A Limiting Analysis on Regularization of Singular SDP and its Implication to Infeasible Interior-point Algorithms [*]

Takashi Tsuchiya[†] Bruno F. Lourenço[‡] Masakazu Muramatsu[§]
Takayuki Okuno[¶]

Revised: May 2021 (Original: December 2019)

Abstract

We consider primal-dual pairs of semidefinite programs and assume that they are singular, i.e., both primal and dual are either weakly feasible or weakly infeasible. Under such circumstances, strong duality may break down and the primal and dual might have a nonzero duality gap. Nevertheless, there are arbitrary small perturbations to the problem data which would make them strongly feasible thus zeroing the duality gap. In this paper, we conduct an asymptotic analysis of the optimal value as the perturbation for regularization is driven to zero. Specifically, we fix two positive definite matrices, I_p and I_d, say, (typically the identity matrices), and regularize the primal and dual problems by shifting their associated affine space by ηI_p and εI_d, respectively, to recover interior feasibility of both problems, where ε and η are positive numbers. Then we analyze the behavior of the optimal value of the regularized problem when the perturbation is reduced to zero keeping the ratio between η and ε constant. A key feature of our analysis is that no further assumptions such as compactness or constraint qualifications are ever made. It will be shown that the optimal value of the perturbed problem converges to a value between the primal and dual optimal values of the original problems. Furthermore, the limiting optimal value changes "monotonically" from the primal optimal value to the dual optimal value as a function of θ, if we parametrize (ε, η) as $(\varepsilon, \eta) = t(\cos\theta, \sin\theta)$ and let $t \to 0$. Finally, the analysis leads us to the relatively surprising consequence that some representative infeasible interior-point algorithms for

[*]The first author is supported in part by MEXT Grant-in-Aid for Scientific Research (B)18H03206, the second author is supported in part by the same grant and MEXT Grant-in-Aid for Young Scientists 19K20217, the third author is supported in part by MEXT Grant-in-Aid for Scientific Research (C)17K00031 and the same grant for Scientific Research (B)20H04145, and the fourth author is supported in part by MEXT Grant-in-Aid for Young Scientists 20K19748 and the same grant for Scientific Research (B)20H04145.

[†]National Graduate Institute for Policy Studies, 7-22-1 Roppongi, Minato-ku, Tokyo 106-8677 Japan, e-mail: tsuchiya@grips.ac.jp

[‡]The Institute of Statistical Mathematics, Midori-cho 10-3, Tachikawa, 190-8562 Tokyo Japan, e-mail: bruno@ism.ac.jp

[§]The University of Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo 182-8585 Japan, e-mail: MasakazuMuramatsu@uec.ac.jp

[¶]Center for Advanced Intelligence Project, RIKEN, 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027 Japan. email: takayuki.okuno.ks@riken.jp

SDP generate sequences converging to a number between the primal and dual optimal values, even in the presence of a nonzero duality gap. Though this result is more of theoretical interest at this point, it might be of some value in the development of infeasible interior-point algorithms that can handle singular problems.

keywords Semidefinite programs singular problems nonzero duality gaps perturbation regularization infeasible interior-point algorithms

1 Introduction

Strong feasibility of primal and dual problems is a standard regularity condition in convex optimization [31][1]. Once this condition is satisfied, powerful algorithms such as interior-point algorithms and the ellipsoid algorithm can be applied to solve them efficiently, at least in theory. On the other hand, if a problem at hand does not satisfy this condition, it can be much harder to solve. Due to the advance of techniques of optimization modelling, there are many problems which do not satisfy primal-dual strong feasibility by nature.

A standard method to deal with such nasty problems in semidefinite programming (SDP) and general convex programming is facial reduction [4–7, 28, 29, 33, 36, 39]. This approach recovers strong feasibility by finding the minimal face of the feasible region, but its implementation is subtle and not easy, being vulnerable to rounding errors. Nevertheless, it is worth mentioning that are several recent works focused on implementational issues regarding facial reduction [9, 24, 25, 42].

In this paper, we will focus on an alternative approach for dealing with these hard problems: *regularization*. The idea is to perturb the problem slightly to recover strong feasibility on both primal and dual sides. Once strong feasibility is recovered, we may, say, apply interior-point algorithms to the regularized problems. However, the resulting approximate optimal solution is not guaranteed to be close to the optimal solution to the original problem, though intuitively we might expect or hope so. In particular, if we consider a SDP problem with a finite and nonzero duality gap, it is not clear what happens with the optimal value and the optimal solutions of the regularized problem as functions of the perturbation when the perturbation is reduced to zero.

Analyzing this problem is one of the main topics of the current paper. We consider primal and dual pairs of semidefinite programs and assume they are singular i.e., either weakly feasible or weakly infeasible (see Section 2.1 for definitions). Under these circumstances, there are arbitrarily small perturbations which make the perturbed pair primal-dual strongly feasible. Then, we fix two positive definite matrices, and shift the associated affine spaces of the primal and dual slightly in the direction of these matrices so that the perturbed problems have interior feasible solutions. Under this setting, we analyze the behavior of the optimal value of the perturbed problem when the perturbation is reduced to zero while keeping the proportion.

First, we demonstrate that, if perturbation is added only to the primal problem to recover strong feasibility, then the optimal value of the perturbed problem converges to the dual optimal value as the perturbation is reduced to zero, even in the presence of nonzero duality gap. An analogous proposition holds for the dual problem. We derive them as a

[1]Strong feasibility is the same as strict feasibility in the literature.

significantly simplified version of the classical asymptotic strong duality theorem (see, for instance, [1, 3, 8, 19, 20, 30] and Chapter 2 of [35]).

Then we analyze the case where perturbation is added to *both* primal and dual sides of the problem. We will demonstrate that in that case the limiting optimal value of the perturbed problems converges to a value between the primal and dual optimal values of the original problem even in the presence of nonzero duality gap. The limiting optimal value is a function of the relative weight of primal and dual perturbations, and reduces monotonically from the primal optimal value to the dual optimal value as the relative weight shifts from the dual side to the primal side.

The result provides an interesting implication to the behavior of infeasible interior-point algorithms applied to general SDPs [12, 13, 21, 23, 26, 37, 41]. In particular, we pick up two well-known polynomial-time infeasible interior-point algorithms by Zhang [41] and Potra and Sheng [26], and prove the following (see Theorems 5 and 6):

1. In the absence of strong infeasibility,
   (a) the algorithms always generate sequences (X^k, S^k, y^k) that are asymptotically primal-dual feasible and such that the duality gap "$X^k \bullet S^k$" converges to zero.
   (b) the sequence of modified (primal and dual) objective values converges to a number between the primal optimal value and the dual optimal value.
2. Otherwise, the algorithms fail to generate a sequence of duality gap converging to zero, so that the duality gap sequence stays uniformly positive. (Needless to say, there's no way to generate an asymptotically primal-dual feasible sequence in this case.)

One implication of the result above is that, at least in theory, these interior-point algorithms generate sequences converging to the optimal value as long as interior feasibility is satisfied at a least one side of the problem. Furthermore, even in the presence of a finite duality gap, they still generate sequences converging to values between the primal and dual optimal values. It is also worth mentioning that our analysis shows that, by setting appropriate initial iterates, it is possible to control how close the limit value will be to the primal or the dual optimal values.

Though this result is more of theoretical interest, this might be of some value if one wants to solve mixed-integer SDP (MISDP) through branch-and-bound and linear SDP relaxations. As discussed in [10], it is quite possible that the relaxations eventually fail to satisfy strong feasibility at a least one of the sides of the problem.

Nevertheless, the solutions obtained by the infeasible interior point methods described above can still be used as bounds to the optimal values of the relaxed linear SDPs regardless of regularity assumptions or constraint qualifications (at least in theory).

This paper is organized as follows. In Section 2, we describe our main results. Section 3 is a preliminary section where we review asymptotic strong duality, infeasible interior-point algorithms, and semialgebraic geometry. In Section 4, we develop a main analysis when both primal and dual problems are perturbed. In Section 5, we apply the developed result to an analysis of the infeasible primal-dual algorithms. In Section 6, illustrative instances will be presented.

2 Main Results

In this section, we introduce our main results after providing the setup and some preliminaries. We also review existing related results.

2.1 Setup and Terminology

First we introduce the notation. The space of $n \times n$ real symmetric matrices will be denoted by $\mathcal{S}^n$. We denote the cone of $n \times n$ real symmetric positive semidefinite matrices and the cone of $n \times n$ real symmetric positive definite matrices by $\mathcal{S}^n_+$ and $\mathcal{S}^n_{++}$. For $U, V \in \mathcal{S}^n$, we define the inner product $U \bullet V$ as $\sum U_{ij} V_{ij}$, and we use $U \succeq 0$ and $U \succ 0$ to denote that $U \in \mathcal{S}^n_+$ and $U \in \mathcal{S}^n_{++}$, respectively. The $n \times n$ identity matrix is denoted by I. We denote the Frobenius norm and the operator norm by $\|X\|_F$ and $\|X\|$. For $v \in \mathbb{R}^k$, we denote by $\|v\|$ its Euclidean norm.

In this paper, we deal with the following standard form primal-dual semidefinite programs

$$\begin{aligned} \mathbf{P}: &\quad \min_X \ C \bullet X \quad \text{s.t. } A_i \bullet X = b_i,\ i = 1, \ldots, m, X \succeq 0 \\ \mathbf{D}: &\quad \max_{y,S} \ b^T y \quad \text{s.t. } C - \sum_{i=1}^m A_i y_i = S,\ S \succeq 0, \end{aligned}$$

where C, $A_i, i = 1, \ldots, m$, X, S are real symmetric $n \times n$ matrices and $y \in \mathbb{R}^m$. For ease of notation, we define the mapping A from $\mathcal{S}^n$ to $\mathbb{R}^m$:

$$A(Y) \equiv (A_1 \bullet Y, \ldots, A_m \bullet Y), \tag{1}$$

and introduce

$$\mathcal{V} \equiv \{X \in \mathcal{S}^n \mid A_i \bullet X = b_i,\ i = 1, \ldots, m\} = \{X \in \mathcal{S}^n \mid A(X) = b\}.$$

We denote by $v(\mathbf{P})$ and $v(\mathbf{D})$ the optimal values of $\mathbf{P}$ and $\mathbf{D}$, respectively. We use analogous notation throughout the paper to denote the optimal value of an optimization problem. For a maximization problem, the optimal value $+\infty$ means that the optimal value is unbounded above and the optimal value $-\infty$ means that the problem is infeasible. For a minimization problem, the optimal value $-\infty$ means that the optimal value is unbounded below and the optimal value $+\infty$ means that the the problem is infeasible.

It is well-known that $v(\mathbf{P}) = v(\mathbf{D})$ holds under suitable regularity conditions, although, in general, we might have $v(\mathbf{P}) \neq v(\mathbf{D})$, i.e., the problem may have a nonzero duality gap. We also note that $v(\mathbf{P})$ and $v(\mathbf{D})$ might not be necessarily attainable.

In general, $\mathbf{P}$ is known to be in one of the following four different mutually exclusive status (see [20]).

1. Strongly feasible: there exists a positive definite matrix satisfying the constraints of $\mathbf{P}$, i.e., $\mathcal{V} \cap \mathcal{S}^n_{++} \neq \emptyset$. This is the same as Slater's condition.

2. Weakly feasible: $\mathbf{P}$ is feasible but not strongly feasible, i.e., $\mathcal{V} \cap \mathcal{S}^n_{++} = \emptyset$ but $\mathcal{V} \cap \mathcal{S}^n_+ \neq \emptyset$.

3. Weakly infeasible: $\mathbf{P}$ is infeasible but the distance between $\mathcal{S}^n_+$ and the affine space $\mathcal{V}$ is zero, i.e., $\mathcal{V} \cap \mathcal{S}^n_+ = \emptyset$ but the zero matrix belongs to the closure of $\mathcal{S}^n_+ - \mathcal{V}$.
4. Strongly infeasible: $\mathbf{P}$ is infeasible but not weakly infeasible. Note that this includes the case where $\mathcal{V} = \emptyset$.

The status of $\mathbf{D}$ is defined analogously by replacing $\mathcal{V}$ by the affine set

$$\{S \in \mathcal{S}^n \mid \exists y \in \mathbb{R}^m, C - \sum_{i=1}^m A_i y_i = S\}.$$

We say that a problem is *asymptotically feasible* if it is either feasible or weakly infeasible. As a reminder, we say that a problem is *singular* if it is either weakly feasible or weakly infeasible.

2.2 Main Result

Now we introduce the main results of this paper. We say that a problem is *asymptotically primal-dual feasible* (or *asymptotically pd-feasible*, in short) if both $\mathbf{P}$ and $\mathbf{D}$ are asymptotically feasible. Evidently, the problem is asymptotically pd-feasible if and only if both $\mathbf{P}$ and $\mathbf{D}$ are feasible or weakly infeasible. The analysis in this paper is conducted mainly under this condition.

Note that asymptotic pd-feasibility is a rather weak condition. All nasty situations such as finite nonzero duality gaps and weak infeasibility of both $\mathbf{P}$ and $\mathbf{D}$ are covered under this condition. Furthermore, since strong infeasibility can be detected by solving auxiliary SDPs that are both primal and dual strongly feasible (see [16]), checking whether a given problem is asymptotically pd-feasible or not can also be checked by solving SDPs that are primal and dual strongly feasible.

We consider the following primal-dual pair $\mathbf{P}(\varepsilon, \eta)$ and $\mathbf{D}(\varepsilon, \eta)$ obtained by perturbing $\mathbf{P}$ and $\mathbf{D}$ with two positive definite matrices I_p and I_d and two nonnegative parameters ε and η:

$$\mathbf{P}(\varepsilon, \eta): \quad \min\ (C + \varepsilon I_d) \bullet X \ \text{ s.t. } A_i \bullet X = b_i + \eta A_i \bullet I_p,\ i = 1, \ldots, m,\ X \succeq 0, \tag{2}$$

and

$$\mathbf{D}(\varepsilon, \eta): \quad \max \sum_{i=1}^m (b_i + \eta A_i \bullet I_p) y_i \ \text{ s.t. } C - \sum_{i=1}^m A_i y_i + \varepsilon I_d = S, \quad S \succeq 0. \tag{3}$$

Using (1), we have

$$\mathbf{P}(\varepsilon, \eta): \quad \min\ (C + \varepsilon I_d) \bullet X \ \text{ s.t. } \mathcal{A}(X) = b + \eta \mathcal{A}(I_p),\ X \succeq 0. \tag{4}$$

While I_p and I_d represent the direction of perturbation, ε and η represent the amount of perturbation. In particular, we could take, for example, $I_p = I_d = I$, where I is the $n \times n$ identity matrix.

If the problem is asymptotically pd-feasible, $\mathbf{D}(\varepsilon, \eta)$ is strongly feasible for any $\varepsilon > 0$ and $\mathbf{P}(\varepsilon, \eta)$ is strongly feasible for any $\eta > 0$. To see the strong feasibility of $\mathbf{P}(\varepsilon, \eta)$, we observe that there always exists $\widetilde{X} \succeq -\eta I_p/2$ satisfying $A_i \bullet \widetilde{X} = b_i, i = 1, \ldots, m$, since $\mathbf{P}$ is weakly infeasible or feasible. Then, we see that the matrix $X = \widetilde{X} + \eta I_p$ is positive definite and a feasible solution to $\mathbf{P}(\varepsilon, \eta)$. We emphasize that the primal-dual pair $\mathbf{P}(\varepsilon, \eta)$ and $\mathbf{D}(\varepsilon, \eta)$ is a natural and possibly one of the simplest regularizations of $\mathbf{P}$ and $\mathbf{D}$ which ensures primal-dual strong feasibility under perturbation.

We define $v(\varepsilon, \eta)$ to be the common optimal value of $\mathbf{P}(\varepsilon, \eta)$ and $\mathbf{D}(\varepsilon, \eta)$ if they coincide. If the optimal values differ, $v(\varepsilon, \eta)$ is not defined. Suppose that $\mathbf{P}$ and $\mathbf{D}$ are asymptotically pd-feasible. In this case, from the the duality theory of convex programs, the function $v(\varepsilon, \eta)$ has the following properties:

1. $v(\varepsilon, \eta)$ is finite if $\varepsilon > 0$ and $\eta > 0$.

2. $v(\varepsilon, 0)$ is well-defined as long as $\varepsilon > 0$ and it takes the value $+\infty$ if $\mathbf{P}$ is infeasible.

3. $v(0, \eta)$ is well-defined as long as $\eta > 0$ and it takes the value $-\infty$ if $\mathbf{D}$ is infeasible.

4. $v(\varepsilon, \eta)$ may not be defined at $(0, 0)$. This is because $\mathbf{P} = \mathbf{P}(0, 0)$ and $\mathbf{D} = \mathbf{D}(0, 0)$ may have different optimal values, i.e., $\mathbf{P}$ and $\mathbf{D}$ may have a nonzero duality gap.

Therefore, although the regularized pair $\mathbf{P}(\varepsilon, \eta)$ and $\mathbf{D}(\varepsilon, \eta)$ satisfies primal-dual strong feasibility if $\varepsilon > 0$ and $\eta > 0$, it is not clear whether this is actually useful in solving SDP under nasty situations such as the presence of nonzero duality gaps. This is precisely one of the main topics of this paper: an analysis on the behavior of the regularized problems without imposing any restrictive assumption.

In this context, it is worth mentioning that the following asymptotic strong duality results

$$\text{(i)}\ \lim_{\varepsilon \downarrow 0} v(\varepsilon, 0) = \lim_{\varepsilon \downarrow 0} v(\mathbf{D}(\varepsilon, 0)) = v(\mathbf{P}) \text{ under dual asymptotic feasibility}$$

and

$$\text{(ii)}\ \lim_{\eta \downarrow 0} v(0, \eta) = \lim_{\eta \downarrow 0} v(\mathbf{P}(0, \eta)) = v(\mathbf{D}) \text{ under primal asymptotic feasibility}$$

are obtained as corollaries of the classical asymptotic strong duality theorem established in the 1950's and 1960's [1, 8]. This theorem received renewed attention with the emergence of conic linear programming; see, for instance, [3, 19, 20, 30] and Chapter 2 of [35]. We will prove (i) and (ii) in the next section (see Theorem 3).

Now we are ready to describe the main results. They are developed to interpolate between (i) and (ii). The first result is the following theorem.

Theorem 1 *Let $\alpha \geq 0$, $\beta \geq 0$ and $(\alpha, \beta) \neq (0, 0)$. If the problem is asymptotically pd-feasible, then $\lim_{t \downarrow 0} v(t\alpha, t\beta)$ exists.*

Here we remark that Theorem 1 includes the case where the limit is $\pm\infty$. Theorem 1 implies that the limit of the optimal value of the perturbed system exists but it is a function of the direction used to approach $(0, 0)$. For $\theta \in [0, \pi/2]$, let us consider the function

$$v_a(\theta) \equiv \lim_{t \downarrow 0} v(t \cos\theta, t \sin\theta),$$

which is the limiting optimal value of $v(\cdot)$ when it approaches zero along the direction making an angle of θ with the ε axis. With that, $v_a(0)$ and $v_a(\pi/2)$ are the special cases corresponding to dual-only perturbation and primal-only perturbation, respectively. So we abuse notation slightly and define

$$v_a(\mathbf{D}) \equiv v_a(0) \quad \text{and} \quad v_a(\mathbf{P}) \equiv v_a(\pi/2). \tag{5}$$

Below is our second main result.

Theorem 2 *If the problem is asymptotically pd-feasible, the following statements hold.*

1. $v_a(0) = v_a(\mathbf{D}) = v(\mathbf{P})$ *and* $v_a(\pi/2) = v_a(\mathbf{P}) = v(\mathbf{D})$.
2. $v_a(\theta)$ *is a monotone decreasing function of* θ.

Now we turn our attention to the connection of these main results to the convergence analysis of the primal-dual infeasible interior-point algorithm. Indeed, the pair (2) and (3) appears often in the analysis of infeasible interior-point algorithms. In particular, primal-dual infeasible interior-point algorithms typically generate a sequence of feasible solutions to $\mathbf{P}(t^k, t^k)$ and $\mathbf{D}(t^k, t^k)$, where I_p and I_d are determined by the initial value of the algorithm and t^k is a positive sequence converging to 0. By Theorem 2 , the common optimal value $v(t^k, t^k)$ of $\mathbf{P}(t^k, t^k)$ and $\mathbf{D}(t^k, t^k)$ converges to $v_a(\pi/4)$ which is between $v(\mathbf{P})$ and $v(\mathbf{D})$. Therefore, if we can show that an infeasible interior-point algorithm generates a sequence which approaches $v(t^k, t^k)$ as $k \to \infty$, we can prove that that sequence converges to $v(\pi/4)$ in the end.

Exploiting this idea, we obtain the following convergence results without any assumption on the feasibility status of the problem. We consider two typical well-known polynomial-time algorithms by Zhang [41] and Potra and Sheng [26]. But the idea can be applied to a broad class of infeasible interior-point algorithms to obtain analogous results. They are stated formally in Theorem 5 and Theorem 6, and summarized as follows:

1. *The algorithms [26, 41] generate asymptotically pd-feasible sequences with the duality gap* $X^k \bullet S^k$ *and* t^k *converging to zero if and only if* $\mathbf{P}$ *and* $\mathbf{D}$ *are asymptotically pd-feasible.*
2. *If* $\mathbf{P}$ *and* $\mathbf{D}$ *are asymptotically pd-feasible, the sequence of modified primal and dual objective values converges to a common value between the primal optimal value* $v(\mathbf{P})$ *and the dual optimal value* $v(\mathbf{D})$ *even in the presence of nonzero duality gap.*

The modified primal and dual objective values mentioned in the statements can be easily computed using the current iterate and do not require any extra knowledge.

If $\mathbf{P}$ and $\mathbf{D}$ are not asymptotically pd-feasible, namely, if one of the problems is strongly infeasible, the algorithms get stuck at a certain point and they fail to generate an asymptotically pd-feasible sequence and fails to drive duality gap and t^k to 0. But the algorithms never fails to generate asymptotically pd-feasible sequences as long as the problems are asymptotically pd-feasible.

We note that Theorems 5 and 6 are to some extent surprising in that infeasible interior-point algorithms work in a meaningful manner without making any restrictive assumptions, at least in theory. This might have interesting implications when solving SDP relaxations arising from hard optimization problems such as MISDP by using infeasible interior-point algorithms. The theorems guarantees that the modified objective function value converges to a value between the primal and dual optimal values. Therefore, the limiting modified objective value can always be used to bound the optimal value of linear SDP relaxations obtained when solving MISDP via, say, branch-and-bound as in [10]. We should mention, however, that if one tries to implement this idea, one would still need to find a way to overcome the severe numerical difficulties that may happen when attempting to solve singular SDPs directly.

2.3 Related Work

There are a number of results on perturbation of semidefinite programs in the literature including [3, 19, 20, 35] which were mentioned in the introduction. The book by Bonnans and Shapiro [3], for instance, has many results on general conic programs that are also applicable to SDPs. However, many of those results require that some sort of constraint qualification holds.

In particular, in Chapter 4 of [3] there is a discussion on a family of optimization problems having the format

$$\min_{x \in X} f(x, u) \quad \text{s.t.} \quad G(x, u) \in \mathcal{K}, \tag{6}$$

where f and G are functions depending on the parameter u and $\mathcal{K}$ is a closed convex set in some Banach space. Denote by $v(u)$, the optimal value of (6). For some fixed u_0, many results are proved about the continuity of $v(\cdot)$ [3, Proposition 4.4], or the directional derivatives of $v(\cdot)$ in a neighborhood of u_0 [3, Theorem 4.24].

However, these existing results do not cover the situations we will deal in this paper. [3, Proposition 4.4], for example, requires a condition called *inf-compactness*, which implies, in particular, that the set of optimal solutions of the problem associated to $v(u_0)$ be compact. [3, Theorem 4.24], on the other hand, requires that the set of optimal solutions associated to $v(u_0)$ be non-empty. In contrast, neither compactness nor non-emptiness is assumed in this paper.

The perturbation we consider is closely related to the infeasible central path appearing in the primal-dual infeasible interior-point algorithms. In fact, we use some properties of the infeasible central path in our proof. The papers [17, 27] showed the analyciticy of the entire trajectory including the end point at the optimal set under the existence of primal-dual optimal solutions satisfying strict complementarity conditions. A very recent paper [34] analyzes the limiting behavior of singular infeasible central paths taking into account the singularity degree. Therein, the authors analyze the speed of convergence under the assumption that the feasible region exists and is bounded. No strong feasibility assumption is made, although we remark that if the feasible region of a primal SDP is non-empty and bounded, then its dual counterpart must satisfy Slater's condition. While their analysis conducts a detailed limiting analysis on the asymptotic behavior of the central path, our analysis deals with the limiting behavior of the optimal value of the perturbed system under weaker assumptions.

In reality, it may be necessary to estimate the error of an approximate optimal solution to a problem with a finite perturbation. In this regard, an interesting and closely related topic to the limiting perturbation analysis is error bounds. The error bound analysis is relatively easy under primal-dual strong feasibility, but it becomes much harder for singular SDPs. See [18,36] for SDP and SOCP, and [15] for a more general class of convex programs. The relationship between forward and backward errors of a semidefinite feasibility system is closely related to its singular degree, which, roughly, is defined as the number of facial reduction steps necessary for regularizing the problem. Recently, some analysis of limiting behaviors of the external (or infeasible) central path involving singularity degree is developed in [34]. Finally, we mention [32] which conducted a sensitivity analysis of SDP under perturbation of the coefficient matrices "A_i".

3 Preliminaries

In this section, we introduce three ingredients of this paper, namely, asymptotic strong duality, infeasible interior-point algorithms and real-algebraic geometry.

3.1 Asymptotic Strong Duality

A main difference between the duality theory in linear programming and general convex programming is that the latter requires some regularity conditions for strong duality to hold. If such regularity condition is violated, then the primal and dual may have nonzero duality gap [28]. Nevertheless, the so-called *asymptotic strong duality* holds even in such singular cases [1, 3, 8, 19, 20, 30, 35]. Here we quickly review the result and work on it a bit to derive a modified and simplified version suitable for our purposes.

Let a-val($\mathbf{P}$) and a-val($\mathbf{D}$) be

$$\begin{aligned} \text{a-val}(\mathbf{P}) &\equiv \lim_{\varepsilon\downarrow 0} \inf_{\|\Delta b\|<\varepsilon} \inf\{C \bullet X|\ A(X) = b + \Delta b,\ X \succeq 0\}, \\ \text{a-val}(\mathbf{D}) &\equiv \lim_{\varepsilon\downarrow 0} \sup_{\|\Delta C\|<\varepsilon} \sup\{b^T y|\ C + \Delta C - \sum_i A_i y_i \succeq 0\}. \end{aligned} \tag{7}$$

Here, a-val($\mathbf{P}$) and a-val($\mathbf{D}$) are called the *asymptotic optimal values of* $\mathbf{P}$ *and* $\mathbf{D}$, respectively [30]. (It is also called *subvalue* in [1, 3, 8, 19, 20].) The following asymptotic duality theorem holds.

Theorem[Asymptotic Duality Theorem, e.g., Theorem 3.2.4 in [30]]

1. *If* $\mathbf{P}$ *is asymptotically feasible, then,* a-val($\mathbf{P}$) $= v(\mathbf{D})$.
2. *If* $\mathbf{D}$ *is asymptotically feasible, then,* a-val($\mathbf{D}$) $= v(\mathbf{P})$.

Note that the Asymptotic Duality Theorem includes the cases where a-val$(\cdot) = \pm\infty$.

Now we develop a simplified version of the Asymptotic Duality Theorem. Let $\varepsilon \geq 0$, and let $\mathbf{D}(\varepsilon)$ be $\mathbf{D}(\varepsilon, 0)$, i.e., the relaxed dual problem

$$\max\ b^T y\ \text{ s.t. } C - \sum_{i=1}^m A_i y_i + \varepsilon I_d = S, \quad S \succeq 0. \tag{8}$$

According to the notation introduced in Section 2.2, the optimal value of (8) is written as $v(\varepsilon, 0)$. Recall also that

$$\lim_{\varepsilon \downarrow 0} v(\varepsilon, 0) = v_a(0) = v_a(\mathbf{D}).$$

Next we consider an analogous relaxation at the primal side. Notice that (8) is obtained by shifting the semidefinite cone by $-\varepsilon I_d$. The analogous perturbation of the primal problem is given by

$$\min\ C \bullet \widetilde{X} \quad \text{s.t. } A(\widetilde{X}) = b,\ \widetilde{X} \succeq -\eta I_p, \tag{9}$$

where $\eta \geq 0$. Letting $X \equiv \widetilde{X} + \eta I_p$, we obtain

$$\min\ C \bullet X - \eta C \bullet I_p \quad \text{s.t. } A(X) = b + \eta A(I_p), \quad X \succeq 0. \tag{10}$$

The optimal value of (9) is monotone decreasing in η, because the feasible region enlarges as η is increased (strictly speaking, it does not shrink). Observe also that this problem is $\mathbf{P}(0, \eta)$ with the objective function shifted by a constant $-\eta C \bullet I_p$. Since this constant vanishes as $\eta \to 0$, we obtain

$$v_a\left(\frac{\pi}{2}\right) = v_a(\mathbf{P}) = \lim_{\eta \downarrow 0} \{\text{The optimal value of (10)}\}.$$

Now we prove the following result.

Theorem 3 *The following statements hold.*

1. *If $\mathbf{D}$ is asymptotically feasible, then $v_a(0) = v_a(\mathbf{D}) = v(\mathbf{P})$.*
2. *If $\mathbf{P}$ is asymptotically feasible, then $v_a(\pi/2) = v_a(\mathbf{P}) = v(\mathbf{D})$.*

Proof. Recall that by definition (see (5)), we have $v_a(0) = v_a(\mathbf{D})$ and $v_a(\pi/2) = v_a(\mathbf{P})$.

First we show that $v_a(\mathbf{D}) = v(\mathbf{P})$. From the Asymptotic Duality Theorem, a-val$(\mathbf{D}) = v(\mathbf{P})$ holds including the special cases where a-val$(\mathbf{D}) = \pm\infty$. We observe that a-val$(\mathbf{D})$ satisfies

$$\text{a-val}(\mathbf{D}) = \lim_{\varepsilon \downarrow 0} \sup_{y, \Delta C} \{b^T y \mid C + \Delta C - \sum_i A_i y_i \succeq 0,\ \|\Delta C\| \leq \varepsilon\},$$

where $\|\Delta C\| < \varepsilon$ in (7) is changed to $\|\Delta C\| \leq \varepsilon$.

Since $v_a(0)$ is obtained by restricting the condition on ΔC from "$\|\Delta C\| \leq \varepsilon$" to "$\Delta C = \varepsilon I_d / \|I_d\|$", we obtain $v_a(0) \leq$ a-val$(\mathbf{D}) = v_a(\mathbf{D})$. We also have the converse inequality $v_a(0) \geq v_a(\mathbf{D})$ because

$$\begin{aligned}
\text{a-val}(\mathbf{D}) &= \lim_{\varepsilon \downarrow 0} \sup\{b^T y \mid C + \Delta C - \sum_i A_i y_i \succeq 0,\ \|\Delta C\| \leq \varepsilon\} \\
&= \lim_{\varepsilon \downarrow 0} \sup\{b^T y \mid C + \Delta C - \sum_i A_i y_i \succeq 0,\ -\varepsilon I \preceq \Delta C \preceq \varepsilon I\} \\
&\leq \lim_{\varepsilon \downarrow 0} \sup\{b^T y \mid C + \Delta C - \sum_i A_i y_i \succeq 0,\ \Delta C \preceq \varepsilon I\} \\
&\leq \lim_{\varepsilon \downarrow 0} \sup\{b^T y \mid C + \Delta C - \sum_i A_i y_i \succeq 0,\ \Delta C \preceq \varepsilon \|I_d^{-1}\| I_d\} \\
&\leq \lim_{\varepsilon \downarrow 0} \sup\{b^T y \mid C + \varepsilon I_d - \sum_i A_i y_i \succeq 0\} = v_a(\mathbf{D}).
\end{aligned}$$

Here we used $I \preceq \|I_d^{-1}\| I_d$ for the second inequality. The proof of item 1 is complete.

We proceed to prove item 2. From the Asymptotic Duality Theorem again, we have $v(\mathbf{D}) = \text{a-val}(\mathbf{P})$. Hence, for the sake of proving assertion 2, it suffices to show that $v_a(\mathbf{P}) = \text{a-val}(\mathbf{P})$. The proof of the inequality $v_a(\mathbf{P}) \geq \text{a-val}(\mathbf{P})$ is analogous to the proof for $v_a(\mathbf{D}) \leq \text{a-val}(\mathbf{D})$. We will now show the converse inequality. If $\text{a-val}(\mathbf{P}) = +\infty$, then $v_a(\mathbf{P}) \geq \text{a-val}(\mathbf{P})$ implies that $v_a(\mathbf{P}) = +\infty$. Therefore, in what follows we assume that $\text{a-val}(\mathbf{P}) < +\infty$.

By assumption, $\mathbf{P}$ is not strongly infeasible (see Section 2.1). By the definition of $\text{a-val}(\mathbf{P})$, for every $\varepsilon > 0$ sufficiently small, there exist X_ε and Δb_ε such that $\|\Delta b_\varepsilon\| \leq \varepsilon$, X_ε is feasible to "$A(X) = b + \Delta b_\varepsilon,\ X \succeq 0$", and

$$\text{a-val}(\mathbf{P}) = \lim_{\varepsilon \downarrow 0} C \bullet X_\varepsilon. \tag{11}$$

Note that this is still valid even when $\text{a-val}(\mathbf{P}) = -\infty$.

In addition, the fact that $\mathbf{P}$ is not strongly infeasible implies the existence of a solution to the system "$A(X') = b$". As a consequence, "$A(Y) = \Delta b_\varepsilon$" too has a solution when Δb_ε is as described above. Otherwise, "$A(X) = b + \Delta b_\varepsilon$" is infeasible, contradicting the existence of X_ε above.

Then, there exists $M > 0$ depending only on A such that "$A(Y) = \Delta b_\varepsilon$" has a solution with norm bounded by $M\|\Delta b_\varepsilon\|$, see this footnote[2]. Let Y_ε be one such solution. Therefore, $\|Y_\varepsilon\| \leq M\|\Delta b_\varepsilon\| \leq M\varepsilon$ for each sufficiently small $\varepsilon > 0$ and hence

$$\lim_{\varepsilon \downarrow 0} \|Y_\varepsilon\| = 0. \tag{12}$$

Observing that $\|I_p^{-1}\| I_p \succeq I$ and $\|Y_\varepsilon\| I - Y_\varepsilon \succeq 0$ yield $\|Y_\varepsilon\| \|I_p^{-1}\| I_p - Y_\varepsilon \succeq 0$, we let

$$X'_\varepsilon \equiv X_\varepsilon + \|Y_\varepsilon\| \|I_p^{-1}\| I_p - Y_\varepsilon.$$

With that, X'_ε is positive semidefinite and is a feasible solution to $\mathbf{P}(0, \eta)$ with $\eta = \|Y_\varepsilon\| \|I_p^{-1}\|$ (see (4)). Furthermore,

$$|C \bullet X_\varepsilon - C \bullet X'_\varepsilon| = |C \bullet (\|Y_\varepsilon\| \|I_p^{-1}\| I_p - Y_\varepsilon)| \leq 2\|C\| \|Y_\varepsilon\| \|I_p^{-1}\| \|I_p\|_F, \tag{13}$$

which approaches 0 by driving $\varepsilon \to 0$ because of (12).

We are now ready to show the desired assertion. Notice that we have $\lim_{\varepsilon \downarrow 0} C \bullet X'_\varepsilon \geq v_a(\mathbf{P})$, since X'_ε is feasible to $\mathbf{P}(0, \|Y_\varepsilon\| \|I_p^{-1}\|)$ and (12) holds. This fact combined with (11) and (13) implies $v_a(\mathbf{P}) \leq \text{a-val}(\mathbf{P})$. The proof is complete. ■

Compared with the classical asymptotic duality results [1,3,8,19,20,30,35], the perturbation considered in (8) and (10) is simpler, but this difference is crucial for our results. In the first version of this article, Theorem 3 was proved directly from the Slater strong duality theorem without employing the Asymptotic Duality Theorem.

[2] Let $\mathcal{V}$ denote the set of solutions to "$A(Y) = \Delta b$" and let S be a symmetric matrix. Denote by $\text{dist}\,(S, \mathcal{V})$ the Euclidean distance between S and $\mathcal{V}$. Hoffman's lemma (e.g., [11, Theorem 11.26]) says that there exists a constant M depending on A but not on Δb such that for every S, we have that $\text{dist}\,(S, \mathcal{V})$ is bounded above by $M\|\Delta b - A(S)\|$. Taking $S = 0$, we conclude the existence of Y satisfying $A(Y) = \Delta b$ and $\|Y\| \leq M\|\Delta b\|$.

3.2 Infeasible Primal-dual Interior-point Algorithms

We introduce some basic concepts of infeasible primal-dual interior-point algorithms for SDP [26, 38, 40, 41]. This is because our analysis leads to a novel convergence property of the infeasible primal-dual interior-point algorithms when applied to singular problems. We also need some theoretical results about infeasible interior-point algorithms in the proof of Theorem 1. In this subsection, we assume that A_i $(i = 1, \ldots, m)$ are linearly independent. This assumption is not essential but to ensure uniqueness of y and Δy in the system of equations of the form $S = \sum_i A_i y_i + C'$ and $\Delta S = \sum_i A_i \Delta y + R'$ with respect to (S, y) and $(\Delta S, \Delta y)$, respectively, where C' and R' are constants, which appear throughout the analysis.

3.2.1 Outline of infeasible primal-dual interior-point algorithms

Primal-dual interior-point methods for **P** and **D** are based on the following optimality conditions:

$$XS = 0, \quad C - \sum_i A_i y_i = S, \quad A(X) = b \quad X \succeq 0, \quad S \succeq 0. \tag{14}$$

Rather than solving this system directly, a relaxed problem

$$XS = \nu I, \quad C - \sum_i A_i y_i = S, \quad A(X) = b, \quad X \succeq 0, \quad S \succeq 0, \tag{15}$$

is considered, where $\nu > 0$. The algorithm solves (14) by solving (15) approximately and reducing ν gradually to zero repeatedly. This amounts to following the central path

$$\{(X_\nu, S_\nu, y_\nu) \mid (X, S, y) = (X_\nu, S_\nu, y_\nu) \text{ is a solution to (15)}, \ \nu \in (0, \infty]\} \tag{16}$$

towards "$\nu = 0$". Let us take a closer look at the algorithm proposed by Zhang [41].

Let (X, S, y) be the current iterate such that $X \succ 0$ and $S \succ 0$. The method employs the Newton direction to solve the system (15). More precisely, the first equation $XS = \nu I$ is replaced with an equivalent symmetric reformulation

$$\Phi(X, S) = \frac{1}{2}(PXSP^{-1} + P^{-1}SXP) = \nu I, \tag{17}$$

where P is a constant nonsingular matrix. In Zhang's algorithm, the constant matrix P is set to $X^{-1/2}$. Then we consider a modified nonlinear system of equations to (15) where $XS = \nu I$ is replaced with (17). The Newton direction $(\Delta X, \Delta S, \Delta y)$ for that modified system at the point (X, S, y) is the unique solution to the following system of linear equations.

$$\Phi(X, S) + L_\Phi(\Delta X, \Delta S) = \nu I, \ C - \sum_i A_i(y_i + \Delta y_i) = S + \Delta S, \ A(X + \Delta X) = b, \tag{18}$$

where L_Φ is a linearization of $\Phi(X, S)$.

Starting from the kth iterate $(X^k, S^k, y^k) = (X, S, y)$, the next iterate $(X^{k+1}, S^{k+1}, y^{k+1})$ is determined as:

$$(X^{k+1}, S^{k+1}, y^{k+1}) = (X^k, S^k, y^k) + s^k(\Delta X, \Delta S, \Delta y). \tag{19}$$

The stepsize $0 < s^k \leq 1$ is chosen not only so that X^{k+1} and S^{k+1} are strictly positive but also carefully so that they stay close to the central path in order to ensure good convergence properties. Then ν is updated appropriately and the iteration continues.

Now we briefly describe another representative polynomial-time infeasible primal-dual interior-point algorithm developed by Potra and Sheng [26]. Let (X^0, S^0, y^0) be a point satisfying $X^0 \succ 0$ and $S^0 \succ 0$ and consider the path defined as follows.

$$\{(X, S, y) \mid XS = tI, \quad C - \sum A_i y_i - S = t(C - \sum A_i y_i^0 - S^0),$$
$$A(X) - b = t(A(X^0) - b),\ X \succeq 0,\ S \succeq 0,\ t \in (0, 1]\}. \tag{20}$$

The algorithm follows this path by driving $t \to 0$ and using a predictor-corrector method.

We note that polynomial-time convergence is proved for both algorithms [26,41] assuming the existence of optimal solutions (X^*, S^*, y^*) to **P** and **D**. In the analysis, the initial iterate (X^0, S^0, y^0) is set to $(\rho_0 I, \rho_1 I, 0)$ where ρ_0 and ρ_1 are selected to be large enough in order to satisfy the conditions $X^0 - X^* \succ 0$ and $S^0 - S^* \succ 0$. Although the polynomial convergence analysis was conducted using this initial iterate, the algorithms themselves can be applied to any SDP problem by choosing (X^0, S^0, y^0) such that $X^0 \succ 0$ and $S^0 \succ 0$ as the initial iterate.

In many practical implementations of the algorithm [38,40], they take different stepsizes in the primal and dual space for the sake of practical efficiency. For simplicity of presentation, we only analyze the case (19) which corresponds to the situation where we take the same stepsize in the primal-dual space.

The following well-known property connects Theorems 1 and 2 to the analysis of infeasible interior-point algorithms.

Proposition 1 *Let $X^0 \succ 0$ and $S^0 \succ 0$, and let $\{(X^k, S^k, y^k)\}$ be a sequence generated by the primal-dual infeasible interior-point algorithms in [26, 41] with initial iterate (X^0, S^0, y^0). Let $I'_d \equiv S^0 - (C - \sum A_i y_i^0)$ and let $I'_p \equiv X^0 - \widetilde{X}$ where $A(\widetilde{X}) = b$. Then, there exists a nonnegative sequence $\{t^k\}$ such that the following equations hold:*

$$(C + t^k I'_d) - \sum_i A_i y^k = S^k, \quad A(X^k) = b + t^k A(I'_p). \tag{21}$$

(*cf. The linear equality constraints of* (2) *and* (3))

Proof. This result is a fundamental tool used in the analysis of the algorithms in [26,41]. For the sake of completeness, here we prove the result only for Zhang's algorithm.

We prove the first relation of (21) by induction. For $k = 0$, the proposition holds by taking $t^0 \equiv 1$. Suppose that the relation (21) holds for k, then, the search direction $(\Delta X, \Delta S, \Delta y)$ is the solution to the linear system of equations (18) with $(X, S, y) = (X^k, S^k, y^k)$. Because of the second equation of (18), we have

$$C - \sum A_i(y_i^k + \Delta y_i) - (S^k + \Delta S) = 0.$$

Therefore,

$$C - \sum A_i(y_i^k + s^k \Delta y_i) - (S^k + s^k \Delta S) = (1 - s^k)(C - \sum A_i y_i^k - S^k).$$

Since $y_i^{k+1} = y_i^k + s^k \Delta y_i, S^{k+1} = S^k + s^k \Delta S$ and $t^{k+1} = (1 - s^k)t^k$, we obtain

$$C - \sum A_i y_i^{k+1} - S^{k+1} = (1 - s^k)t^k I'_d = t^{k+1} I'_d$$

as we desired, because $C - \sum A_i y_i^k - S^k = t^k(C - \sum A_i y_i^0 - S^0) = t^k I'_d$ holds by the induction assumption. The primal relation, i.e., the right side in (21), follows similarly. ■

Remark In view of Proposition 1, by convention, we treat t^k as a part of iterates of the algorithms. By its construction, we have $t^0 = 1$ and

$$t^{k+1} = \prod_{l=0}^{k} (1 - s^l) \tag{22}$$

for $k = 0, 1, \ldots$

3.2.2 Path formed by points on the central path of perturbed problems

We fix ν to be a positive number, and write the solution of the following system as $w_\nu(t) \equiv (X_\nu(t), S_\nu(t), y_\nu(t))$:

$$\begin{aligned} XS - \nu I = 0, \quad C + t\alpha I_d - \sum_i A_i y_i - S = 0, \\ A(X - t\beta I_p) - b = 0, \\ X \succeq 0, \quad S \succeq 0, \ t > 0. \end{aligned} \tag{23}$$

If the problem is asymptotically pd-feasible, for any $t > 0$, then the solution of (23) defines a point on the central path of the primal-dual pair of strongly feasible SDP:

$$\min \ (C + t\alpha I_d) \bullet X \quad \text{s.t. } A(X - t\beta I_p) = b, \quad X \succeq 0 \tag{24}$$

and

$$\max \sum_i (b_i + t\beta A_i \bullet I_p) y_i \quad \text{s.t. } C + t\alpha I_d - \sum_i A_i y_i = S, \quad S \succeq 0. \tag{25}$$

In this case, $w_\nu(t)$ is ensured to exist and is uniquely determined for all $t \in (0, \infty)$ (due to the assumption of linear independence of A_i, $i = 1, \ldots, m$). Moreover, the set

$$\mathcal{C} \equiv \{w_\nu(t) \mid t \in (0, \infty)\} \tag{26}$$

forms an analytic path running through $\mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \times \mathbb{R}^m$. The existence and analyticity of $\mathcal{C}$ is a folklore result (e.g., [17,27]), but we outline a proof in the Appendix A based on a result in [22][3]. We also note that a special case where $\nu = 1$ and $C = 0$ is analyzed in [33] in the context of facial reduction.

[3] Essentially the existence and analyticity of the path just relies on local conditions, so, the existence of optimal solutions of **P** and **D** is not necessary.

Since $A(X_\nu(t)) = b + t\beta A(I_p)$, $C + t\alpha I_d - \sum_i A_i y_{\nu i}(t) = S_\nu(t)$, and $X_\nu(t)S_\nu(t) = \nu I$ hold, we have

$$\begin{aligned}
0 &\leq (C + t\alpha I_d) \bullet X_\nu(t) - \sum_{i=1}^{m} (b_i + t\beta A_i \bullet I_p) y_{\nu i}(t) \\
&= (C + t\alpha I_d) \bullet X_\nu(t) - \sum_{i=1}^{m} A_i \bullet X_\nu(t) y_{\nu i}(t) \\
&= S_\nu(t) \bullet X_\nu(t) = \mathrm{Tr}(X_\nu(t)S_\nu(t)) = \mathrm{Tr}(\nu I) = n\nu. \qquad (27)
\end{aligned}$$

Let us denote by $v_{\mathrm{opt}}(t)$ the common optimal value of (24) and (25). Since $v_{\mathrm{opt}}(t)$ is between $(C + t\alpha I_d) \bullet X_\nu(t)$ and $\sum_{i=1}^{m}(b_i + t\beta A_i \bullet I_p)y_{\nu i}(t)$, i.e.,

$$v_{\mathrm{opt}}(t) \in \left[\sum_{i=1}^{m} (b_i + t\beta A_i \bullet I_p) y_{\nu i}(t), (C + t\alpha I_d) \bullet X_\nu(t) \right] \qquad (28)$$

holds by weak duality, we see, together with (27), that

$$0 \leq (C + t\alpha I_d) \bullet X_\nu(t) - v_{\mathrm{opt}}(t) \leq n\nu \qquad (29)$$

holds for each $t > 0$.

3.3 Semialgebraic sets and the Tarski-Seidenberg Theorem

A set S in $\mathbb{R}^k$ is called *basic semialgebraic* if it can be written as the set of solutions of finitely many polynomial equalities and strict polynomial inequalities. Then, a set is said to be *semialgebraic* if it is a union of finitely many basic semialgebraic sets. In particular, a semialgebraic set in $\mathbb{R}$ is a union of finitely many points and intervals. For $x = (x_1, \ldots, x_k) \in \mathbb{R}^k$, let $T(x)$ be a coordinate projection to $\mathbb{R}^{k-1}$ defined as $T(x) \equiv (x_2, \ldots, x_n)$. The Tarski-Seidenberg Theorem states that a coordinate projection of a semialgebraic set is again a semialgebraic set in the lower-dimensional space, and described as follows.

Tarski-Seidenberg Theorem (e.g. Theorem 2.2.1 of [2])
Let $W \subseteq \mathbb{R}^k$ be a semialgebraic set. Then, $T(W)$ is a semialgebraic set in $\mathbb{R}^{k-1}$.

4 Proof of the Main Results

In this section, we prove Theorems 1 and 2. We start with some basic properties of $v(\varepsilon, \eta)$.

Proposition 2 *If the problem is asymptotically pd-feasible, the following statements hold.*

1. *$v(\varepsilon, \eta)$ is well-defined for all $(\varepsilon, \eta) \geq 0$ except for $(0, 0)$. Furthermore,*

   (i) $\lim_{\varepsilon \downarrow 0} v(\varepsilon, 0) = v(\mathbf{P})$ *and*

   (ii) $\lim_{\eta \downarrow 0} v(0, \eta) = v(\mathbf{D})$

   hold including the cases where their values are $\pm\infty$.

2. $v(\varepsilon, \eta)$ *is a monotone increasing function in* ε*.*

3. $v_P(\varepsilon, \eta) \equiv v(\varepsilon, \eta) - \eta C \bullet I_p - \eta\varepsilon I_d \bullet I_p$ *is a monotone decreasing function in* η*.*

Proof. Item 1 follows directly from Theorem 3. Item 2 follows because the feasible region becomes larger as ε is increased. To see item 3, observe that the optimal value of (10) (or, equivalently, (9)) is monotone decreasing in η. Letting $C \equiv C + \varepsilon I_d$ in (10), we obtain

$$\min\ (C + \varepsilon I_d) \bullet X - \eta(C + \varepsilon I_d) \bullet I_p \ \text{ s.t. } \mathcal{A}(X) = b + \eta\mathcal{A}(I_p),\ X \succeq 0.$$

The optimal value of this problem is monotone decreasing in η. Since this problem differs from $\mathbf{P}(\varepsilon, \eta)$ by a constant $(\eta(C \bullet I_p) + \eta\varepsilon I_d \bullet I_p)$ in the objective function, the statement follows. ∎

In the following, we prove Theorem 1. The theorem claims that, even though $v(0, 0)$ is not well-defined, the limiting value exists when approaching $(0, 0)$ along a straight line emanating from the origin to any direction of the first orthant.

(Proof of Theorem 1)

Although the result holds even if the A_i's are linearly dependent, for simplicity sake, in this proof we assume linear independence of the A_i $(i = 1, \ldots, m)$. In addition, we write $v(t\alpha, t\beta)$ as $v_{\text{opt}}(t)$, since $v(t\alpha, t\beta)$ is the common optimal value to the primal-dual pair $\mathbf{P}(t\alpha, t\beta)$ and $\mathbf{D}(t\alpha, t\beta)$. We also assume that $\alpha > 0$ and $\beta > 0$, since the proof for the case where either of α and β is 0 (but $(\alpha, \beta) \neq 0$) has already been established in Proposition 2.

Recall that we introduced the analytic path $\mathcal{C}$ in Section 3.2.2 (See (23)–(26)). We follow the same notation described therein. The path $\mathcal{C}$ is parametrized by t. We divide the proof into the following two steps:

(Step 1) For any fixed $\nu > 0$, we prove the monotonicity of $(C + t\alpha I_d) \bullet X_\nu(t)$ when $t > 0$ is sufficiently small.
(Step 2) Prove the existence of $\lim_{t\downarrow 0} v_{\text{opt}}(t)$.

(Step 1)

We analyze the behavior of $(C + t\alpha I_d) \bullet X_\nu(t)$ along the path $\mathcal{C}$ as $t \to 0$. Recall that (23) is the system parametrized by t which defines the path $\mathcal{C}$. By differentiating the three equations in (23) with respect to t, we see that the following system of equations in $(t, X, S, y, \delta X, \delta S, \delta y)$ (with semidefinite constraints on X and S) has a unique solution[4]

$$(t, X, S, y, \delta X, \delta S, \delta y) = \left(t, X_\nu(t), S_\nu(t), y_\nu(t), \frac{dX_\nu(t)}{dt}, \frac{dS_\nu(t)}{dt}, \frac{dy_\nu(t)}{dt}\right)$$

[4]By the discussion in Section 3.2.2, for fixed $t > 0$, $X_\nu(t), S_\nu(t)$ are uniquely defined. Since the A_i are linearly independent, $y_{\nu(t)}$ must be unique as well. In order to see that $\delta X, \delta S, \delta y$ are also uniquely determined, we take a look at the first three equations of (30) for fixed positive definite matrices X and S. They become linear equations in $\delta X, \delta S, \delta y$ and determine a unique solution if and only if the kernel of $\phi : (U, V, z) \mapsto (XU + VS, V + \sum_i A_i z_i, \mathcal{A}(U))$ is trivial. Suppose $\phi(U, V, z) = 0$. Then, $U \bullet V = 0$. Considering the first component of ϕ, we have the equation $XU = -VS$, which implies that $\nu U = -SVS$. Taking the inner product with V, we obtain $0 = (SVS) \bullet V = \|S^{1/2}VS^{1/2}\|_F^2$. Therefore, $S^{1/2}VS^{1/2} = 0$ and since S is invertible, $V = 0$. By $\nu U = -SVS$, we have $U = 0$.

for each $t \in (0, \infty)$:

$$
\begin{array}{ll}
X\delta S + \delta X S = 0, & \\
\alpha I_d - \sum_i A_i \delta y_i = \delta S, & \\
A_i \bullet (\delta X - \beta I_p) = 0, \qquad (i = 1, \ldots, m), & \\
XS = \nu I, & \\
C + t\alpha I - \sum_i A_i y_i = S, & \\
A_i \bullet (X - t\beta I_p) = b_i, \qquad (i = 1, \ldots, m), & \\
X \succeq 0, \;\; S \succeq 0, t > 0, &
\end{array} \tag{30}
$$

that is, (30) is a system of equations with semidefinite constraints which determines the curve $(X_\nu(t), S_\nu(t), y_\nu(t))$ and its tangent $\left(\frac{dX_\nu(t)}{dt}, \frac{dS_\nu(t)}{dt}, \frac{dy_\nu(t)}{dt}\right)$.

Let us denote by $\mathcal{D}$ the set of solutions to (30) as follows:

$$\mathcal{D} = \{(t, X, S, y, \delta X, \delta S, \delta y) \mid (t, X, S, y, \delta X, \delta S, \delta y) \text{ satisfies (30).}\}$$

Each element of $\mathcal{D}$ can be seen as a pair consisting of a point on $\mathcal{C}$ and its tangent. Since the semidefinite conditions $S \succeq 0$ and $X \succeq 0$ can be written as the solution set of finitely many polynomial inequalities, $\mathcal{D}$ is a semialgebraic set.

Now we claim that $(C + t\alpha I_d) \bullet X_\nu(t)$ is either monotonically increasing or monotonically decreasing for sufficiently small t. To this end, we analyze the set of local minimum points and local maximum points of $(C + t\alpha I_d) \bullet X_\nu(t)$ over $(0, \infty)$. A necessary condition for local and maximum points is:

$$\frac{d(C + t\alpha I_d) \bullet X_\nu(t)}{dt} = (C + \alpha I_d) \bullet \frac{dX_\nu(t)}{dt} + \alpha I_d \bullet X_\nu(t) = 0.$$

For every $\hat{t} > 0$, $(\frac{dX_\nu}{dt}(\hat{t}), \frac{dS_\nu}{dt}(\hat{t}), \frac{dy_\nu}{dt}(\hat{t}))$ is the tangent part $(\delta X, \delta S, \delta y)$ of the unique solution to (30) with $t = \hat{t}$. Therefore, a necessary condition for $(C + t\alpha I_d) \bullet X_\nu(t)$ to have an extreme value at t is that t is in the set

$$\mathcal{T}_1 \equiv \{t \mid (t, X, S, y, \delta X, \delta S, \delta y) \in \mathcal{T}\},$$

where

$$\mathcal{T} \equiv \{(t, X, S, y, \delta X, \delta S, \delta y) \in \mathcal{D} \mid \;\; (C + t\alpha I_d) \bullet \delta X + \alpha I_d \bullet X = 0\}.$$

Since $\mathcal{D}$ is a semialgebraic set, so is $\mathcal{T}$. Since $\mathcal{T}_1$ is the projection of $\mathcal{T}$ onto the t coordinate, by applying the Tarski-Seidenberg Theorem, we see that $\mathcal{T}_1$ is a semialgebraic set.

Thus, $\mathcal{T}_1$ is a semialgebraic set contained in $\mathbb{R}$, therefore $\mathcal{T}_1$ can be expressed as a union of finitely many points and intervals over $\mathbb{R}$. Since $(C + t\alpha I_d) \bullet X_\nu(t)$ is an analytic function (see Section 3.2.2), the same is true for its derivatives. Therefore, if $\mathcal{T}_1$ contains an interval, then the derivative of $(C + t\alpha I_d) \bullet X_\nu(t)$ with respect to t must, in fact, be zero throughout $(0, \infty)$[5]. In particular, $(C + t\alpha I_d) \bullet X_\nu(t)$ is constant for all $t > 0$. Thus, $(C + t\alpha I_d) \bullet X_\nu(t)$ is a monotonically increasing/decreasing function in this case.

Now we deal with the case where $\mathcal{T}_1$ consists of a finite number of points only. We recall that $(C + t\alpha I_d) \bullet X_\nu(t)$ takes an extreme value at t only if $t \in \mathcal{T}_1$. This implies that the

[5]Here we are using the fact that the zero function is analytic and if two real analytic functions $f : (0, \infty) \to \mathbb{R}$, $g : (0, \infty) \to \mathbb{R}$ coincide in some interval (a, b) with $a < b$, then f and g coincide throughout $(0, \infty)$, e.g., [14, Corollary 1.2.6].

number of extreme points of $(C + t\alpha I_d) \bullet X_\nu(t)$ is finite and hence $(C + t\alpha I_d) \bullet X_\nu(t)$ is monotonically increasing or monotonically decreasing for sufficiently small t.

(Step 2)

It follows from Step 1 that there are three possibilities.

(i) $\lim_{t\downarrow 0}(C + t\alpha I_d) \bullet X_\nu(t) = \infty$,

(ii) $\lim_{t\downarrow 0}(C + t\alpha I_d) \bullet X_\nu(t) = -\infty$,

(iii) $\lim_{t\downarrow 0}(C + t\alpha I_d) \bullet X_\nu(t)$ is a finite value.

First we consider cases (i) and (ii). Recalling (29), we have $|(C+t\alpha I_d)\bullet X_\nu(t)-v_{\rm opt}(t)| \leq n\nu$. Therefore, $v_{\rm opt}(t)$ diverges to $+\infty$ and $-\infty$, respectively. This corresponds to the case of the theorem where the limit is $\pm\infty$.

Next, we proceed to case (iii). In this case, $v_{\rm opt}(t)$ is bounded for sufficiently small $t > 0$ because $|v_{\rm opt}(t) - (C + t\alpha I_d) \bullet X_\nu(t)| \leq n\nu$ and $(C + \alpha I_d) \bullet X_\nu(t)$ is bounded for sufficiently small $t > 0$. Therefore, there exist three constants M_1, M_2, and $\bar{t} > 0$ such that $M_1 < M_2$ and $\bar{t} > 0$ for which

$$v_{\rm opt}(t) \in [M_1, M_2] \quad \text{if } t \in (0, \bar{t}].$$

For the sake of obtaining a contradiction, we assume that $v_{\rm opt}(t)$ does not have a limit as $t \to 0$. Then, there exists an infinite sequence $\{t^k\}$ with $\lim_{k\to\infty} t^k \to 0$ where $\{v_{\rm opt}(t^k)\}$ has two distinct accumulation points, v_1 and v_2, say. Without loss of generality, we let $v_1 > v_2$ and $z \equiv v_1 - v_2$.

Let $\tilde{\nu} \equiv z/(6n)$. By Step 1, it follows that $(C + t\alpha I_d) \bullet X_{\tilde{\nu}}(t)$ is a monotone function for sufficiently small $t > 0$. Furthermore, since $v_{\rm opt}(t)$ is bounded for sufficiently small t, (29) implies that $(C + t\alpha I_d) \bullet X_{\tilde{\nu}}(t)$ does not diverge and has a limit as $t \downarrow 0$. Let us denote by $c^*_{\tilde{\nu}}$ the limit value, and let $\tilde{t} > 0$ be such that

$$|(C + t\alpha I_d) \bullet X_{\tilde{\nu}}(t) - c^*_{\tilde{\nu}}| \leq \frac{z}{6} \tag{31}$$

holds for any $t \in (0, \tilde{t}]$. On the other hand,

$$|(C + t\alpha I_d) \bullet X_{\tilde{\nu}}(t) - v_{\rm opt}(t)| = (C + t\alpha I_d) \bullet X_{\tilde{\nu}}(t) - v_{\rm opt}(t) \leq n\tilde{\nu} = \frac{z}{6} \tag{32}$$

holds due to (29). Adding (31), (32) and using the triangular inequality, we see that

$$|c^*_{\tilde{\nu}} - v_{\rm opt}(t)| \leq \frac{z}{3}, \quad \text{i.e.,} \quad c^*_{\tilde{\nu}} - \frac{1}{3}z \leq v_{\rm opt}(t) \leq c^*_{\tilde{\nu}} + \frac{1}{3}z$$

holds for any $t \in (0, \tilde{t}]$. Together with the fact that $v_1 > v_2$ are the two accumulation points of $\{v_{\rm opt}(t)\}$, the above relation yields

$$c^*_{\tilde{\nu}} - \frac{1}{3}z \leq v_2 < v_1 \leq c^*_{\tilde{\nu}} + \frac{1}{3}z.$$

This implies $z = v_1 - v_2 \leq 2z/3$ and hence $z \leq 0$, which, however, contradicts $z > 0$. Therefore, the accumulation point of $v_{\rm opt}(t)$ is unique and the limit of $v_{\rm opt}(t)$ exists as $t \downarrow 0$.
■

Now we are ready to prove Theorem 2. Let

$$\tilde{v}(\beta) \equiv \lim_{t\downarrow 0} v(t, t\beta) \text{ for } \beta \in [0, \infty), \quad \tilde{v}(\infty) \equiv \lim_{t\downarrow 0} v(0, t).$$

We note that

$$v_a(\theta) = \lim_{t\downarrow 0} v(t\cos\theta, t\sin\theta) = \tilde{v}(\tan\theta).$$

Theorem 2 is a direct consequence of the following theorem.

Theorem 4 *If the problem is asymptotically pd-feasible, then $\tilde{v}(\beta)$ is a monotone decreasing function in β in the interval $[0, +\infty]$ and the following relation holds.*

$$v(\mathbf{D}) = \tilde{v}(\infty) \leq \tilde{v}(\beta) \leq \tilde{v}(0) = v(\mathbf{P}).$$

Proof. We first show that $\tilde{v}$ is a monotone decreasing function in $[0, \infty)$. Suppose that, by contradiction, monotonicity is violated, namely, there exists β_1 and β_2 such that $\beta_1 < \beta_2$ and $\tilde{v}(\beta_1) < \tilde{v}(\beta_2)$. Let $u = \tilde{v}(\beta_2) - \tilde{v}(\beta_1) > 0$. Recall that

$$\tilde{v}(\beta) = \lim_{t\to 0} v(t, t\beta).$$

We show that for sufficiently small t

$$v(t, t\beta_2) - v(t, t\beta_1) \leq u/2$$

holds, which contradicts $\tilde{v}(\beta_2) - \tilde{v}(\beta_1) = \lim_{t\downarrow 0}(v(t, t\beta_2) - v(t, t\beta_1)) = u$. In fact, since $v_P(\varepsilon, \eta) = v(\varepsilon, \eta) - \eta(C \bullet I_p + \varepsilon I_d \bullet I_p)$ is a monotone decreasing function in η (see item 3 of Proposition 2),

$$v(t, t\beta) - t\beta(C \bullet I_p + tI_d \bullet I_p)$$

is a monotone decreasing function in β. Therefore,

$$v(t, t\beta_2) - t\beta_2(C \bullet I_p + tI_d \bullet I_p) \leq v(t, t\beta_1) - t\beta_1(C \bullet I_p + tI_d \bullet I_p)$$

holds. This implies that, for sufficiently small $t > 0$,

$$v(t, t\beta_2) - v(t, t\beta_1) \leq t(\beta_2 - \beta_1)(C \bullet I_p + tI_d \bullet I_p) \leq \frac{u}{2}$$

and hence letting $t \to 0$, we obtain

$$0 < u = \tilde{v}(\beta_2) - \tilde{v}(\beta_1) \leq \frac{u}{2},$$

contradiction.

Now we confirm monotonicity at $\beta = \infty$. Since $\tilde{v}(\infty) = \lim_{t\downarrow 0} v(0, t)$, what we need to show is $\tilde{v}(\infty) \leq \tilde{v}(\beta)$ for any finite β. This is confirmed as follows:

$$\tilde{v}(\infty) = \lim_{t\downarrow 0} v(0, t) \leq \lim_{t\downarrow 0} v(\beta^{-1}t, t) = \lim_{t\downarrow 0} v(t, \beta t) = \tilde{v}(\beta) \ (\beta > 0).$$

The first inequality is due to the item 2 of Proposition 2, and the second equality holds because $\tilde{v}(\gamma) = \tilde{v}(k\gamma)$ for any $k > 0$, i.e., $\tilde{v}$ is a homogeneous function. This completes the proof of the theorem. ■

(Proof of Theorem 2)

Recall that $v_a(\theta) = \lim_{t\downarrow 0} v(t\cos\theta, t\sin\theta)$. We have, for $\theta \in [0, \pi/2]$,

$$v_a(\theta) = \lim_{t\downarrow 0} v(t\cos\theta, t\sin\theta) = \lim_{t\downarrow 0} v(t, t\tan\theta) = \tilde{v}(\tan\theta).$$

Since $v_a(\theta) = \tilde{v}(\tan\theta)$ and tan is a monotone increasing function in θ, Theorem 2 readily follows.

5 Application to Infeasible Interior-point Algorithms

The analysis in the previous section indicates that the limiting common optimal value of $\mathbf{P}(t\alpha, t\beta)$ and $\mathbf{D}(t\alpha, t\beta)$ exists as $t \to 0$ and the value is between $v(\mathbf{D})$ and $v(\mathbf{P})$. In this section, we discuss an application to the convergence analysis of infeasible primal-dual interior-point algorithms.

While the efficiency of infeasible interior-point algorithms is supported by a powerful polynomial-convergence analysis when applied to a primal-dual strongly feasible problems, its behavior for singular problems was not clear. Our analysis leads to a much more clearer picture about what happens when infeasible interior-point algorithms are applied to arbitrary SDP problems. As indicated in Subsection 3.2, we focus on two polynomial-time algorithms by Zhang [41] and Potra and Sheng [26], but the idea and the analysis can be applied to many other variants.

Suppose that $\hat{X}$ is a solution to $A(X) = b$, $(\hat{S}, \hat{y})$ is a solution to $S = C - \sum_i A_i y_i$, and let

$$(X^0, S^0, y^0) \equiv (\hat{X} + \rho\sin\theta I_p, \hat{S} + \rho\cos\theta I_d, 0),$$

where $\theta \in (0, \pi/2)$ and $\rho > 0$ is sufficiently large so that $X^0 \succ 0$ and $S^0 \succ 0$ hold. This is an interior feasible point to the primal-dual pair $\mathbf{P}(\rho\cos\theta, \rho\sin\theta)$ and $\mathbf{D}(\rho\cos\theta, \rho\sin\theta)$, see (2) and (3). In the following, we analyze infeasible primal-dual interior-point algorithms started from this point.

For simplicity of notation, we let $\alpha \equiv \cos\theta$ and $\beta \equiv \sin\theta$. As discussed in Section 3.2.1, in particular as stated in Proposition 1, the infeasible primal-dual interior-point algorithms we are considering generate a sequence (X^k, S^k, y^k) of interior feasible points to the perturbed system

$$C + t^k\alpha I_d - \sum_i A_i y_i^k = S^k, \quad A(X^k - t^k\beta I_p) = b, \quad X^k \succeq 0, \quad S^k \succeq 0. \tag{33}$$

for $t^k \geq 0$. We define

$$(C + t^k\alpha I_d) \bullet X \quad \text{and} \quad \sum_i (b_i + t^k\beta A_i \bullet I_p) y_i^k \tag{34}$$

as the *modified primal objective function* and the *modified dual objective function*, respectively. If (X^k, S^k, y^k, t^k) is a sequence satisfying (33) for every k and $t^k \downarrow 0$, then it is an *asymptotically pd-feasible sequence* in the sense that X^k, S^k satisfy the conic constraints of $\mathbf{P}$ and $\mathbf{D}$ and the distance between (X^k, S^k, y^k) and the set of solutions to the *linear constraints* of $\mathbf{P}$ and $\mathbf{D}$ goes to 0 as $k \to \infty$.[6]

Now we are ready to describe and prove our first result on infeasible interior-point algorithms.

Theorem 5 *Suppose that $\hat{X}$ is a solution to $A(\hat{X}) = b$, $(\hat{S}, \hat{y})$ is a solution to $C - \sum_i A_i y_i = S$, and let $(X^0, S^0, y^0) \equiv (\hat{X} + \rho \sin\theta I_p, \hat{S} + \rho\cos\theta I_d, 0)$, where $\theta \in (0, \pi/2)$ and $\rho > 0$ is sufficiently large so that $X^0 \succ 0$ and $S^0 \succ 0$ hold. Also, let $t^0 \equiv 1$. Apply the algorithm [41] or [26] to solve $\mathbf{P}$ and $\mathbf{D}$, and let $\{(X^k, S^k, y^k, t^k)\}$ be the generated sequence. Then the following statements hold.*

1. *$t^k \to 0$ and $X^k \bullet S^k \to 0$ hold if and only if $\mathbf{P}$ and $\mathbf{D}$ are asymptotically pd-feasible, namely, the algorithms generate an asymptotically pd-feasible sequence with duality gap converging to zero if and only if $\mathbf{P}$ and $\mathbf{D}$ are asymptotically pd-feasible. See the remark after the proof of the theorem for the behavior of the algorithms when $\mathbf{P}$ and $\mathbf{D}$ are not asymptotically pd-feasible.*

2. *If the problem is asymptotically pd-feasible, then the generated sequence of the modified primal and dual objective values (34) converges to the value $v_a(\theta) \in [v(\mathbf{D}), v(\mathbf{P})]$. Here, we include the possibility that $v_a(\theta) = +\infty$ and $v_a(\theta) = -\infty$, interpreting them as divergence to $+\infty$ and $-\infty$, respectively.*

3. *In item 2, if θ is close to 0, this implies that the limiting modified objective values of the infeasible primal-dual algorithm gets closer to the primal optimal value $v(\mathbf{P})$ of the original problem, and the limiting modified objective value approaches the dual optimal value $v(\mathbf{D})$ as θ gets closer to $\pi/2$.*

Proof. First, we discuss item 1. If $\{(X^k, S^k, y^k, t^k)\}$ is an asymptotically pd-feasible sequence, then $\mathbf{P}$ and $\mathbf{D}$ must be asymptotically pd-feasible. Next, we take a look at the converse. In the analysis conducted in [26,41], although both papers assume the existence of a solution to (14), in fact, the existence of a solution is not necessary for showing convergence of t^k and $X^k \bullet S^k$ to zero under asymptotic pd-feasibility. Under asymptotic pd-feasibility, for any $t > 0$ the perturbed problems are strongly feasible. This is enough for showing $t^k \to 0$ and $X^k \bullet S^k \to 0$ in these algorithms. We give more details of the proof in Appendix B.

Now we prove items 2 and 3. The following relations hold at the k-th iteration:

$$(C + t^k \alpha I_d) \bullet X^k - \sum_i (b_i + t^k \beta A_i \bullet I_p) y_i^k = X^k \bullet S^k. \tag{35}$$

$$v(t^k\alpha, t^k\beta) \in \left[\sum_i (b_i + t^k \beta A_i \bullet I_p) y_i^k, (C + t^k \alpha I_d) \bullet X^k \right] \tag{36}$$

(See also (27) and (28) for the derivation of these relations.)

[6] We note, however, that this does **not** imply that, say, the distance between X^k and the feasible region of $\mathbf{P}$ goes to 0 as $k \to \infty$, even if the feasible region of $\mathbf{P}$ is not empty. A similar comment applies to S^k, y^k and the feasible region of $\mathbf{D}$. An instructive example can be seen in [36, Example 1].

Then it follows from (35), (36) and $X^k \bullet S^k \to 0$ that the sets of accumulation points of $\{(C + t^k\alpha I_d) \bullet X^k\}$, $\{v(t^k\alpha, t^k\beta)\}$, and $\{\sum(b_i + t^k\beta A_i \bullet I_p)y_i^k\}$ coincide. Since $t^k \to 0$, this implies that $v(t^k\alpha, t^k\beta) = v(t^k\cos\theta, t^k\sin\theta)$ converges to $v_a(\theta)$. Then the sequences of the modified objective functions (34) also converge to $v_a(\theta)$. ■

Remark When **P** and **D** are not pd-asymptotically feasible, $\lim_{k\to\infty} t^k$ is positive for both algorithms [26, 41]. But the behavior of the duality gap $X^k \bullet S^k$ is a bit different. In the case of Zhang's algorithm, the sequence of $X^k \bullet S^k$ also converges to a positive value as well, but in the case of Potra and Sheng's algorithm, what we can say is that liminf $X^k \bullet S^k$ is positive. This is because the sequence $X^k \bullet S^k$ is not necessarily monotonically decreasing in Potra and Sheng's algorithm.

Now we present the last theorem. A typical choice of the initial iterate (X^0, S^0, y^0) for primal-dual infeasible interior-point algorithms is $(X^0, S^0, y^0) = (\rho_0 I, \rho_1 I, 0)$ with $\rho_0 > 0$ and $\rho_1 > 0$ sufficiently large. This is different from the one adopted in Theorem 5. In concluding this section, we discuss how our results can be adapted to this case.

Let $\hat{X}$ be a solution to $A(X) = b$. If we set $I_p \equiv \rho_0 I - \hat{X}$ and $I_d \equiv \rho_1 I - C$ with ρ_0 and ρ_1 sufficiently large so that $I_p \succ 0$ and $I_d \succ 0$ hold, (X^0, S^0, y^0) is a feasible solution to $\mathbf{P}(1,1)$ and $\mathbf{D}(1,1)$. Now, we are ready to apply an argument analogous to the one we developed earlier to derive Theorem 5 with this choice of I_p and I_d to obtain the following theorem.

Theorem 6 *Let $(X^0, S^0, y^0) \equiv (\rho_0 I, \rho_1 I, 0)$, where $\rho_0 > 0$ and $\rho_1 > 0$ are sufficiently large so that $I_p = \rho_0 I - \hat{X} \succ 0$ and $I_d = \rho_1 I - C \succ 0$ hold, where $\hat{X}$ is a solution to $A(X) = b$. Apply the algorithm [41] or [26] with the initial iterate (X^0, S^0, y^0) and $t^0 = 1$, and let $\{(X^k, S^k, y^k, t^k)\}$ be the generated sequence. Then the following statements hold:*

1. *$t^k \to 0$ and $X^k S^k \to 0$ hold if and only if* **P** *and* **D** *are asymptotically pd-feasible, namely, the algorithm generates an asymptotically pd-feasible sequence with duality gap converging to zero if and only if* **P** *and* **D** *are asymptotically pd-feasible. If* **P** *and* **D** *are not asymptotically pd-feasible, then the same remark after Theorem 5 holds.*

2. *If the problem is asymptotically-pd feasible, then the generated sequence of the modified primal and objective values (34) converges to a value $v_a(\pi/4) \in [v(\mathbf{D}), v(\mathbf{P})]$. Here, we include the possibility that $v_a(\theta) = +\infty$ and $v_a(\theta) = -\infty$, interpreting them as divergence to $+\infty$ and $-\infty$, respectively.*

6 Examples

In this section, we present three examples with nonzero duality gaps to illustrate Theorems 1 and 2. The optimal values of **P** and **D** are both finite in Example 1, the optimal value of **P** is finite but **D** is weakly infeasible in Example 2, and both problems are weakly infeasible in Example 3. In the latter two cases the duality gaps are infinity.

Example 1

We start with a simple instance with a finite nonzero duality gap taken from Ramana's famous paper [28]. The following problem has a duality gap of one.

The problem $\mathbf{D}$ is

$$\max\ y_1 \text{ s.t. } \begin{pmatrix} 1-y_1 & 0 & 0 \\ 0 & -y_2 & -y_1 \\ 0 & -y_1 & 0 \end{pmatrix} \succeq 0.$$

With that, we have

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad b_1 = 1.$$

The optimal value $v(\mathbf{D}) = 0$ for this problem, since $y_1 = 0$ is the only possible value for the lower-right 2×2 submatrix to be positive semidefinite.

The associated primal $\mathbf{P}$ is

$$\min\ x_{11} \text{ s.t. } x_{11} + 2x_{23} = 1,\ x_{22} = 0,\ \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{12} & x_{22} & x_{23} \\ x_{13} & x_{23} & x_{33} \end{pmatrix} \succeq 0.$$

The optimal value $v(\mathbf{P}) = 1$ for this problem, since $x_{23} = 0$ must hold for positive semidefiniteness of the lower-right 2×2 submatrix, which drives x_{11} to be 1.

Now we consider the problem $\mathbf{D}(\varepsilon, \eta)$

$$\max\ (1+\eta)y_1 + \eta y_2 \text{ s.t. } \begin{pmatrix} 1+\varepsilon-y_1 & 0 & 0 \\ 0 & \varepsilon - y_2 & -y_1 \\ 0 & -y_1 & \varepsilon \end{pmatrix} \succeq 0.$$

This is equivalent to

$$\max\ (1+\eta)y_1 + \eta y_2 \text{ s.t. } 1+\varepsilon-y_1 \geq 0, \quad \varepsilon(\varepsilon - y_2) - y_1^2 \geq 0.$$

Since the objective is linear, there is an optimal solution such that at least one of the inequality constraints is active. Taking into account that the second constraint is quadratic, we analyze the following three subproblems, and take the maximum of them.

$$\begin{aligned}
&\text{(Case 1)}\ \max\ (1+\eta)y_1 + \eta y_2 \text{ s.t. } 1+\varepsilon-y_1 = 0, \quad \varepsilon(\varepsilon - y_2) - y_1^2 \geq 0. \\
&\text{(Case 2)}\ \max\ (1+\eta)y_1 + \eta y_2 \text{ s.t. } 1+\varepsilon-y_1 \geq 0, \quad y_1 = \sqrt{\varepsilon(\varepsilon - y_2)}. \\
&\text{(Case 3)}\ \max\ (1+\eta)y_1 + \eta y_2 \text{ s.t. } 1+\varepsilon-y_1 \geq 0, \quad y_1 = -\sqrt{\varepsilon(\varepsilon - y_2)}.
\end{aligned}$$

(Case 1)

In this case, the second constraint yields

$$\varepsilon - \frac{(1+\varepsilon)^2}{\varepsilon} \geq y_2.$$

Together with $y_1 = 1+\varepsilon$, the problem reduces to a linear program, and it follows that the maximum is

$$v_1(\varepsilon, \eta) \equiv (1+\eta)(1+\varepsilon) + \eta\varepsilon - \frac{\eta(1+\varepsilon)^2}{\varepsilon}.$$

(Case 2)

Under this condition, the objective function is written as

$$f(y_2) \equiv (1+\eta)\sqrt{\varepsilon(\varepsilon - y_2)} + \eta y_2.$$

By computing the derivative, we see that the function takes the unique maximum at

$$y_2 = \varepsilon - \frac{\varepsilon(1+\eta)^2}{4\eta^2} \tag{37}$$

and

$$\sqrt{\varepsilon(\varepsilon - y_2)} = \frac{\varepsilon(1+\eta)}{2\eta}. \tag{38}$$

Then, we see that

$$f(y_2) = \varepsilon\eta + \frac{\varepsilon}{4\eta}(1+\eta)^2. \tag{39}$$

But we should recall that this maximum is obtained by ignoring the constraint

$$1 + \varepsilon - y_1 = 1 + \varepsilon - \sqrt{\varepsilon(\varepsilon - y_2)} \geq 0.$$

By substituting (37) and (38) into this constraint, (39) is the maximum only if

$$1 + \varepsilon - \frac{\varepsilon(1+\eta)}{2\eta} \geq 0, \text{or, equivalently, } \frac{2\eta}{1-\eta} \geq \varepsilon \tag{40}$$

is satisfied.

If (40) does not hold, then, the maximum of $f(y_2)$ is taken at the boundary of the constraint $1+\varepsilon - y_1 \geq 0$, i.e., y_2 satisfying the condition

$$1 + \varepsilon = \sqrt{\varepsilon(\varepsilon - y_2)}.$$

Solving this equation with respect to y_2, we obtain

$$y_2 = -2 - \frac{1}{\varepsilon}, \ y_1 = 1 + \varepsilon, \quad f(y_2) = (1+\eta)(1+\varepsilon) - \eta\left(2 + \frac{1}{\varepsilon}\right).$$

In summary, the maximum value in (Case 2) is as follows:

$$v_2(\varepsilon, \eta) \equiv \varepsilon\eta + \frac{\varepsilon}{4\eta}(1+\eta)^2 \ \text{ if } \frac{2\eta}{1-\eta} \geq \varepsilon, \tag{41}$$

$$v_2(\varepsilon, \eta) \equiv (1+\eta)(1+\varepsilon) - \eta\left(2 + \frac{1}{\varepsilon}\right) \ \text{ if } \frac{2\eta}{1-\eta} \leq \varepsilon \tag{42}$$

(Case 3)

In this case, $1+\varepsilon-y_1 \geq 0$ holds trivially. Therefore, the maximization problem in this case is

$$\max -(1+\eta)\sqrt{\varepsilon(\varepsilon-y_2)}+\eta y_2.$$

under the condition that $y_2 \leq \varepsilon$. The function is monotone increasing, so that the maximum is attained when $y_2=\varepsilon$ and the maximum value is

$$v_3(\varepsilon,\eta) \equiv \eta\varepsilon.$$

Now we are ready to combine the three results to complete the evaluation of $\tilde{v}$ and v_a. By letting $\varepsilon = t\alpha$, $\eta = t\beta$ with $t>0$ and letting $t \downarrow 0$, we see that

(Case 1) $\lim_{t\downarrow 0} v_1(t\alpha, t\beta) = 0$.

(Case 2) $\lim_{t\downarrow 0} v_2(t\alpha, t\beta) = \frac{\alpha}{4\beta}$ if $\frac{\beta}{\alpha} \geq \frac{1}{2}$, $\lim_{t\downarrow 0} v_2(t\alpha, t\beta) = 1-\frac{\beta}{\alpha}$ if $\frac{\beta}{\alpha} \leq \frac{1}{2}$

(Case 3) $\lim_{t\downarrow 0} v_3(t\alpha, t\beta) = 0$.

The maximum among the three corresponds to $\tilde{v}$. Comparing the three, we see that (Case 2) always is the maximum. This means

$$\tilde{v}(\beta) = 1-\beta \ (\beta \in [0, \frac{1}{2}]), \quad \tilde{v}(\beta) = \frac{1}{4\beta} \ (\beta \in [\frac{1}{2}, \infty)), \quad \tilde{v}(\infty) = 0.$$

Example 2

The next example is such that $\mathbf{D}$ is weakly infeasible but $\mathbf{P}$ is weakly feasible and has a finite optimal value.

The problem $\mathbf{D}$ is

$$\max \ -y_1 \ \text{ s.t. } \begin{pmatrix} y_2 & 0 & 1 \\ 0 & y_1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \succeq 0.$$

$$C = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad b_1 = -1.$$

This system is weakly infeasible, so $v(\mathbf{D}) = -\infty$.

The associated primal $\mathbf{P}$ is

$$\min \ 2x_{13} \ \text{ s.t. } x_{11} = 0, \ x_{22} = 1, \ \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{12} & x_{22} & x_{23} \\ x_{13} & x_{23} & x_{33} \end{pmatrix} \succeq 0.$$

The optimal value $v(\mathbf{P}) = 0$ for this problem, since $x_{13} = 0$ must hold for feasibility.

Now we consider the problem $\mathbf{D}(\varepsilon, \eta)$

$$\max \ -(1+\eta)y_1 - \eta y_2 \ \text{ s.t. } \begin{pmatrix} y_2+\varepsilon & 0 & 1 \\ 0 & y_1+\varepsilon & 0 \\ 1 & 0 & \varepsilon \end{pmatrix} \succeq 0.$$

It follows that

$$y_1 \geq -\varepsilon, \quad y_2 \geq \frac{1-\varepsilon^2}{\varepsilon}.$$

Therefore, we see that the maximum value is

$$v(\varepsilon,\eta) = (1+\eta)\varepsilon - \frac{1-\varepsilon^2}{\varepsilon}\eta.$$

Now we are ready to evaluate $\tilde{v}$ and v_a. By letting $\varepsilon = t\alpha$, $\eta = t\beta$ with $t > 0$ and letting $t \downarrow 0$, we see that

$$\lim_{t\downarrow 0} v(t\alpha, t\beta) = -\frac{\beta}{\alpha}.$$

and

$$\tilde{v}(\beta) = -\beta \ (\beta \in [0,\infty]).$$

Finally, we deal with a pathological case where both primal and dual are weakly infeasible.

Example 3

The problem $\mathbf{D}$ is

$$\max\ y_1 \ \text{ s.t. } \begin{pmatrix} y_2 & 0 & 1+\frac{1}{2}y_1 \\ 0 & 1+y_1 & 0 \\ 1+\frac{1}{2}y_1 & 0 & 0 \end{pmatrix} \succeq 0.$$

$$C = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & 0 & -\frac{1}{2} \\ 0 & -1 & 0 \\ -\frac{1}{2} & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad b_1 = 1.$$

The optimal value $v(\mathbf{D}) = -\infty$ for this problem, since $y_1 = -2$ should hold for feasibility, but then the (2,2) element becomes -1 and, therefore, the matrix cannot be feasible. By letting y_2 large and $y_1 = 0$, we confirm the problem is weakly infeasible.

The associated primal $\mathbf{P}$ is

$$\min\ 2x_{13} + x_{22} \ \text{ s.t. } x_{13} + x_{22} = -1,\ x_{11} = 0,\ \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{12} & x_{22} & x_{23} \\ x_{13} & x_{23} & x_{33} \end{pmatrix} \succeq 0.$$

This problem is weakly infeasible.

Now we consider the problem $\mathbf{D}(\varepsilon,\eta)$

$$\max\ (1-\eta)y_1 - \eta y_2 \ \text{ s.t. } \begin{pmatrix} \varepsilon + y_1 & 0 & 1+\frac{1}{2}y_1 \\ 0 & 1+\varepsilon+y_1 & 0 \\ 1+\frac{1}{2}y_1 & 0 & \varepsilon \end{pmatrix} \succeq 0.$$

This is equivalent to

$$\max\ (1-\eta)y_1 - \eta y_2 \ \text{ s.t. } \varepsilon + y_2 \geq 0,\ \ \varepsilon(\varepsilon+y_2) - (1+\frac{1}{2}y_1)^2 \geq 0, \quad 1+\varepsilon+y_1 \geq 0.$$

Since the objective is linear, there is an optimal solution such that at least one of the inequality constraints is active. Taking into account that the second constraint is quadratic, we analyze the following three subproblems and take the maximum of them.

$$
\begin{aligned}
&\text{(Case 1)} \quad && \max\ (1-\eta)y_1 - \eta y_2 \ \text{ s.t. } \varepsilon + y_2 = 0,\ \varepsilon(\varepsilon+y_2) - (1+\frac{1}{2}y_1)^2 \geq 0, \\
& && 1+\varepsilon+y_1 \geq 0. \\
&\text{(Case 2)} \quad && \max\ (1-\eta)y_1 - \eta y_2 \ \text{ s.t. } \varepsilon + y_2 \geq 0,\ \varepsilon(\varepsilon+y_2) - (1+\frac{1}{2}y_1)^2 = 0, \\
& && 1+\varepsilon+y_1 \geq 0. \\
&\text{(Case 3)} \quad && \max\ (1-\eta)y_1 - \eta y_2 \ \text{ s.t. } \varepsilon + y_2 \geq 0,\ \varepsilon(\varepsilon+y_2) - (1+\frac{1}{2}y_1)^2 \geq 0, \\
& && 1+\varepsilon+y_1 = 0.
\end{aligned}
$$

(Case 1)

In this case, we have $y_2 = -\varepsilon$, $y_1 = -2$. Then the third constraint becomes $\varepsilon - 1 \geq 0$. Since we are interested in the situation where ε is approaching zero, we may exclude this case.

(Case 2)

In this case, we have

$$\varepsilon(\varepsilon+y_2) = (1+\frac{1}{2}y_1)^2.$$

This implies that

$$y_1 = 2(-1 \pm \sqrt{\varepsilon(\varepsilon+y_2)}).$$

Since the condition $1+\varepsilon+y_1 \geq 0$ yields

$$\pm\sqrt{\varepsilon(\varepsilon+y_2)} \geq 1-\varepsilon,$$

choosing '-' sign is not compatible with our analysis since we are interested in the case where ε is close to zero. Therefore, we pick '+' sign, and seek for the maximum of the objective function

$$2(1-\eta)(-1+\sqrt{\varepsilon(\varepsilon+y_2)} - \eta y_2.$$

By differentiation, we see that the function attains its maximum at

$$y_1 = 2\left(-1+\frac{\varepsilon(1-\eta)}{\eta}\right), \quad y_2 = \frac{\varepsilon}{\eta^2}(1-2\eta).$$

We see that the first constraint is always satisfied at the maximum. The third constraint $1+y_1+\varepsilon \geq 0$ is satisfied if

$$\frac{\varepsilon}{\eta} \geq \frac{1+\varepsilon}{2}.$$

If this condition is not satisfied, then $1+y_1+\varepsilon = 0$ holds at the maximum, so, we can leave the analysis to the third case. Substituting y_1, y_2 to the objective, we conclude that, if $\varepsilon/\eta \geq 1$, then, the maximum is

$$v_2(\varepsilon,\eta) \equiv 2(1-\eta)\left(-1-\varepsilon+\frac{\varepsilon}{\eta}\right) - \frac{\varepsilon}{\eta} + 2\varepsilon,$$

and if the aforementioned condition is not satisfied, then, we can leave the analysis to the third case below.

(Case 3)

We have $y_1 = -1 - \varepsilon$. After simple manipulation, we see that other two inequalities are satisfied iff

$$y_2 \geq \frac{1}{\varepsilon}\left(\frac{1-\varepsilon}{2}\right)^2 - \varepsilon.$$

Therefore, the maximum is

$$v_3(\varepsilon, \eta) \equiv -(1-\eta)(1+\varepsilon) - \frac{\eta}{\varepsilon}\left(\frac{1-\varepsilon}{2}\right)^2 + \varepsilon\eta.$$

Now we are ready to combine the three results to complete evaluation of $\tilde{v}$ and v_a. By letting $\varepsilon = t\alpha$, $\eta = t\beta$ with $t > 0$ and letting $t \downarrow 0$, we see that

(Case 1) Cannot occur.

(Case 2) $\lim_{t\downarrow 0} v_2(t\alpha, t\beta) = -2 + \frac{\alpha}{\beta}$ if $\frac{\alpha}{\beta} \geq \frac{1}{2}$.

(Case 3) $\lim_{t\downarrow 0} v_3(t\alpha, t\beta) = -1 - \frac{1}{4}\frac{\beta}{\alpha}$.

The maximum between the latter two corresponds to $\tilde{v}$. Thus, we obtain that

$$\tilde{v}(\beta) = -2 + \frac{1}{\beta}\ (\beta \in [0, 2]), \quad \tilde{v}(\beta) = -1 - \frac{\beta}{4}\ (\beta \in [2, \infty]),$$

where we used the convention $1/0 = \infty$.

7 Concluding Discussion

In this paper, we developed a perturbation analysis for singular primal-dual semidefinite programs. We assumed that primal and dual problems are asymptotically feasible and added positive definite perturbations to recover strong feasibility. A major innovation was that we considered perturbations of primal and dual problems simultaneously. It was shown that the primal-dual common optimal value of the perturbed problem has a directional limit when the perturbation is reduced to zero along a line. Representing the direction of approach with an angle θ between 0 and $\pi/2$, where the former and latter corresponds to the dual-only perturbation and the primal-only perturbation, respectively, we demonstrated that the limiting objective value is a monotone decreasing function in θ which takes the primal optimal value $v(\mathbf{P})$ at $\theta = 0$ and the dual optimal value $v(\mathbf{D})$ at $\theta = \pi/2$. Based on this result, we could show that the modified objective values of the two infeasible primal-dual interior-point algorithms by Zhang and by Potra and Sheng converge to a value between the optimal values of **P** and **D**. The modified primal and dual objective functions are easily computed from the current iterate. Although we analyzed infeasible interior-point algorithms here, it would not be difficult to extend this result to interior-point algorithms based on homogeneous self-dual embeddings. An interesting further research topic would be to show smoothness of the limiting objective value function $v_a(\theta)$. It would be also interesting to design a robust primal-dual interior-point algorithm based on the theory developed in this paper.

A Outline of a Proof of the Existence and Analyticity of the Path $\mathcal{C} = \{w_\nu(t)|\ 0 < t < \infty\}$

Let

$$\phi_1(X,S,y) = X^{1/2}SX^{1/2} - \nu I, \quad \phi_2(X,S,y) = C - \sum_i A_i y_i - S,$$

$$\phi_3(X,S,y) = \begin{pmatrix} A_1 \bullet X - b_1 \\ \vdots \\ A_m \bullet X - b_m \end{pmatrix}.$$

Then, $w_\nu(t)$ is a unique solution to

$$\Phi(X,S,y,t) \equiv \begin{pmatrix} \phi_1(X,S,y) \\ \phi_2(X,S,y) + t\alpha I \\ \phi_3(X,S,y) - t\beta I \end{pmatrix} = 0.$$

Φ is an analytic mapping from $\{(X,S,y,t) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} \times \mathbb{R}^{(m+1)}\}$ to $\mathcal{S}^n_{++} \times \mathcal{S}^n \times \mathbb{R}^m$, where $\mathcal{S}_{++}$ is the set of symmetric positive definite matrices. Therefore, in order to show the existence and analyticity of the path with the help of the analytic version of the implicit function theorem, it is enough to confirm that the rank of the Jacobian matrix of Φ is $n(n+1)+m$. To this end, we show that the Jacobian matrix of the mapping

$$\begin{pmatrix} \phi_1(X,S,y) \\ \phi_2(X,S,y) \\ \phi_3(X,S,y) \end{pmatrix}$$

is nonsingular. Indeed it is essentially shown in Theorem 2.4 of [22] that the Jacobian matrix is nonsingular if $\phi_1 = 0$, i.e., $XS = \nu I$. (See also the note following the theorem.)

B Outline of a Proof of Item 1 of Theorems 5 and 6

The item claims that the sequence $\{t^k\}$ generated by the algorithm converges to 0 whenever **P** and **D** are asymptotically pd-feasible. We provide an explanation for Zhang's algorithm. A similar argument also holds for Potra and Sheng's algorithm.

If the problems are asymptotically pd-feasible, then for any $t > 0$, $\mathbf{P}(t\alpha, t\beta)$ and $\mathbf{D}(t\alpha, t\beta)$ are strongly feasible. For the sake of obtaining a contradiction, suppose that t^k has a positive limit t^*. Then, we can show that for sufficiently large k, the iterate (X^k, S^k, y^k) is confined in a compact region Ω contained in the neighbourhood of the central path employed by the algorithm. In Ω, the search direction is a continuous function and hence the norm of the search direction is bounded over Ω. This enables us to show that the step s^k in (19) is bounded away from zero as long as $(X^k, S^k, y^k) \in \Omega$. Then there exists $\zeta > 0$ such that $s^k > \zeta$ for all k sufficiently large. This contradicts that $t^k \to t^*$, because, in view of (22), we have $t^* \leq t^{k+1} = (1 - s^k)t^k < (1-\zeta)t^k$, but this cannot hold for k sufficiently large thus leading to a contradiction.

Next, we show that $t^k \to 0$ yields $X^k \bullet S^k \to 0$. The stepsize s^k is controlled in such a way that

$$\frac{X^{k+1} \bullet S^{k+1}}{X^k \bullet S^k} \leq 1 - \eta s^k,$$

holds in the algorithm, where $\eta \in (0,1)$ is a constant. There are two possible cases. The first case is $s^k = 1$ for some $k = \hat{k}$, say. In that case, after the $\hat{k}$th iteration, the algorithm becomes a feasible path following method, and $X^k \bullet S^k$ converges to zero following a standard argument. In the second case, $s^k < 1$ for all k. Since $\lim_{k\to\infty} t^k = \lim_{k\to\infty} \prod_{l=0}^{k}(1 - s^l) = 0$ yields $\lim_{k\to\infty} \prod_{l=0}^{k}(1 - \eta s^l) = 0$, we obtain

$$\lim_{k\to\infty} X^k \bullet S^k \leq \lim_{k\to\infty} \prod_{l=0}^{k}(1 - \eta s^l) X^0 \bullet S^0 = 0.$$

References

[1] Ben-Israel, A., Charnes, A., Kortanek, K.O.: Duality and asymptotic solvability over cones. Bulletin of the American Mathematical Society **75**(2), 318 – 324 (1969)

[2] Bochnak, J., Coste, M., Roy, M.F.: Real Algebraic Geometry. Springer Science (1998)

[3] Bonnans, J.F., Shapiro, A.: Perturbation Analysis of Optimization Problems. Springer-Verlag, New York (2000)

[4] Borwein, J.M., Wolkowicz, H.: Facial reduction for a cone-convex programming problem. Journal of the Australian Mathematical Society (Series A) **30**(03), 369–380 (1981)

[5] Borwein, J.M., Wolkowicz, H.: Regularizing the abstract convex program. Journal of Mathematical Analysis and Applications **83**(2), 495 – 530 (1981)

[6] Cheung, Y.L., Schurr, S., Wolkowicz, H.: Preprocessing and regularization for degenerate semidefinite programs. In: Computational and Analytical Mathematics, *Springer Proceedings in Mathematics & Statistics*, vol. 50, pp. 251–303. Springer New York (2013)

[7] Drusvyatskiy, D., Wolkowicz, H.: The many faces of degeneracy in conic optimization. Tech. rep., University of Washington (2017)

[8] Duffin, R.J.: Infinite programs. In: H.W. Kuhn, A.W. Tucker (eds.) Linear Inequalities and Related Systems. Annals of Mathematics Studies, Volume 38, chap. 6. Princeton University Press (2000)

[9] Friberg, H.A.: A relaxed-certificate facial reduction algorithm based on subspace intersection. Operations Research Letters **44**(6), 718 – 722 (2016)

[10] Gally, T., Pfetsch, M.E., Ulbrich, S.: A framework for solving mixed-integer semidefinite programs. Optimization Methods and Software **33**(3), 594–632 (2018)

[11] Güler, O.: Foundations of Optimization. Graduate Texts in Mathematics. Springer New York (2010)

[12] Helmberg, C., Rendl, F., Vanderbei, R.J., Wolkowicz, H.: An interior-point method for semidefinite programming. SIAM Journal on Optimization **6**(2), 342–361 (1996)

[13] Kojima, M., Shindoh, S., Hara, S.: Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices. SIAM Journal on Optimization **7**(1), 86–125 (1997)

[14] Krantz, S., Parks, H.: A Primer of Real Analytic Functions. Advanced Texts Series. Birkhäuser Boston (2002)

[15] Lourenço, B.F.: Amenable cones: error bounds without constraint qualifications. Mathematical Programming **186**, 1–48 (2021)

[16] Lourenço, B.F., Muramatsu, M., Tsuchiya, T.: Facial reduction and partial polyhedrality. SIAM Journal on Optimization **28**(3), 2304–2326 (2018)

[17] Lu, Z., Monteiro, R.D.C.: Error bounds and limiting behavior of weighted paths associated with the SDP map $X^{1/2}SX^{1/2}$. SIAM Journal on Optimization **15**(2), 348–374 (2004)

[18] Luo, Z.Q., Sturm, J.F.: Error analysis. In: H. Wolkowicz, R. Saigal, L. Vandenberghe (eds.) Handbook of semidefinite programming: theory, algorithms, and applications. Kluwer Academic Publishers (2000)

[19] Luo, Z.Q., Sturm, J.F., Zhang, S.: Duality and self-duality for conic convex programming. Tech. rep., Econometric Institute, Erasmus University Rotterdam, The Netherlands (1996)

[20] Luo, Z.Q., Sturm, J.F., Zhang, S.: Duality results for conic convex programming. Tech. rep., Econometric Institute, Erasmus University Rotterdam, The Netherlands (1997)

[21] Monteiro, R.D.C.: Primal–Dual path-following algorithms for semidefinite programming. SIAM Journal on Optimization **7**(3), 663–678 (1997)

[22] Monteiro, R.D.C., Tsuchiya, T.: Polynomial convergence of a new family of primal-dual algorithms for semidefinite programming **9**(3), 551–577 (1999)

[23] Nesterov, Y., Todd, M.: Self-scaled barriers and interior-point methods for convex programming. Mathematics of Operations research **22**, 1–42 (1997)

[24] Permenter, F., Friberg, H.A., Andersen, E.D.: Solving conic optimization problems via self-dual embedding and facial reduction: A unified approach. SIAM Journal on Optimization **27**(3), 1257–1282 (2017)

[25] Permenter, F., Parrilo, P.: Partial facial reduction: simplified, equivalent SDPs via approximations of the PSD cone. Mathematical Programming (2017)

[26] Potra, F.A., Sheng, R.: A superlinearly convergent primal–dual infeasible–interior–point algorithm for semidefinite programming. SIAM Journal on Optimization **8**, 1007–1028 (1998)

[27] Preiß, M., Stoer, J.: Analysis of infeasible-interior-point paths arising with semidefinite linear complementarity problems. Mathematical Programming **99**, 499–520 (2004)

[28] Ramana, M.V.: An exact duality theory for semidefinite programming and its complexity implications. Mathematical Programming **77** (1995)

[29] Ramana, M.V., Tunçel, L., Wolkowicz, H.: Strong duality for semidefinite programming. SIAM Journal on Optimization **7**(3), 641–662 (1997)

[30] Renegar, J.: A mathematical view of interior-point methods in convex optimization. SIAM, Philadelphia, PA, United States (2001)

[31] Rockafellar, R.T.: Convex Analysis . Princeton University Press (1997)

[32] Sekiguchi, Y., Waki, H.: Perturbation analysis of singular semidefinite program and its application to a control problem. Tech. Rep. 1 (2021)

[33] Sremac, S., Woerdeman, H., Wolkowicz, H.: Complete facial reduction in one step for spectrahedra (2017)

[34] Sremac, S., Woerdeman, H.J., Wolkowicz, H.: Error bounds and singularity degree in semidefinite programming. SIAM Journal on Optimization **31**(1), 812–836 (2021)

[35] Sturm, J.F.: Theory and algorithms of semidefinite programming. In: H. Frenk, K. Roos, T. Terlaky, S. Zhang (eds.) High Performance Optimization, pp. 1–194. Kluwer Academic Publishers (1999)

[36] Sturm, J.F.: Error bounds for linear matrix inequalities. SIAM Journal on Optimization **10**(4), 1228–1248 (2000). DOI 10.1137/S1052623498338606

[37] Todd, M.J., Toh, K.C., Tütüncü, R.H.: On the Nesterov–Todd direction in semidefinite programming. SIAM Journal on Optimization **8**(3), 769–796 (1998)

[38] Toh, K.C., Todd, M.J., Tütüncü, R.H.: SDPT3 — a matlab software package for semidefinite programming, version 1.3. Optimization Methods and Software **11**, 545–581 (1999)

[39] Waki, H., Muramatsu, M.: Facial reduction algorithms for conic optimization problems. Journal of Optimization Theory and Applications **158**(1), 188–215 (2013)

[40] Yamashita, M., Fujisawa, K., Kojima, M.: Implementation and evaluation of SDPA 6.0 (semidefinite programming algorithm 6.0). Optimization Methods and Software **18**, 491–505 (2003)

[41] Zhang, Y.: On extending some Primal–Dual interior-point algorithms from linear programming to semidefinite programming. SIAM Journal on Optimization **8**(2), 365–386 (1998)

[42] Zhu, Y.M., Pataki, G., Tran-Dinh, Q.: Sieve-SDP: a simple facial reduction algorithm to preprocess semidefinite programs. Mathematical Programming Computation **11**(3), 503–586 (2019)