# GENERALIZED CUTS AND GROTHENDIECK COVERS: A PRIMAL-DUAL APPROXIMATION FRAMEWORK EXTENDING THE GOEMANS–WILLIAMSON ALGORITHM

NATHAN BENEDETTO PROENÇA[1¶], MARCEL K. DE CARLI SILVA[2*], CRISTIANE M. SATO[3*], AND LEVENT TUNÇEL[1¶]

Abstract. We provide a primal-dual framework for randomized approximation algorithms utilizing semidefinite programming (SDP) relaxations. Our framework pairs a continuum of APX-complete problems including MaxCut, Max2Sat, MaxDicut, and more generally, Max-Boolean Constraint Satisfaction and MaxQ (maximization of a positive semidefinite quadratic form over the hypercube) with new APX-complete problems which are stated as convex optimization problems with exponentially many variables. These new dual counterparts, based on what we call *Grothendieck covers*, range from fractional cut covering problems (for MaxCut) to tensor sign covering problems (for MaxQ). For each of these problem pairs, our framework transforms the randomized approximation algorithms with the best known approximation factors for the primal problems to randomized approximation algorithms for their dual counterparts with reciprocal approximation factors which are tight with respect to the Unique Games Conjecture. For each APX-complete pair, our algorithms solve a single SDP relaxation and generate feasible solutions for both problems which also provide approximate optimality certificates for each other. Our work utilizes techniques from areas of randomized approximation algorithms, convex optimization, spectral sparsification, as well as Chernoff-type concentration results for random matrices.



---

[1]Department of Combinatorics and Optimization, University of Waterloo
[2]Institute of Mathematics and Statistics, University of São Paulo
[3]Center for Mathematics, Computing and Cognition, Federal University of the ABC Region
[¶]Research of this author was supported in part by a Discovery Grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada .
[*]This work was partially supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

## 1. Introduction

Some of the most impressive successes for randomized approximation algorithms, utilizing semidefinite programming relaxations, have been on problems such as MaxCut [GW95], Max2Sat [LLZ02], and MaxDicut [Bra+23]. We define APX-complete duals for such problems, which involve what we call *Grothendieck covers*. Then, we design a primal-dual framework of randomized approximation algorithms for a wide range of problems, including maximum Boolean constraint satisfaction problems (CSPs) paired with their APX-complete duals, which we call *Boolean CSP covering problems*. Our focus is on 2-CSPs, where each constraint has at most 2 literals; this includes the MaxCut, the Max2Sat, and the MaxDicut problems. For each of these APX-complete problems, our framework transforms the randomized approximation algorithms for the primal problem to randomized approximation algorithms for their (also APX-complete) duals while preserving the approximation factor. In particular, it allows us to recover the same best known approximation factors for the new problems. For example, we provide a randomized (1/0.874)-approximation algorithm for weighted fractional dicut covers. Although the new problems have exponentially many variables, the covers produced have small support and their approximation quality relies on symmetric Grothendieck inequalities; see [FL20]. Our algorithms and analyses utilize Chernoff-type concentration results and spectral sparsification.

We further describe how each APX-complete instance can be paired with a dual APX-complete instance by solving a single semidefinite program, unlike in usual scenarios where the dual is built syntactically from the primal. The SDP solutions yield, via a randomized sampling algorithm, primal and dual feasible solutions along with a simultaneous certificate of the approximation quality of both solutions. Note that such a certificate has two primal-dual pairs involved: one pair intractable, and the other pair tractable. E.g., (i) MaxDicut and weighted fractional dicut cover, (ii) the SDP relaxation of MaxDicut and its SDP dual.

Let $D = (V, A)$ be a digraph. For each $U \subseteq V$, define $\delta^{\text{out}}(U)$ as the set of arcs leaving $U$. A *dicut* is the set $\delta^{\text{out}}(U)$ for some $U \subseteq V$. For arc weights $w \in \mathbb{R}_+^A$, the *maximum dicut number of* $(D, w)$ is

$$\text{md}(D, w) := \max\{\, w^\mathsf{T} \mathbb{1}_{\delta^{\text{out}}(U)} : U \in \mathcal{P}(V)\},$$

where $\mathbb{1}_{\delta^{\text{out}}(U)} \in \{0,1\}^A$ is the incidence vector of $\delta^{\text{out}}(U)$ and $\mathcal{P}(V)$ denotes the power set of $V$. The vector of all-ones is $\mathbb{1}$. The dual problem we consider is fractionally covering the arcs by dicuts: for arc weights $z \in \mathbb{R}_+^A$, the *fractional dicut-covering number of* $(D, z)$ is

$$\text{fdc}(D, z) := \min\left\{ \mathbb{1}^\mathsf{T} y : y \in \mathbb{R}_+^{\mathcal{P}(V)},\ \sum_{U \subseteq V} y_U \mathbb{1}_{\delta^{\text{out}}(U)} \geq z \right\}.$$

[Bra+23] obtained a randomized $\alpha_{\text{BHPZ}}$-approximation for the maximum dicut problem, where $\alpha_{\text{BHPZ}} \approx 0.87446$. Our framework yields the following result.

**Theorem 1** (Fractional Dicut Covering Theorem)**.** Fix $\beta \in (0, \alpha_{\text{BHPZ}})$. There is a randomized polynomial-time algorithm that, given a digraph $D = (V, A)$ and $z \in \mathbb{R}_+^A$, computes $w \in \mathbb{R}_+^A$ and returns $U \subseteq V$ and $y \in \mathbb{R}_+^{\mathcal{P}(V)}$ with support size $|\text{supp}(y)| = O(\log|V|)$ such that $\sum_{S \subseteq V} y_S \mathbb{1}_{\delta^{\text{out}}(S)} \geq z$ holds with high probability (w.h.p.),

$$\mathbb{1}^\mathsf{T} y \leq \tfrac{1}{\beta}\, \text{fdc}(D, z), \qquad \text{and} \qquad w^\mathsf{T} \mathbb{1}_{\delta^{\text{out}}(U)} \geq \beta\, \text{md}(D, w).$$

Moreover, our algorithm returns a simultaneous certificate that each of $U$ and $y$ is within a factor of $\beta$ of the respective optimal value.

**Remark 2** (Primal-Dual Symmetry)**.** Our results also allow one to start from an instance $(D, w)$ of the primal problem (i.e., MaxDicut) and the algorithm computes a dual instance $(D, z)$ of the fractional dicut-covering problem, along with $\beta$-approximate solutions for both and a simultaneous certificate. Analogous claims also apply to Theorems 3 and 4.

Let $(\mathscr{C}, w)$ be an instance of the *maximum 2-satisfiability problem*, i.e., $\mathscr{C}$ is a set of disjunctive 2-clauses on two variables from $x_1, \ldots, x_n$, and $w \in \mathbb{R}_+^{\mathscr{C}}$ is a nonnegative weight vector. Thus, each element of $\mathscr{C}$ has the form $x_i \vee x_j$, $x_i \vee \overline{x_j}$, or $\overline{x_i} \vee \overline{x_j}$. Let $\mathscr{A} \coloneqq \{\texttt{false}, \texttt{true}\}^n$ be the set of all possible assignments for $(x_1, \ldots, x_n)$. For an assignment $a \in \mathscr{A}$, define $\mathrm{val}_{\mathscr{C}}(a) \in \{0,1\}^{\mathscr{C}}$ as the binary vector indexed by $\mathscr{C}$ such that $(\mathrm{val}_{\mathscr{C}}(a))_C = 1$ if $C$ is satisfied by $a$, and 0 otherwise. The goal is to find an assignment $a \in \mathscr{A}$ that maximizes the inner product $\langle w, \mathrm{val}_{\mathscr{C}}(a) \rangle$. Denote

$$\mathrm{max2sat}(\mathscr{C}, w) \coloneqq \max\{\, \langle w, \mathrm{val}_{\mathscr{C}}(a) \rangle : a \in \mathscr{A} \}.$$

The dual problem we consider is fractionally covering the clauses by assignments: for weights $z \in \mathbb{R}_+^{\mathscr{C}}$, the *fractional 2-sat covering number of* $(\mathscr{C}, z)$ is

$$\text{frac-2sat-cover}(\mathscr{C}, z) \coloneqq \min\Big\{\, \mathbb{1}^\mathsf{T} y : y \in \mathbb{R}_+^{\mathscr{A}},\ \sum_{a \in \mathscr{A}} y_a \,\mathrm{val}_{\mathscr{C}}(a) \geq z \Big\}.$$

[LLZ02] provide a randomized $\alpha_{\text{LLZ}}$-approximation algorithm for Max2Sat, where $\alpha_{\text{LLZ}} \approx 0.9401$. Our framework yields the following result.

**Theorem 3** (Fractional 2-Sat Covering Theorem)**.** Fix $\beta \in (0, \alpha_{\text{LLZ}})$. There is a randomized polynomial-time algorithm that, given a set $\mathscr{C}$ of disjunctive 2-clauses on $n$ variables and $z \in \mathbb{R}_+^{\mathscr{C}}$, computes $w \in \mathbb{R}_+^{\mathscr{C}}$ and returns an assignment $a \in \mathscr{A}$ and $y \in \mathbb{R}_+^{\mathscr{A}}$ with $|\mathrm{supp}(y)| = O(\log n)$ such that $\sum_{a \in \mathscr{A}} y_a \,\mathrm{val}_{\mathscr{C}}(a) \geq z$ holds w.h.p.,

$$\mathbb{1}^\mathsf{T} y \leq \tfrac{1}{\beta}\, \text{frac-2sat-cover}(\mathscr{C}, z), \qquad \text{and} \qquad \langle w, \mathrm{val}_{\mathscr{C}}(a) \rangle \geq \beta \,\mathrm{max2sat}(\mathscr{C}, w).$$

Moreover, our algorithm returns a simultaneous certificate that each of $a$ and $y$ is within a factor of $\beta$ of the respective optimal value.

Our results are general enough to include all forms of Boolean 2-CSPs. A Boolean 2-CSP is a CSP where the variables $x_1, \ldots, x_n$ take on Boolean values (i.e., `true` or `false`) and each constraint involves only two variables. Formally, we specify a Boolean constraint satisfaction problem using a set $\mathfrak{P}$ of binary predicate templates, i.e., functions from $\{\texttt{false}, \texttt{true}\}^2$ to $\{\texttt{false}, \texttt{true}\}$. We assume throughout that the constant `false` function is not in $\mathfrak{P}$. Let $(\mathscr{C}, w)$ be an instance of the *(Boolean) maximum 2-CSP problem*, i.e., each element of $\mathscr{C}$ is a function that sends $x \in \{\texttt{false}, \texttt{true}\}^n$ to $f(x_i, x_j)$ for some $f \in \mathfrak{P}$ and $i, j \in [n] \coloneqq \{1, \ldots, n\}$, and $w \in \mathbb{R}_+^{\mathscr{C}}$. We refer to an element of $\mathscr{C}$ as a $\mathfrak{P}$-*constraint* or just as a *constraint*. The *maximum* $\mathfrak{P}$-*satisfiability number* of $(\mathscr{C}, w)$ is

$$\text{max-}\mathfrak{P}\text{-sat}(\mathscr{C}, w) \coloneqq \max\{\, \langle w, \mathrm{val}_{\mathscr{C}}(a) \rangle : a \in \mathscr{A} \}.$$

The dual problem is: for every $z \in \mathbb{R}_+^{\mathscr{C}}$, the *fractional* $\mathfrak{P}$-*constraint covering number* of $(\mathscr{C}, z)$ is

$$\text{frac-}\mathfrak{P}\text{-cover}(\mathscr{C}, z) \coloneqq \min\Big\{\, \mathbb{1}^\mathsf{T} y : y \in \mathbb{R}_+^{\mathscr{A}},\ \sum_{a \in \mathscr{A}} y_a \,\mathrm{val}_{\mathscr{C}}(a) \geq z \Big\}.$$

By choosing distinct sets $\mathfrak{P}$ one can formulate various interesting problems. By setting $\mathfrak{P} \coloneqq \{\overline{x_1} \vee x_2\}$, we recover the MaxDicut problem via max-$\mathfrak{P}$-sat and the fractional dicut-covering problem via frac-$\mathfrak{P}$-cover. Our Max2Sat results are recovered with $\mathfrak{P} \coloneqq \{x_1 \vee x_2, \overline{x_1} \vee x_2, \overline{x_1} \vee \overline{x_2}\}$. Using these choices, Theorems 1 and 3 are special cases a more general result from our framework, which we state next. The approximation factor $\alpha_{\Xi}$ that appears in the statement will be defined shortly in (4); a self-contained version of the result will be stated later as Theorem 17.

**Theorem 4** (Fractional $\mathfrak{P}$-Covering Theorem)**.** Let $\mathfrak{P}$ be a set of predicates in two Boolean variables. Fix $\beta \in (0, \alpha_{\Xi})$. There is a randomized polynomial-time algorithm that, given a set $\mathscr{C}$ of $\mathfrak{P}$-constraints on $n$ variables and $z \in \mathbb{R}_+^{\mathscr{C}}$, computes $w \in \mathbb{R}_+^{\mathscr{C}}$ and returns an assignment $a \in \mathscr{A}$ and $y \in \mathbb{R}_+^{\mathscr{A}}$ with $|\mathrm{supp}(y)| = O(\log n)$ such that $\sum_{a \in \mathscr{A}} y_a \,\mathrm{val}_{\mathscr{C}}(a) \geq z$ holds w.h.p.,

$$\mathbb{1}^\mathsf{T} y \leq \tfrac{1}{\beta}\, \text{frac-}\mathfrak{P}\text{-cover}(\mathscr{C}, z), \qquad \text{and} \qquad \langle w, \mathrm{val}_{\mathscr{C}}(a) \rangle \geq \beta \,\text{max-}\mathfrak{P}\text{-sat}(\mathscr{C}, w).$$

Moreover, our algorithm returns a simultaneous certificate that each of $a$ and $y$ is within a factor of $\beta$ of the respective optimal value.

Our framework builds on works by Goemans and Williamson [GW95], Grothendieck [Gro53], and Nesterov [Nes98], involving approximation results. [GW95; Nes98] both tackle the problem of solving $\max\{\, s_U^{\mathsf{T}} W s_U : U \in \mathcal{P}(V)\}$, where the $V \times V$ matrix $W$ belongs to the positive semidefinite cone $\mathbb{S}_+^V$ and $s_U := 2\mathbb{1}_U - \mathbb{1} \in \{\pm 1\}^V$ is the signed incidence vector of $U \subseteq V$. We introduce a parameterization for both the domain cone of matrices $W$ and the allowed subsets $U$ of $V$. Throughout $V$ denotes a finite set and let $\mathbb{D}, \mathbb{K} \subseteq \mathbb{S}^V$ be closed convex cones, where $\mathbb{S}^V$ is the space of symmetric $V$-by-$V$ matrices. Let $\mathcal{F}(\mathbb{D}) := \{\, U \subseteq V : s_U s_U^{\mathsf{T}} \in \mathbb{D}\}$ encode the feasible/allowed subsets of $V$. Our primal problem involves maximization of a quadratic form:

$$\text{(1)} \qquad \operatorname{maxq}_{\mathbb{D},\mathbb{K}}(W) := \max\{\, s_U^{\mathsf{T}} W s_U : U \in \mathcal{F}(\mathbb{D})\}, \qquad \text{for every } W \in \mathbb{K}.$$

Let $\operatorname{aff}(\mathbb{K})$ denote the smallest affine space containing $\mathbb{K}$. For each $Z$ in the dual cone $\mathbb{K}^* := \{\, X \in \operatorname{aff}(\mathbb{K}) : \langle X, Y\rangle \geq 0 \text{ for each } Y \in \mathbb{K}\}$ (where we use the trace inner product), a vector $y \in \mathbb{R}_+^{\mathcal{F}(\mathbb{D})}$ is a *tensor sign cover for* $Z$ if $\sum_{U \in \mathcal{F}(\mathbb{D})} y_U s_U s_U^{\mathsf{T}} \succeq_{\widehat{\mathbb{K}^*}} Z$, where as usual the notation $A \succeq_{\widehat{\mathbb{K}^*}} B$ means $A - B \in \widehat{\mathbb{K}^*}$. Here, we are denoting by $\widehat{\mathbb{K}^*} := \{\, X \in \mathbb{S}^n : \langle X, Y\rangle \geq 0 \text{ for each } Y \in \mathbb{K}\}$ the dual cone to $\mathbb{K}$ in the potentially larger space of symmetric matrices $\mathbb{S}^n$ — see Appendix A. Our dual problem is to find a tensor sign cover $y$ that minimizes $\mathbb{1}^{\mathsf{T}} y$:

$$\text{(2)} \qquad \operatorname{fevc}_{\mathbb{D},\mathbb{K}}(Z) := \min\left\{\, \mathbb{1}^{\mathsf{T}} y : y \in \mathbb{R}_+^{\mathcal{F}(\mathbb{D})}, \sum_{U \in \mathcal{F}(\mathbb{D})} y_U s_U s_U^{\mathsf{T}} \succeq_{\widehat{\mathbb{K}^*}} Z \right\}, \qquad \text{for every } Z \in \mathbb{K}^*.$$

The notation 'fevc' refers to fractional elliptope vertex cover. Recall that the elliptope is the set $\mathcal{E}^V := \{\, Y \in \mathbb{S}_+^V : \operatorname{diag}(Y) = \mathbb{1}\}$, where $\operatorname{diag} \colon \mathbb{S}^V \to \mathbb{R}^V$ extracts the diagonal, and its vertices are $\{\, s_U s_U^{\mathsf{T}} : U \in \mathcal{P}(V)\}$; see [LP95]. By fixing $\mathbb{D}$ and varying $\mathbb{K}$, it is clear that $\operatorname{maxq}_{\mathbb{D},\mathbb{K}}$ always attributes the same value for an input matrix, whereas $\operatorname{fevc}_{\mathbb{D},\mathbb{K}}$ defines a continuum of relaxations, affecting feasibility via the constraint $\sum_{U \in \mathcal{F}(\mathbb{D})} y_U s_U s_U^{\mathsf{T}} \succeq_{\widehat{\mathbb{K}^*}} Z$ on the tensor sign covers. The smaller $\mathbb{K}$ is, the weaker the constraint on the tensor sign cover becomes.

Theorems 1 and 3 describe SDP-based approximation algorithms for fractional covering problems. Covering problems, in general, proved to be difficult for tractable SDP relaxations. For some negative results on various SDP relaxations of vertex cover problem, see for instance [Ben+11; Geo+10; KG98]. In those settings, the SDP relaxations considered fail to improve on their much simpler LP-based counterparts, in terms of the approximation ratio. Thus, it is noteworthy that in our framework we obtain randomized approximation algorithms that are tight under the UGC. Another interesting feature of our results is that our conic covering problems have an exponential number of variables (and computing their optimal values is NP-hard) but we still are able to treat these covering problems algorithmically, in polynomial time, and obtain approximately optimal sparse covers.

Throughout the paper, we assume that $\mathbb{D}, \mathbb{K} \subseteq \mathbb{S}^n := \mathbb{S}^{[n]}$ are closed convex cones such that the following conditions hold:

$$\text{(3)} \qquad \begin{gathered} \mathbb{D} \subseteq \mathbb{S}_+^n, \qquad \mathbb{K} \subseteq \mathbb{D}^*, \qquad \operatorname{int}(\operatorname{cone}(\mathrm{CUT}^{\mathbb{D}})) \neq \varnothing, \\ \{0\} \neq \mathbb{K} \text{ has a strictly feasible point,} \end{gathered}$$

where $\mathrm{CUT}^{\mathbb{D}} := \operatorname{conv}\{\, s_U s_U^{\mathsf{T}} : U \subseteq [n],\ s_U s_U^{\mathsf{T}} \subseteq \mathbb{D}\}$, conv is the convex hull, cone denotes the generated convex cone containing 0, and int takes the interior. We refer the reader to Appendix A for the definition of strictly feasible point. Set $\mathcal{E}(\mathbb{D}) := \mathcal{E}^{[n]} \cap \mathbb{D}$. A *randomized rounding algorithm* $\Xi$ *for* $\mathbb{D}$ is an indexed set $\Xi = (\Xi_Y)_{Y \in \mathcal{E}(\mathbb{D})}$ of matrix-valued random variables sampled from $\{\, s_U s_U^{\mathsf{T}} : U \in \mathcal{F}(\mathbb{D})\}$. Define

$$\text{(4)} \qquad \alpha_{\mathbb{D},\mathbb{K},\Xi} := \inf_{Y \in \mathcal{E}(\mathbb{D})} \max\{\, \alpha \in \mathbb{R}_+ : \mathbb{E}[\Xi_Y] \succeq_{\widehat{\mathbb{K}^*}} \alpha Y\},$$

which we call the *rounding constant for* $(\mathbb{D}, \mathbb{K}, \Xi)$. We shall drop the pair $\mathbb{D}, \mathbb{K}$ whenever they can be inferred by context; in particular, the rounding constant $\alpha_{\mathbb{D},\mathbb{K},\Xi}$ may appear as $\alpha_\Xi$. We define a *Grothendieck cover* for $Z \in \mathbb{K}^*$ as a tensor sign cover $y$ for $Z$ such that $\mathbb{1}^\mathsf{T} y \leq (1/\alpha_\Xi)\,\mathrm{fevc}(Z)$. Our algorithms produce tensor sign covers $y$ with approximation factor $\beta$ arbitrarily close to $\alpha_\Xi$; we also call such vectors Grothendieck covers.

We show how to pair instances of the problems maxq and fevc so that, given an instance $W \in \mathbb{K}$ of maxq, we obtain an instance $Z \in \mathbb{K}^*$ of fevc and we approximately solve both instances simultaneously and provide a certificate for the approximation factor of both solutions. We do the same by starting with an instance of fevc. Note from (2) that feasible solutions can have exponential support size. The solutions produced by our algorithm have sparse support, with the bound on the support size varying according to geometric properties of the cone $\mathbb{K}$. For the case that $\mathbb{K} \subseteq \mathbb{S}^n_+$, we rely on spectral sparsification results for positive semidefinite matrices from [BSS12; CHS15].

Our main results are the outcome of our framework powered by primal-dual conic relaxations, randomized rounding algorithms together with generalized Chernoff concentration results, and spectral sparsifications methods. We state our main results in Theorem 5 and Remark 6. They output objects called $\beta$-certificates (see Definition 9), where $\beta$ is an approximation factor, which are formed by feasible solutions for both problems, together with a simultaneous certificate of their approximation quality.

**Theorem 5** (Main Semidefinite Theorem)**.** Assume that $\mathbb{K} \subseteq \mathbb{S}^n_+$, and let $\Xi$ be a randomized rounding algorithm for $\mathbb{D}$. Fix $\beta \in (0, \alpha_\Xi)$. There exists a randomized polynomial-time algorithm that, given an instance $Z \in \mathbb{K}^*$ of fevc as input, computes an instance $W \in \mathbb{K}$ of maxq and a $\beta$-certificate for $(W, Z)$ with high probability. Dually, there exists a randomized polynomial-time algorithm that, given an instance $W \in \mathbb{K}$ of maxq as input, computes an instance $Z \in \mathbb{K}^*$ of fevc and a $\beta$-certificate for $(W, Z)$ with high probability. Both algorithms take at most $O(n^2 \log(n))$ samples from $\Xi$, and produce covers with $O(n)$ support. If $\mathbb{K} = \mathbb{S}^n_+$, then $O(n \log n)$ samples suffice.

**Remark 6** (Main Polyhedral Theorem)**.** We state in Theorem 16 our other main result, which is similar to Theorem 5, however, with a slightly different assumption on the cone $\mathbb{K}$ and it obtains better support size. The cone $\mathbb{K}$ is the image $\mathcal{A}(\mathbb{R}^d_+)$ for a linear map $\mathcal{A}\colon \mathbb{R}^d \to \mathbb{S}^n$, and the support size obtained is $O(\log(n) + \log(d))$. Theorem 4 shall follow from this result.

**Additional Related Work.** In addition to the above cited references, here we mention some additional related work. In the continuum of the APX-complete duals, the one for MaxCut, called fractional cut-covering problem, was previously studied: first, in the special case that $z = \mathbb{1}$, i.e. unweighted graphs, see [NB19; Šám06]; then, in general (arbitrary nonnegative weights $z$), see [Ben+23]. We vastly generalize the results of [Ben+23] while keeping all the desired properties. Their results apply to the pair MaxCut and fractional cut covering, which is a single pair of APX-complete problems in the wide swath of APX-complete problem pairs covered here.

Part of the unification and generalization of the primal problems we consider was proposed earlier [FL20]. Their generalization is similar to the way we use the convex cone $\mathbb{K}$ and the generalized Grothendieck constant. However, our framework is more general than that of [FL20] in two ways: (i) we consider, as an additional generalization, a set of convex cones $\mathbb{D}$ restricting the feasible region of the primal problem (this additional generalization helps us achieve the best approximation ratios for the duals of Max Boolean 2-CSPs); (ii) for every primal APX-complete problem in our generalized domain we associate a dual conic covering problem and provide randomized approximation algorithms which provide approximate solutions to both problems.

Part of our development of the underlying theory leading to the APX-complete duals is best explained via gauge duality [Fre87] and its interplay with conic duality. A closely related concept is antiblocking duality theory [Ful71; Ful72]. The corresponding conic generalization of antiblocking duality appeared previously in [Tin74].

## 2. Framework for Generalized Cuts and Tensor Sign Covers and Certificates

This section introduces our framework along with its theoretical foundations. Recall the assumptions (3). We define relaxations for $\mathrm{maxq}_{\mathbb{D},\mathbb{K}}$ and $\mathrm{fevc}_{\mathbb{D},\mathbb{K}}$: for every $W \in \mathbb{K}$, set

$$\nu_{\mathbb{D},\mathbb{K}}(W) \coloneqq \max\{\, \langle W, Y\rangle : Y \in \mathbb{D},\, \mathrm{diag}(Y) = \mathbb{1}\} \tag{5a}$$

$$= \min\{\, \rho : \rho \in \mathbb{R}_+,\, x \in \mathbb{R}^n,\, \mathrm{Diag}(x) \succeq_{\mathbb{D}^*} W,\, \rho \geq \mathbb{1}^\mathsf{T} x\}, \tag{5b}$$

and, for every $Z \in \mathbb{K}^*$,

$$\nu^\circ_{\mathbb{D},\mathbb{K}}(Z) \coloneqq \min\{\, \mu : \mu \in \mathbb{R}_+,\, Y \in \mathbb{D},\, \mathrm{diag}(Y) = \mu\mathbb{1},\, Y \succeq_{\widehat{\mathbb{K}^*}} Z\} \tag{6a}$$

$$= \max\{\, \langle W, Z\rangle : W \in \mathbb{K},\, x \in \mathbb{R}^n,\, W \preceq_{\mathbb{D}^*} \mathrm{Diag}(x),\, \mathbb{1}^\mathsf{T} x \leq 1\}. \tag{6b}$$

Our algorithms rely on solving these relaxations and then sampling using the feasible solutions found. We show the following relation between maxq, fevc, $\nu$, and $\nu^\circ$, and the rounding constant $\alpha_\Xi$:

**Theorem 7.** Let $\Xi$ be a randomized rounding algorithm for $\mathbb{D}$. We have that

$$\alpha_\Xi \cdot \nu(W) \leq \mathrm{maxq}(W) \leq \nu(W) \qquad \text{for every } W \in \mathbb{K}; \tag{7}$$

$$\nu^\circ(Z) \leq \ \mathrm{fevc}(Z) \ \leq \tfrac{1}{\alpha_\Xi} \cdot \nu^\circ(Z) \qquad \text{for every } Z \in \mathbb{K}^*. \tag{8}$$

*Proof.* Note that $s_U^\mathsf{T} W s_U = \langle W, s_U s_U^\mathsf{T}\rangle \leq \nu(W)$ for every $U \subseteq [n]$, so the second inequality in (7) holds. Let $Y$ be a feasible solution of (5a). Since $\Xi_Y$ has finite support, we have that $\mathbb{E}[\Xi_Y] = \sum_{U\subseteq[n]} \mathbb{P}(\Xi_Y = s_U s_U^\mathsf{T}) s_U s_U^\mathsf{T}$, which implies

$$\mathbb{E}[\Xi_Y] \in \mathrm{conv}(\{\, s_U s_U^\mathsf{T} : U \in \mathcal{F}(\mathbb{D})\}) = \mathrm{CUT}^{\mathbb{D}}\,. \tag{9}$$

Thus (7) follows from (5a), as $\mathrm{maxq}(W) \geq \langle W, \mathbb{E}[\Xi_Y]\rangle \geq \alpha_\Xi \langle W, Y\rangle$. Similarly, for every $y \in \mathbb{R}_+^{\mathcal{F}(\mathbb{D})}$ feasible in (2), we have that $\sum_{S\in\mathcal{F}(\mathbb{D})} y_U s_U s_U^\mathsf{T} \in \mathbb{D}$ is feasible in (6a) with the same objective value, so the first inequality in (8) holds. It is immediate from (4) that for every $Y \in \mathcal{E}(\mathbb{D})$, if $\mu Y \succeq_{\widehat{\mathbb{K}^*}} Z$, then $\mu\mathbb{E}[\Xi_Y] \succeq_{\widehat{\mathbb{K}^*}} \alpha_\Xi Z$, so $\mathrm{fevc}(Z) \leq \mu/\alpha_\Xi$. Hence (8) follows from (6a). ☐

We remark that (7) and (8) are equivalent by gauge duality; see Appendix A for more details.

Our discussion so far has focused exclusively on the matrix space. Indeed, the definition (4) of $\alpha_\Xi$, as well as the concentration results we will exploit are naturally expressed in this context. Yet, applications may require results on other spaces. For example, an approximation algorithm for the fractional dicut covering problem on a digraph $D = (V, A)$ is about weights in $\mathbb{R}_+^A$. In our setting, this mapping between vectors and matrices is built into the cone $\mathbb{K}$. This is natural, as the cone $\mathbb{K}$ is central to the covering constraint of (2). Let $\mathcal{A}\colon \mathbb{R}^d \to \mathbb{S}^n$ be a linear map. We assume throughout the paper that $\mathcal{A}(w) \coloneqq \sum_{i\in[d]} w_i A_i$ for nonzero $A_1, \dots, A_d \in \mathbb{S}^n$. Set $\mathbb{K} \coloneqq \mathcal{A}(\mathbb{R}_+^d) = \{\, \sum_{i\in[d]} w_i A_i : w \in \mathbb{R}_+^d\}$. We have that, for every $X, Y \in \mathbb{S}^n$,

$$X \preceq_{\widehat{\mathbb{K}^*}} Y \text{ if and only if } \langle A_i, X\rangle \leq \langle A_i, Y\rangle \text{ for every } i \in [d]. \tag{10}$$

One can succinctly encode the finite set of linear inequalities above with the adjoint linear map $\mathcal{A}^*\colon \mathbb{S}^n \to \mathbb{R}^d$, thus obtaining that $X \preceq_{\widehat{\mathbb{K}^*}} Y$ holds if and only if $\mathcal{A}^*(X) \leq \mathcal{A}^*(Y)$. This is similar to what is done in the entropy maximization setting; see, e.g., [SV14]. In particular, the linear map $\mathcal{A}^*$ recovers relevant marginal probabilities when working with random matrices. With this setup, we move to $\mathbb{R}_+^d$ by defining, for every $w \in \mathbb{R}_+^d$ and $z \in \mathbb{R}_+^d$,

$$\mathrm{maxq}_{\mathbb{D},\mathcal{A}}(w) \coloneqq \mathrm{maxq}_{\mathbb{D},\mathbb{K}}(\mathcal{A}(w)), \qquad \mathrm{fevc}_{\mathbb{D},\mathcal{A}}(z) \coloneqq \min\{\, \mathrm{fevc}_{\mathbb{D},\mathbb{K}}(Z) : Z \in \mathbb{K}^*,\, \mathcal{A}^*(Z) \geq z\},$$
$$\nu_{\mathbb{D},\mathcal{A}}(w) \coloneqq \nu_{\mathbb{D},\mathbb{K}}(\mathcal{A}(w)), \qquad \nu^\circ_{\mathbb{D},\mathcal{A}}(z) \coloneqq \min\{\, \nu^\circ_{\mathbb{D},\mathbb{K}}(Z) : Z \in \mathbb{K}^*,\, \mathcal{A}^*(Z) \geq z\};$$

see Theorem 22 in Appendix B for details. We highlight that

$$\nu^\circ_{\mathbb{D},\mathcal{A}}(z) = \max\{\, z^\mathsf{T} w : w \in \mathbb{R}_+^d,\, x \in \mathbb{R}^n,\, \mathcal{A}(w) \preceq_{\mathbb{D}^*} \mathrm{Diag}(x),\, \mathbb{1}^\mathsf{T} x \leq 1\} \tag{11a}$$

$$= \min\{\, \mu : \mu \in \mathbb{R}_+,\, Y \in \mathbb{D},\, \operatorname{diag}(Y) = \mu\mathbb{1},\, \mathcal{A}^*(Y) \geq z\}. \tag{11b}$$

Hence, given $z \in \mathbb{R}^d_+$ as input, one can compute $Z \in \mathbb{K}^*$ such that $\nu^\circ_{\mathbb{D},\mathcal{A}}(z) = \nu^\circ_{\mathbb{D},\mathbb{K}}(Z)$, as well as solving (6a) for said $Z$, by solving a single convex optimization problem, namely (11b). In this way, we can both "lift" the vector $w \in \mathbb{R}^d_+$ to the matrix $\mathcal{A}(w) \in \mathbb{K}$, and $z \in \mathbb{R}^d_+$ to the matrix $Z \in \mathbb{K}^*$, with no extra algorithmic cost.

Next we discuss how to simultaneously certify the approximation quality for instances $W \in \mathbb{K}$ of maxq and $Z \in \mathbb{K}^*$ of fevc. A key observation is that

$$\operatorname{maxq}(W) \cdot \operatorname{fevc}(Z) \geq \langle W, Z\rangle \qquad \text{for every } W \in \mathbb{K} \text{ and } Z \in \mathbb{K}^*, \tag{12}$$

which holds since $\langle W, Z\rangle \leq \sum_{U\in\mathcal{F}(\mathbb{D})} y_U \langle W, s_U s_U^\mathsf{T}\rangle \leq \operatorname{maxq}(W)\mathbb{1}^\mathsf{T} y$ for every feasible solution $y \in \mathbb{R}^{\mathcal{F}(\mathbb{D})}_+$ to (2). In the context of gauge duality, (12) serves as a *weak duality result*. Assumptions (3) can be used to provide a strong duality result: for every $W \in \mathbb{K}$ there exists $Z \in \mathbb{K}^*$ such that equality holds in (12); and for every $Z \in \mathbb{K}^*$ there exists $W \in \mathbb{K}$ such that equality holds in (12). Motivated by (12), we define $\beta$-pairings and $\beta$-certificates.

**Definition 8.** Let $\beta \in (0, 1]$. A *$\beta$-pairing on* $(\mathbb{D},\mathbb{K})$ is a pair $(W, Z) \in \mathbb{K} \times \mathbb{K}^*$ such that there exist $\rho, \mu \in \mathbb{R}_+$ with

$$\langle W, Z\rangle \overset{(13a)}{=} \rho\mu \qquad \text{and} \qquad \beta\rho\mu \overset{(13b)}{\leq} \operatorname{maxq}_{\mathbb{D},\mathbb{K}}(W)\mu \overset{(13c)}{\leq} \rho\mu \overset{(13d)}{\leq} \rho\operatorname{fevc}_{\mathbb{D},\mathbb{K}}(Z) \overset{(13e)}{\leq} \frac{1}{\beta}\rho\mu. \tag{13}$$

If $\mathbb{K} = \mathcal{A}(\mathbb{R}^d_+)$ for a linear map $\mathcal{A}\colon \mathbb{R}^d \to \mathbb{S}^n$, we say that $(w, z) \in \mathbb{R}^d_+ \times \mathbb{R}^d_+$ is a $\beta$-pairing if $(\mathcal{A}(w), Z)$ is a $\beta$-pairing for some $Z \in \mathbb{K}^*$ such that $\mathcal{A}^*(Z) \geq z$. We define an *exact pairing on* $(\mathbb{D},\mathbb{K})$ to be a 1-pairing on $(\mathbb{D},\mathbb{K})$.

Note that, for nonzero $\rho$ and $\mu$, this definition implies that $\beta\rho \leq \operatorname{maxq}(W) \leq \rho$ and $\mu \leq \operatorname{fevc}(Z) \leq (1/\beta)\mu$. Thus, the definition of $\beta$-pairing establishes the idea of simultaneous approximations. We need objects which algorithmically certify that a pair $(W, Z)$ is a $\beta$-pairing. For this, we use an analogue of (12) for our relaxations:

$$\nu(W) \cdot \nu^\circ(Z) \geq \langle W, Z\rangle \qquad \text{for every } W \in \mathbb{K} \text{ and } Z \in \mathbb{K}^*. \tag{14}$$

Similar to (12), inequality (14) follows from (5a) and (6a). The advantage of (14) over (12) is that the quantities here are computable in polynomial time, and by Theorem 7, they are closely related to the quantities in (12).

**Definition 9.** A *$\beta$-certificate for* $(W, Z) \in \mathbb{K} \times \mathbb{K}^*$ is a tuple $(\rho, \mu, U, y, x)$ such that

(15.i) $\rho, \mu \in \mathbb{R}_+$ are such that $\rho\mu = \langle W, Z\rangle$,

(15.ii) $U \in \mathcal{F}(\mathbb{D})$ is such that $s_U^\mathsf{T} W s_U \geq \beta\rho$,

(15.iii) $y \in \mathbb{R}^{\mathcal{F}(\mathbb{D})}_+$ is such that $\sum_{U'\in\mathcal{F}(\mathbb{D})} y_{U'} s_{U'} s_{U'}^\mathsf{T} \succeq_{\widehat{\mathbb{K}^*}} Z$ and $\mathbb{1}^\mathsf{T} y \leq \frac{1}{\beta}\mu$, and

(15.iv) $x \in \mathbb{R}^n$ is such that $\rho \geq \mathbb{1}^\mathsf{T} x$ and $\operatorname{Diag}(x) \succeq_{\mathbb{D}^*} W$.

If $\mathbb{K} = \mathcal{A}(\mathbb{R}^d_+)$ for a linear map $\mathcal{A}\colon \mathbb{R}^d \to \mathbb{S}^n$, we say that $(\rho, \mu, U, y, x)$ is a $\beta$-certificate for $(w, z) \in \mathbb{R}^d_+ \times \mathbb{R}^d_+$ if $(\rho, \mu, U, y, x)$ is a $\beta$-certificate for $(\mathcal{A}(w), Z)$ for some $Z \in \mathbb{K}^*$ with $\mathcal{A}^*(Z) \geq z$.

**Proposition 10.** If there exists a $\beta$-certificate for $(W, Z)$, then $(W, Z)$ is a $\beta$-pairing.

*Proof.* (13a), (13b), and (13e) follow immediately from (15.i), (15.ii), and (15.iii), resp. The connecting part $\operatorname{maxq}(W)\mu \leq \rho\mu \leq \rho\operatorname{fevc}(Z)$ is a combination of two notions of duality: conic duality (via (5)) and conic gauge duality (via (12)). Item (15.iv) provides a feasible solution to (5b), which implies $\operatorname{maxq}(W) \leq \nu(W) \leq \rho$. Hence

$$\operatorname{maxq}(W)\mu \leq \rho\mu \overset{(15.i)}{=} \langle W, Z\rangle \overset{(12)}{\leq} \operatorname{maxq}(W) \cdot \operatorname{fevc}(Z) \leq \rho\operatorname{fevc}(Z). \qquad \square$$

## 3. Generalized Rounding Framework and Sparsification

Let $\Xi$ be a randomized rounding algorithm for $\mathbb{D}$, and let $Y \in \mathcal{E}(\mathbb{D})$. One can roughly see from (9) that sampling from $\Xi_Y$ provides a feasible solution for fevc, as well as a feasible solution for maxq in expectation. However, such a solution may have exponential support size. Moreover, even in well-studied special cases like the Goemans and Williamson algorithm, it is not known how to compute the marginal probabilities exactly to obtain an expression for $\mathbb{E}[\Xi_Y]$.

We show how to obtain a Grothendieck cover by repeated sampling from a randomized rounding algorithm so that we have polynomial support size and the approximation ratio can be controlled with high probability. We first treat the polyhedral case.

**Proposition 11.** Let $\varepsilon, \gamma \in (0,1)$. Let $\mathbb{K} := \mathcal{A}(\mathbb{R}^d_+)$ for a linear map $\mathcal{A}\colon \mathbb{R}^d \to \mathbb{S}^n$. Let $X : \Omega \to \mathbb{K}$ be a random matrix such that $\mathcal{A}^*(\mathbb{E}[X]) \geq \varepsilon\mathbb{1}$. Let $(X_t)_{t\in[T]}$ be i.i.d. random variables sampled from $X$. There is $\psi_{\varepsilon,\gamma} \in \Theta(1)$ such that, if $T \geq \psi_{\varepsilon,\gamma}(\log(d) + \log(n))$, then $\frac{1}{T}\sum_{t\in[T]} X_t \succeq_{\widehat{\mathbb{K}^*}} (1-\gamma)\mathbb{E}[X]$ with probability at least $1 - 1/n$.

The main argument in the proof of Proposition 11, which appears in Appendix C, relies on Chernoff's bound for each generating ray of the cone, followed by union bound on those rays.

**Proposition 12.** Let $\varepsilon, \gamma \in (0,1)$. Let $\mathbb{K} := \mathcal{A}(\mathbb{R}^d_+)$ for a linear map $\mathcal{A}\colon \mathbb{R}^d \to \mathbb{S}^n$. Let $\Xi$ be a randomized rounding algorithm for $\mathbb{D}$. Let $Y \in \mathcal{E}(\mathbb{D})$ be such that $\mathcal{A}^*(Y) \geq \varepsilon\mathbb{1}$. There exists a randomized polynomial-time algorithm producing a Grothendieck cover $y \in \mathbb{R}^{\mathcal{F}(\mathbb{D})}_+$ for $Y$ w.h.p. such that the algorithm performs at most $T := O(\log(d) + \log(n))$ samples from $\Xi_Y$, the support size $|\mathrm{supp}(y)|$ is at most $T$ and $\mathbb{1}^\mathsf{T} y \leq ((1-\gamma)\alpha_\Xi)^{-1}$.

Beyond polyhedral cones, we present a rounding algorithm under the assumption that $\mathbb{K} \subseteq \mathbb{S}^n_+$. In this case, we leverage matrix Chernoff bounds to ensure correctness of our algorithms with high probability. We refer the reader to Appendix C for a complete proof. The result is an application of [Tro15, Corollary 6.2.1], which exploits results arising from a Matrix Chernoff bound with respect to the positive semidefinite (Löwner) order. We denote by $\|X\| := \max\{|\lambda_{\max}(X)|, |\lambda_{\min}(X)|\}$ the *spectral norm* on $\mathbb{S}^n$.

**Proposition 13.** Let $X$ be a random matrix in $\mathbb{S}^n$ such that $\|X\| \leq \rho$ almost surely, and set $\sigma^2 := \|\mathbb{E}[X^2]\|$. Let $(X_t)_{t\in[T]}$ be i.i.d. random variables sampled from $X$. There is $\psi_\gamma = \Theta(1)$ such that, if $T \geq \psi_\gamma \max\{\sigma^2, \rho\}\log(n)$, then $\mathbb{E}[X] - \gamma I \preceq \frac{1}{T}\sum_{t\in[T]} X_t \preceq \mathbb{E}[X] + \gamma I$ holds w.h.p..

The tensor sign covers we can obtain by directly applying Proposition 13 have polynomial support size. To guarantee linear support size, we rely on the following spectral sparsification result:

**Proposition 14** ([CHS15, Corollary 10])**.** Let $Z \in \mathbb{S}^n_+$. Let $A_1, A_2 \ldots, A_m \in \mathbb{S}^n_+$ and $c \in \mathbb{R}^m_+$. Suppose that the semidefinite program $\min\{c^\mathsf{T} y : y \in \mathbb{R}^m_+, \sum_{i=1}^m y_i A_i \succeq Z\}$ has a feasible solution $y^*$. Let $\zeta \in (0,1)$. There is a deterministic polynomial-time algorithm that, given $y^*$, and the matrices $A_1, A_2 \ldots, A_m$ and $Z$ as input, computes a feasible solution $\bar{y}$ with at most $O(n/\zeta^2)$ nonzero entries and $c^\mathsf{T}\bar{y} \leq (1+\zeta)c^\mathsf{T} y^*$.

Thus, we obtain the following result, which is proved in Appendix C.

**Proposition 15.** Let $\gamma, \varepsilon, \zeta \in (0,1)$. Let $\mathbb{K} \subseteq \mathbb{S}^n_+$. Let $\Xi$ be a randomized rounding algorithm for $\mathbb{D}$. Let $Y \in \mathcal{E}(\mathbb{D})$ be such that $Y \succeq \varepsilon I$. There exists a randomized polynomial time algorithm producing a Grothendieck cover $y \in \mathbb{R}^{\mathcal{F}(\mathbb{D})}_+$ for $Y$ w.h.p. such that the algorithm performs at most $O(n^2\log(n))$ samples from $\Xi_Y$, the support size $|\mathrm{supp}(y)|$ is $O(n/\zeta^2)$ and $\mathbb{1}^\mathsf{T} y \leq (1+\zeta)((1-\gamma)\alpha_\Xi)^{-1}$.

## 4. Simultaneous Approximation Algorithms

The last ingredient of our algorithms is to ensure the feasible solutions behave well with respect to our sampling results. Both Propositions 12 and 15 require a numeric bound $\varepsilon$ on how interior to the cone the feasible solutions are: either by requiring $Y \succeq \varepsilon I$ or $\mathcal{A}^*(Y) \geq \varepsilon \mathbb{1}$. These assumptions are necessary: [Ben+23] exhibits instances of the fractional cut-covering problem and optimal solutions to the SDP relaxation that require, in expectation, exponentially many samples to ensure feasibility. For a fixed element of $\mathbb{K}^*$ which is "central" enough, we define perturbed versions of (5) and (6) whose feasible regions exclude these ill-behaved matrices. For concreteness, we assume that

$$I \in \operatorname{cone}\{\, s_U s_U^\mathsf{T} : U \in \mathcal{F}(\mathbb{D})\} \subseteq \mathbb{D}, \tag{16}$$

which can be easily verified in the examples we will work with.

We now describe one of the algorithms in Theorem 5. Let $\Xi$ be a randomized rounding algorithm for $\mathbb{D}$. Let $\beta \in (0, \alpha_\Xi)$. Assume we are given an instance $Z \in \mathbb{K}^*$ of fevc as input. Then

(1) nearly solve the perturbed version of (6) to compute $(\mu, Y) \in \mathbb{R}_+ \times \mathbb{D}$ and $(W, x) \in \mathbb{K} \times \mathbb{R}^n$;
(2) sample $O(n^2 \log n)$ times from $\Xi_Y$ to obtain a Grothendieck cover $y \in \mathbb{R}_+^{\mathcal{F}(\mathbb{D})}$ for $Z$;
(3) apply Proposition 14 to reduce the support size of $y$ to $O(n)$;
(4) choose $U$ that maximizes $s_{U'}^\mathsf{T} W s_{U'}$ among all $U' \in \operatorname{supp}(y)$;
(5) output $W$ and the $\beta$-certificate $(1, \mu, U, y, x)$.

(Steps (1)–(3) involve errors terms that are chosen small enough to guarantee our desired approximation factor $\beta$.) Proposition 15 proves the correctness of steps (2) and (3). This is where we crucially exploit $\mathbb{K} \subseteq \mathbb{S}^n_+$, so that concentration and sparsification results developed for positive semidefinite matrices can be translated to the cone $\mathbb{K}^*$. That (4) will define a set $U$ which is part of the $\beta$-certificate follows from $y$ being a good enough estimate: we have that $\beta\rho \leq s_U^\mathsf{T} W s_U$ since

$$\rho\mu = \langle W, Z\rangle \leq \sum_{U' \in \mathcal{F}(\mathbb{D})} y_{U'} \langle W, s_{U'} s_{U'}^\mathsf{T}\rangle \leq \langle W, s_U s_U^\mathsf{T}\rangle \mathbb{1}^\mathsf{T} y \leq \left(s_U^\mathsf{T} W s_U\right) \frac{1}{\beta}\mu.$$

Appendix D has the precise proofs.

The algorithm sketched above highlights an important part of our framework. For a given instance $Z \in \mathbb{K}^*$ of fevc, we obtain from the SDP solutions to (6) an instance $W \in \mathbb{K}$ of maxq, and we then certify *the pair* $(W, Z)$. This mapping among instances is something we now make explicit. Define $\mathrm{H}_{\mathbb{D},\mathbb{K}} := \{\, (Z, W) \in \mathbb{K} \times \mathbb{K}^* : \langle W, Z\rangle = \nu_{\mathbb{D},\mathbb{K}}(W) \cdot \nu^\circ_{\mathbb{D},\mathbb{K}}(Z)\}$. One may prove that

$$\mathrm{H} = \left\{ (W, Z) \in \mathbb{K} \times \mathbb{K}^* : \begin{array}{c} \exists(\mu, Y) \text{ feasible in (6a) for } Z, \\ \exists(\rho, x) \text{ feasible in (5b) for } W, \\ \text{and } \langle W, Z\rangle = \rho\mu \end{array} \right\}. \tag{17}$$

We invite the reader to compare the RHS of (17) with the feasible regions of (5) and (6). One may see solving either SDP as fixing one side of the pair of instances and obtaining the other; i.e., as computing an element of $\mathfrak{Z}(W) := \{\, Z \in \mathbb{K}^* : (W, Z) \in \mathrm{H}\}$ when given $W \in \mathbb{K}$ as input, or computing an element of $\mathfrak{W}(Z) := \{\, W \in \mathbb{K} : (W, Z) \in \mathrm{H}\}$ when given $Z \in \mathbb{K}^*$ as input. In both cases, by solving a single (primal-dual pair of) SDP we obtain an element of H and the objects $(\rho, x)$ and $(\mu, Y)$ which witness the membership.

We now address Remark 6, in which the cone $\mathbb{K}$ is polyhedral and not necessarily contained in $\mathbb{S}^n_+$. Here, we do not require the use of sparsification, as the cover produced is already (very) sparse.

**Theorem 16** (Main Polyhedral Theorem)**.** Let $\mathbb{K} := \mathcal{A}(\mathbb{R}^d_+)$ for a linear map $\mathcal{A}\colon \mathbb{R}^d \to \mathbb{S}^n$. Assume (16) and that $\mathcal{A}^*(I) \geq \kappa\mathbb{1}$ for some positive $\kappa \in \mathbb{R}$. Let $\Xi$ be a randomized rounding algorithm for $\mathbb{D}$. Fix $\beta \in (0, \alpha_\Xi)$. There exists a randomized polynomial-time algorithm that, given an instance $z \in \mathbb{R}^d_+$ of fevc as input, computes an instance $w \in \mathbb{R}^d_+$ of maxq and a $\beta$-certificate for $(w, z)$. Dually, there exists a randomized polynomial-time algorithm that, given an instance $w \in \mathbb{R}^d_+$ of maxq as input, computes an instance $z \in \mathbb{R}^d_+$ of fevc and a $\beta$-certificate for $(w, z)$. Both algorithms output covers whose support size is bounded by $C \cdot (\log(d) + \log(n))$, where $C := C(\kappa, \alpha_\Xi, \beta)$ is independent of $d$ and $n$.

## 5. Boolean 2-CSP

Let $U \subseteq \{0\} \cup [n]$ with $0 \in U$. Let $x\colon [n] \to \{\mathtt{false}, \mathtt{true}\}$ be defined such that $x_i = \mathtt{true}$ if and only if $i \in U \setminus \{0\}$. Let $i, j \in [n]$. For any predicate $P$, we let $[P] \in \{0,1\}$ be 1 if the predicate $P$ is $\mathtt{true}$, and 0 otherwise. Appendix E defines matrices $\Delta_{\pm i, \pm j} \in \mathbb{S}^{\{0\}\cup[n]}$ such that

$$[\overline{x_i} \wedge \overline{x_j}] = \tfrac{1}{4}\langle \Delta_{-i,-j}, s_U s_U^{\mathsf{T}} \rangle, \qquad [\overline{x_i} \wedge x_j] \;= \tfrac{1}{4}\langle \Delta_{-i,+j}, s_U s_U^{\mathsf{T}} \rangle,$$
$$[x_i \wedge \overline{x_j}] = \tfrac{1}{4}\langle \Delta_{+i,-j}, s_U s_U^{\mathsf{T}} \rangle, \qquad [x_i \wedge x_j] \;= \tfrac{1}{4}\langle \Delta_{+i,+j}, s_U s_U^{\mathsf{T}} \rangle.$$

By decomposing a predicate as a disjunction of conjunctions, one can write any Boolean function on two variables as a sum of these matrices. Thus, for any set $\mathscr{C}$ of constraints on two variables, one can define a linear map $\mathcal{A}\colon \mathbb{R}^{\mathscr{C}} \to \mathbb{S}^{\{0\}\cup[n]}$ such that

$$\langle \mathcal{A}(e_f), s_U s_U^{\mathsf{T}} \rangle = [f(x)] \text{ for every } f \in \mathscr{C}; \tag{18}$$

here, $e_f \coloneqq \mathbb{1}_{\{f\}} \in \{0,1\}^{\mathscr{C}}$ is a canonical basis vector. With this particular linear map $\mathcal{A}$, we say that $\mathbb{K}$ is the *polyhedral cone defined by* $\mathscr{C}$ if $\mathbb{K} = \mathcal{A}(\mathbb{R}^{\mathscr{C}}_+)$. The definitions are made so that

$$\mathbb{E}[\Xi_Y] \succeq_{\widehat{\mathbb{K}^*}} \alpha Y \text{ if and only if } \mathbb{P}(f(x) = \mathtt{true}) \geq \alpha \langle \mathcal{A}(e_f), Y \rangle \text{ for every } f \in \mathscr{C}, \tag{19}$$

where $x \in \mathscr{A}$ is obtained from $\Xi_Y$ in the following way: let $U \subseteq \{0\} \cup [n]$ be such that $0 \in U$ and $s_U s_U^{\mathsf{T}}$ was sampled from $\Xi_Y$, and define $x \in \mathscr{A}$ by $x_i = \mathtt{true}$ if and only if $i \in U \setminus \{0\}$. If we set $\Delta^n \coloneqq \operatorname{conv}(\bigcup_{i,j\in[n]}\{\Delta_{\pm i,\pm j}\})$, it is immediate that $\mathbb{K} \subseteq \operatorname{cone}(\Delta^n)$. The set $\mathbb{D}_\Delta \coloneqq \mathbb{S}^{\{0\}\cup[n]}_+ \cap \big(\bigcup_{i,j\in[n]}\{\Delta_{\pm i,\pm j}\}\big)^*$ has been studied — see e.g., [LLZ02; Rag08] —, and these additional inequalities are referred to as *triangle inequalities*. Since $\mathrm{CUT}^{\mathbb{D}_\Delta} = \operatorname{conv}\{\, s_U s_U^{\mathsf{T}} : U \subseteq \{0\} \cup [n]\}$ (see Appendix E) we have that $\mathbb{D}_\Delta^* = \mathbb{S}^{\{0\}\cup[n]}_+ + \operatorname{cone}(\Delta^n)$. This then ensures that (3) holds for $\mathbb{D}_\Delta$ and the polyhedral cone $\mathbb{K}$ defined by $\mathscr{C}$.

**Theorem 17** (Fractional $\mathfrak{P}$-Covering Theorem)**.** Let $\mathfrak{P}$ be a set of predicates in two Boolean variables. For every $n \in \mathbb{N}$, let $\Xi_n$ be a randomized rounding algorithm for $\mathbb{D}_\Delta \subseteq \mathbb{S}^{n+1}$. Let

$$\alpha \leq \inf\left\{ \alpha_{\mathbb{D}_\Delta,\mathbb{K},\Xi_n} : \begin{array}{c} \mathscr{C} \text{ set of } \mathfrak{P}\text{-constraints on } n \text{ variables,} \\ \mathbb{K} \text{ polyhedral cone defined by } \mathscr{C}. \end{array} \right\},$$

and fix $\beta \in (0, \alpha)$. There exists a randomized polynomial-time algorithm that, given an instance $(\mathscr{C}, z)$ of frac-$\mathfrak{P}$-cover as input, computes $w \in \mathbb{R}^{\mathscr{C}}_+$ and a $\beta$-certificate for $(w, z)$. Dually, there exists a polynomial-time randomized algorithm that, given an instance $(\mathscr{C}, w)$ of max-$\mathfrak{P}$-sat as input, computes $z \in \mathbb{R}^{\mathscr{C}}_+$ and a $\beta$-certificate for $(w, z)$. Both algorithms take at most $O(\log n)$ samples from $\Xi_n$ and produce covers with $O(\log n)$ support size.

*Proof of Theorem 17.* Set $\mathbb{D} \coloneqq \mathbb{D}_\Delta$ and let $\mathcal{A}\colon \mathbb{R}^{\mathscr{C}} \to \mathbb{S}^{\{0\}\cup[n]}$ be as in (18). From (55) we have that (16) holds. Since $\mathcal{A}^*(s_U s_U^{\mathsf{T}}) = \operatorname{val}_{\mathscr{C}}(x)$, using that $2^{-n} I = \sum\{\, s_{\{0\}\cup U} s_{\{0\}\cup U}^{\mathsf{T}} : U \subseteq [n]\}$, we see that $\mathcal{A}^*(I)$ computes the marginal probability of satisfying each constraint by uniformly sampling an assignment in $\mathscr{A}$. As the constant $\mathtt{false}$ function is not in $\mathfrak{P}$, any constraint is satisfied by at least $1/4$ of the assignments. Hence $\mathcal{A}^*(I) \geq \frac{1}{4}\mathbb{1}$. Note that since $|\mathfrak{P}| \leq 16$, we have that $\log(|\mathscr{C}|) = O(\log n)$. Theorem 16 then ensures we can compute $\beta$-certificates $(\rho, \mu, U, y, x)$ with $|\operatorname{supp}(y)| = O(\log(|\mathscr{C}|) + \log(n)) = O(\log(n))$. ∎

*Proof of Theorem 1.* Set $\mathfrak{P} \coloneqq \{x_1 \wedge \overline{x_2}\}$. For every digraph $D = (V, A)$, each arc $uv$ can be mapped to a constraint $x_u \wedge \overline{x_v}$. Hence, there exists $\mathscr{C}$ such that $\operatorname{md}(D, w) = $ max-$\mathfrak{P}$-sat$(\mathscr{C}, w)$ and $\operatorname{fdc}(\mathscr{C}, z) = $ frac-$\mathfrak{P}$-cover$(\mathscr{C}, z)$ for every $w \in \mathbb{R}^A_+$ and $z \in \mathbb{R}^A_+$. [Bra+23] — see formulation after Proposition 2.4. — define $\Xi$ such that

$$\langle \mathbb{E}[\Xi_Y], \tfrac{1}{4}\Delta_{+u,-v} \rangle \geq \alpha_{\mathrm{BHPZ}} \langle Y, \tfrac{1}{4}\Delta_{+u,-v} \rangle \text{ for every arc } uv \in A \text{ and } Y \in \mathcal{E}(\mathbb{D}).$$

Thus $\mathbb{E}[\Xi_Y] \succeq_{\widehat{\mathbb{K}^*}} \alpha_{\mathrm{BHPZ}} Y$, so Theorem 17 finishes the proof. ∎

*Proof of Theorem 3.* Let $\Xi_Y = s_U s_U^{\mathsf{T}}$ where $U = \{0\} \cup \{\, i \in [n] : x_i = \texttt{true}\}$ for $x \in \mathscr{A}$ being sampled from the algorithm defined by Lewin, Livnat, and Zwick [LLZ02]. Then

$$\begin{aligned}\langle \mathbb{E}[\Xi_Y], \tfrac{1}{4}(\Delta_{-i,+j} + \Delta_{+i,-j} + \Delta_{+i,+j})\rangle &= \mathbb{P}(x_i \vee x_j) && \text{def. of } \Xi \text{ and } \Delta\\ &\geq \alpha_{\mathrm{LLZ}}\langle Y, \tfrac{1}{4}(3 + Y_{0i} + Y_{0j} + Y_{ij})\rangle && \text{by [LLZ02]}\\ &= \alpha_{\mathrm{LLZ}}\langle Y, \tfrac{1}{4}(\Delta_{-i,+j} + \Delta_{+i,-j} + \Delta_{+i,+j})\rangle.\end{aligned}$$

The mismatch between the expression in the second line and the expression in [LLZ02, Section 5] arises from our modelling imposing $x_0 = \texttt{true}$, whereas [LLZ02] impose $x_0 = \texttt{false}$. The case for constraints $\overline{x_i} \vee x_j$, $x_i \vee \overline{x_j}$, and $\overline{x_i} \vee \overline{x_j}$ is analogous. As this holds for every constraint, we have that $\mathbb{E}[\Xi_Y] \succeq_{\widehat{\mathbb{K}^*}} \alpha_{\mathrm{LLZ}} Y$, where $\mathbb{K}$ is the polyhedral cone defined by $\mathscr{C}$. Thus $\alpha_{\mathrm{LLZ}} \leq \alpha_{\mathbb{D},\mathbb{K},\Xi}$ and Theorem 17 implies the statement. $\square$

## 6. Concluding Remarks and Future Directions

Despite its generality, our framework still captures several best possible and best known results. A first aspect concerns the approximation constants of the algorithms presented. We refer to

$$\varrho_{\mathbb{D},\mathbb{K}} := \min\left\{ \frac{\operatorname{maxq}_{\mathbb{D},\mathbb{K}}(W)}{\nu_{\mathbb{D},\mathbb{K}}(W)} : W \in \mathbb{K}\right\} = \min\left\{ \frac{\nu^{\circ}_{\mathbb{D},\mathbb{K}}(Z)}{\operatorname{fevc}_{\mathbb{D},\mathbb{K}}(Z)} : Z \in \mathbb{K}^*\right\} \tag{20}$$

as the *integrality ratio of* $(\mathbb{D},\mathbb{K})$. Equality between the two expressions above follows from gauge duality. One may see that $\varrho_{\mathbb{D},\mathbb{K}}$ is the largest $\beta$ such that every $(W,Z) \in \mathrm{H}(\mathbb{D},\mathbb{K})$ is a $\beta$-pairing. Positivity of $\varrho_{\mathbb{D},\mathbb{K}}$ is a corollary of all norms on a finite-dimensional vector space being equivalent. It is more interesting then to consider families $\mathcal{I}$ of triples $(\mathbb{D},\mathbb{K},\Xi)$ where $\Xi$ is a randomized rounding algorithm for $\mathbb{D}$. Theorem 7 implies that

$$\varrho_{\mathcal{I}} := \inf\{\, \varrho_{\mathbb{D},\mathbb{K}} : (\mathbb{D},\mathbb{K},\Xi) \in \mathcal{I}\} \geq \inf\{\, \alpha_{\mathbb{D},\mathbb{K},\Xi} : (\mathbb{D},\mathbb{K},\Xi) \in \mathcal{I}\} =: \alpha_{\mathcal{I}}. \tag{21}$$

For example, if $\mathcal{I}$ encodes all the cones arising from instances of the maximum cut problem, to say that $\varrho_{\mathcal{I}} \geq \alpha_{\mathrm{GW}}$ is to say we have a $\alpha_{\mathrm{GW}}$-approximation algorithm for the maximum cut problem, and a $1/\alpha_{\mathrm{GW}}$-approximation algorithm for the fractional cut covering problem. Equality in (21) indicates that no better approximation algorithm can be obtained without strengthening the formulation (by changing $\mathbb{D}$) or restricting the input instances (by changing $\mathbb{K}$). Whenever $\mathcal{I}$ arises from instances related to a specific 2-CSP, Raghavendra [Rag08, Corollary 1.5] shows that the triangle inequalities (i.e., $\mathbb{D}_\Delta$) are enough, as there exists a randomized rounding algorithm ensuring equality in (21). It is also known that $2/\pi = \inf\{\, \varrho_{\mathbb{S}^n_+,\mathbb{S}^n_+} : n \in \mathbb{N}\}$ [AN06; FL20; Gro53; Kri77]. In this way, the algorithms in Theorems 5 and 17 all have tight analyses.

One formulation of the "equivalence between separation and optimization" proved by Grötschel, Lovász, and Schrijver [GLS81] is that one can compute a positive definite monotone gauge whenever one can compute its dual. In this way, whenever $\nu_{\mathbb{D},\mathbb{K}}$ is the best polynomial-time computable approximation to $\operatorname{maxq}_{\mathbb{D},\mathbb{K}}$ under the Unique Games Conjecture [Kho+07] (and assuming $\mathrm{P} \neq \mathrm{NP}$), the same immediately holds for $\nu^{\circ}_{\mathbb{D},\mathbb{K}}$ and $\operatorname{fevc}_{\mathbb{D},\mathbb{K}}$. In particular, [Rag08] shows that, assuming the UGC, it is NP-hard to obtain any approximation algorithm for a Boolean 2-CSP with approximation factor better than $\varrho_{\mathbb{D}_\Delta,\mathbb{K}}$. Thus, the UGC implies that Theorem 17 is best possible unless $\mathrm{P} = \mathrm{NP}$.

The support size bounds in Theorems 17 and 5 are asymptotically tight. If $\mathbb{K} = \mathbb{S}^n_+$, it is immediate that any feasible $y$ in (6a) for $Z = I$ has $|\operatorname{supp}(y)| \geq n$. Hence the $O(n)$ support size in Theorem 5 is best possible. [Ben+23] argues that $|\operatorname{supp}(y)| \geq \log(\chi(G))$ for every graph $G = (V,E)$ whenever $\mathbb{K} = \mathcal{L}_G(\mathbb{R}^E_+)$, where $\mathcal{L}_G(w) := \sum_{ij\in E} w_{ij}(e_i - e_j)(e_i - e_j)^{\mathsf{T}}$ is the *Laplacian of* $G$. Hence the $O(\log n)$ support size in Theorem 17 is also best possible.

Tightly related to the support size of the solutions we produce, is the number of samples necessary to ensure a good enough cover with high probability. Although Theorem 5 shows that $O(n^2 \log n)$ samples suffice when $\mathbb{K} \subseteq \mathbb{S}^n_+$, in specific cones we exploited *conic concentration bounds* in Propositions 11 and 24 to obtain better sampling bounds. It is conceivable that other families of cones also admit better bounds. E.g., [SZ22] offers conic concentration for hyperbolicity cones.

Three of the natural generalizations of our framework not discussed here are:

(i) extension to the complex field and Hermitian matrices,
(ii) extension of the intractable pairs defined by exponentially many constraints and exponentially many variables to a semi-infinite setting (infinitely many constraints in the intractable primal and infinitely many variables in the intractable gauge dual),
(iii) extension to handle general CSPs.

The first two generalizations allow the treatment of many applications in continuous mathematics and engineering, including some applications in robust optimization and system and control theory. The underlying theoretical results include as a special case the Extended Matrix Cube Theorem [BNR03].

## Appendix A. maxq and fevc, and their Conic Relaxations

We denote by $\operatorname{aff}(S)$ the *affine hull* of the set $S \subseteq \mathbb{S}^n$, which is the intersection of all affine subspaces of $\mathbb{S}^n$ containing $S$. Define the *dual cone* of $S \subseteq \mathbb{S}^n$ as

$$S^* := \{\, X \in \operatorname{aff}(S) : \langle Y, X\rangle \geq 0 \text{ for all } Y \in S\}.$$

For full-dimensional convex sets, our definition matches the usual definition of dual cone. In general, the dual of a set is taken in its affine hull, analogous to how the relative interior is taken with respect to the topology induced in the affine hull of the set. This becomes most relevant as (3) allows for cones $\mathbb{K}$ which are not full dimensional. Although for a convex cone $\mathbb{K} \subseteq \mathbb{S}^n$ the set $\operatorname{aff}(\mathbb{K}^*)$ may be strictly smaller than $\operatorname{aff}(\mathbb{K})$ — and hence $\mathbb{K}^{**}$ may not be $\mathbb{K}$ —, if $\mathbb{K}$ is a pointed cone, then $\mathbb{K}^{**} = \mathbb{K}$.

Let $\mathbb{D}, \mathbb{K} \subseteq \mathbb{S}^n$ be closed convex cones. Recall that $\mathrm{CUT}^{\mathbb{D}} := \operatorname{conv}\{\, s_U s_U^\mathsf{T} : U \subseteq V,\ s_U s_U^\mathsf{T} \in \mathbb{D}\}$. Assume

$$\mathbb{D} \subseteq \mathbb{S}^n_+, \tag{22a}$$

$$\mathbb{K} \subseteq \mathbb{D}^*, \tag{22b}$$

$$\operatorname{int}(\mathrm{CUT}^{\mathbb{D}}) \neq \varnothing, \tag{22c}$$

$$\mathbb{K} = \widehat{\mathbb{K}} \cap \operatorname{Null}(\mathcal{L}) \text{ and } \exists \mathring{X} \in \operatorname{int}(\widehat{\mathbb{K}}) \setminus \{0\} \text{ s.t. } \mathcal{L}(\mathring{X}) = 0, \tag{22d}$$

where $\widehat{\mathbb{K}} \subseteq \mathbb{S}^n$ is a closed convex cone and $\mathcal{L} \colon \mathbb{S}^n \to \mathbb{R}^k$ is a linear map for some $k \in \mathbb{N}$. Under these assumptions, we have that

$$\mathbb{K} \text{ is pointed and } \operatorname{ri}(\mathbb{K}) \setminus \{0\} \neq \varnothing. \tag{23}$$

Note that $\mathbb{D}^*$ is pointed, since $\operatorname{int}(\mathbb{D}) \neq \varnothing$ by (22c). As $\mathbb{D}^* \supseteq \mathbb{K}$ by (22b), we conclude $\mathbb{K}$ is pointed. The second part follows from (22d), as $\mathring{X} \neq 0$ and $\operatorname{Null}(\mathcal{L})$ is the smallest affine subspace of $\mathbb{S}^n$ containing $\mathbb{K}$.

Let $\mathbb{D}, \mathbb{K}$ be closed convex cones such that (22) holds, and let $\mathcal{L} \colon \mathbb{S}^n \to \mathbb{R}^k$ and $\widehat{\mathbb{K}}$ be the linear transformation and cone appearing in (22d), respectively. We write

$$\widehat{\mathbb{K}^*} := \widehat{\mathbb{K}}^* + \operatorname{Im}(\mathcal{L}^*) \subseteq \mathbb{S}^n. \tag{24}$$

If we denote by $P \colon \mathbb{S}^n \to \mathbb{S}^n$ the orthogonal projector onto $\operatorname{aff}(\mathbb{K}) = \operatorname{Null}(\mathcal{L})$, then

$$\mathbb{K}^* = P(\widehat{\mathbb{K}^*}). \tag{25}$$

This relationship motivates the notation in (24): it shows that $\widehat{\mathbb{K}^*}$ is a lifting of the cone $\mathbb{K}^*$. In our setting, we will have $\operatorname{aff}(\mathbb{K})$ as the *instance space*, where the inputs to our gauges arise from, and $\mathbb{S}^n$ as the lifted space where optimization is performed. In this way, both $\mathbb{K}^*$ and its lifting $\widehat{\mathbb{K}^*}$ appear throughout our developments. From (24) we have that $\widehat{\mathbb{K}^*} \supseteq \operatorname{Im}(\mathcal{L}^*) = \operatorname{Null}(\mathcal{L})^\perp = \operatorname{aff}(\mathbb{K})^\perp$. Hence

$$\operatorname{aff}(\mathbb{K})^\perp \subseteq \widehat{\mathbb{K}^*}. \tag{26}$$

From (22b), (22d), and (24) we have that

$$\mathbb{D} \subseteq \widehat{\mathbb{K}^*}. \tag{27}$$

Since $\mathbb{K}$ is pointed by (23), we have that

$$\mathbb{K}^{**} = \mathbb{K}. \tag{28}$$

Finally, the orthogonal projector gives a convenient map from $\mathbb{D}$ to $\mathbb{K}^*$, since

$$Y \succeq_{\widehat{\mathbb{K}^*}} P(Y) \in \mathbb{K}^* \text{ for every } Y \in \mathbb{D}. \tag{29}$$

Indeed, for every $Y \in \mathbb{D}$ we have that $P(Y) \in \mathbb{K}^*$, as $P(\mathbb{D}) \subseteq P(\widehat{\mathbb{K}^*}) = \mathbb{K}^*$ by (25) and (27). Moreover, $Y - P(Y) \in \operatorname{Null}(\mathcal{L})^\perp = \operatorname{Im}(\mathcal{L}^*) \subseteq \widehat{\mathbb{K}^*}$ by (24), so (29) holds.

Recall the definitions of $\operatorname{maxq}_{\mathbb{D},\mathbb{K}}$, $\operatorname{fevc}_{\mathbb{D},\mathbb{K}}$, $\nu_{\mathbb{D},\mathbb{K}}$, and $\nu^{\circ}_{\mathbb{D},\mathbb{K}}$, along with conic dual formulations, for each $W \in \mathbb{K}$ and $Z \in \mathbb{K}^*$:

$$\operatorname{maxq}_{\mathbb{D},\mathbb{K}}(W) \coloneqq \max\{\, s_U^{\mathsf{T}} W s_U : U \in \mathcal{F}(\mathbb{D})\}, \tag{30}$$

$$\operatorname{fevc}_{\mathbb{D},\mathbb{K}}(Z) \coloneqq \min\left\{\, \mathbb{1}^{\mathsf{T}} y : y \in \mathbb{R}_+^{\mathcal{F}(\mathbb{D})}, \sum_{U\in\mathcal{F}(\mathbb{D})} y_U s_U s_U^{\mathsf{T}} \succeq_{\widehat{\mathbb{K}^*}} Z \right\} \tag{31a}$$

$$= \max\{\, \langle Z, X\rangle : X \in \mathbb{K},\, \langle s_U s_U^{\mathsf{T}}, X\rangle \leq 1 \text{ for every } U \in \mathcal{F}(\mathbb{D})\}; \tag{31b}$$

$$\nu_{\mathbb{D},\mathbb{K}}(W) \coloneqq \max\{\, \langle W, Y\rangle : Y \in \mathbb{D},\, \operatorname{diag}(Y) = \mathbb{1}\} \tag{32a}$$

$$= \min\{\, \mathbb{1}^{\mathsf{T}} x : x \in \mathbb{R}^n,\, \operatorname{Diag}(x) \succeq_{\mathbb{D}^*} W\}, \tag{32b}$$

$$\nu^{\circ}_{\mathbb{D},\mathbb{K}}(Z) \coloneqq \min\{\, \mu : \mu \in \mathbb{R}_+,\, Y \in \mathbb{D},\, \operatorname{diag}(Y) = \mu\mathbb{1},\, Y \succeq_{\widehat{\mathbb{K}^*}} Z\} \tag{33a}$$

$$= \max\{\, \langle Z, X\rangle : X \in \mathbb{K},\, x \in \mathbb{R}^n,\, \operatorname{Diag}(x) \succeq_{\mathbb{D}^*} X,\, \mathbb{1}^{\mathsf{T}} x \leq 1\}. \tag{33b}$$

Our arguments rely on standard results on Conic Programming Duality — see, e.g., [Nem24, Chapter 7]. In particular, a *strictly feasible solution* to an optimization problem is a feasible solution where every conic constraint is satisfied by a point in the interior of the relevant cone. By (22c), there exists

$$\mathring{y} \in \mathbb{R}_+^{\mathcal{F}(\mathbb{D})} \text{ such that } \sum_{S\in\mathcal{F}(\mathbb{D})} \mathring{y}_U s_U s_U^{\mathsf{T}} \eqqcolon \mathring{Y} \in \operatorname{int}(\mathbb{D}) \text{ and } \operatorname{diag}(\mathring{Y}) = \mathbb{1}. \tag{34}$$

We may assume that $\mathring{y} > 0$. Note that $\alpha\mathring{Y} - Z = \alpha(\mathring{Y} - \frac{1}{\alpha}Z) \in \operatorname{int}(\mathbb{D}) \subseteq \operatorname{int}(\widehat{\mathbb{K}^*})$ for large enough $\alpha \in \mathbb{R}_{++}$. Hence (31a) has a strictly feasible solution. From (22d) one may reformulate (31b) into an equivalent problem with a strictly feasible solution. Conic Programming Strong Duality [Nem24, Theorem 7.2] implies equality and attainment in (31). Similarly, note that $\mathring{Y}$ is a strictly feasible solution to (32a), whereas $\mathring{x} \coloneqq 2\lambda_{\max}(W)\mathbb{1}$ is a strictly feasible solution for (32b), as $\mathbb{S}^n_+ \subseteq \mathbb{D}^*$ by (22a). Once again, Strong Duality ensures equality and attainment in (32). A positive multiple of $(1, \mathring{Y})$ is a strictly feasible point to (33a). Let $\mathring{X} \in \mathbb{K}$ be as in (22d). Without loss of generality, assume that $\lambda_{\max}(\mathring{X}) < 1$. Then $(\frac{1}{2n}\mathbb{1}, \frac{1}{2n}\mathring{X})$ is a strictly feasible solution to (33b). Hence equality and attainment holds in (33).

We will look at these functions through the lens of conic gauges, which are defined as follows:

**Definition 18** (Conic Gauges)**.** Let $\mathbb{E}$ be an Euclidean space. Let $\mathbb{K} \subseteq \mathbb{E}$ be a closed convex cone. A function $\varphi \colon \mathbb{K} \to \mathbb{R}_+$ is a *gauge* if $\varphi$ is positively homogeneous, sublinear, and $\varphi(0) = 0$. The gauge $\varphi$ is *positive definite* if $\varphi(x) > 0$ for each nonzero $x \in \mathbb{K}$, and $\varphi$ is monotone if $0 \preceq_{\mathbb{K}} x \preceq_{\mathbb{K}} y$ implies $\varphi(x) \leq \varphi(y)$.

**Definition 19.** Let $\varphi \colon \mathbb{K} \to \mathbb{R}_+$ be a positive definite monotone gauge. The *dual* of $\varphi$ is the positive definite monotone gauge $\varphi^{\circ} \colon \mathbb{K}^* \to \mathbb{R}_+$ defined by

$$\varphi^{\circ}(y) \coloneqq \max\{\, \langle y, x\rangle : x \in \mathbb{K},\, \varphi(x) \leq 1\} \qquad \text{for each } y \in \mathbb{K}^*. \tag{35}$$

Let $\phi \colon \mathbb{K} \to \mathbb{R}_+$ be a positive definite monotone gauge. Whenever $\mathbb{K}^{**} = \mathbb{K}$, — in particular whenever (22) holds — one can prove that $\phi^{\circ\circ} = \phi$. We show that $\operatorname{maxq}_{\mathbb{D},\mathbb{K}}$, $\operatorname{fevc}_{\mathbb{D},\mathbb{K}}$, $\nu_{\mathbb{D},\mathbb{K}}$, and $\nu^{\circ}_{\mathbb{D},\mathbb{K}}$ are positive definite monotone gauges and how they are related.

**Theorem 20.** Let $\mathbb{D}, \mathbb{K} \subseteq \mathbb{S}^n$ be closed convex cones such that (22) holds. Then

(i) $\operatorname{maxq}_{\mathbb{D},\mathbb{K}}$ and $\operatorname{fevc}_{\mathbb{D},\mathbb{K}}$ are positive definite monotone gauges, dual to each other;

(ii) $\nu_{\mathbb{D},\mathbb{K}}$ and $\nu^{\circ}_{\mathbb{D},\mathbb{K}}$ are positive definite monotone gauges, dual to each other;

(iii) $\operatorname{maxq}_{\mathbb{D},\mathbb{K}} \leq \nu_{\mathbb{D},\mathbb{K}}$ and $\nu^{\circ}_{\mathbb{D},\mathbb{K}} \leq \operatorname{fevc}_{\mathbb{D},\mathbb{K}}$ .

*Proof.* The fact that $\operatorname{maxq}_{\mathbb{D},\mathbb{K}}$ and $\nu_{\mathbb{D},\mathbb{K}}$ are gauges follows directly from their definitions in (30) and (32a). The monotonicity of $\operatorname{maxq}_{\mathbb{D},\mathbb{K}}$ and $\nu_{\mathbb{D},\mathbb{K}}$ is a direct consequence of (22b). Next we show that $\operatorname{maxq}_{\mathbb{D},\mathbb{K}}$ is positive definite. Let $\mathring{y} \in \mathcal{F}(\mathbb{D})$ and $\mathring{Y} \in \operatorname{int}(\mathbb{D})$ be as in (34). Then

$$0 < \langle \mathring{Y}, W \rangle = \sum_{U \in \mathcal{F}(\mathbb{D})} \mathring{y}_U \langle s_U s_U^\mathsf{T}, W \rangle \text{ for every nonzero } W \in \mathbb{K} \subseteq \mathbb{D}^*.$$

Thus there exists $U \in \mathcal{F}(\mathbb{D})$ such that $\operatorname{maxq}_{\mathbb{D},\mathbb{K}}(W) \geq s_U^\mathsf{T} W s_U > 0$. This implies that $\operatorname{maxq}_{\mathbb{D},\mathbb{K}}$ is positive definite. Since

$$\operatorname{maxq}_{\mathbb{D},\mathbb{K}}(W) = \max\{\, s_U^\mathsf{T} W s_U : U \in \mathcal{F}(\mathbb{D})\} \leq \max\{\, \langle W, Y \rangle : Y \in \mathbb{D},\, \operatorname{diag}(Y) = \mathbb{1}\} = \nu_{\mathbb{D},\mathbb{K}}(W),$$

it follows that $\nu_{\mathbb{D},\mathbb{K}}$ is positive definite. Thus, $\operatorname{maxq}_{\mathbb{D},\mathbb{K}}$ and $\nu_{\mathbb{D},\mathbb{K}}$ are positive definite monotone gauges such that $\operatorname{maxq}_{\mathbb{D},\mathbb{K}} \leq \nu_{\mathbb{D},\mathbb{K}}$.

The fact that $\operatorname{fevc}_{\mathbb{D},\mathbb{K}}$ and $\nu^\circ_{\mathbb{D},\mathbb{K}}$ are gauges follows directly from (31b) and (33b). We now prove $\operatorname{fevc}_{\mathbb{D},\mathbb{K}}$ to be monotone. Let $P\colon \mathbb{S}^n \to \mathbb{S}^n$ denote the orthogonal projector onto $\operatorname{aff}(\mathbb{K})$. Let $Z_0, Z_1 \in \mathbb{K}^*$ be such that $Z_0 \preceq_{\mathbb{K}^*} Z_1$. Let $y \in \mathbb{R}_+^{\mathcal{F}(\mathbb{D})}$ be such that $\sum_{U \in \mathcal{F}(\mathbb{D})} y_U s_U s_U^\mathsf{T} \succeq_{\widehat{\mathbb{K}^*}} Z_1$. By (25),

$$P\biggl(\sum_{U \in \mathcal{F}(\mathbb{D})} y_U s_U s_U^\mathsf{T}\biggr) \succeq_{\mathbb{K}^*} P(Z_1) = Z_1 \succeq_{\mathbb{K}^*} Z_0 = P(Z_0).$$

By (25), there exists $\hat{Y} \in \widehat{\mathbb{K}^*}$ such that $P\Bigl(\sum_{U \in \mathcal{F}(\mathbb{D})} y_U s_U s_U^\mathsf{T} - Z_0\Bigr) = P(\widehat{Y})$. Since $\operatorname{Null}(P) = \operatorname{aff}(\mathbb{K})^\perp$, from (26) we conclude

$$\sum_{U \in \mathcal{F}(\mathbb{D})} y_U s_U s_U^\mathsf{T} - Z_0 - \widehat{Y} \in \operatorname{aff}(\mathbb{K})^\perp \subseteq \widehat{\mathbb{K}^*},$$

which ensures

$$\sum_{U \in \mathcal{F}(\mathbb{D})} y_U s_U s_U^\mathsf{T} \succeq_{\widehat{\mathbb{K}^*}} Z_0 + \widehat{Y} \succeq_{\widehat{\mathbb{K}^*}} Z_0.$$

Hence, by (31a), we have that $\operatorname{fevc}_{\mathbb{D},\mathbb{K}}(Z_0) \leq \operatorname{fevc}_{\mathbb{D},\mathbb{K}}(Z_1)$. Similarly, let $Y \in \mathbb{D}$ and $\mu \in \mathbb{R}_+$ be such that $\operatorname{diag}(Y) = \mu\mathbb{1}$ and $Y \succeq_{\widehat{\mathbb{K}^*}} Z_1$. Then (25) implies $P(Y) \succeq_{\mathbb{K}^*} Z_1 \succeq_{\mathbb{K}^*} Z_0 = P(Z_0)$, so $P(Y - Z_0) = P(\widehat{X})$ for some $\widehat{X} \in \widehat{\mathbb{K}^*}$ by (25). Hence $Y - Z_0 - \widehat{X} \in \operatorname{aff}(\mathbb{K})^\perp \subseteq \widehat{\mathbb{K}^*}$ by (26), and hence

$$Y \succeq_{\widehat{\mathbb{K}^*}} Z_0 + \widehat{X} \succeq_{\widehat{\mathbb{K}^*}} Z_0.$$

From (33a) we conclude $\nu^\circ_{\mathbb{D},\mathbb{K}}$ is monotone.

By (23), there exists $\mathring{X} \in \operatorname{ri}(\mathbb{K}) \setminus \{0\}$. Then $\nu_{\mathbb{D},\mathbb{K}}(\mathring{X}) > 0$, and hence we may assume $\nu_{\mathbb{D},\mathbb{K}}(\mathring{X}) = 1$. We claim that

$$\langle Z, \mathring{X} \rangle > 0 \text{ for every nonzero } Z \in \mathbb{K}^*. \tag{36}$$

Note that this implies via (33b) that $\nu^\circ_{\mathbb{D},\mathbb{K}}(Z) \geq \langle Z, \mathring{X} \rangle > 0$ for every nonzero $Z \in \mathbb{K}^*$, so $\nu^\circ_{\mathbb{D},\mathbb{K}}$ is positive definite. We now prove (36). Let $\hat{Y} \in \widehat{\mathbb{K}^*}$ and let $P\colon \mathbb{S}^n \to \mathbb{S}^n$ be the orthogonal projector onto $\operatorname{aff}(\mathbb{K})$. Assume that $P(\hat{Y}) \neq 0$. By (25), it suffices to prove $\langle P(\hat{Y}), \mathring{X} \rangle > 0$. Let $\varepsilon \in (0,1)$. Note that $\mathring{X} - \varepsilon P(\hat{Y}) \in \operatorname{aff}(\mathbb{K})$, since $(1-\varepsilon)^{-1}\mathring{X} \in \mathbb{K}$ and $-P(\hat{Y}) \in \operatorname{aff}(\mathbb{K})$, so

$$\mathring{X} - \varepsilon P(\hat{Y}) = (1-\varepsilon)\biggl(\frac{1}{1-\varepsilon}\mathring{X}\biggr) + \varepsilon(-P(\hat{Y})) \in \operatorname{aff}(\mathbb{K}).$$

Since $\mathring{X} \in \operatorname{ri}(\mathbb{K})$, there exists $\bar{\varepsilon} > 0$ such that $\mathring{X} - \bar{\varepsilon} P(\hat{Y}) \in \mathbb{K}$. Using that $\langle \mathring{X}, \hat{Y} \rangle = \langle P(\mathring{X}), \hat{Y} \rangle = \langle \mathring{X}, P(\hat{Y}) \rangle$ and $\langle P(\hat{Y}), \hat{Y} \rangle = \langle P(\hat{Y}), P(\hat{Y}) \rangle$, we use (25) to conclude

$$0 \leq \langle \mathring{X} - \bar{\varepsilon} P(\hat{Y}), \hat{Y} \rangle = \langle \mathring{X}, \hat{Y} \rangle - \bar{\varepsilon}\langle P(\hat{Y}), \hat{Y} \rangle = \langle \mathring{X}, P(\hat{Y}) \rangle - \bar{\varepsilon}\langle P(\hat{Y}), P(\hat{Y}) \rangle,$$

so (36) holds. Since

$$\begin{aligned}\operatorname{fevc}_{\mathbb{D},\mathbb{K}}(Z) &= \min\Big\{ \mathbb{1}^\mathsf{T} y : y \in \mathbb{R}_+^{\mathcal{F}(\mathbb{D})},\ \sum_{U\in\mathcal{F}(\mathbb{D})} y_U s_U s_U^\mathsf{T} \succeq_{\widehat{\mathbb{K}^*}} Z \Big\}\\ &\geq \min\Big\{ \mu \in \mathbb{R}_+ : Y \in \mathbb{D},\ \operatorname{diag}(Y) = \mu\mathbb{1},\ Y \succeq_{\widehat{\mathbb{K}^*}} Z \Big\} = \nu^\circ_{\mathbb{D},\mathbb{K}}(Z).\end{aligned}$$

it follows that $\operatorname{fevc}_{\mathbb{D},\mathbb{K}}$ is positive definite. Thus, $\operatorname{fevc}_{\mathbb{D},\mathbb{K}}$ and $\nu^\circ_{\mathbb{D},\mathbb{K}}$ are positive definite monotone gauges such that $\operatorname{fevc}_{\mathbb{D},\mathbb{K}} \geq \nu^\circ_{\mathbb{D},\mathbb{K}}$.

It is immediate from (31b) that $\operatorname{maxq}^\circ_{\mathbb{D},\mathbb{K}} = \operatorname{fevc}_{\mathbb{D},\mathbb{K}}$. We have that

$$\nu_{\mathbb{D},\mathbb{K}}(W) = \max\{ \langle W, Z\rangle : Z \in \mathbb{K}^*,\ \nu^\circ_{\mathbb{D},\mathbb{K}}(Z) \leq 1\} \text{ for every } W \in \mathbb{K}. \tag{37}$$

Moreover,

$$\begin{aligned}&\max\{ \langle W, Y\rangle : Y \in \mathbb{D},\ \operatorname{diag}(Y) = \mathbb{1}\}\\ &= \max\{ \langle W, P(Y)\rangle : Y \in \mathbb{D},\ \operatorname{diag}(Y) = \mathbb{1}\} && \text{since } W \in \mathbb{K} \subseteq \operatorname{Null}(\mathcal{L})\\ &\leq \max\{ \langle W, Z\rangle : Z \in \mathbb{K}^*,\ Y \in \mathbb{D},\ \operatorname{diag}(Y) = \mathbb{1},\ Y \succeq_{\widehat{\mathbb{K}^*}} Z\} && \text{by (29)}\\ &\leq \max\{ \langle W, Y\rangle : Y \in \mathbb{D},\ \operatorname{diag}(Y) = \mathbb{1}\}. && \text{since } W \in \mathbb{K} \subseteq \widehat{\mathbb{K}}\end{aligned}$$

Hence equality holds throughout, which implies (37) via (33a). ◻

## Appendix B. Polyhedral Cones

Let $\mathcal{A}\colon \mathbb{R}^d \to \mathbb{S}^n$ be a linear map, and set $\mathbb{K} \coloneqq \mathcal{A}(\mathbb{R}^d_+)$. We have that (22d) always hold. Indeed, we may write

$$\mathcal{A}(\mathbb{R}^d_+) = \{ X \in \mathbb{S}^n : \mathcal{L}(X) = 0,\ \mathcal{B}(X) \geq 0\}$$

with $\mathcal{L}\colon \mathbb{S}^n \to \mathbb{R}^k$ and $\mathcal{B}\colon \mathbb{S}^n \to \mathbb{R}^\ell$ linear transformations such that there exists $\mathring{X} \in \mathbb{S}^n$ such that $\mathcal{L}(X) = 0$ and $\mathcal{B}(X) > 0$. Since $\widehat{\mathbb{K}} \coloneqq \{ X \in \mathbb{S}^n : \mathcal{B}(X) \geq 0\}$ has nonempty interior, we have that (22d) holds. Note further that $\operatorname{Null}(\mathcal{L}) = \operatorname{Im}(\mathcal{A})$. We further have that

$$X \succeq_{\widehat{\mathbb{K}^*}} Y \text{ if and only if } \mathcal{A}^*(X) \geq \mathcal{A}^*(Y) \quad \text{for every } X, Y \in \mathbb{S}^n. \tag{38}$$

**Proposition 21.** Let $\mathbb{D}, \mathbb{K}$ be closed convex cones such that (22) holds, where $\mathbb{K} \coloneqq \mathcal{A}(\mathbb{R}^d_+)$ for a linear map $\mathcal{A}\colon \mathbb{R}^d \to \mathbb{S}^n$ such that $\mathcal{A}(e_i) \neq 0$ for each $i \in [d]$. If $w \in \mathbb{R}^d_+$ is such that $\mathcal{A}(w) = 0$, then $w = 0$.

*Proof.* If there exists nonzero $w \in \mathbb{R}^d_+$ and such that $\mathcal{A}(w) = 0$, then $\mathbb{K}$ is not pointed, contradicting (23). ◻

**Theorem 22.** Let $\mathbb{D}, \mathbb{K} \subseteq \mathbb{S}^n$ be closed convex cones such that (22) holds, where $\mathbb{K} \coloneqq \mathcal{A}(\mathbb{R}^d_+)$ for a linear map $\mathcal{A}\colon \mathbb{R}^d \to \mathbb{S}^n$ such that $\mathcal{A}(e_i) \neq 0$ for each $i \in [d]$. Then $\nu_{\mathbb{D},\mathcal{A}}\colon \mathbb{R}^d_+ \to \mathbb{R}$ defined by

$$\nu_{\mathbb{D},\mathcal{A}}(w) \coloneqq \nu_{\mathbb{D},\mathbb{K}}(\mathcal{A}(w)) \qquad \text{for every } w \in \mathbb{R}^d_+ \tag{39}$$

is a positive definite monotone gauge, and its dual is the positive definite monotone gauge

$$\begin{aligned}\nu^\circ_{\mathbb{D},\mathcal{A}}(z) &= \min\{ \nu^\circ_{\mathbb{D},\mathbb{K}}(Z) : Z \in \mathbb{K}^*,\ \mathcal{A}^*(Z) \geq z\}\\ &= \min\{ \mu : \mu \in \mathbb{R}_+,\ Y \in \mathbb{D},\ \operatorname{diag}(Y) = \mu\mathbb{1},\ \mathcal{A}^*(Y) \geq z\}\end{aligned} \tag{40}$$

for every $z \in \mathbb{R}^d_+$. Similarly, $\operatorname{maxq}_{\mathbb{D},\mathcal{A}}\colon \mathbb{R}^d_+ \to \mathbb{R}$ defined by

$$\operatorname{maxq}_{\mathbb{D},\mathcal{A}}(w) \coloneqq \operatorname{maxq}_{\mathbb{D},\mathbb{K}}(\mathcal{A}(w)) \qquad \text{for every } w \in \mathbb{R}^d_+ \tag{41}$$

is a positive definite monotone gauge, and its dual is the positive definite monotone gauge

$$\begin{aligned}\mathrm{fevc}_{\mathbb{D},\mathcal{A}}(z) &= \min\{\,\mathrm{fevc}_{\mathbb{D},\mathbb{K}}(Z) : Z \in \mathbb{K}^*,\, \mathcal{A}^*(Z) \geq z\} \\ &= \min\Big\{\,\mathbb{1}^\mathsf{T} y : y \in \mathbb{R}_+^{\mathcal{F}(\mathbb{D})},\, \sum_{U\in\mathcal{F}(\mathbb{D})} y_U \mathcal{A}^*(s_U s_U^\mathsf{T}) \geq z\Big\} \end{aligned} \tag{42}$$

for every $z \in \mathbb{R}_+^d$.

*Proof.* We first prove (39) to be a positive definite monotone gauge. As the composition of the gauge $\nu_{\mathbb{D},\mathbb{K}}$ with a linear function, it is immediate that $\nu_{\mathbb{D},\mathcal{A}}$ is a gauge. If $0 \leq w \leq v$, then $0 \preceq_{\mathbb{K}} \mathcal{A}(w) \preceq_{\mathbb{K}} \mathcal{A}(v)$, so monotonicity of $\nu_{\mathbb{D},\mathcal{A}}$ follows from the monotonicity part of Theorem 20, item (ii). Let $w \in \mathbb{R}_+^d$ be such that $\nu_{\mathbb{D},\mathcal{A}}(w) = 0$. Then $\nu_{\mathbb{D},\mathbb{K}}(\mathcal{A}(w)) = 0$, so Theorem 20 implies that $\mathcal{A}(w) = 0$. Hence $w = 0$ by Proposition 21. Thus $\nu_{\mathbb{D},\mathcal{A}}$ is a positive definite monotone gauge. Hence

$$\begin{aligned}\nu^\circ_{\mathbb{D},\mathcal{A}}(z) &= \max\{\,z^\mathsf{T} w : w \in \mathbb{R}_+^d,\, \nu_{\mathbb{D},\mathcal{A}}(w) \leq 1\} \\ &= \max\{\,z^\mathsf{T} w : w \in \mathbb{R}_+^d,\, \nu_{\mathbb{D},\mathbb{K}}(\mathcal{A}(w)) \leq 1\} \\ &= \max\{\,z^\mathsf{T} w : w \in \mathbb{R}_+^d,\, x \in \mathbb{R}^n,\, \mathrm{Diag}(x) \succeq_{\mathbb{D}^*} \mathcal{A}(w),\, \mathbb{1}^\mathsf{T} x \leq 1\} \\ &= \min\{\,\mu : \mu \in \mathbb{R}_+,\, Y \in \mathbb{D},\, \mathrm{diag}(Y) = \mu\mathbb{1},\, \mathcal{A}^*(Y) \geq z\}.\end{aligned}$$

Let $\alpha > 0$ be such that $\lambda_{\max}(\mathcal{A}(\alpha\mathbb{1})) < 1$. Then $(\mathring{w}, \mathring{x}) := (\frac{\alpha}{2n}\mathbb{1}, \frac{1}{2n}\mathbb{1})$ is strictly feasible in the second to last optimization problem, since $\mathrm{Diag}(\mathring{x}) - \mathcal{A}(\mathring{w}) = \frac{1}{2n}I - \frac{\alpha}{2n}\mathcal{A}(\mathbb{1}) \in \mathrm{int}(\mathbb{S}_+^n) \subseteq \mathrm{int}(\mathbb{D}^*)$ and $\mathbb{1}^\mathsf{T}\mathring{x} < 1$. For $\mathring{Y}$ as in (34), since $\mathring{Y} \in \mathrm{int}(\mathbb{D}) \subseteq \mathrm{int}(\widehat{\mathbb{K}^*})$, we have that $\mathcal{A}^*(Y) > 0$ from (38), and thus $\mathcal{A}^*(\alpha\mathring{Y}) - z = \alpha(\mathcal{A}^*(\mathring{Y}) - \frac{1}{\alpha}z) > 0$ for $\alpha \in \mathbb{R}_+$ big enough. Thus the last optimization problem is also strictly feasible. Hence

$$\begin{aligned}\nu^\circ_{\mathbb{D},\mathcal{A}}(z) &= \min\{\,\mu : \mu \in \mathbb{R}_+,\, Y \in \mathbb{D},\, \mathrm{diag}(Y) = \mu\mathbb{1},\, \mathcal{A}^*(Y) \geq z\} \\ &= \min\{\,\mu : \mu \in \mathbb{R}_+,\, Y \in \mathbb{D},\, Z \in \mathbb{K}^*,\, \mathrm{diag}(Y) = \mu\mathbb{1},\, Y \succeq_{\widehat{\mathbb{K}^*}} Z,\, \mathcal{A}^*(Z) \geq z\} \\ &= \min\{\,\nu^\circ_{\mathbb{D},\mathbb{K}}(Z) : Z \in \mathbb{K}^*,\, \mathcal{A}^*(Z) \geq z\}.\end{aligned}$$

The second equation holds because $(\mu, Y) \mapsto (\mu, Y, P(Y))$ and $(\mu, Y, Z) \mapsto (\mu, Y)$ map feasible solutions between both problems while preserving objective value by (29).

That (41) is a positive definite monotone gauge follows from Theorem 20 and Proposition 21 as above. Hence

$$\begin{aligned}&\mathrm{maxq}^\circ_{\mathbb{D},\mathcal{A}}(z) \\ &= \max\{\,z^\mathsf{T} w : w \in \mathbb{R}_+^d,\, \mathrm{maxq}_{\mathbb{D},\mathbb{K}}(\mathcal{A}(w)) \leq 1\} \\ &= \max\{\,z^\mathsf{T} w : w \in \mathbb{R}_+^d,\, \langle w, \mathcal{A}^*(s_U s_U^\mathsf{T})\rangle \leq 1 \text{ for every } U \in \mathcal{F}(\mathbb{D})\} \\ &= \min\Big\{\,\mathbb{1}^\mathsf{T} y : y \in \mathbb{R}_+^{\mathcal{F}(\mathbb{D})},\, \sum_{U\in\mathcal{F}(\mathbb{D})} y_U \mathcal{A}^*(s_U s_U^\mathsf{T}) \geq z\Big\} && \text{by LP Strong Duality} \\ &= \min\Big\{\,\mathbb{1}^\mathsf{T} y : Z \in \mathbb{K}^*,\, y \in \mathbb{R}_+^{\mathcal{F}(\mathbb{D})},\, \sum_{U\in\mathcal{F}(\mathbb{D})} y_U s_U s_U^\mathsf{T} \succeq_{\widehat{\mathbb{K}^*}} Z,\, \mathcal{A}^*(Z) \geq z\Big\} && \text{by (29) and (38)} \\ &= \min\{\,\mathrm{fevc}(Z) : Z \in \mathbb{K}^*,\, \mathcal{A}^*(Z) \geq z\}. && \text{by (31a)}\end{aligned}$$

□

## Appendix C. Concentration Results

In this section we prove the concentration results in Section 3. First we prove the results concerning the polyhedral case.

*Proof of Proposition 11.* Set $S \coloneqq \sum_{t\in[T]} X_t$. Also, define $x_t \coloneqq \mathcal{A}^*(X_t)$ for every $t \in [T]$, and set $s \coloneqq \mathcal{A}^*(S)$. We also denote $x \coloneqq \mathcal{A}^*(X)$, where $X$ is the random matrix in the statement. Then, by linearity of expectation,

$$\mathbb{E}[s] = T\mathbb{E}[x] = T\mathbb{E}[\mathcal{A}^*(X)] = T\mathcal{A}^*(\mathbb{E}[X]) \geq T\varepsilon\mathbb{1}.$$

Let $i \in [d]$. Chernoff's bound and the previous inequality imply that

$$\mathbb{P}(s_i \leq (1-\gamma)\mathbb{E}[s]_i) \leq \exp\Big(-\gamma^2\frac{\mathbb{E}[s]_i}{2}\Big) \leq \exp\Big(-\frac{\gamma^2\varepsilon}{2}T\Big).$$

Hence, by the union bound,

$$\begin{aligned}\mathbb{P}(\exists i \in [d],\, s_i \leq (1-\gamma)\mathbb{E}[s]_i) &\leq d\exp\Big(-\frac{\gamma^2\varepsilon}{2}T\Big)\\ &\leq \exp\Big(\log(d) - \frac{\gamma^2\varepsilon}{2}\frac{2(\log(d)+\log(n))}{\gamma^2\varepsilon}\Big) = \frac{1}{n}.\end{aligned}$$

Thus with probability at least $1-1/n$ we have that $s \geq (1-\gamma)\mathbb{E}[s]$. By (38), this event holds if and only if $S \succeq_{\widehat{\mathbb{K}^*}} (1-\gamma)\mathbb{E}[S]$. ☐

*Proof of Proposition 12.* Both (10) and (4) imply that $\mathcal{A}^*(\mathbb{E}[\Xi_Y]) \geq \alpha_\Xi\mathcal{A}^*(Y) \geq \alpha_\Xi\varepsilon\mathbb{1}$. Proposition 11 implies that, with probability at least $1-1/n$,

$$\frac{1}{T}\sum_{t\in[T]}(\Xi_Y)_t \succeq_{\widehat{\mathbb{K}^*}} (1-\gamma)\mathbb{E}[\Xi_Y] \succeq_{\widehat{\mathbb{K}^*}} (1-\gamma)\alpha_\Xi Y.$$

Hence $y \in \mathbb{R}_+^{\mathcal{F}(\mathbb{D})}$ defined by $(1-\gamma)\alpha_\Xi \cdot y_U \coloneqq \frac{1}{T}|\{\, t\in[T] : (\Xi_Y)_t = s_U s_U^\mathsf{T}\}|$ for every $U \in \mathcal{F}(\mathbb{D})$ satisfies the desired properties. ☐

One case we treat separately is when $\mathbb{K} = \mathbb{S}^n_+$. For this case, we use the following result by Tropp:

**Proposition 23** (see [Tro12, Theorem 1.1])**.** Let $\{\, X_t : t \in T\}$ be independent random matrices in $\mathbb{S}^n$. Let $\rho \in \mathbb{R}$ be such that

$$0 \preceq X_t \preceq \rho I \text{ almost surely for every } t \in T.$$

Set $S \coloneqq \sum_{t\in T} X_t$. Then for every $\gamma \in (0,1)$,

$$\mathbb{P}(\lambda_{\min}(S) \leq (1-\gamma)\lambda_{\min}(\mathbb{E}S)) \leq n\exp\Big(-\frac{\gamma^2}{2}\frac{\lambda_{\min}(\mathbb{E}S)}{\rho}\Big).$$

Proposition 23 weakens the upper bound from [Tro12, Theorem 1.1] using that

$$\frac{\exp(-\gamma)}{(1-\gamma)^{1-\gamma}} \leq \exp\Big(-\frac{\gamma^2}{2}\Big),$$

which follows from

$$\Big(1-\frac{\gamma}{2}\Big)\frac{\gamma}{1-\gamma} = \gamma + \sum_{k=2}^\infty \frac{\gamma^k}{2} \geq \gamma + \sum_{k=2}^\infty\frac{\gamma^k}{k} = \log\Big(\frac{1}{1-\gamma}\Big).$$

We prove the following result.

**Proposition 24.** Let $\gamma \in (0,1)$, let $\tau, \rho \in \mathbb{R}_+$, and let $\bar{Y} \in \mathbb{S}^n_{++}$. Let $X \colon \Omega \to \mathbb{S}^n$ be a random matrix such that

$$0 \preceq X \preceq \rho\bar{Y} \text{ almost surely, and } \tau\bar{Y} \preceq \mathbb{E}[X].$$

Let $(X_t)_{t\in[T]}$ be independent identically distributed random variables sampled from $X$, for any

$$T \geq \Big\lceil \frac{4\rho}{\gamma^2\tau}\log(2n)\Big\rceil. \tag{43}$$

Then, with probability at least $1-1/2n$,

$$\frac{1}{T}\sum_{t\in[T]} X_t \succeq (1-\gamma)\tau\bar{Y}.$$

*Proof.* For every $t\in[T]$, set $Y_t := \bar{Y}^{-1/2}X_t\bar{Y}^{-1/2}$. Then

$$0 \preceq Y_t = \bar{Y}^{-1/2}X_t\bar{Y}^{-1/2} \preceq \rho\bar{Y}^{-1/2}\bar{Y}\bar{Y}^{-1/2} = \rho I$$

for every $t\in[T]$ almost surely. Set $Q := \sum_{t\in[T]} Y_t$. Since $\mathbb{E}[X]\succeq\tau\bar{Y}$,

$$\mathbb{E}[Q] = \mathbb{E}\bigg[\sum_{t\in[T]} Y_t\bigg] = \sum_{t\in[T]} \bar{Y}^{-1/2}\mathbb{E}[X_t]\bar{Y}^{-1/2} \succeq T\tau I,$$

which implies that $\lambda_{\min}(\mathbb{E}[Q]) \geq T\tau$. Hence

$$\begin{aligned}
1-\mathbb{P}\bigg(\frac{1}{T}\sum_{t\in[T]} X_t \succeq (1-\gamma)\tau\bar{Y}\bigg) &= 1-\mathbb{P}(Q\succeq T(1-\gamma)\tau I)\\
&\leq 1-\mathbb{P}(Q\succeq(1-\gamma)\lambda_{\min}(\mathbb{E}[Q])I) && \text{as } \lambda_{\min}(\mathbb{E}[Q])\geq T\tau\\
&\leq \mathbb{P}(\lambda_{\min}(Q)\leq(1-\gamma)\lambda_{\min}(\mathbb{E}[Q]))\\
&\leq 2n\exp\bigg(-\gamma^2\frac{\lambda_{\min}(\mathbb{E}[Q])}{2\rho}\bigg) && \text{by Proposition 23}\\
&\leq 2n\exp\bigg(-\frac{\gamma^2\tau}{2\rho}T\bigg) && \text{as } \lambda_{\min}(\mathbb{E}[Q])\geq T\tau\\
&\leq \exp\bigg(\log(2n)-\frac{\gamma^2\tau}{2\rho}\frac{4\rho\log(2n)}{\gamma^2\tau}\bigg) = \frac{1}{2n}. && \text{by (43)}
\end{aligned}$$

□

For the more general case when $\mathbb{K}\subseteq\mathbb{S}^n_+$, we use the following result by Tropp which requires a bound on the spectral norm of the random matrix and uses its second moment:

**Theorem 25** (see [Tro15, Corollary 6.2.1])**.** Let $T\in\mathbb{N}$ be nonzero. Let $X$ be a random matrix in $\mathbb{S}^n$ such that $\|X\|\leq\rho$ almost surely. Let $(X_t)_{t\in[T]}$ be i.i.d. random variables sampled from $X$. Set $\sigma^2 := \|\mathbb{E}[X^2]\|$, set $M := \mathbb{E}[X]$, and set

$$E := \frac{1}{T}\sum_{t=1}^{T} X_t.$$

Then for all $\gamma\geq 0$,

$$\mathbb{P}(\|E-M\|\geq\gamma)\leq 2n\exp\bigg(-T\frac{\gamma^2/2}{\sigma^2+2\rho\gamma/3}\bigg).$$

Using Theorem 25, we are ready to prove our general concentration result.

*Proof of Proposition 13.* If $\sigma^2\geq(2/3)\rho\gamma$, then

$$-\frac{T}{2}\frac{\gamma^2}{\sigma^2+(2/3)\gamma\rho} \leq -\frac{T}{4}\frac{\gamma^2}{\sigma^2} \leq -\frac{8\sigma^2\log(2n)}{\gamma^2}\frac{\gamma^2}{4\sigma^2} = -2\log(2n).$$

On the other hand, if $\sigma^2\leq(2/3)\rho\gamma$, then

$$-\frac{T}{2}\frac{\gamma^2}{\sigma^2+(2/3)\gamma\rho} \leq -\frac{T}{4}\frac{\gamma^2}{(2/3)\gamma\rho} \leq -\frac{16\rho\log(2n)}{3\gamma}\frac{3\gamma}{8\rho} = -2\log(2n).$$

Theorem 25 implies

$$\mathbb{P}\bigg(\bigg\|\frac{1}{T}\sum_{t\in[T]} X_t-\mathbb{E}[X]\bigg\|\geq\gamma\bigg)\leq 2n\exp\bigg(-\frac{T}{2}\frac{\gamma^2}{\sigma^2+(2/3)\rho\gamma}\bigg)\leq 2n\exp(-2\log(2n)) = \frac{1}{2n}.$$

□

Finally, we prove Proposition 15 by combining Proposition 13 with Proposition 14.

*Proof of Proposition 15.* Set $Z \coloneqq Y^{-1/2}\Xi_Y Y^{-1/2}$. For every $U \subseteq [n]$,

$$\|Y^{-1/2}s_U s_U^\mathsf{T} Y^{-1/2}\| = \langle s_U s_U^\mathsf{T}, Y^{-1}\rangle \quad \text{and} \quad \|(Y^{-1/2}s_U s_U^\mathsf{T} Y^{-1/2})^2\| = \langle s_U s_U^\mathsf{T}, Y^{-1}\rangle^2.$$

Thus

$$\operatorname{maxq}(Y^{-1}) \geq \|Z\| \text{ and } \operatorname{maxq}(Y^{-1})^2 I \succeq Z^2 \text{ almost surely.}$$

Set $\sigma^2 \coloneqq (n/\varepsilon)^2$ and $\rho \coloneqq n/\varepsilon$. As $Y \succeq \varepsilon I$, we have that $Y^{-1} \preceq (1/\varepsilon)I$, so $\operatorname{maxq}(Y^{-1}) \leq n/\varepsilon$. Since $n \geq 1 \geq (2/3)\gamma\alpha_\Xi\varepsilon$, we have that

$$T \geq \left\lceil \frac{8}{(\gamma\varepsilon\alpha_\Xi)^2} n^2 \log(2n) \right\rceil \geq \frac{8}{(\gamma\alpha_\Xi)^2}\frac{n^2}{\varepsilon^2}\log(2n) = \frac{8\sigma^2\log(2n)}{(\gamma\alpha_\Xi)^2} \geq \frac{16}{3}\frac{n}{\varepsilon}\frac{1}{\gamma\alpha_\Xi}\log(2n) = \frac{16\rho\log(2n)}{3\gamma\alpha_\Xi}.$$

Let $\{\, Z_t : t \in [T] \}$ be i.i.d. random variables sampled from $Z$. Proposition 13 implies that

$$Y^{-1/2}\mathbb{E}[\Xi_Y]Y^{-1/2} - \gamma\alpha_\Xi I \preceq Y^{-1/2}\Big(\frac{1}{T}\sum_{t\in[T]} Z_t\Big)Y^{-1/2}$$

with probability at least $1 - 1/(2n)$. Assume that this event holds. Let $y \in \mathbb{R}_+^{\mathcal{F}(\mathbb{D})}$ be defined by $y_U \coloneqq \frac{1}{T}|\{\, t \in [T] : Z_t = s_U s_U^\mathsf{T} \}|$. Then $\mathbb{1}^\mathsf{T} y = 1$ and

$$\sum_{U\in\mathcal{F}(\mathbb{D})} y_U s_U s_U^\mathsf{T} = \frac{1}{T}\sum_{t\in[T]} Z_t \succeq \mathbb{E}[\Xi_Y] - \alpha_\Xi\gamma Y.$$

Proposition 14 implies we can compute in polynomial time $\tilde{y} \in \mathbb{R}_+^{\mathcal{F}(\mathbb{D})}$ with support size $|\operatorname{supp}(\tilde{y})| \in O(n/\zeta^2)$, such that $\mathbb{1}^\mathsf{T}\tilde{y} \leq 1 + \zeta$, and $\sum_{U\in\mathcal{F}(\mathbb{D})} \tilde{y}_U s_U s_U^\mathsf{T} \succeq \mathbb{E}[\Xi_Y] - \alpha_\Xi\gamma Y$. As $\mathbb{K} \subseteq \mathbb{S}_+^n$, we have that $\mathbb{S}_+^n \subseteq \widehat{\mathbb{K}^*}$, so by (4) we obtain

$$\sum_{U\in\mathcal{F}(\mathbb{D})} \tilde{y}_U s_U s_U^\mathsf{T} \succeq_{\widehat{\mathbb{K}^*}} \mathbb{E}[\Xi_Y] - \gamma\alpha_\Xi Y \succeq_{\widehat{\mathbb{K}^*}} \alpha_\Xi Y - \gamma\alpha_\Xi Y = (1-\gamma)\alpha_\Xi Y. \qquad \square$$

## Appendix D. Algorithmic Simultaneous Certificates

Let $\mathbb{D}, \mathbb{K} \subseteq \mathbb{S}^n$ be cones such that (22) holds. Let $\varepsilon \in (0,1)$. Set, for every $W \in \mathbb{K}$,

$$\nu_{\varepsilon,\mathbb{D},\mathbb{K}}(W) \coloneqq (1-\varepsilon)\nu_{\mathbb{D},\mathbb{K}}(W) + \varepsilon\langle I, W\rangle \tag{44a}$$

$$= \min\{\, \rho : \rho \in \mathbb{R}_+,\ x \in \mathbb{R}^n,\ \rho \geq (1-\varepsilon)\mathbb{1}^\mathsf{T} x + \varepsilon\langle I, W\rangle,\ \operatorname{Diag}(x) \succeq_{\mathbb{D}^*} W \}. \tag{44b}$$

For every $Z \in \mathbb{K}^*$, set

$$\nu^\circ_{\varepsilon,\mathbb{D},\mathbb{K}}(Z) = \max\{\, \langle Z, W\rangle : x \in \mathbb{R}^n,\ W \in \mathbb{K},\ W \preceq_{\mathbb{D}^*} \operatorname{Diag}(x),\ (1-\varepsilon)\mathbb{1}^\mathsf{T} x + \varepsilon\langle I, W\rangle \leq 1 \} \tag{45a}$$

$$= \min\{\, \mu : \mu \in \mathbb{R}_+,\ Y \in \mathbb{S}^n,\ Y \succeq_{\mathbb{D}} \mu\varepsilon I,\ Y \succeq_{\widehat{\mathbb{K}^*}} Z,\ \operatorname{diag}(Y) = \mu\mathbb{1} \}. \tag{45b}$$

One may check that $\nu_{\varepsilon,\mathbb{D},\mathbb{K}}$ and $\nu^\circ_{\varepsilon,\mathbb{D},\mathbb{K}}$ are positive definite monotone gauges, dual to each other. Let $\sigma \in (0,1)$ and set

$$\mathrm{H}_{\varepsilon,\sigma}(\mathbb{D},\mathbb{K}) \coloneqq \left\{ (W,Z) \in \mathbb{K}\times\mathbb{K}^* : \begin{array}{r} \exists(\mu,Y) \text{ feasible for (45b) for } Z, \\ \exists(\rho,x) \text{ feasible for (44b) for } W, \\ \text{and } \langle W,Z\rangle \geq (1-\sigma)\rho\mu \end{array} \right\}. \tag{46}$$

One may check that

$$\begin{array}{c} \text{if } (\rho,x) \text{ and } (\mu,Y) \text{ witness the membership } (W,Z) \in \mathrm{H}_{\varepsilon,\sigma}(\mathbb{D},\mathbb{K}), \text{ then} \\ (1-\sigma)\rho\mu \leq \langle W,Z\rangle \leq \rho\mu. \end{array} \tag{47}$$

**Theorem 26.** Let $\varepsilon \in [0,1)$. Then,

$$(1-\varepsilon)\nu(W) \leq \nu_\varepsilon(W) \leq \nu(W), \qquad \text{for each } W \in \mathbb{K}, \tag{48a}$$

$$\nu^\circ(Z) \leq \nu^\circ_\varepsilon(Z) \leq \frac{1}{1-\varepsilon}\nu^\circ(Z), \qquad \text{for each } Z \in \mathbb{K}^*. \tag{48b}$$

*Proof.* We have that $(1-\varepsilon)\nu(W) \leq \nu_\varepsilon(W)$ since $\nu_\varepsilon(W) = (1-\varepsilon)\nu(W) + \varepsilon\langle I, W\rangle$ and $\langle I, W\rangle \geq 0$ by (16) since $W \in \mathbb{K} \subseteq \mathbb{D}^*$.

We have that $\nu(W) \geq \langle I, W\rangle$ by (16). Therefore, $\nu_\varepsilon(W) = (1-\varepsilon)\nu(W) + \varepsilon \operatorname{Tr}(W) \leq \nu(W)$. Equation (48b) holds by duality. □

Let $\mathbb{D}, \mathbb{K} \subseteq \mathbb{S}^n$ be closed convex cones such that (3) holds, where $\mathbb{K}$ is the polyhedral cone defined by $\mathcal{A}\colon \mathbb{R}^d \to \mathbb{S}^n$. Assume that (16) holds. Set

$$\nu_{\varepsilon,\mathbb{D},\mathcal{A}}(w) \coloneqq \nu_{\varepsilon,\mathbb{D},\mathbb{K}}(\mathcal{A}(w)) \text{ for every } w \in \mathbb{R}^d_+. \tag{49}$$

Then, for every $z \in \mathbb{R}^d_+$,

$$\begin{aligned} \nu^\circ_{\varepsilon,\mathbb{K},\mathcal{A}}(z) &= \min\{\, \mu \in \mathbb{R}_+ : Y \succeq_{\mathbb{D}} \varepsilon\mu I,\, \operatorname{diag}(Y) = \mu\mathbb{1},\, \mathcal{A}^*(Y) \geq z\} \\ &= \max\{\, z^\mathsf{T} w : w \in \mathbb{R}^d_+,\, x \in \mathbb{R}^n,\, (1-\varepsilon)\mathbb{1}^\mathsf{T} x + \varepsilon \operatorname{Tr}(\mathcal{A}(w)) \leq 1,\, \operatorname{Diag}(x) \succeq_{\mathbb{D}^*} \mathcal{A}(w)\}. \end{aligned} \tag{50}$$

Let $(\mathring{\mu}, \mathring{Y})$ be as in (34). Since $I \in \mathbb{D} \subseteq \mathbb{K}^*$ by (16) and (22), it follows from (10) that $\mathcal{A}^*(I) \geq 0$. Hence

$$\mathcal{A}^*((1-\varepsilon)\mathring{Y} + \varepsilon\mathring{\mu} I) \geq (1-\varepsilon)\mathcal{A}^*(\mathring{Y}) > 0.$$

We thus conclude that $(1-\varepsilon)\mathring{Y} + \varepsilon\mathring{\mu} I$ is strictly feasible in the first optimization problem in (50). The second problem in (50) is also strictly feasible, as one can see by setting $(\mathring{w}, \mathring{x}) \coloneqq (\frac{\alpha}{3n}\mathbb{1}, \frac{1}{3n}\mathbb{1})$ for $\alpha \in \mathbb{R}_{++}$ such that $\lambda_{\max}(\mathcal{A}(\alpha\mathbb{1})) < 1$. Hence the second equality in (50) and attainment of both problems follow from Strong Duality. For the first equality in (50), note that

$$\begin{aligned} &\nu^\circ_\varepsilon(z) \\ &= \min\{\, \nu^\circ_\varepsilon(Z) : Z \in \mathbb{K}^*,\, \mathcal{A}^*(Z) \geq z\} \\ &= \min\{\, \mu \in \mathbb{R}_+ : Z \in \mathbb{K}^*,\, \mathcal{A}^*(Z) \geq z,\, Y \in \mathbb{S}^n,\, Y \succeq_{\mathbb{D}} \varepsilon\mu I,\, Y \succeq_{\widehat{\mathbb{K}^*}} Z,\, \operatorname{diag}(Y) = \mu\mathbb{1}\} \quad \text{by (45b)} \\ &\geq \min\{\, \mu \in \mathbb{R}_+ : Y \in \mathbb{S}^n,\, Y \succeq_{\mathbb{D}} \varepsilon\mu I,\, \operatorname{diag}(Y) = \mu\mathbb{1},\, \mathcal{A}^*(Y) \geq z\}, \end{aligned}$$

as $Y \succeq_{\widehat{\mathbb{K}^*}} Z$ implies $\mathcal{A}^*(Y) \geq \mathcal{A}^*(Z)$ by (38). Equality follows from (29) and $\mathcal{A}^*(Y) = \mathcal{A}^*(P(Y))$ for every $Y \in \mathbb{S}^n$, where $P\colon \mathbb{S}^n \to \mathbb{S}^n$ is the orthogonal projector on $\operatorname{aff}(\mathbb{K})$.

**Proposition 27.** Let $\mathbb{D}, \mathbb{K} \subseteq \mathbb{S}^n$ be closed convex cones such that (3) holds, where $\mathbb{K} \coloneqq \mathcal{A}(\mathbb{R}^d_+)$ for a linear map $\mathcal{A}\colon \mathbb{R}^d \to \mathbb{S}^n$. Assume that (16) holds. Let $z \in \mathbb{R}^d_+$. If $(\mu, Y)$ and $(w, x)$ are feasible solutions to (50) such that

$$(1-\sigma)\mu \leq z^\mathsf{T} w \leq \mu,$$

then

$$(1, x) \text{ and } (\mu, Y) \text{ witness the membership } (\mathcal{A}(w), Y) \in \mathrm{H}_{\varepsilon,\sigma}(\mathbb{D}, \mathbb{K}).$$

*Proof.* It is immediate that $(1, x)$ is feasible in (44b) for $W \coloneqq \mathcal{A}(w)$. It is also clear that $(\mu, Y)$ is feasible in (45b) for $Z \coloneqq Y$. The proof follows from

$$(1-\sigma)\mu \leq w^\mathsf{T} z \leq \langle w, \mathcal{A}^*(Y)\rangle = \langle \mathcal{A}(w), Y\rangle.$$ □

**Proposition 28.** Let $\varepsilon, \sigma, \zeta \in [0,1)$. Let $\mathbb{D}, \mathbb{K} \subseteq \mathbb{S}^n$ be such that (3) holds. Let $W, Z \in \mathbb{S}^n$ be nonzero. Let $(\bar{\rho}, x)$ and $(\bar{\mu}, Y)$ witness the membership $(W, Z) \in \mathrm{H}_{\varepsilon,\sigma}(\mathbb{D}, \mathbb{K})$. Set $\rho \coloneqq (1-\varepsilon)^{-1}\bar{\rho}$ and $\mu \coloneqq \rho^{-1}\langle Z, W\rangle$. Let $p \in \mathbb{R}^{\mathcal{F}(\mathbb{D})}_+$ be such that $\mathbb{1}^\mathsf{T} p \leq 1 + \zeta$ and

$$\sum_{U \in \mathcal{F}(\mathbb{D})} p_U s_U s_U^\mathsf{T} \succeq_{\widehat{\mathbb{K}^*}} \frac{1}{\bar{\mu}} Y.$$

Set $\beta \coloneqq (1-\varepsilon)(1-\sigma)/(1+\zeta)$. Then

$$s_V^{\mathsf{T}} W s_V \geq \beta\rho \text{ and } Z \preceq_{\widehat{\mathbb{K}^*}} \frac{1}{\beta} \sum_{U \in \mathcal{F}(\mathbb{D})} p_U s_U s_U^{\mathsf{T}}$$

for $V \coloneqq \arg\max\{\, s^{\mathsf{T}} W s_U : U \in \operatorname{supp}(p)\}$.

*Proof.* Set $V \coloneqq \arg\max\{\, s_U^{\mathsf{T}} W s_U : U \in \operatorname{supp}(p)\}$. Then

$$\begin{aligned}
\langle s_V s_V^{\mathsf{T}}, W\rangle &\geq \sum_{U \in \mathcal{F}(\mathbb{D})} \frac{p_U}{\mathbb{1}^{\mathsf{T}} p} \langle s_U s_U^{\mathsf{T}}, W\rangle \\
&\geq \frac{1}{\bar{\mu}\mathbb{1}^{\mathsf{T}} p} \langle Y, W\rangle \\
&\geq \frac{1}{\bar{\mu}\mathbb{1}^{\mathsf{T}} p} \langle Z, W\rangle && \text{by (45b)} \\
&\geq \frac{1-\sigma}{\mathbb{1}^{\mathsf{T}} p} \bar{\rho} && \text{by (46)} \\
&= \frac{(1-\sigma)(1-\varepsilon)}{\mathbb{1}^{\mathsf{T}} p} \rho \\
&\geq \frac{(1-\sigma)(1-\varepsilon)}{1+\zeta} \rho.
\end{aligned}$$

Moreover,

$$\mu = \rho^{-1}\langle Z, W\rangle \geq \rho^{-1}(1-\sigma)\bar{\rho}\bar{\mu} = (1-\sigma)(1-\varepsilon)\bar{\mu}.$$

Hence

$$\begin{aligned}
Z &\preceq_{\widehat{\mathbb{K}^*}} Y && \text{by (45b)} \\
&\preceq_{\widehat{\mathbb{K}^*}} \bar{\mu} \sum_{U \in \mathcal{F}(\mathbb{D})} p_U s_U s_U^{\mathsf{T}} \\
&\preceq_{\widehat{\mathbb{K}^*}} \frac{\mu}{(1-\sigma)(1-\varepsilon)} \sum_{U \in \mathcal{F}(\mathbb{D})} p_U s_U s_U^{\mathsf{T}} \\
&\preceq_{\widehat{\mathbb{K}^*}} \mu \frac{1+\zeta}{(1-\sigma)(1-\varepsilon)} \sum_{U \in \mathcal{F}(\mathbb{D})} p_U s_U s_U^{\mathsf{T}}.
\end{aligned}$$

□

**Proposition 29.** Let $\varepsilon, \sigma, \gamma \in (0,1)$. Let $\mathbb{D}, \mathbb{K} \subseteq \mathbb{S}^n$ be closed convex cones such that (3) holds, where $\mathbb{K} \coloneqq \mathcal{A}(\mathbb{R}^d_+)$ for a linear map $\mathcal{A}\colon \mathbb{R}^d \to \mathbb{S}^n$. Assume (16) and that

(51) $$\mathcal{A}^*(I) \geq \kappa \mathbb{1},$$

for some $\kappa \in \mathbb{R}_{++}$. Let $\Xi$ be a randomized rounding algorithm for $\mathbb{D}$. Set $\beta \coloneqq \alpha_\Xi(1-\gamma)(1-\sigma)(1-\varepsilon)$. Let $(W, Z) \in \mathrm{H}_{\varepsilon,\sigma}(\mathbb{D}, \mathbb{K})$ be such that $W \neq 0 \neq Z$. Let $(\bar{\rho}, x)$ and $(\bar{\mu}, Y)$ witness the membership $(W, Z) \in \mathrm{H}_{\varepsilon,\sigma}(\mathbb{D}, \mathbb{K})$. There exists a randomized polynomial-time algorithm that takes $(\bar{\rho}, x)$ and $(\bar{\mu}, Y)$ as input and outputs a $\beta$-certificate $(\rho, \mu, y, U, x)$ for $(W, Z)$ with high probability, and such that

$$|\operatorname{supp}(y)| \leq \left\lceil \frac{2(\log(d) + \log(n))}{\kappa\varepsilon\alpha_\Xi\gamma^2} \right\rceil \text{ almost surely.}$$

In particular, $(W, Z)$ is a $\beta$-pairing.

*Proof.* Note that $\bar{\rho}, \bar{\mu} > 0$ as $W \neq 0 \neq Z$. Set $\rho \coloneqq (1-\varepsilon)^{-1}\bar{\rho}$ and $\mu \coloneqq (1/\rho)\langle W, Z\rangle$. Note that (15.i) holds trivially. We also have (15.iv), since $\operatorname{Diag}(x) \succeq_{\mathbb{D}^*} W$ and

$$\rho = \frac{\bar{\rho}}{1-\varepsilon} \geq \frac{1}{1-\varepsilon}\Big((1-\varepsilon)\mathbb{1}^{\mathsf{T}} x + \varepsilon \operatorname{Tr}(W)\Big) \geq \mathbb{1}^{\mathsf{T}} x$$

as $I \in \mathbb{D}$ by (16) and $\mathbb{D} \subseteq \widehat{\mathbb{K}^*}$ by (27).

We now prove (15.iii). Set $\bar{Y} := \bar{\mu}^{-1} Y$. Since $(\bar{\mu}, Y)$ is feasible in (45b) and $\mathbb{D} \subseteq \widehat{\mathbb{K}^*}$, we have that $\bar{Y} \succeq_{\widehat{\mathbb{K}^*}} \varepsilon I$. Thus $\mathcal{A}^*(\bar{Y}) \geq \varepsilon \mathcal{A}^*(I) \geq \varepsilon\kappa \mathbb{1}$ by (10) and (51). Proposition 12 ensures that one can compute $\bar{y} \in \mathbb{R}_+^{\mathcal{F}(\mathbb{D})}$ such that, with probability at least $1 - 1/n$,

$$\frac{1}{(1-\gamma)} \frac{1}{\alpha_\Xi} \sum_{U \in \mathcal{F}(\mathbb{D})} \bar{y}_U s_U s_U^\mathsf{T} \succeq_{\widehat{\mathbb{K}^*}} \frac{1}{\bar{\mu}} Y. \tag{52}$$

Setting $p := (\alpha_\Xi(1-\gamma))^{-1}\bar{y}$ and $\zeta := 0$, Proposition 28 finishes the proof. □

*Proof of Theorem 16.* Set $\tau := 1 - \beta/\alpha_\Xi$, and $\sigma := \gamma := \varepsilon := \tau/3$. If we are given $w \in \mathbb{R}_+^d$ as input, nearly solve (44a) with $W := \mathcal{A}(w)$ and set $z := \mathcal{A}^*(Y)$ and $Z := Y$. If we are given $z \in \mathbb{R}_+^d$, nearly solve (50) and set $Z := Y$. In both cases, we obtain $(\mathcal{A}(w), Z) \in \mathrm{H}_{\varepsilon,\sigma}(\mathbb{D}, \mathbb{K})$ such that $\mathcal{A}^*(Z) \geq z$, as well as the appropriate witnesses of this membership. By definition, it suffices to obtain a $\beta$-certificate for $(\mathcal{A}(w), Z)$. Proposition 29 implies one can compute, in polynomial time, a $\hat{\beta}$-certificate $(\rho, \mu, U, y, x)$ for $(\mathcal{A}(w), Z)$ with $\hat{\beta} = \alpha_\Xi(1-\gamma)(1-\sigma)(1-\varepsilon)$. Since $0 < \tau < 1 \leq 9$, we have that

$$\hat{\beta} = \alpha_\Xi\left(1 - \frac{\tau}{3}\right)^3 = \alpha_\Xi\left(1 - \tau + \frac{1}{3}\tau^2 - \frac{1}{27}\tau^3\right) \geq \alpha_\Xi(1-\tau) = \beta.$$

This implies that $(\rho, \mu, U, y, x)$ is a $\beta$-certificate. By Proposition 29, we have that

$$|\mathrm{supp}(y)| \leq \left\lceil \frac{2}{\kappa\varepsilon\alpha_\Xi\gamma^2}(\log(d) + \log(n)) \right\rceil = \left\lceil \frac{54}{\kappa\alpha_\Xi\tau^3}(\log(d) + \log(n)) \right\rceil. \quad \square$$

**Proposition 30.** Let $\varepsilon, \sigma, \gamma, \zeta \in (0,1)$. Let $\mathbb{D}, \mathbb{K} \subseteq \mathbb{S}^n$ be such that (3) holds, and assume that $\mathbb{K} \subseteq \mathbb{S}_+^n$. Let $\Xi$ be a randomized rounding algorithm for $\mathbb{D}$. Set $\beta := \alpha_\Xi(1-\gamma)(1-\sigma)(1-\varepsilon)/(1+\zeta)$. Let $(W, Z) \in \mathrm{H}_{\varepsilon,\sigma}(\mathbb{D}, \mathbb{K})$ be such that $W \neq 0 \neq Z$. Let $(\bar{\rho}, x)$ and $(\bar{\mu}, Y)$ witness the membership $(W, Z) \in \mathrm{H}_{\varepsilon,\sigma}(\mathbb{D}, \mathbb{K})$. There exists a polynomial time algorithm that takes $(\bar{\rho}, x)$ and $(\bar{\mu}, Y)$ as input and outputs a $\beta$-certificate $(\rho, \mu, y, U, x)$ with high probability. Almost surely, we have that $|\mathrm{supp}(y)| \in O(n/\zeta^2)$ and that the algorithm takes at most

$$\left\lceil \frac{8n^2\log(2n)}{(\varepsilon\alpha_\Xi\gamma)^2} \right\rceil,$$

samples from $\Xi_{\bar{\mu}^{-1}Y}$. In particular, $(W, Z)$ is a $\beta$-pairing.

*Proof.* Let $(W, Z) \in \mathrm{H}_{\varepsilon,\sigma}(\mathbb{D}, \mathbb{K})$. Let $(\bar{\rho}, x)$ and $(\bar{\mu}, Y)$ witness the membership $(W, Z) \in \mathrm{H}_{\varepsilon,\sigma}(\mathbb{D}, \mathbb{K})$. Note that $\bar{\rho}, \bar{\mu} > 0$ as $W \neq 0 \neq Z$. Set $\rho := (1-\varepsilon)^{-1}\bar{\rho}$ and $\mu := \rho^{-1}\langle Z, W\rangle$. Then (15.i) holds trivially. We also have (15.iv), since $\mathrm{Diag}(x) \succeq_{\mathbb{D}^*} W$ and

$$\rho = \frac{\bar{\rho}}{1-\varepsilon} \geq \frac{1}{1-\varepsilon}((1-\varepsilon)\mathbb{1}^\mathsf{T} x + \varepsilon \operatorname{Tr}(W)) \geq \mathbb{1}^\mathsf{T} x,$$

as $I \in \mathbb{D}$ by (16) and $\mathbb{D} \subseteq \mathbb{K}^*$ by (3).

We now prove (15.iii). Set $\bar{Y} := \bar{\mu}^{-1} Y$. Since $(\bar{\mu}, Y)$ is feasible in (45b) and $\mathbb{D} \subseteq \widehat{\mathbb{K}^*}$, we have that $\bar{Y} \succeq_{\widehat{\mathbb{K}^*}} \varepsilon I$. Proposition 15 ensures one can compute, in polynomial time, $\bar{y} \in \mathbb{R}_+^{\mathcal{F}(\mathbb{D})}$ such that $\mathbb{1}^\mathsf{T}\bar{y} \leq 1 + \zeta$ and $|\mathrm{supp}(\bar{y})| \in O(n/\zeta^2)$, and with probability at least $1 - 1/(2n)$,

$$\sum_{U \in \mathcal{F}(\mathbb{D})} y_U s_U s_U^\mathsf{T} \succeq_{\widehat{\mathbb{K}^*}} (1-\gamma)\alpha_\Xi \bar{Y}.$$

Assume this event holds, and set $p := (\alpha_\Xi(1-\gamma))^{-1} y$. Proposition 28 finishes the proof. □

**Proposition 31.** Let $\xi$, $\gamma$ be such that $\xi \in (0,1]$ and $\gamma \in (0,1)$. Let $\Xi$ be a randomized rounding algorithm for $\mathbb{D}$. Let $Z \in \mathbb{S}^n_+$. Let $(\mu, Y)$ be feasible in (6a) with $\mu > 0$. Let $T \geq \left\lceil \frac{2\pi}{\gamma^2 \xi} \log(2n) \right\rceil$, and let $(X_t)_{t\in[T]}$ be i.i.d. random variables sampled from $\Xi_Y$. If

$$Y \succ 0 \quad \text{and} \quad s_U^\mathsf{T} Y^{-1} s_U \leq \frac{1}{\xi\mu} \text{ for every } U \subseteq [n] \text{ with } \mathbb{P}(\Xi_Y = U) > 0, \tag{53}$$

then

$$\mathbb{P}\Big( \frac{\mu}{(1-\gamma)T} \sum_{t\in[T]} X_t \succeq \alpha_\Xi Z \Big) \geq 1 - \frac{1}{2n}.$$

*Proof.* Set $\bar{Y} := \mu^{-1} Y$. Since $\mu > 0$ and $\operatorname{diag}(Y) = \mu \mathbb{1}$, we have that $\bar{Y} \succ 0$ and $\operatorname{diag}(\bar{Y}) = 1$. Note that

$$\lambda_{\max}(\bar{Y}^{-1/2} s_U s_U^\mathsf{T} \bar{Y}^{-1/2}) = \mu s_U^\mathsf{T} Y^{-1} s_U \leq \mu \frac{1}{\mu\xi} \leq \frac{1}{\xi}. \tag{54}$$

Hence $0 \preceq \Xi_Y \preceq \frac{1}{\xi}\bar{Y}$ almost surely. From (4), we may apply Proposition 24 with $\rho := 1/\xi$ and $\tau := \alpha_{\mathrm{N}} = 2/\pi$ to conclude that with

$$T \geq \left\lceil \frac{4\rho}{\gamma^2\tau} \log(2n) \right\rceil = \left\lceil \frac{4}{\gamma^2 \xi \alpha_{\mathrm{N}}} \log(2n) \right\rceil = \left\lceil \frac{2\pi}{\gamma^2\xi} \log(2n) \right\rceil$$

we have that $\frac{1}{T}\sum_{t\in[T]} X_t \succeq (1-\gamma)\alpha_\Xi \bar{Y}$ with probability at least $1 - 1/(2n)$. The result follows from $(\mu, Y)$ being feasible in (6a). ∎

*Proof of Theorem 5.* Set $\tau := 1 - \beta/\alpha_\Xi$, and set $\gamma := \sigma := \varepsilon := \tau/4$. Set $\zeta := \tau/(4-\tau)$, so $(1+\zeta)^{-1} = 1 - \tau/4$. By nearly solving either (44a) or (45), depending on whether $W \in \mathbb{K}$ or $Z \in \mathbb{K}^*$ was given as input, one can compute $(\bar{\rho}, x)$ and $(\bar{\mu}, Y)$ witnessing the membership $(W, Z) \in \mathrm{H}_{\varepsilon,\sigma}(\mathbb{D}, \mathbb{K})$. Proposition 30 ensures one can compute, in randomized polynomial time, a $\hat{\beta}$-certificate $(\rho, \mu, U, y, x)$ for $(W, Z)$ with high probability, where $\hat{\beta} = (1-\tau/4)^4$. Since the function $x \mapsto (1 - x/4)^4$ is convex, it overestimates $1 - x$, which is its best linear approximation at $x := 0$. Hence

$$\hat{\beta} = \alpha_\Xi (1-\tau/4)^4 \geq \alpha_\Xi(1-\tau) = \beta.$$

Thus $(\rho, \mu, U, y, x)$ is a $\beta$-certificate. Note that $\zeta$ is independent of $n$, so from Proposition 30 we can conclude that $|\operatorname{supp}(y)| \in O(n)$, the hidden constant depending only on $\beta$. Similarly, since $\varepsilon$ and $\gamma$ do not depend on $n$ either, we have that the algorithm takes at most

$$\left\lceil \frac{8}{(\varepsilon\alpha_\Xi\gamma)^2} n^2 \log(2n) \right\rceil = \left\lceil \frac{2048}{\alpha_\Xi^2 \tau^4} n^2 \log(2n) \right\rceil \in O(n^2 \log(n))$$

samples from $\Xi$, the hidden constant depending only on $\beta$.

Now assume $\mathbb{K} = \mathbb{S}^n_+$. It is immediate that (3) holds. Set $\bar{Y} := \bar{\mu}^{-1} Y$. We have that $\bar{Y} \succeq \varepsilon I$, so $\bar{Y}^{-1} \preceq \frac{1}{\varepsilon} I$. Hence $s_U^\mathsf{T} \bar{Y}^{-1} s_U \leq n/\varepsilon$ for every $U \subseteq [n]$, so (53) holds for $\xi := \varepsilon/n$. Propositions 31 and 28 imply that we can compute, with high probability, a $\hat{\beta}$-certificate $(\rho, \mu, U, \bar{y}, x)$ for $(W, Z)$ with

$$\left\lceil \frac{2\pi}{\xi\gamma^2} \log(2n) \right\rceil = \left\lceil \frac{2\pi}{\varepsilon\gamma^2} n \log(2n) \right\rceil = \left\lceil \frac{128\pi}{\tau^3} n \log(2n) \right\rceil$$

samples from GW. Since $\tau$ is independent of $n$, Proposition 14 implies that we can sparsify $\bar{y} \in \mathbb{R}^{\mathcal{F}(\mathbb{D})}_+$, which potentially has support $O(n \log n)$, into $y \in \mathbb{R}^{\mathcal{F}(\mathbb{D})}_+$ with $|\operatorname{supp}(y)| \in O(n)$. ∎

## Appendix E. Boolean 2-CSP

Let $i, j \in [n]$. Define the following matrices in $\mathbb{S}^{\{0\} \cup [n]}$:

$$\Delta_{-i,-j} \coloneqq \tfrac{1}{2}((e_0 - e_i)(e_0 - e_j)^\mathsf{T} + (e_0 - e_j)(e_0 - e_i)^\mathsf{T}),$$
$$\Delta_{-i,+j} \coloneqq \tfrac{1}{2}((e_0 - e_i)(e_0 + e_j)^\mathsf{T} + (e_0 + e_j)(e_0 - e_i)^\mathsf{T}),$$
$$\Delta_{+i,-j} \coloneqq \tfrac{1}{2}((e_0 + e_i)(e_0 - e_j)^\mathsf{T} + (e_0 - e_j)(e_0 + e_i)^\mathsf{T}),$$
$$\Delta_{+i,+j} \coloneqq \tfrac{1}{2}((e_0 + e_i)(e_0 + e_j)^\mathsf{T} + (e_0 + e_j)(e_0 + e_i)^\mathsf{T}).$$

Direct computation shows that, for every $B \in \mathbb{R}^{\{0,\dots,n\} \times \{0,\dots,n\}}$,

$$\langle \Delta_{-i,-j}, B^\mathsf{T} B\rangle = \langle B(e_0 - e_i), B(e_0 - e_i)\rangle, \quad \langle \Delta_{-i,+j}, B^\mathsf{T} B\rangle = \langle B(e_0 - e_i), B(e_0 + e_i)\rangle$$
$$\langle \Delta_{+i,-j}, B^\mathsf{T} B\rangle = \langle B(e_0 + e_i), B(e_0 - e_i)\rangle, \quad \langle \Delta_{+i,+j}, B^\mathsf{T} B\rangle = \langle B(e_0 + e_i), B(e_0 + e_i)\rangle.$$

It is then routine to check that

$$\operatorname{conv}\{\, s_U s_U^\mathsf{T} : U \subseteq \{0\} \cup [n]\} \cap \mathbb{D}_\Delta = \operatorname{conv}\{\, s_U s_U^\mathsf{T} : U \subseteq \{0\} \cup [n]\} = \mathrm{CUT}^{\mathbb{D}_\Delta} . \tag{55}$$

## References


[AN06] N. Alon and A. Naor. "Approximating the Cut-Norm via Grothendieck's Inequality". In: *SIAM Journal on Computing* 35.4 (2006), pages 787–803.

[Ben+11] S. Benabbas, S. O. Chan, K. Georgiou, and A. Magen. "Tight gaps for vertex cover in the Sherali-Adams SDP hierarchy". In: *31st International Conference on Foundations of Software Technology and Theoretical Computer Science*. Volume 13. LIPIcs. Leibniz Int. Proc. Inform. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2011, pages 41–54.

[Ben+23] N. Benedetto Proença, M. K. de Carli Silva, C. M. Sato, and L. Tunçel. *A Primal-Dual Extension of the Goemans–Williamson Algorithm for the Weighted Fractional Cut-Covering Problem*. November 26, 2023. arXiv: 2311.15346 [math.OC].

[BNR03] A. Ben-Tal, A. Nemirovski, and C. Roos. "Extended matrix cube theorems with applications to $\mu$-theory in control". In: *Math. Oper. Res.* 28.3 (2003), pages 497–523.

[Bra+23] J. Brakensiek, N. Huang, A. Potechin, and U. Zwick. "Separating MAX 2-AND, MAX DI-CUT and MAX CUT". In: *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*. Los Alamitos, CA, USA: IEEE Computer Society, November 2023, pages 234–252.

[BSS12] J. Batson, D. A. Spielman, and N. Srivastava. "Twice-Ramanujan Sparsifiers". In: *SIAM Journal on Computing* 41.6 (2012), pages 1704–1721.

[CHS15] M. K. de Carli Silva, N. J. A. Harvey, and C. M. Sato. "Sparse Sums of Positive Semidefinite Matrices". In: *ACM Trans. Algorithms* 12.1 (December 9, 2015), Article No. 9, pp. 1–17.

[FL20] S. Friedland and L.-H. Lim. *Symmetric Grothendieck inequality*. March 16, 2020. arXiv: 2003.07345 [math.FA].

[Fre87] R. M. Freund. "Dual gauge programs, with applications to quadratic programming and the minimum-norm problem". In: *Math. Programming* 38.1 (1987), pages 47–67.

[Ful71] D. R. Fulkerson. "Blocking and anti-blocking pairs of polyhedra". In: *Math. Programming* 1 (1971), pages 168–194.

[Ful72] D. R. Fulkerson. "Anti-blocking polyhedra". In: *J. Combinatorial Theory Ser. B* 12 (1972), pages 50–71.

[Geo+10] K. Georgiou, A. Magen, T. Pitassi, and I. Tourlakis. "Integrality gaps of $2 - o(1)$ for vertex cover SDPs in the Lovász-Schrijver hierarchy". In: *SIAM J. Comput.* 39.8 (2010), pages 3553–3570.

[GLS81] M. Grötschel, L. Lovász, and A. Schrijver. “The ellipsoid method and its consequences in combinatorial optimization”. In: *Combinatorica* 1.2 (1981), pages 169–197.

[Gro53] A. Grothendieck. “Résumé de la théorie métrique des produits tensoriels topologiques”. In: *Bol. Soc. Mat. São Paulo* 8 (1953), pages 1–79.

[GW95] M. X. Goemans and D. P. Williamson. “Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming”. In: *J. ACM* 42.6 (1995), pages 1115–1145.

[KG98] J. Kleinberg and M. X. Goemans. “The Lovász theta function and a semidefinite programming relaxation of vertex cover”. In: *SIAM J. Discrete Math.* 11.2 (1998), pages 196–204.

[Kho+07] S. Khot, G. Kindler, E. Mossel, and R. O’Donnell. “Optimal inapproximability results for MAX-CUT and other 2-variable CSPs?” In: *SIAM J. Comput.* 37.1 (2007), pages 319–357.

[Kri77] J.-L. Krivine. “Sur la constante de Grothendieck”. In: *C. R. Acad. Sci. Paris Sér. A-B* 284.8 (1977), A445–A446.

[LLZ02] M. Lewin, D. Livnat, and U. Zwick. “Improved rounding techniques for the MAX 2-SAT and MAX DI-CUT problems”. In: *Integer Programming and Combinatorial Optimization. IPCO 2002.* Edited by W. J. Cook and A. S. Schulz. Volume 2337. Lecture Notes in Comput. Sci. Springer, Berlin, 2002, pages 67–82.

[LP95] M. Laurent and S. Poljak. “On a positive semidefinite relaxation of the cut polytope”. In: *Linear Algebra Appl.* 223/224 (1995). Special issue honoring Miroslav Fiedler and Vlastimil Pták, pages 439–461.

[NB19] J. Neto and W. Ben-Ameur. “On fractional cut covers”. In: *Discrete Appl. Math.* 265 (2019), pages 168–181.

[Nem24] A. Nemirovski. *Introduction to Linear Optimization.* 2024. URL: http://www2.isye.gatech.edu/~nemirovs/OPTILN2024Spring.pdf (visited on 05/01/2024).

[Nes98] Y. Nesterov. “Semidefinite relaxation and nonconvex quadratic optimization”. In: *Optim. Methods Softw.* 9.1-3 (1998), pages 141–160.

[Rag08] P. Raghavendra. “Optimal algorithms and inapproximability results for every CSP?” In: *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing. STOC’08.* ACM, New York, 2008, pages 245–254.

[Šám06] R. Šámal. “On *XY* Mappings”. PhD thesis. Charles University, 2006. vi+146 pages. URL: https://dspace.cuni.cz/handle/20.500.11956/6160 (visited on 06/27/2023).

[SV14] M. Singh and N. K. Vishnoi. “Entropy, optimization and counting”. In: *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing. STOC’14.* New York, NY, USA: ACM, New York, 2014, pages 50–59.

[SZ22] Z. Song and R. Zhang. “Hyperbolic Concentration, Anti-Concentration, and Discrepancy”. In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2022).* Edited by A. Chakrabarti and C. Swamy. Volume 245. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022, 10:1–10:19.

[Tin74] J. Tind. “Blocking and antiblocking sets”. In: *Math. Programming* 6 (1974), pages 157–166.

[Tro12] J. A. Tropp. “User-Friendly Tail Bounds for Sums of Random Matrices”. In: *Found. Comput. Math.* 12.4 (2012), pages 389–434.

[Tro15] J. A. Tropp. “An Introduction to Matrix Concentration Inequalities”. In: *Found. Trends Mach. Learn.* 8.1-2 (2015), pages 1–230.