

Radial-Type Error Bounds for Semidefinite Feasibility Problems without Strict Feasibility: Qualitative Estimates and Asymptotic Tightness

Mitsuhiro Nishijima¹

Guoyin Li²

Bruno F. Lourenço³

August 27, 2026

Abstract

In this paper, we develop a systematic framework for deriving explicit error bounds for semidefinite feasibility problems without assuming strict feasibility (Slater's condition), a setting in which existing results are limited. Our main technical contribution is the introduction of radial-type Hölder error bounds, where the error bound constant depends explicitly on the norm of the reference matrix through radial modulus functions. By combining facial reduction with recently developed facial residual functions, we obtain explicit descriptions of these modulus functions, yielding qualitative radial-type Hölder error bounds *without imposing any constraint qualifications*. Our results complement the classical work of Sturm by providing explicit estimates for the constants involved in Sturm's local Hölder error bounds over bounded sets with a given size. We further analyze the asymptotic behavior of these bounds as the dimension of the underlying matrix space grows, identifying regimes in which they can be asymptotically tight up to a dimension-free constant. As an application, we establish explicit error bounds for the optimality system of semidefinite programs by reformulating them as feasibility problems, a setting where Slater's condition typically fails. Under the generically satisfied strict complementarity condition, we derive radial-type error bounds without assuming the usual solution uniqueness requirement, and demonstrate their asymptotic tightness through an explicit example.

Key words Error bounds, Semidefinite feasibility problems, Facial residual functions, Semidefinite programming

Mathematics Subject Classification (2020) 15B48, 52A20, 90C22, 90C25

1 Introduction

The problem of finding an element of the intersection of an affine space $\mathcal{V}$ and a closed convex cone $\mathcal{K}$ in a finite-dimensional vector space $\mathcal{E}$ is called a *conic linear feasibility problem*. We write this problem as $\text{Feas}(\mathcal{V}, \mathcal{K})$ and, throughout this paper, we always assume that $\mathcal{V} \cap \mathcal{K} \neq \emptyset$, that is, $\text{Feas}(\mathcal{V}, \mathcal{K})$ admits a feasible solution. *Error bounds* provide us upper bounds on the distance between a point $x \in \mathcal{E}$ and the feasible set $\mathcal{V} \cap \mathcal{K}$ by using the distances between x and $\mathcal{V}$ as well as that between x and $\mathcal{K}$. Assuming that $\mathcal{E}$ is equipped with the norm $\|\cdot\|$ induced by an inner product, we define $\text{dist}(x, \mathcal{C}) := \min\{\|x - y\| \mid y \in \mathcal{C}\}$ as the distance between x and a closed convex set $\mathcal{C}$ in $\mathcal{E}$. Following [29, 48], the problem $\text{Feas}(\mathcal{V}, \mathcal{K})$ is said to admit a *(local) Hölder error bound* with exponent $\gamma \in (0, 1]$ if for any bounded set B (sometimes referred to as test sets), it follows that

$$\text{dist}(x, \mathcal{V} \cap \mathcal{K}) \leq \kappa_B \max\{\text{dist}(x, \mathcal{V}), \text{dist}(x, \mathcal{K})\}^\gamma \text{ for all } x \in B \quad (1.1)$$

for some nonnegative modulus κ_B .^{*1} If $\gamma = 1$, then (1.1) is called a *Lipschitz error bound*. Instead of restricting the set B to be bounded, if (1.1) holds with B being the whole space, then we say that a *global error bound* holds for the problem.

¹Department of Industrial and Systems Engineering, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama-shi, 2238522, Kanagawa, Japan. (nishijima@keio.jp).

²Department of Applied Mathematics, University of New South Wales, 2052, Sydney, Australia. (g.li@unsw.edu.au)

³Department of Fundamental Statistical Mathematics, The Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa-shi, 1908562, Tokyo, Japan. (bruno@ism.ac.jp).

^{*1}In some literature, this property is stated in an equivalent way. For example, in [16], the quantity $\max\{\text{dist}(x, \mathcal{V}), \text{dist}(x, \mathcal{K})\}^\gamma$ is replaced by $\text{dist}(x, \mathcal{V})^\gamma + \text{dist}(x, \mathcal{K})^\gamma$ and the pair $(\mathcal{V}, \mathcal{K})$ is referred to as being γ -Hölder regular.

In the definition of Hölder error bounds, the quantity $\text{dist}(x, \mathcal{V} \cap \mathcal{K})$ is referred to as the *forward error*, which measures the distance to the solution set of the feasibility problem, and is hard to estimate or compute directly, and hence is regarded as unknown. On the other hand, in numerical computation, what is often readily available to users is the quantity $\mathcal{E}_b(x) := \max\{\text{dist}(x, \mathcal{V}), \text{dist}(x, \mathcal{K})\}$ (or equivalently, $\text{dist}(x, \mathcal{V}) + \text{dist}(x, \mathcal{K})$), which is called the *backward error*. Thus, the associated error bounds provide us a good way to quantify the unknown forward error in terms of the computable backward error.

In the case where $\mathcal{K}$ is a polyhedral cone, the feasibility problem $\text{Feas}(\mathcal{V}, \mathcal{K})$ has a close connection with linear programming, and error bound results for $\text{Feas}(\mathcal{V}, \mathcal{K})$ have been studied extensively. For example, in this case, it is known that the problem $\text{Feas}(\mathcal{V}, \mathcal{K})$ admits a global Lipschitz error bound [21], and its modulus is often called a *Hoffman constant* [23, 41, 53]. The Hoffman constant is closely related to other important quantities such as the chi measure χ [52] and the distance to ill-posedness [40, 43].

We refer to a conic linear feasibility problem associated with a positive semidefinite cone as a *semidefinite feasibility problem*. The semidefinite feasibility problem and its associated error bounds have received considerable attention due to their natural connections with the stability and conditioning of the semidefinite programs (SDPs), as well as the powerful modeling capabilities of SDPs [51]. For example, for semidefinite feasibility problems satisfying the strict feasibility condition (often known as *Slater's condition*), they admit Lipschitz error bounds; see [5, Corollary 5] and [7, Lemma 3.1]. If a semidefinite feasibility problem satisfies a stronger version of Slater's condition, the problem admits a global Lipschitz error bound and its modulus can be interpreted geometrically [12, 22, 25, 26, 52].

While Slater's condition is a standard constraint qualification in convex optimization, verifying it can require nontrivial effort. More importantly, it may easily fail in several key settings, such as SDPs arising from relaxations of quadratic assignment problems and semidefinite feasibility problems describing the optimal solution sets of SDPs (see, for example, the survey [17] and the references therein).

In the absence of Slater's condition, [48, Theorem 3.3] established that semidefinite feasibility problems admit (local) Hölder error bounds. Notably, [48] showed that the Hölder exponent γ can be characterized entirely by the *singular degree*, which is determined by the number of steps required by the facial reduction algorithm [10]. This exponent also plays a crucial role in quantifying the convergence rate of alternating projection methods for ill-posed semidefinite feasibility problems [16]. More recently, Lourenço [31] and Lindstrom, Lourenço, and Pong [29] introduced *facial residual functions* and established error bounds for general conic linear feasibility problems, thereby extending Sturm's error bound approach to other classes of cones [27–30, 50].

However, these results are primarily existential and do not provide explicit or computable estimates of the associated error bound constants. To the best of our knowledge, the problem of systematically estimating the constant κ_B in Hölder error bounds for semidefinite feasibility problems, particularly in the absence of constraint qualifications such as Slater's condition, remains largely open. Given the success of explicit estimates for Hoffman constants in analyzing linear convergence rates and studying sensitivity in polyhedral convex programs [2, 6, 24, 34, 49], analogous explicit estimates of error bound constants could similarly facilitate more informative convergence analysis and perturbation bounds for numerical algorithms solving semidefinite feasibility problems or SDPs. This naturally leads to the following research question: *In the absence of constraint qualifications such as Slater's condition, can one obtain effective explicit estimates of the error bound constants in Hölder error bounds for semidefinite feasibility problems? If so, how tight are these estimates, particularly as the dimension of the underlying matrix space increases?*

Let $\mathcal{S}_+^n$ denote the cone of positive semidefinite matrices in the space $\mathcal{S}^n$ of real $n \times n$ symmetric matrices. The purpose of this paper is to address the above questions by studying the following *radial-type Hölder error bounds* with exponent $\gamma \in (0, 1]$:

$$\text{dist}(\mathbf{X}, \mathcal{V} \cap \mathcal{S}_+^n) \leq \mu(\|\mathbf{X}\|_F) \mathcal{E}_b(\mathbf{X})^\gamma \text{ for all } \mathbf{X} \in \mathcal{Q},$$

or, more generally, with the form

$$\text{dist}(\mathbf{X}, \mathcal{V} \cap \mathcal{S}_+^n) \leq \sum_{j=0}^{\ell} \mu_j(\|\mathbf{X}\|_F) \mathcal{E}_b(\mathbf{X})^{\gamma_j} \text{ for all } \mathbf{X} \in \mathcal{Q}. \quad (1.2)$$

Here, $\|\cdot\|_F$ denotes the Frobenius norm, $\mathcal{Q}$ is a given subset of $\mathcal{S}^n$, ℓ is a nonnegative integer, γ_j is a positive constant with $\gamma_j \leq \gamma$ for every $j = 0, \dots, \ell$, and μ_j , $j = 0, \dots, \ell$, are nonnegative nondecreasing functions, which we call *radial modulus functions*. In particular, our main aim is to derive radial-type Hölder error bounds with explicit exponents $\gamma_0, \dots, \gamma_\ell$ and explicit form of the radial modulus functions $\mu_0, \dots, \mu_\ell$ in the absence of constraint qualifications.

It is worth emphasizing that the notion of a radial-type error bound unifies both local and global error bounds. In particular, when $\ell = 0$, $\gamma_0 = \gamma$, $\mathcal{Q} = \mathcal{S}^n$, and μ_0 is a constant function, the above condition reduces to a global error bound. On the other hand, since each radial modulus function μ_j is nondecreasing, the bound immediately yields a local error bound with an explicit constant (depending on ρ) over any bounded test set contained in the ball of radius ρ . We also mention that we do not restrict ourselves to the case where $\mathcal{Q}$ is bounded. In fact, our main focus and result are on the cases where the test set $\mathcal{Q}$ is the ambient space $\mathcal{S}^n$. In line with existing literature, we also consider the cases where $\mathcal{Q} = \{\mathbf{X} \in \mathcal{S}^n \mid \mathcal{E}_b(\mathbf{X}) \leq \lambda\}$ where λ is a given number.

The main contributions of this paper are as follows.

- (1) Firstly, we provide explicit computable forms for the radial modulus functions $\mu_0, \dots, \mu_\ell$. We achieved these by exploiting the facial reduction process together with recently developed facial residual functions [29]. This approach yields new qualitative radial-type Hölder error bounds in the form of (1.2) where the test set is $\mathcal{Q} = \mathcal{S}^n$, without assuming any constraint qualifications (Theorems 3.17 and 3.19). As a consequence, in line with the existing literature, we derive radial-type Hölder error bounds in the form of (1.2) where the test set is $\mathcal{Q} = \{\mathbf{X} \in \mathcal{S}^n \mid \mathcal{E}_b(\mathbf{X}) \leq 1\}$ (Corollary 3.20). This complements the classical work of Sturm [48] by providing explicit estimates for the constants involved in the Hölder error bounds over bounded sets with given sizes.

Interestingly, the derived radial modulus functions depend explicitly on a quantity d^* called the *distance to partial polyhedral Slater's (PPS) condition*, which is a variant of the concept of singularity degree in the literature and relates to the number of steps of a certain facial reduction algorithm.

- (2) Secondly, we investigate the asymptotic tightness of the derived error bounds. To formalize this notion, we introduce in Definition 2.4 the concept of being *asymptotically tight up to a dimension-free constant*. Consider a family of semidefinite feasibility problems in which the dimension of the underlying matrix space varies. Roughly speaking, a derived error bound is asymptotically tight up to a dimension-free constant $C \in (0, 1]$ if it is essentially optimal up to the multiplicative factor C . This means that one cannot obtain a strictly sharper error bound by uniformly scaling the radial modulus functions by any dimension-independent constant factor strictly smaller than C .

Our investigation of asymptotic tightness is based on the values of d^* , the distance to the PPS condition of the feasibility problem. We first focus on instances where d^* are 0 and 1, respectively, and show in Sections 4.1 and 4.2 that the corresponding error bounds can be asymptotically tight up to a dimension-free constant. On the other hand, for the semidefinite feasibility problem presented in [48, Example 2] (which satisfies $d^* = n - 1$), our explicit error bound is not asymptotically tight up to a dimension-free constant.

- (3) Thirdly, as an application, we derive explicit error bounds for the optimality system of SDPs by reformulating it as a semidefinite feasibility problem, a setting in which Slater's condition typically fails. Without assuming solution uniqueness that the existing literature requires, see, for example, [14, Section 1] and [15, Theorem 4.1] (and also the related reference [33]), we establish explicit radial-type error bounds under the strict complementarity condition, which holds generically [1, 38]. Moreover, we show, through an explicit example, that the derived bound can be asymptotically tight in this setting.

The organization of this paper is as follows. In Section 2, we recall and introduce notation and concepts used in the subsequent sections. In Section 3, we derive qualitative radial-type Hölder error bounds for semidefinite feasibility problems. In Section 4, we discuss the asymptotic tightness of the error bound presented in the previous section. In Section 5, as an application of the result of Section 3, we obtain explicit error bounds for the optimality system of SDPs satisfying the strict complementarity condition. In Section 6, we finish with concluding remarks.

2 Preliminaries

In this section, we recall some preliminaries and basic definitions as well as existing tools that will be used in the next few sections.

2.1 Notation

For an integer n , we write $\mathbb{Z}_{\geq n}$ for the set of integers greater than or equal to n . Let V be a finite-dimensional Euclidean space and denote the corresponding norm by $\|\cdot\|$. For $S \subseteq V$, we use $\text{int } S$, $\text{ri } S$, $\text{span } S$, and $S^\perp$ to denote

the interior of S , the relative interior of S , the smallest subspace containing S , and the orthogonal complement of S , respectively. For a closed convex set S in V and $x \in V$, we define $\text{dist}(x, S) := \min\{\|x - y\| \mid y \in S\}$ and the projection of x onto S as $P_S(x) := \arg \min\{\|x - y\| \mid y \in S\}$. For $c \in V$ and $r \geq 0$, the closed ball with center c and radius r is denoted by $\mathbb{B}(c, r) := \{x \in V \mid \|x - c\| \leq r\}$. For a linear mapping $\mathcal{A}$, we use $\mathcal{A}^*$ to denote the adjoint of $\mathcal{A}$.

Vectors are denoted by boldface lowercase letters such as $\mathbf{a}$. The space of n -dimensional real vectors is denoted by $\mathbb{R}^n$ and the set of entrywise nonnegative vectors in $\mathbb{R}^n$ is denoted by $\mathbb{R}_+^n$. For a vector $\mathbf{a} \in \mathbb{R}^n$, we write $\mathbf{a}^\top$ for the transpose of $\mathbf{a}$ and we use $\|\mathbf{a}\|_2 := \sqrt{\mathbf{a}^\top \mathbf{a}}$ to denote the 2-norm of $\mathbf{a}$.

Matrices are denoted by boldface uppercase letters such as $\mathbf{A}$. The space of real $m \times n$ matrices is denoted by $\mathbb{R}^{m \times n}$ and the space of real $n \times n$ symmetric matrices is denoted by $\mathcal{S}^n$. For $\mathbf{A} \in \mathcal{S}^n$, we define $\mathbb{R}\mathbf{A} := \{t\mathbf{A} \mid t \in \mathbb{R}\}$ and $\mathbb{R}_+\mathbf{A} := \{t\mathbf{A} \mid t \in \mathbb{R}_+\}$, respectively. We write the (i, j) th element of a matrix $\mathbf{A}$ as $A_{i,j}$. When there is no ambiguity we will abbreviate $A_{i,j}$ as A_{ij} . The zero matrix is denoted by $\mathbf{O}$ and the identity matrix of order n is denoted by $\mathbf{I}_n$. We use $\mathbf{E}_{i,j}$ or $\mathbf{E}_{ij}$ to denote the matrix whose (i, j) th element is 1 and all other elements are 0. Its order is determined from the context. For $a_1, \dots, a_n \in \mathbb{R}$, we use $\text{Diag}(a_1, \dots, a_n)$ to denote the diagonal matrix whose (i, i) th element is a_i for each $i = 1, \dots, n$. Similarly, for real symmetric matrices $\mathbf{A}_1, \dots, \mathbf{A}_n$, the block diagonal matrix whose i th block is $\mathbf{A}_i$ for each $i = 1, \dots, n$ is denoted by $\text{Diag}(\mathbf{A}_1, \dots, \mathbf{A}_n)$. For a subset $\mathcal{C}$ of $\mathcal{S}^m$, we define $\mathcal{C} \oplus \{0\}^n := \{\text{Diag}(\mathbf{A}, \mathbf{O}) \in \mathcal{S}^{m+n} \mid \mathbf{A} \in \mathcal{C}\}$. The set $\{0\}^n \oplus \mathcal{C}$ is also defined analogously. For a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, we use $\mathbf{A}^\top$ to denote the transpose of $\mathbf{A}$ and write $\text{tr}(\mathbf{A})$ for the trace of $\mathbf{A}$. For matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$, we define $\langle \mathbf{A}, \mathbf{B} \rangle := \text{tr}(\mathbf{B}^\top \mathbf{A})$, which is an inner product on $\mathbb{R}^{m \times n}$. The Frobenius norm, which is the norm induced by this inner product, is denoted by $\|\cdot\|_F$. For a matrix $\mathbf{A} \in \mathcal{S}^n$, we use $\lambda_{\max}(\mathbf{A})$ and $\lambda_{\min}(\mathbf{A})$ to denote the maximum and minimum eigenvalues of $\mathbf{A}$, respectively. In addition, $\lambda_{\min}^+(\mathbf{A})$ denotes the smallest positive eigenvalue of $\mathbf{A}$ if $\mathbf{A}$ has at least one positive eigenvalue. Otherwise, we let $\lambda_{\min}^+(\mathbf{A}) := +\infty$. Furthermore, we use the convention that $\frac{a}{+\infty} = 0$ for every $a \in \mathbb{R}$.

2.2 Positive semidefinite cones and their faces

Similar to [29, 31], our development heavily relies on the facial structure of the underlying cone. Here we present a brief review of positive semidefinite cones and their faces.

Recall that $\mathcal{S}_+^n$ denotes the cone of positive semidefinite matrices in $\mathcal{S}^n$, which we refer to as the positive semidefinite cone. The formula for the distance from a point $\mathbf{X} \in \mathcal{S}^n$ to the cone $\mathcal{S}_+^n$ is well known [11, page 399] and is given by

$$\text{dist}(\mathbf{X}, \mathcal{S}_+^n) = \sqrt{\sum_{i=1}^n \max\{-\lambda_i, 0\}^2}, \quad (2.1)$$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of $\mathbf{X}$. Moreover, it is easy to verify that, for any matrix $\mathbf{X} \in \mathcal{S}^n$ and any principal submatrix $\widehat{\mathbf{X}} \in \mathcal{S}^d$, the following inequality holds:

$$\text{dist}(\widehat{\mathbf{X}}, \mathcal{S}_+^d) \leq \text{dist}(\mathbf{X}, \mathcal{S}_+^n). \quad (2.2)$$

We recall that a nonempty (closed) convex subcone $\mathcal{F}$ of $\mathcal{S}_+^n$ is called a *face* if for all $\mathbf{A}, \mathbf{B} \in \mathcal{S}_+^n$, we have $\mathbf{A}, \mathbf{B} \in \mathcal{F}$ whenever $\mathbf{A} + \mathbf{B} \in \mathcal{F}$ holds. For a face $\mathcal{F}$ of $\mathcal{S}_+^n$, we define $\mathcal{F}^*$ as the *dual cone* of $\mathcal{F}$, i.e., the set of $\mathbf{A} \in \mathcal{S}^n$ such that $\langle \mathbf{A}, \mathbf{B} \rangle \geq 0$ for all $\mathbf{B} \in \mathcal{F}$. A *chain of faces* of $\mathcal{S}_+^n$ is a sequence of faces $\mathcal{F}_1, \dots, \mathcal{F}_\ell$ of $\mathcal{S}_+^n$, denoted by $\mathcal{F}_\ell \subsetneq \dots \subsetneq \mathcal{F}_1$, such that the strict inclusion $\mathcal{F}_{i+1} \subsetneq \mathcal{F}_i$ holds for every $i = 1, \dots, \ell - 1$.

A set of the form $\mathcal{S}_+^n \cap \{\mathbf{Z}\}^\perp$ for some $\mathbf{Z} \in \mathcal{S}_+^n$ is a face of the cone $\mathcal{S}_+^n$, and we say that the face is *exposed* by the matrix $\mathbf{Z}$. Every face of the positive semidefinite cone is linearly isomorphic to a positive semidefinite cone of smaller order; see [3, Lemma 4], [20, Theorem 3.6], and [35, Example 3.2.2]. In particular, for a face $\mathcal{F}$ of $\mathcal{S}_+^n$, there exist an integer d with $0 \leq d \leq n$ and an orthogonal matrix $\mathbf{P}$ of order n such that

$$\mathcal{F} = \mathbf{P}(\mathcal{S}_+^d \oplus \{0\}^{n-d})\mathbf{P}^\top.$$

When $d \geq 1$, by dropping the last $(n - d)$ columns of $\mathbf{P}$, we can also represent the face $\mathcal{F}$ as $\mathbf{Q}\mathcal{S}_+^d\mathbf{Q}^\top$, where $\mathbf{Q} \in \mathbb{R}^{n \times d}$ is a matrix whose column vectors are orthonormal. When $d = 0$, the face $\mathcal{F}$ is the singleton $\{\mathbf{O}\}$.

Recall that $\text{Feas}(\mathcal{V}, \mathcal{S}_+^n)$ denotes the problem of finding an element of the intersection of an affine space $\mathcal{V}$ in $\mathcal{S}^n$ and the positive semidefinite cone $\mathcal{S}_+^n$. The following result, which was established in [31, Proposition 5], describes a chain of faces associated with a feasible semidefinite feasibility problem.

Theorem 2.1. *For any feasible semidefinite feasibility problem $\text{Feas}(\mathcal{V}, \mathcal{S}_+^n)$, there exists a chain of faces of $\mathcal{S}_+^n$, denoted by*

$$\mathcal{F}_\ell \subsetneq \cdots \subsetneq \mathcal{F}_1 = \mathcal{S}_+^n, \quad (2.3)$$

together with matrices $\mathbf{Z}_1, \dots, \mathbf{Z}_{\ell-1} \in \mathcal{S}^n$ such that the following conditions hold:

- (i) $\mathbf{Z}_i \in \mathcal{F}_i^* \cap \mathcal{V}^\perp$ and $\mathcal{F}_{i+1} = \mathcal{F}_i \cap \{\mathbf{Z}_i\}^\perp$ hold for each $i = 1, \dots, \ell - 1$.
- (ii) The face $\mathcal{F}_\ell$ is polyhedral or the set $\mathcal{V} \cap \text{ri } \mathcal{F}_\ell$ is nonempty.

Moreover, any such chain necessarily satisfies $\ell \leq n$.

The inequality $\ell \leq n$ given in Theorem 2.1 follows from [32, Example 1]. For the chain of faces in (2.3), it follows from (i) of Theorem 2.1 that $\mathcal{V} \cap \mathcal{S}_+^n = \mathcal{V} \cap \mathcal{F}_\ell$. This implies that, after the facial reduction steps, we can find a face $\mathcal{F}_\ell$ such that the constraint qualification shown in (ii) of Theorem 2.1 is satisfied and the set of feasible solutions is invariant. The condition in (ii) is called the *partial polyhedral Slater's (PPS) condition* in [29, 31]. The *distance to the PPS condition* of the feasibility problem $\text{Feas}(\mathcal{V}, \mathcal{S}_+^n)$, denoted by $d_{\text{PPS}}(\mathcal{V}, \mathcal{S}_+^n)$, is defined as $\ell - 1$, where ℓ is the length of a shortest chain of the form (2.3) satisfying the conditions in Theorem 2.1. By Theorem 2.1, the quantity $d_{\text{PPS}}(\mathcal{V}, \mathcal{S}_+^n)$ satisfies $d_{\text{PPS}}(\mathcal{V}, \mathcal{S}_+^n) \leq n - 1$. In addition, it is a variant of the *singularity degree* of the problem $\text{Feas}(\mathcal{V}, \mathcal{S}_+^n)$ used in literature. More precisely, the singularity degree of the problem $\text{Feas}(\mathcal{V}, \mathcal{S}_+^n)$, denoted by d_s , is defined as $\bar{\ell} - 1$, where $\bar{\ell}$ is the length of a shortest chain in (2.3) satisfying the condition in (i) and the condition that the set $\mathcal{V} \cap \text{ri } \mathcal{F}_\ell$ is nonempty. From the definitions, it is clear that $d_{\text{PPS}}(\mathcal{V}, \mathcal{S}_+^n) \leq d_s$ always holds. Indeed, for semidefinite feasibility problems with $n \geq 2$, the equation

$$d_{\text{PPS}}(\mathcal{V}, \mathcal{S}_+^n) = d_s \quad (2.4)$$

holds as shown in Appendix A. Using d_{PPS} is typically more advantageous when the underlying cone is a direct product. Nevertheless, we use d_{PPS} here in order to maintain consistency with [27, 31]. For more details on the distance to the PPS condition and the singularity degree, see [31, Section 2.4.1] and references therein.

For each $i = 1, \dots, \ell - 1$, the matrix $\mathbf{Z}_i$ shown above can be taken to have norm 1. By the strict inclusion $\mathcal{F}_{i+1} \subsetneq \mathcal{F}_i$, we have $\mathbf{Z}_i \neq \mathbf{O}$, i.e., $\|\mathbf{Z}_i\|_F \neq 0$. Then it follows from the condition in (i) of Theorem 2.1 that $\mathbf{Z}_i / \|\mathbf{Z}_i\|_F \in \mathcal{F}_i^* \cap \mathcal{V}^\perp$ and $\mathcal{F}_{i+1} = \mathcal{F}_i \cap \{\mathbf{Z}_i / \|\mathbf{Z}_i\|_F\}^\perp$ hold. Therefore, by normalizing if necessary, we may assume that $\|\mathbf{Z}_i\|_F = 1$ without loss of generality.

Let $\mathcal{F}$ be a face of $\mathcal{S}_+^n$ and let $\mathbf{Z} \in \mathcal{F}^*$. A *one-step facial residual function* for $\mathcal{F}$ and $\mathbf{Z}$ is a function $\psi_{\mathcal{F}, \mathbf{Z}}: \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ satisfying the following three conditions:

- (i) $\psi_{\mathcal{F}, \mathbf{Z}}$ is nondecreasing in each argument.
- (ii) $\psi_{\mathcal{F}, \mathbf{Z}}(0, t) = 0$ for any $t \in \mathbb{R}_+$.
- (iii) For any $\mathbf{X} \in \text{span } \mathcal{F}$ and any $\epsilon \in \mathbb{R}_+$, if $\text{dist}(\mathbf{X}, \mathcal{F}) \leq \epsilon$ and $\langle \mathbf{Z}, \mathbf{X} \rangle \leq \epsilon$, then we have

$$\text{dist}(\mathbf{X}, \mathcal{F} \cap \{\mathbf{Z}\}^\perp) \leq \psi_{\mathcal{F}, \mathbf{Z}}(\epsilon, \|\mathbf{X}\|_F).$$

As mentioned in [29, Example 3.6], there exist nonnegative α and β , which depend on $\mathcal{F}$ and $\mathbf{Z}$, such that

$$\psi_{\mathcal{F}, \mathbf{Z}}(s, t) = \alpha s + \beta \sqrt{st} \quad (2.5)$$

is a one-step facial residual function for $\mathcal{F}$ and $\mathbf{Z}$. In Section 3.2, we will compute the constants α and β explicitly, and discuss how these constants relate to the choice of $\mathbf{Z}$. Indeed, the one-step facial residual function introduced in [29] is of the form $\alpha s + \alpha \sqrt{st}$ for some positive α , corresponding to the case $\alpha = \beta$. Here, we extend this definition by allowing the two constants to be independent, which enables us to derive sharper estimates later.

2.3 Error bounds

In this subsection, we review the results of Lipschitz error bounds for polyhedra, and introduce the notion of asymptotic tightness of error bounds. Without loss of generality, we only consider the case that polyhedra are subsets of $\mathbb{R}^n$. The celebrated Hoffman's error bound result [21] states that for a given matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ whose

i th row is denoted by $\mathbf{a}_i^\top$, there exists a nonnegative constant α such that for any $\mathbf{b} = (b_1, \dots, b_m)^\top \in \mathbf{A}\mathbb{R}^n + \mathbb{R}_+^m$ (namely $\mathbf{b}$ such that the linear inequality system $\mathbf{A}\mathbf{y} \leq \mathbf{b}$ has a solution), it follows that

$$\text{dist}(\mathbf{x}, \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{y} \leq \mathbf{b}\}) \leq \alpha \sqrt{\sum_{i=1}^m \max\{\mathbf{a}_i^\top \mathbf{x} - b_i, 0\}^2} \text{ for all } \mathbf{x} \in \mathbb{R}^n. \quad (2.6)$$

We call the smallest nonnegative constant α for which the error bound in (2.6) holds the *Hoffman constant* and write it as $H(\mathbf{A})$.

Let $\mathbf{A}_J$ be the submatrix obtained by extracting the rows of $\mathbf{A}$ indexed by a subset J . In addition, we use $\mathbb{R}^J$ and $\mathbb{R}_+^J$ to denote the space of real vectors whose entries are indexed by the elements in J and the set of nonnegative vectors in $\mathbb{R}^J$, respectively. It is shown in [40, Proposition 2] that if we set

$$H(\mathbf{A}) = \max_{\substack{J \subseteq \{1, \dots, m\}, \\ \mathbf{A}_J \mathbb{R}^n + \mathbb{R}_+^J = \mathbb{R}^J}} \frac{1}{\min_{\substack{\mathbf{v} \in \mathbb{R}_+^J, \\ \|\mathbf{v}\|_2 = 1}} \|(\mathbf{A}_J)^\top \mathbf{v}\|_2}, \quad (2.7)$$

with the convention that the fraction inside the maximum is zero when $J = \emptyset$, then the constant $H(\mathbf{A})$ satisfies the error bound in (2.6) and it is indeed the smallest constant for which the error bound in (2.6) holds. The computation of the Hoffman constant shown in (2.7) is also discussed in [40, Section 3]. Although various explicit representations of a Hoffman constant, for example, (2.7) and those provided in [19, 23], are known, calculating the exact value of the Hoffman constant typically involves significant computational difficulty as discussed in [41]. Meanwhile, some studies discuss algorithmic aspects for calculating them, and effective upper bounds can be found in [23, 40–42].

For linear equality systems, effective upper bounds for Hoffman constants can be computed via the singular values of the associated coefficient matrix. For $\mathbf{A} \in \mathbb{R}^{m \times n}$, we define $\sigma_{\min}^+(\mathbf{A})$ as the smallest positive singular value of $\mathbf{A}$ when $\mathbf{A}$ is nonzero, and set $\sigma_{\min}^+(\mathbf{A}) := +\infty$ when $\mathbf{A}$ is the zero matrix. Then for any $\mathbf{b} \in \mathbb{R}^m$ such that the linear equality system $\mathbf{A}\mathbf{y} = \mathbf{b}$ has a solution, it follows that

$$\text{dist}(\mathbf{x}, \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{y} = \mathbf{b}\}) \leq \frac{1}{\sigma_{\min}^+(\mathbf{A})} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2 \text{ for all } \mathbf{x} \in \mathbb{R}^n. \quad (2.8)$$

This inequality follows from the result on the best approximate solution of a linear equality system [44] and the description of the singular values of a generalized inverse; see also [34, Section 4].

It is well known that a global Lipschitz error bound for the intersection of polyhedral sets always holds, e.g., see [4, Corollary 5.26]. Let C_1 and C_2 be polyhedral sets in $\mathbb{R}^n$ and suppose that $C_1 \cap C_2 \neq \emptyset$. Then there exists a nonnegative constant α such that

$$\text{dist}(\mathbf{x}, C_1 \cap C_2) \leq \alpha \max\{\text{dist}(\mathbf{x}, C_1), \text{dist}(\mathbf{x}, C_2)\} \text{ for all } \mathbf{x} \in \mathbb{R}^n. \quad (2.9)$$

We write $\eta(C_1, C_2)$ for the smallest nonnegative constant α for which the error bound in (2.9) holds. Although computing $\eta(C_1, C_2)$ is also computationally challenging, we can obtain computable upper bounds on $\eta(C_1, C_2)$ by exploiting algebraic descriptions of C_1 and C_2 . We provide a proof of the following lemma in Appendix B.

Lemma 2.2. *Let C_1 and C_2 be polyhedral sets in $\mathbb{R}^n$ with $C_1 \cap C_2 \neq \emptyset$. Let $\eta(C_1, C_2)$ be the smallest nonnegative constant α for which the error bound in (2.9) holds.*

- (i) *Suppose that for each $i = 1, 2$, the polyhedral set C_i is described by a linear inequality system $\mathbf{A}_i \mathbf{x} \leq \mathbf{b}_i$ for some $\mathbf{A}_i \in \mathbb{R}^{m_i \times n}$ and $\mathbf{b}_i \in \mathbb{R}^{m_i}$. Let $\mathbf{A}$ be the matrix obtained by concatenating the matrix $\mathbf{A}_1$ with the matrix $\mathbf{A}_2$. Then we have $\eta(C_1, C_2) \leq H(\mathbf{A}) \|\mathbf{A}\|_F$.*
- (ii) *Suppose that for each $i = 1, 2$, the polyhedral set C_i is described by a linear equality system $\mathbf{A}_i \mathbf{x} = \mathbf{b}_i$ for some $\mathbf{A}_i \in \mathbb{R}^{m_i \times n}$ and $\mathbf{b}_i \in \mathbb{R}^{m_i}$. Let $\mathbf{A}$ be the matrix constructed in a way as in (i). Then we have $\eta(C_1, C_2) \leq \|\mathbf{A}\|_F / \sigma_{\min}^+(\mathbf{A})$.*

Remark 2.3. We note that $\eta(C_1, C_2) \geq 1$ unless $C_1 \cap C_2 = \mathbb{R}^n$, i.e., $C_1 = C_2 = \mathbb{R}^n$. Moreover, we also note that $\eta(C_1, C_2) = 1$ if one of the two sets C_1 and C_2 is included in the other set. It follows that $\eta(C_1, C_2) = 0$ if $C_1 \cap C_2 = \mathbb{R}^n$.

The error bounds studied in this paper apply to semidefinite feasibility problems of arbitrary matrix order n . Consequently, the notion of tightness we adopt below is inherently asymptotic, involving sequences of problems defined over matrix spaces of increasing dimension. Our objective is to characterize the behavior of the associated error bound inequalities as the matrix order n tends to infinity.

More precisely, let $\mathcal{V}_n$ be an affine space in $\mathcal{S}^n$. We consider a family of semidefinite feasibility problems $\text{Feas}(\mathcal{V}_n, \mathcal{S}_+^n)$ in which the order of the matrix (and so, the dimension of the underlying matrix space) varies. The corresponding backward error is $\mathcal{E}_b^{(n)}(\mathbf{X}) := \max\{\text{dist}(\mathbf{X}, \mathcal{V}_n), \text{dist}(\mathbf{X}, \mathcal{S}_+^n)\}$. For each $n \in \mathbb{Z}_{\geq 1}$, define the residual function g_n by

$$g_n(\mathbf{X}) := \sum_{j=0}^{\ell_n} \mu_{j,n}(\|\mathbf{X}\|_F) \mathcal{E}_b^{(n)}(\mathbf{X})^{\gamma_{j,n}},$$

where $\ell_n \in \mathbb{Z}_{\geq 0}$, $\mu_{j,n}: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is a nondecreasing function for each $j = 0, \dots, \ell_n$, and $\gamma_{j,n} > 0$. Suppose that there exist a positive integer n_0 and a set $\mathcal{Q}_n$ in $\mathcal{S}^n$ parametrized by every $n \in \mathbb{N}_{\geq n_0}$ such that the following error bounds hold:

$$\text{dist}(\mathbf{X}, \mathcal{V}_n \cap \mathcal{S}_+^n) \leq g_n(\mathbf{X}) \text{ for all } n \in \mathbb{N}_{\geq n_0} \text{ and for all } \mathbf{X} \in \mathcal{Q}_n. \quad (2.10)$$

Under this assumption, g_n is a nonnegative function such that $g_n(\mathbf{X}) = 0$ if and only if $\mathbf{X} \in \mathcal{V}_n \cap \mathcal{S}_+^n$. As we will see later in Corollary 3.20, error bounds as in (2.10) always hold for semidefinite feasibility problems. With that, we have the following definition of tightness.

Definition 2.4 (Asymptotic tightness for error bounds). We say that the error bound in (2.10) is *asymptotically tight up to a dimension-free constant* $C_0 \in (0, 1]$ if there exist a sequence (n_k) in $\mathbb{Z}_{\geq n_0}$ satisfying $\lim_{k \rightarrow \infty} n_k = \infty$ and a sequence $(\mathbf{X}_k)$ with $\mathbf{X}_k \in \mathcal{Q}_{n_k} \setminus (\mathcal{V}_{n_k} \cap \mathcal{S}_+^{n_k})$ for every k such that

$$\liminf_{k \rightarrow \infty} \frac{\text{dist}(\mathbf{X}_k, \mathcal{V}_{n_k} \cap \mathcal{S}_+^{n_k})}{g_{n_k}(\mathbf{X}_k)} \geq C_0.$$

Moreover, we say that the error bound in (2.10) is *asymptotically tight up to a dimension-free constant* if there exists a constant $C_0 \in (0, 1]$, independent of the dimension, such that the bound is asymptotically tight with constant C_0 .

Remark 2.5 (Intuition behind Definition 2.4). If the error bound in (2.10) is asymptotically tight up to a dimension-free constant $C_0 \in (0, 1]$, this then implies that for any constant $C \in (0, C_0)$, the following inequality cannot be true:

$$\text{dist}(\mathbf{X}, \mathcal{V}_n \cap \mathcal{S}_+^n) \leq C g_n(\mathbf{X}) \text{ for all } n \in \mathbb{Z}_{\geq n_0} \text{ and for all } \mathbf{X} \in \mathcal{Q}_n.$$

In other words, in this case, for any $C \in (0, C_0)$, an improved version of error bound results with the residual function g_n scaling with a constant contracting factor C is not possible.

3 Radial-type error bounds for semidefinite feasibility problems

In this section, we derive a qualitative radial-type Hölder error bound for the semidefinite feasibility problem $\text{Feas}(\mathcal{V}, \mathcal{S}_+^n)$ *without assuming any constraint qualifications*. Our qualitative explicit estimate relies on the facial reduction algorithm. For simplicity, we write d^* for $d_{\text{PPS}}(\mathcal{V}, \mathcal{S}_+^n)$, the distance to the PPS condition of the feasibility problem $\text{Feas}(\mathcal{V}, \mathcal{S}_+^n)$. Then, from Theorem 2.1 there exists a chain

$$\mathcal{F}_{d^*+1} \subsetneq \dots \subsetneq \mathcal{F}_1 \quad (3.1)$$

of faces of $\mathcal{S}_+^n$, such that $\mathcal{F}_1 = \mathcal{S}_+^n$, $\mathbf{Z}_i \in \mathcal{F}_i^* \cap \mathcal{V}^\perp$ and $\mathcal{F}_{i+1} = \mathcal{F}_i \cap \{\mathbf{Z}_i\}^\perp$ for each $i = 1, \dots, d^*$, and the problem $\text{Feas}(\mathcal{V}, \mathcal{F}_{d^*+1})$ satisfies the PPS condition. As mentioned in Section 2.3, we may assume that $\|\mathbf{Z}_i\|_F = 1$ for all $i = 1, \dots, d^*$. Throughout this section, we assume that the $\mathcal{F}_i$'s and $\mathbf{Z}_i$'s are given in (3.1), and we shall fix them and utilize them to derive the desired radial-type error bound.

The underlying ideas of the derivation are intuitive, though the argument is somewhat technical. To streamline the exposition, we first present a roadmap of the proof. *To derive a qualitative radial-type Hölder error bound, we first consider the facially reduced problem with minimal face, $\text{Feas}(\mathcal{V}, \mathcal{F}_{d^*+1})$, and compute an explicit modulus of a Lipschitz error bound for this problem in Section 3.1. Then we compute the constants associated to a one-step facial residual function for $\mathcal{F}_i$ and $\mathbf{Z}_i$ when one moves from a lower layer face $\mathcal{F}_i$ to the higher layer face $\mathcal{F}_{i+1}$ for every $i = 1, \dots, d^*$ in Section 3.2. Combining these results, we then obtain the final qualitative radial-type Hölder error bound for the problem $\text{Feas}(\mathcal{V}, \mathcal{S}_+^n)$ in Section 3.3.*

3.1 Radial modulus function for facially reduced problem with minimal face

In this subsection, we consider the problem $\text{Feas}(\mathcal{V}, \mathcal{F}_{d^*+1})$, the feasibility problem with respect to the subspace $\mathcal{V}$ and the minimal face $\mathcal{F}_{d^*+1}$. Our main aim in this subsection is to establish an explicit form of a radial modulus function κ (see Theorem 3.6 later) such that

$$\text{dist}(\mathbf{X}, \mathcal{V} \cap \mathcal{F}_{d^*+1}) \leq \kappa(\|\mathbf{X}\|_F) \max\{\text{dist}(\mathbf{X}, \mathcal{V}), \text{dist}(\mathbf{X}, \mathcal{F}_{d^*+1})\} \text{ for all } \mathbf{X} \in \mathcal{S}^n. \quad (3.2)$$

Recall that the problem $\text{Feas}(\mathcal{V}, \mathcal{F}_{d^*+1})$ satisfies the PPS condition, so $\mathcal{V} \cap \text{ri} \mathcal{F}_{d^*+1} \neq \emptyset$ holds or $\mathcal{F}_{d^*+1}$ is polyhedral. First, we consider the case where $\mathcal{V} \cap \text{ri} \mathcal{F}_{d^*+1} \neq \emptyset$. Direct verification shows that the following equations hold (this can also be seen by the proof of [7, Lemma 3.2]):

$$(\text{span } \mathcal{F}_{d^*+1}) \cap (\mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp) = \mathcal{F}_{d^*+1}, \quad (3.3)$$

$$(\text{span } \mathcal{F}_{d^*+1}) \cap \text{int}(\mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp) = \text{ri } \mathcal{F}_{d^*+1}. \quad (3.4)$$

Lemma 3.1. *Let $\mathbf{U} \in \text{int}(\mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp)$. Then we have*

$$\lambda_{\min}^+(P_{\mathcal{F}_{d^*+1}}(\mathbf{U})) = \max\{\epsilon \mid \mathbb{B}(\mathbf{U}, \epsilon) \subseteq \mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp\} \in (0, +\infty]. \quad (3.5)$$

Proof. When $\mathcal{F}_{d^*+1} = \{\mathbf{O}\}$, both the left- and right-hand sides of (3.5) are $+\infty$, and the desired equality holds. In what follows, we assume that $\mathcal{F}_{d^*+1} \neq \{\mathbf{O}\}$. Since $\mathcal{F}_{d^*+1}$ is a nonzero face of $\mathcal{S}_+^n$, there exist an orthogonal matrix $\mathbf{P}$ of order n and a positive integer r such that

$$\mathcal{F}_{d^*+1} = \mathbf{P}(\mathcal{S}_+^r \oplus \{0\}^{n-r})\mathbf{P}^\top. \quad (3.6)$$

Then we see that

$$\begin{aligned} \mathcal{F}_{d^*+1}^\perp &= \mathbf{P} \left\{ \begin{pmatrix} \mathbf{O} & \mathbf{A}_{[12]} \\ \mathbf{A}_{[12]}^\top & \mathbf{A}_{[22]} \end{pmatrix} \mid \mathbf{A}_{[12]} \in \mathbb{R}^{r \times (n-r)}, \mathbf{A}_{[22]} \in \mathcal{S}^{n-r} \right\} \mathbf{P}^\top, \\ \mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp &= \mathbf{P} \left\{ \begin{pmatrix} \mathbf{A}_{[11]} & \mathbf{A}_{[12]} \\ \mathbf{A}_{[12]}^\top & \mathbf{A}_{[22]} \end{pmatrix} \mid \mathbf{A}_{[11]} \in \mathcal{S}_+^r, \mathbf{A}_{[12]} \in \mathbb{R}^{r \times (n-r)}, \mathbf{A}_{[22]} \in \mathcal{S}^{n-r} \right\} \mathbf{P}^\top. \end{aligned} \quad (3.7)$$

By $\mathbf{U} \in \text{int}(\mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp)$, there exist $\mathbf{A}_{[11]} \in \text{int} \mathcal{S}_+^r$, $\mathbf{A}_{[12]} \in \mathbb{R}^{r \times (n-r)}$, and $\mathbf{A}_{[22]} \in \mathcal{S}^{n-r}$ such that

$$\mathbf{U} = \mathbf{P} \begin{pmatrix} \mathbf{A}_{[11]} & \mathbf{A}_{[12]} \\ \mathbf{A}_{[12]}^\top & \mathbf{A}_{[22]} \end{pmatrix} \mathbf{P}^\top.$$

Then we have $P_{\mathcal{F}_{d^*+1}}(\mathbf{U}) = \mathbf{P} \text{Diag}(\mathbf{A}_{[11]}, \mathbf{O}) \mathbf{P}^\top$ and $\lambda_{\min}^+(P_{\mathcal{F}_{d^*+1}}(\mathbf{U})) = \lambda_{\min}(\mathbf{A}_{[11]}) \in (0, +\infty)$.

First, we show the inclusion $\mathbb{B}(\mathbf{U}, \lambda_{\min}(\mathbf{A}_{[11]})) \subseteq \mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp$. Suppose that $\mathbf{X} \in \mathcal{S}^n$ satisfies $\|\mathbf{U} - \mathbf{X}\|_F \leq \lambda_{\min}(\mathbf{A}_{[11]})$. We divide $\mathbf{P}^\top \mathbf{X} \mathbf{P}$ into

$$\mathbf{P}^\top \mathbf{X} \mathbf{P} = \begin{pmatrix} \mathbf{Z}_{[11]} & \mathbf{Z}_{[12]} \\ \mathbf{Z}_{[12]}^\top & \mathbf{Z}_{[22]} \end{pmatrix},$$

where $\mathbf{Z}_{[11]} \in \mathcal{S}^r$, $\mathbf{Z}_{[12]} \in \mathbb{R}^{r \times (n-r)}$, and $\mathbf{Z}_{[22]} \in \mathcal{S}^{n-r}$. Then, since

$$\|\mathbf{A}_{[11]} - \mathbf{Z}_{[11]}\|_F \leq \left\| \begin{pmatrix} \mathbf{A}_{[11]} & \mathbf{A}_{[12]} \\ \mathbf{A}_{[12]}^\top & \mathbf{A}_{[22]} \end{pmatrix} - \begin{pmatrix} \mathbf{Z}_{[11]} & \mathbf{Z}_{[12]} \\ \mathbf{Z}_{[12]}^\top & \mathbf{Z}_{[22]} \end{pmatrix} \right\|_F = \|\mathbf{U} - \mathbf{X}\|_F \leq \lambda_{\min}(\mathbf{A}_{[11]}),$$

we have $|\lambda_{\max}(\mathbf{A}_{[11]} - \mathbf{Z}_{[11]})| \leq \lambda_{\min}(\mathbf{A}_{[11]})$. For any $\mathbf{x} \in \mathbb{R}^r$ with $\|\mathbf{x}\|_2 = 1$, we see that

$$\mathbf{x}^\top \mathbf{Z}_{[11]} \mathbf{x} = \mathbf{x}^\top \mathbf{A}_{[11]} \mathbf{x} - \mathbf{x}^\top (\mathbf{A}_{[11]} - \mathbf{Z}_{[11]}) \mathbf{x} \geq \lambda_{\min}(\mathbf{A}_{[11]}) - \lambda_{\max}(\mathbf{A}_{[11]} - \mathbf{Z}_{[11]}) \geq 0,$$

from which we obtain $\mathbf{Z}_{[11]} \in \mathcal{S}_+^r$. Therefore, from (3.7), we have $\mathbf{X} \in \mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp$.

Next, we show the inclusion $\mathbb{B}(\mathbf{U}, \epsilon) \subseteq \mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp$ does not hold for any ϵ such that $\epsilon > \lambda_{\min}(\mathbf{A}_{[11]})$. Let $\mathbf{Q} \text{Diag}(\lambda_1, \dots, \lambda_r) \mathbf{Q}^\top$ be an eigendecomposition of $\mathbf{A}_{[11]}$, where $\mathbf{Q}$ is an orthogonal matrix of order r and $\lambda_1, \dots, \lambda_r$ are the eigenvalues of $\mathbf{A}_{[11]}$ satisfying $\lambda_1 \geq \dots \geq \lambda_r = \lambda_{\min}(\mathbf{A}_{[11]})$. We define

$$\begin{aligned} \mathbf{R} &:= \mathbf{P} \text{Diag}(\mathbf{Q} \text{Diag}(0, \dots, 0, \epsilon) \mathbf{Q}^\top, \mathbf{O}) \mathbf{P}^\top, \\ \mathbf{X} &:= \mathbf{U} - \mathbf{R} = \mathbf{P} \begin{pmatrix} \mathbf{Q} \text{Diag}(\lambda_1, \dots, \lambda_{r-1}, \lambda_r - \epsilon) \mathbf{Q}^\top & \mathbf{A}_{[12]} \\ \mathbf{A}_{[12]}^\top & \mathbf{A}_{[22]} \end{pmatrix} \mathbf{P}^\top. \end{aligned}$$

On the one hand, we have $\|U - X\|_F = \|R\|_F = \epsilon$, which implies that $X \in \mathbb{B}(U, \epsilon)$. On the other hand, since $\lambda_r - \epsilon < 0$, the matrix $Q \text{Diag}(\lambda_1, \dots, \lambda_{r-1}, \lambda_r - \epsilon) Q^\top$ is not positive semidefinite and X does not belong to $\mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp$. $\square$

In the next lemma, we show that the distance from a given matrix X to $\mathcal{V} \cap \mathcal{F}_{d^*+1}$ can be bounded by the maximum distances to the two sets $\mathcal{V} \cap \text{span } \mathcal{F}_{d^*+1}$ and $\mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp$.

Lemma 3.2. *Suppose that $\mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1} \neq \emptyset$. For every $X \in \mathcal{S}^n$, we have*

$$\text{dist}(X, \mathcal{V} \cap \mathcal{F}_{d^*+1}) \leq \left(1 + 2 \inf_{U \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}} \frac{\|U - X\|_F}{\lambda_{\min}^+(U)}\right) \max\{\text{dist}(X, \mathcal{V} \cap \text{span } \mathcal{F}_{d^*+1}), \text{dist}(X, \mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp)\}. \quad (3.8)$$

Proof. When $\mathcal{F}_{d^*+1} = \{\mathbf{O}\}$, both the left- and right-hand sides of (3.8) are $\|X\|_F$, and the desired inequality holds. In what follows, we assume that $\mathcal{F}_{d^*+1} \neq \{\mathbf{O}\}$. For convenience, we define $\delta(X) := \max\{\text{dist}(X, \mathcal{V} \cap \text{span } \mathcal{F}_{d^*+1}), \text{dist}(X, \mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp)\}$. Let $U \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}$ be arbitrary and let

$$V := U + \frac{\lambda_{\min}^+(U)}{2\delta(X)} (P_{\mathcal{V} \cap \text{span } \mathcal{F}_{d^*+1}}(X) - P_{\mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp}(P_{\mathcal{V} \cap \text{span } \mathcal{F}_{d^*+1}}(X))). \quad (3.9)$$

We note that $\lambda_{\min}^+(U)$ is a real number since U is a nonzero positive semidefinite matrix. Then $\|V - U\|_F = \frac{\lambda_{\min}^+(U)}{2\delta(X)} \text{dist}(P_{\mathcal{V} \cap \text{span } \mathcal{F}_{d^*+1}}(X), \mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp)$, and hence, the matrix V satisfies

$$\begin{aligned} \|V - U\|_F &\leq \frac{\lambda_{\min}^+(U)}{2\delta(X)} \|P_{\mathcal{V} \cap \text{span } \mathcal{F}_{d^*+1}}(X) - P_{\mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp}(X)\|_F \\ &\leq \frac{\lambda_{\min}^+(U)}{2\delta(X)} (\text{dist}(X, \mathcal{V} \cap \text{span } \mathcal{F}_{d^*+1}) + \text{dist}(X, \mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp)) \\ &\leq \lambda_{\min}^+(U) \\ &= \lambda_{\min}^+(P_{\mathcal{F}_{d^*+1}}(U)), \end{aligned}$$

i.e., $V \in \mathbb{B}(U, \lambda_{\min}^+(P_{\mathcal{F}_{d^*+1}}(U)))$, where the first inequality follows from $P_{\mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp}(X) \in \mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp$, the second inequality follows from the triangle inequality, the third inequality follows from the definition of $\delta(X)$, and the last equality follows from $U \in \mathcal{F}_{d^*+1}$. In addition, it follows from (3.4) that $U \in \text{int}(\mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp)$. Therefore, from Lemma 3.1, we see that $V \in \mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp$.

We let

$$Y := \frac{2\delta(X)}{\lambda_{\min}^+(U) + 2\delta(X)} V + \frac{\lambda_{\min}^+(U)}{\lambda_{\min}^+(U) + 2\delta(X)} P_{\mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp}(P_{\mathcal{V} \cap \text{span } \mathcal{F}_{d^*+1}}(X)).$$

The matrix Y belongs to $\mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp$ since it is a convex combination of the two elements in $\mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp$. In addition, it also belongs to $\mathcal{V} \cap \text{span } \mathcal{F}_{d^*+1}$ since by substituting V defined in (3.9) into the definition of Y , the matrix Y can be also represented as

$$Y = \frac{2\delta(X)}{\lambda_{\min}^+(U) + 2\delta(X)} U + \frac{\lambda_{\min}^+(U)}{\lambda_{\min}^+(U) + 2\delta(X)} P_{\mathcal{V} \cap \text{span } \mathcal{F}_{d^*+1}}(X), \quad (3.10)$$

which is a linear combination of the two elements in $\mathcal{V} \cap \text{span } \mathcal{F}_{d^*+1}$. Then it follows from (3.3) that $Y \in \mathcal{V} \cap \mathcal{F}_{d^*+1}$, and so we have

$$\begin{aligned} \text{dist}(X, \mathcal{V} \cap \mathcal{F}_{d^*+1}) &\leq \|X - Y\|_F \\ &\leq \text{dist}(X, \mathcal{V} \cap \text{span } \mathcal{F}_{d^*+1}) + \|Y - P_{\mathcal{V} \cap \text{span } \mathcal{F}_{d^*+1}}(X)\|_F \\ &\leq \delta(X) + \frac{2\delta(X)}{\lambda_{\min}^+(U)} \|U - P_{\mathcal{V} \cap \text{span } \mathcal{F}_{d^*+1}}(X)\|_F \\ &\leq \left(1 + \frac{2\|U - X\|_F}{\lambda_{\min}^+(U)}\right) \delta(X), \end{aligned} \quad (3.11)$$

where the second inequality follows from the triangle inequality, the third inequality follows from the definition of $\delta(\mathbf{X})$ and (3.10), and the fourth inequality follows from $\mathbf{U} \in \mathcal{V} \cap \text{span } \mathcal{F}_{d^*+1}$ and the nonexpansiveness of the projection. Since $\mathbf{U} \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}$ is arbitrary, by taking the infimum over $\mathbf{U} \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}$ in (3.11), we obtain the desired result. $\square$

The proof of Lemma 3.2 follows the general strategy of [7, Lemma 3.1], which addresses convex feasibility problems. In [7], one fixes a matrix $\mathbf{U} \in \text{int}(\mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp)$ and chooses $\epsilon > 0$ such that $\mathbb{B}(\mathbf{U}, \epsilon) \subseteq \mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp$, leading to an error bound that depends on this particular choice. In our setting, however, Lemma 3.1 allows us to compute the maximal admissible ϵ explicitly. Moreover, by taking the infimum over all $\mathbf{U} \in \text{int}(\mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp)$, we obtain the following sharper estimate.

Proposition 3.3. *Suppose that $\mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1} \neq \emptyset$. For each nonnegative ρ , we let*

$$\kappa(\rho) := \max\{\eta(\mathcal{V}, \text{span } \mathcal{F}_{d^*+1}), 1\}\theta(\rho), \quad (3.12)$$

where η is given as in Lemma 2.2 and $\theta: \mathbb{R}_+ \rightarrow [1, +\infty)$ is given by

$$\theta(\rho) := 1 + 2 \sup_{\mathbf{Y} \in \mathbb{B}(\mathbf{O}, \rho)} \inf_{\mathbf{U} \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}} \frac{\|\mathbf{U} - \mathbf{Y}\|_F}{\lambda_{\min}^+(\mathbf{U})}. \quad (3.13)$$

Then we have

$$\text{dist}(\mathbf{X}, \mathcal{V} \cap \mathcal{F}_{d^*+1}) \leq \kappa(\|\mathbf{X}\|_F) \max\{\text{dist}(\mathbf{X}, \mathcal{V}), \text{dist}(\mathbf{X}, \mathcal{F}_{d^*+1})\} \text{ for all } \mathbf{X} \in \mathcal{S}^n.$$

Proof. Using Lemma 3.2, we have

$$\begin{aligned} \text{dist}(\mathbf{X}, \mathcal{V} \cap \mathcal{F}_{d^*+1}) &\leq \left(1 + 2 \inf_{\mathbf{U} \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}} \frac{\|\mathbf{U} - \mathbf{X}\|_F}{\lambda_{\min}^+(\mathbf{U})}\right) \max\{\text{dist}(\mathbf{X}, \mathcal{V} \cap \text{span } \mathcal{F}_{d^*+1}), \text{dist}(\mathbf{X}, \mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp)\} \\ &\leq \theta(\|\mathbf{X}\|_F) \max\{\text{dist}(\mathbf{X}, \mathcal{V} \cap \text{span } \mathcal{F}_{d^*+1}), \text{dist}(\mathbf{X}, \mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp)\}. \end{aligned} \quad (3.14)$$

In addition, since $\mathcal{V}$ and $\text{span } \mathcal{F}_{d^*+1}$ are polyhedral, it follows from the definition of $\eta(\mathcal{V}, \text{span } \mathcal{F}_{d^*+1})$ that

$$\text{dist}(\mathbf{X}, \mathcal{V} \cap \text{span } \mathcal{F}_{d^*+1}) \leq \eta(\mathcal{V}, \text{span } \mathcal{F}_{d^*+1}) \max\{\text{dist}(\mathbf{X}, \mathcal{V}), \text{dist}(\mathbf{X}, \text{span } \mathcal{F}_{d^*+1})\}. \quad (3.15)$$

From (3.14) and (3.15), we obtain

$$\begin{aligned} &\text{dist}(\mathbf{X}, \mathcal{V} \cap \mathcal{F}_{d^*+1}) \\ &\leq \max\{\eta(\mathcal{V}, \text{span } \mathcal{F}_{d^*+1}), 1\}\theta(\|\mathbf{X}\|_F) \max\{\text{dist}(\mathbf{X}, \mathcal{V}), \text{dist}(\mathbf{X}, \text{span } \mathcal{F}_{d^*+1}), \text{dist}(\mathbf{X}, \mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp)\} \\ &\leq \max\{\eta(\mathcal{V}, \text{span } \mathcal{F}_{d^*+1}), 1\}\theta(\|\mathbf{X}\|_F) \max\{\text{dist}(\mathbf{X}, \mathcal{V}), \text{dist}(\mathbf{X}, \mathcal{F}_{d^*+1})\} \\ &= \kappa(\|\mathbf{X}\|_F) \max\{\text{dist}(\mathbf{X}, \mathcal{V}), \text{dist}(\mathbf{X}, \mathcal{F}_{d^*+1})\}, \end{aligned}$$

where the second inequality holds since $\mathcal{F}_{d^*+1}$ is included in both $\text{span } \mathcal{F}_{d^*+1}$ and $\mathcal{F}_{d^*+1} + \mathcal{F}_{d^*+1}^\perp$, and the equality follows from the definition of the function κ . Therefore, we obtain the desired result. $\square$

In general, it may be difficult to directly compute the exact value of $\theta(\rho)$. However, an upper bound can be obtained through the triangle inequality as follows:

$$\theta(\rho) \leq 1 + 2 \inf_{\mathbf{U} \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}} \frac{\|\mathbf{U}\|_F + \rho}{\lambda_{\min}^+(\mathbf{U})}. \quad (3.16)$$

In addition, by taking $\mathbf{U}_0 \in \mathcal{V} \cap \mathcal{F}_{d^*+1}$, it follows from (3.16) that

$$\theta(\rho) \leq \bar{\theta}(\rho; \mathbf{U}_0) := 1 + 2 \frac{\|\mathbf{U}_0\|_F + \rho}{\lambda_{\min}^+(\mathbf{U}_0)}. \quad (3.17)$$

We now discuss situations where the radial modulus function κ can be bounded uniformly by a constant, which leads to global error bounds. Recall that the function κ defined in (3.12) is composed of the constant $\eta(\mathcal{V}, \text{span } \mathcal{F}_{d^*+1})$ and the function θ . Firstly, we note that the constant $\eta(\mathcal{V}, \text{span } \mathcal{F}_{d^*+1})$ can be effectively upper

bounded by computing the smallest positive singular value of a linear mapping determined by the two affine spaces $\mathcal{V}$ and $\text{span } \mathcal{F}_{d^*+1}$, as shown in (ii) of Lemma 2.2. By [5, Corollary 11], we can also bound it by a constant using the angle between these two affine spaces. Secondly, we discuss how $\theta(\rho)$ can be bounded by a more concise quantity. As shown in the following corollary, the upper bound in (3.16) can be bounded by an explicit quantity that does not depend on ρ when $\mathcal{V}$ is a subspace, so that Proposition 3.3 implies a global Lipschitz error bound for $\text{Feas}(\mathcal{V}, \mathcal{F}_{d^*+1})$. We define the *condition number* of the set $\mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}$ by

$$\text{cond}(\mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}) := \inf \left\{ \frac{\lambda_{\max}(\mathbf{X})}{\lambda_{\min}^+(\mathbf{X})} \mid \mathbf{X} \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1} \right\}.$$

Note that various versions of the condition number for a general conic system have been introduced and studied in the literature, see for example [8, 39, 45]. The one we used here is a variant of those in the literature, restricted to the setting of a semidefinite system, and expressed explicitly in terms of the extreme eigenvalues of the matrices.

Corollary 3.4. *Let $\mathcal{V}$ be a subspace, and let the matrix $\mathbf{P}$ and the face $\mathcal{F}_{d^*+1}$ be as in (3.6), in which case $\mathcal{F}_{d^*+1}$ is linearly isomorphic to $\mathcal{S}_+^r$. Suppose that $\mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1} \neq \emptyset$. Then, for all $\rho \in \mathbb{R}_+$,*

$$\theta(\rho) \leq \theta_0 := 1 + 2\sqrt{r} \text{cond}(\mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}).$$

Moreover, if there exists $\alpha > 0$ such that $\mathbf{P} \text{Diag}(\alpha \mathbf{I}_r, \mathbf{O}) \mathbf{P}^\top \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}$, then $\theta_0 = 1 + 2\sqrt{r}$, and so, for every $\rho \in \mathbb{R}_+$, we have

$$\theta(\rho) \leq 1 + 2\sqrt{r}. \quad (3.18)$$

Proof. To begin with, we show that

$$\inf_{\mathbf{U} \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}} \frac{\|\mathbf{U}\|_{\text{F}} + \rho}{\lambda_{\min}^+(\mathbf{U})} = \inf_{\mathbf{U} \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}} \frac{\|\mathbf{U}\|_{\text{F}}}{\lambda_{\min}^+(\mathbf{U})}. \quad (3.19)$$

Since ρ is nonnegative, we see that the right-hand side of (3.19) is less than or equal to the left-hand side of (3.19). To prove the equality, let $\mathbf{U}_0 \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}$ be arbitrary. The set $\mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}$ is a cone (not necessarily containing the origin) since $\mathcal{V}$ is a subspace, so $t\mathbf{U}_0 \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}$ holds for all $t > 0$. Then we have for every $t > 0$

$$\inf_{\mathbf{U} \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}} \frac{\|\mathbf{U}\|_{\text{F}} + \rho}{\lambda_{\min}^+(\mathbf{U})} \leq \frac{\|t\mathbf{U}_0\|_{\text{F}} + \rho}{\lambda_{\min}^+(t\mathbf{U}_0)} = \frac{\|\mathbf{U}_0\|_{\text{F}}}{\lambda_{\min}^+(\mathbf{U}_0)} + \frac{\rho}{t\lambda_{\min}^+(\mathbf{U}_0)}.$$

Since $t > 0$ is arbitrary, taking the limit $t \rightarrow \infty$, we have

$$\inf_{\mathbf{U} \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}} \frac{\|\mathbf{U}\|_{\text{F}} + \rho}{\lambda_{\min}^+(\mathbf{U})} \leq \frac{\|\mathbf{U}_0\|_{\text{F}}}{\lambda_{\min}^+(\mathbf{U}_0)}.$$

Since $\mathbf{U}_0 \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}$ is also arbitrary, we obtain

$$\inf_{\mathbf{U} \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}} \frac{\|\mathbf{U}\|_{\text{F}} + \rho}{\lambda_{\min}^+(\mathbf{U})} \leq \inf_{\mathbf{U} \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}} \frac{\|\mathbf{U}\|_{\text{F}}}{\lambda_{\min}^+(\mathbf{U})},$$

and the equality in (3.19) holds.

For any $\mathbf{U} \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}$, since $\|\mathbf{U}\|_{\text{F}}$ is the square root of the sum of squares of the nonzero eigenvalues of $\mathbf{U}$ and the number of its nonzero eigenvalues is r , it follows from the definition of $\text{cond}(\mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1})$ that

$$\inf_{\mathbf{U} \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}} \frac{\|\mathbf{U}\|_{\text{F}}}{\lambda_{\min}^+(\mathbf{U})} \leq \sqrt{r} \text{cond}(\mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}). \quad (3.20)$$

Therefore, we obtain

$$\theta(\rho) \stackrel{(a)}{\leq} 1 + 2 \inf_{\mathbf{U} \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}} \frac{\|\mathbf{U}\|_{\text{F}} + \rho}{\lambda_{\min}^+(\mathbf{U})} \stackrel{(b)}{=} 1 + 2 \inf_{\mathbf{U} \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}} \frac{\|\mathbf{U}\|_{\text{F}}}{\lambda_{\min}^+(\mathbf{U})} \stackrel{(c)}{\leq} 1 + 2\sqrt{r} \text{cond}(\mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}),$$

where we use (3.16) to derive (a), use (3.19) to derive (b), and use (3.20) to derive (c).

Finally, suppose, in addition, that there exists $\alpha > 0$ such that $\mathbf{P} \text{Diag}(\alpha \mathbf{I}_r, \mathbf{O}) \mathbf{P}^\top \in \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1}$. Then the condition number $\text{cond}(\mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1})$ equals 1, and so, the conclusion follows. $\square$

As we will see in Section 4, the radial modulus function in the error bound for the problem $\text{Feas}(\mathcal{V}, \mathcal{S}_+^n)$ cannot, in general, be uniformly bounded by a constant. Moreover, the asymptotic tightness of the derived error bound depends on the quantity d^* , which measures the distance to the PPS condition of the semidefinite feasibility problem.

In contrast, the following example demonstrates that one can construct simple instances for which the radial modulus function in the error bound for $\text{Feas}(\mathcal{V}, \mathcal{F}_{d^*+1})$, defined in (3.12), is uniformly bounded by a dimension-free constant. In this case, the resulting estimate is asymptotically tight in the sense of Definition 2.4. A more detailed investigation of tightness in a broader setting will be presented in Section 4.

Example 3.5 (An instance where the error bound for the facially reduced problem is asymptotically tight). Let $n \in \mathbb{Z}_{\geq 3}$ and

$$\begin{aligned}\mathcal{K}_n &:= \mathcal{S}_+^2 \oplus \{0\}^{n-2}, \\ \mathcal{V}_n &:= \text{span } \mathcal{K}_n = \mathcal{S}^2 \oplus \{0\}^{n-2}.\end{aligned}$$

Note that $\mathcal{K}_n$ is a face of $\mathcal{S}_+^n$ exposed by the matrix $\frac{1}{\sqrt{n-2}} \sum_{i=3}^n \mathbf{E}_{ii} \in \mathcal{S}_+^n \cap \mathcal{V}_n^\perp$ and

$$\mathcal{V}_n \cap \text{ri } \mathcal{K}_n = (\text{int } \mathcal{S}_+^2) \oplus \{0\}^{n-2} \neq \emptyset.$$

Then Proposition 3.3 implies that

$$\text{dist}(\mathbf{X}, \mathcal{V}_n \cap \mathcal{K}_n) \leq \kappa(\|\mathbf{X}\|_F) \max\{\text{dist}(\mathbf{X}, \mathcal{V}_n), \text{dist}(\mathbf{X}, \mathcal{K}_n)\} \text{ for all } \mathbf{X} \in \mathcal{S}^n, \quad (3.21)$$

where $\kappa(\|\mathbf{X}\|_F) = \max\{\eta(\mathcal{V}_n, \text{span } \mathcal{K}_n), 1\}\theta(\|\mathbf{X}\|_F)$.

To see the asymptotic tightness of the error bound in (3.21), we take a sequence $(n_k) \subseteq \mathbb{Z}_{\geq 2}$ satisfying $\lim_{k \rightarrow \infty} n_k = \infty$ and a sequence $(\mathbf{X}_k)$ satisfying $\mathbf{X}_k \in \mathcal{S}^{n_k}$ for all k arbitrarily. It follows from $\mathcal{V}_{n_k} = \text{span } \mathcal{K}_{n_k}$ and Remark 2.3 that $\eta(\mathcal{V}_{n_k}, \text{span } \mathcal{K}_{n_k}) = 1$. In addition, since $\mathcal{V}_{n_k}$ is a subspace and $\text{Diag}(\mathbf{I}_2, \mathbf{O}) \in \mathcal{V}_{n_k} \cap \text{ri } \mathcal{K}_{n_k}$ holds, it follows from (3.18) that $\theta(\|\mathbf{X}_k\|_F) \leq 1 + 2\sqrt{2}$, and so $\kappa(\|\mathbf{X}_k\|_F) \leq 1 + 2\sqrt{2}$. Therefore, we have

$$\frac{\text{dist}(\mathbf{X}_k, \mathcal{V}_{n_k} \cap \mathcal{K}_{n_k})}{\kappa(\|\mathbf{X}_k\|_F) \max\{\text{dist}(\mathbf{X}_k, \mathcal{V}_{n_k}), \text{dist}(\mathbf{X}_k, \mathcal{K}_{n_k})\}} \geq \frac{\text{dist}(\mathbf{X}_k, \mathcal{K}_{n_k})}{(1 + 2\sqrt{2}) \text{dist}(\mathbf{X}_k, \mathcal{K}_{n_k})} = \frac{1}{1 + 2\sqrt{2}}.$$

Thus, we see that the error bound in (3.21) is asymptotically tight with a dimension-free constant $\frac{1}{1+2\sqrt{2}}$.

Considering the case where $\mathcal{F}_{d^*+1}$ is polyhedral leads to the following theorem.

Theorem 3.6 (Radial modulus function for facially reduced problem with minimal face). Let θ be defined as in (3.13) and let η be given as in Lemma 2.2. Then

$$\text{dist}(\mathbf{X}, \mathcal{V} \cap \mathcal{F}_{d^*+1}) \leq \kappa(\|\mathbf{X}\|_F) \max\{\text{dist}(\mathbf{X}, \mathcal{V}), \text{dist}(\mathbf{X}, \mathcal{F}_{d^*+1})\} \text{ for all } \mathbf{X} \in \mathcal{S}^n,$$

where the function $\kappa(\rho)$ can be set to

$$\kappa(\rho) = \begin{cases} \max\{\eta(\mathcal{V}, \text{span } \mathcal{F}_{d^*+1}), 1\}\theta(\rho) & (\text{if } \mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1} \neq \emptyset), \\ \eta(\mathcal{V}, \mathcal{F}_{d^*+1}) & (\text{if } \mathcal{F}_{d^*+1} \text{ is polyhedral}). \end{cases}$$

Proof. If $\mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1} \neq \emptyset$ holds, then the statement holds by Proposition 3.3. If $\mathcal{F}_{d^*+1}$ is polyhedral, it follows from the definition of $\eta(\mathcal{V}, \mathcal{F}_{d^*+1})$ that

$$\text{dist}(\mathbf{X}, \mathcal{V} \cap \mathcal{F}_{d^*+1}) \leq \eta(\mathcal{V}, \mathcal{F}_{d^*+1}) \max\{\text{dist}(\mathbf{X}, \mathcal{V}), \text{dist}(\mathbf{X}, \mathcal{F}_{d^*+1})\} \text{ for all } \mathbf{X} \in \mathcal{S}^n.$$

Therefore, (3.2) holds by setting $\kappa(\rho) := \eta(\mathcal{V}, \mathcal{F}_{d^*+1})$ for any $\rho \in \mathbb{R}_+$. $\square$

When $\mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1} \neq \emptyset$, the upper bound for the function θ given in (3.17) yields an upper bound for the function κ . Indeed, by taking $\mathbf{U}_0 \in \mathcal{V} \cap \mathcal{F}_{d^*+1}$ and letting

$$\bar{\kappa}(\rho; \mathbf{U}_0) := \max\{\eta(\mathcal{V}, \text{span } \mathcal{F}_{d^*+1}), 1\}\bar{\theta}(\rho; \mathbf{U}_0), \quad (3.22)$$

we then obtain $\kappa(\rho) \leq \bar{\kappa}(\rho; \mathbf{U}_0)$.

In passing, we note that the constant $\eta(\mathcal{V}, \mathcal{F}_{d^*+1})$ can also be upper bounded by using a Hoffman constant (as shown in (i) of Lemma 2.2), and so, can be effectively estimated. Also, the two situations $\mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1} \neq \emptyset$ and $\mathcal{F}_{d^*+1}$ being polyhedral can hold simultaneously; in that case, we can set

$$\kappa(\rho) := \min\{\max\{\eta(\mathcal{V}, \text{span } \mathcal{F}_{d^*+1}), 1\}\theta(\rho), \eta(\mathcal{V}, \mathcal{F}_{d^*+1})\}. \quad (3.23)$$

3.2 Constants for one-step facial residual functions

Recall that a one-step facial residual function for a face of the positive semidefinite cone can be written as (2.5). This means that for $i = 1, \dots, d^*$, a one-step facial residual function $\psi_{\mathcal{F}_i, \mathbf{Z}_i}(s, t)$ for $\mathcal{F}_i$ and $\mathbf{Z}_i$ takes the form of

$$\psi_{\mathcal{F}_i, \mathbf{Z}_i}(s, t) = \alpha_i s + \beta_i \sqrt{st}$$

for some nonnegative α_i and β_i , where the $\mathcal{F}_i$'s and $\mathbf{Z}_i$'s are as in (3.1). In this subsection, we compute the constants α_i and β_i explicitly. We also discuss how these constants relate to the choices of $\mathcal{F}_i$ and $\mathbf{Z}_i$.

The following lemma provides a one-step facial residual function for the positive semidefinite cone $\mathcal{S}_+^n$ associated with a block diagonal matrix having one positive definite block and all other blocks equal to zero.

Lemma 3.7. *Let $r < n$ and $\mathbf{\Lambda} = \text{Diag}(\lambda_1, \dots, \lambda_{n-r})$ be a diagonal matrix with positive diagonal elements $\lambda_1 \geq \dots \geq \lambda_{n-r} > 0$. Let $\mathbf{Z} := \text{Diag}(\mathbf{O}, \mathbf{\Lambda}) \in \mathcal{S}_+^n$. Define*

$$\begin{aligned} \gamma &:= \frac{\sqrt{\sum_{j=1}^{n-r} \lambda_j^2} + 1}{\lambda_{n-r}} + 1 = \frac{\|\mathbf{\Lambda}\|_F + 1}{\lambda_{\min}(\mathbf{\Lambda})} + 1, \\ \alpha &:= \sqrt{1 + 2(\gamma + 1)r + \gamma^2}, \\ \beta &:= \sqrt{2(\gamma + 1)\sqrt{r}}. \end{aligned}$$

Then for any $\mathbf{X} \in \mathcal{S}^n$ and $\epsilon \in \mathbb{R}_+$ such that $\text{dist}(\mathbf{X}, \mathcal{S}_+^n) \leq \epsilon$ and $\langle \mathbf{Z}, \mathbf{X} \rangle \leq \epsilon$, we have

$$\text{dist}(\mathbf{X}, \mathcal{S}_+^n \cap \{\mathbf{Z}\}^\perp) \leq \alpha\epsilon + \beta\sqrt{\epsilon\|\mathbf{X}\|_F}. \quad (3.24)$$

In other words, for α and β defined above, $\alpha s + \beta\sqrt{st}$ is a one-step facial residual function for $\mathcal{S}_+^n$ and $\mathbf{Z}$.

Proof. We partition $\mathbf{X}$ into

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_{[11]} & \mathbf{X}_{[12]} \\ \mathbf{X}_{[12]}^\top & \mathbf{X}_{[22]} \end{pmatrix},$$

where $\mathbf{X}_{[11]} \in \mathcal{S}^r$, $\mathbf{X}_{[12]} \in \mathbb{R}^{r \times (n-r)}$, and $\mathbf{X}_{[22]} \in \mathcal{S}^{n-r}$. To prove (3.24), we show the following inequalities:

$$\text{dist}(\mathbf{X}_{[11]}, \mathcal{S}_+^r) \leq \epsilon, \quad (3.25)$$

$$\|\mathbf{X}_{[22]}\|_F \leq \gamma\epsilon, \quad (3.26)$$

$$\|\mathbf{X}_{[12]}\|_F \leq \sqrt{(\gamma + 1)(\|\mathbf{X}\|_F + \sqrt{r}\epsilon)\sqrt{r}\epsilon}. \quad (3.27)$$

Indeed, if (3.25), (3.26), and (3.27) hold, then it follows that

$$\begin{aligned} \text{dist}(\mathbf{X}, \mathcal{S}_+^n \cap \{\mathbf{Z}\}^\perp) &= \text{dist}(\mathbf{X}, \mathcal{S}_+^r \oplus \{0\}^{n-r}) \\ &= \sqrt{\text{dist}(\mathbf{X}_{[11]}, \mathcal{S}_+^r)^2 + \|\mathbf{X}_{[22]}\|_F^2 + 2\|\mathbf{X}_{[12]}\|_F^2} \\ &\leq \sqrt{\epsilon^2 + \gamma^2\epsilon^2 + 2(\gamma + 1)(\|\mathbf{X}\|_F + \sqrt{r}\epsilon)\sqrt{r}\epsilon} \\ &= \sqrt{(1 + 2(\gamma + 1)r + \gamma^2)\epsilon^2 + 2(\gamma + 1)\sqrt{r}\epsilon\|\mathbf{X}\|_F} \\ &\leq \alpha\epsilon + \beta\sqrt{\epsilon\|\mathbf{X}\|_F}, \end{aligned}$$

so (3.24) holds. We next turn to the proofs of (3.25), (3.26), and (3.27).

Proof of (3.25) (3.25) follows from $\text{dist}(\mathbf{X}, \mathcal{S}_+^n) \leq \epsilon$ and (2.2).

Proof of (3.26) It follows from $\text{dist}(\mathbf{X}, \mathcal{S}_+^n) \leq \epsilon$ and (2.2) that $\text{dist}(\mathbf{X}_{[22]}, \mathcal{S}_+^{n-r}) \leq \epsilon$. $\mathbf{R} := -\mathbf{X}_{[22]} + P_{\mathcal{S}_+^{n-r}}(\mathbf{X}_{[22]})$ satisfies $\mathbf{X}_{[22]} + \mathbf{R} \in \mathcal{S}_+^{n-r}$ and $\|\mathbf{R}\|_F \leq \epsilon$. In addition, we see that

$$\|\mathbf{X}_{[22]} + \mathbf{R}\|_F \leq \text{tr}(\mathbf{X}_{[22]} + \mathbf{R}) \leq \frac{\langle \mathbf{\Lambda}, \mathbf{X}_{[22]} + \mathbf{R} \rangle}{\lambda_{n-r}},$$

where we use $\mathbf{X}_{[22]} + \mathbf{R} \in \mathcal{S}_+^{n-r}$ to derive the first inequality. Therefore, we have

$$\|\mathbf{X}_{[22]}\|_F \leq \|\mathbf{R}\|_F + \frac{\langle \mathbf{A}, \mathbf{X}_{[22]} \rangle}{\lambda_{n-r}} + \frac{\langle \mathbf{A}, \mathbf{R} \rangle}{\lambda_{n-r}} \leq \|\mathbf{R}\|_F + \frac{\epsilon}{\lambda_{n-r}} + \frac{\|\mathbf{A}\|_F \|\mathbf{R}\|_F}{\lambda_{n-r}} \leq \gamma \epsilon,$$

where we use $\langle \mathbf{A}, \mathbf{X}_{[22]} \rangle = \langle \mathbf{Z}, \mathbf{X} \rangle \leq \epsilon$ and the Cauchy–Schwarz inequality to derive the second inequality and use $\|\mathbf{R}\|_F \leq \epsilon$ to derive the third inequality.

Proof of (3.27) By the assumption $\text{dist}(\mathbf{X}, \mathcal{S}_+^n) \leq \epsilon$ and the distance formula in (2.1), every eigenvalue of $\mathbf{X}$ is at least $-\epsilon$, and hence $\mathbf{X} + \epsilon \mathbf{I}_n \in \mathcal{S}_+^n$. This implies that

$$\mathbf{X} + (\delta + \epsilon) \mathbf{I}_n = \begin{pmatrix} \mathbf{X}_{[11]} + (\delta + \epsilon) \mathbf{I}_r & \mathbf{X}_{[12]} \\ \mathbf{X}_{[12]}^\top & \mathbf{X}_{[22]} + (\delta + \epsilon) \mathbf{I}_{n-r} \end{pmatrix} \in \text{int } \mathcal{S}_+^n$$

for all $\delta > 0$. Therefore, the Schur complement

$$\mathbf{X}_{[11]} + (\delta + \epsilon) \mathbf{I}_r - \mathbf{X}_{[12]} (\mathbf{X}_{[22]} + (\delta + \epsilon) \mathbf{I}_{n-r})^{-1} \mathbf{X}_{[12]}^\top \quad (3.28)$$

is positive definite. Taking the inner product of (3.28) with $\mathbf{I}_r$, we have

$$\begin{aligned} \langle \mathbf{X}_{[11]} + (\delta + \epsilon) \mathbf{I}_r, \mathbf{I}_r \rangle &\geq \langle \mathbf{X}_{[12]} (\mathbf{X}_{[22]} + (\delta + \epsilon) \mathbf{I}_{n-r})^{-1} \mathbf{X}_{[12]}^\top, \mathbf{I}_r \rangle \\ &= \langle (\mathbf{X}_{[22]} + (\delta + \epsilon) \mathbf{I}_{n-r})^{-1}, \mathbf{X}_{[12]}^\top \mathbf{X}_{[12]} \rangle \\ &\geq \lambda_{\min}((\mathbf{X}_{[22]} + (\delta + \epsilon) \mathbf{I}_{n-r})^{-1}) \text{tr}(\mathbf{X}_{[12]}^\top \mathbf{X}_{[12]}) \\ &= \frac{\|\mathbf{X}_{[12]}\|_F^2}{\lambda_{\max}(\mathbf{X}_{[22]} + (\delta + \epsilon) \mathbf{I}_{n-r})} \\ &\geq \frac{\|\mathbf{X}_{[12]}\|_F^2}{\|\mathbf{X}_{[22]}\|_F + \delta + \epsilon} \\ &\geq \frac{\|\mathbf{X}_{[12]}\|_F^2}{(\gamma + 1)\epsilon + \delta}, \end{aligned} \quad (3.29)$$

where we use the following well-known inequality $\langle \mathbf{A}, \mathbf{B} \rangle \geq \lambda_{\min}(\mathbf{A}) \text{tr}(\mathbf{B})$ for $\mathbf{A} \in \mathcal{S}^n$ and $\mathbf{B} \in \mathcal{S}_+^n$ (see, for example, [9, Proposition 8.4.13]) to derive the second inequality and use (3.26) to derive the last inequality. In addition, the left-hand side of (3.29) can be bounded by

$$\langle \mathbf{X}_{[11]} + (\delta + \epsilon) \mathbf{I}_r, \mathbf{I}_r \rangle \leq \|\mathbf{X}_{[11]} + (\delta + \epsilon) \mathbf{I}_r\|_F \sqrt{r} \leq (\|\mathbf{X}\|_F + \sqrt{r}(\delta + \epsilon)) \sqrt{r}.$$

Therefore, we have

$$\|\mathbf{X}_{[12]}\|_F \leq \sqrt{((\gamma + 1)\epsilon + \delta)(\|\mathbf{X}\|_F + \sqrt{r}(\delta + \epsilon))\sqrt{r}}.$$

Since $\delta > 0$ is arbitrary, we have

$$\|\mathbf{X}_{[12]}\|_F \leq \sqrt{(\gamma + 1)(\|\mathbf{X}\|_F + \sqrt{r}\epsilon)\sqrt{r}\epsilon}.$$

□

Now, for each $i = 1, \dots, d^*$, we provide a one-step facial residual function for the face $\mathcal{F}_i$ and the matrix $\mathbf{Z}_i \in \mathcal{F}_i^*$. Each face $\mathcal{F}_i$ of $\mathcal{S}_+^n$ is linearly isomorphic to $\mathcal{S}_+^{r_i}$ for some $r_i \leq n$. When $\mathcal{F}_i = \{\mathbf{O}\}$, the index i must be $d^* + 1$ and r_{d^*+1} is set to 0. Overall, for every $i = 1, \dots, d^*$, there exists $\mathbf{Q}_i \in \mathbb{R}^{n \times r_i}$ whose column vectors are orthonormal such that

$$\mathcal{F}_i = \mathbf{Q}_i \mathcal{S}_+^{r_i} \mathbf{Q}_i^\top. \quad (3.30)$$

In addition, since $\mathcal{F}_{i+1}$ is also a face of $\mathcal{F}_i$, there exists an orthogonal matrix $\mathbf{P}$ of order r_i such that

$$\mathcal{F}_{i+1} = \mathbf{Q}_i \mathbf{P} (\mathcal{S}_+^{r_{i+1}} \oplus \{0\}^{r_i - r_{i+1}}) \mathbf{P}^\top \mathbf{Q}_i^\top. \quad (3.31)$$

Lemma 3.8. *Let $\mathbf{Q}_i \in \mathbb{R}^{n \times r_i}$ be given as in (3.30). Then for any $\mathbf{Z} \in \text{ri}(\mathcal{F}_i^* \cap \mathcal{F}_{i+1}^\perp)$, the matrix $\mathbf{Q}_i^\top \mathbf{Z} \mathbf{Q}_i$ is positive semidefinite and has rank $r_i - r_{i+1}$.*

Proof. From (3.31), it follows that

$$\mathcal{F}_{i+1}^\perp = \left\{ \mathbf{X} \in \mathcal{S}^n \mid \mathbf{P}^\top \mathbf{Q}_i^\top \mathbf{X} \mathbf{Q}_i \mathbf{P} \in \left\{ \begin{pmatrix} \mathbf{O} & \mathbf{A}_{[12]} \\ \mathbf{A}_{[12]}^\top & \mathbf{A}_{[22]} \end{pmatrix} \mid \begin{array}{l} \mathbf{A}_{[12]} \in \mathbb{R}^{r_{i+1} \times (r_i - r_{i+1})} \\ \mathbf{A}_{[22]} \in \mathcal{S}^{r_i - r_{i+1}} \end{array} \right\} \right\}. \quad (3.32)$$

In addition, we see that

$$\mathcal{F}_i^* = \{ \mathbf{X} \in \mathcal{S}^n \mid \mathbf{Q}_i^\top \mathbf{X} \mathbf{Q}_i \in \mathcal{S}_+^{r_i} \} = \{ \mathbf{X} \in \mathcal{S}^n \mid \mathbf{P}^\top \mathbf{Q}_i^\top \mathbf{X} \mathbf{Q}_i \mathbf{P} \in \mathcal{S}_+^{r_i} \}. \quad (3.33)$$

From (3.32) and (3.33), we have

$$\text{ri}(\mathcal{F}_i^* \cap \mathcal{F}_{i+1}^\perp) = \{ \mathbf{X} \in \mathcal{S}^n \mid \mathbf{P}^\top \mathbf{Q}_i^\top \mathbf{X} \mathbf{Q}_i \mathbf{P} \in \{0\}^{r_{i+1}} \oplus \text{int } \mathcal{S}_+^{r_i - r_{i+1}} \}. \quad (3.34)$$

Let $\mathbf{Z} \in \text{ri}(\mathcal{F}_i^* \cap \mathcal{F}_{i+1}^\perp)$. By (3.34), there exists $\mathbf{A}_{[22]} \in \text{int } \mathcal{S}_+^{r_i - r_{i+1}}$ such that $\mathbf{P}^\top \mathbf{Q}_i^\top \mathbf{Z} \mathbf{Q}_i \mathbf{P} = \text{Diag}(\mathbf{O}, \mathbf{A}_{[22]})$. Then it follows that $\mathbf{Q}_i^\top \mathbf{Z} \mathbf{Q}_i = \mathbf{P} \text{Diag}(\mathbf{O}, \mathbf{A}_{[22]}) \mathbf{P}^\top$. By the choice of $\mathbf{A}_{[22]}$, the matrix $\mathbf{Q}_i^\top \mathbf{Z} \mathbf{Q}_i$ is positive semidefinite and has rank $r_i - r_{i+1}$. Therefore, we obtain the desired result. $\square$

In particular, by the following lemma, Lemma 3.8 can be applied to the matrices $\mathbf{Z}_1, \dots, \mathbf{Z}_{d^*}$.

Lemma 3.9. *For every $i = 1, \dots, d^*$, $\mathbf{Z}_i \in \text{ri}(\mathcal{F}_i^* \cap \mathcal{F}_{i+1}^\perp)$ holds.*

Proof. We prove this lemma using the notion of *nice*ness. Recall that a closed convex cone $\mathcal{K}$ in $\mathcal{S}^n$ is said to be nice if $\mathcal{K}^* + \mathcal{F}^\perp$ is closed for every face $\mathcal{F}$ of $\mathcal{K}$.

We note that $\mathcal{F}_i^\perp$ is the lineality space of $\mathcal{F}_i^*$, namely, the largest subspace included in $\mathcal{F}_i^*$. It is known that $\mathcal{F}_i^*$ can be written as the sum of its lineality space $\mathcal{F}_i^\perp$ and its pointed component $\mathcal{F}_i^* \cap (\mathcal{F}_i^\perp)^\perp$ [46, page 165]. Moreover, $\mathcal{F}_i^*$ is nice if and only if its pointed component $\mathcal{F}_i^* \cap (\mathcal{F}_i^\perp)^\perp$ is nice [47, Proposition 2.6.ii]. We observe that

$$\mathcal{F}_i^* \cap (\mathcal{F}_i^\perp)^\perp = \mathcal{F}_i^* \cap \text{span } \mathcal{F}_i = \mathcal{F}_i = \mathcal{S}_+^n \cap \text{span } \mathcal{F}_i.$$

Here, the second equality follows from the self-duality of $\mathcal{F}_i$ in the subspace $\text{span } \mathcal{F}_i$. This self-duality follows from the fact that $\mathcal{F}_i$ is linearly isomorphic to $\mathcal{S}_+^{r_i}$ as shown in (3.30) and that $\mathcal{S}_+^{r_i}$ is self-dual. Since $\mathcal{S}_+^n$ is nice [36, Section 2.5] and the subspace $\text{span } \mathcal{F}_i$ is nice as a polyhedral cone [36, Section 1.1], [37, Proposition 5] implies that their intersection $\mathcal{S}_+^n \cap \text{span } \mathcal{F}_i$, which is equal to $\mathcal{F}_i^* \cap (\mathcal{F}_i^\perp)^\perp$, is nice. This means that the entire cone $\mathcal{F}_i^*$ is also nice. Therefore, by [31, Proposition 1], we conclude that $\mathbf{Z}_i \in \text{ri}(\mathcal{F}_i^* \cap \mathcal{F}_{i+1}^\perp)$. $\square$

Proposition 3.10 (Constants for the one-step facial residual functions). *For each $i = 1, \dots, d^*$, define the following constants:*

$$\gamma_i := \frac{\|\mathbf{Q}_i^\top \mathbf{Z}_i \mathbf{Q}_i\|_F + 1}{\lambda_{\min}^+(\mathbf{Q}_i^\top \mathbf{Z}_i \mathbf{Q}_i)} + 1, \quad (3.35)$$

$$\alpha_i := \sqrt{1 + 2(\gamma_i + 1)r_{i+1} + \gamma_i^2}, \quad (3.36)$$

$$\beta_i := \sqrt{2(\gamma_i + 1)\sqrt{r_{i+1}}}. \quad (3.37)$$

Then, for any $\mathbf{X} \in \text{span } \mathcal{F}_i$ and $\epsilon \in \mathbb{R}_+$ such that $\text{dist}(\mathbf{X}, \mathcal{F}_i) \leq \epsilon$ and $\langle \mathbf{Z}_i, \mathbf{X} \rangle \leq \epsilon$, we have $\text{dist}(\mathbf{X}, \mathcal{F}_{i+1}) \leq \alpha_i \epsilon + \beta_i \sqrt{\epsilon} \|\mathbf{X}\|_F$. In other words, for α_i and β_i defined above, $\alpha_i s + \beta_i \sqrt{st}$ is a one-step facial residual function for $\mathcal{F}_i$ and $\mathbf{Z}_i$.

Proof. Since $\mathbf{Q}_i^\top \mathbf{Z}_i \mathbf{Q}_i$ is a positive semidefinite matrix of rank $r_i - r_{i+1}$ by Lemmas 3.8 and 3.9, there exist an orthogonal matrix $\mathbf{P}$ of order r_i and a diagonal matrix $\mathbf{\Lambda}$ of order $r_i - r_{i+1}$ with positive diagonal elements such that $\mathbf{Q}_i^\top \mathbf{Z}_i \mathbf{Q}_i = \mathbf{P} \text{Diag}(\mathbf{O}, \mathbf{\Lambda}) \mathbf{P}^\top$.

Let $\mathbf{X} \in \text{span } \mathcal{F}_i$ and $\epsilon \in \mathbb{R}_+$ be such that $\text{dist}(\mathbf{X}, \mathcal{F}_i) \leq \epsilon$ and $\langle \mathbf{Z}_i, \mathbf{X} \rangle \leq \epsilon$. We have

$$\begin{aligned} \text{dist}(\mathbf{P}^\top \mathbf{Q}_i^\top \mathbf{X} \mathbf{Q}_i \mathbf{P}, \mathcal{S}_+^{r_i}) &= \text{dist}(\mathbf{X}, \mathcal{F}_i) \leq \epsilon, \\ \langle \text{Diag}(\mathbf{O}, \mathbf{\Lambda}), \mathbf{P}^\top \mathbf{Q}_i^\top \mathbf{X} \mathbf{Q}_i \mathbf{P} \rangle &= \langle \mathbf{Z}_i, \mathbf{X} \rangle \leq \epsilon, \\ \|\mathbf{P}^\top \mathbf{Q}_i^\top \mathbf{X} \mathbf{Q}_i \mathbf{P}\|_F &= \|\mathbf{X}\|_F. \end{aligned} \quad (3.38)$$

Applying Lemma 3.7 to $\mathbf{P}^\top \mathbf{Q}_i^\top \mathbf{X} \mathbf{Q}_i \mathbf{P}$ and $\text{Diag}(\mathbf{O}, \mathbf{\Lambda})$, the resulting constants coincide with γ_i , α_i , β_i as defined in (3.35), (3.36), (3.37) and we have

$$\begin{aligned} \text{dist}(\mathbf{X}, \mathcal{F}_{i+1}) &= \text{dist}(\mathbf{X}, \mathcal{F}_i \cap \{\mathbf{Z}_i\}^\perp) \\ &= \text{dist}(\mathbf{P}^\top \mathbf{Q}_i^\top \mathbf{X} \mathbf{Q}_i \mathbf{P}, \mathcal{S}_+^{r_i} \cap \{\text{Diag}(\mathbf{O}, \mathbf{\Lambda})\}^\perp) \\ &\leq \alpha_i \epsilon + \beta_i \sqrt{\epsilon \|\mathbf{P}^\top \mathbf{Q}_i^\top \mathbf{X} \mathbf{Q}_i \mathbf{P}\|_F} \\ &= \alpha_i \epsilon + \beta_i \sqrt{\epsilon \|\mathbf{X}\|_F}, \end{aligned}$$

where the final equality follows from (3.38). $\square$

Remark 3.11. We discuss lower bounds for the constants α_i and β_i introduced in (3.36) and (3.37), respectively. From (3.36) and (3.37), we see that $\alpha_i \geq \max\{\beta_i, 1\}$. The constant β_i depends on how the affine space $\mathcal{V}$ intersects $\mathcal{S}_+^n$. When $\mathcal{V} \cap \mathcal{S}_+^n = \{\mathbf{O}\}$, it follows from the proof of [31, Proposition 27] that we can take a matrix in $\mathcal{S}_+^n \cap \mathcal{V}^\perp$ that exposes the face $\{\mathbf{O}\}$ of $\mathcal{S}_+^n$, so that $d^* \leq 1$. If $d^* = 1$, then we have $r_2 = 0$, and (3.37) implies $\beta_1 = 0$. When $\mathcal{V} \cap \mathcal{S}_+^n \neq \{\mathbf{O}\}$, for every $i = 1, \dots, d^* + 1$, the face $\mathcal{F}_i$ strictly includes $\{\mathbf{O}\}$ and $r_i \geq 1$ holds, so we see from (3.37) that $\beta_i \geq 1$.

Remark 3.12 (Comparison to earlier work). In [31, Theorem 35] it was shown that facial residual functions for symmetric cones can be taken to be of the form $\alpha s + \beta \sqrt{st}$ for nonnegative constants α and β . This result also applies to $\mathcal{S}_+^n$ since it is a symmetric cone [18, Section I.2]. However, the constants α and β are never computed explicitly in [31]. In contrast, Proposition 3.10 leads to a computable formula for α_i and β_i and clarifies how the face $\mathcal{F}_i$ and the matrix $\mathbf{Z}_i$ affect these constants.

Specifically, the constant γ_i defined in (3.35) depends on the choice of the matrix $\mathbf{Z}_i$. Once the faces $\mathcal{F}_1, \dots, \mathcal{F}_{d^*+1}$ are fixed, any matrix $\mathbf{Z} \in \text{ri}(\mathcal{F}_i^* \cap \mathcal{F}_{i+1}^\perp)$ satisfies $\mathcal{F}_{i+1} = \mathcal{F}_i \cap \{\mathbf{Z}\}^\perp$. Conversely, any matrix $\mathbf{Z} \in \mathcal{F}_i^*$ satisfying $\mathcal{F}_{i+1} = \mathcal{F}_i \cap \{\mathbf{Z}\}^\perp$ belongs to $\text{ri}(\mathcal{F}_i^* \cap \mathcal{F}_{i+1}^\perp)$. The latter can be shown by the same argument as in the proof of Lemma 3.9. Hence, the matrix $\mathbf{Z}_i \in \text{ri}(\mathcal{F}_i^* \cap \mathcal{F}_{i+1}^\perp) \cap \mathcal{V}^\perp \cap \{\mathbf{Z} \in \mathcal{S}^n \mid \|\mathbf{Z}\|_F = 1\}$ can be chosen in order to minimize γ_i . The following example shows that the “deeper” $\mathbf{Z}_i$ lies in this set, the smaller the resulting γ_i becomes.

Recall that $\mathbf{E}_{11}$ is the matrix whose $(1, 1)$ th element is 1 and the other elements are 0. Let $\mathcal{F}_1 := \mathcal{S}_+^n$, $\mathcal{F}_2 := \mathbb{R}_+ \mathbf{E}_{11}$, and $\mathcal{V} := \{\mathbf{Z} \in \mathcal{S}^n \mid Z_{22} = \dots = Z_{nn} = 0\}$. Then, $\mathcal{F}_2$ is a face of $\mathcal{F}_1$, and we have

$$\text{ri}(\mathcal{F}_1^* \cap \mathcal{F}_2^\perp) \cap \mathcal{V}^\perp \cap \{\mathbf{Z} \in \mathcal{S}^n \mid \|\mathbf{Z}\|_F = 1\} = \{\text{Diag}(0, Z_{22}, \dots, Z_{nn}) \mid Z_{22}, \dots, Z_{nn} > 0, \sqrt{Z_{22}^2 + \dots + Z_{nn}^2} = 1\}.$$

For each $\mathbf{Z} \in \text{ri}(\mathcal{F}_1^* \cap \mathcal{F}_2^\perp) \cap \mathcal{V}^\perp \cap \{\mathbf{Z} \in \mathcal{S}^n \mid \|\mathbf{Z}\|_F = 1\}$, γ_1 becomes

$$\frac{2}{\min\{Z_{22}, \dots, Z_{nn}\}} + 1. \quad (3.39)$$

Let us consider (3.39) as a function of $Z_{22}, \dots, Z_{nn}$. It attains its global minimum at $Z_{22} = \dots = Z_{nn} = \frac{1}{\sqrt{n-1}}$ under the condition that $Z_{22}, \dots, Z_{nn} > 0$ and $\sqrt{Z_{22}^2 + \dots + Z_{nn}^2} = 1$.

We can interpret this result in a geometrical way. For each $Z_{22}, \dots, Z_{nn}$ such that $Z_{22}, \dots, Z_{nn} > 0$ and $\sqrt{Z_{22}^2 + \dots + Z_{nn}^2} = 1$, the smallest perturbation that makes the vector $(Z_{22}, \dots, Z_{nn})^\top$ not be in $\text{int } \mathbb{R}_+^{n-1}$ is

$$\inf_{\mathbf{a} \in \mathbb{R}_+^{n-1}} \{\|\mathbf{a}\|_2 \mid (Z_{22}, \dots, Z_{nn})^\top + \mathbf{a} \notin \text{int } \mathbb{R}_+^{n-1}\} = \min\{Z_{22}, \dots, Z_{nn}\},$$

which is maximized at $Z_{22} = \dots = Z_{nn} = \frac{1}{\sqrt{n-1}}$.

3.3 Qualitative error bound for the underlying semidefinite feasibility problem

In this subsection, by combining the ingredients derived in Sections 3.1 and 3.2, we obtain the final qualitative radial-type Hölder error bound for the problem $\text{Feas}(\mathcal{V}, \mathcal{S}_+^n)$. For each $i = 1, \dots, d^*$, let $\alpha_i s + \beta_i \sqrt{st}$ be the one-step facial residual function for $\mathcal{F}_i$ and $\mathbf{Z}_i$ derived in Proposition 3.10 and let

$$\hat{\psi}_i(s, t) := \begin{cases} \alpha_1 s + \beta_1 \sqrt{st} & (i = 1), \\ (2\alpha_i + 1)s + \beta_i \sqrt{2st} & (i \geq 2). \end{cases}$$

For each $i = 0, \dots, d^*$, we define $\phi_i: \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ recursively as follows:^{*2}

$$\phi_i(s, t) := \begin{cases} s & (i = 0), \\ \widehat{\psi}_1(s, t) & (i = 1), \\ \widehat{\psi}_i(s + \phi_{i-1}(s, t), t) & (i \geq 2). \end{cases} \quad (3.40)$$

When $i \geq 1$, the nonnegativity of $\phi_i(s, t)$ follows from that of $\widehat{\psi}_i(s, t)$. In the following lemma, we show that the function ϕ_i defined in (3.40) satisfies $\phi_i(s, t) \geq s$.

Lemma 3.13. *For each $i = 0, \dots, d^*$, we have $\phi_i(s, t) \geq s$ for all $s, t \in \mathbb{R}_+$.*

Proof. When $i = 0$, by definition, we have $\phi_0(s, t) = s$. When $i = 1$, it follows that

$$\phi_1(s, t) = \widehat{\psi}_1(s, t) = \alpha_1 s + \beta_1 \sqrt{st} \geq s,$$

where we use $\alpha_1 \geq 1$ and $\beta_1 \geq 0$ (see Remark 3.11) to derive the inequality. When $i \geq 2$, we see that

$$\phi_i(s, t) = \widehat{\psi}_i(s + \phi_{i-1}(s, t), t) = (2\alpha_i + 1)(s + \phi_{i-1}(s, t)) + \beta_i \sqrt{2(s + \phi_{i-1}(s, t))t} \geq s,$$

where we use $\alpha_i, \beta_i \geq 0$ and $\phi_{i-1}(s, t) \geq 0$ to derive the inequality. $\square$

In the following theorem, we show that the distance to the face $\mathcal{F}_{d^*+1}$ can be bounded using the function ϕ_{d^*} . To ease the notation, we recall the definition of κ in Theorem 3.6, and define

$$\mathcal{E}_b(\mathbf{X}) := \max\{\text{dist}(\mathbf{X}, \mathcal{V}), \text{dist}(\mathbf{X}, \mathcal{S}_+^n)\}. \quad (3.41)$$

Theorem 3.14. *Let the function ϕ_i be given as in (3.40) and $\mathcal{E}_b(\mathbf{X})$ be defined as in (3.41). Then, for every $i = 0, \dots, d^*$, we have*

$$\text{dist}(\mathbf{X}, \mathcal{F}_{i+1}) \leq \phi_i(\mathcal{E}_b(\mathbf{X}), \|\mathbf{X}\|_F) \text{ for all } \mathbf{X} \in \mathcal{S}^n. \quad (3.42)$$

In particular, it follows that

$$\text{dist}(\mathbf{X}, \mathcal{V} \cap \mathcal{S}_+^n) \leq \kappa(\|\mathbf{X}\|_F) \phi_{d^*}(\mathcal{E}_b(\mathbf{X}), \|\mathbf{X}\|_F) \text{ for all } \mathbf{X} \in \mathcal{S}^n. \quad (3.43)$$

Proof. First, we prove (3.42) by induction on $i \geq 0$. Firstly, it follows from the definition of ϕ_0 that

$$\text{dist}(\mathbf{X}, \mathcal{F}_1) = \text{dist}(\mathbf{X}, \mathcal{S}_+^n) \leq \mathcal{E}_b(\mathbf{X}) = \phi_0(\mathcal{E}_b(\mathbf{X}), \|\mathbf{X}\|_F) \quad (3.44)$$

holds for all $\mathbf{X} \in \mathcal{S}^n$, so the inequality in (3.42) holds for $i = 0$. Secondly, we consider the case $i = 1$. It follows from $\mathbf{Z}_1 \in \mathcal{V}^\perp$ that $\mathcal{V} \subseteq \{\mathbf{Z}_1\}^\perp$. Then we have

$$\langle \mathbf{Z}_1, \mathbf{X} \rangle \leq |\langle \mathbf{Z}_1, \mathbf{X} \rangle| = \text{dist}(\mathbf{X}, \{\mathbf{Z}_1\}^\perp) \stackrel{(a)}{\leq} \text{dist}(\mathbf{X}, \mathcal{V}) \stackrel{(b)}{\leq} \mathcal{E}_b(\mathbf{X}), \quad (3.45)$$

where we use $\mathcal{V} \subseteq \{\mathbf{Z}_1\}^\perp$ to derive (a) and use the definition of $\mathcal{E}_b(\mathbf{X})$ to derive (b). Combining (3.44) and (3.45) with the one-step facial residual function for $\mathcal{F}_1$ and $\mathbf{Z}_1$, we have

$$\text{dist}(\mathbf{X}, \mathcal{F}_2) \leq \alpha_1 \mathcal{E}_b(\mathbf{X}) + \beta_1 \sqrt{\mathcal{E}_b(\mathbf{X}) \|\mathbf{X}\|_F} = \widehat{\psi}_1(\mathcal{E}_b(\mathbf{X}), \|\mathbf{X}\|_F) = \phi_1(\mathcal{E}_b(\mathbf{X}), \|\mathbf{X}\|_F).$$

Therefore, the inequality in (3.42) holds for $i = 1$.

Now, we assume that the inequality in (3.42) holds for $i = k - 1 \in \{1, \dots, d^* - 1\}$. Let $\mathbf{X} \in \mathcal{S}^n$. Then we have

$$\text{dist}(P_{\text{span } \mathcal{F}_k}(\mathbf{X}), \mathcal{F}_k) \leq \|P_{\text{span } \mathcal{F}_k}(\mathbf{X}) - P_{\mathcal{F}_k}(\mathbf{X})\|_F \leq \text{dist}(\mathbf{X}, \text{span } \mathcal{F}_k) + \text{dist}(\mathbf{X}, \mathcal{F}_k) \leq 2 \text{dist}(\mathbf{X}, \mathcal{F}_k) + \mathcal{E}_b(\mathbf{X}), \quad (3.46)$$

where we use $P_{\mathcal{F}_k}(\mathbf{X}) \in \mathcal{F}_k$ to derive the first inequality, use the triangle inequality to derive the second inequality, and use $\mathcal{F}_k \subseteq \text{span } \mathcal{F}_k$ to derive the third inequality. Moreover, we have

$$\text{dist}(P_{\text{span } \mathcal{F}_k}(\mathbf{X}), \mathcal{V}) \leq \|P_{\text{span } \mathcal{F}_k}(\mathbf{X}) - P_{\mathcal{V}}(\mathbf{X})\|_F \leq \text{dist}(\mathbf{X}, \text{span } \mathcal{F}_k) + \text{dist}(\mathbf{X}, \mathcal{V}) \leq 2 \text{dist}(\mathbf{X}, \mathcal{F}_k) + \mathcal{E}_b(\mathbf{X}), \quad (3.47)$$

^{*2}The operation of constructing ϕ_i for $i \geq 2$ from $\widehat{\psi}_i$ and ϕ_{i-1} is called *diamond composition* in [29].

where we use $P_{\mathcal{V}}(\mathbf{X}) \in \mathcal{V}$ to derive the first inequality, use the triangle inequality to derive the second inequality, and use $\mathcal{F}_k \subseteq \text{span } \mathcal{F}_k$ and the definition of $\mathcal{E}_b(\mathbf{X})$ to derive the third inequality. By (3.47) and $\mathbf{Z}_k \in \mathcal{V}^\perp$, in a manner similar to that used in (3.45), it follows that

$$\langle \mathbf{Z}_k, P_{\text{span } \mathcal{F}_k}(\mathbf{X}) \rangle \leq \text{dist}(P_{\text{span } \mathcal{F}_k}(\mathbf{X}), \{\mathbf{Z}_k\}^\perp) \leq 2 \text{dist}(\mathbf{X}, \mathcal{F}_k) + \mathcal{E}_b(\mathbf{X}). \quad (3.48)$$

Combining (3.46) and (3.48) with the one-step facial residual function for $\mathcal{F}_k$ and $\mathbf{Z}_k$, we have

$$\begin{aligned} \text{dist}(P_{\text{span } \mathcal{F}_k}(\mathbf{X}), \mathcal{F}_{k+1}) &\leq \alpha_k(2 \text{dist}(\mathbf{X}, \mathcal{F}_k) + \mathcal{E}_b(\mathbf{X})) + \beta_k \sqrt{(2 \text{dist}(\mathbf{X}, \mathcal{F}_k) + \mathcal{E}_b(\mathbf{X})) \|P_{\text{span } \mathcal{F}_k}(\mathbf{X})\|_F} \\ &\leq \alpha_k(2 \text{dist}(\mathbf{X}, \mathcal{F}_k) + \mathcal{E}_b(\mathbf{X})) + \beta_k \sqrt{(2 \text{dist}(\mathbf{X}, \mathcal{F}_k) + \mathcal{E}_b(\mathbf{X})) \|\mathbf{X}\|_F}, \end{aligned} \quad (3.49)$$

where we use the fact that the projection to the subspace $\text{span } \mathcal{F}_k$ is nonexpansive to derive the second inequality. Then it follows that

$$\begin{aligned} \text{dist}(\mathbf{X}, \mathcal{F}_{k+1}) &\stackrel{(a)}{\leq} \|\mathbf{X} - P_{\mathcal{F}_{k+1}}(P_{\text{span } \mathcal{F}_k}(\mathbf{X}))\|_F \\ &\stackrel{(b)}{\leq} \text{dist}(\mathbf{X}, \text{span } \mathcal{F}_k) + \text{dist}(P_{\text{span } \mathcal{F}_k}(\mathbf{X}), \mathcal{F}_{k+1}) \\ &\stackrel{(c)}{\leq} \text{dist}(\mathbf{X}, \mathcal{F}_k) + \alpha_k(2 \text{dist}(\mathbf{X}, \mathcal{F}_k) + \mathcal{E}_b(\mathbf{X})) + \beta_k \sqrt{(2 \text{dist}(\mathbf{X}, \mathcal{F}_k) + \mathcal{E}_b(\mathbf{X})) \|\mathbf{X}\|_F} \\ &\leq (2\alpha_k + 1)(\text{dist}(\mathbf{X}, \mathcal{F}_k) + \mathcal{E}_b(\mathbf{X})) + \beta_k \sqrt{2(\text{dist}(\mathbf{X}, \mathcal{F}_k) + \mathcal{E}_b(\mathbf{X})) \|\mathbf{X}\|_F} \\ &\stackrel{(d)}{=} \hat{\psi}_k(\text{dist}(\mathbf{X}, \mathcal{F}_k) + \mathcal{E}_b(\mathbf{X}), \|\mathbf{X}\|_F) \\ &\stackrel{(e)}{\leq} \hat{\psi}_k(\mathcal{E}_b(\mathbf{X}) + \phi_{k-1}(\mathcal{E}_b(\mathbf{X}), \|\mathbf{X}\|_F), \|\mathbf{X}\|_F) \\ &\stackrel{(f)}{=} \phi_k(\mathcal{E}_b(\mathbf{X}), \|\mathbf{X}\|_F), \end{aligned}$$

where we use $P_{\mathcal{F}_{k+1}}(P_{\text{span } \mathcal{F}_k}(\mathbf{X})) \in \mathcal{F}_{k+1}$ to derive (a), use the triangle inequality to derive (b), use $\mathcal{F}_k \subseteq \text{span } \mathcal{F}_k$ and (3.49) to derive (c), use the definition of $\hat{\psi}_k$ to derive (d), use the inductive hypothesis to derive (e), and use the definition of ϕ_k to derive (f). Therefore, the inequality in (3.42) also holds for $i = k$.

Next, we prove (3.43). It follows that

$$\begin{aligned} \text{dist}(\mathbf{X}, \mathcal{V} \cap \mathcal{S}_+^n) &\stackrel{(a)}{=} \text{dist}(\mathbf{X}, \mathcal{V} \cap \mathcal{F}_{d^*+1}) \\ &\stackrel{(b)}{\leq} \kappa(\|\mathbf{X}\|_F) \max\{\text{dist}(\mathbf{X}, \mathcal{V}), \text{dist}(\mathbf{X}, \mathcal{F}_{d^*+1})\} \\ &\stackrel{(c)}{\leq} \kappa(\|\mathbf{X}\|_F) \max\{\mathcal{E}_b(\mathbf{X}), \phi_{d^*}(\mathcal{E}_b(\mathbf{X}), \|\mathbf{X}\|_F)\} \\ &\stackrel{(d)}{\leq} \kappa(\|\mathbf{X}\|_F) \phi_{d^*}(\mathcal{E}_b(\mathbf{X}), \|\mathbf{X}\|_F), \end{aligned}$$

where (a) holds because $\mathcal{V} \cap \mathcal{S}_+^n = \mathcal{V} \cap \mathcal{F}_{d^*+1}$, (b) follows from Theorem 3.6, (c) results from (3.42) and $\text{dist}(\mathbf{X}, \mathcal{V}) \leq \mathcal{E}_b(\mathbf{X})$, and (d) is a consequence of Lemma 3.13. $\square$

In what follows, we bound the function $\phi_{d^*}(s, t)$ by explicit power functions, thereby deriving a qualitative radial-type Hölder error bound for the problem $\text{Feas}(\mathcal{V}, \mathcal{S}_+^n)$. When $d^* = 0$, since $\phi_0(s, t) = s$ by definition, see (3.40), the error bound in (3.43) reduces to

$$\text{dist}(\mathbf{X}, \mathcal{V} \cap \mathcal{S}_+^n) \leq \kappa(\|\mathbf{X}\|_F) \mathcal{E}_b(\mathbf{X}) \text{ for all } \mathbf{X} \in \mathcal{S}^n, \quad (3.50)$$

which is a Lipschitz error bound.

When $d^* \geq 1$, the radial-type error bound in (3.43) is described in terms of the function ϕ_{d^*} , which is constructed recursively by using the functions $\hat{\psi}_1, \dots, \hat{\psi}_{d^*}$. In what follows, we derive upper bounds of the function $\phi_i(s, t)$ for every $i = 1, \dots, d^*$ by simple bivariate power functions with respect to s and t . To do this, for $i = 1, \dots, d^*$, we let

$$A_i := \begin{cases} \alpha_1 & (i = 1), \\ 2\alpha_i + 1 & (i \geq 2) \end{cases} \text{ and } B_i := \begin{cases} \beta_1 & (i = 1), \\ \sqrt{2}\beta_i & (i \geq 2). \end{cases}$$

Lemma 3.15. Let $d^* \geq 1$. For $i = 1, \dots, d^*$ and $j = 0, \dots, i$, we define $c_{i,j}$ recursively as

$$\begin{cases} c_{1,0} := A_1 = \alpha_1, \\ c_{1,1} := B_1 = \beta_1, \\ c_{i,0} := A_i(1 + c_{i-1,0}) & (i \geq 2), \\ c_{i,1} := A_i c_{i-1,1} + B_i(1 + \sqrt{c_{i-1,0}}) & (i \geq 2), \\ c_{i,j} := A_i c_{i-1,j} + B_i \sqrt{c_{i-1,j-1}} & (i \geq 2, 2 \leq j \leq i-1), \\ c_{i,i} := B_i \sqrt{c_{i-1,i-1}} & (i \geq 2). \end{cases}$$

Then, for every $i = 1, \dots, d^*$, we have

$$\phi_i(s, t) \leq \sum_{j=0}^i c_{i,j} s^{\frac{1}{2^j}} t^{1-\frac{1}{2^j}}. \quad (3.51)$$

Proof. We note that $\widehat{\psi}_i(s, t) = A_i s + B_i \sqrt{st}$ holds for every $i = 1, \dots, d^*$ by the definitions of A_i and B_i . We prove (3.51) by induction on $i \geq 1$. When $i = 1$, both the left- and right-hand sides of (3.51) are $\alpha_1 s + \beta_1 \sqrt{st}$, so (3.51) holds. Next, we assume that (3.51) holds for $i = k-1 \geq 1$. For $i = k$, we have

$$\begin{aligned} \phi_k(s, t) &= A_k(s + \phi_{k-1}(s, t)) + B_k \sqrt{(s + \phi_{k-1}(s, t))t} \\ &\stackrel{(a)}{\leq} A_k \left(s + \sum_{j=0}^{k-1} c_{k-1,j} s^{\frac{1}{2^j}} t^{1-\frac{1}{2^j}} \right) + B_k \sqrt{\left(s + \sum_{j=0}^{k-1} c_{k-1,j} s^{\frac{1}{2^j}} t^{1-\frac{1}{2^j}} \right) t} \\ &\stackrel{(b)}{\leq} A_k s + \sum_{j=0}^{k-1} A_k c_{k-1,j} s^{\frac{1}{2^j}} t^{1-\frac{1}{2^j}} + B_k \sqrt{st} + \sum_{j=0}^{k-1} B_k \sqrt{c_{k-1,j}} s^{\frac{1}{2^{j+1}}} t^{1-\frac{1}{2^{j+1}}} \\ &= A_k(1 + c_{k-1,0})s + (A_k c_{k-1,1} + B_k(1 + \sqrt{c_{k-1,0}}))\sqrt{st} \\ &\quad + \sum_{j=2}^{k-1} (A_k c_{k-1,j} + B_k \sqrt{c_{k-1,j-1}}) s^{\frac{1}{2^j}} t^{1-\frac{1}{2^j}} + B_k \sqrt{c_{k-1,k-1}} s^{\frac{1}{2^k}} t^{1-\frac{1}{2^k}} \\ &\stackrel{(c)}{=} \sum_{j=0}^k c_{k,j} s^{\frac{1}{2^j}} t^{1-\frac{1}{2^j}}, \end{aligned}$$

where (a) is a consequence of the inductive hypothesis, (b) holds by the subadditivity of the square root function, and (c) follows from the definition of $c_{k,j}$. Thus, (3.51) also holds for $i = k$. $\square$

In the following lemma, we provide an explicit upper bound for the constants $c_{i,j}$ in (3.51) using both $\alpha_1, \dots, \alpha_i$ and $\beta_1, \dots, \beta_i$.

Lemma 3.16. Let $d^* \geq 1$. For each $i = 1, \dots, d^*$, we define

$$\alpha^{(i)} := \max_{1 \leq k \leq i} \alpha_k, \quad \beta^{(i)} := \max_{1 \leq k \leq i} \beta_k.$$

Then, for $i = 1, \dots, d^*$ and $j = 0, \dots, i$, we have

$$c_{i,j} \leq 6^{i-1} (\alpha^{(i)})^{i-j} (\beta^{(i)})^{2-\frac{1}{2^{j-1}}}. \quad (3.52)$$

Proof. For simplicity, let $p_j := 2 - \frac{1}{2^{j-1}}$. We prove that (3.52) holds for any $j = 0, \dots, i$ by induction on $i \geq 1$. For $i = 1$, we have

$$\begin{aligned} c_{1,0} &= \alpha_1 = 6^{1-1} (\alpha^{(1)})^{1-0} (\beta^{(1)})^{p_0}, \\ c_{1,1} &= \beta_1 = 6^{1-1} (\alpha^{(1)})^{1-1} (\beta^{(1)})^{p_1}, \end{aligned}$$

so (3.52) holds for $i = 1$ and $j = 0, 1$.

In what follows, we suppose that $d^* \geq 2$, because the proof finishes at the previous paragraph when $d^* = 1$. Under the inductive hypothesis that (3.52) holds for $i = k-1 \geq 1$ and $j = 0, \dots, k-1$, we shall prove that (3.52)

holds for $i = k$ and $j = 0, \dots, k$. The discussion in Remark 3.11 tells us that if $d^* \geq 2$ holds, then $\mathcal{V} \cap \mathcal{S}_+^n \neq \{\mathbf{O}\}$ holds as well. Furthermore, in this case, $\alpha_i, \beta_i \geq 1$ hold for all $i = 1, \dots, d^*$. By the definitions of A_k and B_k , we have

$$A_k \leq 3\alpha_k, \quad B_k \leq \sqrt{2}\beta_k. \quad (3.53)$$

Combining $\alpha_i, \beta_i \geq 1$ with the recursive definition of $c_{k-1,j}$, we see that

$$c_{k-1,j} \geq 1 \quad (3.54)$$

for any $j = 0, \dots, k-1$.

Firstly, we have

$$c_{k,0} = A_k(1 + c_{k-1,0}) \leq 6\alpha_k c_{k-1,0} \leq 6^{k-1} \alpha_k (\alpha^{(k-1)})^{k-1} (\beta^{(k-1)})^{p_0} \leq 6^{k-1} (\alpha^{(k)})^k (\beta^{(k)})^{p_0},$$

where we use (3.53) and (3.54) to derive the first inequality, use the inductive hypothesis to derive the second inequality, and use $\alpha_k, \alpha^{(k-1)} \leq \alpha^{(k)}$ to derive the third inequality. Secondly, we have

$$\begin{aligned} c_{k,1} &= A_k c_{k-1,1} + B_k(1 + \sqrt{c_{k-1,0}}) \\ &\leq 3\alpha_k c_{k-1,1} + 2\sqrt{2}\beta_k \sqrt{c_{k-1,0}} \\ &\leq 3 \cdot 6^{k-2} \alpha_k (\alpha^{(k-1)})^{k-2} (\beta^{(k-1)})^{p_1} + 2\sqrt{2}\beta_k \sqrt{6^{k-2} (\alpha^{(k-1)})^{k-1} (\beta^{(k-1)})^{p_0}} \\ &\leq (3 + 2\sqrt{2}) 6^{k-2} (\alpha^{(k)})^{k-1} (\beta^{(k)})^{p_1} \\ &\leq 6^{k-1} (\alpha^{(k)})^{k-1} (\beta^{(k)})^{p_1}, \end{aligned}$$

where the first inequality holds by (3.53) and (3.54), the second inequality follows by the inductive hypothesis, and the third inequality holds because $\alpha_k \leq \alpha^{(k)}$, $\beta_k \leq \beta^{(k)}$, $(\alpha^{(i)})_i$ and $(\beta^{(i)})_i$ are nondecreasing sequences. Thirdly, for every $j = 2, \dots, i-1$, we have

$$\begin{aligned} c_{k,j} &= A_k c_{k-1,j} + B_k \sqrt{c_{k-1,j-1}} \\ &\leq 3\alpha_k c_{k-1,j} + \sqrt{2}\beta_k \sqrt{c_{k-1,j-1}} \\ &\leq 3 \cdot 6^{k-2} \alpha_k (\alpha^{(k-1)})^{k-1-j} (\beta^{(k-1)})^{p_j} + \sqrt{2}\beta_k \sqrt{6^{k-2} (\alpha^{(k-1)})^{k-j} (\beta^{(k-1)})^{p_{j-1}}} \\ &\leq (3 + \sqrt{2}) 6^{k-2} (\alpha^{(k)})^{k-j} (\beta^{(k)})^{p_j} \\ &\leq 6^{k-1} (\alpha^{(k)})^{k-j} (\beta^{(k)})^{p_j}, \end{aligned}$$

where we use $\alpha_k, \alpha^{(k-1)} \leq \alpha^{(k)}$, $\beta_k, \beta^{(k-1)} \leq \beta^{(k)}$, and $p_j = 1 + p_{j-1}/2$ to derive the third inequality. Fourthly, we have

$$c_{k,k} = B_k \sqrt{c_{k-1,k-1}} \leq \sqrt{2}\beta_k \sqrt{6^{k-2} (\beta^{(k-1)})^{p_{k-1}}} \leq 6^{k-1} (\beta^{(k)})^{p_k},$$

where we use $\beta_k, \beta^{(k-1)} \leq \beta^{(k)}$ and $p_k = 1 + p_{k-1}/2$ to derive the final inequality. Therefore, (3.52) with $i = k$ also holds for any $j = 0, \dots, k$. $\square$

Applying Lemmas 3.15 and 3.16 to the error bound in (3.43), we obtain a radial-type Hölder error bound of the form (1.2) involving $\alpha_1, \dots, \alpha_{d^*}$ and $\beta_1, \dots, \beta_{d^*}$.

Theorem 3.17 (Radial-type error bound over the whole space). *Let $d^* \geq 1$. Then*

$$\text{dist}(\mathbf{X}, \mathcal{V} \cap \mathcal{S}_+^n) \leq \kappa(\|\mathbf{X}\|_F) 6^{d^*-1} \sum_{j=0}^{d^*} (\alpha^{(d^*)})^{d^*-j} (\beta^{(d^*)})^{2-\frac{1}{2^{j-1}}} \mathcal{E}_b(\mathbf{X})^{\frac{1}{2^j}} \|\mathbf{X}\|_F^{1-\frac{1}{2^j}} \text{ for all } \mathbf{X} \in \mathcal{S}^n.$$

In other words, a radial-type Hölder error bound in (1.2) holds with

$$\mu_j(\rho) := 6^{d^*-1} (\alpha^{(d^*)})^{d^*-j} (\beta^{(d^*)})^{2-\frac{1}{2^{j-1}}} \kappa(\rho) \rho^{1-\frac{1}{2^j}} \text{ and } \gamma_j = \frac{1}{2^j} \text{ for every } j = 0, \dots, \ell, \text{ and } \ell = d^*.$$

Next, we show that the function ϕ_i can be alternatively bounded by another simpler bivariate power function with fractional exponents, up to an explicit constant C_i defined recursively.

Lemma 3.18. *Let $d^* \geq 1$. For each $i = 1, \dots, d^*$, we define*

$$C_i := \begin{cases} \alpha_1 & (i = 1), \\ 2(C_{i-1} + 1)(2\alpha_i + 1) & (i \geq 2). \end{cases}$$

Then we have $1 \leq C_i \leq 12^{i-1}\alpha_1 \cdots \alpha_i$ and

$$\phi_i(s, t) \leq C_i \sum_{j=0}^i s^{\frac{1}{2^j}} t^{1-\frac{1}{2^j}}. \quad (3.55)$$

Proof. First, we derive bounds for the constant C_i . By Remark 3.11, $\alpha_i \geq 1$ holds for $i = 1, \dots, d^*$. Then, the definition of C_i implies $C_i \geq 1$ for $i = 1, \dots, d^*$.

We next prove $C_i \leq 12^{i-1}\alpha_1 \cdots \alpha_i$ by induction on $i \geq 1$. When $i = 1$ both the left- and right-hand sides are α_1 , so the inequality holds. We assume that the inequality holds for $i = k-1 \geq 1$. For $i = k$, it follows from the definition of C_k that

$$C_k = 2(C_{k-1} + 1)(2\alpha_k + 1) \leq 12C_{k-1}\alpha_k \leq 12^{k-1}\alpha_1 \cdots \alpha_k,$$

where we use $C_{k-1} \geq 1$ (and so, $C_{k-1} + 1 \leq 2C_{k-1}$) and $\alpha_k \geq 1$ (and so, $2\alpha_k + 1 \leq 3\alpha_k$) to derive the first inequality, and use the inductive hypothesis to derive the second inequality. Therefore, $C_i \leq 12^{i-1}\alpha_1 \cdots \alpha_i$ holds for $i = k$.

Next, we show (3.55) by induction on $i \geq 1$. When $i = 1$, it follows from $\alpha_1 \geq \beta_1$ (Remark 3.11) that

$$\phi_1(s, t) = \widehat{\psi}_1(s, t) = \alpha_1 s + \beta_1 \sqrt{st} \leq C_1 \sum_{j=0}^1 s^{\frac{1}{2^j}} t^{1-\frac{1}{2^j}},$$

so (3.55) holds. We assume that (3.55) holds for $i = k-1 \geq 1$. Recalling (3.40), when $i = k$, we have

$$\begin{aligned} \phi_k(s, t) &= (2\alpha_k + 1)(s + \phi_{k-1}(s, t)) + \beta_k \sqrt{2(s + \phi_{k-1}(s, t))t} \\ &\leq (2\alpha_k + 1) \left((C_{k-1} + 1)s + C_{k-1} \sum_{j=1}^{k-1} s^{\frac{1}{2^j}} t^{1-\frac{1}{2^j}} \right) + \alpha_k \sqrt{2 \left((C_{k-1} + 1)s + C_{k-1} \sum_{j=1}^{k-1} s^{\frac{1}{2^j}} t^{1-\frac{1}{2^j}} \right) t} \\ &\leq (C_{k-1} + 1)(2\alpha_k + 1)s + C_{k-1}(2\alpha_k + 1) \sum_{j=1}^{k-1} s^{\frac{1}{2^j}} t^{1-\frac{1}{2^j}} \\ &\quad + \sqrt{2(C_{k-1} + 1)\alpha_k} \sqrt{st} + \sqrt{2C_{k-1}\alpha_k} \sum_{j=1}^{k-1} s^{\frac{1}{2^{j+1}}} t^{1-\frac{1}{2^{j+1}}} \\ &= (C_{k-1} + 1)(2\alpha_k + 1)s + \underbrace{(C_{k-1}(2\alpha_k + 1))}_{\leq (C_{k-1}+1)(2\alpha_k+1)} + \underbrace{\sqrt{2(C_{k-1} + 1)\alpha_k}}_{\leq (C_{k-1}+1)(2\alpha_k+1)} \sqrt{st} \\ &\quad + \underbrace{(C_{k-1}(2\alpha_k + 1))}_{\leq (C_{k-1}+1)(2\alpha_k+1)} + \underbrace{\sqrt{2C_{k-1}\alpha_k}}_{\leq (C_{k-1}+1)(2\alpha_k+1)} \sum_{j=2}^{k-1} s^{\frac{1}{2^j}} t^{1-\frac{1}{2^j}} + \sqrt{2C_{k-1}\alpha_k} s^{\frac{1}{2^k}} t^{1-\frac{1}{2^k}} \\ &\leq C_k \sum_{j=0}^k s^{\frac{1}{2^j}} t^{1-\frac{1}{2^j}}, \end{aligned}$$

where the first inequality is a consequence of the inductive hypothesis and $\alpha_k \geq \beta_k$, the second inequality holds by the subadditivity of the square root function, and the third inequality follows from the fact that $C_{k-1} \geq 1$ holds, which implies that the coefficients of $s^{\frac{1}{2^j}} t^{1-\frac{1}{2^j}}$ for every $j = 0, \dots, k$ are bounded by $2(C_{k-1} + 1)(2\alpha_k + 1)$. Therefore, (3.55) also holds for $i = k$. $\square$

Applying Lemma 3.18 to the error bound in (3.43), we obtain an alternative radial-type Hölder error bound of the form (1.2) involving only $\alpha_1, \dots, \alpha_{d^*}$. We note that this form is, in general, independent and simpler than the one given in Theorem 3.17.

Theorem 3.19 (Alternative radial-type error bound over the whole space). *Let $d^* \geq 1$. Then*

$$\text{dist}(\mathbf{X}, \mathcal{V} \cap \mathcal{S}_+^n) \leq \kappa(\|\mathbf{X}\|_F) 12^{d^*-1} \alpha_1 \cdots \alpha_{d^*} \sum_{j=0}^{d^*} \mathcal{E}_b(\mathbf{X})^{\frac{1}{2^j}} \|\mathbf{X}\|_F^{1-\frac{1}{2^j}} \text{ for all } \mathbf{X} \in \mathcal{S}^n.$$

In other words, the radial-type Hölder error bound in (1.2) holds with

$$\mu_j(\rho) := 12^{d^*-1} \alpha_1 \cdots \alpha_{d^*} \kappa(\rho) \rho^{1-\frac{1}{2^j}} \text{ and } \gamma_j = \frac{1}{2^j} \text{ for every } j = 0, \dots, \ell, \text{ and } \ell = d^*.$$

Next, for the ease of comparison and discussion for asymptotic tightness in Section 4, we summarize and simplify the derived error bounds in the corollary below, according to the value of $d^* \in \{0, 1, \dots, n-1\}$. In particular, in line with Sturm's error bound [48], we consider the radial-type error bound where the test set $\mathcal{R}_n$ is given by

$$\mathcal{R}_n := \{\mathbf{X} \in \mathcal{S}^n \mid \mathcal{E}_b(\mathbf{X}) \leq 1\}. \quad (3.56)$$

Naturally, the conclusion can be extended to test sets of the form $\{\mathbf{X} \in \mathcal{S}^n \mid \mathcal{E}_b(\mathbf{X}) \leq \lambda\}$ for some $\lambda > 0$.

Corollary 3.20 (Radial-type error bounds over test sets with a fixed upper bound of the backward error). *For all $\mathbf{X} \in \mathcal{R}_n$, we have*

$$\text{dist}(\mathbf{X}, \mathcal{V} \cap \mathcal{S}_+^n) \leq \begin{cases} \kappa(\|\mathbf{X}\|_F) \mathcal{E}_b(\mathbf{X}) & (d^* = 0), \\ \kappa(\|\mathbf{X}\|_F) (\alpha_1 \mathcal{E}_b(\mathbf{X}) + \beta_1 \sqrt{\mathcal{E}_b(\mathbf{X}) \|\mathbf{X}\|_F}) & (d^* = 1), \\ \kappa(\|\mathbf{X}\|_F) 12^{d^*-1} \alpha_1 \cdots \alpha_{d^*} (d^* \max\{\|\mathbf{X}\|_F^{1-\frac{1}{2^{d^*}}}, \sqrt{\|\mathbf{X}\|_F}\} + 1) \mathcal{E}_b(\mathbf{X})^{\frac{1}{2^{d^*}}} & (d^* \geq 2), \end{cases} \quad (3.57)$$

where the function κ is given as in Theorem 3.6.

Proof. The result follows from (3.50) when $d^* = 0$ and from Theorem 3.17 when $d^* = 1$. In what follows, we consider the case where $d^* \geq 2$. Let $\mathbf{X} \in \mathcal{R}_n$ be arbitrary. It follows from $\mathcal{E}_b(\mathbf{X}) \leq 1$ that

$$\sum_{j=0}^{d^*} \mathcal{E}_b(\mathbf{X})^{\frac{1}{2^j}} \|\mathbf{X}\|_F^{1-\frac{1}{2^j}} \leq \left(\sum_{j=0}^{d^*} \|\mathbf{X}\|_F^{1-\frac{1}{2^j}} \right) \mathcal{E}_b(\mathbf{X})^{\frac{1}{2^{d^*}}} = \left(\sum_{j=1}^{d^*} \|\mathbf{X}\|_F^{1-\frac{1}{2^j}} + 1 \right) \mathcal{E}_b(\mathbf{X})^{\frac{1}{2^{d^*}}}. \quad (3.58)$$

If $\|\mathbf{X}\|_F \geq 1$, then since $\|\mathbf{X}\|_F^{1-\frac{1}{2^j}} \leq \|\mathbf{X}\|_F^{1-\frac{1}{2^{d^*}}}$ for all $j = 1, \dots, d^*$, we have

$$\sum_{j=1}^{d^*} \|\mathbf{X}\|_F^{1-\frac{1}{2^j}} \leq d^* \|\mathbf{X}\|_F^{1-\frac{1}{2^{d^*}}}. \quad (3.59)$$

If $\|\mathbf{X}\|_F < 1$, then since $\|\mathbf{X}\|_F^{1-\frac{1}{2^j}} \leq \sqrt{\|\mathbf{X}\|_F}$ for all $j = 1, \dots, d^*$, we have

$$\sum_{j=1}^{d^*} \|\mathbf{X}\|_F^{1-\frac{1}{2^j}} \leq d^* \sqrt{\|\mathbf{X}\|_F}. \quad (3.60)$$

By (3.59) and (3.60), we have

$$(3.58) \leq (d^* \max\{\|\mathbf{X}\|_F^{1-\frac{1}{2^{d^*}}}, \sqrt{\|\mathbf{X}\|_F}\} + 1) \mathcal{E}_b(\mathbf{X})^{\frac{1}{2^{d^*}}}.$$

Therefore, the conclusion directly follows from Theorem 3.19. $\square$

Remark 3.21 (Comparison with Sturm's (local) Hölder error bound). Let d_s be the singularity degree of the semidefinite feasibility problem $\text{Feas}(\mathcal{V}, \mathcal{S}_+^n)$. In a pioneering work, Sturm [48] showed that for any $\rho > 0$ there exists $\alpha_\rho > 0$ such that

$$\text{dist}(\mathbf{X}, \mathcal{V} \cap \mathcal{S}_+^n) \leq \alpha_\rho \epsilon^{\frac{1}{2^{d_s}}} \text{ for any } \mathbf{X} \in \mathcal{S}^n \text{ with } \|\mathbf{X}\|_F \leq \rho \text{ and } \mathcal{E}_b(\mathbf{X}) \leq \epsilon,$$

where $\epsilon \in [0, 1]$.

On the other hand, Corollary 3.20 implies that for any $\mathbf{X} \in \mathcal{S}^n$ with $\|\mathbf{X}\|_F \leq \rho$ and $\mathcal{E}_b(\mathbf{X}) \leq \epsilon$ where $\epsilon \in [0, 1]$, one has

$$\text{dist}(\mathbf{X}, \mathcal{V} \cap \mathcal{S}_+^n) \leq \alpha_\rho \epsilon^{\frac{1}{2^{d^*}}},$$

where α_ρ is given by

$$\alpha_\rho = \begin{cases} \kappa(\rho) & (d^* = 0), \\ \kappa(\rho)(\alpha_1 + \beta_1 \sqrt{\rho}) & (d^* = 1), \\ \kappa(\rho) 12^{d^*-1} \alpha_1 \cdots \alpha_{d^*} (d^* \max\{\rho^{1-\frac{1}{2^{d^*}}}, \sqrt{\rho}\} + 1) & (d^* \geq 2). \end{cases}$$

Since $d^* = d_{\text{PPS}}(\mathcal{V}, \mathcal{S}_+^n) \leq d_s$ always holds, we complement Sturm's result by providing explicit expressions for the constant α_ρ .

4 Asymptotic tightness of the derived radial-type error bound

In this section, we discuss the asymptotic tightness of the derived error bound. For ease of comparison, we focus on the most explicit form of radial-type error bounds given in (3.57). We consider three cases depending on the quantity $d^* = d_{\text{PPS}}(\mathcal{V}, \mathcal{S}_+^n)$, the distance to the PPS condition, of the problem $\text{Feas}(\mathcal{V}, \mathcal{S}_+^n)$.

We present two complementary perspectives on this aspect. On the one hand, we first examine the cases where $d^* \in \{0, 1\}$, for which d^* is independent of the ambient dimension. In these cases, we provide families of instances where the error bound derived in the previous section is asymptotically tight up to a dimension-free constant. On the other hand, when $d^* = n - 1$, corresponding to the most singular case, we identify a family of instances with varying dimension for which the derived error bound need not be asymptotically tight, as we will see at the end of the section.

4.1 The case $d^* = 0$

In this subsection, we examine a family of semidefinite feasibility problems $\text{Feas}(\mathcal{V}, \mathcal{S}_+^n)$ with $d^* := d_{\text{PPS}}(\mathcal{V}, \mathcal{S}_+^n) = 0$ for which the error bound in (3.57) is asymptotically tight with respect to a dimension-free constant.

Let $n \in \mathbb{Z}_{\geq 2}$ and

$$\mathcal{V}_n := \{\text{Diag}(-t+1, t+1, \dots, t+1) \mid t \in \mathbb{R}\}.$$

Noting that

$$\begin{aligned} \mathcal{V}_n \cap \mathcal{S}_+^n &= \{\text{Diag}(-t+1, t+1, \dots, t+1) \mid -1 \leq t \leq 1\}, \\ \mathcal{V}_n \cap \text{int } \mathcal{S}_+^n &= \{\text{Diag}(-t+1, t+1, \dots, t+1) \mid -1 < t < 1\}, \end{aligned} \quad (4.1)$$

we see that the problem $\text{Feas}(\mathcal{V}_n, \mathcal{S}_+^n)$ satisfies Slater's condition, so $d^* = 0$ holds. Moreover, Corollary 3.20 implies that $\text{dist}(\mathbf{X}, \mathcal{V}_n \cap \mathcal{S}_+^n) \leq \kappa(\|\mathbf{X}\|_F) \mathcal{E}_b(\mathbf{X})$ for all $\mathbf{X} \in \mathcal{R}_n$. For convenience, let $n_k := k + 1$ and

$$\mathbf{X}_k := \text{Diag}\left(-\frac{1}{k+1}, 2 + \frac{1}{k+1}, \dots, 2 + \frac{1}{k+1}\right) \in \mathcal{S}^{n_k}$$

for each $k \in \mathbb{Z}_{\geq 1}$.

For this example, the radial modulus function κ scales with the norm of the matrix (and so, cannot be a constant function). Using (2.1), we have $\text{dist}(\mathbf{X}_k, \mathcal{V}_{n_k}) = 0$ and $\text{dist}(\mathbf{X}_k, \mathcal{S}_+^{n_k}) = \frac{1}{k+1}$. In particular, $\mathcal{E}_b(\mathbf{X}_k) = \frac{1}{k+1} \leq 1$. Moreover, by (4.1),

$$\text{dist}(\mathbf{X}_k, \mathcal{V}_{n_k} \cap \mathcal{S}_+^{n_k}) = \frac{1}{\sqrt{k+1}}. \quad (4.2)$$

Thus, for an error bound as in the first case of (3.57) to hold, $\kappa(\|\mathbf{X}_k\|_F)$ must be at least on the order of $\sqrt{k}$.

Next, let us compute an upper bound for $\kappa(\|\mathbf{X}_k\|_F)$ according to (3.22). It follows from $\mathcal{V}_{n_k} \subseteq \text{span } \mathcal{S}_+^{n_k}$ and Remark 2.3 that $\max\{\eta(\mathcal{V}_{n_k}, \text{span } \mathcal{S}_+^{n_k}), 1\} = 1$. In addition, letting $\mathbf{U}_0 := \mathbf{I}_{n_k} \in \mathcal{V}_{n_k} \cap \mathcal{S}_+^{n_k}$ and using the bound in (3.17), we have

$$\theta(\|\mathbf{X}_k\|_F) \leq \bar{\theta}(\|\mathbf{X}_k\|_F; \mathbf{U}_0) = 1 + 2\sqrt{k+1} + 2\sqrt{4(k+1) - \frac{3}{k+1}}.$$

It then follows from (3.22) that

$$\kappa(\|\mathbf{X}_k\|_F) \leq \bar{\kappa}(\|\mathbf{X}_k\|_F; \mathbf{U}_0) = 1 + 2\sqrt{k+1} + 2\sqrt{4(k+1) - \frac{3}{k+1}}.$$

Thus, combining this with the fact that $\kappa(\|\mathbf{X}_k\|_F)$ is at least on the order of $\sqrt{k}$, we conclude that $\kappa(\|\mathbf{X}_k\|_F)$ grows proportionally to $\sqrt{k}$.

Next, we show that the error bound estimate is asymptotically tight. Recall that $\mathcal{E}_b(\mathbf{X}_k) = \frac{1}{k+1}$. Then the right-hand side of (3.57) is bounded by

$$\kappa(\|\mathbf{X}_k\|_F) \mathcal{E}_b(\mathbf{X}_{n_k}) \leq \left(1 + 2\sqrt{k+1} + 2\sqrt{4(k+1) - \frac{3}{k+1}}\right) \frac{1}{k+1}. \quad (4.3)$$

It then follows from (4.2) and (4.3) that

$$\liminf_{k \rightarrow \infty} \frac{\text{dist}(\mathbf{X}_k, \mathcal{V}_{n_k} \cap \mathcal{S}_+^{n_k})}{\kappa(\|\mathbf{X}_k\|_F) \mathcal{E}_b(\mathbf{X}_k)} \geq \liminf_{k \rightarrow \infty} \frac{\sqrt{k+1}}{1 + 2\sqrt{k+1} + 2\sqrt{4(k+1) - \frac{3}{k+1}}} = \frac{1}{6}.$$

Thus, the derived error bound is asymptotically tight up to the dimension-free constant 1/6 for this instance.

4.2 The case $d^* = 1$

For every $n \in \mathbb{Z}_{\geq 3}$, we define

$$\mathcal{V}_n := \left\{ \mathbf{A} \in \mathcal{S}^n \mid \sum_{i=1}^n A_{ii} = 1, \sum_{i=3}^n A_{ii} = 0 \right\}. \quad (4.4)$$

We note that $\mathcal{V}_n \cap \mathcal{S}_+^n = \{\text{Diag}(\mathbf{A}, \mathbf{O}) \in \mathcal{S}^n \mid \mathbf{A} \in \mathcal{S}_+^2, A_{11} + A_{22} = 1\}$. The cone $\mathcal{S}_+^n$ is not polyhedral since $n \geq 3$. Furthermore, $\mathcal{V}_n \cap \text{int } \mathcal{S}_+^n$ is empty. Let

$$\begin{aligned} \mathbf{Z}_1 &:= \text{Diag}(0, 0, 1, \dots, 1) \in \mathcal{S}_+^n \cap \mathcal{V}_n^\perp, \\ \mathcal{F}_2 &:= \mathcal{S}_+^n \cap \{\mathbf{Z}_1\}^\perp = \mathcal{S}_+^2 \oplus \{0\}^{n-2}. \end{aligned} \quad (4.5)$$

Then, the problem $\text{Feas}(\mathcal{V}_n, \mathcal{F}_2)$ satisfies the PPS condition because $\mathcal{V}_n \cap \text{ri } \mathcal{F}_2 \neq \emptyset$. Therefore, $d^* = 1$ holds. For this problem, we show that the error bound in (3.57) is asymptotically tight up to a dimension-free constant.

For every $k \in \mathbb{Z}_{\geq 1}$, let $n_k := k + 3$ and

$$\mathbf{X}_k := \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{k} & \frac{1}{k} & \cdots & \frac{1}{k} & \frac{1}{k} \\ 0 & 0 & \frac{1}{k} & 0 & \frac{1}{k} & \cdots & \frac{1}{k} & \frac{1}{k} \\ 0 & 0 & \frac{1}{k} & \frac{1}{k} & 0 & \cdots & \frac{1}{k} & \frac{1}{k} \\ 0 & 0 & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \frac{1}{k} & \frac{1}{k} & \frac{1}{k} & \cdots & 0 & \frac{1}{k} \\ 0 & 0 & \frac{1}{k} & \frac{1}{k} & \frac{1}{k} & \cdots & \frac{1}{k} & 0 \end{pmatrix} \in \mathcal{S}^{n_k}. \quad (4.6)$$

Since the upper left 2×2 block of $\mathbf{X}_k$ is positive semidefinite and $(\mathbf{X}_k)_{11} + (\mathbf{X}_k)_{22} = 1$ holds, the left-hand side of (3.57) is

$$\text{dist}(\mathbf{X}_k, \mathcal{V}_{n_k} \cap \mathcal{S}_+^{n_k}) = \sqrt{\sum_{i,j=3}^{n_k} (\mathbf{X}_k)_{ij}^2} = \sqrt{\frac{k+1}{k}} \rightarrow 1 \quad (\text{as } k \rightarrow \infty). \quad (4.7)$$

We next evaluate the right-hand side of (3.57). Since $\mathbf{X}_k \in \mathcal{V}_{n_k}$ and $\mathbf{X}_k$ has eigenvalues 1 with multiplicity 2, 0 with multiplicity 1, and $-1/k$ with multiplicity k , by (2.1), we have $\text{dist}(\mathbf{X}_k, \mathcal{V}_{n_k}) = 0$ and $\text{dist}(\mathbf{X}_k, \mathcal{S}_+^{n_k}) = 1/\sqrt{k}$. Therefore, $\mathcal{E}_b(\mathbf{X}_k) = 1/\sqrt{k}$. Moreover, by choosing $\mathbf{X}^* := \text{Diag}(1/2, 1/2, 0, \dots, 0) \in \mathcal{V}_{n_k} \cap \text{ri } \mathcal{F}_2$, it follows from (3.22) that $\kappa(\|\mathbf{X}_k\|_F) \leq \bar{\kappa}(\|\mathbf{X}_k\|_F; \mathbf{X}^*)$. As shown in Appendix C, we have $\bar{\kappa}(\rho; \mathbf{X}^*) = \sqrt{2}(1 + 2\sqrt{2} + 4\rho)$, $\alpha_1 =$

$\sqrt{n+8\sqrt{n-2}+15}$, and $\beta_1 = \sqrt{2\sqrt{2(n-2)}+6\sqrt{2}}$, where $n = n_k = k+3$. Hence, by combining these estimates, the right-hand side of (3.57) at $\mathbf{X} = \mathbf{X}_k$ is bounded by

$$\begin{aligned} & \kappa(\|\mathbf{X}_k\|_F)(\alpha_1 \mathcal{E}_b(\mathbf{X}_k) + \beta_1 \sqrt{\mathcal{E}_b(\mathbf{X}_k)\|\mathbf{X}_k\|_F}) \\ & \leq \bar{\kappa}(\|\mathbf{X}_k\|_F; \mathbf{X}^*)(\alpha_1 \mathcal{E}_b(\mathbf{X}_k) + \beta_1 \sqrt{\mathcal{E}_b(\mathbf{X}_k)\|\mathbf{X}_k\|_F}) \\ & \leq \sqrt{2} \left(1 + 2\sqrt{2} + 4\sqrt{\frac{2k+1}{k}} \right) \left(\sqrt{\frac{k+8\sqrt{k+1}+18}{k}} + \sqrt{\frac{(2\sqrt{2(k+1)}+6\sqrt{2})\sqrt{2k+1}}{k}} \right) \\ & \rightarrow 3(12+\sqrt{2}) \quad (\text{as } k \rightarrow \infty). \end{aligned} \tag{4.8}$$

It follows from (4.7) and (4.8) that

$$\liminf_{k \rightarrow \infty} \frac{\text{dist}(\mathbf{X}_k, \mathcal{V}_{n_k} \cap \mathcal{S}_+^{n_k})}{\kappa(\|\mathbf{X}_k\|_F)(\alpha_1 \mathcal{E}_b(\mathbf{X}_k) + \beta_1 \sqrt{\mathcal{E}_b(\mathbf{X}_k)\|\mathbf{X}_k\|_F})} \geq \frac{1}{3(12+\sqrt{2})}.$$

Thus, in this instance, the error bound in (3.57) is asymptotically tight up to the dimension-free constant $\frac{1}{3(12+\sqrt{2})}$.

We will revisit this instance in Example 5.4 as the set of optimal solutions of an SDP satisfying the strict complementarity condition.

4.3 The case $d^* = n - 1$

For every $n \in \mathbb{Z}_{\geq 3}$, let

$$\mathcal{V}_n := \{\mathbf{A} \in \mathcal{S}^n \mid A_{nn} = 0, A_{ii} = A_{1,i+1} \ (i = 2, \dots, n-1)\}. \tag{4.9}$$

The vector space $\mathcal{V}_n$ defined in (4.9) is obtained from the instance in [48, Example 2] by reordering rows and columns. The singularity degree of this instance is $n - 1$, and hence (2.4) yields $d^* = n - 1$. After facial reduction, we have a chain $\mathcal{F}_n \subsetneq \dots \subsetneq \mathcal{F}_1$ of faces of $\mathcal{S}_+^n$ such that $\mathcal{F}_n = \mathbb{R}_+ \mathbf{E}_{11}$. In what follows, we show that, for this particular family of instances $\mathcal{V}_n$, the error bound in (3.57) is not asymptotically tight up to a dimension-free constant.

Take a sequence $(n_k) \subseteq \mathbb{Z}_{\geq 3}$ satisfying $\lim_{k \rightarrow \infty} n_k = \infty$ and a sequence $(\mathbf{X}_k)$ satisfying $\mathbf{X}_k \in \mathcal{R}_{n_k}$ arbitrarily, where we recall that $\mathcal{R}_{n_k}$ is defined in (3.56). For these sequences, we prove that the limit inferior of the ratio of the left-hand side of (3.57) to the right-hand side of (3.57), i.e.,

$$\rho_k := \frac{\text{dist}(\mathbf{X}_k, \mathcal{V}_{n_k} \cap \mathcal{S}_+^{n_k})}{\kappa(\|\mathbf{X}_k\|_F) 12^{n_k-2} \alpha_1 \dots \alpha_{n_k-1} ((n_k-1) \max\{\|\mathbf{X}_k\|_F^{1-\frac{1}{2^{n_k-1}}}, \sqrt{\|\mathbf{X}_k\|_F}\} + 1) \mathcal{E}_b(\mathbf{X}_k)^{\frac{1}{2^{n_k-1}}}}$$

is 0. For simplicity, we write ϵ_k for $\mathcal{E}_b(\mathbf{X}_k)$.

First, we provide a lower bound for the right-hand side of (3.57). The value of $\kappa(\|\mathbf{X}_k\|_F)$ is

$$\kappa(\|\mathbf{X}_k\|_F) = \min\{\max\{\eta(\mathcal{V}_{n_k}, \text{span } \mathcal{F}_{n_k}), 1\} \theta(\|\mathbf{X}_k\|_F), \eta(\mathcal{V}_{n_k}, \mathcal{F}_{n_k})\} = 1,$$

where the first equality holds since $\mathcal{V}_{n_k} \cap \text{ri } \mathcal{F}_{n_k} \neq \emptyset$ and $\mathcal{F}_{n_k} = \mathbb{R}_+ \mathbf{E}_{11}$ is polyhedral (see (3.23)), and the second equality follows from $\eta(\mathcal{V}_{n_k}, \mathcal{F}_{n_k}) = 1$, which is implied by $\mathcal{F}_{n_k} \subseteq \mathcal{V}_{n_k}$. In addition, it follows from Remark 3.11 that $\alpha_i \geq 1$ for each $i = 1, \dots, n_k - 1$. Therefore, we have

$$\begin{aligned} & \kappa(\|\mathbf{X}_k\|_F) 12^{n_k-2} \alpha_1 \dots \alpha_{n_k-1} ((n_k-1) \max\{\|\mathbf{X}_k\|_F^{1-\frac{1}{2^{n_k-1}}}, \sqrt{\|\mathbf{X}_k\|_F}\} + 1) \epsilon_k^{\frac{1}{2^{n_k-1}}} \\ & \geq 12^{n_k-2} (\max\{\|\mathbf{X}_k\|_F^{1-\frac{1}{2^{n_k-1}}}, \sqrt{\|\mathbf{X}_k\|_F}\} + 1) \epsilon_k^{\frac{1}{2^{n_k-1}}}. \end{aligned} \tag{4.10}$$

Next, we provide an upper bound for $\text{dist}(\mathbf{X}_k, \mathcal{V}_{n_k} \cap \mathcal{S}_+^{n_k})$, the left-hand side of (3.57). By $\mathcal{V}_{n_k} \cap \mathcal{S}_+^{n_k} = \mathbb{R}_+ \mathbf{E}_{11}$ and the subadditivity of the square root function, we see that

$$\begin{aligned} \text{dist}(\mathbf{X}_k, \mathcal{V}_{n_k} \cap \mathcal{S}_+^{n_k}) & \leq |(\mathbf{X}_k)_{n_k n_k}| + \text{dist}((\mathbf{X}_k)_{11}, \mathbb{R}_+) + \sqrt{2} \sum_{s=0}^{n_k-2} |(\mathbf{X}_k)_{1, n_k-s}| \\ & \quad + \sum_{s=0}^{n_k-3} |(\mathbf{X}_k)_{n_k-s-1, n_k-s-1}| + \sqrt{2} \sum_{0 \leq t < s \leq n_k-2} |(\mathbf{X}_k)_{n_k-s, n_k-t}|. \end{aligned} \tag{4.11}$$

We bound each term of the right-hand side of (4.11). Firstly, since

$$\text{dist}(\mathbf{X}_k, \mathcal{V}_{n_k}) = \sqrt{\frac{2}{3} \sum_{i=2}^{n_k-1} ((\mathbf{X}_k)_{ii} - (\mathbf{X}_k)_{1,i+1})^2 + (\mathbf{X}_k)_{n_k n_k}^2}$$

and it does not exceed ϵ_k , we have

$$\sqrt{2/3} |(\mathbf{X}_k)_{ii} - (\mathbf{X}_k)_{1,i+1}| \leq \epsilon_k \quad (i = 2, \dots, n_k - 1), \quad (4.12)$$

$$|(\mathbf{X}_k)_{n_k n_k}| \leq \epsilon_k. \quad (4.13)$$

In particular, (4.13) implies that, by using $\epsilon_k \leq 1$ and $1 \leq (\|\mathbf{X}_k\|_F + 1)^{1 - \frac{1}{2^{n_k-1}}}$, the first term of the right-hand side of (4.11), i.e., $|(\mathbf{X}_k)_{n_k n_k}|$ can be bounded by

$$|(\mathbf{X}_k)_{n_k n_k}| \leq (\|\mathbf{X}_k\|_F + 1)^{1 - \frac{1}{2^{n_k-1}}} \epsilon_k^{\frac{1}{2^{n_k-1}}}. \quad (4.14)$$

Secondly, the second term of the right-hand side of (4.11) admits the bound

$$\text{dist}((\mathbf{X}_k)_{11}, \mathbb{R}_+) \leq \epsilon_k \leq (\|\mathbf{X}_k\|_F + 1)^{1 - \frac{1}{2^{n_k-1}}} \epsilon_k^{\frac{1}{2^{n_k-1}}}, \quad (4.15)$$

where the first inequality holds by $\text{dist}(\mathbf{X}_k, \mathcal{S}_+^{n_k}) \leq \epsilon_k$ and (2.2), and the second inequality follows in the same manner as (4.14). In the following lemma, we bound the third and fourth terms of the right-hand side of (4.11).

Lemma 4.1. *It follows that*

$$|(\mathbf{X}_k)_{1, n_k-s}| \leq (2 + (\sqrt{3/2} + 1)s)(\|\mathbf{X}_k\|_F + 1)^{1 - \frac{1}{2^{s+1}}} \epsilon_k^{\frac{1}{2^{s+1}}} \quad (s = 0, \dots, n_k - 2), \quad (4.16)$$

$$|(\mathbf{X}_k)_{n_k-s-1, n_k-s-1}| \leq (2 + \sqrt{3/2} + (\sqrt{3/2} + 1)s)(\|\mathbf{X}_k\|_F + 1)^{1 - \frac{1}{2^{s+1}}} \epsilon_k^{\frac{1}{2^{s+1}}} \quad (s = 0, \dots, n_k - 3). \quad (4.17)$$

In particular, there exist constants a_{n_k} and b_{n_k} (whose dependency on n_k is polynomial) such that

$$\sqrt{2} \sum_{s=0}^{n_k-2} |X_{1, n_k-s}| \leq a_{n_k} (\|\mathbf{X}_k\|_F + 1)^{1 - \frac{1}{2^{n_k-1}}} \epsilon_k^{\frac{1}{2^{n_k-1}}}, \quad (4.18)$$

$$\sum_{s=0}^{n_k-3} |X_{n_k-s-1, n_k-s-1}| \leq b_{n_k} (\|\mathbf{X}_k\|_F + 1)^{1 - \frac{1}{2^{n_k-1}}} \epsilon_k^{\frac{1}{2^{n_k-1}}}. \quad (4.19)$$

Proof. We note that

$$|(\mathbf{X}_k)_{1i}| \leq \sqrt{((\mathbf{X}_k)_{11} + \epsilon_k)((\mathbf{X}_k)_{ii} + \epsilon_k)} \quad (4.20)$$

holds for every $i = 1, \dots, n_k$. Indeed, $\mathbf{X}_k + \epsilon_k \mathbf{I}_{n_k}$ is positive semidefinite by $\text{dist}(\mathbf{X}_k, \mathcal{S}_+^{n_k}) \leq \epsilon_k$ and (2.1). Hence, the determinant of the principal submatrix of $\mathbf{X}_k + \epsilon_k \mathbf{I}$ obtained by extracting the rows and columns indexed by 1 and i is nonnegative, which leads to (4.20).

First, we show the inequalities in (4.16) and (4.17) by induction. The inequality in (4.16) holds for $s = 0$ since

$$|(\mathbf{X}_k)_{1n_k}| \leq \sqrt{((\mathbf{X}_k)_{11} + \epsilon_k)((\mathbf{X}_k)_{n_k n_k} + \epsilon_k)} \leq \sqrt{2(\|\mathbf{X}_k\|_F + 1)\epsilon_k} \leq 2\sqrt{(\|\mathbf{X}_k\|_F + 1)\epsilon_k},$$

where we use (4.20) to derive the first inequality and use $|(\mathbf{X}_k)_{11}| \leq \|\mathbf{X}_k\|_F$, $\epsilon_k \leq 1$, and (4.13) to derive the second inequality. We assume that (4.16) holds for $s = t \in \{0, \dots, n_k - 3\}$. Then it follows that

$$\begin{aligned} |(\mathbf{X}_k)_{n_k-t-1, n_k-t-1}| &\leq |(\mathbf{X}_k)_{1, n_k-t}| + \sqrt{3/2}\epsilon_k \\ &\leq (2 + (\sqrt{3/2} + 1)t)(\|\mathbf{X}_k\|_F + 1)^{1 - \frac{1}{2^{t+1}}} \epsilon_k^{\frac{1}{2^{t+1}}} + \sqrt{3/2}\epsilon_k \\ &\leq (2 + \sqrt{3/2} + (\sqrt{3/2} + 1)t)(\|\mathbf{X}_k\|_F + 1)^{1 - \frac{1}{2^{t+1}}} \epsilon_k^{\frac{1}{2^{t+1}}}, \end{aligned} \quad (4.21)$$

where the first inequality holds by (4.12), the second inequality results from the inductive hypothesis, and the third inequality holds by $\epsilon_k \leq (\|\mathbf{X}_k\|_F + 1)^{1-\frac{1}{2^{t+1}}} \epsilon_k^{\frac{1}{2^{t+1}}}$. Therefore, (4.17) holds for $s = t$. Next, we have

$$\begin{aligned} |(\mathbf{X}_k)_{1, n_k - (t+1)}| &\leq \sqrt{((\mathbf{X}_k)_{11} + \epsilon_k)((\mathbf{X}_k)_{n_k - (t+1), n_k - (t+1)} + \epsilon_k)} \\ &\leq \sqrt{(\|\mathbf{X}_k\|_F + 1)((2 + \sqrt{3/2} + (\sqrt{3/2} + 1)t)(\|\mathbf{X}_k\|_F + 1)^{1-\frac{1}{2^{t+1}}} \epsilon_k^{\frac{1}{2^{t+1}}} + \epsilon_k)} \\ &\leq \sqrt{(2 + (\sqrt{3/2} + 1)(t+1))(\|\mathbf{X}_k\|_F + 1)^{1-\frac{1}{2^{t+2}}} \epsilon_k^{\frac{1}{2^{t+2}}}} \\ &\leq (2 + (\sqrt{3/2} + 1)(t+1))(\|\mathbf{X}_k\|_F + 1)^{1-\frac{1}{2^{t+2}}} \epsilon_k^{\frac{1}{2^{t+2}}}, \end{aligned}$$

where we use (4.20) to derive the first inequality, use $|(\mathbf{X}_k)_{11}| \leq \|\mathbf{X}_k\|_F$, $\epsilon_k \leq 1$, and (4.21) to derive the second inequality, use $\epsilon_k \leq (\|\mathbf{X}_k\|_F + 1)^{1-\frac{1}{2^{t+1}}} \epsilon_k^{\frac{1}{2^{t+1}}}$ to derive the third inequality. Therefore, (4.16) holds for $s = t + 1$.

Next, using (4.16), we have

$$\begin{aligned} \sqrt{2} \sum_{s=0}^{n_k-2} |X_{1, n_k-s}| &\leq \sqrt{2} \sum_{s=0}^{n_k-2} (2 + (\sqrt{3/2} + 1)s)(\|\mathbf{X}_k\|_F + 1)^{1-\frac{1}{2^{s+1}}} \epsilon_k^{\frac{1}{2^{s+1}}} \\ &\leq \sqrt{2} \underbrace{\left(\sum_{s=0}^{n_k-2} (2 + (\sqrt{3/2} + 1)s) \right)}_{=: a_{n_k}} (\|\mathbf{X}_k\|_F + 1)^{1-\frac{1}{2^{n_k-1}}} \epsilon_k^{\frac{1}{2^{n_k-1}}}, \end{aligned}$$

so we obtain (4.18). Similarly, using (4.17), by setting $b_{n_k} := \sum_{s=0}^{n_k-3} (2 + \sqrt{3/2} + (\sqrt{3/2} + 1)s)$, we obtain (4.19). $\square$

In the following lemma, we bound the fifth term of the right-hand side of (4.11).

Lemma 4.2. *There exists a positive constant c_{n_k} (whose dependency on n_k is polynomial) such that*

$$\sqrt{2} \sum_{0 \leq t < s \leq n_k-2} |(\mathbf{X}_k)_{n_k-s, n_k-t}| \leq c_{n_k} (\|\mathbf{X}_k\|_F + 1)^{1-\frac{1}{2^{n_k-1}}} \epsilon_k^{\frac{1}{2^{n_k-1}}}.$$

Proof. It follows from (4.17) that

$$(\mathbf{X}_k)_{n_k-s, n_k-s} + \epsilon_k \leq (2 + (\sqrt{3/2} + 1)s)(\|\mathbf{X}_k\|_F + 1)^{1-\frac{1}{2^s}} \epsilon_k^{\frac{1}{2^s}} \quad (4.22)$$

for every $s = 1, \dots, n_k-2$. The inequality in (4.22) also holds for $s = 0$ since (4.13) implies that $(\mathbf{X}_k)_{n_k, n_k} + \epsilon_k \leq 2\epsilon_k$ holds. Therefore, for each $0 \leq t < s \leq n_k-2$, we have

$$\begin{aligned} |(\mathbf{X}_k)_{n_k-s, n_k-t}| &\leq \sqrt{((\mathbf{X}_k)_{n_k-s, n_k-s} + \epsilon_k)((\mathbf{X}_k)_{n_k-t, n_k-t} + \epsilon_k)} \\ &\leq \sqrt{(2 + (\sqrt{3/2} + 1)s)(2 + (\sqrt{3/2} + 1)t)(\|\mathbf{X}_k\|_F + 1)^{1-(\frac{1}{2^{s+1}} + \frac{1}{2^{t+1}})} \epsilon_k^{\frac{1}{2^{s+1}} + \frac{1}{2^{t+1}}}} \\ &\leq \sqrt{(2 + (\sqrt{3/2} + 1)s)(2 + (\sqrt{3/2} + 1)t)(\|\mathbf{X}_k\|_F + 1)^{1-\frac{1}{2^{n_k-1}}} \epsilon_k^{\frac{1}{2^{n_k-1}}}}, \end{aligned}$$

where we use the positive semidefiniteness of $\mathbf{X}_k + \epsilon_k \mathbf{I}_{n_k}$ to derive the first inequality, use (4.22) to derive the second inequality, and use $0 \leq t < s \leq n_k-2$ to derive the third inequality. Therefore, by letting

$$c_{n_k} := \sqrt{2} \sum_{0 \leq t < s \leq n_k-2} \sqrt{(2 + (\sqrt{3/2} + 1)s)(2 + (\sqrt{3/2} + 1)t)},$$

whose order is polynomial in n_k , we obtain the desired result. $\square$

By (4.11), (4.14), (4.15), and Lemmas 4.1 and 4.2, we obtain

$$\text{dist}(\mathbf{X}_k, \mathcal{V}_{n_k} \cap \mathcal{S}_+^{n_k}) \leq (2 + a_{n_k} + b_{n_k} + c_{n_k})(\|\mathbf{X}_k\|_F + 1)^{1-\frac{1}{2^{n_k-1}}} \epsilon_k^{\frac{1}{2^{n_k-1}}}. \quad (4.23)$$

When $\|\mathbf{X}_k\|_F \geq 1$, since $\max\{\|\mathbf{X}_k\|_F^{1-\frac{1}{2^{n_k-1}}}, \sqrt{\|\mathbf{X}_k\|_F}\} = \|\mathbf{X}_k\|_F^{1-\frac{1}{2^{n_k-1}}}$ holds by $n_k \geq 3$, it follows from (4.10) and (4.23) that

$$\rho_k \leq \frac{(2 + a_{n_k} + b_{n_k} + c_{n_k})(\|\mathbf{X}_k\|_F + 1)^{1-\frac{1}{2^{n_k-1}}}}{12^{n_k-1}(\|\mathbf{X}_k\|_F^{1-\frac{1}{2^{n_k-1}}} + 1)} \leq \frac{2 + a_{n_k} + b_{n_k} + c_{n_k}}{12^{n_k-1} \left(\frac{\|\mathbf{X}_k\|_F}{\|\mathbf{X}_k\|_F + 1} \right)^{1-\frac{1}{2^{n_k-1}}}} \leq \frac{2(2 + a_{n_k} + b_{n_k} + c_{n_k})}{12^{n_k-1}}.$$

When $\|\mathbf{X}_k\|_F < 1$, since $\max\{\|\mathbf{X}_k\|_F^{1-\frac{1}{2^{n_k-1}}}, \sqrt{\|\mathbf{X}_k\|_F}\} = \sqrt{\|\mathbf{X}_k\|_F}$ holds by $n_k \geq 3$, it follows from (4.10) and (4.23) that

$$\rho_k \leq \frac{(2 + a_{n_k} + b_{n_k} + c_{n_k})(\|\mathbf{X}_k\|_F + 1)^{1-\frac{1}{2^{n_k-1}}}}{12^{n_k-1}(\sqrt{\|\mathbf{X}_k\|_F} + 1)} \leq \frac{2(2 + a_{n_k} + b_{n_k} + c_{n_k})}{12^{n_k-1}},$$

where the second inequality follows from $(\|\mathbf{X}_k\|_F + 1)^{1-\frac{1}{2^{n_k-1}}} \leq 2$ and $\sqrt{\|\mathbf{X}_k\|_F} \geq 0$. Therefore, regardless of the value of $\|\mathbf{X}_k\|_F$, we have

$$\rho_k \leq \frac{2(2 + a_{n_k} + b_{n_k} + c_{n_k})}{12^{n_k-1}},$$

and so $\liminf_{k \rightarrow \infty} \rho_k = 0$. Since the sequences (n_k) and $(\mathbf{X}_k)$ are arbitrary, we see that the error bound in (3.57) is not asymptotically tight up to a dimension-free constant.

5 Explicit error bounds for optimality system of SDPs

Consider the following standard SDP

$$\begin{aligned} & \underset{\mathbf{X} \in \mathcal{S}^n}{\text{minimize}} && \langle \mathbf{C}, \mathbf{X} \rangle \\ & \text{subject to} && \mathcal{A}(\mathbf{X}) = \mathbf{b}, \mathbf{X} \in \mathcal{S}_+^n \end{aligned} \quad (\text{PSDP})$$

and its dual problem

$$\begin{aligned} & \underset{\mathbf{y} \in \mathbb{R}^m}{\text{maximize}} && \mathbf{b}^\top \mathbf{y} \\ & \text{subject to} && \mathbf{C} - \mathcal{A}^*(\mathbf{y}) \in \mathcal{S}_+^n, \end{aligned} \quad (\text{DSDP})$$

where $\mathbf{C} \in \mathcal{S}^n$, $\mathcal{A}: \mathcal{S}^n \rightarrow \mathbb{R}^m$ is a linear mapping, and $\mathbf{b} \in \mathbb{R}^m$. Denote the set of optimal solutions of (PSDP) by $\mathcal{X}^*$, which we assume to be nonempty, and its optimal value by α^* . Let

$$\mathcal{V}^* := \{\mathbf{X} \in \mathcal{S}^n \mid \mathcal{A}(\mathbf{X}) = \mathbf{b}, \langle \mathbf{C}, \mathbf{X} \rangle = \alpha^*\}. \quad (5.1)$$

Then we have $\mathcal{X}^* = \mathcal{V}^* \cap \mathcal{S}_+^n$. It is worth noting that, even assuming the standard Slater's condition for (PSDP), that is, $\mathcal{A}^{-1}(\mathbf{b}) \cap \text{int } \mathcal{S}_+^n \neq \emptyset$, the strict feasibility condition for the semidefinite feasibility problem describing the optimal solution set, that is $\text{Feas}(\mathcal{V}^*, \mathcal{S}_+^n)$, *typically fails*.

Below, as an application, we provide an error bound result with an explicit estimate on the modulus of the radial-type error bound, under the commonly assumed strict complementarity condition. We note that, in the special case where (PSDP) has a unique solution (with other additional suitable assumptions), error estimates for feasible solutions can be deduced from the existing literature on the study of the conditioning of the so-called simple SDPs [14, 15] (see also [33]). See the discussions in Remark 5.3 later for details.

Recall that for a feasible point $\mathbf{X}^*$ of (PSDP) and a feasible point $\mathbf{y}^*$ of (DSDP) with $\mathbf{S}^* = \mathbf{C} - \mathcal{A}^*(\mathbf{y}^*)$, they satisfy *strict complementarity* if it follows that

$$\mathbf{X}^* \in \text{ri } \mathcal{F}^*, \quad (5.2)$$

where $\mathcal{F}^* := \mathcal{S}_+^n \cap \{\mathbf{S}^*\}^\perp$. In the above strict complementarity condition, $\mathbf{X}^*$ and $\mathbf{y}^*$ are indeed optimal solutions of (PSDP) and (DSDP), respectively since (5.2) implies strong duality, i.e., $\langle \mathbf{C}, \mathbf{X}^* \rangle = \mathbf{b}^\top \mathbf{y}^*$. In addition, by the strong duality between (PSDP) and (DSDP), it follows that

$$\mathcal{V}^* = \{\mathbf{X} \in \mathcal{S}^n \mid \mathcal{A}(\mathbf{X}) = \mathbf{b}, \langle \mathbf{S}^*, \mathbf{X} \rangle = 0\}.$$

It can be readily verified that the above strict complementarity condition is equivalent to the statement that the sum of the ranks of $\mathbf{X}^*$ and $\mathbf{S}^*$ is equal to n . The strict complementarity condition is a widely used condition in studying SDPs, and it is known that it holds generically [1, 38]. Let d be the rank of $\mathbf{S}^*$. We note that the face $\mathcal{F}^*$ is linearly isomorphic to $\mathcal{S}_+^{n-d}$. Throughout this section, we always assume the following.

Assumption 5.1.

- (A1) The problem (PSDP) and its dual (DSDP) admit a pair of feasible points $(\mathbf{X}^*, \mathbf{y}^*)$ with $\mathbf{S}^* = \mathbf{C} - \mathcal{A}^*(\mathbf{y}^*)$ such that the strict complementarity condition (5.2) holds.
- (A2) For the matrix $\mathbf{S}^* = \mathbf{C} - \mathcal{A}^*(\mathbf{y}^*)$, we assume that $\mathbf{S}^* \neq \mathbf{O}$ (that is, the rank d of $\mathbf{S}^*$ is at least 1).

Note that if $\mathbf{S}^* = \mathbf{O}$, then $\mathcal{V}^* = \{\mathbf{X} \in \mathcal{S}^n \mid \mathcal{A}(\mathbf{X}) = \mathbf{b}\}$, and the assumption in (A1) implies that $\mathbf{X}^* \in \mathcal{V}^* \cap \text{int } \mathcal{S}_+^n$. Thus, the optimal solution set satisfies Slater's condition, and an explicit Lipschitz error bound holds (for example, by the case $d^* = 0$ in Corollary 3.20). Therefore, below we only consider the case where $\mathbf{S}^* \neq \mathbf{O}$.

To formulate the explicit error bound, we first define

$$\begin{aligned} \gamma &:= \frac{\|\mathbf{S}^*\|_F + 1}{\lambda_{\min}^+(\mathbf{S}^*)} + 1, \\ \alpha &:= \sqrt{1 + 2(\gamma + 1)(n - d) + \gamma^2}, \end{aligned} \quad (5.3)$$

$$\beta := \sqrt{2(\gamma + 1)\sqrt{n - d}}. \quad (5.4)$$

Let $\mathcal{M}: \mathcal{S}^n \rightarrow \mathbb{R}^p$ be a linear mapping with $p \geq 1$ whose kernel is $\text{span } \mathcal{F}^*$. We define linear mappings $\bar{\mathcal{A}}: \mathcal{S}^n \rightarrow \mathbb{R}^{m+1}$ and $\hat{\mathcal{A}}: \mathcal{S}^n \rightarrow \mathbb{R}^{m+p+1}$ respectively by

$$\bar{\mathcal{A}}(\mathbf{X}) := \begin{pmatrix} \mathcal{A}(\mathbf{X}) \\ \langle \mathbf{S}^*, \mathbf{X} \rangle \end{pmatrix} \quad \text{and} \quad \hat{\mathcal{A}}(\mathbf{X}) := \begin{pmatrix} \mathcal{A}(\mathbf{X}) \\ \mathcal{M}(\mathbf{X}) \\ \langle \mathbf{S}^*, \mathbf{X} \rangle \end{pmatrix}. \quad (5.5)$$

By using the linear mapping $\bar{\mathcal{A}}$, the affine space $\mathcal{V}^*$ can be represented as

$$\mathcal{V}^* = \left\{ \mathbf{X} \in \mathcal{S}^n \mid \bar{\mathcal{A}}(\mathbf{X}) = \begin{pmatrix} \mathbf{b} \\ 0 \end{pmatrix} \right\}. \quad (5.6)$$

Let $\sigma_{\min}^+(\bar{\mathcal{A}})$ and $\sigma_{\min}^+(\hat{\mathcal{A}})$ be the smallest positive singular values of $\bar{\mathcal{A}}$ and $\hat{\mathcal{A}}$, respectively.^{*3} Recall that $\mathbb{R}\mathbf{S}^*$ is the line along the matrix $\mathbf{S}^*$, and, for a linear mapping $\mathcal{L}$ between two finite-dimensional normed spaces, we use $\|\mathcal{L}\|_{\text{op}}$ and $\|\mathcal{L}\|_{\text{HS}}$ to denote the operator norm of $\mathcal{L}$ and the Hilbert–Schmidt norm of $\mathcal{L}$, respectively. Finally, the function $\bar{\kappa}(\rho; \mathbf{X}^*)$ with $\mathbf{X}^* \in \mathcal{V}^* \cap \text{ri } \mathcal{F}^*$ is defined as in (3.22) for $\mathcal{V} = \mathcal{V}^*$ and $\mathcal{F}_{d^*+1} = \mathcal{F}^*$, that is,

$$\begin{aligned} \bar{\kappa}(\rho; \mathbf{X}^*) &= \max\{\eta(\mathcal{V}^*, \text{span } \mathcal{F}^*), 1\} \left(1 + 2 \frac{\|\mathbf{X}^*\|_F + \rho}{\lambda_{\min}^+(\mathbf{X}^*)} \right) \\ &\leq \frac{\|\hat{\mathcal{A}}\|_{\text{HS}}}{\sigma_{\min}^+(\hat{\mathcal{A}})} \left(1 + 2 \frac{\|\mathbf{X}^*\|_F + \rho}{\lambda_{\min}^+(\mathbf{X}^*)} \right), \end{aligned} \quad (5.7)$$

where the inequality follows from (ii) of Lemma 2.2 and the fact that the Hilbert–Schmidt norm of a linear operator coincides with the Frobenius norm of its matrix representation.

Theorem 5.2 (Error bounds for optimality system of SDPs with strict complementarity). *Consider (PSDP) and its dual problem (DSDP). Suppose that Assumption 5.1 holds. For each $\mathbf{X} \in \mathcal{S}^n$, let the optimality measure $\epsilon_{\text{OM}}(\mathbf{X})$ be given by*

$$\epsilon_{\text{OM}}(\mathbf{X}) := \max\{|\langle \mathbf{S}^*, \mathbf{X} \rangle|, \|\mathcal{A}(\mathbf{X}) - \mathbf{b}\|_2, \text{dist}(\mathbf{X}, \mathcal{S}_+^n)\}. \quad (5.8)$$

- (i) For all $\mathbf{X} \in \mathcal{S}^n$, we have

$$\text{dist}(\mathbf{X}, \mathcal{X}^*) \leq \alpha_{\text{SDP}}(\|\mathbf{X}\|_F) \epsilon_{\text{OM}}(\mathbf{X}) + \beta_{\text{SDP}}(\|\mathbf{X}\|_F) \sqrt{\epsilon_{\text{OM}}(\mathbf{X})},$$

where $\alpha_{\text{SDP}}(\cdot)$ and $\beta_{\text{SDP}}(\cdot)$ are defined as

$$\alpha_{\text{SDP}}(\rho) := \bar{\kappa}(\rho; \mathbf{X}^*) \max \left\{ \frac{1}{\sigma_{\min}^+(\bar{\mathcal{A}})} \left(1 + \frac{\|\mathcal{A}|_{\mathbb{R}\mathbf{S}^*}\|_{\text{op}}}{\|\mathbf{S}^*\|_F} \right), \alpha + \frac{1}{\|\mathbf{S}^*\|_F} \right\} + \frac{1}{\|\mathbf{S}^*\|_F}, \quad (5.9)$$

$$\beta_{\text{SDP}}(\rho) := \beta \sqrt{\rho} \bar{\kappa}(\rho; \mathbf{X}^*), \quad (5.10)$$

respectively, and $\bar{\kappa}$, α , and β are given as in (5.7), (5.3), and (5.4), respectively.

^{*3}The linear mappings $\bar{\mathcal{A}}$ and $\hat{\mathcal{A}}$ each have at least one positive singular value since the matrix $\mathbf{S}^*$ is nonzero.

(ii) For all $\mathbf{X} \in \mathcal{A}^{-1}(\mathbf{b}) \cap \mathcal{S}_+^n$, we have

$$\text{dist}(\mathbf{X}, \mathcal{X}^*) \leq \bar{\alpha}_{\text{SDP}}(\|\mathbf{X}\|_{\text{F}})(\langle \mathbf{C}, \mathbf{X} \rangle - \alpha^*) + \beta_{\text{SDP}}(\|\mathbf{X}\|_{\text{F}})\sqrt{\langle \mathbf{C}, \mathbf{X} \rangle - \alpha^*},$$

where

$$\bar{\alpha}_{\text{SDP}}(\rho) := \bar{\kappa}(\rho; \mathbf{X}^*) \max \left\{ \frac{\|\mathcal{A}|_{\mathbb{R}\mathbf{S}^*}\|_{\text{op}}}{\sigma_{\min}^+(\bar{\mathcal{A}})\|\mathbf{S}^*\|_{\text{F}}}, \alpha + \frac{1}{\|\mathbf{S}^*\|_{\text{F}}} \right\} + \frac{1}{\|\mathbf{S}^*\|_{\text{F}}}.$$

Proof. First, we prove (i). Let $\mathbf{X} \in \mathcal{S}^n$. We decompose it as

$$\mathbf{X} = \underbrace{\mathbf{X} - \frac{\langle \mathbf{S}^*, \mathbf{X} \rangle}{\|\mathbf{S}^*\|_{\text{F}}^2} \mathbf{S}^*}_{=P_{\{\mathbf{S}^*\}^\perp}(\mathbf{X})} + \underbrace{\frac{\langle \mathbf{S}^*, \mathbf{X} \rangle}{\|\mathbf{S}^*\|_{\text{F}}^2} \mathbf{S}^*}_{=P_{\mathbb{R}\mathbf{S}^*}(\mathbf{X})}.$$

Then we have

$$\begin{aligned} & \text{dist}(\mathbf{X}, \mathcal{X}^*) \\ & \stackrel{(a)}{\leq} \text{dist}(P_{\{\mathbf{S}^*\}^\perp}(\mathbf{X}), \mathcal{X}^*) + \frac{\epsilon_{\text{OM}}(\mathbf{X})}{\|\mathbf{S}^*\|_{\text{F}}} \\ & \stackrel{(b)}{=} \text{dist}(P_{\{\mathbf{S}^*\}^\perp}(\mathbf{X}), \mathcal{V}^* \cap \mathcal{F}^*) + \frac{\epsilon_{\text{OM}}(\mathbf{X})}{\|\mathbf{S}^*\|_{\text{F}}} \\ & \stackrel{(c)}{\leq} \bar{\kappa}(\|P_{\{\mathbf{S}^*\}^\perp}(\mathbf{X})\|_{\text{F}}; \mathbf{X}^*) \max\{\text{dist}(P_{\{\mathbf{S}^*\}^\perp}(\mathbf{X}), \mathcal{V}^*), \text{dist}(P_{\{\mathbf{S}^*\}^\perp}(\mathbf{X}), \mathcal{F}^*)\} + \frac{\epsilon_{\text{OM}}(\mathbf{X})}{\|\mathbf{S}^*\|_{\text{F}}} \\ & \stackrel{(d)}{\leq} \bar{\kappa}(\|\mathbf{X}\|_{\text{F}}; \mathbf{X}^*) \max \left\{ \frac{\|\mathcal{A}(P_{\{\mathbf{S}^*\}^\perp}(\mathbf{X})) - \mathbf{b}\|_2}{\sigma_{\min}^+(\bar{\mathcal{A}})}, \text{dist}(P_{\{\mathbf{S}^*\}^\perp}(\mathbf{X}), \mathcal{F}^*) \right\} + \frac{\epsilon_{\text{OM}}(\mathbf{X})}{\|\mathbf{S}^*\|_{\text{F}}} \\ & \stackrel{(e)}{\leq} \bar{\kappa}(\|\mathbf{X}\|_{\text{F}}; \mathbf{X}^*) \max \left\{ \frac{1}{\sigma_{\min}^+(\bar{\mathcal{A}})} \left(\|\mathcal{A}(\mathbf{X}) - \mathbf{b}\|_2 + \|\mathcal{A}|_{\mathbb{R}\mathbf{S}^*}\|_{\text{op}} \frac{\epsilon_{\text{OM}}(\mathbf{X})}{\|\mathbf{S}^*\|_{\text{F}}} \right), \text{dist}(P_{\{\mathbf{S}^*\}^\perp}(\mathbf{X}), \mathcal{F}^*) \right\} + \frac{\epsilon_{\text{OM}}(\mathbf{X})}{\|\mathbf{S}^*\|_{\text{F}}} \quad (5.11) \\ & \stackrel{(f)}{\leq} \bar{\kappa}(\|\mathbf{X}\|_{\text{F}}; \mathbf{X}^*) \max \left\{ \frac{\epsilon_{\text{OM}}(\mathbf{X})}{\sigma_{\min}^+(\bar{\mathcal{A}})} \left(1 + \frac{\|\mathcal{A}|_{\mathbb{R}\mathbf{S}^*}\|_{\text{op}}}{\|\mathbf{S}^*\|_{\text{F}}} \right), \text{dist}(P_{\{\mathbf{S}^*\}^\perp}(\mathbf{X}), \mathcal{F}^*) \right\} + \frac{\epsilon_{\text{OM}}(\mathbf{X})}{\|\mathbf{S}^*\|_{\text{F}}}, \quad (5.12) \end{aligned}$$

where (a) follows from the triangular inequality and the definition of $\epsilon_{\text{OM}}(\mathbf{X})$, (b) follows from $\mathcal{X}^* = \mathcal{V}^* \cap \mathcal{F}^*$, (c) is obtained by applying Proposition 3.3 with $\mathcal{V} = \mathcal{V}^*$ and $\mathcal{F}_{d^*+1} = \mathcal{F}^*$ and then using the inequality $\kappa(\|P_{\{\mathbf{S}^*\}^\perp}(\mathbf{X})\|_{\text{F}}) \leq \bar{\kappa}(\|P_{\{\mathbf{S}^*\}^\perp}(\mathbf{X})\|_{\text{F}}; \mathbf{X}^*)$, (d) follows from (2.8), (5.6), and the nonexpansiveness of the projection $P_{\{\mathbf{S}^*\}^\perp}$, (e) follows from the triangular inequality and the definition of $\epsilon_{\text{OM}}(\mathbf{X})$, and (f) also follows from the definition of $\epsilon_{\text{OM}}(\mathbf{X})$. On the quantity $\text{dist}(P_{\{\mathbf{S}^*\}^\perp}(\mathbf{X}), \mathcal{F}^*)$ that appears in (5.12), it follows from Proposition 3.10 with $\epsilon = \epsilon_{\text{OM}}(\mathbf{X})$ that

$$\begin{aligned} \text{dist}(P_{\{\mathbf{S}^*\}^\perp}(\mathbf{X}), \mathcal{F}^*) &= \text{dist}(P_{\{\mathbf{S}^*\}^\perp}(\mathbf{X}), \mathcal{S}_+^n \cap \{\mathbf{S}^*\}^\perp) \\ &\leq \text{dist}(\mathbf{X}, \mathcal{S}_+^n \cap \{\mathbf{S}^*\}^\perp) + \|\mathbf{X} - P_{\{\mathbf{S}^*\}^\perp}(\mathbf{X})\|_{\text{F}} \\ &\leq \left(\alpha + \frac{1}{\|\mathbf{S}^*\|_{\text{F}}} \right) \epsilon_{\text{OM}}(\mathbf{X}) + \beta \sqrt{\epsilon_{\text{OM}}(\mathbf{X}) \|\mathbf{X}\|_{\text{F}}}, \quad (5.13) \end{aligned}$$

where α and β are given as in (5.3) and (5.4), respectively. By (5.12) and (5.13), we obtain

$$\begin{aligned} & \text{dist}(\mathbf{X}, \mathcal{X}^*) \\ & \leq \bar{\kappa}(\|\mathbf{X}\|_{\text{F}}; \mathbf{X}^*) \max \left\{ \frac{\epsilon_{\text{OM}}(\mathbf{X})}{\sigma_{\min}^+(\bar{\mathcal{A}})} \left(1 + \frac{\|\mathcal{A}|_{\mathbb{R}\mathbf{S}^*}\|_{\text{op}}}{\|\mathbf{S}^*\|_{\text{F}}} \right), \left(\alpha + \frac{1}{\|\mathbf{S}^*\|_{\text{F}}} \right) \epsilon_{\text{OM}}(\mathbf{X}) + \beta \sqrt{\epsilon_{\text{OM}}(\mathbf{X}) \|\mathbf{X}\|_{\text{F}}} \right\} + \frac{\epsilon_{\text{OM}}(\mathbf{X})}{\|\mathbf{S}^*\|_{\text{F}}} \\ & \leq \alpha_{\text{SDP}}(\|\mathbf{X}\|_{\text{F}}) \epsilon_{\text{OM}}(\mathbf{X}) + \beta_{\text{SDP}}(\|\mathbf{X}\|_{\text{F}}) \sqrt{\epsilon_{\text{OM}}(\mathbf{X})}. \end{aligned}$$

Next, we prove (ii). Let $\mathbf{X} \in \mathcal{A}^{-1}(\mathbf{b}) \cap \mathcal{S}_+^n$. Then it follows from $\mathbf{X} \in \mathcal{A}^{-1}(\mathbf{b}) \cap \mathcal{S}_+^n$ and $\mathbf{S}^* \in \mathcal{S}_+^n$ that $\epsilon_{\text{OM}}(\mathbf{X}) = \langle \mathbf{C}, \mathbf{X} \rangle - \alpha^*$. Moreover, by $\mathbf{X} \in \mathcal{A}^{-1}(\mathbf{b}) \cap \mathcal{S}_+^n$, (5.11) can be simplified as

$$\text{dist}(\mathbf{X}, \mathcal{X}^*) \leq \bar{\kappa}(\|\mathbf{X}\|_{\text{F}}; \mathbf{X}^*) \max \left\{ \frac{\|\mathcal{A}|_{\mathbb{R}\mathbf{S}^*}\|_{\text{op}} \epsilon_{\text{OM}}(\mathbf{X})}{\sigma_{\min}^+(\bar{\mathcal{A}})\|\mathbf{S}^*\|_{\text{F}}}, \text{dist}(P_{\{\mathbf{S}^*\}^\perp}(\mathbf{X}), \mathcal{F}^*) \right\} + \frac{\epsilon_{\text{OM}}(\mathbf{X})}{\|\mathbf{S}^*\|_{\text{F}}}.$$

Combining this inequality with (5.13), we obtain the desired result. $\square$

Remark 5.3 (Links to existing works). We note that, in addition to Assumption 5.1, if we further assume that (PSDP) has a unique optimal solution and $\mathcal{A}^{-1}(\mathbf{b}) \cap \text{int } \mathcal{S}_+^n \neq \emptyset$ (that is, standard Slater's condition holds for (PSDP)), similar error estimates for feasible solutions can be deduced from the existing literature on the study of the conditioning of the so-called simple SDPs [14, 15].

More precisely, suppose that (PSDP) has a unique optimal solution, i.e., $\mathcal{X}^* = \{\mathbf{X}^*\}$. Recall that $\mathcal{F}^* = \mathcal{S}_+^n \cap \{\mathbf{S}^*\}^\perp$. Then the mapping $\mathcal{A}|_{\text{span } \mathcal{F}^*}$ is injective. (Otherwise, we can take $\mathbf{U} \in (\text{span } \mathcal{F}^*) \setminus \{\mathbf{O}\}$ such that $\mathcal{A}(\mathbf{U})$ is zero. Then the matrix $\mathbf{X}(t) := \mathbf{X}^* + t\mathbf{U}$ defined for every $t > 0$ belongs to $\mathcal{A}^{-1}(\mathbf{b}) \cap \text{span } \mathcal{F}^*$. As $\mathbf{X}^* \in \text{ri } \mathcal{F}^*$, for sufficiently small $t > 0$, we have $\mathbf{X}(t) \in \mathcal{F}^*$. This implies that $\mathbf{X}(t) \in \mathcal{X}^*$, which contradicts the assumption that $\mathcal{X}^*$ is a singleton.) Thus, the minimum singular value of $\mathcal{A}|_{\text{span } \mathcal{F}^*}$, which is denoted by $\sigma_{\min}(\mathcal{A}|_{\text{span } \mathcal{F}^*})$ and calculated as

$$\sigma_{\min}(\mathcal{A}|_{\text{span } \mathcal{F}^*}) = \min\{\|\mathcal{A}(\mathbf{X})\|_2 \mid \mathbf{X} \in \text{span } \mathcal{F}^*, \|\mathbf{X}\|_F = 1\}, \quad (5.14)$$

is positive.

Let $\mathbf{X} \in \mathcal{A}^{-1}(\mathbf{b}) \cap \mathcal{S}_+^n$ be arbitrary. It follows that

$$\begin{aligned} \text{dist}(\mathbf{X}, \mathcal{X}^*) &\stackrel{(a)}{=} \|\mathbf{X} - \mathbf{X}^*\|_F \\ &\leq \|P_{\text{span } \mathcal{F}^*}(\mathbf{X}) - \mathbf{X}^*\|_F + \|\mathbf{X} - P_{\text{span } \mathcal{F}^*}(\mathbf{X})\|_F \\ &\stackrel{(b)}{\leq} \frac{\|\mathcal{A}(P_{\text{span } \mathcal{F}^*}(\mathbf{X})) - \mathbf{b}\|_2}{\sigma_{\min}(\mathcal{A}|_{\text{span } \mathcal{F}^*})} + \|\mathbf{X} - P_{\text{span } \mathcal{F}^*}(\mathbf{X})\|_F \\ &\stackrel{(c)}{=} \frac{\|\mathcal{A}(\mathbf{X} - P_{\text{span } \mathcal{F}^*}(\mathbf{X}))\|_2}{\sigma_{\min}(\mathcal{A}|_{\text{span } \mathcal{F}^*})} + \|\mathbf{X} - P_{\text{span } \mathcal{F}^*}(\mathbf{X})\|_F \\ &\leq \left(1 + \frac{\|\mathcal{A}|_{(\text{span } \mathcal{F}^*)^\perp}\|_{\text{op}}}{\sigma_{\min}(\mathcal{A}|_{\text{span } \mathcal{F}^*})}\right) \|\mathbf{X} - P_{\text{span } \mathcal{F}^*}(\mathbf{X})\|_F, \end{aligned} \quad (5.15)$$

where we use the assumption that $\mathcal{X}^* = \{\mathbf{X}^*\}$ to derive (a), use (5.14) and $\mathcal{A}(\mathbf{X}^*) = \mathbf{b}$ to derive (b), use $\mathcal{A}(\mathbf{X}) = \mathbf{b}$ to derive (c). Now, from [15, Lemma 4.3] and $\mathbf{X} \in \mathcal{S}_+^n$, it follows that

$$\|\mathbf{X} - P_{\text{span } \mathcal{F}^*}(\mathbf{X})\|_F \leq \frac{\langle \mathbf{S}^*, \mathbf{X} \rangle}{\lambda_{\min}^+(\mathbf{S}^*)} + \sqrt{2 \frac{\langle \mathbf{S}^*, \mathbf{X} \rangle}{\lambda_{\min}^+(\mathbf{S}^*)} \lambda_{\max}(\mathbf{X})}. \quad (5.16)$$

By (5.15) and (5.16), we obtain

$$\text{dist}(\mathbf{X}, \mathcal{X}^*) \leq \left(1 + \frac{\|\mathcal{A}|_{(\text{span } \mathcal{F}^*)^\perp}\|_{\text{op}}}{\sigma_{\min}(\mathcal{A}|_{\text{span } \mathcal{F}^*})}\right) \left(\frac{\epsilon_{\text{OM}}(\mathbf{X})}{\lambda_{\min}^+(\mathbf{S}^*)} + \sqrt{2 \frac{\epsilon_{\text{OM}}(\mathbf{X})}{\lambda_{\min}^+(\mathbf{S}^*)} \|\mathbf{X}\|_F}\right).$$

Different from the error estimate provided in (ii) of Theorem 5.2, the above estimate relies heavily on the fact that $\sigma_{\min}(\mathcal{A}|_{\text{span } \mathcal{F}^*}) > 0$, which holds due to the uniqueness of the optimal solution of (PSDP).

In the following example, we illustrate the error bound shown in (i). For this instance, we see that the optimal solution set of the primal problem is not a singleton (and so, the results in the existing literature mentioned above are not applicable). Moreover, this example shows that the error bound can be asymptotically tight up to a dimension-free constant.

Example 5.4 (Asymptotic tightness for explicit error bound for the optimality system of SDP). For $n \in \mathbb{Z}_{\geq 3}$, we consider the following SDP:

$$\begin{aligned} &\underset{\mathbf{X} \in \mathcal{S}_+^n}{\text{minimize}} && \sum_{i=3}^n X_{ii} \\ &\text{subject to} && \sum_{i=1}^n X_{ii} = 1, \mathbf{X} \in \mathcal{S}_+^n. \end{aligned} \quad (5.17)$$

By letting $\mathbf{C} := \text{Diag}(0, 0, 1, \dots, 1)$, $\mathcal{A}(\mathbf{X}) := \sum_{i=1}^n X_{ii}$, and $b := 1$, (PSDP) reduces to (5.17). The dual problem of Problem (5.17) can be written as

$$\begin{aligned} &\underset{y \in \mathbb{R}}{\text{maximize}} && y \\ &\text{subject to} && \text{Diag}(-y, -y, 1 - y, \dots, 1 - y) \in \mathcal{S}_+^n. \end{aligned}$$

It can be seen that an optimal solution of the primal problem is $\mathbf{X}^* := \text{Diag}(1/2, 1/2, 0, \dots, 0)$, and the unique optimal solution of the dual problem is $y^* = 0$ with the corresponding slack matrix $\mathbf{S}^* = \text{Diag}(0, 0, 1, \dots, 1)$. The optimal value of these problems is $\alpha^* = 0$. We note that the matrix $\mathbf{X}^*$ is the same as that used in Section 4.2. For Problem (5.17), by $\alpha^* = 0$, the affine space $\mathcal{V}_n^*$ corresponding to (5.1) can be written as

$$\mathcal{V}_n^* = \left\{ \mathbf{A} \in \mathcal{S}^n \mid \sum_{i=1}^n A_{ii} = 1, \sum_{i=3}^n A_{ii} = 0 \right\}.$$

We note that the affine space $\mathcal{V}_n^*$ is the same as that introduced in (4.4). The set $\mathcal{X}^* = \mathcal{V}_n^* \cap \mathcal{S}_+^n$ of optimal solutions of the primal problem is

$$\{\text{Diag}(\mathbf{A}, \mathbf{O}) \in \mathcal{S}^n \mid \mathbf{A} \in \mathcal{S}_+^2, A_{11} + A_{22} = 1\},$$

and the primal problem has nonunique optimal solutions. The optimal solution $\mathbf{X}^*$ of the primal problem and the optimal solution y^* of the dual problem satisfy the strict complementarity condition in (5.2). Using $\mathbf{X}^*$, we see from (i) of Theorem 5.2 that

$$\begin{aligned} \text{dist}(\mathbf{X}, \mathcal{X}^*) &\leq \left(\sqrt{2}(1 + 2\sqrt{2} + 4\|\mathbf{X}\|_F) \left(\sqrt{n + 8\sqrt{n-2} + 15} + \frac{1}{\sqrt{n-2}} \right) + \frac{1}{\sqrt{n-2}} \right) \epsilon_{\text{OM}}(\mathbf{X}) \\ &\quad + 2\sqrt{\sqrt{2(n-2)} + 3\sqrt{2}(1 + 2\sqrt{2} + 4\|\mathbf{X}\|_F)} \sqrt{\|\mathbf{X}\|_F} \sqrt{\epsilon_{\text{OM}}(\mathbf{X})} \text{ for all } \mathbf{X} \in \mathcal{S}^n, \end{aligned} \quad (5.18)$$

where $\epsilon_{\text{OM}}(\mathbf{X})$ defined in (5.8) is

$$\epsilon_{\text{OM}}(\mathbf{X}) = \max \left\{ \left| \sum_{i=3}^n X_{ii} \right|, \left| \sum_{i=1}^n X_{ii} - 1 \right|, \text{dist}(\mathbf{X}, \mathcal{S}_+^n) \right\}$$

and see Appendix D for the calculations of $\alpha_{\text{SDP}}(\rho)$ and $\beta_{\text{SDP}}(\rho)$ defined in (5.9) and (5.10), respectively. In what follows, we show that the error bound in (5.18) is asymptotically tight up to a dimension-free constant.

For every $k \in \mathbb{Z}_{\geq 1}$, let $n_k := k + 3$ and $\mathbf{X}_k$ be the matrix defined in (4.6). As calculated in (4.7), the left-hand side of (5.18) with $\mathbf{X} = \mathbf{X}_k$ is

$$\text{dist}(\mathbf{X}_k, \mathcal{X}^*) = \text{dist}(\mathbf{X}_k, \mathcal{V}_{n_k} \cap \mathcal{S}_+^{n_k}) = \sqrt{\frac{k+1}{k}} \rightarrow 1 \quad (\text{as } k \rightarrow \infty).$$

We then calculate the right-hand side of (5.18). It can be seen that $|\sum_{i=3}^{n_k} (\mathbf{X}_k)_{ii}| = 0$ and $|\sum_{i=1}^{n_k} (\mathbf{X}_k)_{ii} - 1| = 0$. In addition, it follows that $\text{dist}(\mathbf{X}_k, \mathcal{S}_+^{n_k}) = 1/\sqrt{k}$ as seen in Section 4.2. By combining them, we have $\epsilon_{\text{OM}}(\mathbf{X}_k) = 1/\sqrt{k}$, and the right-hand side of (5.18) with $\mathbf{X} = \mathbf{X}_k$ and $n = n_k$ is

$$\begin{aligned} &\left(\sqrt{2} \left(1 + 2\sqrt{2} + 4\sqrt{\frac{2k+1}{k}} \right) \left(\sqrt{k + 8\sqrt{k+1} + 18} + \frac{1}{\sqrt{k+1}} \right) + \frac{1}{\sqrt{k+1}} \right) \frac{1}{\sqrt{k}} \\ &\quad + 2\sqrt{\sqrt{2(k+1)} + 3\sqrt{2}} \left(1 + 2\sqrt{2} + 4\sqrt{\frac{2k+1}{k}} \right) \sqrt{\frac{2k+1}{k}} \frac{1}{\sqrt[4]{k}} \rightarrow 3(12 + \sqrt{2}) \quad (\text{as } k \rightarrow \infty). \end{aligned}$$

Thus, the error bound in (5.18) is asymptotically tight up to the dimension-free constant $\frac{1}{3(12+\sqrt{2})}$.

6 Conclusions and remarks

In this paper, we investigated radial-type Hölder error bounds for semidefinite feasibility problems $\text{Feas}(\mathcal{V}_n, \mathcal{S}_+^n)$ without assuming any constraint qualifications. The final objective was to furnish error bounds for which the associated constants have explicit expressions. This was accomplished via facial reduction and the framework of facial residual functions [29].

Furthermore, we introduced the notion of asymptotic tightness of error bounds and analyzed the asymptotic tightness of the qualitative bounds derived in this paper in terms of $d^* = d_{\text{PPS}}(\mathcal{V}_n, \mathcal{S}_+^n)$, the distance to the PPS condition of $\text{Feas}(\mathcal{V}_n, \mathcal{S}_+^n)$. Our analysis showed that the derived radial-type error bounds can be asymptotically

tight up to a dimension-free constant factor, when d^* equals 0 or 1. We also discussed the more challenging case where $d^* = n - 1$.

As an application, we derived an explicit error bound for the optimality system of SDPs satisfying strict complementarity. Interestingly, this explicit bound was obtained without assuming the usual uniqueness of the optimal solution. Moreover, we presented an example showing that it can be asymptotically tight up to a dimension-free constant factor.

A limitation of our results is that the notion of tightness considered here is asymptotic and is defined in terms of sequences of instances; see Definition 2.4. For $d^* \in \{0, 1\}$, the error bound (3.57) is asymptotically tight in the sense that there exist sequences of instances whose dimensions tend to infinity, for which the condition in Definition 2.4 is satisfied. As discussed in Remark 2.5, this shows that, for these families of instances, the derived error bound constants are best possible up to a dimension-free constant factor.

In order to dispel potential sources of confusion, it is important to emphasize what this asymptotic tightness statement does and does not imply. In particular, we do not claim that, for *every* family of instances with $d^* \in \{0, 1\}$, the error bound in (3.57) is asymptotically tight. Nor do we claim that, for a *single specific instance*, that is, for fixed $\mathcal{V}$ and $\mathcal{S}_+^n$, the derived constants in the error bound (3.57) are necessarily best possible. The error bound holds for each individual instance, but its constants may be conservative for particular instances. Indeed, as observed in Section 4.3, in the most degenerate case where $d^* = n - 1$, the derived error bound appears to be rather conservative.

Getting a reasonable expression for the best possible constants for one specific semidefinite feasibility problem remains a significant research challenge but we hope this work may be a useful first step. It would be of interest to investigate whether our bounds can be improved, for example, by exploiting more refined geometric properties of the positive semidefinite cone.

On the algorithmic side, it would be interesting to understand how the explicit error bounds obtained in this paper can be used to derive explicit convergence rate estimates for some of the relevant numerical algorithms in the literature.

Acknowledgments The first author is supported by Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Research Activity Start-up JP25K23344. The second author is partially supported by the Australian Research Council under Discovery Projects DP190100555 and DP250101112. The third author is supported by JSPS Grant-in-Aid for Early-Career Scientists JP23K16844 and Japan Science and Technology Agency ASPIRE Grant Number JPMJAP2520.

Statements and declarations

- Competing interests: The authors declare that there are no competing interests.
- Availability of data and materials: No data were used for the research described in the article.
- Use of AI: Preliminary versions of Lemmas 3.15 and 3.16 were developed with the assistance of ChatGPT. The proofs in the present version were subsequently rewritten and verified by the authors, who assume full responsibility for their correctness.

Appendix A The proof of (2.4)

Proof. For simplicity, we write d^* for $d_{\text{PPS}}(\mathcal{V}, \mathcal{S}_+^n)$. It is sufficient to show that $d_s \leq d^*$. Let $\mathcal{F}_{d^*+1} \subsetneq \cdots \subsetneq \mathcal{F}_1$ be a chain of faces of $\mathcal{S}_+^n$. By the definition of d^* , the face $\mathcal{F}_{d^*+1}$ is either polyhedral or satisfies $\mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1} \neq \emptyset$. If $\mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1} \neq \emptyset$, then, by the definition of d_s , we have $d_s \leq d^*$. It remains to consider the case in which $\mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1} = \emptyset$. Then $\mathcal{F}_{d^*+1}$ must be polyhedral. The polyhedral faces of $\mathcal{S}_+^n$ are precisely the 0-dimensional face $\{\mathbf{O}\}$ and the 1-dimensional faces $\mathbb{R}_+ \mathbf{A}$ with $\mathbf{A}$ being a positive semidefinite matrix of rank 1. If $\mathcal{F}_{d^*+1} = \{\mathbf{O}\}$, we have $\mathcal{V} \cap \mathcal{S}_+^n = \{\mathbf{O}\}$. As shown in the proof of [31, Proposition 27], we can reach the polyhedral face $\{\mathbf{O}\}$ of $\mathcal{S}_+^n$ in only one facial reduction step, and $\mathcal{V} \cap \text{ri } \{\mathbf{O}\} \neq \emptyset$. Under the assumption that $\mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1} = \emptyset$, we have $\mathcal{V} \cap \text{int } \mathcal{S}_+^n = \emptyset$. Moreover, $\mathcal{S}_+^n$ is not polyhedral since $n \geq 2$. Therefore, we obtain $d_s = d^* = 1$. If $\mathcal{F}_{d^*+1} = \mathbb{R}_+ \mathbf{A}$ with $\mathbf{A}$ being a positive semidefinite matrix of rank 1, then $\text{ri } \mathcal{F}_{d^*+1} = \{t\mathbf{A} \mid t > 0\}$. Since $\mathcal{V} \cap \text{ri } \mathcal{F}_{d^*+1} = \emptyset$ and the problem $\text{Feas}(\mathcal{V}, \mathcal{S}_+^n)$ is feasible, we have $\mathcal{V} \cap \mathcal{S}_+^n = \mathcal{V} \cap \mathcal{F}_{d^*+1} = \{\mathbf{O}\}$. Therefore, by the same argument as in the case where $\mathcal{F}_{d^*+1} = \{\mathbf{O}\}$, we again obtain $d_s = d^* = 1$. $\square$

Appendix B The proof of Lemma 2.2

Proof. Here, we only prove (i). The statement in (ii) can be proven in a similar fashion. Let $\mathbf{a}_{ij}^\top$ denote the j th row of $\mathbf{A}_i$, and let b_{ij} denote the j th element of $\mathbf{b}_i$. For notational convenience, we define $C_{ij} := \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_{ij}^\top \mathbf{x} \leq b_{ij}\}$ for each $i = 1, 2$ and $j = 1, \dots, m_i$. We note that

$$C_1 \cap C_2 = \bigcap_{i=1}^2 \bigcap_{j=1}^{m_i} C_{ij} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_{ij}^\top \mathbf{x} \leq b_{ij} \ (i = 1, 2, \ j = 1, \dots, m_i)\} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \leq \mathbf{b}\},$$

where $\mathbf{b}$ is the vector obtained by concatenating the vector $\mathbf{b}_1$ with the vector $\mathbf{b}_2$. Then for any $\mathbf{x} \in \mathbb{R}^n$, we have

$$\begin{aligned} \text{dist}(\mathbf{x}, C_1 \cap C_2) &\leq H(\mathbf{A}) \sqrt{\sum_{i=1}^2 \sum_{j=1}^{m_i} \max\{\mathbf{a}_{ij}^\top \mathbf{x} - b_{ij}, 0\}^2} \\ &= H(\mathbf{A}) \sqrt{\sum_{i=1}^2 \sum_{j=1}^{m_i} \|\mathbf{a}_{ij}\|_2^2 \text{dist}(\mathbf{x}, C_{ij})^2} \\ &\leq H(\mathbf{A}) \|\mathbf{A}\|_F \max_{\substack{i=1,2, \\ j=1, \dots, m_i}} \text{dist}(\mathbf{x}, C_{ij}) \\ &\leq H(\mathbf{A}) \|\mathbf{A}\|_F \max\{\text{dist}(\mathbf{x}, C_1), \text{dist}(\mathbf{x}, C_2)\}, \end{aligned}$$

where the first inequality follows from the definition of the Hoffman constant $H(\mathbf{A})$, the equality follows from the formula for the distance to a half-space [13, Theorem 9.39], and the last inequality follows from the fact that $C_i \subseteq C_{ij}$ for each $i = 1, 2$ and $j = 1, \dots, m_i$. Thus, the conclusion holds. $\square$

Appendix C Computation of the quantities in Section 4.2

In this appendix, we illustrate the computations of $\bar{\kappa}(\rho; \mathbf{X}^*)$, α_1 , and β_1 used in Section 4.2. For simplicity, we consider the case where the ambient space is $\mathcal{S}^n$ instead of $\mathcal{S}^{n_k}$ with $n_k = k + 3$, because we can obtain the result for the case of $\mathcal{S}^{n_k}$ by substituting n_k into n .

$\bar{\kappa}(\rho; \mathbf{X}^*) = \sqrt{2}(1 + 2\sqrt{2} + 4\rho)$ Recall that $\bar{\kappa}(\rho; \mathbf{X}^*) = \max\{\eta(\mathcal{V}_n, \text{span } \mathcal{F}_2), 1\} \bar{\theta}(\rho; \mathbf{X}^*)$. The value $\bar{\theta}(\rho; \mathbf{X}^*)$, which is defined in (3.17), is $1 + 2\sqrt{2} + 4\rho$. In what follows, we show that $\eta(\mathcal{V}_n, \text{span } \mathcal{F}_2) = \sqrt{2}$. Before computing $\eta(\mathcal{V}_n, \text{span } \mathcal{F}_2)$, we prepare an additional lemma regarding the direct sum decomposition of a subspace of $\mathcal{S}^n$.

Lemma C.1. *Let V_1 and V_2 be subspaces of $\mathcal{S}^n$. If $P_{V_2}(V_1)$ is included in V_1 , then we have*

$$V_1 = (V_1 \cap V_2) \oplus (V_1 \cap V_2^\perp).$$

The proof of Lemma C.1 is omitted because it is straightforward. We note that the statement does not hold if $P_{V_2}(V_1)$ is not included in V_1 .

The face $\mathcal{F}_2$ of $\mathcal{S}_+^n$ exposed by the matrix $\mathbf{Z}_1$ is provided by (4.5). Let

$$\mathcal{W} := \mathcal{V}_n - \mathbf{X}^* = \left\{ \mathbf{A} \in \mathcal{S}^n \mid \sum_{i=1}^n A_{ii} = 0, \sum_{i=3}^n A_{ii} = 0 \right\}$$

for simplicity. We observe that $\eta(\mathcal{V}_n, \text{span } \mathcal{F}_2)$, denoted by η for brevity, is the minimum positive number satisfying the following inequality:

$$\text{dist}(\mathbf{X}, \mathcal{W} \cap \text{span } \mathcal{F}_2) \leq \eta \max\{\text{dist}(\mathbf{X}, \mathcal{W}), \text{dist}(\mathbf{X}, \text{span } \mathcal{F}_2)\} \text{ for all } \mathbf{X} \in \mathcal{S}^n.$$

Let $\mathbf{D}_1 := \mathbf{E}_{11} + \mathbf{E}_{22}$ and $\mathbf{D}_2 := \sum_{i=3}^n \mathbf{E}_{ii}$. We note that $\mathcal{W} = \{\mathbf{D}_1, \mathbf{D}_2\}^\perp$. It follows that

$$\mathcal{S}^n = \mathcal{W} \oplus \mathcal{W}^\perp = \underbrace{(\mathcal{W} \cap \text{span } \mathcal{F}_2) \oplus (\mathcal{W} \cap (\text{span } \mathcal{F}_2)^\perp)}_{=\mathcal{W}} \oplus \underbrace{\mathbb{R}\mathbf{D}_1 \oplus \mathbb{R}\mathbf{D}_2}_{=\mathcal{W}^\perp},$$

where the second inequality follows from $P_{\text{span } \mathcal{F}_2}(\mathcal{W}) \subseteq \mathcal{W}$ and Lemma C.1. Following this (orthogonal) direct sum decomposition, we decompose a given matrix $\mathbf{X} \in \mathcal{S}^n$ into $\mathbf{X} = \mathbf{Y} + \mathbf{Z} + s\mathbf{D}_1 + t\mathbf{D}_2$, where $\mathbf{Y} \in \mathcal{W} \cap \text{span } \mathcal{F}_2$, $\mathbf{Z} \in \mathcal{W} \cap (\text{span } \mathcal{F}_2)^\perp$,

and $s, t \in \mathbb{R}$. Let $a := \sqrt{2}|s|$, $b := \|\mathbf{Z}\|_{\text{F}}$, and $c := \sqrt{n-2}|t|$. By $\mathbb{R}\mathbf{D}_1 = \mathcal{W}^\perp \cap \text{span } \mathcal{F}_2$ and $\mathbb{R}\mathbf{D}_2 = \mathcal{W}^\perp \cap (\text{span } \mathcal{F}_2)^\perp$, we see that

$$\begin{aligned}\text{dist}(\mathbf{X}, \mathcal{W} \cap \text{span } \mathcal{F}_2) &= \|\mathbf{Z} + s\mathbf{D}_1 + t\mathbf{D}_2\|_{\text{F}} = \sqrt{a^2 + b^2 + c^2}, \\ \text{dist}(\mathbf{X}, \mathcal{W}) &= \|s\mathbf{D}_1 + t\mathbf{D}_2\|_{\text{F}} = \sqrt{a^2 + c^2}, \\ \text{dist}(\mathbf{X}, \text{span } \mathcal{F}_2) &= \|\mathbf{Z} + t\mathbf{D}_2\|_{\text{F}} = \sqrt{b^2 + c^2}.\end{aligned}$$

The numbers a , b , and c range over all nonnegative numbers as $\mathbf{X}$ ranges over the set $\mathcal{S}^n$. Therefore, η is indeed the minimum positive number satisfying the following inequality:

$$\sqrt{a^2 + b^2 + c^2} \leq \eta \max\{\sqrt{a^2 + c^2}, \sqrt{b^2 + c^2}\} \quad \text{for all } a, b, c \in \mathbb{R}_+.$$

When $c = 0$ and $a = b > 0$, since

$$\frac{\sqrt{a^2 + b^2 + c^2}}{\max\{\sqrt{a^2 + c^2}, \sqrt{b^2 + c^2}\}} = \sqrt{2},$$

we see that $\eta \geq \sqrt{2}$. Conversely, for any $a, b, c \in \mathbb{R}_+$ with $(a, b, c) \neq (0, 0, 0)$, we have

$$\frac{\sqrt{a^2 + b^2 + c^2}}{\max\{\sqrt{a^2 + c^2}, \sqrt{b^2 + c^2}\}} \leq \sqrt{2},$$

i.e., $\eta \leq \sqrt{2}$. Thus, we obtain $\eta = \sqrt{2}$.

$\alpha_1 = \sqrt{n + 8\sqrt{n-2} + 15}$ and $\beta_1 = \sqrt{2\sqrt{2(n-2)} + 6\sqrt{2}}$ Recall that α_1 and β_1 are defined as (3.36) and (3.37), respectively. These definitions require r_2 and γ_1 . First, we see that $r_2 = 2$ since $\mathcal{F}_2$ is linearly isomorphic to $\mathcal{S}_+^2$. Second, since $\|\mathbf{Z}_1\|_{\text{F}} = \sqrt{n-2}$ and $\lambda_{\min}^+(\mathbf{Z}_1) = 1$, the value γ_1 defined in (3.35) is $\sqrt{n-2} + 2$. Substituting these components into (3.36) and (3.37), we obtain the desired result.

Appendix D Computation of the radial modulus functions in Example 5.4

In this appendix, we illustrate the computations of $\alpha_{\text{SDP}}(\rho)$ and $\beta_{\text{SDP}}(\rho)$. To compute them, we need to calculate the ingredients involved with them, i.e., $\bar{\kappa}(\rho; \mathbf{X}^*)$, $\|\mathcal{A}|_{\mathbb{R}\mathcal{S}^*}\|_{\text{op}}$, $\sigma_{\min}^+(\bar{\mathcal{A}})$, α , and β . Among them, $\bar{\kappa}(\rho; \mathbf{X}^*)$, α , and β have been in fact already calculated in Appendix C; α is the same as α_1 and β is the same as β_1 . In what follows, we compute $\sigma_{\min}^+(\bar{\mathcal{A}})$ and $\|\mathcal{A}|_{\mathbb{R}\mathcal{S}^*}\|_{\text{op}}$.

$\sigma_{\min}^+(\bar{\mathcal{A}}) = \sqrt{n-1-\sqrt{n^2-4n+5}}$ For the instance discussed in Example 5.4, the linear mapping $\bar{\mathcal{A}}$ defined in (5.5) can be described as

$$\bar{\mathcal{A}}(\mathbf{X}) = \left(\sum_{i=1}^n X_{ii} \right)$$

for every $\mathbf{X} \in \mathcal{S}^n$. The representation matrix of $\bar{\mathcal{A}}\bar{\mathcal{A}}^*$ with respect to the standard basis of $\mathbb{R}^2$ is $\begin{pmatrix} n & n-2 \\ n-2 & n-2 \end{pmatrix}$, so we have $\sigma_{\min}^+(\bar{\mathcal{A}}) = \sqrt{n-1-\sqrt{n^2-4n+5}}$.

$\|\mathcal{A}|_{\mathbb{R}\mathcal{S}^*}\|_{\text{op}} = \sqrt{n-2}$ Recall that $\mathcal{A}$ is the linear mapping that maps $\mathbf{X} \in \mathcal{S}^n$ to $\langle \mathbf{I}_n, \mathbf{X} \rangle$. Then $\mathcal{A}|_{\mathbb{R}\mathcal{S}^*}$ is the linear mapping that maps $t\mathbf{S}^*$ for $t \in \mathbb{R}$ to $\langle \mathbf{I}_n, t\mathbf{S}^* \rangle = t(n-2)$. Therefore, we have

$$\|\mathcal{A}|_{\mathbb{R}\mathcal{S}^*}\|_{\text{op}} = \sup_{\substack{\mathbf{X} \in \mathbb{R}\mathcal{S}^*, \\ \mathbf{X} \neq \mathbf{0}}} \frac{\|\mathcal{A}|_{\mathbb{R}\mathcal{S}^*}(\mathbf{X})\|_2}{\|\mathbf{X}\|_{\text{F}}} = \sup_{t \neq 0} \frac{|t(n-2)|}{\|t\mathbf{S}^*\|_{\text{F}}} = \sqrt{n-2}.$$

We can directly calculate $\beta_{\text{SDP}}(\rho)$ defined in (5.10) by using β and $\bar{\kappa}(\rho; \mathbf{X}^*)$. The function $\alpha_{\text{SDP}}(\rho)$ defined in (5.9) is calculated as

$$\begin{aligned}\alpha_{\text{SDP}}(\rho) &= \sqrt{2}(1 + 2\sqrt{2} + 4\rho) \max \left\{ \frac{2}{\sqrt{n-1-\sqrt{n^2-4n+5}}}, \sqrt{n+8\sqrt{n-2}+15} + \frac{1}{\sqrt{n-2}} \right\} + \frac{1}{\sqrt{n-2}} \\ &= \sqrt{2}(1 + 2\sqrt{2} + 4\rho) \left(\sqrt{n+8\sqrt{n-2}+15} + \frac{1}{\sqrt{n-2}} \right) + \frac{1}{\sqrt{n-2}},\end{aligned}$$

where the second equality follows because

$$\sqrt{n+8\sqrt{n-2}+15} + \frac{1}{\sqrt{n-2}} \geq \frac{2}{\sqrt{n-1-\sqrt{n^2-4n+5}}}$$

holds under the assumption that $n \geq 3$.

References

- [1] F. Alizadeh, J.-P.A. Haeberly, and M.L. Overton. Complementarity and nondegeneracy in semidefinite programming. *Math. Program.*, 77(2):111–128, 1997. doi:[10.1007/BF02614432](https://doi.org/10.1007/BF02614432).
- [2] D. Applegate, O. Hinder, H. Lu, and M. Lubin. Faster first-order primal-dual methods for linear programming using restarts and sharpness. *Math. Program.*, 201(1–2):133–184, 2023. doi:[10.1007/s10107-022-01901-9](https://doi.org/10.1007/s10107-022-01901-9).
- [3] G.P. Barker and D.H. Carlson. Cones of diagonally dominant matrices. *Pac. J. Math.*, 57(1):15–32, 1975. doi:[10.2140/pjm.1975.57.15](https://doi.org/10.2140/pjm.1975.57.15).
- [4] H.H. Bauschke and J.M. Borwein. On projection algorithms for solving convex feasibility problems. *SIAM Rev.*, 38(3):367–426, 1996. doi:[10.1137/S0036144593251710](https://doi.org/10.1137/S0036144593251710).
- [5] H.H. Bauschke, J.M. Borwein, and W. Li. Strong conical hull intersection property, bounded linear regularity, Jameson’s property (G), and error bounds in convex optimization. *Math. Program.*, 86(1):135–160, 1999. doi:[10.1007/s101070050083](https://doi.org/10.1007/s101070050083).
- [6] A. Beck and S. Shtern. Linearly convergent away-step conditional gradient for non-strongly convex functions. *Math. Program.*, 164(1–2):1–27, 2017. doi:[10.1007/s10107-016-1069-4](https://doi.org/10.1007/s10107-016-1069-4).
- [7] A. Beck and M. Teboulle. Convergence rate analysis and error bounds for projection algorithms in convex feasibility problems. *Optim. Methods Softw.*, 18(4):377–394, 2003. doi:[10.1080/10556780310001604977](https://doi.org/10.1080/10556780310001604977).
- [8] A. Belloni and R.M. Freund. A geometric analysis of Renegar’s condition number, and its interplay with conic curvature. *Math. Program.*, 119(1):95–107, 2009. doi:[10.1007/s10107-007-0203-8](https://doi.org/10.1007/s10107-007-0203-8).
- [9] D.S. Bernstein. *Matrix Mathematics: Theory, Facts, and Formulas*. Princeton University Press, Princeton, NJ, second edition, 2009.
- [10] J. Borwein and H. Wolkowicz. Regularizing the abstract convex program. *J. Math. Anal. Appl.*, 83(2):495–530, 1981. doi:[10.1016/0022-247X\(81\)90138-4](https://doi.org/10.1016/0022-247X(81)90138-4).
- [11] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, UK, 2004.
- [12] S. Deng and H. Hu. Computable error bounds for semidefinite programming. *J. Glob. Optim.*, 14(2):105–115, 1999. doi:[10.1023/A:1008298700376](https://doi.org/10.1023/A:1008298700376).
- [13] F. Deutsch. *Best Approximation in Inner Product Spaces*. Springer, New York, NY, 2001. doi:[10.1007/978-1-4684-9298-9](https://doi.org/10.1007/978-1-4684-9298-9).
- [14] L. Ding and M. Udell. On the simplicity and conditioning of low rank semidefinite programs. *SIAM J. Optim.*, 31(4):2614–2637, 2021. doi:[10.1137/20M1346262](https://doi.org/10.1137/20M1346262).
- [15] L. Ding, A. Yurtsever, V. Cevher, J.A. Tropp, and M. Udell. An optimal-storage approach to semidefinite programming using approximate complementarity. *SIAM J. Optim.*, 31(4):2695–2725, 2021. doi:[10.1137/19M1244603](https://doi.org/10.1137/19M1244603).
- [16] D. Drusvyatskiy, G. Li, and H. Wolkowicz. A note on alternating projections for ill-posed semidefinite feasibility problems. *Math. Program.*, 162(1–2):537–548, 2017. doi:[10.1007/s10107-016-1048-9](https://doi.org/10.1007/s10107-016-1048-9).
- [17] D. Drusvyatskiy and H. Wolkowicz. The many faces of degeneracy in conic optimization. *Found. Trends Optim.*, 3(2):77–170, 2017. doi:[10.1561/24000000011](https://doi.org/10.1561/24000000011).
- [18] J. Faraut and A. Korányi. *Analysis on Symmetric Cones*. Clarendon Press, Oxford, UK, 1994.
- [19] O. Güler, A.J. Hoffman, and U.G. Rothblum. Approximations to solutions to systems of linear inequalities. *SIAM J. Matrix Anal. Appl.*, 16(2):688–696, 1995. doi:[10.1137/S0895479892237744](https://doi.org/10.1137/S0895479892237744).
- [20] R.D. Hill and S.R. Waters. On the cone of positive semidefinite matrices. *Linear Algebra Appl.*, 90:81–88, 1987. doi:[10.1016/0024-3795\(87\)90307-7](https://doi.org/10.1016/0024-3795(87)90307-7).
- [21] A.J. Hoffman. On approximate solutions of systems of linear inequalities. *J. Res. Natl. Bur. Stand.*, 49(4):263–265, 1952.
- [22] H. Hu. Geometric condition measures and smoothness condition measures for closed convex sets and linear regularity of infinitely many closed convex sets. *J. Optim. Theory Appl.*, 126(2):287–308, 2005. doi:[10.1007/s10957-005-4715-1](https://doi.org/10.1007/s10957-005-4715-1).

- [23] D. Klatte and G. Thiere. Error bounds for solutions of linear equations and inequalities. *Math. Methods Oper. Res.*, 41(2):191–214, 1995. doi:[10.1007/BF01432655](https://doi.org/10.1007/BF01432655).
- [24] D. Leventhal and A.S. Lewis. Randomized methods for linear constraints: convergence rates and conditioning. *Math. Oper. Res.*, 35(3):641–654, 2010. doi:[10.1287/moor.1100.0456](https://doi.org/10.1287/moor.1100.0456).
- [25] G. Li, A.K.C. Ma, and T.K. Pong. Robust least square semidefinite programming with applications. *Comput. Optim. Appl.*, 58(2):347–379, 2014. doi:[10.1007/s10589-013-9634-8](https://doi.org/10.1007/s10589-013-9634-8).
- [26] X.-B. Li, S. Al-Homidan, Q.H. Ansari, and C.-F. Wen. Global error bounds for linear semi-infinite system over polyhedral constraints. *J. Nonlinear Convex Anal.*, 26(3):495–514, 2025.
- [27] Y. Lin, S.B. Lindstrom, B.F. Lourenço, and T.K. Pong. Generalized power cones: optimal error bounds and automorphisms. *SIAM J. Optim.*, 34(2):1316–1340, 2024. doi:[10.1137/22M1542921](https://doi.org/10.1137/22M1542921).
- [28] Y. Lin, S.B. Lindstrom, B.F. Lourenço, and T.K. Pong. Tight error bounds for log-determinant cones without constraint qualifications. *J. Optim. Theory Appl.*, 205(3):45, 2025. doi:[10.1007/s10957-025-02644-1](https://doi.org/10.1007/s10957-025-02644-1).
- [29] S.B. Lindstrom, B.F. Lourenço, and T.K. Pong. Error bounds, facial residual functions and applications to the exponential cone. *Math. Program.*, 200(1):229–278, 2023. doi:[10.1007/s10107-022-01883-8](https://doi.org/10.1007/s10107-022-01883-8).
- [30] S.B. Lindstrom, B.F. Lourenço, and T.K. Pong. Optimal error bounds in the absence of constraint qualifications with applications to p -cones and beyond. *Math. Oper. Res.*, 50(2):1204–1232, 2025. doi:[10.1287/moor.2022.0135](https://doi.org/10.1287/moor.2022.0135).
- [31] B.F. Lourenço. Amenable cones: error bounds without constraint qualifications. *Math. Program.*, 186(1–2):1–48, 2021. doi:[10.1007/s10107-019-01439-3](https://doi.org/10.1007/s10107-019-01439-3).
- [32] B.F. Lourenço, M. Muramatsu, and T. Tsuchiya. Facial reduction and partial polyhedrality. *SIAM J. Optim.*, 28(3):2304–2326, 2018. doi:[10.1137/15M1051634](https://doi.org/10.1137/15M1051634).
- [33] M.V. Nayakkankuppam and M.L. Overton. Conditioning of semidefinite programs. *Math. Program.*, 85(3):525–540, 1999. doi:[10.1007/s101070050070](https://doi.org/10.1007/s101070050070).
- [34] I. Necoara, Y. Nesterov, and F. Glineur. Linear convergence of first order methods for non-strongly convex optimization. *Math. Program.*, 175(1–2):69–107, 2019. doi:[10.1007/s10107-018-1232-1](https://doi.org/10.1007/s10107-018-1232-1).
- [35] G. Pataki. The geometry of semidefinite programming. In H. Wolkowicz, R. Saigal, and L. Vandenbergh, editors, *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*, pages 29–65. Kluwer Academic Publishers, Norwell, MA, 2000. doi:[10.1007/978-1-4615-4381-7_3](https://doi.org/10.1007/978-1-4615-4381-7_3).
- [36] G. Pataki. On the closedness of the linear image of a closed convex cone. *Math. Oper. Res.*, 32(2):395–412, 2007. doi:[10.1287/moor.1060.0242](https://doi.org/10.1287/moor.1060.0242).
- [37] G. Pataki. On the connection of facially exposed and nice cones. *J. Math. Anal. Appl.*, 400(1):211–221, 2013. doi:[10.1016/j.jmaa.2012.10.033](https://doi.org/10.1016/j.jmaa.2012.10.033).
- [38] G. Pataki and L. Tunçel. On the generic properties of convex optimization problems in conic form. *Math. Program.*, 89(3):449–457, 2001. doi:[10.1007/PL00011408](https://doi.org/10.1007/PL00011408).
- [39] J. Peña and V. Roshchina. A data-independent distance to infeasibility for linear conic systems. *SIAM J. Optim.*, 30(2):1049–1066, 2020. doi:[10.1137/18M1189464](https://doi.org/10.1137/18M1189464).
- [40] J. Peña, J.C. Vera, and L.F. Zuluaga. New characterizations of Hoffman constants for systems of linear constraints. *Math. Program.*, 187(1–2):79–109, 2021. doi:[10.1007/s10107-020-01473-6](https://doi.org/10.1007/s10107-020-01473-6).
- [41] J.F. Peña. An easily computable upper bound on the Hoffman constant for homogeneous inequality systems. *Comput. Optim. Appl.*, 87(1):323–335, 2024. doi:[10.1007/s10589-023-00514-y](https://doi.org/10.1007/s10589-023-00514-y).
- [42] J.F. Peña, J. Vera, and L.F. Zuluaga. An algorithm to compute the Hoffman constant of a system of linear constraints. *arXiv e-prints*, 2018. doi:[10.48550/arXiv.1804.08418](https://doi.org/10.48550/arXiv.1804.08418).
- [43] J.F. Peña, J.C. Vera, and L.F. Zuluaga. Equivalence and invariance of the chi and Hoffman constants of a matrix. *arXiv e-prints*, 2020. doi:[10.48550/arXiv.1905.06366](https://doi.org/10.48550/arXiv.1905.06366).
- [44] R. Penrose. On best approximate solutions of linear matrix equations. *Math. Proc. Camb. Philos. Soc.*, 52(1):17–19, 1956. doi:[10.1017/S0305004100030929](https://doi.org/10.1017/S0305004100030929).

- [45] J. Renegar. Some perturbation theory for linear programming. *Math. Program.*, 65(1):73–91, 1994. doi:[10.1007/BF01581690](https://doi.org/10.1007/BF01581690).
- [46] R.T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, NJ, 1970.
- [47] V. Roshchina and L. Tunçel. Facially dual complete (nice) cones and lexicographic tangents. *SIAM J. Optim.*, 29(3):2363–2387, 2019. doi:[10.1137/17M1126643](https://doi.org/10.1137/17M1126643).
- [48] J.F. Sturm. Error bounds for linear matrix inequalities. *SIAM J. Optim.*, 10(4):1228–1248, 2000. doi:[10.1137/S1052623498338606](https://doi.org/10.1137/S1052623498338606).
- [49] P.-W. Wang and C.-J. Lin. Iteration complexity of feasible descent methods for convex optimization. *J. Mach. Learn. Res.*, 15(45):1523–1548, 2014. URL: <http://jmlr.org/papers/v15/wang14a.html>.
- [50] X. Wang, B.F. Lourenço, and T.K. Pong. Error bounds for perspective cones of a class of nonnegative Legendre functions. *arXiv e-prints*, 2025. doi:[10.48550/arXiv.2509.26289](https://doi.org/10.48550/arXiv.2509.26289).
- [51] H. Wolkowicz, R. Saigal, and L. Vandenberghe. *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*. Kluwer Academic Publishers, Norwell, MA, 2000. doi:[10.1007/978-1-4615-4381-7](https://doi.org/10.1007/978-1-4615-4381-7).
- [52] S. Zhang. Global error bounds for convex conic problems. *SIAM J. Optim.*, 10(3):836–851, 2000. doi:[10.1137/S105262349834429X](https://doi.org/10.1137/S105262349834429X).
- [53] C. Zălinescu. Sharp estimates for Hoffman’s constant for systems of linear inequalities and equalities. *SIAM J. Optim.*, 14(2):517–533, 2003. doi:[10.1137/S1052623402403505](https://doi.org/10.1137/S1052623402403505).