes \mathbb{C}^r$, is given by

$$\sum_{e=(u,v)\in E} |v\rangle\langle u| \otimes a_e.$$

It can be helpful to think to think of A in matrix form, as a $|V| \times |V|$ matrix whose entries are themselves $r \times r$ edge-weight matrices. Note that if G is undirected then A will be self-adjoint, $A = A^*$.

Definition 2.2 (Extension of a matrix-weighted graph)**.** In addition to viewing a matrix-weighted graph A as an $n \times n$ matrix of $r \times r$ edge weights, one may also think of it as a $nr \times nr$ matrix with scalar edge weights. We call this the *extension* of A, and denote this $nr \times nr$ matrix as $\widetilde{A}$.

2.2 Matrix polynomials

Definition 2.3 (Matrix polynomial)**.** Let $Y_1, \dots, Y_d$ be formal indeterminates that are their own inverses, and let $Z_1, \dots, Z_e$ be formal indeterminates with formal inverses $Z_1^*, \dots, Z_e^*$. For a fixed r, we define a *matrix polynomial* to be a formal noncommutative polynomial over the indeterminates $Y_1, \dots, Z_e^*$, with coefficients in $\mathbb{C}^{r\times r}$. In particular, it is a finite sum of terms, each of which looks like $aX_1 \cdots X_k$, with $a \in \mathbb{C}^{r\times r}$ and each X_k being one of the indeterminates $Y_1, \dots, Z_e^*$.

We also allow an 'empty' term, which we write as $a_0\mathbb{1}$. We will be substituting unitary matrices in for our indeterminates, so one can think of Z_i^* as standing for both the adjoint and the inverse of Z_i.

Definition 2.4 (Adjoint of a polynomial)**.** Given a matrix polynomial p, we define its *adjoint* to be p^*, where p^* is the sum of the adjoints of each term of p. Here, the adjoint of a term $aX_1 \cdots X_k$ is $a^*X_k^* \cdots X_1^*$, where a^* is the standard adjoint operation on matrices in $\mathbb{C}^{r\times r}$, and $Y_i^* := Y_i$ and $Z_i^{**} := Z_i$.

Definition 2.5 (Reduced matrix polynomial)**.** We call a matrix polynomial p *reduced* if we are unable to cancel out any pairs of Y_is and any Z_i/Z_i^* pairs in any term. Any specific term with this property is also called reduced term or a reduced word.

In this work, we will only be considering self-adjoint, reduced polynomials. In any such polynomial, some terms will be self-adjoint and the others will come in pairs T and T^*.

2.3 Lifts of matrix polynomials

Definition 2.6 (n-lift)**.** Given some self-adjoint matrix polynomial p over the indeterminates $Y_1, \dots, Z_e^*$, we define a n-lift as a substitution of matching matrices (adjacency matrices of perfect matchings) M_i on $[n]$ for indeterminates Y_i, and permutation matrices P_i on $[n]$ for the indeterminates Z_i, and P_i^* for the indeterminates Z_i^*. If p contains any instances of Y_is, we require that n is even. We substitute the $n \times n$ identity matrix for $\mathbb{1}$. Note that the matching matrices M_i and the permutation matrices P_i that we substitute are unitary and self-adjoint as desired.

For some fixed matching and permutation matrices, we'll sometimes specify a lift of p as $\mathcal{L} = (M_1, \dots, M_d, P_1, \dots, P_e)$.

Definition 2.7 (Lift of a word)**.** Given a polynomial p over indeterminates $Y_1, \dots Z_e^*$, some n-lift $\mathcal{L} = (M_1, \dots, M_d, P_1, \dots, P_e)$ of p, and a (reduced) word $w = X_1 \dots X_k$ over the indeterminates $Y_1, \dots Z_e^*$, we denote the lift of the word w as $\mathcal{L}^w = X_1 \dots X_k$ under the substitution $\mathcal{L}$.

Definition 2.8 (± 1-signed n-lift)**.** A signed n-lift of a polynomial p is an n-lift that substitutes signed matching and permutation matrices in for the indeterminates, where a signed matrix allows entries in $\{\pm 1\}$ as opposed to just 1.

Remark 2.9. Regarding edge-signs: If we were to substitute *non-edge-signed* matching matrix, this would give a familiar model for generating random 3-regular graphs: superimposing three random matchings. The random edge-signs are to model "random negated literals", and are technically convenient to ensure that the optimization problem arising from any random graph model is nontrivial. (E.g., in a random d-regular bipartite graph, Max-Cut is trivial, but with random edge-signs we get random bipartite Max-2XOR, which is not trivial.)

Definition 2.10 (Evaluating a matrix polynomial)**.** Given a matrix polynomial p over indeterminates $Y_1, \ldots, Z_e^*$, where

$$p = a_0 \mathbb{1} + \sum_i a_i X_{i_1} \ldots X_{i_k}$$

we "evaluate" p on lift $\mathcal{L}$ in the following way:

$$p(\mathcal{L}) = a_0 \otimes \mathbb{1} + \sum_i \mathcal{L}^{X_{i_1} \ldots X_{i_k}} \otimes a_i$$

There are two equivalent interpretations here. One interpretation is that substituting $n \times n$ (edge-signed) matching/permutation matrices yields us the adjacency matrix of n-vertex graph with edge-weights that are $r \times r$ matrices. This is recovered by viewing this adjacency matrix in block form, as an $n \times n$ matrix with $r \times r$ entries. Alternatively, the interpretation of a term like $a_i Z_i$ (where a_i is an explicit constant $r \times r$ matrix, and Z_i is an indeterminate) is that we should form the tensor/Kronecker product of a_i with the signed $n \times n$ permutation matrix substituted for Z_i. This yields the adjacency matrix of an nr-vertex graph with scalar edge-weights.

The following definition of the "graph of an n-lift" describes the graph we get via the first interpretation: an n-vertex graph with $r \times r$ edge weights.

Definition 2.11 (Graph of an n-lift)**.** Let $\mathcal{L} = (M_1, \ldots, M_d, P_1, \ldots, P_e)$ be a (signed) n-lift of a matrix polynomial p. $\mathcal{L}$ naturally defines a graph G_n on $[n]$ in the following way: for every vertex $i \in [n]$ and for every term $T = aX_1 \ldots X_k$, let $j \in [n]$ be the vertex such that $|j\rangle = \mathcal{L}^{X_1 \ldots X_k} |i\rangle$. Then, there is a directed edge of weight $a \in \mathbb{C}^{r \times r}$ from i to j in G.

Because p is self adjoint, for any directed edge (i, j) in G_n associated to weight a, there is also a directed edge (j, i) associated with weight a^*.

More formally, every edge is labeled by a self-adjoint (Hermitian) edge weight, and these edges come in directed pairs.

This graph exactly corresponds to the evaluation of $p(\mathcal{L})$. That is, if A_n is the adjacency matrix of this G_n, then we have that $A_n = p(\mathcal{L})$.

In this way, a "matrix polynomial" p may be identified with a random graph model, or as a "recipe" for creating large random (symmetric, edge-weighted) graphs (equivalently, homogeneous degree-2 polynomials).

Example 2.12. A simple example is the matrix polynomial

$$p(Y_1, Y_2, Y_3) = Y_1 + Y_2 + Y_3.$$

Given this p, one obtains (the adjacency matrix of) an n-vertex random graph by substituting independent, uniformly randomly chosen $n \times n$ (edge-signed) matching matrices for each Y_i. If the terms have scalar coefficients, the resulting graph will have scalar edge-weights.

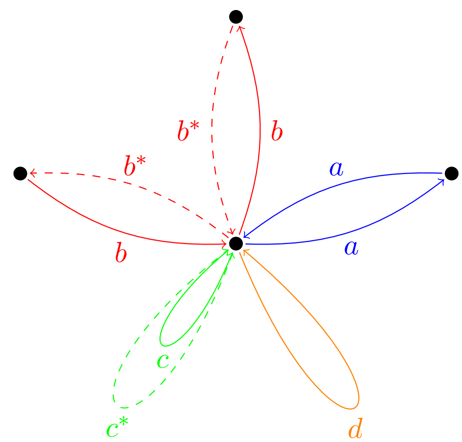

Figure 2: A possible edge distribution from a vertex in a graph of some n-lift. Edge weights a and d correspond to self-adjoint terms and are self adjoint themselves, whereas b and c are not self adjoint and appear with their adjoint pair.

By allowing matrix coefficients, one can get the random (signed) graph model given by randomly n-lifting any base r-vertex graph H.

Example 2.13. As a simple example,

$$p(Z_1, Z_2, Z_3) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} Z_1 + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} Z_1^* + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} Z_2 + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} Z_2^* + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} Z_3 + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} Z_3^*$$

is the recipe for random 3-regular (edge-signed) $(n+n)$-vertex bipartite graphs. The reader may like to view this as a 2×2 matrix of polynomials,

$$p(Z_1, Z_2, Z_3) = \begin{pmatrix} 0 & Z_1 + Z_2 + Z_3 \\ Z_1^* + Z_2^* + Z_3^* & 0 \end{pmatrix}.$$

However we remark that we actually Kronecker-product the coefficients "on the other side". So rather than as a 2×2 block-matrix with $n \times n$ blocks, we think of the resulting adjacency matrix as an $n \times n$ block-matrix with 2×2 blocks; equivalently, an n-vertex graph with 2×2 matrix edge-weights.

Finally, as suggested by their names, a matrix polynomial p may also have degree higher than 1 in its indeterminates Y_i, Z_i, and Z_i^* (which, naturally, should be treated as *noncommuting* indeterminates). When combined with matrix coefficients, this allows one to create a very wide range of random graphs.

Example 2.14. As just one example, the matrix polynomial mentioned previously in Section 1,

$$p = \begin{pmatrix} 0 & Z_{1,1}Z_{1,2}^* + Z_{2,1}Z_{2,2}^* + Z_{3,1}Z_{3,2}^* & Z_{1,1}Z_{1,3}^* + Z_{2,1}Z_{2,3}^* + Z_{3,1}Z_{3,3}^* \\ Z_{1,2}Z_{1,1}^* + Z_{2,2}Z_{2,1}^* + Z_{3,2}Z_{3,1}^* & 0 & Z_{1,2}Z_{1,3}^* + Z_{2,2}Z_{2,3}^* + Z_{3,2}Z_{3,3}^* \\ Z_{1,3}Z_{1,1}^* + Z_{2,3}Z_{2,1}^* + Z_{3,3}Z_{3,1}^* & Z_{1,3}Z_{1,2}^* + Z_{2,3}Z_{2,2}^* + Z_{3,3}Z_{3,2}^* & 0 \end{pmatrix}$$

can be shown to produce random graphs that locally resemble (indeed, are "covered by") the graph depicted in Figure 1; i.e., 6-regular n-vertex graphs in which each vertex participates in 3 triangles.

Definition 2.15 (∞-lift)**.** Formally, we extend the definition of an n-lift to the case of $n = \infty$, as follows. Let V_∞ denote the group $\mathbb{Z}_2^{\star d} \star \mathbb{Z}^{\star e}$, with its components generated by $g_1, \dots, g_{d+e}$. Each of these generators acts as a permutation on V_∞ by left-multiplication; we write $\sigma_1, \dots, \sigma_{d+e}$ for these permutations. Writing also σ_0 for the identity permutation on V_∞, and $\sigma_{j^*} = \sigma_j^{-1}$ for $d < j \le d+e$, we define $\mathfrak{L}_\infty = (\sigma_0, \dots, \sigma_{d+2e})$ to be "the" ∞-lift associated to p.

Less formally, we can think of the ∞-lift as the following graph $\mathfrak{X}_p$:

- The vertices of G are every possible reduced word over $Y_1, \dots, Z_e^*$.
- From any vertex W, the term $aX_1 \dots X_k$ corresponds with an out-edge of weight a from W to a vertex whose name is the reduced form of $X_1 \dots X_k W$. This exactly corresponds to each vertex having one out-edge per term of p.

Example 2.16. For the polynomial $p = Y_1 + \dots + Y_d$, the graph $\mathfrak{X}_p$ corresponding to the infinite lift is the infinite d-regular tree.

Throughout this paper, we will be substituting *random* signed matchings and permutations into our matrix polynomials as a way to generate finite graphs that should locally look like the infinite lift.

In this setting, we have the following theorem. This is essentially the main result ("Theorem 2") of [BC19]. The small difference is that substituting *signed* rather than unsigned matchings/permutations gets rid of the "trivial" eigenvalues, as verified in [OW20, Thm. 1.9,10.10].

Theorem 2.17. *Let p be a self-adjoint matrix polynomial with coefficients from $\mathbb{C}^{r \times r}$ on indeterminates $\mathbb{1}, Y_1, \dots, Z_e^*$. Then for all $\epsilon, \beta > 0$ and sufficiently large n, the following holds:*

Let $\boldsymbol{A}_n$ be the operator on $\mathbb{C}^n \otimes \mathbb{C}^r$ obtained by a random ± 1-signed n-lift of p, and let A_∞ be the operator acting on $\ell_2(V_\infty) \otimes \mathbb{C}^r$ obtained from the ∞-lift on p.

Then except with probability at most β, the spectra $\sigma(\boldsymbol{A}_n)$ and $\sigma(A_\infty)$ are at Hausdorff distance at most ϵ

2.4 Maximizing the value of a matrix polynomial

We recall some standard notation, which will be used in this section. First, we write $\rho \ge 0$ to denote that $\rho \in \mathbb{C}^{n \times n}$ is positive semidefinite (and Hermitian). Similarly, we write that $\langle \rho, A \rangle$ to denote $\mathrm{tr}(\rho A)$, and note that when ρ and A are Hermitian this number is real and is equal to $\sum_{ij} \overline{\rho}_{ij} A_{ij}$, where $\bar{\cdot}$ denotes complex conjugate. One can restrict their attention to real matrices A, in which case the objects we will optimize over in this section may also be assumed to be real.

Throughout this paper, we will be discussing the problem of maximizing $\frac{1}{n} x^\top A x$ over $x \in \{\pm 1\}^n$, where A is an $n \times n$ Hermitian matrix. Since A is Hermitian, $x^\top A x = \langle x | A | x \rangle$ will always be real, and so it makes sense to maximize its value.

Definition 2.18 (Optimal Value)**.** Given an Hermitian matrix $A \in \mathbb{C}^{n \times n}$, we will be trying to compute the following value.

$$\textsc{Opt}(A) = \max_{x \in \{\pm 1\}^n} \frac{1}{n} x^\top A x = \max_{x \in \left\{\pm \frac{1}{\sqrt{n}}\right\}^n} x^\top A x$$

Note that $\textsc{Opt}(A) = \max_{\rho \in \mathrm{Cut}_n} \frac{1}{n} \langle \rho, A \rangle$, where Cut_n is the "cut polytope", the convex hull of all matrices of the form $xx^\top$ for $x \in \{\pm 1\}^n$. Since $\langle \rho, A \rangle$ is linear in ρ, maximizing over the convex hull is the same as maximizing over the extreme points, which are just those matrices of the form $xx^\top$.

Remark 2.19. The above problem is the same as maximizing a homogenous degree-2 polynomial defined by $P = \sum_{i,j} A_{ij} X_i X_j$ over the Boolean cube, and it is also the same as solving Max-Cut on an edge-weighted graph defined by $-A$.

The above problem has a natural relaxation to trying to maximize $x^\mathsf{T} A x$ over unit vectors x, which leads to the following efficiently computable upper bound on $\textsc{Opt}(A)$.

Definition 2.20 (Eigenvalue Bound)**.** Given an Hermitian matrix $A \in \mathbb{C}^{n\times n}$, we define the *eigenvalue bound* in the following way:

$$\textsc{Eig}(A) = \sup\{\langle \rho, A\rangle : \rho \geq 0,\ \operatorname{tr}\rho = 1\}.$$

This bound is maximizing the inner product between A and a "density matrix" ρ. Note that if $\varrho \in \mathrm{Cut}_n$, then $\rho = \frac{1}{n}\varrho$ is a density matrix. Thus, $\textsc{Eig}(A)$ is a relaxation of $\textsc{Opt}(A)$, or in other words, $\textsc{Opt}(A) \leq \textsc{Eig}(A)$.

The set of density matrices is convex, and it's well known that its extreme points are all the rank-1 density matrices; i.e., those ρ of the form xx^T for $x \in \mathbb{C}^n$ with $\|x\|_2^2 = 1$. $\textsc{Eig}(A)$ as defined above is therefore equivalent to maximizing over these ρ, and so we have the following equivalent definition of $\textsc{Eig}(A)$:

$$\textsc{Eig}(A) = \sup_{\substack{x\in\mathbb{C}^n \\ \|x\|_2^2=1}} \langle xx^\mathsf{T}, A\rangle = \sup_{\substack{x\in\mathbb{C}^n \\ \|x\|_2^2=1}} x^\mathsf{T} A x.$$

This formula is clearly equivalent to $\lambda_{\max}(A)$, which is why we refer to this bound as the "eigenvalue bound". One may also think of $\textsc{Eig}(A)$ and $\lambda_{\max}(A)$ as SDP duals of one another.

We now mention another well known, tighter, upper bound on $\textsc{Opt}(A)$.

Definition 2.21 (Basic SDP Bound)**.** Given a Hermitian matrix $A \in \mathbb{C}^{n\times n}$, we refer to the basic SDP bound as

$$\textsc{Sdp}(A) = \sup_{\substack{\rho\geq 0 \\ \rho\in\mathbb{R}^{n\times n}}} \langle \rho, A\rangle \quad \text{subject to } \rho_{ii} = \tfrac{1}{n}, \forall i$$

In this SDP, we are essentially optimizing over "correlation matrices". Formally, a "correlation matrix" [Sty73] is a PSD matrix whose diagonal entries are 1, so we're equivalently maximizing $\frac{1}{n}\langle \rho, A\rangle$ over correlation matrices.

We also note that any cut matrix is a correlation matrix, and any correlation matrix is a density matrix, and so $\textsc{Opt}(A) \leq \textsc{Sdp}(A) \leq \textsc{Eig}(A)$.

The above bound has the following dual SDP [DP93]:

$$\textsc{Sdp-Dual}(A) = \inf_{\substack{u\in\mathbb{R}^n \\ \operatorname{avg}(u_1,\ldots,u_n)=0}} \lambda_{\max}(A + \operatorname{diag}(u)).$$

It is clear that $\textsc{Sdp-Dual}(A) \leq \lambda_{\max}(A)$.

Despite the fact that the usual "Slater condition" for strong SDP duality fails because the set of correlation matrices isn't full-dimensional, one can still show that $\textsc{Sdp}(A) = \textsc{Sdp-Dual}(A)$ holds for finite A. [PR95]

In the above definitions and discussions, we assume that the matrices we are operating on are scalar valued. It is still of interest to us to define an analogous definition for matrices with matrix valued entries.

Definition 2.22. For any A with $r \times r$ matrix entries and $\widetilde{A}$ its extension, we define the above bounds Opt, Eig, and Sdp in the following way:

$$\textsc{Opt}(A) = \textsc{Opt}(\widetilde{A})$$

and similarly

$$\textsc{Eig}(A) = \textsc{Eig}(\widetilde{A})$$

For finite A, we can have that

$$\textsc{Sdp}(A) = \textsc{Sdp}(\widetilde{A})$$

As mentioned in Section 1, we see that our definition of $\textsc{Eig}(A)$ makes sense for infinite graphs of finite degree, and more generally for operators A with finite operator norm $\|A\|$. However, $\textsc{Sdp}(A)$ does not extend as well, as the number "n" appearing in its definition is not finite.

The infinite graphs/operators we consider in this paper have the special property that they are $r \times r$ edge-weighted. For such graphs, we introduce a new quantity that goes *part of the way* in the cut→correlation, Eig→SDP improvement, which still makes sense for infinite graphs.

We define the following value as an infinite version of the basic SDP bound:

Definition 2.23 (Partitioned SDP Bound)**.** Let us write $M_n(\mathbb{C}^{r\times r})$ for the $n \times n$ matrices with entries from $\mathbb{C}^{r\times r}$. (In Section 3 we will suitably generalize to the "$n = \infty$" case.) The trace of such a matrix is the sum of the diagonal $r \times r$ matrices. For Hermitian $A \in M_n(\mathbb{C}^{r\times r})$, we define

$$\textsc{PartSdp}(A) = \sup_{\substack{\rho \geq 0 \\ \rho \in M_n(\mathbb{C}^{r\times r})}} \langle \rho, A \rangle \quad \text{subject to } \operatorname{tr}(\rho) = 1/r \cdot c_{r\times r}$$

where $c_{r\times r}$ denotes some correlation matrix of dimension $r \times r$. This is equivalent to requiring that $\operatorname{tr}(\rho)$ has $1/r$ as its diagonal entries.

This has the following SDP dual:

$$\textsc{PartSdp-Dual}(A) = \inf_{\substack{u \in \mathbb{R}^r \\ \operatorname{avg}(u_1,\dots,u_r)=0}} \lambda_{\max}(A + \mathbb{1} \otimes \operatorname{diag}(u))$$

The goal of Section 3 is to show that strong duality holds for this program as well. This section also treats the case of infinite A more formally.

Remark 2.24. Given some $n \times n$ matrix A, we can think of A as an $r \times r$ matrix A_r, $r = n$, with 1×1 entries; *or*, we can think of it as a 1×1 matrix A_1 with a single $n \times n$ entry. Then, we have that $\textsc{PartSdp}(A_r) = \textsc{Eig}(A_r)$ and $\textsc{PartSdp}(A_1) = \textsc{Sdp}(A_1)$.

We recall our main theorem.

Theorem 1.2. *Let p be a matrix polynomial with $r \times r$ coefficients. Let A_∞ be the adjacency operator (with $r \times r$ entries) of the associated infinite lift $\mathfrak{X}_p$, and write $s_p^* = \textsc{PartSdp}(A_\infty) = \textsc{PartSdp-Dual}(A_\infty)$. Then for any $\epsilon > 0$, if $\boldsymbol{A}_n$ is the adjacency matrix of a random edge-signed n-lift of p, it holds that*

$$s_p^* - \epsilon \leq \textsc{Sdp}(\boldsymbol{A}_n) \leq s_p^* + \epsilon$$

except with probability at most $o_{n\to\infty}(1)$.

We prove our statement in two parts, by first proving the upper bound $\textsc{Sdp}(\widetilde{A_n}) \leq \textsc{PartSdp}(A_\infty) + \epsilon$ in Section 4 followed by the lower bound $\textsc{Sdp}(\widetilde{A_n}) \geq \textsc{PartSdp}(A_\infty) - \epsilon$ in Section 4.2.

3 The infinite SDPs

This technical section has two goals. First, in Section 3.1 we show that strong duality $\textsc{PartSdp}(A) = \textsc{PartSdp-Dual}(A)$ holds, even for infinite matrices A (with $r \times r$ entries). Even in the finite case this is not trivial, as the feasible region for the SDP $\textsc{PartSdp}(A)$ is not full-dimensional, and hence the Slater condition ensuring strong duality does not apply. (This difficulty even appears for the Basic SDP.) The infinite case involves some additional technical considerations, needed so that we may eventually apply the strong duality theorem for conic linear programming of Bonnans and Shapiro [BS00, Thm. 2.187]. Second, in Section 3.2, we show that in the optimization problem $\textsc{PartSdp}(A)$, values arbitrarily close to the optimum can be achieved by matrices ρ of finite support (i.e., only finitely many nonzero entries). Indeed (though we don't need this fact), these finite-rank ρ need only have rank at most r. Finally, in Section 3.3 we consolidate all these results into a theorem statement suitable for use with infinite lifts of matrix polynomial graphs.

3.1 SDP duality

Let V be a countable set and write $\mathcal{H} = \ell_2(V)$ for the associated (complex, separable) Hilbert space of square-summable functions $f : V \to \mathbb{C}$. We write $B_{00}(\mathcal{H})$, $B_1(\mathcal{H})$, $B(\mathcal{H})$ for the spaces of finite-rank, trace-class, and bounded operators on $\mathcal{H}$, respectively. Focusing on $B_1(\mathcal{H})$ (with the weak topology) and $B(\mathcal{H})$ (with the σ-weak topology), these are locally convex topological vector spaces forming a dual pair with bilinear map $\langle \cdot, \cdot \rangle : B_1(\mathcal{H}) \times B(\mathcal{H}) \to \mathbb{C}$ defined by $\langle \rho, a \rangle = \operatorname{tr}(\rho a)$ (see [RS80, Thm. VI.26], [Lan17, Thm. B.146], or [BR02, Prop. 2.4.3]). Recall [Lan17, Lem. B.142] that $\langle \rho, a \rangle \le \|\rho\|_1 \|a\|$, where $\|\rho\|_1 = \operatorname{tr}\sqrt{\rho^\dagger \rho}$ is the trace-norm.

Write $B_1(\mathcal{H})_{\mathrm{sa}}$ (respectively, $B(\mathcal{H})_{\mathrm{sa}}$) for the (closed) real subspace of self-adjoint operators in $B_1(\mathcal{H})$ (respectively, $B(\mathcal{H})_{\mathrm{sa}}$); note that $B_1(\mathcal{H})_{\mathrm{sa}}$ and $B(\mathcal{H})_{\mathrm{sa}}$ continue to form a dual pair (see, e.g., [Mey06, p. 212]). Recall that $a \in B(\mathcal{H})_{\mathrm{sa}}$ is positive semidefinite if and only if $\langle \varphi | a | \varphi \rangle \ge 0$ for all $|\varphi\rangle \in \mathcal{H}$ (for such a we have $\sqrt{a^\dagger a} = a$); as usual we write $b \ge a$ to mean that $b - a$ is positive semidefinite. Letting $B_1(\mathcal{H})_+$ (respectively, $B(\mathcal{H})_+$) denote the positive semidefinite operators in $B_1(\mathcal{H})_{\mathrm{sa}}$ (respectively, $B(\mathcal{H})_{\mathrm{sa}}$), we have that these are both (nonempty) closed, convex cones (for $B(\mathcal{H})_+$, see [Con90, Prop. 3.7], [Lan17, Prop. C.51]; for $B_1(\mathcal{H})_+$ see [Fri19, Sec. 4], [vDE20, App. A.2], [SH08, Sec. 2]); further, they are topologically dual cones [vDE20, App. A.2].

Our SDP. Fix any $a \in B(\mathcal{H})_{\mathrm{sa}}$. Let $V = V_1 \sqcup V_2 \sqcup \cdots \sqcup V_r$ be a partition of V into $r \in \mathbb{N}_+$ nonempty parts, and for $1 \le j \le r$ let $I_j \in B(\mathcal{H})_{\mathrm{sa}}$ be the operator $\sum_{v \in V_j} |v\rangle\langle v|$. Consider the following semidefinite program:

$$\sup_{\rho \in B_1(\mathcal{H})_+} \langle \rho, a \rangle \quad \text{subject to} \quad \langle \rho, I_j \rangle = \tfrac{1}{r},\ j = 1 \dots r. \tag{SDP-P}$$

For our application, suppose A is a self-adjoint matrix, indexed by countable vertex V_∞ and with $r \times r$ entries (finitely many nonzero ones per row/column). Let $a = \widetilde{A}$, with rows/columns indexed by $V = V_\infty \times [r]$, and let $V_j = V_\infty \times \{j\}$. Then (SDP-P) is precisely how we define $\textsc{PartSdp}(A)$.

We remark that (SDP-P) is always feasible. By summing the constraints on ρ we get $\operatorname{tr}\rho = 1$; i.e., ρ is required to be a density operator. In particular, $\langle \rho, a \rangle \le \operatorname{tr}(\rho)\|a\|$ always so the optimum value of (SDP-P) is finite.

We would like to show that the optimum value of (SDP-P) is equal to that of the following dual semidefinite program which, in our application mentioned above, is $\textsc{PartSdp-Dual}(A)$:

$$\inf_{u \in \mathbb{R}^r} \lambda_{\max}(a + u_1 I_1 + \cdots + u_r I_r) \quad \text{subject to} \quad \operatorname{avg}(u_1, \dots, u_r) = 0. \tag{SDP-D}$$

Showing this will take a couple of steps. The first is to mechanically write down the Lagrangian dual of (SDP-P), which is:

$$\inf_{y \in \mathbb{R}^r} \tfrac{1}{r} \sum_j y_j \quad \text{subject to} \quad a \leq y_1 I_1 + \cdots + y_r I_r. \tag{D1}$$

We first show that (D1) is equivalent to (SDP-D). We may reparameterize all $y \in \mathbb{R}^r$ by defining $\overline{y} = \frac{1}{r}\sum_j y_j$ and writing $y_j = \overline{y} - \frac{1}{r}$; in this way, every $y \in \mathbb{R}^r$ corresponds to a pair $\overline{y} \in \mathbb{R}$ and $u \in \mathbb{R}^r$ satisfying $\operatorname{avg}(u_1, \dots, u_r) = 0$, and vice versa. Under this reparameterization, (D1) becomes

$$\inf_{\substack{\overline{y} \in \mathbb{R} \\ u \in \mathbb{R}^r}} \overline{y} \quad \text{subject to} \quad \operatorname{avg}(u_1, \dots, u_r) = 0 \text{ and } a \leq \overline{y} I - \sum_j u_j I_j. \tag{D2}$$

The second constraint here is equivalent to $\lambda_{\max}(a + \sum_j u_j I_j) \leq \overline{y}$, and hence the optimal choice of $\overline{y}$ given u is achieved by $\lambda_{\max}(a + \sum_j u_j I_j)$. Thus (D1) is equivalent to (D2) is equivalent to (SDP-D), as claimed.

We would next like to claim that strong duality holds for the dual pair (SDP-P) and (D1); however, there is a difficulty because the feasible region of (SDP-P) only has nonempty *relative interior*, not interior. Alternatively, one might say the difficulty is that the feasible region for (D1) is not bounded. However we can fix this by introducing the following equivalent bounded variant:

$$\inf_{y \in \mathbb{R}^r} \tfrac{1}{r} \sum_j y_j \quad \text{subject to} \quad a \leq y_1 I_1 + \cdots + y_r I_r \text{ and } |y_j| \leq C, \; j = 1 \dots r, \tag{D3}$$

where $C = 2r\|a\|$.

Claim 3.1. (D3) *is equivalent to* (D1).

Proof. Suppose we have any feasible solution y for (D1). First observe that for any $1 \leq j \leq r$, if we take an arbitrary $v_j \in V_j$ we have

$$y_1 I_1 + \cdots + y_r I_r \geq a \implies y_j \geq \langle v_j | a | v_j \rangle \geq -\|a\|. \tag{2}$$

Next observe that the optimum value of (D1) is at most $\|a\|$ (since we could take $y_j = \|a\|$ for all j); thus we may restrict attention to y's that achieve objective value at most $\|a\|$. But now we claim any such feasible y will have $y_k \leq 2r\|a\|$ for any $1 \leq k \leq r$. To see this, observe (using Inequality (2)) that

$$\tfrac{1}{r} \sum_j y_i \geq \tfrac{1}{r}(y_k - \sum_{j \neq k} \|a\|) \geq y_k / r - \|a\|.$$

Thus if $y_k > 2r\|a\|$, it is strictly not optimal. We conclude that we may add the constraints $-\|a\| \leq y_j \leq 2r\|a\|$ for $j = 1 \dots r$ to (D1) without changing it. Clearly it doesn't hurt to widen this interval to $-C \leq y_j \leq C$ (for the sake of symmetry), and so we conclude that (D3) is indeed equivalent to (D1). $\square$

We can now write the Lagrangian dual of (D3), which is

$$\sup_{\rho \in B_1(\mathcal{H})_+} \langle \rho, a \rangle - C \sum_{j=1}^{r} |\langle \rho, I_j \rangle - \tfrac{1}{r}|. \tag{P3}$$

Claim 3.2. (P3) *is equivalent to* (SDP-P).

Proof. Clearly the objective value of (SDP-P) is at most that of (P3), so it suffices to show the converse inequality. To that end, let ρ be any feasible solution for (P3), and write the objective value as $o(\rho) - p(\rho)$, where $o(\rho) = \langle \rho, a\rangle$ ("objective") and $p(\rho) = C\sum_j |\langle \rho, I_j\rangle - \frac{1}{r}|$ ("penalty").

Let $\epsilon = (\max_j\{r\langle \rho, I_j\rangle\} - 1)_+$, so that $\langle \rho, I_j\rangle \le \frac{1}{r}(1+\epsilon)$ for all j, and write $\widetilde{\rho} = \rho/(1+\epsilon) \in B_1(\mathcal{H})_+$. Note that $p(\widetilde{\rho}) = C\sum_j \delta_j$, where $\delta_j = \frac{1}{r} - \langle \widetilde{\rho}, I_j\rangle \ge 0$. Finally, define $\rho' = \widetilde{\rho} + \sum_j \delta_j |v_j\rangle\langle v_j| \in B_1(\mathcal{H})_+$, where each v_j is an arbitrary element of V_j. By construction, $p(\rho') = 0$. We now want to show

$$o(\rho') \ge o(\rho) - p(\rho) \qquad \Longleftrightarrow \qquad p(\rho) \ge o(\rho) - o(\rho'). \tag{3}$$

We will then have that ρ' is feasible for (SDP-P), with an objective value at least that of ρ's in (P3); this will complete the proof.

First, we have

$$p(\rho) = C\sum_j |\langle \rho, I_j\rangle - \tfrac{1}{r}| = C\sum_j |(1+\epsilon)\langle \widetilde{\rho}, I_j\rangle - \tfrac{1}{r}| = 2r\|a\|\sum_j |\epsilon/r - (1+\epsilon)\delta_j|.$$

On the other hand, we have

$$o(\rho) - o(\rho') = \langle \rho, a\rangle - \tfrac{1}{1+\epsilon}\langle \rho, a\rangle - \delta_j \sum_j \langle v_j|a|v_j\rangle \le \tfrac{\epsilon}{1+\epsilon}\langle \rho, a\rangle + \|a\|\sum_j \delta_j \le \|a\|(\tfrac{\epsilon}{1+\epsilon}\operatorname{tr}\rho + \sum_j \delta_j).$$

Summing the inequality $\langle \rho, I_j\rangle \le \frac{1}{r}(1+\epsilon)$ yields $\operatorname{tr}\rho \le 1+\epsilon$, so to establish Inequality (3) it remains to show

$$2r\sum_j |r\epsilon - (1+\epsilon)\delta_j| \ge \epsilon + \sum_j \delta_j \qquad \Longleftrightarrow \qquad \sum_j (2r|r\epsilon - (1+\epsilon)\delta_j| - \delta_j) \ge \epsilon. \tag{4}$$

This is certainly true if $\epsilon = 0$, so it remains to assume $\epsilon > 0$. Then writing j_0 for the j achieving $\max_j\{r\langle \rho, I_j\rangle\}$, we have $\delta_{j_0} = 0$ by construction. This gives us a contribution of 2ϵ to the sum on the right in Inequality (4). On the other hand, for $j \ne j_0$, the worst possible value for δ_j is $\frac{\epsilon}{r(1+\epsilon)}$, in which case the contribution to the sum is $-\frac{\epsilon}{r(1+\epsilon)}$; summing this over all $j \ne j_0$ yields at worst $-\frac{\epsilon}{1+\epsilon}$. Thus we have shown the sum in Inequality (4) is at least $2\epsilon - \frac{\epsilon}{1+\epsilon} \ge \epsilon$, as needed. ☐

By inspecting the above proof we may extract the following lemma, which will be useful later:

Lemma 3.3. *Suppose $\rho \in B_1(\mathcal{H})_+$ has $|\langle \rho, I_j\rangle - \frac{1}{r}| \le \eta$ for each $1 \le j \le r$. Then there is $\rho' \in B_1(\mathcal{H})_+$ satisfying $\langle \rho', I_j\rangle = \frac{1}{r}$ for each j and with $|\langle \rho, a\rangle - \langle \rho', a\rangle| \le Cr\eta$. Furthermore, if ρ has finite rank (respectively, support) then so too does ρ'.*

Returning to (P3), formally speaking we are regarding it as the conic linear program

$$\sup_{(\rho,\lambda^+,\lambda^-)\in B_1(\mathcal{H})_+\oplus\mathbb{R}^r_{\ge 0}\oplus\mathbb{R}^r_{\ge 0}} \langle \rho, a\rangle - C\sum_{j=1}^r (\lambda_j^+ + \lambda_j^-) \quad \text{subject to} \quad \langle \rho, I_j\rangle - \tfrac{1}{r} = \lambda_j^+ - \lambda_j^-,\ j = 1\ldots r. \tag{P3}$$

Not only is this feasible, meaning the vector $(\frac{1}{r}, \ldots, \frac{1}{r})$ is in the region

$$D = \Big\{\big(\langle \sigma, I_1\rangle - (\lambda_1^+ - \lambda_1^-), \ldots, \langle \sigma, I_r\rangle - (\lambda_1^+ - \lambda_1^-)\big) : \sigma \in B_1(\mathcal{H})_+,\ \lambda^+ \in \mathbb{R}^r_{\ge 0},\ \lambda^- \in \mathbb{R}^r_{\ge 0}\Big\} \subseteq \mathbb{R}^r,$$

but it is even in the interior of the region, since D is in fact all of $\mathbb{R}^r$. We are therefore finally in a position to apply a strong duality theorem for conic linear programming [BS00, Thm. 2.187] to the pair (P3), (D3). We conclude:

Theorem 3.4. (SDP-P) *and* (SDP-D) *have a common value s^*, and s^* is achieved in* (SDP-D) *by some $\hat{u} \in \mathbb{R}^r$. We may therefore write $s^* = \lambda_{\max}(\hat{a})$, where $\hat{a} = a + \sum_j \hat{u}_j I_j$.*

3.2 Nearly optimal finite, rank-r solutions

Although the optimal value s^* is achieved in (SDP-D) (that is, its "inf" may be written "min"), the value s^* might not be achieved in (SDP-P). However for any $\delta > 0$, we can find $\rho \in B_1(\mathcal{H})_+$ satisfying $\langle \rho, I_j \rangle = \frac{1}{r}$ for all j and with $\langle \rho, a \rangle \geq s^* - \delta$. Let us further simplify this ρ.

Finite rank. First, recall that $B_{00}(\mathcal{H})_+$ (the finite-rank positive semidefinite operators on $\mathcal{H}$) is dense in $B_1(\mathcal{H})_+$ with respect to the trace-norm (see [Con85, Thm. 1.11(d)], with the proof clearly holding under the positive semidefinite restriction). Thus we can find a finite-rank $\rho_0 \in B_{00}(\mathcal{H})_+$ satisfying $|\langle \rho_0, I_j \rangle - \frac{1}{r}| \leq \delta$ for all j and with $\langle \rho_0, a \rangle \geq s^* - \delta - \|a\|\delta$. Next, using Lemma 3.3 we can convert this to another finite-rank $\rho_1 \in B_{00}(\mathcal{H})_+$ satisfying $\langle \rho_1, I_j \rangle = \frac{1}{r}$ for all j and with $\langle \rho_1, a \rangle \geq s^* - \delta_1$, where $\delta_1 = \delta + \|a\|\delta + C\delta \to 0$ as $\delta \to 0$.

Finite support. Since $\rho_1 \in B_{00}(\mathcal{H})_+$ has trace 1 we can write $\rho_1 = \sum_{i=1}^{k_1} \lambda_i |\varphi_i\rangle\langle\varphi_i|$ for some $\lambda = (\lambda_1, \ldots, \lambda_{k_1}) \in \mathbb{R}_+^{k_1}$ forming a probability distribution on $[k_1]$ and some orthonormal vectors $|\varphi_i\rangle \in \ell_2(V)$. We can approximate each $|\varphi_i\rangle$ by a unit vector $|\varphi_i'\rangle$ of finite support satisfying $|\langle \varphi_i | \varphi_i' \rangle|^2 \geq 1 - \delta_1$, from which it follows (an easy calculation [NC10, (9.60), (9.99)]) that $\|e_i\|_1 \leq 2\sqrt{\delta_1}$, where $e_i = |\varphi_i\rangle\langle\varphi_i| - |\varphi_i'\rangle\langle\varphi_i'|$. Now let $\rho_2 = \sum_{i=1}^{k_1} \lambda_i |\varphi_i'\rangle\langle\varphi_i'|$, a positive semidefinite matrix of finite support. Then

$$\langle \rho_2, a \rangle = \langle \rho_1, a \rangle - \sum_{i=1}^{k_1} \lambda_i \langle e_i, a \rangle \geq s^* - \delta_1 - \sum_{i=1}^{k_1} \lambda_i \cdot 2\sqrt{\delta_1}\|a\| = s^* - \delta_1 - 2\sqrt{\delta_1}\|a\|,$$

and one can similarly check that $|\langle \rho_2, I_j \rangle - \frac{1}{r}| \leq 2\sqrt{\delta_1}$ for all j. Again applying Lemma 3.3, we can convert this to another finite-support $\hat{\rho} \in B_{00}(\mathcal{H})_+$ satisfying $\langle \hat{\rho}, I_j \rangle = \frac{1}{r}$ for all j and with $\langle \rho_1, a \rangle \geq s^* - \delta_2$, where $\delta_2 = \delta_1 + 2\sqrt{\delta_1}\|a\| + 2C\sqrt{\delta_1} \to 0$ as $\delta \to 0$. Since δ_2 can be made arbitrarily small by taking $\delta \to 0$, we finally conclude:

Proposition 3.5. *For any $\epsilon > 0$, there is a feasible finite-support solution $\hat{\rho}$ to* (SDP-P) *achieving objective value at least $s^* - \epsilon = \lambda_{\max}(\hat{a}) - \epsilon$.*

Rank r. In general, any solution ρ to (SDP-P) need not have rank more than r. Although we don't strictly need this fact, it is particularly easy to show for ρ of finite support, and we do so now.

Proposition 3.6. *Let $\hat{\rho}$ be a feasible solution to* (SDP-P)*, with nonzero entries only in rows/columns indexed by a finite set $F \subseteq V$. Then there is another $\tilde{\rho}$ feasible for* (SDP-P)*, supported on F and with rank at most r such that $\langle \tilde{\rho}, a \rangle \geq \langle \hat{\rho}, a \rangle$.*

Proof. We think of $\hat{\rho}$ as a matrix indexed just by $\hat{F}$, and write $a_{\hat{F}}$ for the submatrix of a on rows/columns indexed by $\hat{F}$. Suppose $\hat{\rho}$ has eigenvalues $\mu_1, \ldots, \mu_{|\hat{F}|} \geq 0$ (nonnegative, since $\hat{\rho} \geq 0$), with corresponding orthonormal eigenvectors $|\psi_1\rangle, \ldots, |\psi_{|\hat{F}|}\rangle$. Then $\hat{\rho}$ achieves objective value

$$c := \langle \hat{\rho}, a \rangle = \langle \hat{\rho}, a_{\hat{F}} \rangle = \mathrm{tr}(\hat{\rho} a_{\hat{F}}) = \mathrm{tr}\left(\left(\sum_i \mu_i |\psi_i\rangle\langle\psi_i|\right) a_{\hat{F}}\right) = \sum_i \mu_i \langle \psi_i | a_{\hat{F}} | \psi_i \rangle .$$

Also, writing I_j for the identity matrix restricted to $V_j \cap \hat{F}$, feasibility of ρ implies

$$\tfrac{1}{r} = \langle \hat{\rho}, I_j \rangle = \sum_i \mu_i \langle \psi_i | I_j | \psi_i \rangle , \quad j = 1 \ldots r.$$

Now if we imagine $|\psi_1\rangle, \dots, |\psi_{|\hat{F}|}\rangle$ are fixed, and $\lambda_1, \dots, \lambda_{|\hat{F}|}$ are real variables, the above tells us that the following linear program —

$$\max_{\lambda \in \mathbb{R}^{\geq 0}_{|\hat{F}|}} \left\{ \sum_i \lambda_i \langle \psi_i | a_{\hat{F}} | \psi_i \rangle : \sum_i \lambda_i \langle \psi_i | I_j | \psi_i \rangle = \tfrac{1}{r},\ j = 1 \dots r \right\} \tag{LP}$$

— has optimal value at least c, since we may take $\lambda = \mu$. By standard theory of finite linear programs, the optimal value occurs at a vertex where $|\hat{F}|$ linearly independent constraints are tight. Since there are only r equality constraints in (LP), there is an optimal solution λ^* with at least $|\hat{F}| - r$ zero entries. Thus $\tilde{\rho} = \sum_i \lambda_i^* |\psi_i\rangle\langle\psi_i|$ has rank at most r, achieves objective value at least c, and is feasible for (SDP-P). □

3.3 Conclusion for matrix edge-weighted graphs

Let V_∞ be a countable set of nodes and let G_∞ be a bounded-degree graph on V_∞ with matrix edge-weights from $\mathbb{C}^{r\times r}$. Let A_∞ be the adjacency operator for G_∞, acting on $\ell_2(V_\infty) \otimes \mathbb{C}^r$ and assumed self-adjoint; we may think of it as an infinite matrix with rows and columns indexed by V_∞, and with entries from $\mathbb{C}^{r\times r}$. As in Definition 2.2, we write $\widetilde{A}_\infty$ for its "extension"; this is a self-adjoint bounded operator on $\ell_2(V_\infty \times [r])$. Conversely, given any solution $\hat{\rho}$ for (SDP-P) from Section 3, we may "unextend" it and think of it as an infinite matrix ρ with rows and columns indexed by V_∞, and entries from $\mathbb{C}^{r\times r}$.

We apply the the SDP duality theory from Sections 3 and 3.2, with "V" being $V_\infty \times [r]$, "a" being $\widetilde{A}_\infty$, and "V_j" being $V_\infty \times \{j\}$ for all $1 \leq j \leq r$. The conclusion is that for any $\epsilon > 0$, there exists:

- $\hat{u} \in \mathbb{R}^r$ with $\operatorname{avg}_j(\hat{u}_j) = 0$;
- a finite subset $F \subset V_\infty$;
- a PSD matrix ρ (the "unextension" of $\tilde{\rho}$ of rank at most r from Proposition 3.6) with rows/columns indexed by V_∞ and entries from $\mathbb{C}^{r\times r}$, supported on the rows/columns F, with
$$\operatorname{tr}(\rho)_{jj} = \tfrac{1}{r}, \quad j = 1 \dots r;$$

such that for
$$\hat{A} = \widetilde{A}_\infty + \mathbb{1}_{V_\infty} \otimes \operatorname{diag}(\hat{u}),$$
we have
$$s^* := \lambda_{\max}(\hat{A}) = \textsc{PartSdp-Dual}(A_\infty) = \textsc{PartSdp}(A_\infty) \geq \langle \rho, A_\infty \rangle = \langle \rho_F, A_F \rangle \geq s^* - \epsilon, \tag{5}$$
where ρ_F, A_F denote ρ, A_∞ (respectively) restricted to the rows/columns F.

4 The SDP value of random matrix polynomial lifts

In this section we prove our main Theorem 1.2. To that end, let p be any self-adjoint matrix polynomial over indeterminates $Y_1, \dots, Y_d, Z_1, \dots, Z_e^*$ with $r \times r$ coefficients. Let A_∞ denote the adjacency operator of the infinite lift $\mathfrak{X}_p$, and write $s^* = \textsc{PartSdp}(A_\infty) = \textsc{PartSdp-Dual}(A_\infty)$

as in Equation (5). Fix any $\epsilon > 0$, and let $\boldsymbol{A}_n$ denote the adjacency matrix of a random edge-signed n-lift of p, with $\boldsymbol{A}_n = p(M_1, ..., P_1, ...)$ for random edge-signed matchings $\boldsymbol{M}_i$ and random edge-signed permutations $\boldsymbol{P}_i$. Our goal is to show that except with probability $o_{n\to\infty}(1)$,

$$s^* - \epsilon \leq \textsc{Sdp}(\boldsymbol{A}_n) = \textsc{Sdp-Dual}(\boldsymbol{A}_n) \leq s^* + \epsilon.$$

Given our setup, the upper bound follows easily from prior work, namely Theorem 2.17. Let $\hat{u}$ and $\hat{A}$ be as in Section 3.3, and consider the matrix polynomial p' defined by

$$p' = p + \operatorname{diag}(\hat{u})\mathbb{1}.$$

Then on one hand, the ∞-lift of p' has adjacency operator precisely $\hat{A}$; on the other hand,

$$\boldsymbol{A}'_n := p'(\boldsymbol{M}_1, \ldots, \boldsymbol{P}_1, \ldots) = \boldsymbol{A}_n + \mathbb{1}_{n\times n} \otimes \operatorname{diag}(\hat{u}).$$

Thus Theorem 2.17 tells us that except with probability $o_{n\to\infty}(1)$, the spectra $\sigma(\boldsymbol{A}'_n)$ and $\sigma(\hat{A})$ are at Hausdorff distance at most ϵ, from which it follows that

$$\lambda_{\max}(\boldsymbol{A}'_n) \leq \lambda_{\max}(\hat{A}) + \epsilon = s^* + \epsilon.$$

But this indeed proves $\textsc{Sdp-Dual}(\boldsymbol{A}_n) \leq s^* + \epsilon$, because $\hat{u}$ has $\operatorname{avg}(\hat{u}) = 0$ and hence is feasible for $\textsc{Sdp-Dual}(\boldsymbol{A}_n)$.

It therefore remains to prove $\textsc{Sdp}(\boldsymbol{A}_n) \geq s^* - \epsilon$.

4.1 Adapting the partitioned SDP solution

Recalling the discussion in Section 3.3, we have that given any solution $\hat{\rho}$ for (SDP-P), we can "unextend" it and think of it as an infinite matrix ρ with rows and columns indexed by V_∞, and entries from $\mathbb{C}^{r\times r}$. Then, we have some finite subset $F \subset V_\infty$ such that ρ is supported on the rows/columns of F, and for any $\epsilon > 0$,

$$\rho \text{ is PSD}, \quad \sum_{v\in F} \rho_{v,v} = \tfrac{1}{r}\mathbb{1}, \quad s^* := \textsc{PartSdp}(A_\infty) \geq \langle \hat{\rho}, A_\infty \rangle = \langle \rho, A_F \rangle \geq s^* - \epsilon, \tag{6}$$

where A_F denotes A_∞ restricted to the rows/columns F, and $\langle \rho, A_F \rangle$ continues to denote $\operatorname{tr}(\rho A_F)$. Because F is finite, we are able to assume that without loss of generality, G_F is a ball of radius ℓ for some finite ℓ.

Let us now suppose we have a finite graph G_n on vertex set $[n]$, again with matrix edge-weights from $\mathbb{C}^{r\times r}$. (One should think of n as growing "large", with $|F|$ being "constant".) We use the notation A_n for G_n's adjacency matrix (an operator on $\mathbb{C}^n \otimes \mathbb{C}^r$), and $\widetilde{A}_n$ for A_n's "extension" (an operator on $\mathbb{C}^{nr}$). We write $(|s,j\rangle)_{s\in[n],j\in[r]}$ for the standard orthonormal basis on $\mathbb{C}^{nr}$. The basic "Goemans–Williamson" SDP for $\widetilde{A}_n$, normalized by factor $\frac{1}{nr}$, is

$$\max_{\substack{\widetilde{X}\in\mathbb{C}^{nr\times nr} \\ \text{PSD}}} \langle \widetilde{X}, \widetilde{A}_n \rangle \quad \text{subject to} \quad \langle s,j|\widetilde{X}|s,j\rangle = \tfrac{1}{nr} \ \forall s \in [n], j \in [r]. \tag{SDP}$$

We recall that every vertex in A_∞ (and therefore A_F) can be referred to as a reduced word over the indeterminates of p. For every $t \in [n]$ we define $\Pi_t = \sum_{v\in F} \mathcal{L}^v |t\rangle \langle v| \otimes \mathbb{1}$, where $\mathbb{1}$ is the identity operator on $\mathbb{C}^r$. $\mathcal{L}^v |t\rangle$ ends up being equivalent to $s_{u,t} |u\rangle$ for some $u \in [n]$, where $s_{u,t} \in \{\pm 1\}$ based off of the signings of the substitutions in $\mathcal{L}$.

We may then define the operator X_t on $\mathbb{C}^n \otimes \mathbb{C}^r$ by

$$X_t = (\Pi_t)\rho(\Pi_t)^\top.$$

In other words, $(X_t)_{a,b} = \sum\{\rho_{v,w} : v, w \in F,\ \pm|a\rangle = \mathcal{L}^v|t\rangle,\ \pm|b\rangle = \mathcal{L}^w|t\rangle\}$. This operator is PSD (since ρ is), and it satisfies the following:

Lemma 4.1. *Let $t \in [n]$ and suppose there is no reduced, nonempty word w of length at most $2|F|$ with $\mathcal{L}^w|t\rangle$ evaluating to $\pm|t\rangle$. Then*

$$(\Pi_t)^\top A_n(\Pi_t) = A_F$$

and $\langle X_t, A_n\rangle = \langle \rho, A_F\rangle$.

Proof. Consider some $t \in [n]$ which satisfies the condition in the lemma statement. For any $a, b \in F$, we can express the corresponding entry of $(\Pi_t)^\top A_n(\Pi_t)$ in the following way:

$$\left((\Pi_t)^\top A_n(\Pi_t)\right)_{a,b} = \sum\{s_{v,t}s_{w,t}(A_n)_{v,w} : v, w \in [n],\ s_{v,t}|v\rangle = \mathcal{L}^a|t\rangle,\ s_{w,t}|w\rangle = \mathcal{L}^b|t\rangle\}$$

There's exactly one pair $v, w \in [n]$ which satisfies the condition we are summing over. In addition, because there's no nonempty word w of length at most $2|F|$ with $\pm|t\rangle = \mathcal{L}^w|t\rangle$, we know that for any vertex $v \in [n]$, there's exactly one $a \in F$ such that $\pm|v\rangle = \mathcal{L}^a|t\rangle$.

This means that the value of $(A_n)_{v,w}$ exactly corresponds to the weight of the edge $(a, b) \in A_F$. While this edge in A_n would be signed depending on the signs of substitutions in $\mathcal{L}$, we note that the factor of $s_{v,t}s_{w,t}$ cancels these signs out.

Therefore, we have that for any $a, b \in F$,

$$((\Pi_t)^\top A_n(\Pi_t))_{a,b} = (A_F)_{a,b}$$

and so $(\Pi_t)^\top A_n(\Pi_t) = A_F$.

From this, it follows that

$$\langle \rho, A_F\rangle = \mathrm{tr}(\rho A_F) = \mathrm{tr}(\rho(\Pi_t)^\top A_n(\Pi_t)) = \mathrm{tr}((\Pi_t)\rho(\Pi_t)^\top A_n) = \mathrm{tr}(X_t A_n) = \langle X_t, A_n\rangle$$

□

As a consequence, if this condition holds true for all $t \in [n]$, then

$$X = \operatorname*{avg}_{t\in[n]} X_t$$

is PSD and has $\langle X, A_n\rangle = \langle \rho, A_F\rangle$.

Ideally, we would like if X had $X_{s,s} = \frac{1}{nr}\mathbb{1}$ for each $s \in [n]$, as then its extension $\widetilde{X}$ would be a solution to (SDP) of value $\langle \rho, A_F\rangle \geq \|\hat{A}\| - \epsilon$. We see that we can achieve this for $s \in [n]$ where the same condition in Lemma 4.1 is held:

Lemma 4.2. *Let $s \in [n]$ and suppose there is no reduced, nonempty word $w \in F$ of length at most $2|F|$ with $\mathcal{L}^w|s\rangle$ evaluating to $\pm|s\rangle$. Then $X_{s,s} = \frac{1}{nr}\mathbb{1}$.*

Proof. We expand the definition of X to get the value of $X_{s,s}$.

$$\begin{aligned} X_{s,s} &= \operatorname*{avg}_{t\in[n]} \left(\sum \{\rho_{v,w} : v, w \in F,\ \pm|s\rangle = \mathcal{L}^v |t\rangle,\ \pm|s\rangle = \mathcal{L}^w |t\rangle\} \right) \\ &= \frac{1}{n} \sum_{t\in[n]} \sum \{\rho_{v,w} : v, w \in F,\ \pm|s\rangle = \mathcal{L}^v |t\rangle,\ \pm|s\rangle = \mathcal{L}^w |t\rangle\} \end{aligned}$$

If there is no reduced, nonempty word w of length $2|F|$ such that $\mathcal{L}^w|s\rangle$ evaluates to $\pm|s\rangle$, then we know that for any vertex $t \in [n]$, there cannot be two words $v, w \in F$ such that $\pm|s\rangle = \mathcal{L}^v|t\rangle = \mathcal{L}^w|t\rangle$, as this would imply that $\pm|s\rangle = \mathcal{L}^{v^*w}|s\rangle$. Therefore,

$$X_{s,s} = \frac{1}{n} \sum_t \sum \{\rho_{v,v} : v \in F,\ \pm|s\rangle = \mathcal{L}^v |t\rangle\}$$

For every $v \in F$, there is exactly one $t \in [n]$ satisfying $\pm|t\rangle = \mathcal{L}^{v^*}|s\rangle$, or equivalently, $\pm|s\rangle = \mathcal{L}^v|t\rangle$. We can swap the sums and by Equation (6), we have

$$X_{s,s} = \frac{1}{n} \sum_{v\in F} \{\rho_{v,v}\} = \frac{1}{nr}\mathbb{1}$$

□

Lemma 4.3. *Let p be a matrix polynomial of degree $a = d + 2e$ and $\boldsymbol{A}_n$ a randomly signed n-lift of p formed by substitution $\mathcal{L}$. Then for any ℓ, except with probability $1/\log n$, there are at most $O(\log n)$ vertices $v \in [n]$ for which there is a reduced, nonempty word w with $\mathcal{L}^w|v\rangle$ evaluating to $\pm|v\rangle$, where $|w| \le \ell$.*

Proof. We consider the matrix polynomial $p' = \sum_{i=1}^d Y_i + \sum_{i=1}^e Z_i + Z_i^*$, and note that for A'_n formed from p' under the same substitution $\mathcal{L}$, there is a reduced word w with $\pm|v\rangle = \mathcal{L}^w|v\rangle$ if and only if v participates in a cycle in $\boldsymbol{A}'_n$. By [BC19, Lemma 23], the expected number of cycles in $\boldsymbol{A}'_n$ of length at most ℓ is $O(a^\ell)$. Markov's inequality implies that except with probability $1/\log(n)$, the number of cycles in $\boldsymbol{A}_n$ of length at most ℓ is at most $a^\ell \log n$. There are ℓ vertices in such cycles, and a^ℓ is a constant, so the lemma holds. □

4.2 A lower bound on the basic SDP value

In this section we complete the proof of our main theorem by showing that $\textsc{Sdp}(\boldsymbol{A}_n) \ge s^* - \epsilon$ whp.

Lemma 4.4. *For all $\epsilon > 0$, as $n \to \infty$,*

$$\textsc{PartSdp}(A_\infty) - \epsilon \le \textsc{Sdp}(\widetilde{\boldsymbol{A}_n})$$

except with probability $o(1)$.

Proof. Let ρ be a finite rank PSD matrix achieving $\textsc{PartSdp}(A_\infty) = \langle \rho, A_\infty \rangle$ with finite support F, which we know exists from Section 3.

Consider how we defined X_t for every $t \in [n]$ in Section 4.1. Say

$$X = \operatorname*{avg}_{t\in[n]} X_t$$

and note that

$$\langle X, \boldsymbol{A}_n \rangle = \frac{1}{n} \sum_{t \in [n]} \langle X_t, \boldsymbol{A}_n \rangle$$

By Lemma 4.1 and Lemma 4.2, we have that if $t \in [n]$ satisfies the condition that there are no reduced, nonempty words w of length at most $2|F|$ such that $\pm |t\rangle = \mathcal{L}^w |t\rangle$, then $\langle X_t, \boldsymbol{A}_n \rangle = \langle \rho, A_F \rangle$ and $X_{t,t} = \frac{1}{nr}\mathbb{1}$. Lemma 4.3 tells us that except with probability $1/\log n$, there are at most $O(\log n)$ vertices $t \in [n]$ for which this condition doesn't hold.

Note that for any $t \in [n]$ for which this condition doesn't hold, we have that X_t has only constantly many nonzero entries, which are each bounded by constants. Thus, $\langle X_t, \boldsymbol{A}_n \rangle$ differs from $\langle \rho, A_F \rangle$ by at most some constant amount.

Then with high probability,

$$\langle X, \boldsymbol{A}_n \rangle = \frac{1}{n} \sum_{t \in [n]} \langle X_t, \boldsymbol{A}_n \rangle \geq \frac{1}{n} \sum_{t \in [n]} \langle \rho, A_\infty \rangle - O\left(\frac{\log n}{n}\right) = \langle \rho, A_\infty \rangle - o(1)$$

and so we have that X is a PSD matrix such that $\langle X, \boldsymbol{A}_n \rangle$ loses at most an ϵ factor from $\textsc{PartSdp}(A_\infty)$. However, because not all of the diagonal entries have value $\frac{1}{nr}\mathbb{1}$, technically X is not a matrix we would maximize over in $\textsc{Sdp}(\boldsymbol{A}_n)$.

We'll now fix this matrix X to have all $\frac{1}{nr}$ on the diagonals while only changing the value achieved by an ϵ factor.

Say $V = [n]$ denotes the set of vertices of $\boldsymbol{A}_n$, and let S be the set of $s \in [n]$ such that $X_{s,s} = \frac{1}{nr}\mathbb{1}$.

We define a matrix X' such that

$$X'_{a,b} = \begin{cases} X_{a,b} & a, b \in S \\ \frac{1}{nr}\mathbb{1} & a, b \in V \setminus S, a = b \\ 0 & \text{otherwise} \end{cases}$$

X is PSD, so for all $w \in \mathbb{R}^n$ (where the entries of w are vectors of length r), $w^\mathsf{T} X w \geq 0$. In particular, if we consider any $w \in \mathbb{R}^n$ such that $w_i = 0$ when $i \notin S$, it is clear that X restricted to S is PSD as well. For any $w \in \mathbb{R}^n$,

$$w^\mathsf{T} X' w = \sum_{i,j \in [n]} X'_{i,j} w_i w_j = \sum_{i \notin S} \frac{1}{nr} w_i^2 + \sum_{i,j \in S} X_{i,j} w_i w_j \geq 0$$

X' is PSD, and it remains to show that $\langle X', \boldsymbol{A}_n \rangle \geq \langle X, \boldsymbol{A}_n \rangle - \epsilon$. We do this by showing that $\langle X - X', \boldsymbol{A}_n \rangle \leq \epsilon$. We use the below fact to upper bound $\langle X - X', \boldsymbol{A}_n \rangle$.

$$\langle X - X', \boldsymbol{A}_n \rangle \leq \|X - X'\|_1 \|\boldsymbol{A}_n\|_\infty$$

where $\|\boldsymbol{A}_n\|_\infty$ is the spectral norm of $\boldsymbol{A}_n$, or the maximum absolute value of an eigenvalue of A, and $\|X - X'\|_1$ is the sum of the absolute values of the eigenvalues.

Note that $\|\boldsymbol{A}_n\|_\infty$ is at most the maximum absolute value of any entry, and so by repeated applications of the triangle property and the fact that $\|P\| \leq 1$ for P a permutation matrix:

$$\|\boldsymbol{A}_n\| = \left\| \sum_i T_i \otimes a_i \right\| \leq \sum_i \|T_i \otimes a_i\| \leq \sum_i \|a_i\| \prod_{P_j \in T_i} \|P_j\| \leq \sum_i \|a_i\|$$

with T_i denoting the ith term of p. For a fixed p, this value is a constant, and it suffices to bound $\|X - X'\|$.

Let Π be the projection matrix preserving the rows and columns in S, and let $\mathbb{1}_{\overline{S}}$ be the $nr \times nr$ matrix with ones only on the diagonal entries corresponding with $V \setminus S$. We have

$$X' = \Pi X \Pi + \frac{1}{nr}\mathbb{1}_{\overline{S}}$$

where $\left\|\frac{1}{nr}\mathbb{1}_{\overline{S}}\right\| = |V \setminus S|/nr = o(1)$ by Lemma 4.3.

Then,

$$\|X - X'\| = \left\|X - \Pi X \Pi + \frac{1}{nr}\mathbb{1}_{\overline{S}}\right\| \leq \|X - \Pi X \Pi\| + \left\|\frac{1}{nr}\mathbb{1}_{\overline{S}}\right\| = \|X - \Pi X \Pi\| + o(1)$$

We can apply the Gentle Measurement Lemma [Win99, Lem. 9], which says that for PSD matrices Σ with trace 1 and symmetric projection matrices P, $\|\Sigma - P\Sigma P\|_1 \leq c\sqrt{\epsilon}$, where $\epsilon := 1 - \mathrm{tr}(P\Sigma P)$ In our case,

$$\|X - \Pi X \Pi\|_1 \leq c\sqrt{1 - \mathrm{tr}(\Pi X \Pi)}$$

We observe that $\Pi X \Pi$ looks like X' restricted to the rows and columns of S, and so $\mathrm{tr}(\Pi X \Pi) = \sum_{s \in S} X'_{s,s} = \frac{|S|}{nr}$ by definition. Again by Lemma 4.3, as n approaches infinity, $|S| \geq 1 - \log n$ with high probability, and so

$$c\sqrt{1 - \frac{|S|}{nr}} = c\sqrt{\frac{\log n}{nr}} = o(1)$$

This completes the proof that with high probability as n approaches infinity,

$$\langle X - X', \boldsymbol{A}_n \rangle = o(1)$$

The matrix X' therefore achieves an SDP value only $o(1)$ smaller than the original X, which we recall is only $o(1)$ than the value achieved by $\textsc{PartSdp}(A_\infty)$. Since X' will be maximized over in the basic SDP, this clearly shows that the value $\langle X', \boldsymbol{A}_n \rangle$ is achievable by the basic SDP on $\boldsymbol{A}_n$. So, for every $\epsilon > 0$,

$$\textsc{PartSdp}(A_\infty) - \epsilon \leq \langle X', \boldsymbol{A}_n \rangle \leq \textsc{Sdp}(\widetilde{\boldsymbol{A}_n})$$

except with probability at most $1/\log n = o_{n \to \infty}(1)$. □

Our main theorem, Theorem 1.2, now follows as a corollary from our discussion in the beginning of this section, which proves the upper bound on $\textsc{Sdp}(\boldsymbol{A}_n)$, and Lemma 4.4 proving the lower bound on $\textsc{Sdp}(\boldsymbol{A}_n)$.

References

[AA18] Scott Aaronson and Andris Ambainis. Forrelation: a problem that optimally separates quantum from classical computing. *SIAM J. Comput.*, 47(3):982–1038, 2018. 1.1

[BC19] Charles Bordenave and Benoît Collins. Eigenvalues of random lifts and polynomials of random permutation matrices. *Annals of Mathematics*, 190(3):811–875, 2019. 1.1, 1.2, 1.2, 2.3, 4.1

[BGT13] Mohsen Bayati, David Gamarnik, and Prasad Tetali. Combinatorial approach to the interpolation method and scaling limits in sparse random graphs. *Annals of Probability*, 41(6):4080–4115, 2013. 1

[Bop87] Ravi Boppana. Eigenvalues and graph bisection: An average-case analysis. In *Proceedings of the 28th Annual IEEE Symposium on Foundations of Computer Science*, pages 280–285, 1987. 1.1

[Bor20] Charles Bordenave. A new proof of Friedman's second eigenvalue theorem and its extension to random lifts. *Ann. Sci. Éc. Norm. Supér. (4)*, 53(6):1393–1439, 2020. 1.1, 1.1

[BR02] Ola Bratelli and Derek Robinson. *Operator algebras and quantum statistical mechanics.* Springer, 2002. 3.1

[BS00] J. Frédéric Bonnans and Alexander Shapiro. *Perturbation Analysis of Optimization Problems.* Springer Series in Operations Research. Springer-Verlag, New York, 2000. 3, 3.1

[CGHV15] Endre Csóka, Balázs Gerencsér, Viktor Harangi, and Bálint Virág. Invariant Gaussian processes and independent sets on regular graphs of large girth. *Random Structures Algorithms*, 47(2):284–303, 2015. 1.1

[CHSH69] John Clauser, Michael Horne, Abner Shimony, and Richard Holt. Proposed experiment to test local hidden-variable theories. *Physical Review Letters*, 23(15):880–884, 1969. 1.1

[CHTW04] Richard Cleve, Peter Høyer, Benjamin Toner, and John Watrous. Consequences and limits of nonlocal strategies. In *Proceedings of the 19th Annual Computational Complexity Conference*, pages 236–249. IEEE, 2004. 1.1

[Con85] John B. Conway. *The Theory of Subnormal Operators.* American Mathematical Society, 1985. 3.2

[Con90] John B. Conway. *A Course in Functional Analysis.* Springer, 1990. 3.1

[DH73] Wilm Donath and Alan Hoffman. Lower bounds for the partitioning of graphs. *IBM J. Res. Develop.*, 17:420–425, 1973. 1.1

[DLS14] Amit Daniely, Nati Linial, and Shai Shalev-Shwartz. From average case complexity to improper learning complexity. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, pages 441–448, 2014. 1

[DMO+19] Yash Deshpande, Andrea Montanari, Ryan O'Donnell, Tselil Schramm, and Subhabrata Sen. The threshold for sdp-refutation of random regular nae-3sat. In *Proceedings of the 30th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2305–2321. SIAM, 2019. 1.1, 1.1

[DMS+17] Amir Dembo, Andrea Montanari, Subhabrata Sen, et al. Extremal cuts of sparse random graphs. *Annals of Probability*, 45(2):1190–1217, 2017. 1

[DP93] Charles Delorme and Svatopluk Poljak. Laplacian eigenvalues and the maximum cut problem. *Mathematical Programming*, 62:557–574, 1993. 1.1, 1.1, 2.4

[Elo09] Yehonatan Elon. Gaussian waves on the regular tree. Technical Report 0907.5065, arXiv, 2009. 1.1

[FL92] Uriel Feige and László Lovász. Two-prover one-round proof systems: Their power and their problems. In *Proceedings of the 24th Annual ACM Symposium on Theory of Computing*, pages 733–744, 1992. 1

[Fri08] Joel Friedman. A proof of Alon's second eigenvalue conjecture and related problems. *Memoirs of the American Mathematical Society*, 195(910):viii+100, 2008. 1.1

[Fri19] Shmuel Friedland. Infinite dimensional generalizations of Choi's theorem. *Spec. Matrices*, 7:67–77, 2019. 3.1

[GVK21] Jorge Garza-Vargas and Archit Kulkarni. Spectra of infinite graphs via freeness with amalgamation. Technical Report 1912.10137, arXiv, 2021. 1.2

[GW95] Michel Goemans and David Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. Assoc. Comput. Mach.*, 42(6):1115–1145, 1995. 1.1

[HV15] Viktor Harangi and Bálint Virág. Independence ratio and random eigenvectors in transitive graphs. *Ann. Probab.*, 43(5):2810–2840, 2015. 1.1

[JLS20] Aayush Jain, Huijia Lin, and Amit Sahai. Indistinguishability Obfuscation from well-founded assumptions. Technical Report 2008.09317, arXiv, 2020. 1

[Lan17] Klaas Landsman. *Foundations of Quantum Theory*. Springer Open, 2017. 3.1

[Leh99] Franz Lehner. Computing norms of free operators with matrix coefficients. *Amer. J. Math.*, 121(3):453–486, 1999. 1.2

[Mey06] Paul-André Meyer. *Quantum probability for probabilists*. Springer, second edition, 2006. 3.1

[MM09] Marc Mézard and Andrea Montanari. *Information, physics, and computation*. Oxford Graduate Texts. Oxford University Press, Oxford, 2009. 1

[Mon21] Andrea Montanari. Optimization of the Sherrington–Kirkpatrick hamiltonian. *SIAM Journal on Computing*, FOCS19, 2021. 1

[MOP20] Sidhanth Mohanty, Ryan O'Donnell, and Pedro Paredes. The SDP value for random two-eigenvalue CSPs. In Christophe Paul and Markus Bläser, editors, *Proceedings of the 37th Annual Symposium on Theoretical Aspects of Computer Science*, volume 154 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 50:1–50:45, Dagstuhl, Germany, 2020. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. 1.1

[MS16] Andrea Montanari and Subhabrata Sen. Semidefinite programs on sparse random graphs and their application to community detection. In *Proceedings of the 48th Annual ACM Symposium on Theory of Computing*, pages 814–827. ACM, New York, 2016. 1.1, 1.1

[NC10] Michael Nielsen and Isaac Chung. *Quantum Computation and Quantum Information*. Cambridge University Press, 10th anniversary edition edition, 2010. 3.2

[OW20] Ryan O'Donnell and Xinyu Wu. Explicit near-fully X-Ramanujan graphs. In *Proceedings of the 61st Annual IEEE Symposium on Foundations of Computer Science*, pages 1045–1056, 2020. 1.1, 1.2, 1.2, 2.3

[PR95] Svatopluk Poljak and Franz Rendl. Nonpolyhedral relaxations of graph-bisection problems. *SIAM Journal on Optimization*, 5(3):467–487, 1995. 1.1, 2.4

[Rag09] Prasad Raghavendra. *Approximating NP-hard problems: efficient algorithms and their limits*. PhD thesis, University of Washington, 2009. 1

[RS80] Michael Reed and Barry Simon. *Functional Analysis*. Academic Press, 1980. 3.1

[RW95] Franz Rendl and Henry Wolkowicz. A projection technique for partitioning the nodes of a graph. *Ann. Oper. Res.*, 58:155–179, 1995. Applied mathematical programming and modeling, II (APMOD 93) (Budapest, 1993). 1.1

[SH08] Maksim Shirokov and Alexander Holevo. On the approximation of infinite-dimensional quantum channels. *Problemy Peredachi Informatsii*, 44(2):3–22, 2008. 3.1

[Sty73] George Styan. Hadamard products and multivariate statistical analysis. *Linear Algebra Appl.*, 6:217–240, 1973. 2.21

[vDE20] Anne van Delft and Michael Eichler. A note on Herglotz's theorem for time series on function spaces. *Stochastic Process. Appl.*, 130(6):3687–3710, 2020. 3.1

[Win99] Andreas Winter. Coding theorem and strong converse for quantum channels. *IEEE Trans. Inform. Theory*, 45(7):2481–2485, 1999. 4.2