SIAM J. DISCRETE MATH.
Vol. 34, No. 1, pp. 470–496

A NOTION OF TOTAL DUAL INTEGRALITY FOR CONVEX, SEMIDEFINITE, AND EXTENDED FORMULATIONS*

MARCEL K. DE CARLI SILVA† AND LEVENT TUNÇEL‡

Abstract. Total dual integrality is a powerful and unifying concept in polyhedral combinatorics and integer programming that enables the refinement of geometric min-max relations given by linear programming strong duality into combinatorial min-max theorems. The definition of a linear inequality system being totally dual integral (TDI) revolves around the existence of optimal dual solutions that are integral and thus naturally applies to a host of combinatorial optimization problems that are cast as integer programs whose linear program (LP) relaxations have the TDIness property. However, when combinatorial problems are formulated using more general convex relaxations, such as semidefinite programs (SDPs), it is not at all clear what an appropriate notion of integrality in the dual program is, thus inhibiting the generalization of the theory to more general forms of structured convex optimization. (In fact, we argue that the rank-one constraint usually added to SDP relaxations is not adequate in the dual SDP.) In this paper, we propose a notion of total dual integrality for SDPs that generalizes the notion for LPs, by relying on an "integrality constraint" for SDPs that is primal-dual symmetric. A key ingredient for the theory is a generalization to compact convex sets of a result of Hoffman for polytopes, fundamental for generalizing the polyhedral notion of total dual integrality introduced by Edmonds and Giles. We study the corresponding theory applied to SDP formulations for stable sets in graphs using the Lovász theta function and show that total dual integrality in this case corresponds to the underlying graph being perfect. We also relate dual integrality of an SDP formulation for the maximum cut problem to bipartite graphs. Total dual integrality for extended formulations naturally comes into play in this context.

Key words. total dual integrality, semidefinite programming, Lovász theta function

AMS subject classifications. 90C10, 90C22, 90C25, 90C27

DOI. 10.1137/18M1169710

1. Introduction.

In the polyhedral approach to combinatorial optimization one usually starts by formulating a combinatorial problem as an integer linear program (ILP) of the form $\max\{ c^{\mathsf T} x : Ax \le b,\, x \ge 0,\, x \in \mathbb{Z}^n\}$, which is relaxed into a linear program (LP) and then studied in the light of LP duality. This basic approach of polyhedral combinatorics can be summarized by the following simple yet fundamental result.

THEOREM 1.1. *If $A \in \mathbb{Q}^{m\times n}$ is a matrix and $b \in \mathbb{Q}^m$ and $c \in \mathbb{Q}^n$ are vectors,*

*Received by the editors February 8, 2018; accepted for publication (in revised form) October 21, 2019; published electronically February 19, 2020. Part of the results in this paper appeared in the Ph.D. thesis [7] of the first author.
https://doi.org/10.1137/18M1169710
Funding: Research of the first author was supported in part by a Sinclair Scholarship, a Tutte Scholarship, Discovery Grants from NSERC, and by the U.S. Office of Naval Research under award N00014-12-1-0049 while at the Department of Combinatorics and Optimization, University of Waterloo, and by Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (423833/2018-9, 456792/2014-7, and 477203/2012-4), by grant 2013/03447-6 of the São Paulo Research Foundation (FAPESP), and by CAPES while at the Institute of Mathematics and Statistics, University of São Paulo. Research of the second author was supported in part by Discovery Grants from NSERC, and by the U.S. Office of Naval Research under awards N00014-12-1-0049 and N00014-15-1-2171. Part of this work was done while the second author was visiting the Simons Institute for the Theory of Computing, supported in part by the DIMACS/Simons Collaboration on Bridging Continuous and Discrete Optimization through NSF grant CCF-1740425.
†Instituto de Matemática e Estatística, Universidade de São Paulo, 05508-090 São Paulo, Brazil (mksilva@ime.usp.br).
‡Department of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, Waterloo, ON N2L 3G1, Canada (ltuncel@uwaterloo.ca).

then

$$
\begin{aligned}
&\text{(ILP)} && \sup\{\, c^{\mathsf T}x : Ax \le b,\, x \ge 0,\, x \in \mathbb{Z}^n \}\\
&\text{(LP)} && \le \sup\{\, c^{\mathsf T}x : Ax \le b,\, x \ge 0,\, x \in \mathbb{R}^n \}\\
&\text{(LD)} && \le \inf\{\, b^{\mathsf T}y : A^{\mathsf T}y \ge c,\, y \ge 0,\, y \in \mathbb{R}^m \}\\
&\text{(ILD)} && \le \inf\{\, b^{\mathsf T}y : A^{\mathsf T}y \ge c,\, y \ge 0,\, y \in \mathbb{Z}^m \}.
\end{aligned}
$$

If (ILP) *and* (ILD) *are both feasible, the suprema and infima are attained (by Meyer's theorem* [23]*), and the middle (second) inequality holds with equality.*

Usually, in the formulations of combinatorial problems, the feasible region of (ILP) is contained in $\{0,1\}^n$ and some optimal solution of (ILD) lies in $\{0,1\}^m$. For instance, suppose that we are given a graph $G = (V, E)$ with vertices V and edges E, and both b and c are equal to the vector $\mathbb{1}$ of all-ones. Then, when A is the $V \times E$ incidence matrix of G, (ILP) formulates the maximum cardinality matching problem, and (ILD) formulates the minimum cardinality vertex cover problem. Alternatively, if A is the $E \times V$ incidence matrix of G, we obtain the maximum cardinality stable set problem and the minimum cardinality edge cover problem. If A is the clique-vertex incidence matrix of G, then (ILP) still formulates the maximum cardinality stable set problem, but now (ILD) formulates the minimum cardinality coloring problem.

What makes the conceptual framework brought forth by Theorem 1.1 so fundamental is the fact that, in many interesting and important cases [31], equality holds throughout in the chain from Theorem 1.1, which allows us to refine a *geometric* min-max relation (equality between (LP) and (LD) given by LP strong duality) into a *combinatorial* min-max relation (equality between (ILP) and (ILD)). For instance, equality throughout holds for the first two cases above when G is bipartite (and has no isolated vertices in the second case), thus proving very strong, weighted forms of Kőnig's matching theorem and the Kőnig–Rado edge cover theorem.

Total dual integrality is arguably the most powerful and unifying sufficient condition for equality throughout the chain from Theorem 1.1. A vector in $\mathbb{R}^n$ is *integral* if each of its components is an integer, and a rational system of linear inequalities $Ax \le b$ is *totally dual integral* (TDI) if, for each integral vector $c \in \mathbb{Z}^n$, the linear program dual to $\sup\{\, c^{\mathsf T}x : Ax \le b\}$ has an integral optimal solution whenever it has an optimal solution at all. In this case, if b itself is integral, then the polyhedron P determined by $Ax \le b$ is *integral*, i.e., each nonempty face of P has an integral vector; thus, equality holds throughout in the chain from Theorem 1.1. This was proved in seminal work of Edmonds and Giles [11] as a consequence of the following fundamental result.

THEOREM 1.2 (Edmonds–Giles [11]). *Let* $A \in \mathbb{Q}^{m\times n}$ *and* $b \in \mathbb{Q}^m$. *If* $P := \{\, x \in \mathbb{R}^n : Ax \le b\}$ *is such that* $\sup_{x\in P} c^{\mathsf T}x \in \mathbb{Z} \cup \{\pm\infty\}$ *for each* $c \in \mathbb{Z}^n$, *then* P *is integral.*

COROLLARY 1.3 (Hoffman [15]). *Let* $A \in \mathbb{Q}^{m\times n}$, *and let* $b \in \mathbb{Q}^m$. *If* $P := \{\, x \in \mathbb{R}^n : Ax \le b\}$ *is bounded and* $\sup_{x\in P} c^{\mathsf T}x \in \mathbb{Z} \cup \{-\infty\}$ *for each* $c \in \mathbb{Z}^n$, *then* P *is integral.*

In the past couple of decades, it has become popular to formulate combinatorial optimization problems using more general models of convex optimization, with semidefinite programs (SDPs) playing a key role. Let us introduce some basic notation for SDPs. The real vector space of symmetric $n \times n$ matrices is $\mathbb{S}^n$. The set of *positive semidefinite* matrices is $\mathbb{S}^n_+ := \{\, X \in \mathbb{S}^n : h^{\mathsf T}Xh \ge 0\, \forall h \in \mathbb{R}^n\}$. The *(trace)* *inner product* of $X, Y \in \mathbb{S}^n$ is $\langle X, Y \rangle := \mathrm{Tr}(XY) = \sum_{i=1}^n \sum_{j=1}^n X_{ij} Y_{ij}$. Denote $[n] := \{1, \ldots, n\}$ for each $n \in \mathbb{N}$. We refer the reader to subsection 1.1 for the rest of the notation used throughout the text.

When a combinatorial problem is formulated as in (ILP), the combinatorial objects are usually embedded in the (geometric) space $\mathbb{R}^n$ as incidence vectors; i.e., we consider feasible solutions of the form $x = \mathbb{1}_U \in \mathbb{R}^n$, for certain subsets $U \subseteq [n]$, where the ith coordinate of $\mathbb{1}_U$ is 1 if $i \in U$ and 0 otherwise. Having a correct ILP formulation for a combinatorial optimization problem typically means that the feasible solutions for (ILP) are in *exact* correspondence with the combinatorial objects of interest in the problem. One then considers the LP relaxation (LP) by dropping the nonconvex constraint "$x \in \mathbb{Z}^n$." Note that the "integer dual" (ILD) is obtained from the dual (LD) of (LP) by adding back the nonconvex constraint "$y \in \mathbb{Z}^m$" of the same form.

When embedding combinatorial objects into matrix space $\mathbb{S}^n$ for an SDP formulation, one may embed a subset $U \subseteq [n]$ as the rank-one matrix $X = \mathbb{1}_U \mathbb{1}_U^\mathsf{T} \in \mathbb{S}^n_+$. It is also common to use rank-one matrices arising from signed incidence vectors, e.g., $X = s_U s_U^\mathsf{T}$ where $s_U = 2\mathbb{1}_U - \mathbb{1} \in \{\pm 1\}^n$ for some $U \subseteq [n]$. (We shall argue later that there is a "better" embedding, which we shall adopt.) One then obtains the following optimization problems, partially mimicking the chain from Theorem 1.1:

$$\sup\{ \langle C, X \rangle : \langle A_i, X \rangle \le b_i \, \forall i \in [m],\, X \in \mathbb{S}^n_+,\, \mathrm{rank}(X) = 1 \} \tag{1.1a}$$
$$\le \sup\{ \langle C, X \rangle : \langle A_i, X \rangle \le b_i \, \forall i \in [m],\, X \in \mathbb{S}^n_+ \} \tag{1.1b}$$
$$\le \inf\{ b^\mathsf{T} y : y \in \mathbb{R}^m_+,\, \textstyle\sum_{i=1}^m y_i A_i - C \in \mathbb{S}^n_+ \}, \tag{1.1c}$$

where $A_1, \ldots, A_m, C \in \mathbb{S}^n$ and $b \in \mathbb{R}^m$. Here usually the feasible solutions for (1.1a) correspond *exactly* to the combinatorial objects of interest, as is the case for (ILP). Similarly as in Theorem 1.1, the SDP relaxation (1.1b) is obtained from (1.1a) by dropping the nonconvex constraint "$\mathrm{rank}(X) = 1$," (1.1c) is the SDP dual of (1.1b), and the last inequality is SDP weak duality. There are many instances of the chain (1.1) in the literature; see, e.g., [14, 13, 25]. Some of this work is in copositive programming (see, for instance, [3]).

Conspicuously missing from (1.1) is a fourth optimization problem, that is, an "integer dual SDP" corresponding to (ILD). In fact, it is not even clear what the right notion of integrality is for (1.1c), i.e., which nonconvex constraint to add to (1.1c) to obtain a sensible combinatorial problem. One could argue that we may just add back the nonconvex constraint from (1.1a), by requiring the dual slack $\sum_{i=1}^m y_i A_i - C$ to have rank one, and it might also make sense to require y to be integral. Unfortunately, as we describe in section 2, the "integer dual SDP" thus obtained is not very satisfactory: whereas it can be made to generalize the corresponding notion for LPs, it fails to provide sensible "integer duals" for the SDP formulations of some of the most classical combinatorial problems, namely the Lovász theta function and the Max Cut SDP. Thus, we require our notion of "integrality constraints in the dual" to provide meaningful combinatorial min-max theorems at least for Max Cut SDP and more importantly, for SDP formulations of the Lovász theta function.

The theta function introduced by Lovász [20] was one of the earliest applications of SDPs to combinatorial optimization. The theta function of a graph, which can be computed efficiently (to within any given precision), lies sandwiched between its stability and clique-covering numbers. More importantly, the theta function has a rich and elegant duality theory (see, e.g., [8]). This is why we take the underlying SDPs as the main test case for any generalization of TDI theory. The other SDP mentioned above, the Max Cut SDP, was famously exploited in a breakthrough approximation algorithm and its analysis by Goemans and Williamson [13] and helped popularize SDPs in the discrete optimization and theoretical computer science communities.

In this paper, we introduce a notion of integrality for SDPs that

(i) generalizes the usual rank-one constraint in primal SDPs;
(ii) allows us to extend the chain (1.1) so as to generalize Theorem 1.1 for LPs in the natural, diagonal embedding of $Ax \leq b$ into matrix space;
(iii) is primal-dual symmetric;
(iv) yields sensible "integer duals" for the SDPs for the Lovász theta function and the Max Cut SDP.

We use this integrality condition for SDPs to define the notion of *total dual integrality* for the defining system of an SDP. We connect this new notion to Corollary 1.3 by extending the latter to compact convex sets, using basic tools from convex analysis and ILP theory, such as the Gomory–Chvátal closure. We prove that the total dual integrality of an SDP formulation for the Lovász theta function is equivalent to the underlying graph being perfect. We also study a close relative of TDIness for the Max Cut SDP and relate it to bipartiteness of the underlying graph. Along the way, we discuss an intermediate generalization of TDIness for LPs in terms of lifted (extended) formulations. Finally, we discuss future research directions along these lines, inspired by integrality (and other exactness) notions in convex optimization.

In order to achieve this, several obstacles must be overcome. First, we must choose a specific format for SDPs that makes it natural to work with integral solutions; that is, we must settle for a specific embedding of combinatorial objects into matrix space. We solve this partially by restricting ourselves to binary integer programs, i.e., where variables can only take values in $\{0, 1\}$; this is the usual case in combinatorial optimization. Our choice of embedding and our focus on the combinatorial aspects of the dual SDP require us to rewrite SDP constraints in a slightly unusual way; this happens because other works in the literature do not focus on integrality for the dual SDP. Finally, SDP formulations for combinatorial problems are usually lifted formulations, so we must generalize the (algebraic) notion of TDIness to these (geometric) extended formulations.

Our work is related to previous abstract notions of duality in integer programming; we highlight [4, 27].

1.1. Notation. The set of nonnegative integers (resp., reals) is denoted by $\mathbb{Z}_+$ (resp., $\mathbb{R}_+$). The set of positive reals is $\mathbb{R}_{++}$. We use Iverson's notation: for a predicate P, we denote

$$[P] := \begin{cases} 1 & \text{if } P \text{ holds,} \\ 0 & \text{otherwise.} \end{cases}$$

Throughout the text, V should be considered a finite set, usually taken to be the vertex set of a graph. The set of k-subsets of V is $\binom{V}{k} := \{ U \subseteq V : |U| = k \}$. The collection of subsets of V that contain some $i \in V$ (resp., that contain some $i \in V$ and $j \in V$) is denoted $\binom{V}{i\in}$ (resp., $\binom{V}{ij\subseteq}$).

All graphs in this paper are simple. Let $G = (V, E)$ be a graph. The *complement* of G is $\overline{G} := (V, \overline{E})$, where $\overline{E} := \binom{V}{2} \setminus E$. The subgraph of G *induced* by $U \subseteq V$ is $G[U] := \big(U, E \cap \binom{U}{2}\big)$. The complete graph on vertex set V is K_V. A subset U of V is a *clique* in G if $G[U] = K_U$; we say that U is *stable* in G if U is a clique in $\overline{G}$. Denote the set of cliques of G by $\mathcal{K}(G)$. The *clique number* of G is $\omega(G) := \max\{ |K| : K \in \mathcal{K}(G)\}$.

The *chromatic number* $\chi(G)$ of G is the smallest size of a partition of V into stable sets of G. A graph $G = (V, E)$ is *perfect* if $\omega(G[U]) = \chi(G[U])$ for every $U \subseteq V$.

The canonical basis of $\mathbb{R}^V$ is $\{ e_i : i \in V\}$. The *incidence vector* of $U \subseteq V$ is $\mathbb{1}_U \in \mathbb{R}^V$, i.e., $(\mathbb{1}_U)_i := [i \in U]$ for each $i \in V$. Let W be a finite set. The direct sum of $x \in \mathbb{R}^V$ and $y \in \mathbb{R}^W$ is $x \oplus y$. The componentwise product of $x, y \in \mathbb{R}^V$ is denoted by $x \odot y$. The *support* of $x \in \mathbb{R}^V$ is $\operatorname{supp}(x) := \{ i \in V : x_i \neq 0\}$.

The real vector space of symmetric $V \times V$ matrices is $\mathbb{S}^V$. The cone of symmetric positive semidefinite (resp., entrywise nonnegative) $V \times V$ matrices is denoted by $\mathbb{S}^V_+$ (resp., $\mathbb{S}^V_{\geq 0}$). For $A, B \in \mathbb{S}^V$, we write $A \succeq B$ if $A - B \in \mathbb{S}^V_+$, and for $A, B \in \mathbb{R}^{V \times W}$, we write $A \geq B$ if $A_{ij} \geq B_{ij}$ for every $(i, j) \in V \times W$. The identity matrix in appropriate space is denoted by I. The map $\operatorname{diag}\colon \mathbb{R}^{V\times V} \to \mathbb{R}^V$ extracts the diagonal of a matrix, and its adjoint $\operatorname{Diag}\colon \mathbb{R}^V \to \mathbb{R}^{V\times V}$ builds diagonal matrices. The *principal submatrix* of $X \in \mathbb{R}^{V\times V}$ indexed by $U \subseteq V$ is the restriction $X[U] := X\restriction_{U\times U} \in \mathbb{R}^{U\times U}$. The orthogonal projection of $A \in \mathbb{R}^{V\times V}$ onto $\mathbb{S}^V$ is $\operatorname{Sym}(A) := \frac{1}{2}(A + A^{\mathsf{T}})$. The convex hull of a subset $\mathscr{U}$ of a Euclidean space is denoted by $\operatorname{conv}(\mathscr{U})$.

1.2. Organization. The rest of this text is organized as follows. We discuss dual integrality constraints for SDPs in section 2, including drawbacks of the rank-one constraint usually added to the primal SDP, as well as embedding issues. There, we show that our notion of dual integrality befits nicely with the Lovász theta function. In section 3, we generalize Corollary 1.3, which motivates us to define a notion of *total* dual integrality for SDPs in section 4; we show that the latter is sufficient for *primal* integrality. In section 5, we characterize total dual integrality for formulations of the Lovász theta function, and we study dual integrality for the Max Cut SDP with nonnegative weight functions in section 6. We conclude our paper with several open problems and future research directions in section 7.

2. Fundamental framework and integrality constraint for dual SDP.

We discuss in subsection 2.1 the shortcomings of the rank constraint as an "integrality constraint" for the dual SDP (1.1c), and we propose a suitable replacement in subsection 2.2. Throughout the discussion, a few somewhat unusual choices will be made which are not common in the SDP literature; e.g., whenever appropriate we are careful when writing linear inequalities of the form $\langle A, X\rangle \leq \beta$ on a matrix variable X with an *integral symmetric* matrix A and *integer* β. The reason we insist on symmetry of A is to properly set up the dual SDP, and we want A and β to be integral so as to simplify combinatorial interpretation of the linear system; this is also the case when one studies the ILP chain from Theorem 1.1 in the context of classical TDIness theory.

2.1. Drawbacks of the rank-one constraint as a dual integrality constraint. In order to discuss integrality constraints for SDPs, we must first choose a standard form to embed combinatorial objects (e.g., subsets of some finite ground set V) into matrix space $\mathbb{S}^V$. The format we shall choose actually embeds subsets of a finite set V as matrices in $\mathbb{S}^{\{0\}\cup V}$; i.e., the index set has one extra element, which we call 0, *assumed throughout not to be in* V. Each subset U of V is embedded as the rank-one matrix

$$\hat{X} := \begin{bmatrix} 1 \\ \mathbb{1}_U \end{bmatrix} \begin{bmatrix} 1 \\ \mathbb{1}_U \end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} 1 & \mathbb{1}_U^{\mathsf{T}} \\ \mathbb{1}_U & \mathbb{1}_U \mathbb{1}_U^{\mathsf{T}} \end{bmatrix} \in \mathbb{S}^{\{0\}\cup V}_+; \tag{2.1}$$

as a convention, *we decorate matrices in this lifted space with a hat*, e.g., $\hat{X}$ in (2.1). Similarly, since we use the lifted matrix space so often, we shall abbreviate

$$\widehat{\mathbb{S}}^V := \mathbb{S}^{\{0\}\cup V} \quad \text{and} \quad \widehat{\mathbb{S}}^V_+ := \mathbb{S}^{\{0\}\cup V}_+, \tag{2.2}$$

and we also decorate subsets of $\widehat{\mathbb{S}}^V$ with a hat, e.g., $\widehat{\mathscr{C}} \subseteq \widehat{\mathbb{S}}^V$. By writing any matrix $\hat{X}$ from (2.1) in the form

$$\hat{X} = \begin{bmatrix} 1 & x^\mathsf{T} \\ x & X \end{bmatrix} \in \widehat{\mathbb{S}}^V, \tag{2.3}$$

with $X \in \mathbb{S}^V$, one sees that it satisfies $\hat{X}_{00} = 1$ and $x_j = X_{jj} \geq 0$ for each $j \in V$, which we shall write as

$$\langle e_0 e_0^\mathsf{T}, \hat{X}\rangle = 1, \tag{2.4a}$$

$$\big\langle 2\,\mathrm{Sym}(e_j(e_j - e_0)^\mathsf{T}), \hat{X}\big\rangle = 0 \qquad \forall j \in V, \tag{2.4b}$$

$$\big\langle e_j e_j^\mathsf{T}, \hat{X}\big\rangle \geq 0 \qquad \forall j \in V. \tag{2.4c}$$

The constraints (2.4), together with the constraint $\mathrm{rank}(\hat{X}) = 1$, ensure that $\hat{X}$ has the form (2.1) for some $U \subseteq V$. Throughout the rest of the text, one may think that every system of linear inequalities on $\hat{X}$ arising from combinatorial problems includes the constraints (2.4), just as one usually considers the linear constraints $Ax \leq b$, $x \geq 0$ from (ILP) to include $0 \leq x \leq \mathbb{1}$.

Another constraint satisfied by $\hat{X}$ of the form (2.1), using the notation of (2.3), is $X = \hat{X}[V] \geq 0$. Sometimes it will make sense to add this extra constraint to (2.4), leading to the following constraints:

$$\langle e_0 e_0^\mathsf{T}, \hat{X}\rangle = 1, \tag{2.5a}$$

$$\big\langle 2\,\mathrm{Sym}(e_j(e_j - e_0)^\mathsf{T}), \hat{X}\big\rangle = 0 \qquad \forall j \in V, \tag{2.5b}$$

$$\big\langle e_j e_j^\mathsf{T}, \hat{X}\big\rangle \geq 0 \qquad \forall j \in V, \tag{2.5c}$$

$$\big\langle 2\,\mathrm{Sym}(e_i e_j^\mathsf{T}), \hat{X}\big\rangle \geq 0 \qquad \forall i, j \in V, \text{ such that } i \neq j. \tag{2.5d}$$

The embedding described above is used in some formulations of the theta function (see [14, 31]), in the lift-and-project hierarchies of Lovász and Schrijver [21] and Lasserre [18], and in copositive formulations for mixed integer linear programs by Burer [3].

A simple, natural way to obtain an SDP relaxation for (ILP) is to formulate

$$\text{Maximize} \quad \langle \mathrm{Diag}(0 \oplus c), \hat{X}\rangle \tag{2.6a}$$

$$\text{subject to} \quad \hat{X} \in \widehat{\mathbb{S}}^n_+ \text{ satisfies (2.5) with } V := [n], \tag{2.6b}$$

$$\langle \mathrm{Diag}(-b_i \oplus A^\mathsf{T} e_i), \hat{X}\rangle \leq 0 \qquad \forall i \in [m]. \tag{2.6c}$$

In this case, to obtain an exact reformulation of (ILP), corresponding to (1.1a), one may add the rank constraint $\mathrm{rank}(\hat{X}) \leq 1$ to (2.6). Note, however, that (2.6) is a potentially tighter relaxation for (ILP) than (LP). The SDP dual to (2.6) may be written as

(2.7a) Minimize η

(2.7b) subject to $\begin{bmatrix} \eta & -u^{\mathsf{T}} \\ -u & \mathrm{Diag}(2u) - Z \end{bmatrix} + \sum_{i \in [m]} y_i \begin{bmatrix} -b_i & 0^{\mathsf{T}} \\ 0 & \mathrm{Diag}(A^{\mathsf{T}} e_i) \end{bmatrix} - \hat{S} = \begin{bmatrix} 0 & 0^{\mathsf{T}} \\ 0 & \mathrm{Diag}(c) \end{bmatrix},$

(2.7c) $\hat{S} \in \widehat{\mathbb{S}}^n_+, \eta \in \mathbb{R}, u \in \mathbb{R}^n, y \in \mathbb{R}^m_+, Z \in \mathbb{S}^n_{\geq 0}.$

If (2.6b) is weakened to "$\hat{X} \in \widehat{\mathbb{S}}^n_+$ satisfies (2.4)," again with $V = [n]$, then the variable Z in (2.7) would be required to take the form $Z = \mathrm{Diag}(z)$ for some $z \in \mathbb{R}^n_+$; that is, the dual feasible region is smaller.

It is easy to check that if y is feasible in (ILD), then

$$(\eta, Z, y, \hat{S}, u) := (b^{\mathsf{T}} y, \mathrm{Diag}(A^{\mathsf{T}} y - c), y, 0, 0)$$

is feasible in (2.7) with the same objective value as that of y in (ILD). Thus, the rank constraint $\mathrm{rank}(\hat{S}) \leq 1$ seems reasonable as an integrality constraint for (2.7). In fact, we may even consider the tighter rank constraint $\mathrm{rank}(\hat{S}) = 1$, as long as we allow η to take on real values (rather than only integral ones), possibly at the cost of nonattainment. Note that we had to be very permissive for the rank constraint to make any sense at all.

Now we move on to the SDP formulation for ϑ, the Lovász theta function. In fact, we will also consider variations of ϑ usually denoted by ϑ' and ϑ^+, which were introduced independently in [22, 28, 32]. These parameters have subtle, slightly different properties, and hence it is important to study all three parameters. For instance, ϑ is multiplicative with respect to certain graph products whereas ϑ' is not; see, e.g., [24, sect. 4] and [1, Example 4.5]. *We shall show that the rank constraint is very inadequate for the dual SDP in this setting, for the theta function and its two variants.*

Let $G = (V, E)$ be a graph, and let $w \colon V \to \mathbb{R}$. There are several equivalent formulations for the *weighted theta number* $\vartheta(G; w)$ (see, e.g., [8]), and similarly for its variations $\vartheta'(G; w)$ and $\vartheta^+(G; w)$. In view of our choice of format for SDPs that includes the constraints (2.4), we define

(2.8a) $\vartheta'(G, w) :=$ Maximize $\langle \mathrm{Diag}(0 \oplus w), \hat{X} \rangle$

(2.8b) subject to $\hat{X} \in \widehat{\mathbb{S}}^V_+$ satisfies (2.4),

(2.8c) $\langle 2\,\mathrm{Sym}(e_i e_j^{\mathsf{T}}), \hat{X} \rangle = 0 \qquad \forall ij \in E,$

(2.8d) $\langle 2\,\mathrm{Sym}(e_i e_j^{\mathsf{T}}), \hat{X} \rangle \geq 0 \qquad \forall ij \in \overline{E},$

where $\overline{E} := \binom{V}{2} \setminus E$. Note that, if $U \subseteq V$ is stable in G, then the matrix $\hat{X}$ defined in (2.1) is feasible in (2.8) with objective value $w^{\mathsf{T}} \mathbb{1}_U = \sum_{u \in U} w_u$. Define also

(2.9) $\vartheta(G, w) := \max\{ \langle \mathrm{Diag}(0 \oplus w), \hat{X} \rangle : \hat{X} \in \widehat{\mathbb{S}}^V_+ \text{ satisfies (2.4) and (2.8c)} \},$

(2.10) $\vartheta^+(G, w) := \max\{ \langle \mathrm{Diag}(0 \oplus w), \hat{X} \rangle : \hat{X} \in \widehat{\mathbb{S}}^V_+ \text{ satisfies (2.4) and } \langle 2\,\mathrm{Sym}(e_i e_j^{\mathsf{T}}), \hat{X} \rangle \leq 0\, \forall ij \in E \}.$

The dual SDP of (2.8) is

(2.11a)
$$\text{Minimize} \quad \eta$$
(2.11b)
$$\text{subject to} \quad \begin{bmatrix} \eta & -u^{\mathsf{T}} \\ -u & \operatorname{Diag}(2u - z) \end{bmatrix} + \sum_{ij \in \binom{V}{2}} y_{ij} \begin{bmatrix} 0 & 0^{\mathsf{T}} \\ 0 & 2\operatorname{Sym}(e_i e_j^{\mathsf{T}}) \end{bmatrix} - \hat{S} = \begin{bmatrix} 0 & 0^{\mathsf{T}} \\ 0 & \operatorname{Diag}(w) \end{bmatrix},$$
(2.11c)
$$\hat{S} \in \widehat{\mathbb{S}}_{+}^{V},\ \eta \in \mathbb{R},\ u \in \mathbb{R}^{V},\ z \in \mathbb{R}_{+}^{V},\ y \in \mathbb{R}^{E} \oplus -\mathbb{R}_{+}^{\overline{E}}.$$

Note that the dual for the formulation (2.9) of $\vartheta(G; w)$ is similar, except that it requires $y\restriction_{\overline{E}} = 0$, and the dual for the formulation (2.10) of $\vartheta^{+}(G; w)$ furthermore has the sign constraint $y\restriction_{E} \geq 0$. We claim that

(2.12)
$$\text{if (2.11) has a feasible solution with } \operatorname{rank}(\hat{S}) \leq 1 \text{ and } w \in \mathbb{R}_{++}^{V}, \text{ then } G = K_V.$$

Indeed, suppose that $\operatorname{rank}(\hat{S}) \leq 1$. We have $\eta > 0$ by weak duality, so $\operatorname{rank}(\hat{S}) = 1$ and $\hat{S}[V] = \frac{1}{\eta} u u^{\mathsf{T}}$. Then,

$$\operatorname{Diag}(2u - z - w) + \sum_{ij \in \binom{V}{2}} 2y_{ij} \operatorname{Sym}(e_i e_j^{\mathsf{T}}) = \tfrac{1}{\eta} u u^{\mathsf{T}}. \tag{2.13}$$

By applying diag to both sides of (2.13), we get $2u - z - w = u \odot u$, so $2u = (u \odot u) + z + w \in \mathbb{R}_{++}^{V}$. Next let $i, j \in V$ be distinct. The ijth entry of (2.13) is $y_{ij} = \frac{1}{\eta} u_i u_j > 0$ whence $ij \in E$. This proves (2.12). Since the set of dual feasible slacks for ϑ' is larger than those for ϑ and ϑ^{+}, it follows that the dual SDPs for the formulations of ϑ and its two variants only have feasible solutions with rank-one slacks if G is complete.

One might argue that we have chosen an inappropriate formulation for the rank constraint. However, given the mandatory constraints (2.4), the formulation above is the most natural one. For completeness, we show in Appendix A that the rank constraint is not adequate either for another, more popular formulation of ϑ; in Appendix B.1, we also treat the rank constraint for the dual of the Max Cut SDP.

2.2. An improved dual integrality constraint. In view of our adopted embedding (2.1), let us draft the complete version of the (partial) chain of inequalities (1.1) as

(2.14a)
$$\sup\{\langle \hat{C}, \hat{X} \rangle : \langle \hat{A}_i, \hat{X} \rangle \leq b_i\ \forall i \in [m],\ \hat{X} \in \widehat{\mathbb{S}}_{+}^{n},\ \text{“}\hat{X} \text{ integral”}\}$$
(2.14b)
$$\leq \sup\{\langle \hat{C}, \hat{X} \rangle : \langle \hat{A}_i, \hat{X} \rangle \leq b_i\ \forall i \in [m],\ \hat{X} \in \widehat{\mathbb{S}}_{+}^{n}\}$$
(2.14c)
$$\leq \inf\{b^{\mathsf{T}} y : y \in \mathbb{R}_{+}^{m},\ \hat{S} = \textstyle\sum_{i=1}^{m} y_i \hat{A}_i - \hat{C} \in \widehat{\mathbb{S}}_{+}^{n}\}$$
(2.14d)
$$\leq \inf\{b^{\mathsf{T}} y : y \in \mathbb{Z}_{+}^{m},\ \hat{S} = \textstyle\sum_{i=1}^{m} y_i \hat{A}_i - \hat{C} \in \widehat{\mathbb{S}}_{+}^{n},\ \text{“}\hat{S} \text{ integral”}\}.$$

Assume that the system $\langle \hat{A}_i, \hat{X} \rangle \leq b_i$, $i \in [m]$, includes the constraints (2.4).

To define the *integrality constraint* for (2.14d), we shall consider the dual slack $\hat{S} = \sum_{i=1}^{m} y_i \hat{A}_i - \hat{C}$.

Definition 2.1. *Let $\hat{S}$ be feasible in* (2.14c). *We say that $\hat{S}$ is* integral *if $\hat{S}$ is a sum $\hat{S} = \sum_{k=1}^{N} \hat{S}_k$ of rank-one matrices $\hat{S}_1, \dots, \hat{S}_N \in \widehat{\mathbb{S}}^n_+$ such that, for each $k \in [N]$, we have*

$$\langle e_0 e_0^\mathsf{T}, \hat{S}_k \rangle = 1, \tag{2.15a}$$

$$\big\langle 2\,\mathrm{Sym}(e_j(e_j + e_0)^\mathsf{T}), \hat{S}_k \big\rangle = 0 \qquad \forall j \in V. \tag{2.15b}$$

Note that this is almost identical to the constraints in (2.4), except for the sign of e_0 in (2.15b). Equivalently, each $\hat{S}_k$ must have the form

$$\hat{S}_k = \begin{bmatrix} 1 & -s_k^\mathsf{T} \\ -s_k & S_k \end{bmatrix}$$

and satisfy $\mathrm{diag}(S_k) = s_k$. Since $\hat{S}_k$ has rank one, we must have $S_k = s_k s_k^\mathsf{T}$. Hence, the condition "$\hat{S}$ is *integral*" may be interpreted with a more combinatorial flavor as requiring $\hat{S}$ to have the form

$$\hat{S} = \sum_{K \in \mathcal{K}} \begin{bmatrix} -1 \\ \mathbb{1}_K \end{bmatrix} \begin{bmatrix} -1 \\ \mathbb{1}_K \end{bmatrix}^\mathsf{T}$$

for some *family* (i.e., multiset) $\mathcal{K}$ of subsets of $[n]$. Denote the power set of V by $\mathcal{P}(V)$. By denoting by $m\colon \mathcal{P}(V) \to \mathbb{Z}_+$ the multiplicity of each subset $K \subseteq V := [n]$ in $\mathcal{K}$, the condition "$\hat{S}$ is integral" becomes

$$\hat{S} = \sum_{A \subseteq V} m_A \begin{bmatrix} 1 & -\mathbb{1}_A^\mathsf{T} \\ -\mathbb{1}_A & \mathbb{1}_A \mathbb{1}_A^\mathsf{T} \end{bmatrix} \text{ for some } m\colon \mathcal{P}(V) \to \mathbb{Z}_+. \tag{D$\mathbb{Z}$}$$

The *integrality constraint* for (2.14a) is analogous.

Definition 2.2. *Let $\hat{X}$ be feasible in* (2.14b). *We say that $\hat{X}$ is* integral *if $\hat{X}$ is a sum $\hat{X} = \sum_{k=1}^{N} \hat{X}_k$ of rank-one matrices $\hat{X}_1, \dots, \hat{X}_N \in \widehat{\mathbb{S}}^n_+$ such that $\hat{X}_k$ satisfies* (2.4) *for each $k \in [N]$.*

As before, this integrality constraint can be described as

$$\hat{X} = \sum_{A \subseteq V} m_A \begin{bmatrix} 1 & \mathbb{1}_A^\mathsf{T} \\ \mathbb{1}_A & \mathbb{1}_A \mathbb{1}_A^\mathsf{T} \end{bmatrix} \text{ for some } m\colon \mathcal{P}(V) \to \mathbb{Z}_+. \tag{P$\mathbb{Z}$}$$

The usual rank constraint "$\mathrm{rank}(\hat{X}) = 1$" can be simply enforced by the linear constraint $\hat{X}_{00} = 1$.

Fact 2.3. *Let $\hat{X}$ be feasible in* (2.14b). *Then $\hat{X}$ is* integral *if and only if $\hat{X}_{00} = 1$.*

With these "semidefinite integrality" conditions in mind, we can state a semidefinite analogue of Theorem 1.1. To make the theorems syntactically more similar, we shall adopt a more compact notation for SDPs via linear maps: define $\mathcal{A}\colon \widehat{\mathbb{S}}^n \to \mathbb{R}^m$ by setting $[\mathcal{A}(\hat{X})]_i := \langle \hat{A}_i, \hat{X} \rangle$ for each $i \in [m]$, so that $\mathcal{A}(\hat{X}) \le b$ is equivalent to $\langle \hat{A}_i, \hat{X} \rangle \le b_i \ \forall i \in [m]$. Then the adjoint $\mathcal{A}^*\colon \mathbb{R}^m \to \widehat{\mathbb{S}}^n$ satisfies $\mathcal{A}^*(y) = \sum_{i=1}^m y_i \hat{A}_i$ for every $y \in \mathbb{R}^m$.

Theorem 2.4. *If $\hat{C} \in \widehat{\mathbb{S}}^n$ is a matrix, $\mathcal{A}\colon \widehat{\mathbb{S}}^V \to \mathbb{R}^m$ is a linear map, and $b \in \mathbb{R}^m$*

is a vector, then

$$\begin{aligned}
&\text{(ISDP)} && \sup\{\, \langle \hat{C}, \hat{X}\rangle : \mathcal{A}(\hat{X}) \le b,\ \hat{X} \textit{ satisfies } (\mathrm{P}\mathbb{Z})\} \\
&\text{(SDP)} && \le \sup\{\, \langle \hat{C}, \hat{X}\rangle : \mathcal{A}(\hat{X}) \le b,\ \hat{X} \in \widehat{\mathbb{S}}^n_+ \} \\
&\text{(SDD)} && \le \inf\{\, b^\mathsf{T} y : y \in \mathbb{R}^m_+,\ \hat{S} = \mathcal{A}^*(y) - \hat{C} \in \widehat{\mathbb{S}}^n_+ \} \\
&\text{(ISDD)} && \le \inf\{\, b^\mathsf{T} y : y \in \mathbb{Z}^m_+,\ \hat{S} = \mathcal{A}^*(y) - \hat{C} \textit{ satisfies } (\mathrm{D}\mathbb{Z})\},
\end{aligned}$$

and the middle (second) inequality holds with equality if either one of (SDP) *and* (SDD) *has a positive definite feasible solution and finite optimal value.*

The equality in Theorem 2.4 follows from the usual constraint qualification for SDP, namely the fact that the SDP satisfies the *relaxed Slater condition*; see, e.g., [7, Theorem 1.1].

We shall refer to (ISDD) as the *integer dual SDP* of (SDP). For convenience, we shall say that a feasible solution $(y, \hat{S})$ for (SDD) is *integral* if it is actually feasible in (ISDD), that is, if y is integral and $\hat{S}$ satisfies (D$\mathbb{Z}$). Integrality of y in (ISDD) shows why it is important to use integral matrices $\hat{A}_i$.

Let us set up the integer dual SDP of the SDP formulation (2.6) of LPs. If we require integrality from feasible solutions of (2.7), that is, if we add the constraint (D$\mathbb{Z}$) and further constrain η, u, y, and Z to be integral, then (2.7b) becomes equivalent to

$$\eta - b^\mathsf{T} y = \mathbb{1}^\mathsf{T} m, \tag{2.16a}$$

$$-u = -\sum_{A \subseteq V} m_A \mathbb{1}_A, \tag{2.16b}$$

$$\mathrm{Diag}(2u + A^\mathsf{T} y - c) = \sum_{A \subseteq V} m_A \mathbb{1}_A \mathbb{1}_A^\mathsf{T} + Z. \tag{2.16c}$$

At each feasible solution we have $Z \ge 0$, which implies that $\mathrm{supp}(m) \subseteq \binom{V}{1}$; we may always set $m_\varnothing \coloneqq 0$. Thus, the integer dual SDP of (2.6) can be written as

$$\min\{\, \mathbb{1}^\mathsf{T} u + b^\mathsf{T} y : A^\mathsf{T} y + u \ge c,\ y \in \mathbb{Z}^m_+,\ u \in \mathbb{Z}^n_+\}, \tag{2.17}$$

assuming A, b, and c to be integral. Hence, every feasible solution y for (ILD) yields a feasible solution for (2.16) with the same objective value by setting $u \coloneqq 0$. In the case of binary ILPs, we can say more.

FACT 2.5. *If* (ILP) *from Theorem* 1.1 *is* $\sup\{\, c^\mathsf{T} x : Ax \le b,\ 0 \le x \le \mathbb{1},\ x \in \mathbb{Z}^n\}$ *with A, b, and c integral, then the integer dual SDP of* (2.6) *is equivalent to* (ILD).

From our previous discussion after (2.7), our new notion of dual integrality passes the test of behaving nicely with respect to ILPs. Next we will see that it surpasses the rank-one constraint by showing that it yields the "natural" combinatorial dual for the theta function.

Let $G = (V, E)$ be a graph, and let $w \colon V \to \mathbb{Z}$. The *clique covering number* $\overline{\chi}(G, w)$ is defined as

$$\overline{\chi}(G, w) \coloneqq \min\Big\{\, \mathbb{1}^\mathsf{T} m : m \in \mathbb{Z}^{\mathcal{K}(G)}_+,\ \sum_{K \in \mathcal{K}(G)} m_K \mathbb{1}_K \ge w \Big\}. \tag{2.18}$$

Every feasible solution of (2.18) is a *clique cover* of G with respect to w. We now show that the integer dual SDPs for each of the SDP formulations (2.9), (2.8), and (2.10) are extended formulations for $\overline{\chi}(G, w)$.

Proposition 2.6. *Let $G = (V, E)$ be a graph, and let $w\colon V \to \mathbb{Z}$. Then the following hold:*

(i) *If $m\colon \mathcal{P}(V) \to \mathbb{Z}_+$ is a clique cover of G with respect to w, then there exists an integral dual solution $(\hat{S}, \eta, u, y, z)$ for (2.11) such that (D$\mathbb{Z}$) holds for $\hat{S}$ and $m\colon \mathcal{P}(V) \to \mathbb{Z}_+$, $\eta = \mathbb{1}^{\mathsf{T}} m$, and $y \in \mathbb{R}^E_+ \oplus 0$.*

(ii) *If $(\hat{S}, \eta, u, y, z)$ is an integral dual solution for (2.11) and (D$\mathbb{Z}$) holds for $\hat{S}$ and $m\colon \mathcal{P}(V) \to \mathbb{Z}_+$, then $\eta = \mathbb{1}^{\mathsf{T}} m$ and m is a clique cover of G with respect to w.*

Proof. To restrict ourselves to integral dual solutions for (2.11), we require the dual slack $\hat{S}$ to satisfy (D$\mathbb{Z}$), and η, u, y, and z to be integral. In this case, (2.11b) can be rewritten as $\eta = \mathbb{1}^{\mathsf{T}} m$, $u = \sum_{A \subseteq V} m_A \mathbb{1}_A$, and

$$\text{Diag}(2u - z - w) + \sum_{ij \in \binom{V}{2}} 2y_{ij} \,\text{Sym}(e_i e_j^{\mathsf{T}}) = \sum_{A \subseteq V} m_A \mathbb{1}_A \mathbb{1}_A^{\mathsf{T}}. \tag{2.19}$$

Applying diag to both sides of (2.19) yields $2u - z - w = \sum_{A \subseteq V} m_A \mathbb{1}_A = u$. Let $i, j \in V$ be distinct. The ijth entry of (2.19) is $y_{ij} = \mathbb{1}^{\mathsf{T}}_{\binom{V}{ij \subseteq}} m$. Hence, the integer dual SDP of (2.8) can be written as

$$\text{Minimize} \quad \mathbb{1}^{\mathsf{T}} m \tag{2.20a}$$

$$\text{subject to} \quad m\colon \mathcal{P}(V) \to \mathbb{Z}_+, \tag{2.20b}$$

$$w + z = u = \sum_{A \subseteq V} m_A \mathbb{1}_A, \tag{2.20c}$$

$$y_{ij} = \mathbb{1}^{\mathsf{T}}_{\binom{V}{ij \subseteq}} m \qquad \forall ij \in \tbinom{V}{2}, \tag{2.20d}$$

$$\hat{S} = \sum_{A \subseteq V} m_A \begin{bmatrix} 1 & -\mathbb{1}_A^{\mathsf{T}} \\ -\mathbb{1}_A & \mathbb{1}_A \mathbb{1}_A^{\mathsf{T}} \end{bmatrix}, \tag{2.20e}$$

$$\hat{S} \in \widehat{\mathbb{S}}^V_+, \ \eta \in \mathbb{Z}, \ u \in \mathbb{Z}^V, \ z \in \mathbb{Z}^V_+, \ y \in \mathbb{Z}^E \oplus -\mathbb{Z}^{\overline{E}}_+. \tag{2.20f}$$

We may now prove the result. We start with (i). Suppose $m\colon \mathcal{P}(V) \to \mathbb{Z}_+$ is a clique cover of G with respect to w. Set $u := \sum_{A \subseteq V} m_A \mathbb{1}_A$, $z := u - w \geq 0$, and $\eta := \mathbb{1}^{\mathsf{T}} m$. Define y and $\hat{S}$ as in (2.20d) and (2.20e), respectively. Since $\text{supp}(m) \subseteq \mathcal{K}(G)$, we get $y \in \mathbb{Z}^E_+ \oplus 0$. Hence, $(\hat{S}, \eta, u, y, z)$ is feasible in (2.20) and satisfies the desired properties in (i).

For (ii), let $(\hat{S}, \eta, u, y, z)$ be feasible in (2.20). If $ij \in \overline{E}$, then $y_{ij} \leq 0$ together with (2.20d) yield $m_A = 0$ for each $A \subseteq V$ such that $i, j \in A$. Hence, $m_A > 0$ and $i, j \in A \subseteq V$ imply $ij \in E$, i.e., $\text{supp}(m) \subseteq \mathcal{K}(G)$, whence m is a clique cover of G. This proves (ii). ∎

The statement of the above result makes it clear that the integer dual SDPs of ϑ, ϑ', and ϑ^+ are all equivalent to the clique covering problem.

We have just seen that not only does the integer dual SDP have a feasible solution for every graph, but it is actually equivalent to a natural combinatorial optimization problem. In fact, the clique covering problem is the *right* dual problem for the maximum stable set problem at least for the very rich class of perfect graphs; see, e.g., [31, Chap. 67].

Now that we have a sensible notion of integrality for the dual SDP, we go back to the chain from Theorem 2.4. Motivated by the notion of total dual integrality that was so powerful for proving equality throughout in the chain from Theorem 1.1, and which was based on Theorem 1.2 and Corollary 1.3, we shall prove a generalized version of the latter corollary in the next section.

3. Integrality in convex relaxations. In this section, we generalize Corollary 1.3 to compact convex sets, which will motivate the definition of total dual integrality for SDPs in the next section. Denote the *support function* of a convex set $\mathscr{C} \subseteq \mathbb{R}^n$ by

$$\sigma_{\mathscr{C}}(w) \coloneqq \sup_{x\in\mathscr{C}} \langle w, x\rangle \in [-\infty, +\infty] \qquad \forall w \in \mathbb{R}^n. \tag{3.1}$$

THEOREM 3.1. *If $\mathscr{C} \subseteq \mathbb{R}^n$ is a compact convex set, then*
$\mathscr{C} = \{\, x \in \mathbb{R}^n : w^{\mathsf{T}} x \le \sigma_{\mathscr{C}}(w)\, \forall w \in \mathbb{Z}^n \}$.

Proof. We may assume that $\mathscr{C} \neq \varnothing$. The inclusion $\subseteq$ is obvious. For the reverse inclusion, we start by noting that the right-hand side (RHS) is equal to $\mathscr{C}' \coloneqq \{\, x \in \mathbb{R}^n : w^{\mathsf{T}} x \le \sigma_{\mathscr{C}}(w)\, \forall w \in \mathbb{Q}^n \}$ by positive homogeneity of $\sigma_{\mathscr{C}}(\cdot)$. Let $\bar{x} \in \mathscr{C}'$. Let $\bar{w} \in \mathbb{R}^n$, and let $(w_k)_{k\in\mathbb{N}}$ be a sequence in $\mathbb{Q}^n$ converging to $\bar{w}$. Then $w_k^{\mathsf{T}}\bar{x} \le \sigma_{\mathscr{C}}(w_k)$ for every $k \in \mathbb{N}$, which in the limit yields $\bar{w}^{\mathsf{T}}\bar{x} \le \sigma_{\mathscr{C}}(\bar{w})$ by the (Lipschitz) continuity of the support function (apply Corollary 13.3.3 of [26] to the function $\sigma_{\mathscr{C}}(\cdot)$, where $\mathscr{C}$ is a compact convex set). Hence $\mathscr{C}' \subseteq \{\, x \in \mathbb{R}^n : w^{\mathsf{T}} x \le \sigma_{\mathscr{C}}(w)\, \forall w \in \mathbb{R}^n \} = \mathscr{C}$, where the latter equation follows from Theorem 13.1 of [26]. ∎

Theorem 3.1 holds more generally for pointed closed convex sets; a proof using elementary convex analysis can be found in [9]. The obvious generalization of Theorem 3.1 to unbounded (even polyhedral) convex sets is false; e.g., let $\mathscr{C}$ be any closed halfspace with a normal vector containing both rational and irrational entries.

Let $\mathscr{C} \subseteq \mathbb{R}^n$ be a convex set. The *Gomory–Chvátal closure* and the *integer hull* of $\mathscr{C}$ are, respectively,

$$\mathrm{CG}(\mathscr{C}) \coloneqq \left\{\, x \in \mathbb{R}^n : w^{\mathsf{T}} x \le \lfloor \sigma_{\mathscr{C}}(w) \rfloor\, \forall w \in \mathbb{Z}^n \right\}, \tag{3.2}$$

$$\mathscr{C}_I \coloneqq \mathrm{conv}(\mathscr{C} \cap \mathbb{Z}^n). \tag{3.3}$$

THEOREM 3.2 ([29]). *If $\mathscr{C} \subseteq \mathbb{R}^n$ is a bounded convex set, then $\mathrm{CG}^k(\mathscr{C}) = \mathscr{C}_I$ for some integer $k \ge 1$.*

We now generalize Corollary 1.3 (see [2, 6, 5] for recent generalizations in similar directions).

COROLLARY 3.3. *If $\mathscr{C} \subseteq \mathbb{R}^n$ is a nonempty compact convex set, then $\mathscr{C} = \mathscr{C}_I$ if and only if $\sigma_{\mathscr{C}}(w) \in \mathbb{Z}$ for every $w \in \mathbb{Z}^n$.*

Proof. Necessity is clear. For sufficiency, note that
$\mathscr{C} = \{\, x \in \mathbb{R}^n : w^{\mathsf{T}} x \le \lfloor \sigma_{\mathscr{C}}(w) \rfloor\, \forall w \in \mathbb{Z}^n \} = \mathrm{CG}(\mathscr{C})$
by Theorem 3.1. Hence, $\mathrm{CG}^k(\mathscr{C}) = \mathscr{C}$ for every $k \ge 1$, so $\mathscr{C} = \mathscr{C}_I$ by Theorem 3.2. ∎

Corollary 3.3 provides a blueprint to define sensible, algebraic notions of total dual integrality for convex formulations in certain formats, which depend on some arbitrary choice of embedding. We shall do this in the next section for SDPs, but it is plausible that similar notions could be useful for conic optimization problems over other cones, e.g., the second-order cone. And whereas Corollary 3.3 is a purely geometric result, independent of algebraic representations (if any!), it does fully characterize integrality through a total criterion using duality. Thus, it seems fair to regard it as a very general, geometric notion of total dual integrality, which can be seen as a precursor to algebraic notions of total dual integrality for specific embeddings.

Characterizations of exactness of convex relaxations for sets of integer points can naturally involve (convex) geometry in general, boundary structure of convex sets in particular (including polyhedral combinatorics), diophantine equations (number theory), and convex analysis and optimization. Next, we summarize some of the consequences of our geometric characterization (Corollary 3.3) of exactness for convex relaxations of integral polytopes. The next theorem, well known in the special case of LP relaxations, provides equivalent characterizations of integrality in terms of the facial structure of the convex relaxation, optimum values of linear functions over the relaxation, optimal solutions of the linear optimization problems over the relaxation, diophantine equations, and gauge functions in convex optimization and analysis.

In the next result, the *polar* and the *gauge* functions of a subset $\mathscr{C}$ of a Euclidean space $\mathbb{E}$ are, respectively,

$$\mathscr{C}^\circ := \{\, y \in \mathbb{E} : \langle y, x\rangle \le 1 \,\forall x \in \mathscr{C} \},$$
$$\gamma_{\mathscr{C}}(x) := \inf\{\, \eta \in \mathbb{R}_{++} : \tfrac{1}{\eta} x \in \mathscr{C} \}.$$

COROLLARY 3.4. *Let $\mathscr{C} \subseteq \mathbb{R}^n$ be a closed convex set with the origin in its interior. If $\gamma_{\mathscr{C}}(w) \in \mathbb{Z}$ for each $w \in \mathbb{Z}^n$, then $\mathscr{C}$ is a polyhedron.*

Proof. Suppose that $\gamma_{\mathscr{C}}(w) \in \mathbb{Z}$ for each $w \in \mathbb{Z}^n$. Since $\mathscr{D} := \mathscr{C}^\circ$ is a compact convex set and $\gamma_{\mathscr{C}} = \sigma_{\mathscr{D}}$ (see, e.g., [26, Theorem 14.5]), Corollary 3.3 shows that $\mathscr{D}$ is a polytope. Hence, $\mathscr{C} = \mathscr{D}^\circ$ is a polyhedron. ∎

A convex subset $\mathscr{F}$ of a convex set $\mathscr{C}$ is a *face* of $\mathscr{C}$ if, for every $x, y \in \mathscr{C}$ such that the open line segment $(x, y) := \{\, \lambda x + (1-\lambda) y : \lambda \in (0,1)\}$ between x and y meets $\mathscr{F}$, we have $x, y \in \mathscr{F}$. A nonempty face of $\mathscr{C}$ which does not contain another nonempty face of $\mathscr{C}$ is a *minimal face* of $\mathscr{C}$. If $w \in \mathbb{R}^n \setminus \{0\}$ and $\beta \in \mathbb{R}$, we say that $\mathscr{H} := \{\, x \in \mathbb{R}^n : w^\mathsf{T} x \le \beta\}$ is a *supporting halfspace* of $\mathscr{C}$ if $\mathscr{C} \subseteq \mathscr{H}$ and $w^\mathsf{T} x = \beta$ for at least one $x \in \mathscr{C}$; in this case we also say that $\{\, x \in \mathbb{R}^n : w^\mathsf{T} x = \beta\}$ is a *supporting hyperplane* of $\mathscr{C}$. The intersection of $\mathscr{C}$ with any of its supporting hyperplanes is a face of $\mathscr{C}$; such faces (as well as the empty face) are *exposed*. Faces of polyhedra are well known to be exposed, but this need not be the case for compact convex sets.

THEOREM 3.5. *Let $\mathscr{C}$ be a nonempty compact convex set in $\mathbb{R}^n$. Then, the following are equivalent:*

(i) $\mathscr{C} = \mathscr{C}_I$;
(ii) *every nonempty face of $\mathscr{C}$ contains an integral point;*
(iii) *every minimal face of $\mathscr{C}$ contains an integral point;*
(iv) *for every $w \in \mathbb{R}^n$, we have that $\max\{\, \langle w, x\rangle : x \in \mathscr{C}\}$ is attained by an integral point;*
(v) *for every $w \in \mathbb{Z}^n$, we have $\max\{\, \langle w, x\rangle : x \in \mathscr{C}\} \in \mathbb{Z}$;*
(vi) *every rational supporting hyperplane for $\mathscr{C}$ contains integral points;*
(vii) *for each $x_0 \in \mathscr{C}$ and for each $w \in \mathbb{Z}^n$, we have $\langle w, x_0\rangle + \gamma_{(\mathscr{C}-x_0)^\circ}(w) \in \mathbb{Z}$;*
(viii) *there exists $x_0 \in \mathscr{C}$ such that, for each $w \in \mathbb{Z}^n$, $\langle w, x_0\rangle + \gamma_{(\mathscr{C}-x_0)^\circ}(w) \in \mathbb{Z}$.*

Proof. (i) ⇒ (ii): Since $\mathscr{C}$ is compact, it is bounded. Therefore, $\mathscr{C} = \mathscr{C}_I$ implies that $\mathscr{C}$ is a polytope. Every nonempty face of $\mathscr{C}$ contains an extreme point of $\mathscr{C}$ and every extreme point of $\mathscr{C} = \mathscr{C}_I$ is integral.

(ii) ⇒ (iii): Immediate.

(iii) ⇒ (iv): Suppose every minimal face of $\mathscr{C}$ contains an integral point. Let $w \in \mathbb{R}^n$. Then, since $\mathscr{C}$ is nonempty, compact, and convex, $\arg\max_{x\in\mathscr{C}} \langle w, x\rangle =: \mathscr{F}$

is a nonempty (exposed) face of $\mathscr{C}$. Every minimal face contained in $\mathscr{F}$ contains an integral point (by part (iii)); hence, $\mathscr{F}$ contains an integral point.

(iv) $\Rightarrow$ (v): Suppose $\mathscr{C}$ satisfies (iv). Let $w \in \mathbb{Z}^n$. Then, by (iv), there exists $\bar{x} \in \mathscr{C} \cap \mathbb{Z}^n$ such that $\max_{x\in\mathscr{C}} \langle w, x\rangle = \langle w, \bar{x}\rangle$. Since w and $\bar{x}$ are integral, it follows that $\max_{x\in\mathscr{C}} \langle w, x\rangle \in \mathbb{Z}$.

(v) $\Rightarrow$ (vi): Suppose $\mathscr{C}$ has the property (v). Let $w \in \mathbb{Q}^n$. Define $\mathscr{F} \coloneqq \arg\max_{x\in\mathscr{C}} \langle w, x\rangle$. Let μ be a positive rational such that $\mu w \in \mathbb{Z}^n$ and $\gcd(\mu w_1, \dots, \mu w_n) = 1$. Then, $\arg\max_{x\in\mathscr{C}} \langle \mu w, x\rangle = \mathscr{F}$. By property (v), $\beta \coloneqq \max_{x\in\mathscr{C}} \langle \mu w, x\rangle \in \mathbb{Z}$. Since

$$\{\, x \in \mathbb{Z}^n : \langle \mu w, x\rangle = \beta \,\} \neq \varnothing \iff \gcd(\mu w_1, \dots, \mu w_n) \text{ divides } \beta,$$

and we have $\gcd(\mu w_1, \dots, \mu w_n) = 1$, we are done.

(vi) $\iff$ (i): Suppose $\mathscr{C}$ has property (vi). Then, for every $w \in \mathbb{Z}^n$, $\sigma_{\mathscr{C}}(w) \in \mathbb{Z}$. Therefore, by Corollary 3.3, $\mathscr{C} = \mathscr{C}_I$. The converse also follows from Corollary 3.3.

(v) $\iff$ (vii) $\iff$ (viii): Let $x_0 \in \mathscr{C}$ and $w \in \mathbb{Z}^n$. Set $\widetilde{\mathscr{C}} \coloneqq \mathscr{C} - x_0$. Then

$$\sigma_{\mathscr{C}}(w) = \langle w, x_0\rangle + \sigma_{\widetilde{\mathscr{C}}}(w) = \langle w, x_0\rangle + \min\{\, \eta \in \mathbb{R}_+ : \langle w, x\rangle \leq \eta \,\forall x \in \widetilde{\mathscr{C}} \,\},$$

where in the last equation we use the fact that $0 \in \widetilde{\mathscr{C}}$ to add the constraint $\eta \in \mathbb{R}_+$. Finally, note that

$$\min\{\, \eta \in \mathbb{R}_+ : \langle w, x\rangle \leq \eta \,\forall x \in \widetilde{\mathscr{C}} \,\} = \inf\{\, \eta \in \mathbb{R}_{++} : \tfrac{1}{\eta} w \in \widetilde{\mathscr{C}}^{\circ} \,\} = \gamma_{\widetilde{\mathscr{C}}}(w). \qquad \square$$

In the quite common case that $0 \in \mathscr{C}$, Theorem 3.5 shows that $\mathscr{C}_I = \mathscr{C}$ if and only if, for each $w \in \mathbb{Z}^n$, we have $\gamma_{\mathscr{C}^\circ}(w) \in \mathbb{Z}$. Thus, integrality of $\mathscr{C}$ can be characterized using the integrality of the gauge function of its polar; this is related to Corollary 3.4.

Just as Theorem 1.2 motivates the definition of total dual integrality for LP formulations, one may use Corollary 3.3 to define total dual integrality more generally. Next, we shall define it for SDP formulations.

4. Total dual integrality for SDPs. The concept of total dual integrality in LPs is an *algebraic* notion, rather than a *geometric* one, in the following sense: for a rational polyhedron $P \coloneqq \{\, x \in \mathbb{R}^n : Ax \leq b \,\}$, it is not the geometric object P that is TDI, but rather the defining system $Ax \leq b$, which is not uniquely determined by P. For instance, the perfect matching polytope of a graph $G = (V, E)$ can be described either by the system

$$x_e \geq 0 \qquad \forall e \in E, \tag{4.1a}$$

$$\mathbb{1}_{\delta(v)}^{\mathsf{T}} x = 1 \qquad \forall v \in V, \tag{4.1b}$$

$$\mathbb{1}_{\delta(U)}^{\mathsf{T}} x \geq 1 \qquad \forall U \subseteq V, \text{ such that } |U| \text{ is odd}, \tag{4.1c}$$

or by the system obtained from (4.1) by replacing (4.1c) with the constraints $\mathbb{1}_{E[U]}^{\mathsf{T}} x \leq \left\lfloor \frac{1}{2}|U| \right\rfloor$ for each odd $U \subseteq V$. However, the latter system is TDI, whereas the former is not; see [31, sect. 25.4]. In respect to this algebraic dependence, our choice of embedding for SDPs using (2.4) is the *only* arbitrary choice that we make, though as argued, this choice is well justified and natural. The total dual integrality results from section 3 do not rely on any algebraic representation, since they are *geometric*. In this respect, our convex theory of total dual integrality is *complete*, at least for compact sets.

Whereas integrality of a polyhedron and total dual integrality of a defining system are not the same, they are related as follows.

THEOREM 4.1 ([30, Theorem 22.6]). *Every rational polyhedron is defined by a TDI system $Ax \leq b$ with A integral. Moreover, if P is integral, then b may be chosen integral.*

Transforming *any* rational system $Ax \leq b$ into a TDI system is simple: Giles and Pulleyblank [12] showed that there is a positive integer t such that the system $(\frac{1}{t}A)x \leq \frac{1}{t}b$ is TDI. Thus, the "moreover" part of Theorem 4.1 is its most critical part, since without integrality of b we do not recover integrality from Theorem 1.2. The proof of this part relies on a sharper way of making a system TDI, namely on completing Hilbert bases. Giles and Pulleyblank's result shows, however, that TDIness is rather unintuitive and not robust.

Next we move on to define a notion of total dual integrality for SDP formulations. We want to define when the system $\mathcal{A}(\hat{X}) \leq b$, $\hat{X} \succeq 0$ for (SDP) is TDI, but there is a further complication. We may not need the dual SDP to have an "integral solution" for *every* integral objective function $\hat{X} \mapsto \langle \hat{C}, \hat{X} \rangle$. As the formulation (2.9) shows, for the Lovász ϑ function we are only interested in objective functions of the form $\hat{X} \mapsto \langle \mathrm{Diag}(0 \oplus w), \hat{X} \rangle$, perhaps with $w \in \mathbb{R}^V$ integral. The same remark can be made about the diagonal embedding (2.6) of LPs as SDPs. In these cases, one is interested only in the *diagonal* part of the variable $\hat{X}$, and the lifting $w \mapsto \mathrm{Diag}(0 \oplus w)$ embeds in matrix space only the objective functions that matter to us. This arises from the fact that we are essentially dealing with *extended formulations*. However, when we look at the Max Cut SDP in section 6, we shall only be interested in objective functions of the form $\hat{X} \mapsto \langle 0 \oplus \mathcal{L}_G(w), \hat{X} \rangle$, where $\mathcal{L}_G(w) \in \mathbb{S}^V$ is a weighted Laplacian matrix of the input graph G on vertex set V, to be defined later; as before, $\hat{X} \in \widehat{\mathbb{S}}^V_+$ is the variable. In this case, one might argue that we are only interested in the *off-diagonal* (!) entries of the variable $\hat{X}$. Thus, when defining semidefinite TDIness, we shall need to refer to which objective functions (that is, which projection of the feasible region) we care about. (This notion of TDIness coupled with extended formulations already leads to an interesting generalization of TDIness in the polyhedral case, as we discuss in section 7.)

We may now define a semidefinite notion of total dual integrality which plays well with extended formulations. Below, the map $\mathcal{L}$ is a lifting map, such as $w \mapsto \mathrm{Diag}(0 \oplus w)$ and $w \mapsto 0 \oplus \mathcal{L}_G(w)$ from above. The corresponding projection, which will be the adjoint $\mathcal{L}^*$ of the lifting $\mathcal{L}$, will appear in Theorem 4.3 below.

DEFINITION 4.2. *Let $\mathcal{L}\colon \mathbb{R}^k \to \widehat{\mathbb{S}}^n$ be a linear map. The system $\mathcal{A}(\hat{X}) \leq b$, $\hat{X} \succeq 0$ is* totally dual integral (TDI) through $\mathcal{L}$ *if, for every integral $c \in \mathbb{Z}^k$, the SDP dual to $\sup\{\, \langle \mathcal{L}(c), \hat{X} \rangle : \mathcal{A}(\hat{X}) \leq b,\ \hat{X} \succeq 0\}$ has an integral optimal solution whenever it has an optimal solution.*

For convenience, we use the term "TDI" to refer to two separate notions, one for linear inequality systems of the form $Ax \leq b$, and another one for semidefinite systems of the form $\mathcal{A}(\hat{X}) \leq b$, $\hat{X} \succeq 0$; the context shall make it clear to which notion we are referring.

THEOREM 4.3. *Let $\mathcal{A}(\hat{X}) \leq b$, $\hat{X} \succeq 0$ be totally dual integral (TDI) through a linear map $\mathcal{L}\colon \mathbb{R}^k \to \widehat{\mathbb{S}}^n$. Set $\widehat{\mathscr{C}} := \{\, \hat{X} \in \widehat{\mathbb{S}}^n_+ : \mathcal{A}(\hat{X}) \leq b\}$ and $\mathscr{C} := \mathcal{L}^*(\widehat{\mathscr{C}}) \subseteq \mathbb{R}^k$. If b is integral, $\mathscr{C}$ is compact, and $\widehat{\mathscr{C}}$ has a positive definite matrix, then $\mathscr{C} = \mathscr{C}_I$.*

Proof. Let $w \in \mathbb{Z}^k$. Then

$$\sigma_{\mathscr{C}}(w) = \max_{\hat{X} \in \widehat{\mathscr{C}}} \langle w, \mathcal{L}^*(\hat{X}) \rangle = \max_{\hat{X} \in \widehat{\mathscr{C}}} \langle \mathcal{L}(w), \hat{X} \rangle. \tag{4.2}$$

The latter SDP satisfies the relaxed Slater condition by assumption and its optimal value is finite and attained by compactness of $\mathscr{C}$. By SDP strong duality, the dual SDP has an optimal solution. Since $\mathcal{A}(\hat{X}) \le b$, $\hat{X} \succeq 0$ is TDI through $\mathcal{L}$, the dual SDP has an integral optimal solution $(y^*, \hat{S}^*)$. Hence, $\sigma_{\mathscr{C}}(w) = b^{\mathsf{T}} y^*$ and so $\sigma_{\mathscr{C}}(w) \in \mathbb{Z}$, since b is integral. It follows from Corollary 3.3 that $\mathscr{C} = \mathscr{C}_I$. □

We have established that total dual integrality is sufficient for exact (primal) representations. We next describe conditions under which the chain of inequalities in Theorem 2.4 holds with equality throughout, thus completing our discussion in section 1 regarding equality throughout in Theorem 1.1.

Again there is a more involved setup due to our choice of embedding (2.1). Let $\mathscr{C} \subseteq [0,1]^k$ be a convex set. Let $\mathcal{L}\colon \mathbb{R}^k \to \widehat{\mathbb{S}}^n$ be a linear map, and let $\widehat{\mathscr{C}} \subseteq \widehat{\mathbb{S}}^n$. We say that $\widehat{\mathscr{C}}$ is a *rank-one embedding* of $\mathscr{C}_I$ via $\mathcal{L}$ if for each $\bar{x} \in \{0,1\}^k$ there exists $\hat{X} \in \widehat{\mathscr{C}}$ such that $\bar{x} = \mathcal{L}^*(\hat{X})$ and $\hat{X}$ has the form (2.1) for some $U \subseteq V := [n]$. One may think of $\widehat{\mathscr{C}}$ as a convex set in (lifted) matrix space, e.g., the feasible region of an SDP, described algebraically by a linear system $\mathcal{A}(\hat{X}) \le b$, $\hat{X} \succeq 0$ that includes (2.4). Then to have the (lifted) rank-constrained SDP formulation $\sup\{\, \langle \mathcal{L}(w), \hat{X}\rangle : \hat{X} \in \widehat{\mathscr{C}},\ \mathrm{rank}(\hat{X}) = 1\}$ be a correct relaxation for the combinatorial optimization problem $\max\{\, w^{\mathsf{T}} x : x \in \mathscr{C} \cap \{0,1\}^k\}$ requires the conditions for $\widehat{\mathscr{C}}$ to be a rank-one embedding of $\mathscr{C}_I$.

If $\mathcal{L}\colon w \in \mathbb{R}^V \mapsto 0 \oplus \mathrm{Diag}(w)$ and $\mathscr{C} \subseteq [0,1]^V$, to say that the set $\widehat{\mathscr{C}} \subseteq \widehat{\mathbb{S}}^n$ defined by a system $\mathcal{A}(\hat{X}) \le b$, $\hat{X} \succeq 0$ is a rank-one embedding of $\mathscr{C}_I$ via $\mathcal{L}$ requires that, for each $\bar{x} \in \mathscr{C} \cap \{0,1\}^V$, we have

$$\mathcal{A}\left(\begin{bmatrix} 1 & \bar{x}^{\mathsf{T}} \\ \bar{x} & \bar{x}\bar{x}^{\mathsf{T}} \end{bmatrix}\right) \le b.$$

THEOREM 4.4. *Let $\mathcal{A}(\hat{X}) \le b$, $\hat{X} \succeq 0$ be totally dual integral (TDI) through a linear map $\mathcal{L}\colon \mathbb{R}^k \to \widehat{\mathbb{S}}^n$ such that b is integral. Suppose that $\widehat{\mathscr{C}} := \{\, \hat{X} \in \widehat{\mathbb{S}}^n_+ : \mathcal{A}(\hat{X}) \le b\}$ has a positive definite matrix and that $\mathscr{C} := \mathcal{L}^*(\widehat{\mathscr{C}}) \subseteq [0,1]^k$ is compact. If $\widehat{\mathscr{C}}$ is a rank-one embedding of $\mathscr{C}_I$ via $\mathcal{L}$, then for every $w \in \mathbb{Z}^k$, equality holds throughout in the chain of inequalities from Theorem 2.4 for $\hat{C} := \mathcal{L}(w)$, all optimum values are equal to*

$$\max\{\, w^{\mathsf{T}} x : x \in \mathscr{C}_I\}, \tag{4.3}$$

and all suprema and infima are attained.

Proof. Fix $w \in \mathbb{Z}^k$ and set $\hat{C} := \mathcal{L}(w)$ throughout the proof. Note that the optimal value of (SDP) is bounded above, since each $\hat{X} \in \widehat{\mathscr{C}}$ has objective value $\langle \mathcal{L}(w), \hat{X}\rangle = \langle w, \mathcal{L}^*(\hat{X})\rangle \le \sigma_{\mathscr{C}}(w) < \infty$ by compactness. Since the relaxed Slater condition holds by assumption, SDP strong duality shows that (SDD) has an optimal solution and hence is feasible. Together with the TDI assumption, this shows that (SDP), (SDD), and (ISDD) have the same optimal values and the latter two are attained.

It remains to prove that (SDP), (ISDP), and (4.3) have the same optimal values and are attained. Let $\bar{x}$ be an optimal solution for $\max\{\, w^{\mathsf{T}} x : x \in \mathscr{C} \cap \{0,1\}^k\}$. Then there exists $\bar{X} \in \widehat{\mathscr{C}}$ that satisfies (P$\mathbb{Z}$) such that $\bar{x} = \mathcal{L}^*(\bar{X})$. Then the optimal value of (4.3) is $w^{\mathsf{T}}\bar{x} = \langle w, \mathcal{L}^*(\bar{X})\rangle = \langle \hat{C}, \bar{X}\rangle$, which is upper bounded by the optimal value of (ISDP). On the other hand, as shown above, the optimal value of (SDP) is upper bounded by $\sigma_{\mathscr{C}}(w) = \sigma_{\mathscr{C}_I}(w) = w^{\mathsf{T}}\bar{x}$ since $\mathscr{C} = \mathscr{C}_I$ by Theorem 4.3. Hence, $\bar{x}$ is optimal in (4.3), and $\bar{X}$ is optimal in (ISDP) and (SDP), all with the same objective values. ∎

Naturally, any other choice of (i) embedding in some lifted space and (ii) integrality conditions would require an adaptation of the definition of "rank-one embedding" of $\mathscr{C}_I$ via a lifting map, if only to ensure that the lifted representation $\widehat{\mathscr{C}}$ is a correct formulation of (4.3).

The next result characterizes TDIness for the diagonal embedding (2.6) of LPs. It shows that our notion of semidefinite TDIness is the same as the polyhedral notion, at least in $[0,1]^n$.

THEOREM 4.5. *Let $Ax \leq b$ be a rational system of linear inequalities. The system defining* (2.6) *is TDI through $w \in \mathbb{R}^V \mapsto \mathrm{Diag}(0 \oplus w)$ if and only if the system $Ax \leq b$, $0 \leq x \leq \mathbb{1}$ is TDI.*

Proof. Immediate from Fact 2.5. ∎

Together with Theorem 4.5, Theorem 4.4 yields a richer version of equality throughout the chain from Theorem 1.1, since it includes the LP case via the diagonal embedding (2.6) as well as other, lifted formulations; see, e.g., Theorem 5.2 in the next section. Theorem 4.4 yields further results when the lifting map involves the Laplacian of a graph G, i.e., when $\mathcal{L}$ has the form $w \mapsto 0 \oplus \mathcal{L}_G(w)$ as discussed before Definition 4.2. In this case, we leave it to the reader to check exactly how the set $\mathscr{C}$ must be related to the cuts of G.

COROLLARY 4.6. *Let $P \subseteq [0,1]^V$ be an integral polytope such that $P \neq \{0\}$. Then there is $U \subseteq V$ and a system $\mathcal{A}(\hat{X}) \leq b$, $\hat{X} \succeq 0$ on $\widehat{\mathbb{S}}^U$ that is TDI through $\mathcal{L}\colon w \in \mathbb{R}^V \mapsto \mathrm{Diag}(0 \oplus w\restriction_U) \in \widehat{\mathbb{S}}^U$ and with at least one positive definite solution such that, for $\widehat{\mathscr{C}} \coloneqq \{\, \hat{X} \in \widehat{\mathbb{S}}^U_+ : \mathcal{A}(\hat{X}) \leq b\}$, we have $\mathcal{L}^*(\widehat{\mathscr{C}}) = P$.*

Proof. Write $P = Q \oplus 0$ where $Q \subseteq \mathbb{R}^U$ for some $U \subseteq V$ is such that, for each $i \in U$, there is $x \in P$ such that $x_i > 0$. Hence, there is $\bar{x} \in P$ such that $\mathrm{supp}(\bar{x}) = U$. The result now follows from Theorem 4.1 and Theorem 4.5. ∎

Hence, our semidefinite notion of total dual integrality is *complete* for polytopes in $[0,1]^V$; note also that Corollary 4.6 provides a relaxed Slater point as prescribed in Theorem 4.3. One may ask whether we gain anything in terms of efficiency, e.g., compact representations. The next section provides a positive answer to this question.

5. Integrality in the theta function formulation.

In this section, we prove that the formulation (2.9) for the Lovász ϑ function of a graph G is TDI through the appropriate lifting if and only if G is perfect.

Let $G = (V,E)$ be a graph. For each $w\colon V \to \mathbb{R}$, the *weighted stability number* is $\alpha(G,w) \coloneqq \max\{\, w^{\mathsf{T}}\mathbb{1}_U : U \subseteq V \text{ stable}\}$. A subset $\mathscr{C}$ of $\mathbb{R}^n_+$ is a *convex corner* if $\mathscr{C}$ is a compact convex set with nonempty interior and such that $0 \leq y \leq x \in \mathscr{C}$ implies $y \in \mathscr{C}$. Associate with each graph $G = (V,E)$ the following convex corners:

$$\begin{aligned}
\mathrm{STAB}(G) &\coloneqq \mathrm{conv}\{\, \mathbb{1}_U : U \subseteq V \text{ stable}\},\\
\mathrm{TH}'(G) &\coloneqq \{\, \mathrm{diag}(\hat{X}[V]) : \hat{X} \text{ feasible in (2.8)}\},\\
\mathrm{TH}(G) &\coloneqq \{\, \mathrm{diag}(\hat{X}[V]) : \hat{X} \text{ feasible in (2.9)}\},\\
\mathrm{TH}^+(G) &\coloneqq \{\, \mathrm{diag}(\hat{X}[V]) : \hat{X} \text{ feasible in (2.10)}\},\\
\mathrm{QSTAB}(G) &\coloneqq \{\, x \in \mathbb{R}^V_+ : \mathbb{1}_K^{\mathsf{T}} x \leq 1\, \forall K \in \mathcal{K}(G)\}.
\end{aligned}$$

A strong form of the Lovász sandwich theorem [20] is that

$$\mathrm{STAB}(G) \subseteq \mathrm{TH}'(G) \subseteq \mathrm{TH}(G) \subseteq \mathrm{TH}^{+}(G) \subseteq \mathrm{QSTAB}(G). \tag{5.1}$$

The following result is well known; we include a sketch of its proof for completeness.

THEOREM 5.1. *Let G be a graph. The following are equivalent:*

(i) *G is perfect;*
(ii) *$\overline{G}$ is perfect;*
(iii) $\mathrm{STAB}(G) = \mathrm{QSTAB}(G)$*;*
(iv) *the system $x \geq 0$, $\mathbb{1}_K^\mathsf{T} x \leq 1$ $\forall K \in \mathcal{K}(G)$ defining* $\mathrm{QSTAB}(G)$ *is TDI;*
(v) *$\alpha(G, w) = \overline{\chi}(G, w)$ for each $w \colon V \to \mathbb{Z}$;*
(vi) $\mathrm{TH}(G)$ *is a polytope;*
(vii) $\mathrm{TH}'(G)$ *is a polytope;*
(viii) $\mathrm{TH}^{+}(G)$ *is a polytope.*

Proof. Most equivalences can be seen in [14, Chap. 9], except for (vii) and (viii), involving $\mathrm{TH}'(G)$ and $\mathrm{TH}^{+}(G)$. It is clear that (iii) and (5.1) imply both (vii) and (viii). When proving that (vi) implies (iii), [14, Cor. 9.3.27] relies on the facts that the antiblocker of $\mathrm{TH}(G)$ is $\mathrm{TH}(\overline{G})$ and that the nontrivial facets of $\mathrm{TH}(G)$ are determined by the clique inequalities $\mathbb{1}_K^\mathsf{T} x \leq 1$ for each $K \in \mathcal{K}(G)$. It is well known that the antiblocker of $\mathrm{TH}'(G)$ is $\mathrm{TH}^{+}(\overline{G})$ and that the nontrivial facets of both $\mathrm{TH}'(G)$ and $\mathrm{TH}^{+}(G)$ are determined by the same clique inequalities above; one may find complete, unified proofs in [8, Theorem 24]. These facts are sufficient to adapt the proof from [14, Cor. 9.3.27] to show that each of (vii) and (viii), separately, implies (iii). ∎

We can now characterize TDIness for ϑ via the underlying graph being perfect. In the proof below we comment on the modifications to obtain analogous results for the SDP formulations (2.8) and (2.10), of ϑ' and ϑ^{+}, respectively. Hence, total dual integrality works in a robust manner for these formulations.

THEOREM 5.2. *Let $G = (V, E)$ be a graph. The defining system for the SDP formulation of the Lovász ϑ function in* (2.9) *is TDI through $w \in \mathbb{R}^V \mapsto \mathrm{Diag}(0 \oplus w)$ if and only if G is perfect. Analogous statements hold for the SDP formulations of ϑ' and ϑ^{+} in* (2.8) *and* (2.10)*, respectively.*

Proof. We start with sufficiency. Suppose G is perfect. Let $w \colon V \to \mathbb{Z}$. Let $U \subseteq V$ be a stable set of G such that $\alpha(G, w) = w^\mathsf{T} \mathbb{1}_U$, so that $\hat{X}$ defined as in (2.1) is feasible in (2.9) with objective value $\alpha(G, w)$. Then by item (v) in Theorem 5.1 there exists a clique cover m of G with respect to w such that $\mathbb{1}^\mathsf{T} m = \alpha(G, w)$. Hence, Proposition 2.6 shows that there is an integral dual solution $(\hat{S}, \eta, u, y, z)$ for the dual SDP of (2.9) with objective value $\eta = \mathbb{1}^\mathsf{T} m = \alpha(G, w)$, which is the same as the objective value of $\hat{X}$. Hence, $(\hat{S}, \eta, u, y, z)$ is optimal for the dual SDP of (2.9) by weak duality. Note in fact that Proposition 2.6 shows that $(\hat{S}, \eta, u, y, z)$ is an integer dual solution also for the dual SDPs of (2.8) and (2.10).

Now we move to necessity. Suppose the defining system is TDI through $\mathrm{Diag}(0 \oplus \cdot)$. By Theorem 4.3, it follows that $\mathrm{TH}(G) = \mathrm{TH}(G)_I$; hence $\mathrm{TH}(G)$ is a polytope and G is perfect by Theorem 5.1. Note that the equivalences (vii) and (viii) in Theorem 5.1 also show that the defining systems for ϑ' and ϑ^{+} can only be TDI if G is perfect. ∎

By Theorem 3.5, $\mathrm{TH}(G) = \mathrm{STAB}(G)$ if and only if the gauge function of $[\mathrm{TH}(G)]^{\circ}$ is integer-valued for every integral vector $w \in \mathbb{Z}^V$. Since $\mathrm{TH}(\overline{G}) = [\mathrm{TH}(G)]^{\circ} \cap \mathbb{R}^V_{+}$, Theorem 3.5 also relates exactness (as a relaxation of the stable set polytope) of the theta body of a graph to the integrality of the gauge function of its complement.

6. Dual integrality for the Max Cut SDP. Let $G = (V, E)$ be a graph. A *cut* in G is a set of edges of the form $\delta(U) \coloneqq \{\, e \in E : |e \cap U| = 1 \}$ for some $U \subseteq V$ such that $\varnothing \neq U \neq V$. If $U = \{i\} \subseteq V$ is a singleton, write $\delta(i) \coloneqq \delta(\{i\})$. The *maximum cut problem* (or Max Cut problem) is to find, given a graph $G = (V, E)$ and $w \colon E \to \mathbb{R}_+$, an optimal solution for $\max\{\, w^{\mathsf{T}} \mathbb{1}_{\delta(U)} : \varnothing \neq U \subsetneq V \}$. (We shall discuss nonnegativity of w and related issues in Appendix B.3.) By using the embedding $U \in \mathcal{P}(V) \mapsto s_U s_U^{\mathsf{T}} \in \mathbb{S}^V$ with $s_U \coloneqq 2\mathbb{1}_U - \mathbb{1}$, i.e., $(s_U)_i = (-1)^{[i \notin U]}$ for each $i \in V$, one may reformulate the Max Cut problem exactly by adding the constraint "$\operatorname{rank}(Y) = 1$" to the SDP

$$\max\{\, \langle \tfrac{1}{4}\mathcal{L}_G(w), Y \rangle : \operatorname{diag}(Y) = \mathbb{1},\, Y \in \mathbb{S}_+^V \}; \tag{6.1}$$

here, $\mathcal{L}_G \colon \mathbb{R}^E \to \mathbb{S}^V$ is the *Laplacian* of the graph G, defined as

$$\mathcal{L}_G(w) \coloneqq \sum_{ij \in E} w_{ij}(e_i - e_j)(e_i - e_j)^{\mathsf{T}} \qquad \forall w \in \mathbb{R}^E. \tag{6.2}$$

It is not hard to check that $\mathbb{1}_U^{\mathsf{T}} \mathcal{L}_G(w) \mathbb{1}_U = \frac{1}{4} s_U^{\mathsf{T}} \mathcal{L}_G(w) s_U = w^{\mathsf{T}} \mathbb{1}_{\delta(U)}$ for each $U \subseteq V$. We call (6.1) the *Max Cut SDP*. It is one of the most famous SDPs, since it was used by Goemans and Williamson [13] in their seminal approximation algorithm and its analysis.

We postpone discussing the drawbacks of the rank-one constraint for the dual SDP of (6.1) to Appendix B.1. Here we shall study the integer dual SDP for the Max Cut SDP with objective functions of the form $X \mapsto \langle \frac{1}{4}\mathcal{L}_G(w), X \rangle$ for every $w \in \mathbb{R}_+^E$. To apply our theory to the Max Cut SDP, we formulate (6.1) in our format. First we rewrite it as $\max\{\, \langle 0 \oplus \frac{1}{4}\mathcal{L}_G(w), \hat{Y} \rangle : \operatorname{diag}(\hat{Y}) = \mathbb{1},\, \hat{Y} \in \widehat{\mathbb{S}}_+^V \}$ and then perform the change of variable

$$\hat{Y} \mapsto \hat{B}\hat{Y}\hat{B}^{\mathsf{T}} = \hat{X}, \qquad \text{where } \hat{B} \coloneqq \frac{1}{2}\begin{bmatrix} 2 & 0^{\mathsf{T}} \\ \mathbb{1} & I \end{bmatrix},$$

and add the redundant constraints $\operatorname{diag}(\hat{X}[V]) \geq 0$ to get the equivalent SDP

$$\begin{aligned} &\text{Maximize} && \langle 0 \oplus \mathcal{L}_G(w), \hat{X} \rangle \\ &\text{subject to} && \hat{X} \in \widehat{\mathbb{S}}_+^V \text{ satisfies (2.4)}, \end{aligned} \tag{6.3}$$

called the *homogeneous Max Cut SDP*. The change of variable is a linear automorphism of $\widehat{\mathbb{S}}_+^V$ that preserves rank, so we are not giving ourselves any undue advantage by choosing this embedding.

The dual SDP of (6.3) is

$$\begin{aligned} &\text{Minimize} && \eta \\ &\text{subject to} && \begin{bmatrix} \eta & -u^{\mathsf{T}} \\ -u & \operatorname{Diag}(2u - z) \end{bmatrix} - \hat{S} = \begin{bmatrix} 0 & 0^{\mathsf{T}} \\ 0 & \mathcal{L}_G(w) \end{bmatrix}, \\ &&& \hat{S} \in \widehat{\mathbb{S}}_+^V, \eta \in \mathbb{R},\, u \in \mathbb{R}^V,\, z \in \mathbb{R}_+^V. \end{aligned} \tag{6.4}$$

Upon adding the integrality constraint to (6.4), assuming integrality of $w \in \mathbb{Z}_+^E$, and simplifying, we get

$$\min\left\{ \mathbb{1}^{\mathsf{T}} m : \mathbb{1}_{\delta(i)}^{\mathsf{T}} w \leq \mathbb{1}_{\binom{V}{i \in}}^{\mathsf{T}} m \,\forall i \in V,\, \mathbb{1}_{\binom{V}{ij \subseteq}}^{\mathsf{T}} m = w_{ij} \,\forall ij \in E,\, m \colon \mathcal{K}(G) \setminus \{\varnothing\} \to \mathbb{Z}_+ \right\}. \tag{6.5}$$

The next result, whose proof we postpone to Appendix B.2, describes the unique optimal solution of (6.5).

THEOREM 6.1. *Let $G = (V, E)$ be a graph, and let $w\colon E \to \mathbb{Z}_+$. Then the optimization problem* (6.5) *has a unique optimal solution m^*, and it satisfies* $\operatorname{supp}(m^*) \subseteq E$ *and* $m^*\restriction_E = w$.

Note that Theorem 6.1 does not characterize total dual integrality of the Max Cut SDP (6.1) since it only identifies integral dual optimal solutions when the weight function w on the edges is nonnegative. We postpone the discussion of dual integrality for not necessarily nonnegative weight functions to Appendix B.3.

7. Conclusion and future directions. We have introduced a primal-dual symmetric notion of integrality in SDPs in Definitions 2.1 and 2.2; see also conditions (P$\mathbb{Z}$) and (D$\mathbb{Z}$). This enabled the SDP version in Theorem 2.4 of the LP-based Theorem 1.1. Then, by relying on our generalization of Corollary 1.3 in Corollary 3.3, and the notion of *total dual integrality* through a linear map in Definition 4.2, we described sufficient conditions for exactness of the (primal) SDP formulation in Theorem 4.3 and equality throughout the chain from Theorem 2.4 in Theorem 4.4. We also characterized the semidefinite notions of TDIness in the LP case (Theorem 4.5) and the theta function and its two variants (Theorem 5.2) via natural conditions. Finally, in Theorem 6.1, we determined the optimal solutions for the integer dual SDP for the Max Cut SDP when the weight function on the edges of the graph is nonnegative.

Our approach leads to several other interesting research directions. We start with the following problem.

PROBLEM 7.1. *Obtain a primal-dual symmetric integrality condition for SDPs that applies to arbitrary ILPs, not just binary ones.*

The theory of total dual integrality for LPs is considered well understood. Our work raises new issues, related to the interplay between total dual integrality and extended formulations in LP; the latter area has received a lot of attention recently. More concretely, one may define a system of linear inequalities $Ax \leq b$ on $\mathbb{R}^n$ to be TDI through a linear map $L\colon \mathbb{R}^k \to \mathbb{R}^n$ if, for every integral $c \in \mathbb{Z}^k$, the LP dual to $\sup\{\, \langle L(c), x\rangle : Ax \leq b\}$ has an integral optimal solution if its optimal value is finite.

PROBLEM 7.2. *Are there compact extended formulations for classical combinatorial optimization problems (e.g., maximum weight r-arborescences, minimum spanning trees) that are TDI through the corresponding lifting maps? Do these lead to new min-max theorems?*

PROBLEM 7.3. *Let $Ax \leq b$ be a system of linear inequalities on $\mathbb{R}^n$ and $L\colon \mathbb{R}^k \to \mathbb{R}^n$ be a linear map such that for $P := \{\, x \in \mathbb{R}^n : Ax \leq b\}$ the projection $L^*(P)$ is integral. Does there exist a TDI system $Cx \leq d$ in $\mathbb{R}^n$ with d integral such that $L^*(P) = L^*(\{\, x \in \mathbb{R}^n : Cx \leq d\})$?*

Recall that a polyhedron $P \subseteq \mathbb{R}^n$ has the *integer decomposition property* if for every $k \in \mathbb{Z}_+$, each integral point in $kP = \{\, kx : x \in P\}$ is a sum of k integral points in P. The next problem is more open-ended.

PROBLEM 7.4. *What is the relation between total dual integrality and the integer decomposition property (see* [30, sec 22.10]*), of which our dual integrality condition in Definition* 2.1 *is reminiscent?*

In section 6 we studied dual integrality of Max Cut SDP with nonnegative weight functions, and we discuss in Appendix B.3 the issues that arise when we allow weights of arbitrary signs. These issues suggest further research directions. One may define a refinement of the notion of total dual integrality restricted to a rational polyhedral cone $\mathbb{K} \subseteq \mathbb{R}^k$; there, one would only require the dual SDP to have an integral optimal solution for primal objective functions of the form $\hat{X} \mapsto \langle \mathcal{L}(c), \hat{X} \rangle$ with integral $c \in \mathbb{K}$. In this context, it seems misleading to use the term *total* dual integrality; $\mathbb{K}$-dual integrality seems more adequate.

PROBLEM 7.5. *Adapt Theorem* 4.3 *to some notion of* $\mathbb{K}$*-dual integrality. Similarly, can Corollary* 3.3 *be modified by taking only integral* $w \in \mathbb{K}$ *to obtain integrality of some modification of* $\mathscr{C}$ *using* $\mathbb{K}$*?*

Concerning the semidefinite notion of TDIness, one may ask for a characterization of total dual integrality of other SDP formulations, such as the application of lift-and-project hierarchies (see [19]) to ILP formulations of combinatorial optimization problems. One possible instance is the following.

PROBLEM 7.6. *Given* $k \geq 1$ *and the* LS_+ *operator of Lovász and Schrijver* [21] *(called* N_+ *in their paper), determine the class of graphs for which the* k*th iterate of the* LS_+ *operator applied to the system*

$$x \geq 0, \qquad x_i + x_j \leq 1 \quad \forall ij \in E \tag{7.1}$$

yields a TDI system through the appropriate lifting, leading to a minmax relation involving stable sets in such graphs.

Still in the realm of SDPs, one may ask for notions of *exactness* other than integrality, as well as their dual counterparts. For instance, many problems in continuous mathematics, such as control theory, lead to nonconvex optimization problems where the variable matrix is required to be rank-one or of restricted rank. However, the entries of such a matrix may define a continuous curve rather than taking on only finitely many values. For a general convex relaxation framework working with such formulations, see [17].

PROBLEM 7.7. *Obtain systematic, primal-dual symmetric conditions for exactness in SDP relaxations for continuous problems.*

Finally, one may consider the problem of defining integrality in a systematic and primal-dual symmetric way for convex optimization problems in other forms. This is especially challenging since a dual integrality notion, even in the polyhedral case, is inherently dependent on the algebraic representation of the problem, not only on its geometry.

Appendix A. Rank constraint in dual SDP of trace formulation for theta. In subsection 2.1 we showed that the rank-one constraint for the dual SDP of a formulation of the theta function is not very interesting. There, the formulation we used was based on our chosen embedding into the lifted space $\widehat{\mathbb{S}}^V$, which requires the constraints (2.4). One might argue that the rank-one constraint might make more sense for the dual SDP of the probably more popular formulation of $\vartheta(G, w)$ for a graph $G = (V, E)$ and $w\colon V \to \mathbb{R}_+$ obtained by dropping the nonnegativity constraint from

$$\vartheta'(G, w) = \max\Big\{ \langle \sqrt{w}\sqrt{w}^{\mathsf{T}}, X \rangle : \mathrm{Tr}(X) = 1,\ X_{ij} = 0\, \forall ij \in E,\ X_{ij} \geq 0\, \forall ij \in \overline{E},\ X \in \mathbb{S}^V_+ \Big\}. \tag{A.1}$$

Here, $\sqrt{w} \in \mathbb{R}^V_+$ denotes the componentwise square root of w. We will show that the rank-one constraint is not very meaningful even in the dual of the SDP in (A.1);

as in subsection 2.1, this dual feasible region contains those of the corresponding formulations for ϑ and ϑ^+. Denote $\mathscr{A}_G \coloneqq \{\, A \in \mathbb{S}^V : A_{ij} \neq 0 \implies ij \in E\}$. The dual SDP of (A.1) can be written as

$$\text{(A.2)} \qquad \min\{\, \lambda : \lambda I + A - \overline{A} - S = \sqrt{w}\sqrt{w}^{\mathsf{T}},\, S \succeq 0,\, A \in \mathscr{A}_G, \overline{A} \in \mathscr{A}_{\overline{G}} \cap \mathbb{S}^V_{\geq 0}\}.$$

The embedding of stable sets in G as feasible solutions of (A.1) goes as follows: if $U \subseteq V$ is a stable set in G with positive weight $w^{\mathsf{T}}\mathbb{1}_U$, then $X \coloneqq (w^{\mathsf{T}}\mathbb{1}_U)^{-1}(\sqrt{w} \odot \mathbb{1}_U)(\sqrt{w} \odot \mathbb{1}_U)^{\mathsf{T}}$ is feasible in (A.1), with objective value $w^{\mathsf{T}}\mathbb{1}_U$. The normalization factor and the square root in the definition of X already hint that this formulation does not play so well with integrality.

Proposition A.1. *Let $G = (V, E)$ be a graph, and let $w \in \mathbb{R}^V_{++}$. If there exists a feasible solution $(\lambda, A, \overline{A}, S)$ for* (A.2) *such that* $\operatorname{rank}(S) \leq 1$*, then $\overline{G}$ is bipartite.*

Proof. Suppose $S = ss^{\mathsf{T}}$ for some $s \in \mathbb{R}^V$. Then

$$\text{(A.3)} \qquad \lambda I + A = ss^{\mathsf{T}} + \sqrt{w}\sqrt{w}^{\mathsf{T}} + \overline{A}.$$

Apply diag to both sides of (A.3) to get $\lambda\mathbb{1} = (s \odot s) + w$. Hence, $\lambda\mathbb{1} \geq w$ and there exists $U \subseteq V$ such that $s = \operatorname{Diag}(2\mathbb{1}_U - \mathbb{1})\sqrt{\lambda\mathbb{1} - w}$. Let $ij \in \overline{E}$. Specialize (A.3) to the ijth entry to get

$$\text{(A.4)} \quad 0 = s_i s_j + \sqrt{w_i w_j} + \overline{A}_{ij} \geq (-1)^{[i \notin U] + [j \notin U]}(\lambda - w_i)^{1/2}(\lambda - w_j)^{1/2} + \sqrt{w_i w_j}.$$

If $i, j \in U$ or $i, j \in \overline{U} \coloneqq V \setminus U$, then the RHS of (A.4) is positive, since $w \in \mathbb{R}^V_{++}$. This contradiction shows that $G[U] = K_U$ and $G[\overline{U}] = K_{\overline{U}}$, so $\overline{G}$ is bipartite with color classes U and $\overline{U}$. ∎

By our previous discussion, the dual SDPs of the above formulations of ϑ, ϑ', and ϑ^+ only have rank-one slacks when $\overline{G}$ is bipartite (whence G is perfect).

We note, however, that another low-rank constraint for the dual SDP for ϑ' does in fact yield a useful and almost exact formulation for the chromatic number of a graph $G = (V, E)$, via the circular chromatic number. Suppose G has at least one edge. We first describe the vector chromatic number $\chi_v(G)$, introduced in [16]:

$$\text{(A.5)} \qquad \chi_v(G) \coloneqq \min\{\, \tau : \operatorname{diag}(Y) = \mathbb{1},\, Y_{ij} \leq -\tfrac{1}{\tau - 1}\, \forall ij \in E,\, Y \in \mathbb{S}^V_+,\, \tau \geq 2\}.$$

The map $(S, \lambda) \mapsto \frac{1}{\lambda - 1}S$ maps bijectively the feasible region of (A.2) applied to $\overline{G}$ to the feasible region of (A.5) and preserves objective values. Hence, $\chi_v(G) = \vartheta'(\overline{G})$. Any optimal solution σ^* for the SDP

$$\text{(A.6)} \qquad \min\{\, \sigma : \operatorname{diag}(Y) = \mathbb{1},\, Y_{ij} \leq \sigma\, \forall ij \in E,\, Y \in \mathbb{S}^V_+,\, \tau \geq 2\}$$

lies in $[-1, 0)$ and leads to the optimal value $\tau^* \coloneqq 1 - 1/\sigma^*$ for (A.5).

Consider next the *circular chromatic number* of G, which can be defined as (see [10])

$$\text{(A.7)} \qquad \chi_c(G) \coloneqq \min\{\, \tau : y \colon V \to S^1,\, \phi_{ij} \geq 2\pi/\tau\, \forall ij \in E,\, \tau \geq 2\},$$

where S^1 denotes the unit sphere in $\mathbb{R}^2$ and $\phi_{ij} \in [0, \pi]$ is the angle between y_i and y_j. Since cos is monotone decreasing on $[0, \pi]$, we can rewrite (A.7) using Gram matrices as

$$\text{(A.8)} \quad \min\{\, \tau : \operatorname{diag}(Y) = \mathbb{1},\, Y_{ij} \leq \cos(2\pi/\tau)\, \forall ij \in E,\, Y \in \mathbb{S}^V_+,\, \operatorname{rank}(Y) = 2,\, \tau \geq 2\}.$$

Finally, since $f\colon \tau \in [2,\infty) \mapsto \cos\frac{2\pi}{\tau} \in [-1,1)$ is a monotone increasing bijection, we see that, if σ^* is the optimal value of (A.6) with the extra constraint $\operatorname{rank}(Y) = 2$, then $\chi_c(G) = f^{-1}(\sigma^*)$. One can then read off the chromatic number of G since $\chi(G) = \lceil \chi_c(G) \rceil$; see [33].

Note, however, that this dual formulation required quite a lot of ad hoc treatment.

Appendix B. Further analysis of the Max Cut integer dual SDP.

B.1. Rank-one constraint in dual of the Max Cut SDP. In this section, we show that the dual of the Max Cut SDP has a feasible solution with a rank-one slack only if the weight function on the edges comes from a very restricted (though rather interesting) class of weight functions. Let $G = (V,E)$ be a graph, and let $w\colon E \to \mathbb{R}$. The dual of the Max Cut SDP (6.1) is

$$\min\{\, \mathbb{1}^\mathsf{T} y : S = \operatorname{Diag}(y) - \tfrac{1}{4}\mathcal{L}_G(w),\ S \in \mathbb{S}^V_+,\ y \in \mathbb{R}^V \,\}. \tag{B.1}$$

Proposition B.1. *Let $G = (V,E)$ be a graph without isolated vertices. Let $w\colon E \to \mathbb{R}\setminus\{0\}$. If* (B.1) *has a feasible solution (S,y) such that* $\operatorname{rank}(S) \le 1$, *then $G = K_V$, and there exists $u\colon V \to \mathbb{R}\setminus\{0\}$ such that $w_{ij} = u_i u_j$ for each $ij \in E$.*

Proof. Set $L := \mathcal{L}_G(w)$. Suppose there exists $u \in \mathbb{R}^V$ such that $S = uu^\mathsf{T}$. Then, for each $i \in V$, we have $y_i - \frac{1}{4}L_{ii} = S_{ii} = u_i^2 \ge 0$. If $u_i = 0$ for some $i \in V$, then $Se_i = 0$ whence i is an isolated vertex; recall that $S = \operatorname{Diag}(y) - L$. Hence, $\operatorname{supp}(u) = V$. Now the off-diagonal entries of the equality constraint of (B.1) show that $G = K_V$ and that $w_{ij} = 4u_iu_j$ for each $ij \in E = \binom{V}{2}$. ∎

Instances of Max Cut of the form described by Proposition B.1 are still NP-hard. Indeed, they may be reformulated as $\max\{\, (\mathbb{1}_U^\mathsf{T} u)(\mathbb{1}_{V\setminus U}^\mathsf{T} u) : \varnothing \ne U \subsetneq V \,\}$. The latter can be seen to include the partition problem.

B.2. Optimal solution for Max Cut integer dual SDP. Theorem 6.1 follows immediately from the following slightly more general result.

Theorem B.2. *Let $G = (V,E)$ be a graph, and let $w\colon E \to \mathbb{Z}_+$. Then the optimization problem*

$$\begin{aligned}
&\text{(B.2a)} && \textit{Minimize} && \mathbb{1}^\mathsf{T} m \\
&\text{(B.2b)} && \textit{subject to} && m\colon \mathcal{P}(V)\setminus\{\varnothing\} \to \mathbb{Z}_+, \\
&\text{(B.2c)} && && \operatorname{supp}(m) \subseteq \mathcal{K}(G), \\
&\text{(B.2d)} && && \mathbb{1}_{\delta(i)}^\mathsf{T} w \le \mathbb{1}_{\binom{V}{i\in}}^\mathsf{T} m && \forall i \in V, \\
&\text{(B.2e)} && && \mathbb{1}_{\binom{V}{ij\subseteq}}^\mathsf{T} m \le w_{ij} && \forall ij \in E
\end{aligned}$$

has a unique optimal solution m^*, and it satisfies $\operatorname{supp}(m^*) \subseteq E$ *and* $m^*\restriction_E = w$.

Proof. Let $m_w\colon \mathcal{P}(V) \to \mathbb{Z}_+$ such that $\operatorname{supp}(m_w) \subseteq E$ and $m_w\restriction_E = w$. It is easy to check that m_w is feasible in (B.2). Let $m^*\colon \mathcal{P}(V) \to \mathbb{Z}_+$ be an optimal solution for (B.2); one exists since there exist feasible solutions and the objective value of every feasible solution is a nonnegative integer. We will prove that $m^* = m_w$.

The key part of the proof is to show that

$$\operatorname{supp}(m^*) \subseteq \tbinom{V}{1} \cup \tbinom{V}{2}. \tag{B.3}$$

Let $C \in \operatorname{supp}(m^*)$. We claim that

$$\tilde{m} := m^* - d + \mathbb{1}_{E[C]} \text{ is feasible for (B.2), where } d := e_C + \big(|C|-2\big)\mathbb{1}_{\binom{C}{1}}. \tag{B.4}$$

For each $i \in V$, we have $\mathbb{1}^{\mathsf{T}}_{\binom{V}{i\in}} d = [i \in C]\big(|C|-1\big) = \mathbb{1}^{\mathsf{T}}_{\binom{V}{i\in}} \mathbb{1}_{E[C]}$, so (B.2d) holds for $\tilde{m}$. For every $ij \in E$ we have $\mathbb{1}^{\mathsf{T}}_{\binom{V}{ij\subseteq}} d = \big[ij \in E[C]\big] = \mathbb{1}^{\mathsf{T}}_{\binom{V}{ij\subseteq}} \mathbb{1}_{E[C]}$, so (B.2e) holds for $\tilde{m}$. In verifying (B.2b) for $\tilde{m}$, we may assume $|C| \geq 2$. We will prove that (B.2b) holds for $\tilde{m}$ by showing that

$$\bar{m} := m^* - e_C \geq \big(|C| - 2\big)\mathbb{1}_{\binom{C}{1}}; \tag{B.5}$$

then (B.2c) for $\tilde{m}$ will also follow, thus completing the proof of (B.4).

Note that $\bar{m} \geq 0$. Let $i \in V$, and let $N(i)$ denote the set of its neighbors. Then

$$\begin{aligned}
\mathbb{1}^{\mathsf{T}}_{\delta(i)} w &\leq \mathbb{1}^{\mathsf{T}}_{\binom{V}{i\in}} m^* && \text{by (B.2d)}\\
&= \mathbb{1}^{\mathsf{T}}_{\binom{V}{i\in}} \bar{m} + [i \in C] && \text{since } m^* = \bar{m} + e_C\\
&\leq \bar{m}_{\{i\}} + \sum_{j \in V \setminus \{i\}} \mathbb{1}^{\mathsf{T}}_{\binom{V}{ij\subseteq}} \bar{m} + [i \in C] && \text{since } \mathbb{1}_{\binom{V}{i\in}} \leq e_{\{i\}} + \sum_{j \in V\setminus\{i\}} \mathbb{1}_{\binom{V}{ij\subseteq}}\\
&= \bar{m}_{\{i\}} + \sum_{j \in N(i)} \mathbb{1}^{\mathsf{T}}_{\binom{V}{ij\subseteq}} \bar{m} + [i \in C] && \text{by (B.2c)}\\
&= \bar{m}_{\{i\}} + \sum_{j \in N(i)} \mathbb{1}^{\mathsf{T}}_{\binom{V}{ij\subseteq}} m^* - \sum_{j \in N(i)} \mathbb{1}^{\mathsf{T}}_{\binom{V}{ij\subseteq}} e_C + [i \in C] && \text{since } \bar{m} = m^* - e_C\\
&\leq \bar{m}_{\{i\}} + \sum_{j \in N(i)} w_{ij} - |\delta(i) \cap E[C]| + [i \in C] && \text{by (B.2e)}\\
&= \bar{m}_{\{i\}} + \mathbb{1}^{\mathsf{T}}_{\delta(i)} w - [i \in C]\big(|C| - 2\big) && \text{since } |\delta(i) \cap E[C]| = [i \in C]\big(|C| - 1\big).
\end{aligned}$$

This proves (B.5) and thus completes the proof of (B.4).

We have $\mathbb{1}^{\mathsf{T}} m^* - \mathbb{1}^{\mathsf{T}} \tilde{m} = \mathbb{1}^{\mathsf{T}} d - \mathbb{1}^{\mathsf{T}} \mathbb{1}_{E[C]} = 1 + |C|\big(|C| - 2\big) - \binom{|C|}{2} = \frac{1}{2}\big(|C| - 1\big)\big(|C| - 2\big)$. Optimality of m^* and (B.4) imply that $|C| \in \{1, 2\}$. This concludes the proof of (B.3).

By summing the vertex constraints (B.2d) and using (B.3), we obtain

$$2\mathbb{1}^{\mathsf{T}} w \leq \Bigg(\sum_{A \subseteq V} |A| e_A\Bigg)^{\mathsf{T}} m^* = \mathbb{1}^{\mathsf{T}}_{\binom{V}{1}} m^* + 2\mathbb{1}^{\mathsf{T}}_{\binom{V}{2}} m^*. \tag{B.6}$$

By summing the edge constraints (B.2e) and using (B.3), we obtain

$$\mathbb{1}^{\mathsf{T}}_{\binom{V}{2}} m^* = \Bigg(\sum_{A \subseteq V} \binom{|A|}{2} e_A\Bigg)^{\mathsf{T}} m^* \leq \mathbb{1}^{\mathsf{T}} w. \tag{B.7}$$

It follows from (B.3), (B.6), and (B.7) that

$$\mathbb{1}^{\mathsf{T}} m_w = \mathbb{1}^{\mathsf{T}} w \leq \mathbb{1}^{\mathsf{T}}_{\binom{V}{1}} m^* + \mathbb{1}^{\mathsf{T}}_{\binom{V}{2}} m^* = \mathbb{1}^{\mathsf{T}} m^*. \tag{B.8}$$

Equality throughout in (B.8) implies that each constraint in (B.2d) and (B.2e) holds with equality for m^*, so that m^* is feasible for (6.5). The latter fact, together with (B.3), easily implies that $m^* = m_w$. ∎

B.3. The Max Cut problem and nonnegative weights. One may wonder whether Theorem 6.1 may be extended to arbitrary weight functions $w \colon E \to \mathbb{Z}$, not just nonnegative weights. Such an extension might be used to characterize the graphs G for which the system defining the Max Cut SDP (6.1) is TDI through $w \in$

$\mathbb{R}^E \mapsto 0 \oplus \mathcal{L}_G(w) =: \mathcal{L}(w)$; by Theorem 6.1 such graphs form a subset of the bipartite graphs. Then we would be able to obtain the *cut polytope* $\operatorname{conv}\{ \mathbb{1}_{\delta(S)} : \varnothing \neq S \subsetneq V\}$ of any such graph G as a projection of the feasible region of (6.1) via $\mathcal{L}^*$. However, due to constraints (B.2e), if $w\colon E \to \mathbb{Z}$ has a negative entry, problem (B.2) is infeasible. One may attempt to "fix" this issue by adding to (6.3) the redundant constraint $\mathcal{L}^*(\hat{X}) = \mathcal{L}_G^*(\hat{X}[V]) \geq 0$. Note that this is similar to the redundant constraint (2.4c) added in our chosen embedding, which is fundamental for dealing with $w \in \mathbb{R}^V \backslash \mathbb{R}^V_+$ for the ϑ function; in both cases, the redundant constraint comes from the projection $\mathcal{L}^*$. The dual SDP is then obtained from (6.4) by replacing the occurrence of $\mathcal{L}_G(w)$ in the RHS with $\mathcal{L}_G(w + y)$, where $y \in \mathbb{R}^E_+$ is a new variable. Optimal solutions for the corresponding integer dual SDP are described by the next result.

Corollary B.3. *Let $G = (V, E)$ be a graph, and let $w\colon E \to \mathbb{Z}$. Then the optimization problem*

$$\text{(B.9a)} \qquad \text{Minimize} \quad \mathbb{1}^\mathsf{T} m$$

$$\text{(B.9b)} \qquad \text{subject to} \quad m\colon \mathcal{K}(G) \setminus \{\varnothing\} \to \mathbb{Z}_+,\ y \in \mathbb{R}^E_+,$$

$$\text{(B.9c)} \qquad \mathbb{1}^\mathsf{T}_{\delta(i)}(w + y) \leq \mathbb{1}^\mathsf{T}_{\binom{V}{i\in}} m \qquad \forall i \in V,$$

$$\text{(B.9d)} \qquad \mathbb{1}^\mathsf{T}_{\binom{V}{ij\subseteq}} m \leq w_{ij} + y_{ij} \qquad \forall ij \in E$$

has a unique optimal solution (m^*, y^*), and it satisfies $\operatorname{supp}(m^*) \subseteq E$*, and for each $e \in E$,*

$$m^*_e = [w_e \geq 0] w_e, \qquad y^*_e = -[w_e < 0] w_e.$$

Proof. Let $(\bar{m}, \bar{y})$ be feasible. By (B.9d), we have $w + \bar{y} \geq 0$ so $y \geq y^*$. By Theorem 6.1, the optimization problem (B.9) with the extra constraint $y = \bar{y}$ has a unique optimal solution, and its optimal value is $\mathbb{1}^\mathsf{T}(w + \bar{y})$, which is greater than or equal to $\mathbb{1}^\mathsf{T}(w + y^*)$, the objective value of the feasible solution (m^*, y^*). ∎

Even though Corollary B.3 shows how the dual SDP for Max Cut with an extra (redundant) constraint may have integral solutions, the optimal value is always non-negative. The deeper problem here is that the Max Cut SDP (6.1) is not tight for arbitrary weights w, even if the underlying graph is bipartite. Hence, if $\mathscr{C} \subseteq \mathbb{R}^E$ is the projection of the feasible region of (6.1) via $\mathcal{L}^*$, we cannot even expect $\mathscr{C} = \mathscr{C}_I$, let alone total dual integrality of the defining system.

To see this, first note that, for a graph $G = (V, E)$ and weights $w\colon E \to \mathbb{R}$, we should redefine the maximum cut problem as the optimization problem $\sup\{ w^\mathsf{T} \mathbb{1}_{\delta(U)} : \varnothing \neq U \subsetneq V\}$; when $w \geq 0$, since $\delta(\varnothing) = \delta(V) = \varnothing$, it was harmless to keep both trivial sets $U = \varnothing$ and $U = V$ in the feasible set. Correspondingly, in the Max Cut SDP (6.1), the feasible solution $X := \mathbb{1}\mathbb{1}^\mathsf{T}$ shows that the optimal value is always nonnegative, *even when w is negative and G is connected*! To prevent these trivial solutions from being feasible in a modified Max Cut SDP, one may add the constraint $\langle \mathbb{1}\mathbb{1}^\mathsf{T}, X\rangle \leq (|V| - 2)^2$, since $\max\{ \langle \mathbb{1}\mathbb{1}^\mathsf{T}, s_U s_U^\mathsf{T}\rangle : \varnothing \neq U \subsetneq V\} = (|V| - 2)^2$, where $s_U := 2\mathbb{1}_U - \mathbb{1}$ for each $U \subseteq V$. These considerations lead us to strengthen (6.1) as

$$\text{(B.10)} \qquad \max\{ \langle \tfrac{1}{4}\mathcal{L}_G(w), Y\rangle : \operatorname{diag}(Y) = \mathbb{1},\ \langle \mathbb{1}\mathbb{1}^\mathsf{T}, Y\rangle \leq (|V| - 2)^2,\ Y \in \mathbb{S}^V_+\}.$$

Even this strengthened formulation is not exact for connected bipartite graphs if we allow weights of arbitrary signs. Consider, for instance, the path of length 3 given by $G = ([4], \{12, 23, 34\})$, with weights $w = -\mathbb{1}$. Then Max Cut is really a minimum cut problem and the optimal value is clearly -1. However, the feasible solution xx^{T} in (B.10) where $x := \begin{bmatrix} 1 & 1 & -2^{-1/2} & -2^{-1/2} \end{bmatrix}^{\mathsf{T}}$ has objective value $-3/4$.

These issues motivate the study of dual integrality for weight functions in a cone, as described in Problem 7.5.

REFERENCES

[1] I. M. Bomze, F. Frommlet, and M. Locatelli, *Gap, cosum and product properties of the ϑ' bound on the clique number*, Optimization, 59 (2010), pp. 1041–1051, https://doi.org/10.1080/02331930903395634.

[2] G. Braun and S. Pokutta, *A short proof for the polyhedrality of the Chvátal–Gomory closure of a compact convex set*, Oper. Res. Lett., 42 (2014), pp. 307–310, https://doi.org/10.1016/j.orl.2014.05.004.

[3] S. Burer, *On the copositive representation of binary and continuous nonconvex quadratic programs*, Math. Program., 120 (2009), pp. 479–495, https://doi.org/10.1007/s10107-008-0223-z.

[4] P. C. P. Carvalho and L. E. Trotter, Jr., *An abstract linear duality model*, Math. Oper. Res., 14 (1989), pp. 639–663, https://doi.org/10.1287/moor.14.4.639.

[5] D. Dadush, S. S. Dey, and J. P. Vielma, *The split closure of a strictly convex body*, Oper. Res. Lett., 39 (2011), pp. 121–126, https://doi.org/10.1016/j.orl.2011.02.002.

[6] D. Dadush, S. S. Dey, and J. P. Vielma, *On the Chvátal–Gomory closure of a compact convex set*, Math. Program., 145 (2014), pp. 327–348, https://doi.org/10.1007/s10107-013-0649-9.

[7] M. K. de Carli Silva, *Geometric Ramifications of the Lovász Theta Function and Their Interplay with Duality*, Ph.D. thesis, Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Canada, 2013, http://hdl.handle.net/10012/7812 (accessed 2019-10-30).

[8] M. K. de Carli Silva and L. Tunçel, *An axiomatic duality framework for the theta body and related convex corners*, Math. Program., 162 (2017), pp. 283–323, https://doi.org/10.1007/s10107-016-1041-3.

[9] M. K. de Carli Silva and L. Tunçel, *Pointed Closed Convex Sets Are the Intersection of All Rational Supporting Closed Halfspaces*, preprint, https://arxiv.org/abs/1802.03296, 2018.

[10] M. DeVos, *Some Chromatic Numbers*, http://www.sfu.ca/~mdevos/notes/semidef/chrom.pdf (accessed 2019-10-30).

[11] J. Edmonds and R. Giles, *A min-max relation for submodular functions on graphs*, in Studies in Integer Programming (Proceedings of a Workshop on Integer Programming, Bonn, 1975), P. L. Hammer, E. L. Johnson, B. H. Korte, and G. L. Nemhauser, eds., Ann. Discrete Math. 1, North-Holland, Amsterdam, 1977, pp. 185–204, https://doi.org/10.1016/S0167-5060(08)70734-9.

[12] F. R. Giles and W. R. Pulleyblank, *Total dual integrality and integer polyhedra*, Linear Algebra Appl., 25 (1979), pp. 191–196, https://doi.org/10.1016/0024-3795(79)90018-1.

[13] M. X. Goemans and D. P. Williamson, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, J. ACM, 42 (1995), pp. 1115–1145, https://doi.org/10.1145/227683.227684.

[14] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, 2nd ed., Algorithms and Combinatorics 2, Springer-Verlag, Berlin, 1993.

[15] A. J. Hoffman, *A generalization of max flow–min cut*, Math. Program., 6 (1974), pp. 352–359, https://doi.org/10.1007/BF01580250.

[16] D. Karger, R. Motwani, and M. Sudan, *Approximate graph coloring by semidefinite programming*, J. ACM, 45 (1998), pp. 246–265, https://doi.org/10.1145/274787.274791.

[17] M. Kojima and L. Tunçel, *Cones of matrices and successive convex relaxations of nonconvex sets*, SIAM J. Optim., 10 (2000), pp. 750–778, https://doi.org/10.1137/S1052623498336450.

[18] J. B. Lasserre, *An explicit exact SDP relaxation for nonlinear* 0–1 *programs*, in Integer Programming and Combinatorial Optimization (Utrecht, 2001), Lecture Notes in Comput. Sci. 2081, Springer, Berlin, 2001, pp. 293–303, https://doi.org/10.1007/3-540-45535-3_23.

[19] M. Laurent, *A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for* 0–1 *programming*, Math. Oper. Res., 28 (2003), pp. 470–496, https://doi.org/10.1287/moor.28.3.470.16391.

[20] L. Lovász, *On the Shannon capacity of a graph*, IEEE Trans. Inform. Theory, 25 (1979),

pp. 1–7, https://doi.org/10.1109/TIT.1979.1055985.
[21] L. Lovász and A. Schrijver, *Cones of matrices and set-functions and* 0–1 *optimization*, SIAM J. Optim., 1 (1991), pp. 166–190, https://doi.org/10.1137/0801013.
[22] R. J. McEliece, E. R. Rodemich, and H. C. Rumsey, Jr., *The Lovász bound and some generalizations*, J. Combin. Inform. System Sci., 3 (1978), pp. 134–152.
[23] R. R. Meyer, *On the existence of optimal solutions to integer and mixed-integer programming problems*, Math. Program., 7 (1974), pp. 223–235, https://doi.org/10.1007/BF01585518.
[24] D. Miklós, *On product of association schemes and Shannon capacity*, Discrete Math., 150 (1996), pp. 441–447, https://doi.org/10.1016/0012-365X(95)00213-G.
[25] Y. Nesterov, *Semidefinite relaxation and nonconvex quadratic optimization*, Optim. Methods Softw., 9 (1998), pp. 141–160, https://doi.org/10.1080/10556789808805690.
[26] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1997.
[27] J. Ryan and L. E. Trotter, Jr., *Weyl-Minkowski duality for integral monoids*, Optimization, 28 (1994), pp. 315–331, https://doi.org/10.1080/02331939408843925.
[28] A. Schrijver, *A comparison of the Delsarte and Lovász bounds*, IEEE Trans. Inform. Theory, 25 (1979), pp. 425–429, https://doi.org/10.1109/TIT.1979.1056072.
[29] A. Schrijver, *On cutting planes*, Ann. Discrete Math., 9 (1980), pp. 291–296, https://doi.org/10.1016/S0167-5060(08)70085-2.
[30] A. Schrijver, *Theory of Linear and Integer Programming*, Wiley-Interscience Series in Discrete Mathematics, John Wiley & Sons, Chichester, UK, 1986.
[31] A. Schrijver, *Combinatorial Optimization*, Algorithms and Combinatorics 24, Springer-Verlag, Berlin, 2003.
[32] M. Szegedy, *A note on the theta number of Lovász and the generalized Delsarte bound*, in Proceedings of the 35th Annual IEEE Symposium on Foundations of Computer Science, IEEE, Washington, DC, 1994, https://doi.org/10.1109/SFCS.1994.365707.
[33] X. Zhu, *Circular chromatic number: A survey*, Discrete Math., 229 (2001), pp. 371–410, https://doi.org/10.1016/S0012-365X(00)00217-X.