Department of Policy and Planning Sciences

Discussion Paper Series

No.1376

Evaluating approximations of the semidefinite cone with trace normalized distance

by

Yuzhu WANG and Akiko YOSHISE

May 2021

Revised July 2022

UNIVERSITY OF TSUKUBA
Tsukuba, Ibaraki 305-8573
JAPAN

Evaluating approximations of the semidefinite cone with trace normalized distance *

Yuzhu Wang† and Akiko Yoshise‡

May 2021
Revised July 2022

Abstract

We evaluate the dual cone of the set of diagonally dominant matrices (resp., scaled diagonally dominant matrices), namely $\mathcal{DD}_n^*$ (resp., $\mathcal{SDD}_n^*$), as an approximation of the semidefinite cone. We prove that the norm normalized distance, proposed by Blekherman et al. (2022), between a set $\mathcal{S}$ and the semidefinite cone has the same value whenever $\mathcal{SDD}_n^* \subseteq \mathcal{S} \subseteq \mathcal{DD}_n^*$. This implies that the norm normalized distance is not a sufficient measure to evaluate these approximations. As a new measure to compensate for the weakness of that distance, we propose a new distance, called the trace normalized distance. We prove that the trace normalized distance between $\mathcal{DD}_n^*$ and $\mathcal{S}_+^n$ has a different value from the one between $\mathcal{SDD}_n^*$ and $\mathcal{S}_+^n$ and give the exact values of these distances.

Key words: Semidefinite optimization problem; Diagonally dominant matrix; Scaled diagonally dominant matrix.

1 Introduction

Semidefinite programming (SDP) can provide powerful convex relaxations for combinatorial and non-convex optimization problems ([10], [13], [17], [19], etc.). SDP has the following standard form:

$$\begin{aligned} &\min\ \langle C, X\rangle \\ &\text{s.t. } \langle A_j, X\rangle = b_j, j = 1, 2, \ldots, m, \\ &\quad\ X \in \mathcal{S}_+^n, \end{aligned}$$

where $C \in \mathbb{S}^n$, $A_j \in \mathbb{S}^n$, $b_j \in \mathbb{R}$ $(j = 1, 2, \ldots, m)$, and $\langle A, B\rangle := \sum_{i,j=1}^n A_{i,j}B_{i,j}$ is the inner product over $\mathbb{S}^n$. The space of symmetric matrices is denoted as $\mathbb{S}^n := \{X \in \mathbb{R}^{n\times n} \mid X_{i,j} = X_{j,i}\ (1 \le i < j \le n)\}$ and the semidefinite cone is defined as $\mathcal{S}_+^n := \{X \in \mathbb{S}^n \mid d^T X d \ge 0 \text{ for any } d \in \mathbb{R}^n\}$.

SDP is theoretically attractive because it can be solved in polynomial time to any desired precision. However, it is difficult to solve large-scale SDP instances even using state-of-the-art solvers, such as

*This research was supported by the Japan Society for the Promotion of Science through a Grant-in-Aid for Scientific Research ((B)19H02373) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

†Graduate School of Systems and Information Engineering, University of Tsukuba, Tsukuba, Ibaraki 305-8573, Japan. email: wangyuzhu820@yahoo.co.jp

‡Corresponding author. Faculty of Engineering, Information and Systems, University of Tsukuba, Tsukuba, Ibaraki 305-8573, Japan. email: yoshise@sk.tsukuba.ac.jp

Mosek [14]. One technique to overcome this deficiency is to relax the semidefinite constraint and solve the resulting easier problem by using, e.g., linear programming (LP) or second order cone programming (SOCP). A typical relaxation technique is to replace the constraint $X \in \mathcal{S}^n_+$ with a relaxed constraint $X \in \mathcal{S}$, where $\mathcal{S}$ is a subset of $\mathbb{S}^n$ containing $\mathcal{S}^n_+$, i.e., $\mathcal{S}^n_+ \subseteq \mathcal{S} \subseteq \mathbb{S}^n$. If $X \in \mathcal{S}$ is described by linear constraints, the resulting problem becomes an LP problem and if $X \in \mathcal{S}$ is described by second-order constraints, the resulting problem becomes an SOCP problem. Such a set $\mathcal{S}$ is called an outer approximation of $\mathcal{S}^n_+$. There are two kinds of approximation, i.e., inner approximation and outer approximation. An inner approximation (outer approximation) of $\mathcal{S}^n_+$ can be obtained by constructing the dual cone of an outer approximation (inner approximation). We will focus on the outer approximations of $\mathcal{S}^n_+$ and refer to them as approximations of $\mathcal{S}^n_+$ throughout this paper.

Several sets have been used as approximations of the semidefinite cone, including the k-PSD closure, namely $\mathcal{S}^{n,k}$ ([5]) and the dual cone of the set of diagonally dominant matrices (resp., scaled diagonally dominant matrices), namely $\mathcal{DD}^*_n$ (resp., $\mathcal{SDD}^*_n$) ([2]). Multiple experiments have shown the efficiency of cutting-plane methods using these approximations ([2], [4], [18]). Although the inclusive relationship of these approximations has been given (see, e.g. [2], [4]), theoretical analyses of how well these sets approximate the semidefinite cone have been limited.

Fawzi [8] evaluated how polytopes can approximate a compact slice of the semidefinite cone by using a measure called extension complexity. Bertsimas and Cory-Wright [4] evaluated $\mathcal{DD}^*_n$ and $\mathcal{SDD}^*_n$ as approximations of the semidefinite cone by comparing the lower bounds on the minimum eigenvalues of matrices from these two sets. Blekherman et al. [5] evaluated $\mathcal{S}^{n,k}$, as an approximation of the semidefinite cone, by using an evaluation method called the norm normalized distance. The norm normalized distance between a given approximation $\mathcal{S} \subseteq \mathbb{S}^n$ and $\mathcal{S}^n_+$ is the maximum Frobenius distance from a matrix $X \in \mathcal{S}$ to $\mathcal{S}^n_+$, where the Frobenius norm of the matrix X is assumed to be one. They obtained several upper bounds and lower bounds on the norm normalized distance between $\mathcal{S}^{n,k}$ and $\mathcal{S}^n_+$.

In this paper, we first show that the norm normalized distance between a set $\mathcal{S}$ and $\mathcal{S}^n_+$ has the same value whenever $\mathcal{SDD}^*_n \subseteq \mathcal{S} \subseteq \mathcal{DD}^*_n$. This implies that the norm normalized distance is not a sufficient measure to differentiate these approximations. As a new measure to compensate for the weakness of that distance, we introduce a new distance, called the trace normalized distance. We prove that the trace normalized distance between $\mathcal{DD}^*_n$ and $\mathcal{S}^n_+$ has a different value from the one between $\mathcal{SDD}^*_n$ and $\mathcal{S}^n_+$ and give the exact values of these distances.

The organization of this paper is as follows. Section 2 introduces approximations of the semidefinite cone $\mathcal{S}^{n,k}$, $\mathcal{DD}^*_n$ and $\mathcal{SDD}^*_n$ and their inclusive relationship. In Section 3, the norm normalized distance proposed by Blekherman et al. [5] is used to evaluate $\mathcal{DD}^*_n$ and $\mathcal{SDD}^*_n$. In Section 4, the trace normalized distance is proposed and used to evaluate $\mathcal{DD}^*_n$ and $\mathcal{SDD}^*_n$. We conclude our work in Section 5.

2 Approximations of the semidefinite cone

Consider the following three sets as approximations of the semidefinite cone. Let k and n be positive integers satisfying $2 \leq k \leq n$ and

$$\mathcal{S}^{n,k} := \{X \in \mathbb{S}^n \mid \text{ All } k \times k \text{ principal submatrices of } X \text{ are positive semidefinite}\}. \tag{1}$$

$$\mathcal{DD}^*_n := \{X \in \mathbb{S}^n \mid X_{i,i} + X_{j,j} \pm 2X_{i,j} \geq 0 \ (1 \leq i \leq j \leq n)\}. \tag{2}$$

$$\mathcal{SDD}^*_n := \{X \in \mathbb{S}^n \mid X_{i,i}X_{j,j} \geq X^2_{i,j} \ (1 \leq i < j \leq n), X_{i,i} \geq 0 \ (i = 1, \ldots, n)\}. \tag{3}$$

$\mathcal{S}^{n,k}$ is called the *k-PSD closure*, whose properties are discussed in [5]. It is obvious from the definition (1) that $\mathcal{S}^n_+ = \mathcal{S}^{n,n} \subseteq \mathcal{S}^{n,k_1} \subseteq \mathcal{S}^{n,k_2}$ when $n \geq k_1 \geq k_2 \geq 2$.

$\mathcal{DD}^*_n$ (resp. $\mathcal{SDD}^*_n$) is the dual cone of the set of diagonally dominant matrices (resp. scaled diagonally dominant matrices). These duality relationships imply that $\mathcal{SDD}^*_n$ is a subset of $\mathcal{DD}^*_n$. It is worth noting that $\mathcal{DD}^*_n$ and $\mathcal{SDD}^*_n$ are used as approximations of the semidefinite cone in cutting-plane methods ([1],[2],[4],[18]) and facial reduction methods [15].

According to definitions (1) and (3), it is clear that $\mathcal{S}^{n,2} = \mathcal{SDD}^*_n$. The relations among $\mathcal{S}^{n,k}$, $\mathcal{DD}^*_n$ and $\mathcal{SDD}^*_n$ can be concluded to be

$$\mathcal{S}^n_+ = \mathcal{S}^{n,n} \subseteq \cdots \subseteq \mathcal{S}^{n,2} = \mathcal{SDD}^*_n \subseteq \mathcal{DD}^*_n. \tag{4}$$

3 The norm normalized distance between the semidefinite cone and its approximation

Blekherman et al. [5] proposed a method of evaluating approximations of the semidefinite cone, which is based on the maximum distance from a matrix in a given approximation $\mathcal{S}^n_+ \subseteq \mathcal{S} \subseteq \mathbb{S}^n$ to $\mathcal{S}^n_+$. A feature of their method is that the distance is evaluated under the constraint that the value of the Frobenius norm is one. The norm normalized distance between a set $\mathcal{S}$ and $\mathcal{S}^n_+$ is defined as

$$\overline{\mathrm{dist}}_F(\mathcal{S}, \mathcal{S}^n_+) := \sup_{X \in \mathcal{S}, \|X\|_F = 1} \|X - \mathrm{P}_{\mathcal{S}^n_+}(X)\|_F, \tag{5}$$

where $\mathrm{P}_{\mathcal{S}^n_+}(X) := \mathrm{argmin}_{Y \in \mathcal{S}^n_+} \|X - Y\|_F$ is the metric projection of X on $\mathcal{S}^n_+$.

The authors of [5] showed that $\overline{\mathrm{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \leq \frac{n-k}{n+k-2}$. Through a similar discussion to the one given there, we can prove the following theorem:

Theorem 1. *For $n \geq 4$,*

$$\overline{\mathrm{dist}}_F(\mathcal{DD}^*_n, \mathcal{S}^n_+) = \overline{\mathrm{dist}}_F(\mathcal{SDD}^*_n, \mathcal{S}^n_+) = \frac{n-2}{n}.$$

Note that the second equality (i.e., $\overline{\mathrm{dist}}_F(\mathcal{SDD}^*_n, \mathcal{S}^n_+) = \frac{n-2}{n}$) is a corollary of Blekherman et al.'s lower and upper bounds on $\mathcal{S}^{n,k}$ (Theorem 1 and 3 of [5]). The proof for $\overline{\mathrm{dist}}_F(\mathcal{DD}^*_n, \mathcal{S}^n_+)$ is provided in Appendix A.

Theorem 1 shows, unfortunately, that the norm normalized distance (5) gives the same value $\overline{\mathrm{dist}}_F(\mathcal{S}, \mathcal{S}^n_+) = \frac{n-2}{n}$ for any approximation $\mathcal{S} \subseteq \mathbb{S}^n$ whenever it satisfies $\mathcal{SDD}^*_n \subseteq \mathcal{S} \subseteq \mathcal{DD}^*_n$. In the next section, we introduce a new distance, called the trace normalized distance. We show that the new distance between $\mathcal{SDD}^*_n$ and $\mathcal{S}^n_+$ has a different value from the one between $\mathcal{DD}^*_n$ and $\mathcal{S}^n_+$.

4 The trace normalized distance between the semidefinite cone and its approximation

The Frobenius norm normalized distance can be generalized by expanding the normalization method and the distance function. For example, let $\mathcal{S}$ be an approximation such that $\mathcal{S}^n_+ \subseteq \mathcal{S} \subseteq \mathbb{S}^n$, $p \in [1, \infty]$, and $f : \mathbb{S}^n \to \mathbb{R}$ be a normalization function which requires the set $\{X \in \mathcal{S} \mid f(X) = 1\}$ to be bounded. One can define the $(f(\cdot), p)$ distance from $\mathcal{S}$ to $\mathcal{S}^n_+$ as:

$$\overline{\mathrm{dist}}_{(f(\cdot),p)}(\mathcal{S}, \mathcal{S}^n_+) := \sup_{X \in \mathcal{S}, f(X)=1} \|X - \mathrm{P}_{\mathcal{S}^n_+}(X)\|_p, \tag{6}$$

where $\|X\|_p := \sqrt[p]{\sum_{i=1}^n |\lambda_i|^p}$, where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of X, denotes the Schatten p-norm of $X \in \mathbb{S}^n$. Note that $\|X\|_\infty := \max\{|\lambda_1|, \ldots, |\lambda_n|\}$. In this notation, the Frobenius norm normalized distance (5) can be rewritten as an $(\|\cdot\|_F, 2)$ distance (6).

Recently, Blekherman et al. [6] studied a hyperbolic relaxation of $\mathcal{S}^{n,k}$ and provided an upper bound on the $(f(\cdot), \infty)$ distance from $\mathcal{S}^{n,k}$ to $\mathcal{S}^n_+$, where f can be any unitarily invariant matrix norm or the trace function. It can be observed that the infinity distance of a matrix X to $\mathcal{S}^n_+$, i.e., $\|X - \mathrm{P}_{\mathcal{S}^n_+}(X)\|_\infty$, is the absolute value of the most negative eigenvalue of X. In this paper, instead of evaluating the most negative eigenvalue of the matrices in the approximation $\mathcal{S}$, we try to figure out how $\mathcal{S}$ approximates $\mathcal{S}^n_+$ from a geometric point of view; i.e., we set $p = 2$ and stick with the Euclidean norm $\|\cdot\|_F$ of $\mathbb{S}^n$.

One reason why the Frobenius normalized distance (5) fails to distinguish $\mathcal{DD}^*_n$ and $\mathcal{SDD}^*_n$ might be that the constraint $\|X\|_F = 1$ is restrictive and makes the set $\{X \in \mathbb{S}^n \mid \|X\|_F = 1\}$ bounded. The required properties for normalization methods (e.g., $f(\cdot)$) here are only to make $\{X \in \mathcal{DD}^*_n \mid f(X) = 1\}$ and $\{X \in \mathcal{SDD}^*_n \mid f(X) = 1\}$ bounded.

There are some choices of the normalization method. For example, one may consider bounding another norm (i.e., $\|X\| = 1$), the determinant (i.e., $|X| = 1$), or the trace (i.e., $\mathrm{Tr}(X) := \sum_{i=1}^n X_{i,i} = 1$) of all matrices $X \in \mathbb{S}^n$. In the case of using other norms, we know from the equivalence of norms (e.g., Corollary 5.4.5 [11]) that for any norm $\|\cdot\|$ on $\mathbb{S}^n$ and for every matrix $X \in \mathbb{S}^n$, if $\|X\| = 1$, then $\|X\|_F$ is bounded from above. This shows that the set $\{X \in \mathbb{S}^n \mid \|X\| = 1\}$ is also bounded. Although it is not known whether distances defined with other norms can successfully distinguish $\mathcal{DD}^*_n$ and $\mathcal{SDD}^*_n$, studying these distances might be challenging, since the norm (e.g., the Schatten p-norm with $p \neq 2$) would be a complicated function in terms of entries of the matrix. As for the choice of determinant, although $\{X \in \mathbb{S}^n \mid |X| = 1\}$ is unbounded, one may notice that $\{X \in \mathcal{DD}^*_n \mid |X| = 1\}$ is also unbounded, which shows that the determinant is unavailable as a normalization method. Similarly, $\{X \in \mathbb{S}^n \mid \mathrm{Tr}(X) = 1\}$ is unbounded, but one can show that $\{X \in \mathcal{DD}^*_n \mid \mathrm{Tr}(X) = 1\}$ and $\{X \in \mathcal{SDD}^*_n \mid \mathrm{Tr}(X) = 1\}$ are bounded. Note as well that since $\mathrm{Tr}(X) = 1$ is a linear constraint, the subset of the polyhedral cone $\mathcal{DD}^*_n$ with trace equal to 1, i.e., $\{X \in \mathcal{DD}^*_n \mid \mathrm{Tr}(X) = 1\}$, is still polyhedral. In fact, we used this fact to derive the trace normalized distance between $\mathcal{DD}^*_n$ and $\mathcal{S}^n_+$. From the above discussion, we consider that a distance using the trace is effective for identifying the sets $\mathcal{DD}^*_n$ and $\mathcal{SDD}^*_n$.

Remark 1. *There actually is a very interesting norm that can be regarded as a normalization method equivalent to Tr$(\cdot)$ on $\mathcal{DD}^*_n$ and $\mathcal{SDD}^*_n$. Let $\mathcal{K}$ be a regular cone (i.e., $\mathcal{K}$ is convex closed pointed with nonempty interior) where $I \in \mathrm{int}K^*$ and $\|X\|_I := \min\{\mathrm{Tr}(X_1 + X_2) \mid X_1 - X_2 = X,\ X_1, X_2 \in \mathcal{K}\}$ be the norm induced by I, which was introduced in [9]. Proposition 1 of [9] implies that $\|X\|_I = \mathrm{Tr}(X)$ if $X \in \mathcal{K}$. By letting $\mathcal{K} = \mathcal{DD}^*_n$, it is straightforward to see that $\{X \in \mathcal{DD}^*_n \mid \|X\|_I = 1\} = \{X \in \mathcal{DD}^*_n \mid \mathrm{Tr}(X) = 1\}$. Since $\mathcal{SDD}^*_n \subseteq \mathcal{DD}^*_n$, $\|\cdot\|_I$ and Tr$(\cdot)$ are also equivalent on $\mathcal{SDD}^*_n$.*

We are now ready to use the Frobenius norm as the distance function and the trace function as the normalization method in (6). To simplify the notation, we will rewrite the $(\mathrm{Tr}(\cdot), 2)$ distance (6) into the following trace normalized distance from a set $\mathcal{S}$ and $\mathcal{S}^n_+$:

$$\overline{\mathrm{dist}}_T(\mathcal{S}, \mathcal{S}^n_+) := \sup_{X \in \mathcal{S}, \mathrm{Tr}(X)=1} \|X - \mathrm{P}_{\mathcal{S}^n_+}(X)\|_F, \tag{7}$$

As will be shown in the sections below, $\overline{\mathrm{dist}}_T(\mathcal{SDD}_n^*, \mathcal{S}_+^n)$ and $\overline{\mathrm{dist}}_T(\mathcal{DD}_n^*, \mathcal{S}_+^n)$ are different, i.e., $\overline{\mathrm{dist}}_T(\mathcal{SDD}_n^*, \mathcal{S}_+^n) = \frac{n-2}{n}$ (Theorem 2) and $\overline{\mathrm{dist}}_T(\mathcal{DD}_n^*, \mathcal{S}_+^n) = \frac{\sqrt{n-1}}{2}$ (Theorem 3).

Table 1 compares the proof techniques used by Blekherman et al. [5] and those in this paper. As can be seen in Table 1, the results for $\overline{\mathrm{dist}}_F(\mathcal{SDD}_n^*, \mathcal{S}_+^n)$, $\overline{\mathrm{dist}}_F(\mathcal{DD}_n^*, \mathcal{S}_+^n)$ and $\overline{\mathrm{dist}}_T(\mathcal{SDD}_n^*, \mathcal{S}_+^n)$ in this paper are proved by using techniques presented by Blekherman et al. [5]. Specifically, the "averaging technique" in Theorem 1 [5] (resp., the "specific matrix construction technique" in Theorem 3 [5]) is modified and used to prove Theorem 1 and Lemma 2 (resp., Lemma 1) in this paper. As for Theorem 3 in this paper, we obtained the exact result of $\overline{\mathrm{dist}}_T(\mathcal{DD}_n^*, \mathcal{S}_+^n)$ by analyzing the structures of extreme points of $\{X \in \mathcal{DD}_n^* \mid \mathrm{Tr}(X) = 1\}$ and by adopting the Bauer maximum principle. This new approach benefits from the polyhedral structure of $\{X \in \mathcal{DD}_n^* \mid \mathrm{Tr}(X) = 1\}$.

Note that in the proof of Theorem 2 [5], Blekherman et al. used the Cauchy interlacing theorem to calculate the number of negative eigenvalues of an arbitrary matrix $X \in \mathcal{S}^{n,k}$, which is at most $n-k$. They showed that $\|X - P_{\mathcal{S}_+^n}(X)\| \leq \sqrt{(n-k)\lambda_1^2}$, where λ_1 is the most negative eigenvalue of X. Finally, to get an upper bound of $\overline{\mathrm{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}_+^n)$, they only need to bound λ_1. However, in the setting of this paper (e.g., to analyze $\mathcal{SDD}_n^*$ and $\mathcal{DD}_n^*$), this approach may not be efficient. For example, because $\mathcal{SDD}_n^* = \mathcal{S}^{n,2}$, the matrices in $\mathcal{SDD}_n^*$ could have at most $n-2$ negative eigenvalues. If we bound the most negative eigenvalue of a matrix in $\mathcal{SDD}_n^*$ and use the Cauchy interlace theorem, we would inflate the bound by $\sqrt{n-2}$, making the upper bound of $\overline{\mathrm{dist}}_F(\mathcal{SDD}_n^*, \mathcal{S}_+^n)$ unnecessarily large.

Proof techniques of each theorem		Blekherman et al. 2022 [5]
Object		$\mathcal{S}^{n,k}$
$\overline{\mathrm{dist}}_F(\cdot, \mathcal{S}_+^n)$	Upper bound	(Theorem 1 [5]) Averaging technique
		(Theorem 2 [5]) Bound the most negative eigenvalue
	Lower bound	(Theorem 3 [5]) Construct a matrix far from $\mathcal{S}_+^n$
		(Theorem 4 [5]) Restricted Isometry Property

Proof techniques of each theorem		This paper	
Object		$\mathcal{SDD}_n^*$	$\mathcal{DD}_n^*$
$\overline{\mathrm{dist}}_F(\cdot, \mathcal{S}_+^n)$	Upper bound	(Theorem 1) $\overline{\mathrm{dist}}_F(\mathcal{SDD}_n^*, \mathcal{S}_+^n) \leq \overline{\mathrm{dist}}_F(\mathcal{DD}_n^*, \mathcal{S}_+^n)$	(Theorem 1) Averaging technique
	Lower bound	(Theorem 1) Corollary of Theorem 3 [5]	(Theorem 1) $\overline{\mathrm{dist}}_F(\mathcal{DD}_n^*, \mathcal{S}_+^n) \geq \overline{\mathrm{dist}}_F(\mathcal{SDD}_n^*, \mathcal{S}_+^n)$
$\overline{\mathrm{dist}}_T(\cdot, \mathcal{S}_+^n)$	Upper bound	(Lemma 2) Averaging technique	(Theorem 3) Analyze extreme points of $\{X \in \mathcal{DD}_n^* \mid \mathrm{Tr}(X) = 1\}$ and use the Bauer maximum principle
	Lower bound	(Lemma 1) Construct a matrix far from $\mathcal{S}_+^n$	

Table 1: Comparison of proof techniques in Blekherman et al. [5] and those in this paper.

4.1 The trace normalized distance between $\mathcal{SDD}_n^*$ and $\mathcal{S}_+^n$

Theorem 2. *For all $n \geq 2$,*

$$\overline{\mathrm{dist}}_T(\mathcal{SDD}_n^*, \mathcal{S}_+^n) = \frac{n-2}{n}.$$

To prove this theorem, we need Lemmas 1 and 2. Lemma 1 gives a lower bound on $\overline{\mathrm{dist}}_T(\mathcal{SDD}_n^*, \mathcal{S}_+^n)$ and Lemma 2 gives an upper bound on $\overline{\mathrm{dist}}_T(\mathcal{SDD}_n^*, \mathcal{S}_+^n)$. Here, we assume that $n \geq 3$. If $n = 2$, from the fact that $\mathcal{SDD}_2 = \mathcal{S}_+^2$, we can easily see that $\overline{\mathrm{dist}}_T(\mathcal{SDD}_2^*, \mathcal{S}_+^2) = \frac{n-2}{n} = 0$.

Lemmas 1 and 2 are based on the proofs of Theorems 3 and 1 in [5].

Lemma 1. *For all $n \geq 3$,*

$$\overline{\mathrm{dist}}_T(\mathcal{SDD}_n^*, \mathcal{S}_+^n) \geq \frac{n-2}{n}.$$

Proof. Let $I \in \mathbb{S}^n$ be the identity matrix and $e := (1, \ldots, 1)^T \in \mathbb{R}^n$. Given scalars $a, b \geq 0$, we define a matrix,

$$G(a, b, n) := (a + b)I - aee^T. \tag{8}$$

If $G(a, b, n) \in \mathcal{SDD}_n^* \setminus \mathcal{S}_+^n$ and $\mathrm{Tr}(G(a, b, n)) = 1$, then by definition (7), $\|G(a, b, n) - \mathrm{P}_{\mathcal{S}_+^n}(G(a, b, n))\|_F$ gives a lower bound on $\overline{\mathrm{dist}}_T(\mathcal{SDD}_n^*, \mathcal{S}_+^n)$. To find a tighter lower bound, we consider the following problem (9) on the parameters a and b:

$$\begin{aligned} \max\nolimits_{a,b\geq 0} \quad & \|G(a, b, n) - \mathrm{P}_{\mathcal{S}_+^n}(G(a, b, n))\|_F && \text{(9a)} \\ \text{s.t.} \quad & G(a, b, n) \notin \mathcal{S}_+^n, && \text{(9b)} \\ & G(a, b, n) \in \mathcal{SDD}_n^*, && \text{(9c)} \\ & \mathrm{Tr}(G(a, b, n)) = 1. && \text{(9d)} \end{aligned}$$

Problem (9) can be equivalently written as:

$$\begin{aligned} \max\nolimits_{a,b\geq 0} \quad & (n-1)a - b && \text{(10a)} \\ \text{s.t.} \quad & b < (n-1)a, && \text{(10b)} \\ & b \geq a, && \text{(10c)} \\ & nb = 1. && \text{(10d)} \end{aligned}$$

To prove the equivalence between (9) and (10), we first show that the constraints (9b) and (10b) are equivalent. Proposition 4 in [5] ensures that the eigenvalues of $G(a, b, n)$ are $a + b$ with multiplicity $n - 1$ and $b - (n-1)a$ with multiplicity 1. Note that $a, b \geq 0$; hence,

$$G(a, b, n) \notin \mathcal{S}_+^n \text{ if and only if } b < (n-1)a. \tag{11}$$

Next, we verify that (9c) and (10c) are equivalent. It follows from definition (3) that $G(a, b, n) \in \mathcal{SDD}_n^*$ if and only if all the 2×2 submatrices of $G(a, b, n)$ are positive semidefinite. It is obvious from (8) that any 2×2 submatrix of $G(a, b, n)$ is $G(a, b, 2)$. (11) ensures that $G(a, b, 2) \in \mathcal{S}_+^2$ if and only if $b \geq a$ and we can conclude that

$$G(a, b, n) \in \mathcal{SDD}_n^* \text{ if and only if } b \geq a. \tag{12}$$

The equivalence between (9d) and (10d) comes from the fact that the definition (8) implies that

$$\mathrm{Tr}(G(a, b, n)) = nb. \tag{13}$$

Finally, we show that the objective functions (9a) and (10a) are equivalent. Since (9b) implies that $G(a,b,n) \notin \mathcal{S}_+^n$, it is apparent from (11) that $b-(n-1)a<0$. Then, $b-(n-1)a$ is the only negative eigenvalue of $G(a,b,n)$, and hence,

$$\|G(a,b,n) - \mathrm{P}_{\mathcal{S}_+^n}(G(a,b,n))\|_F = (n-1)a - b. \tag{14}$$

One can see that problems (9) and (10) are equivalent from (11), (12), (13) and (14). Then for any feasible solution (a,b) of (10), $(n-1)a-b$ gives a lower bound on $\overline{\mathrm{dist}}_T(\mathcal{SDD}_n^*, \mathcal{S}_+^n)$. The optimal solution of problem (10) is $\bar{a} = \bar{b} = \frac{1}{n}$; hence, we have

$$\overline{\mathrm{dist}}_T(\mathcal{SDD}_n^*, \mathcal{S}_+^n) \geq \|G(\bar{a},\bar{b},n) - \mathrm{P}_{\mathcal{S}_+^n}(G(\bar{a},\bar{b},n))\|_F = \frac{n-2}{n}.$$

$\square$

Lemma 2. *For all $n \geq 3$,*

$$\overline{\mathrm{dist}}_T(\mathcal{SDD}_n^*, \mathcal{S}_+^n) \leq \frac{n-2}{n}.$$

Proof. If a scalar U satisfies $\|X - \mathrm{P}_{\mathcal{S}_+^n}(X)\|_F \leq U$ for every $X \in \mathcal{SDD}_n^*$ with $\mathrm{Tr}(X) = 1$, then U is an upper bound on $\overline{\mathrm{dist}}_T(\mathcal{SDD}_n^*, \mathcal{S}_+^n)$. Below, we find such a scalar U.

Let X be a matrix in $\mathcal{SDD}_n^*$ satisfying $\mathrm{Tr}(X) = 1$. We construct a matrix $\tilde{X} \in \mathcal{S}_+^n$ and a scalar $\tilde{\alpha} \geq 0$ in a way such that $\|X - \tilde{\alpha}\tilde{X}\|_F \leq U$, which then shows that $\|X - P_{\mathcal{S}_+^n}(X)\|_F \leq U$.

Define a matrix $X^{(i,j)} \in \mathbb{S}^n$ for every $1 \leq i < j \leq n$:

$$X_{p,q}^{(i,j)} := \begin{cases} X_{i,i} & (\text{if } p = q = i), \\ X_{j,j} & (\text{if } p = q = j), \\ X_{i,j} & (\text{if } (p,q) \in \{(i,j),(j,i)\}), \\ 0 & (\text{otherwise}). \end{cases} \tag{15}$$

Let $C_n^k := \frac{n!}{(n-k)!k!}$ and let $\tilde{X} = \frac{1}{C_n^2}\sum_{1\leq i<j\leq n} X^{(i,j)}$. Then, (15) implies that

$$\begin{aligned} \tilde{X}_{i,i} =& \frac{C_n^2 - C_{n-1}^2}{C_n^2} X_{i,i} = \frac{2}{n} X_{i,i} \ (i = 1, \ldots, n), \\ \tilde{X}_{i,j} =& \frac{1}{C_n^2} X_{i,j} = \frac{2}{n(n-1)} X_{i,j} \ (1 \leq i < j \leq n). \end{aligned}$$

By (3), we know that $X^{(i,j)} \in \mathcal{S}_+^n$ for all $1 \leq i < j \leq n$ and hence $\tilde{X} \in \mathcal{S}_+^n$.

Let $\alpha \geq 0$ be any scalar. Then,

$$
\begin{aligned}
\|X - \alpha\tilde{X}\|_F &= \sqrt{\sum_{i=1}^{n}(X - \alpha\tilde{X})_{i,i}^2 + \sum_{i\neq j}(X - \alpha\tilde{X})_{i,j}^2} \\
&= \sqrt{\sum_{i=1}^{n}(1 - \frac{2\alpha}{n})^2 X_{i,i}^2 + \sum_{i\neq j}(1 - \frac{2\alpha}{n(n-1)})^2 X_{i,j}^2} \\
&\leq \sqrt{(1 - \frac{2\alpha}{n})^2 \sum_{i=1}^{n} X_{i,i}^2 + (1 - \frac{2\alpha}{n(n-1)})^2 \sum_{i\neq j} X_{i,i}X_{j,j}} \\
&= \sqrt{(1 - \frac{2\alpha}{n})^2 \sum_{i=1}^{n} X_{i,i}^2 + (1 - \frac{2\alpha}{n(n-1)})^2 (\mathrm{Tr}(X)^2 - \sum_{i=1}^{n} X_{i,i}^2)} \\
&= \sqrt{\left((1 - \frac{2\alpha}{n})^2 - (1 - \frac{2\alpha}{n(n-1)})^2\right) \sum_{i=1}^{n} X_{i,i}^2 + (1 - \frac{2\alpha}{n(n-1)})^2 \mathrm{Tr}(X)^2} \\
&= \sqrt{\left(1 - \frac{4\alpha}{n} + \frac{4\alpha^2}{n^2} - (1 - \frac{4\alpha}{n(n-1)} + \frac{4\alpha^2}{n^2(n-1)^2})\right) \sum_{i=1}^{n} X_{i,i}^2 + (1 - \frac{2\alpha}{n(n-1)})^2 \mathrm{Tr}(X)^2} \\
&= \sqrt{\left(\frac{4\alpha^2(n-2)}{n(n-1)^2} - \frac{4\alpha(n-2)}{n(n-1)}\right) \sum_{i=1}^{n} X_{i,i}^2 + (1 - \frac{2\alpha}{n(n-1)})^2 \mathrm{Tr}(X)^2}. \qquad (16)
\end{aligned}
$$

Note that $\mathrm{Tr}(X) = 1$ and $\tilde{\alpha} := n - 1 \geq 0$ satisfies that $\frac{4\tilde{\alpha}^2(n-2)}{n(n-1)^2} - \frac{4\tilde{\alpha}(n-2)}{n(n-1)} = 0$. By substituting $\tilde{\alpha}$ into (16), we have

$$\|X - \tilde{\alpha}\tilde{X}\|_F \leq \sqrt{(1 - \frac{2\tilde{\alpha}}{n(n-1)})^2} = \frac{n-2}{n}.$$

Since $\tilde{\alpha} \geq 0$ and $\tilde{X} \in \mathcal{S}_+^n$, by letting $U = \frac{n-2}{n}$, we can see that

$$\|X - \mathrm{P}_{\mathcal{S}_+^n}(X)\|_F \leq \|X - \tilde{\alpha}\tilde{X}\|_F \leq U = \frac{n-2}{n},$$

and hence,

$$\overline{\mathrm{dist}}_T(\mathcal{SDD}_n^*, \mathcal{S}_+^n) = \sup_{X \in \mathcal{SDD}_n^*, \mathrm{Tr}(X)=1} \|X - \mathrm{P}_{\mathcal{S}_+^n}(X)\|_F \leq U = \frac{n-2}{n}.$$

$\square$

4.2 The trace normalized distance between $\mathcal{DD}_n^*$ and $\mathcal{S}_+^n$

In this section, we prove the following theorem:

Theorem 3. *For all $n \geq 2$,*

$$\overline{\mathrm{dist}}_T(\mathcal{DD}_n^*, \mathcal{S}_+^n) = \frac{\sqrt{n} - 1}{2}.$$

The idea behind Theorem 3 is as follows. Define

$$\mathcal{DDT}_n^* := \mathcal{DD}_n^* \cap \{X \in \mathbb{S}^n \mid \mathrm{Tr}(X) = 1\}. \tag{17}$$

Definition (7) ensures that

$$\overline{\mathrm{dist}}_T(\mathcal{DD}_n^*, \mathcal{S}_+^n) = \max_{X \in \mathcal{DDT}_n^*} \|X - \mathrm{P}_{\mathcal{S}_+^n}(X)\|_F.$$

Note that $\|X - \mathrm{P}_{\mathcal{S}_+^n}(X)\|_F$ is continuous and convex on $\mathbb{S}^n$ and $\mathcal{DDT}_n^*$ is closed, bounded, and convex. The Bauer maximum principle [3] states that any continuous convex function defined on a compact convex set in $\mathbb{R}^n$ attains its maximum at some extreme point of the set. As a corollary, we have the following:

Corollary 1. $\max_{X \in \mathcal{DDT}_n^*} \|X - \mathrm{P}_{\mathcal{S}_+^n}(X)\|_F$ *attains its maximum at some extreme point of* $\mathcal{DDT}_n^*$.

Proposition 1 shows that every extreme point of $\mathcal{DDT}_n^*$ has a special structure. Lemma 3 uses this special structure to show that the distance from each extreme point of $\mathcal{DDT}_n^*$ to $\mathcal{S}_+^n$ is the same. The exact distance is also given in Lemma 3. Theorem 3 follows directly from Corollary 1 and Lemma 3.

Proposition 1. *For $n \geq 2$, let X be an extreme point of $\mathcal{DDT}_n^*$. There exists an integer q satisfying $1 \leq q \leq n$ such that*

$$X_{i,j} = \begin{cases} 1 & (\text{ if } i = j = q), \\ \frac{1}{2} \text{ or } -\frac{1}{2} & (\text{ if either } i = q \text{ or } j = q), \\ 0 & (\text{ otherwise}). \end{cases} \tag{18}$$

Proof. Let $X \in \mathcal{DDT}_n^*$. By (2) and (17), we see that for every $i = 1, \ldots, n$,

$$X_{i,i} \geq 0,$$

and for every $1 \leq i < j \leq n$,

$$X_{i,i} + X_{j,j} + 2X_{i,j} \geq 0, \tag{19}$$
$$X_{i,i} + X_{j,j} - 2X_{i,j} \geq 0. \tag{20}$$

Thus, the set $\mathcal{DDT}_n^*$ can be written as

$$\begin{aligned} \mathcal{DDT}_n^* = \{X \in \mathbb{S}^n \mid & \mathrm{Tr}(X) = 1, \\ & X_{i,i} \geq 0 \ (i = 1, \ldots, n), & (21) \\ & X_{i,i} + X_{j,j} + 2X_{i,j} \geq 0 \ (1 \leq i < j \leq n), & (22) \\ & X_{i,i} + X_{j,j} - 2X_{i,j} \geq 0 \ (1 \leq i < j \leq n)\}. & (23) \end{aligned}$$

Let $\bar{X}$ be an extreme point of $\mathcal{DDT}_n^*$ and let $N(X)$ be the number of linearly independent inequalities in (21), (22) and (23) that are active (i.e., the equalities hold) at $X \in \mathcal{DDT}_n^*$. From a characterization of the extreme points of a polyhedron (see, e.g., Theorem 5.7, [16]), we know that

$$N(\bar{X}) = \frac{n(n+1)}{2} - 1. \tag{24}$$

Below, we prove that $\bar{X}$ satisfies (18) by observing the active inequalities at $\bar{X}$.

It follows from $\mathrm{Tr}(\bar{X}) = 1$ that $\bar{X}$ has at least one nonzero diagonal element. This implies that the number of active inequalities in (21) at $\bar{X}$ is at most $n-1$. Suppose that $n-k$ inequalities in (21) are active at $\bar{X}$, where k is an integer and $1 \le k \le n$. Below, we show that $k \neq n$ by contradiction.

Assume that $k = n$. Then we have $\bar{X}_{i,i} > 0$ for each $1 \le i \le n$. At most one of (19) and (20) can be active at $\bar{X}$ for each $1 \le i < j \le n$. In fact, suppose that (19) and (20) are simultaneously active for some $1 \le i < j \le n$:

$$\bar{X}_{i,i} + \bar{X}_{j,j} + 2\bar{X}_{i,j} = 0,\ \ \bar{X}_{i,i} + \bar{X}_{j,j} - 2\bar{X}_{i,j} = 0.$$

Then, $\bar{X}_{i,j} = \bar{X}_{i,i} + \bar{X}_{j,j} = 0$ and since $\bar{X}_{i,i}, \bar{X}_{j,j} \ge 0$, we obtain $\bar{X}_{i,i} = \bar{X}_{j,j} = 0$, which is a contradiction to the assumption $\bar{X}_{i,i}, \bar{X}_{j,j} > 0$. This implies that $N(\bar{X})$ is at most $\frac{n(n-1)}{2}$, which is strictly less than the number $\frac{n(n+1)}{2} - 1$ in (24). This contradiction implies that $k \neq n$.

Since we have shown that $1 \le k \le n-1$, there exists a permutation matrix $P \in \mathbb{R}^{n \times n}$ such that the matrix $X^* := P\bar{X}P^T$ satisfies

$$\begin{aligned} &X^*_{i,i} = 0 \quad (1 \le i \le n-k), \\ &X^*_{i,i} > 0 \quad (n-k+1 \le i \le n). \end{aligned} \tag{25}$$

Note that $X^* \in \mathcal{DDT}^*_n$ and $N(X^*) = \frac{n(n+1)}{2} - 1$. Below, we show that X^* satisfies (18) by observing the active inequalities at X^* instead of $\bar{X}$.

Next, we show that $k = 1$; i.e., exactly $n-1$ inequalities in (21) are active at X^*. It follows from (22), (23) and (25) that $X^*_{i,j} = 0$ for each $1 \le i < j \le n-k$. This implies that all inequalities (19) and (20) with $1 \le i < j \le n-k$ at X^* are active. For each pair of (i,j) where $X_{j,j} > 0$ and $1 \le i < j$, one can show again by contradiction that at most one of (19) and (20) can be active at X^*. Consider the case when the number of active inequalities at X^* attains its maximum; i.e., exactly one of (19) and (20) is active at X^* for each pair of (i,j), where $n-k+1 \le j \le n$ and $1 \le i < j$. The following system,

$$\begin{cases} 0 = X^*_{i,i} \quad (1 \le i \le n-k), \\ 0 = X^*_{i,i} + X^*_{j,j} + 2X^*_{i,j} \quad (1 \le i < j \le n-k), \\ 0 = X^*_{i,i} + X^*_{j,j} - 2X^*_{i,j} \quad (1 \le i < j \le n-k), \\ \text{either (19) or (20) is active at } X^* \quad (n-k+1 \le j \le n,\ 1 \le i < j), \end{cases}$$

includes exactly $(n-k) + \frac{n(n-1)}{2} = \frac{n(n+1)}{2} - k$ linearly independent active inequalities. This implies that $N(X^*) \le \frac{n(n+1)}{2} - k$. By (24), we know that $k = 1$ and $\frac{n(n+1)}{2} - 1$ in (24) is attained only if the number of active inequalities in (22) and (23) attains its maximum.

$k = 1$ implies that $X^*_{n,n} = 1$ and $X^*_{i,i} = 0$ for each $1 \le i \le n-1$; and hence, $X^*_{i,j} = 0$ for each $1 \le i < j \le n-1$. Since the number of active inequalities in (22) and (23) attains its maximum, we know that either (19) or (20) is active at X^* for each (i,j) satisfying $j = n$ and $1 \le i < j$, which implies that $X^*_{i,n} \in \{\frac{1}{2}, -\frac{1}{2}\}$ for each $1 \le i < n$.

Finally, by applying the permutation $\bar{X} = P^T X^* P$, we know that there exists an integer q satisfying $1 \le q \le n$ for which $\bar{X}$ satisfies (18).

□

Lemma 3. *For $n \ge 2$, let X be an extreme point of $\mathcal{DDT}^*_n$. There exist scalars $\alpha_1, \dots, \alpha_{n-1} \in \{\frac{1}{2}, -\frac{1}{2}\}$*

such that the following matrix,

$$X^* := \begin{pmatrix} 0 & & & \alpha_1 \\ & \ddots & & \vdots \\ & & 0 & \alpha_{n-1} \\ \alpha_1 & \dots & \alpha_{n-1} & 1 \end{pmatrix} \tag{26}$$

satisfies

$$\|X - \mathrm{P}_{\mathcal{S}^n_+}(X)\|_F = \|X^* - \mathrm{P}_{\mathcal{S}^n_+}(X^*)\|_F = \frac{\sqrt{n}-1}{2}.$$

Proof. Let X be an extreme point of $\mathcal{DDT}^*_n$. By Proposition 1, there exists an integer q such that $1 \leq q \leq n$ for which X satisfies (18). Note that X only has one nonzero diagonal element $X_{q,q} = 1$. Let $P \in \mathbb{R}^{n \times n}$ be a permutation matrix such that $(PXP^T)_{n,n} = 1$. It is easy to see that there are scalars $\alpha_1, \dots, \alpha_{n-1} \in \{\frac{1}{2}, -\frac{1}{2}\}$ such that the matrix X^* defined in (26) satisfies $X^* = PXP^T$. Since the permutation matrix P is orthogonal, we see that $Y \in \mathcal{S}^n_+$ if and only if $PYP^T \in \mathcal{S}^n_+$ for any $Y \in \mathbb{S}^n$. This fact implies that

$$\begin{aligned} \|X - \mathrm{P}_{\mathcal{S}^n_+}(X)\|_F &= \inf_{Y \in \mathcal{S}^n_+} \|X - Y\|_F \\ &= \inf_{Y \in \mathcal{S}^n_+} \|PXP^T - PYP^T\|_F \\ &= \inf_{PYP^T \in \mathcal{S}^n_+} \|X^* - PYP^T\|_F \\ &= \|X^* - \mathrm{P}_{\mathcal{S}^n_+}(X^*)\|_F. \end{aligned} \tag{27}$$

By solving the eigenvalue equation $0 = |\lambda I - X^*|$ with respect to the scalar λ, we obtain that:

1. If $n = 2$, the eigenvalues of X^* are $\frac{1+\sqrt{n}}{2}$ with multiplicity 1 and $\frac{1-\sqrt{n}}{2}$ with multiplicity 1.

2. If $n \geq 3$, the eigenvalues of X^* are $\frac{1+\sqrt{n}}{2}$ with multiplicity 1, $\frac{1-\sqrt{n}}{2}$ with multiplicity 1 and 0 with multiplicity $n-2$.

From these observations, for every $n \geq 2$, X^* has only one negative eigenvalue $\lambda_{\min} := \frac{1-\sqrt{n}}{2}$; hence,

$$\|X^* - \mathrm{P}_{\mathcal{S}^n_+}(X^*)\|_F = \sqrt{\lambda^2_{\min}} = \frac{\sqrt{n}-1}{2}. \tag{28}$$

We conclude from (27) and (28) that

$$\|X - \mathrm{P}_{\mathcal{S}^n_+}(X)\|_F = \|X^* - \mathrm{P}_{\mathcal{S}^n_+}(X^*)\|_F = \frac{\sqrt{n}-1}{2}.$$

□

5 Concluding remarks

We showed that the norm normalized distance $\overline{\mathrm{dist}}_F(\mathcal{S}, \mathcal{S}^n_+)$ has the same value whenever $\mathcal{SDD}^*_n \subseteq \mathcal{S} \subseteq \mathcal{DD}^*_n$, since $\overline{\mathrm{dist}}_F(\mathcal{DD}^*_n, \mathcal{S}^n_+) = \overline{\mathrm{dist}}_F(\mathcal{SDD}^*_n, \mathcal{S}^n_+)$ holds. This implies that the norm normalized distance is not a sufficient measure to evaluate these approximations. Moreover, as a new measure to compensate for the weakness of that distance, we proposed a new distance, the trace normalized distance $\overline{\mathrm{dist}}_T(\mathcal{S}, \mathcal{S}^n_+)$. Using this new measure, we proved that $\overline{\mathrm{dist}}_T(\mathcal{DD}^*_n, \mathcal{S}^n_+)$ and $\overline{\mathrm{dist}}_T(\mathcal{SDD}^*_n, \mathcal{S}^n_+)$ are different, i.e., $\overline{\mathrm{dist}}_T(\mathcal{DD}^*_n, \mathcal{S}^n_+) = \frac{\sqrt{n}-1}{2}$ and $\overline{\mathrm{dist}}_T(\mathcal{SDD}^*_n, \mathcal{S}^n_+) = \frac{n-2}{n}$.

In [18], the authors proposed a class of polyhedral approximations of the semidefinite cone, denoted as $\mathcal{SDB}^*_n$. The experimental results on cutting-plane methods, where $\mathcal{SDB}^*_n$ is used as an approximation of $\mathcal{S}^n_+$ for solving SDP instances are promising. It is an interesting but also challenging issue to analyze the value of $\overline{\mathrm{dist}}_T(\mathcal{SDB}^*_n, \mathcal{S}^n_+)$. To evaluate approximations of the semidefinite cone, one may also study general properties of the set of eigenvalues that are attained by matrices in the approximation sets. For example, Kozhasov [12] proved that the set of eigenvalue vectors of matrices in $\mathcal{S}^{4,2} = \mathcal{SDD}^*_4$ is not convex. It is also an interesting and challenging direction to study the set of eigenvalues that are attained by matrices in $\mathcal{DD}^*_n$ and $\mathcal{SDD}^*_n$ respectively.

6 Acknowledgments

This research was supported by the Japan Society for the Promotion of Science through a Grant-in-Aid for Scientific Research ((B)19H02373) from the Ministry of Education, Culture, Sports, Science and Technology of Japan. The authors would like to sincerely thank the anonymous reviewers for their thoughtful and valuable comments which have significantly improved the paper.

References

[1] Ahmadi, A.A., Dash, S., Hall, G.: *Optimization over structured subsets of positive semidefinite matrices via column generation.* Discrete Optimization, 24, pp. 129–151 (2017).

[2] Ahmadi, A.A., Majumdar, A.: *DSOS and SDSOS optimization: more tractable alternatives to sum of squares and semidefinite optimization.* SIAM Journal on Applied Algebra and Geometry, 3, pp. 193–230 (2019).

[3] Bauer, H.: *Minimalstellen von Funktionen und Extremalpunkte.* Archiv der Mathematik, 9, 389–393 (1958). https://doi.org/10.1007/BF01898615

[4] Bertsimas, D., Cory-Wright, R.: *On polyhedral and second-order cone decompositions of semidefinite optimization problems.* Operations Research Letters, 48(1), pp. 78–85 (2020).

[5] Blekherman, G., Dey, S.S., Molinaro, M., Sun, S.: *Sparse PSD approximation of the PSD cone.* Mathematical Programming, 191, pp. 981-1004 (2022).

[6] Blekherman, G., Dey, S.S., Shu, K., Sun, S.: *Hyperbolic Relaxation of k-Locally Positive Semidefinite Matrices.* arXiv preprint arXiv:2012.04031 (2021).

[7] Boman, E.G., Chen, D., Parekh, O., Toledo, S.: *On factor width and symmetric H-matrices.* Linear Algebra and Its Applications, 405, pp. 239–248 (2005).

[8] Fawzi, H.: *On Polyhedral Approximations of the Positive Semidefinite Cone.* Mathematics of Operations Research, 46(4), pp. 1479-1486 (2021).

[9] Freund, R. M.: *On the behavior of the homogeneous self-dual model for conic convex optimization.* Mathematical programming, 106, pp. 527-545 (2006).

[10] Goemans, M.X., Williamson, D.P.: *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming.* Journal of the ACM, 42, pp. 1115–1145 (1995).

[11] R. A. Horn and C. R. Johnson: *Matrix Analysis*, Cambridge university press (1990).

[12] Kozhasov, K.: *On eigenvalues of symmetric matrices with PSD principal submatrices.* arXiv preprint arXiv:2103.15811 (2021).

[13] Lasserre, J.B.: *An explicit exact SDP relaxation for nonlinear 0-1 programs*, in Integer Programming and Combinatorial Optimization, K. Aardal and B. Gerards, eds., Berlin, Heidelberg, 2001, Springer Berlin Heidelberg, pp. 293–303.

[14] MOSEK ApS: *The MOSEK optimization toolbox for MATLAB manual. Version 8.1.*, available at `http://docs.mosek.com/8.1/toolbox/index.html`,2017, [Accessed: 2021-3-12]

[15] Permenter, F., Parrilo, P.: *Partial facial reduction: simplified, equivalent SDPs via approximations of the PSD cone.* Mathematical Programming, 171, pp. 1–54 (2018).

[16] Schrijver, A.: *Combinatorial optimization: polyhedra and efficiency.*vol. 24, Springer Science & Business Media (2003).

[17] Todd, M.J.: *Semidefinite optimization*, Acta Numerica, 10, pp. 515–560 (2001).

[18] Wang, Y., Tanaka, A., Yoshise, A.: *Polyhedral approximations of the semidefinite cone and their application*, Comput Optim Appl, 78, pp. 893–913 (2021). https://doi.org/10.1007/s10589-020-00255-2

[19] Wolkowicz, H., Saigal, R., and Vandenberghe, L.: *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*, vol. 27, Springer Science & Business Media (2012).

Appendix

A Proof of Theorem 1

Proof. We prove the following inequalities:

$$\frac{n-2}{n} \leq \overline{\mathrm{dist}}_F(\mathcal{SDD}_n^*, \mathcal{S}_+^n) \leq \overline{\mathrm{dist}}_F(\mathcal{DD}_n^*, \mathcal{S}_+^n) \leq \frac{n-2}{n}. \tag{29}$$

The relation $\mathcal{S}^{n,2} = \mathcal{SDD}_n^*$ in (4) implies that $\overline{\mathrm{dist}}_F(\mathcal{SDD}_n^*, \mathcal{S}_+^n) = \overline{\mathrm{dist}}_F(\mathcal{S}^{n,2}, \mathcal{S}_+^n)$. By Theorem 3 in [5], we know that $\overline{\mathrm{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}_+^n) \geq \frac{n-k}{\sqrt{(k-1)^2 n + n(n-1)}}$ and hence that,

$$\overline{\mathrm{dist}}_F(\mathcal{SDD}_n^*, \mathcal{S}_+^n) = \overline{\mathrm{dist}}_F(\mathcal{S}^{n,2}, \mathcal{S}_+^n) \geq \frac{n-2}{\sqrt{(2-1)^2 n + n(n-1)}} = \frac{n-2}{n}. \tag{30}$$

The relation $\mathcal{SDD}_n^* \subseteq \mathcal{DD}_n^*$ in (4) ensures that

$$\overline{\mathrm{dist}}_F(\mathcal{SDD}_n^*, \mathcal{S}_+^n) \leq \overline{\mathrm{dist}}_F(\mathcal{DD}_n^*, \mathcal{S}_+^n). \tag{31}$$

Next, we prove that $\overline{\mathrm{dist}}_F(\mathcal{DD}_n^*, \mathcal{S}_+^n) \leq \frac{n-2}{n}$ with the following idea. If a scalar U satisfies $\|X - \mathrm{P}_{\mathcal{S}_+^n}(X)\|_F \leq U$ for every $X \in \mathcal{DD}_n^*$ with $\|X\|_F = 1$, then U is an upper bound on $\overline{\mathrm{dist}}_F(\mathcal{DD}_n^*, \mathcal{S}_+^n)$. We can find such a scalar U by constructing a matrix $\tilde{X} \in \mathcal{S}_+^n$ and a scalar $\tilde{\alpha} \geq 0$ for every $X \in \mathcal{DD}_n^*$ with $\|X\|_F = 1$ in such a way that $\|X - \mathrm{P}_{\mathcal{S}_+^n}(X)\|_F \leq \|X - \tilde{\alpha}\tilde{X}\|_F$.

Let X be a matrix in $\mathcal{DD}_n^*$ satisfying $\|X\|_F = 1$. Define a matrix $X^{(i,j)} \in \mathbb{S}^n$ for every $1 \leq i < j \leq n$:

$$X^{(i,j)}_{p,q} := \begin{cases} \frac{X_{i,i}+X_{j,j}}{2} & (\text{if } p = q \in \{i,j\}), \\ X_{i,j} & (\text{if } (p,q) \in \{(i,j),(j,i)\}), \\ 0 & (\text{otherwise}). \end{cases} \tag{32}$$

Let $\bar{X} := \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} X^{(i,j)}$. By definitions (2) and (32), one can verify that $X^{(i,j)} \in \mathcal{S}_+^n$ for all $1 \leq i < j \leq n$ and hence that $\bar{X} \in \mathcal{S}_+^n$. Let α be a scalar satisfying $\alpha \geq \frac{2n(n-1)}{3n-4} > 0$. For all $1 \leq i < j \leq n$, we can obtain from (32) that $\bar{X}_{i,j} = \bar{X}_{j,i} = \frac{2}{n(n-1)} X_{i,j}$ and hence that

$$\sum_{i \neq j} (X_{i,j} - \alpha \bar{X}_{i,j})^2 = \sum_{i \neq j} (1 - \frac{2\alpha}{n(n-1)})^2 X_{i,j}^2. \tag{33}$$

For all $i = 1, \dots, n$, (32) implies that $\bar{X}_{i,i} = \frac{2}{n(n-1)} \left(\frac{n-2}{2} X_{i,i} + \frac{1}{2}\mathrm{Tr}(X) \right)$ and hence that

$$\begin{aligned} \sum_{i=1}^n (X_{i,i} - \alpha \bar{X}_{i,i})^2 =& \sum_{i=1}^n ((1 - \frac{\alpha(n-2)}{n(n-1)}) X_{i,i} - \frac{\alpha}{n(n-1)} \mathrm{Tr}(X))^2 \\ =& \sum_{i=1}^n ((1 - \frac{\alpha(n-2)}{n(n-1)})^2 X_{i,i}^2 - 2(1 - \frac{\alpha(n-2)}{n(n-1)}) X_{i,i} \frac{\alpha}{n(n-1)} \mathrm{Tr}(X) \\ &+ \frac{\alpha^2}{n^2(n-1)^2} \mathrm{Tr}(X)^2) \\ =& (1 - \frac{\alpha(n-2)}{n(n-1)})^2 \sum_{i=1}^n X_{i,i}^2 - (\frac{2\alpha}{n(n-1)} - \frac{2\alpha^2(n-2)}{n^2(n-1)^2}) \mathrm{Tr}(X) \sum_{i=1}^n X_{i,i} \\ &+ \frac{\alpha^2 n}{n^2(n-1)^2} \mathrm{Tr}(X)^2 \\ =& (1 - \frac{\alpha(n-2)}{n(n-1)})^2 \sum_{i=1}^n X_{i,i}^2 + (\frac{\alpha^2(3n-4)}{n^2(n-1)^2} - \frac{2\alpha}{n(n-1)}) \mathrm{Tr}(X)^2. \end{aligned} \tag{34}$$

The assumption $\alpha \geq \frac{2n(n-1)}{3n-4}$ ensures that $\frac{\alpha^2(3n-4)}{n^2(n-1)^2} - \frac{2\alpha}{n(n-1)} \geq 0$. One can verify that $\mathrm{Tr}(X)^2 \leq$

$n\sum_{i=1}^n X_{i,i}^2$ by using the Cauchy-Schwarz inequality. Then, it follows from (34) that

$$
\begin{aligned}
\sum_{i=1}^n (X_{i,i}-\alpha\bar{X}_{i,i})^2 \leq & (1-\frac{\alpha(n-2)}{n(n-1)})^2\sum_{i=1}^n X_{i,i}^2 + (\frac{\alpha^2(3n-4)}{n^2(n-1)^2}-\frac{2\alpha}{n(n-1)})n\sum_{i=1}^n X_{i,i}^2 \\
&= \left((1-\frac{\alpha(n-2)}{n(n-1)})^2+\frac{\alpha^2 n(3n-4)}{n^2(n-1)^2}-\frac{2\alpha n}{n(n-1)}\right)\sum_{i=1}^n X_{i,i}^2 \\
&= \left(1-\frac{2\alpha(n-2)}{n(n-1)}+\frac{\alpha^2(n-2)^2}{n^2(n-1)^2}+\frac{\alpha^2 n(3n-4)}{n^2(n-1)^2}-\frac{2\alpha n}{n(n-1)}\right)\sum_{i=1}^n X_{i,i}^2 \\
&= \left(1-\frac{2\alpha(n-2+n)}{n(n-1)}+\frac{\alpha^2(n^2-4n+4+3n^2-4n)}{n^2(n-1)^2}\right)\sum_{i=1}^n X_{i,i}^2 \\
&= \left(1-\frac{4\alpha(n-1)}{n(n-1)}+\frac{4\alpha^2(n-1)^2}{n^2(n-1)^2}\right)\sum_{i=1}^n X_{i,i}^2 \\
&= \left(1-\frac{2\alpha}{n}\right)^2\sum_{i=1}^n X_{i,i}^2. \qquad (35)
\end{aligned}
$$

Combining (33) and (35) gives

$$
\|X-\alpha\bar{X}\|_F \leq \sqrt{\sum_{i\neq j}(1-\frac{2\alpha}{n(n-1)})^2 X_{i,j}^2+\left(1-\frac{2\alpha}{n}\right)^2\sum_{i=1}^n X_{i,i}^2}. \qquad (36)
$$

Note that $\bar{\alpha} := n-1$ satisfies $\bar{\alpha} \geq \frac{2n(n-1)}{3n-4}$ when $n \geq 4$, and the coefficients in (36) satisfy

$$
1-\frac{2\bar{\alpha}}{n(n-1)} = -(1-\frac{2\bar{\alpha}}{n}) = \frac{n-2}{n}.
$$

Since $\|X\|_F = 1$, by substituting $\bar{\alpha}$ into (36), we have

$$
\begin{aligned}
\|X-\bar{\alpha}\bar{X}\|_F \leq & \sqrt{\left(\frac{n-2}{n}\right)^2\sum_{i\neq j}X_{i,j}^2+\left(\frac{n-2}{n}\right)^2\sum_{i=1}^n X_{i,i}^2} \\
=&\frac{n-2}{n}\|X\|_F^2 \\
=&\frac{n-2}{n}.
\end{aligned}
$$

Because $\bar{X} \in \mathcal{S}_+^n$ and $\bar{\alpha} \geq 0$, by letting $U = \frac{n-2}{n}$, we have

$$
\|X-\mathrm{P}_{\mathcal{S}_+^n}(X)\|_F \leq \|X-\bar{\alpha}\bar{X}\|_F \leq U = \frac{n-2}{n}
$$

and hence,

$$
\overline{\mathrm{dist}}_F(\mathcal{DD}_n^*, \mathcal{S}_+^n) = \sup_{X\in\mathcal{DD}_n^*, \|X\|_F=1}\|X-\mathrm{P}_{\mathcal{S}_+^n}(X)\|_F \leq U = \frac{n-2}{n}. \qquad (37)
$$

(30), (31) and (37) imply that (29) holds, which proves this theorem.

□