Infeasibility detection in the alternating direction method of multipliers for convex optimization

Goran Banjac, Paul Goulart, Bartolomeo Stellato, and Stephen Boyd

June 20, 2018

Abstract

The alternating direction method of multipliers (ADMM) is a powerful operator splitting technique for solving structured optimization problems. For convex optimization problems, it is well-known that the iterates generated by ADMM converge to a solution provided that it exists. If a solution does not exist, then the ADMM iterates diverge. Nevertheless, we show that the ADMM iterates yield conclusive information regarding problem infeasibility for optimization problems with linear or quadratic objective functions and conic constraints, which includes quadratic, second-order cone and semidefinite programs. In particular, we show that in the limit the ADMM iterates either satisfy a set of first-order optimality conditions or produce a certificate of either primal or dual infeasibility. Based on these results, we propose termination criteria for detecting primal and dual infeasibility in ADMM.

1 Introduction

Operator splitting methods can be used to solve structured optimization problems of the form

$$\text{minimize} \quad f(w) + g(w), \tag{1}$$

where f and g are convex, closed and proper functions. These methods encompass algorithms such as the proximal gradient method (PGM), Douglas-Rachford splitting (DRS) and the alternating direction method of multipliers (ADMM) [PB13], and have been applied to problems ranging from feasibility and best approximation problems [BB96, BCL04] to quadratic and conic programs [Bol13, OCPB16, ZFP$^+$17]. Due to their relatively low per-iteration computational cost and ability to exploit sparsity in the problem data [ZFP$^+$17], splitting methods are suitable for embedded [OSB13, JGR$^+$14, BSM$^+$17] and large-scale optimization [BT09], and have increasingly been applied for solving problems arising in signal processing [CW05, CP11], machine learning [BPC$^+$11] and optimal control [SSS$^+$16].

In order to solve problem (1), PGM requires differentiability of one of the two functions. If a fixed step-size is used in the algorithm, then one also requires a bound on the Lipschitz constant of the function's gradient [BT09]. On the other hand, ADMM and DRS, which turn out to be equivalent to each other, do not require any additional assumptions on the problem beyond convexity, making them more robust to the problem data.

The growing popularity of ADMM has triggered a strong interest in understanding its theoretical properties. Provided that problem (1) is solvable and satisfies certain constraint qualification (see [BC11, Cor. 26.3] for more details), both ADMM and DRS are known to converge to an optimal solution [BPC+11, BC11]. The use of ADMM for solving convex quadratic programs (QPs) was analyzed in [Bol13] and was shown to admit an asymptotic linear convergence rate. The authors in [GTSJ15] analyzed global linear convergence of ADMM for solving strongly convex QPs with inequality constraints that are linearly independent, and the authors in [GB17] extended these results to a wider class of optimization problems involving a strongly convex objective function. A particularly convenient framework for analyzing asymptotic behavior of such method is by representing it as a fixed-point iteration of an averaged operator [BC11, GB17, BG18].

The ability to detect infeasibility of an optimization problem is very important in many applications, *e.g.* in any embedded application or in mixed-integer optimization when branch-and-bound techniques are used [NB17]. It is well-known that for infeasible convex optimization problems some of the iterates of ADMM and DRS diverge [EB92]. However, terminating the algorithm when the iterates become large is unreliable in practice for several reasons. First, an upper bound on the norm of iterates should be big enough in order to reduce the number of false detections of infeasibility. Second, divergence of the algorithm's iterates is observed to be very slow in practice. Finally, such termination criterion is just an indication that a problem might be infeasible, and not a certificate of infeasibility.

Aside from [EB92], the asymptotic behavior of ADMM and DRS for infeasible problems has been studied only in some special cases. DRS for solving feasibility problems involving two convex sets that do not necessarily intersect was studied in [BCL04, BDM16, BM16, BM17, Mou16]. The authors in [RDC14] study the asymptotic behavior of ADMM for solving convex QPs when the problem is infeasible, but impose full rank assumptions on certain matrices derived from the problem data. The authors in [OCPB16] apply ADMM to the homogeneous self-dual embedding of a convex conic program, thereby producing a larger problem which is always feasible and whose solutions can be used either to produce a primal-dual solution or a certificate of infeasibility for the original problem. A disadvantage of this approach in application to optimization problems with quadratic objective functions is that the problem needs to be transformed into an equivalent conic program which is harder to solve than the original problem in general [Toh08, HM12].

In this paper we consider a class of convex optimization problems that includes linear programs (LPs), QPs, second-order cone programs (SOCPs) and semidefinite programs (SDPs) as special cases. We use a particular version of ADMM introduced in [SBG+18] that imposes no conditions on the problem data such as strong convexity of the objective function or full rank of the constraint matrix. We show that the method either generates iterates for which the violation of the optimality conditions goes to zero, or produces a certificate of primal or dual infeasibility. These results are directly applicable to infeasibility detection in ADMM for the considered class of problems.

1.1 Contents

We introduce the problem of interest in Section 2 and present a particular ADMM algorithm for solving it in Section 3. Section 4 analyzes the asymptotic behavior of ADMM and shows that the algorithm can detect primal and dual infeasibility of the problem. Section 5 demonstrates these results on several small numerical examples. Finally, Section 6 concludes the paper.

1.2 Notation

We introduce some definitions and notation that will be used in the rest of the paper. All definitions here are standard, and can be found *e.g.* in [RW98, BC11].

Let $\mathbf{N}$ denote the set of natural numbers, $\mathbf{R}$ the set of real numbers, $\mathbf{R}_+$ the set of nonnegative real numbers, $\tilde{\mathbf{R}} := \mathbf{R} \cup \{+\infty\}$ the extended real line and $\mathbf{R}^n$ the n-dimensional real space equipped with inner product $\langle \cdot, \cdot \rangle$, induced norm $\|\cdot\|$ and identity operator $\mathrm{Id} : x \mapsto x$. We denote by $\mathbf{R}^{m \times n}$ the set of real m-by-n matrices and by $\mathbf{S}^n$ ($\mathbf{S}^n_+$) the set of real n-by-n symmetric (positive semidefinite) matrices. Let $\mathrm{vec} : \mathbf{S}^n \mapsto \mathbf{R}^{n^2}$ be the operator mapping a matrix to the stack of its columns, $\mathrm{mat} = \mathrm{vec}^{-1}$ its inverse operator, and $\mathrm{diag} : \mathbf{R}^n \mapsto \mathbf{S}^n$ the operator mapping a vector to a diagonal matrix. For a sequence $\{x^k\}_{k \in \mathbf{N}}$ we define $\delta x^{k+1} := x^{k+1} - x^k$. The *proximal operator* of a convex, closed and proper function $f : \mathbf{R}^n \to \tilde{\mathbf{R}}$ is given by

$$\mathrm{prox}_f(x) := \underset{y}{\mathrm{argmin}}\{f(y) + \tfrac{1}{2}\|y - x\|^2\}.$$

For a nonempty, closed and convex set $\mathcal{C} \subseteq \mathbf{R}^n$, we denote the *indicator function* of $\mathcal{C}$ by

$$\mathcal{I}_{\mathcal{C}}(x) := \begin{cases} 0 & x \in \mathcal{C} \\ +\infty & \text{otherwise,} \end{cases}$$

the *distance* of $x \in \mathbf{R}^n$ to $\mathcal{C}$ by

$$\mathrm{dist}_{\mathcal{C}}(x) := \min_{y \in \mathcal{C}} \|x - y\|,$$

the *projection* of $x \in \mathbf{R}^n$ onto $\mathcal{C}$ by

$$\Pi_{\mathcal{C}}(x) := \underset{y \in \mathcal{C}}{\mathrm{argmin}} \|x - y\|,$$

the *support function* of $\mathcal{C}$ by

$$S_{\mathcal{C}}(x) := \sup_{y\in\mathcal{C}} \langle x, y\rangle,$$

the *recession cone* of $\mathcal{C}$ by

$$\mathcal{C}^{\infty} := \{y \in \mathbf{R}^n \,|\, x + \tau y \in \mathcal{C},\, x \in \mathcal{C},\, \tau \geq 0\},$$

and the *normal cone* of $\mathcal{C}$ at $x \in \mathcal{C}$ by

$$N_{\mathcal{C}}(x) := \{y \in \mathbf{R}^n \mid \sup_{x'\in\mathcal{C}} \langle x' - x, y\rangle \leq 0\}.$$

Note that $\Pi_{\mathcal{C}}$ is the proximal operator of the indicator function of $\mathcal{C}$. For a convex cone $\mathcal{K} \subseteq \mathbf{R}^n$, we denote its *polar cone* by

$$\mathcal{K}^{\circ} := \{y \in \mathbf{R}^n \mid \sup_{x\in\mathcal{K}} \langle x, y\rangle \leq 0\},$$

and for any $b \in \mathbf{R}^n$ we denote a translated cone by $\mathcal{K}_b := \mathcal{K} + \{b\}$.

Let $\mathcal{D}$ be a nonempty subset of $\mathbf{R}^n$. We denote the closure of $\mathcal{D}$ by $\overline{\mathcal{D}}$. For an operator $T : \mathcal{D} \mapsto \mathbf{R}^n$ we define its *fixed-point set* as $\operatorname{Fix} T := \{x \in \mathcal{D} \,|\, Tx = x\}$ and denote its range by $\operatorname{ran}(T)$. We say that T is *nonexpansive* if $(\forall x \in \mathcal{D})(\forall y \in \mathcal{D})$

$$\|Tx - Ty\| \leq \|x - y\|,$$

and T is α-*averaged* with $\alpha \in (0, 1)$ if there exists a nonexpansive operator $R : \mathcal{D} \to \mathbf{R}^n$ such that $T = (1-\alpha)\operatorname{Id} + \alpha R$.

2 Problem description

Consider the following convex optimization problem:

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \frac{1}{2}x^T P x + q^T x \\ \text{subject to} & Ax \in \mathcal{C}, \end{array} \tag{2}$$

with $P \in \mathbf{S}_{+}^{n}$, $q \in \mathbf{R}^n$, $A \in \mathbf{R}^{m\times n}$, and $\mathcal{C} \subseteq \mathbf{R}^m$ a nonempty, closed and convex set. We make the following assumption on the set $\mathcal{C}$:

Assumption 1. The set $\mathcal{C}$ is the Cartesian product of a convex compact set $\mathcal{B} \subseteq \mathbf{R}^{m_1}$ and a translated closed convex cone $\mathcal{K}_b \subseteq \mathbf{R}^{m_2}$, where m_1 and m_2 are nonnegative integers and $m_1 + m_2 = m$, *i.e.* $\mathcal{C} = \mathcal{B} \times \mathcal{K}_b$.

Many convex problems of practical interest, including LPs, QPs, SOCPs and SDPs, can be written in the form of problem (2) with $\mathcal{C}$ satisfying the conditions of Assumption 1. We are interested in finding either an optimal solution to problem (2) or a certificate of either primal or dual infeasibility.

2.1 Optimality conditions

We will find it convenient to rewrite problem (2) in an equivalent form by introducing a variable $z \in \mathbf{R}^m$ to obtain

$$\begin{array}{ll} \underset{x,z}{\text{minimize}} & \frac{1}{2}x^T P x + q^T x \\ \text{subject to} & Ax = z, \quad z \in \mathcal{C}. \end{array} \tag{3}$$

We can then write the optimality conditions for problem (3) as:

$$Ax - z = 0, \tag{4a}$$
$$Px + q + A^T y = 0, \tag{4b}$$
$$z \in \mathcal{C}, \quad y \in N_{\mathcal{C}}(z), \tag{4c}$$

where $y \in \mathbf{R}^m$ is a Lagrange multiplier associated with the constraint $Ax = z$. If there exist $x \in \mathbf{R}^n$, $z \in \mathbf{R}^m$ and $y \in \mathbf{R}^m$ that satisfy the conditions (4), then we say that (x, z) is a *primal* and y is a *dual* solution of problem (3). For completeness, we derive the optimality conditions in Lemma A.1 of the Appendix.

2.2 Infeasibility certificate

In this section we derive conditions for primal and dual infeasibility. The dual problem associated with problem (2) is

$$\begin{array}{ll} \underset{x,y}{\text{maximize}} & -\frac{1}{2}x^T P x - S_{\mathcal{C}}(y) \\ \text{subject to} & Px + A^T y = -q, \quad y \in (\mathcal{C}^\infty)^\circ \end{array} \tag{5}$$

and its derivation is included in Lemma A.2 of the Appendix.

We will use the following pair of results to certify infeasibility of (2) in cases where it is primal and/or dual *strongly infeasible*; we refer the reader to [LMT16] for more details on strong and weak infeasibility.

Proposition 1.

(i) If there exists some $\bar{y} \in \mathbf{R}^m$ such that

$$A^T \bar{y} = 0 \quad \text{and} \quad S_{\mathcal{C}}(\bar{y}) < 0, \tag{6}$$

then the primal problem (2) is infeasible.

(ii) If there exists some $\bar{x} \in \mathbf{R}^n$ such that

$$P\bar{x} = 0, \quad A\bar{x} \in \mathcal{C}^\infty \quad \text{and} \quad \langle q, \bar{x} \rangle < 0, \tag{7}$$

then the dual problem (5) is infeasible.

Proof. (i): The first condition in (6) implies

$$\inf_x \langle \bar{y}, Ax \rangle = \inf_x \left\langle A^T \bar{y}, x \right\rangle = 0,$$

and the second condition is equivalent to

$$\sup_{z\in\mathcal{C}} \langle \bar{y}, z \rangle < 0.$$

Therefore, $\{z \in \mathbf{R}^m \mid \langle \bar{y}, z \rangle = 0\}$ is a hyperplane that separates the sets $\{Ax \,|\, x \in \mathbf{R}^n\}$ and $\mathcal{C}$ strongly [Roc70, Thm. 11.1], meaning that the problem (2) is infeasible.

(ii): Define the set $\mathcal{Q} := \{Px + A^T y \,|\, (x, y) \in \mathbf{R}^n \times (\mathcal{C}^\infty)^\circ\}$. The first two conditions in (7) imply

$$\begin{aligned}
\sup_{s\in\mathcal{Q}} \langle \bar{x}, s \rangle &= \sup \left\{ \left\langle \bar{x}, Px + A^T y \right\rangle \,|\, x \in \mathbf{R}^n,\, y \in (\mathcal{C}^\infty)^\circ \right\} \\
&= \sup_x \langle P\bar{x}, x \rangle + \sup \left\{ \langle A\bar{x}, y \rangle \,|\, y \in (\mathcal{C}^\infty)^\circ \right\} \\
&\le 0,
\end{aligned}$$

where we used the fact that the inner product between vectors in a cone and its polar is nonpositive. Since the third condition in (7) can be written as $\langle \bar{x}, -q \rangle > 0$, this means that $\{x \in \mathbf{R}^n \mid \langle \bar{x}, x \rangle = 0\}$ is a hyperplane that separates the sets $\mathcal{Q}$ and $\{-q\}$ strongly, and thus the dual problem (5) is infeasible. □

Note that if the condition (6) in Proposition 1 holds, then $\bar{y}$ also represents an unbounded direction in the dual problem assuming it is feasible. Likewise, $\bar{x}$ in condition (7) represents an unbounded direction for the primal problem if it is feasible. However, since we cannot exclude the possibility of simultaneous primal and dual infeasibility, we will refer to the condition (6) as *primal infeasibility* rather than *dual unboundedness*, and *vice versa* for (7).

In some cases, *e.g.* when $\mathcal{C}$ is compact or polyhedral, conditions (6) and (7) in Proposition 1 are also necessary for infeasibility, and we say that (6) and (7) are *strong alternatives* for primal and dual feasibility, respectively. When $\mathcal{C}$ is a convex cone, additional assumptions are required for having strong alternatives; see *e.g.* [BV04, §5.9.4].

3 Alternating Direction Method of Multipliers (ADMM)

ADMM is an operator splitting method that can be used for solving structured optimization problems [BPC+11]. The iterates of ADMM in application to problem (1) can be written as

$$\tilde{w}^{k+1} \leftarrow \mathrm{prox}_f(w^k - u^k) \tag{8a}$$
$$w^{k+1} \leftarrow \mathrm{prox}_g\left(\alpha\tilde{w}^{k+1} + (1-\alpha)w^k + u^k\right) \tag{8b}$$
$$u^{k+1} \leftarrow u^k + \alpha\tilde{w}^{k+1} + (1-\alpha)w^k - w^{k+1}, \tag{8c}$$

Algorithm 1 ADMM for problem (2).

1: **given** initial values x^0, z^0, y^0 and parameters $\rho > 0, \sigma > 0, \alpha \in (0, 2)$
2: Set $k = 0$
3: **repeat**
4: $\quad (\tilde{x}^{k+1}, \tilde{z}^{k+1}) \leftarrow \underset{(\tilde{x},\tilde{z}):A\tilde{x}=\tilde{z}}{\operatorname{argmin}} \frac{1}{2}\tilde{x}^T P\tilde{x} + q^T\tilde{x} + \frac{\sigma}{2}\|\tilde{x} - x^k\|^2 + \frac{\rho}{2}\|\tilde{z} - z^k + \rho^{-1}y^k\|^2$
5: $\quad x^{k+1} \leftarrow \alpha\tilde{x}^{k+1} + (1-\alpha)x^k$
6: $\quad z^{k+1} \leftarrow \Pi_{\mathcal{C}}\left(\alpha\tilde{z}^{k+1} + (1-\alpha)z^k + \rho^{-1}y^k\right)$
7: $\quad y^{k+1} \leftarrow y^k + \rho\left(\alpha\tilde{z}^{k+1} + (1-\alpha)z^k - z^{k+1}\right)$
8: $\quad k \leftarrow k+1$
9: **until** termination condition is satisfied

where $\alpha \in (0, 2)$ is a *relaxation parameter*.

We can write problem (3) in the general form (1) by setting

$$\begin{aligned} f(x,z) &= \tfrac{1}{2}x^T Px + q^T x + \mathcal{I}_{Ax=z}(x,z), \\ g(x,z) &= \mathcal{I}_{\mathcal{C}}(z). \end{aligned} \tag{9}$$

If we use the norm $\|(x,z)\| = \sqrt{\sigma\|x\|_2^2 + \rho\|z\|_2^2}$ with $(\sigma, \rho) > 0$ in the proximal operators of functions $f(x,z)$ and $g(x,z)$, then ADMM reduces to Algorithm 1, which was first introduced in [SBG+18]. The scalars ρ and σ are called the *penalty parameters*. Note that the strict positivity of both ρ and σ ensure that the equality constrained QP in step 4 of Algorithm 1 has a unique solution for any $P \in \mathbf{S}_+^n$ and $A \in \mathbf{R}^{m\times n}$.

3.1 Reformulation as the Douglas-Rachford splitting (DRS)

It is well-known that ADMM and DRS are equivalent methods [Gab83]. The authors in [GFB16] show that the ADMM algorithm can be described alternatively in terms of the fixed-point iteration of the Douglas-Rachford operator which is known to be averaged [GB17]. In particular, the algorithm given by iteration (8) can alternatively be implemented as

$$w^k \leftarrow \operatorname{prox}_g(s^k) \tag{10a}$$
$$\tilde{w}^k \leftarrow \operatorname{prox}_f(2w^k - s^k) \tag{10b}$$
$$s^{k+1} \leftarrow s^k + \alpha(\tilde{w}^k - w^k). \tag{10c}$$

Similarly, an iteration of Algorithm 1 is equivalent to

$$(\tilde{x}^k, \tilde{z}^k) \leftarrow \underset{(\tilde{x},\tilde{z}):A\tilde{x}=\tilde{z}}{\operatorname{argmin}} \tfrac{1}{2}\tilde{x}^T P\tilde{x} + q^T\tilde{x} + \tfrac{\sigma}{2}\|\tilde{x} - x^k\|^2 + \tfrac{\rho}{2}\|\tilde{z} - (2\Pi_{\mathcal{C}} - \mathrm{Id})(v^k)\|^2 \tag{11a}$$
$$x^{k+1} \leftarrow x^k + \alpha\left(\tilde{x}^k - x^k\right) \tag{11b}$$
$$v^{k+1} \leftarrow v^k + \alpha\left(\tilde{z}^k - \Pi_{\mathcal{C}}(v^k)\right) \tag{11c}$$

where

$$z^k = \Pi_{\mathcal{C}}(v^k) \quad \text{and} \quad y^k = \rho(\operatorname{Id} - \Pi_{\mathcal{C}})(v^k). \tag{12}$$

We will exploit the following result in the next section to analyze asymptotic behavior of the algorithm.

Fact 1. The iteration described in (11) amounts to

$$(x^{k+1}, v^{k+1}) \leftarrow T(x^k, v^k),$$

where $T : \mathbf{R}^{n+m} \mapsto \mathbf{R}^{n+m}$ is an $(\alpha/2)$-averaged operator.

Proof. The iteration (10) is a special case of the iteration (49)–(51) in [GFB16, §IV-C] with $A = \operatorname{Id}$, $B = -\operatorname{Id}$ and $c = 0$, which is equivalent to

$$s^{k+1} \leftarrow T s^k,$$

where T is the Douglas-Rachford operator which is known to be $(\alpha/2)$-averaged [GB17]. The result follows from the fact that the iteration given by (11) is a special case of the iteration (10) with f and g given by (9). □

Due to [BC11, Prop. 6.46], the identities in (12) imply that at every iteration the pair (z^k, y^k) satisfies the optimality condition (4c) by construction. The solution of the equality constrained QP in (11a) satisfies the pair of optimality conditions

$$0 = A\tilde{x}^k - \tilde{z}^k \tag{13a}$$

$$0 = (P + \sigma I)\tilde{x}^k + q - \sigma x^k + \rho A^T \left(\tilde{z}^k - (2\Pi_{\mathcal{C}} - \operatorname{Id})(v^k)\right). \tag{13b}$$

If we rearrange (11b) and (11c) to isolate $\tilde{x}^k$ and $\tilde{z}^k$, *i.e.* write

$$\tilde{x}^k = x^k + \alpha^{-1}\delta x^{k+1} \tag{14a}$$

$$\tilde{z}^k = z^k + \alpha^{-1}\delta v^{k+1}, \tag{14b}$$

and substitute them into (13), we obtain the following relations between the iterates:

$$Ax^k - \Pi_{\mathcal{C}}(v^k) = -\alpha^{-1}\left(A\delta x^{k+1} - \delta v^{k+1}\right) \tag{15a}$$

$$Px^k + q + \rho A^T(\operatorname{Id} - \Pi_{\mathcal{C}})(v^k) = -\alpha^{-1}\left((P + \sigma I)\delta x^{k+1} + \rho A^T \delta v^{k+1}\right). \tag{15b}$$

Observe that the right-hand terms of (15) are a direct measure of how far the iterates (x^k, z^k, y^k) are from satisfying the optimality conditions (4a) and (4b). We refer to the left-hand terms of (15a) and (15b) as the *primal* and *dual residuals*, respectively. In the next section, we will show that the successive differences $(\delta x^k, \delta v^k)$ appearing in the right-hand side of (15) converge and can be used to test for primal and dual infeasibility.

4 Asymptotic behavior of ADMM

In order to analyze the asymptotic behavior of the iteration (11), which is equivalent to Algorithm 1, we will rely heavily on the following results:

Lemma 1. Let $\mathcal{D}$ be a nonempty, closed and convex subset of $\mathbf{R}^n$ and suppose that $T : \mathcal{D} \mapsto \mathcal{D}$ is an averaged operator. Let $s^0 \in \mathcal{D}$, $s^k = T^k s^0$ and δs be the projection of the zero vector onto $\overline{\operatorname{ran}(T - \mathrm{Id})}$. Then

(i) $\frac{1}{k} s^k \to \delta s$.
(ii) $\delta s^k \to \delta s$.
(iii) If $\operatorname{Fix} T \neq \emptyset$, then s^k converges to a point in $\operatorname{Fix} T$.

Proof. The first result is [Paz71, Cor. 3], the second is [BBR78, Cor. 2.3] and the third is [BC11, Thm. 5.14]. □

Note that since $\operatorname{ran}(T - \mathrm{Id})$ is not necessarily closed, the projection onto this set may not exist, but the projection onto its closure always exists. Moreover, since $\overline{\operatorname{ran}(T - \mathrm{Id})}$ is convex [Paz71, Lem. 4], the projection is unique. Due to Fact 1, Lemma 1 ensures that $(\frac{1}{k} x^k, \frac{1}{k} v^k) \to (\delta x, \delta v)$ and $(\delta x^k, \delta v^k) \to (\delta x, \delta v)$.

The core results of this paper are contained within the following two propositions, which establish various relationships between the limits δx and δv; we include several supporting results required to prove these results in the Appendix. Given these two results, it will then be straightforward to extract certificates of optimality or infeasibility in Section 4.1. For both of these central results, and in the remainder of the paper, we define

$$\delta z := \Pi_{\mathcal{C}^\infty}(\delta v) \quad \text{and} \quad \delta y := \rho \Pi_{(\mathcal{C}^\infty)^\circ}(\delta v).$$

Proposition 2. Suppose that Assumption 1 holds. Then the following relations hold between the limits δx, δz and δy:

(i) $A \delta x = \delta z$.
(ii) $P \delta x = 0$.
(iii) $A^T \delta y = 0$.
(iv) $\frac{1}{k} z^k \to \delta z$ and $\delta z^k \to \delta z$.
(v) $\frac{1}{k} y^k \to \delta y$ and $\delta y^k \to \delta y$.

Proof. Commensurate with our partitioning of the constraint set as $\mathcal{C} = \mathcal{B} \times \mathcal{K}_b$, we partition the matrix A and the iterates $(v^k,\ z^k,\ y^k)$ into components of appropriate dimension. We use subscript 1 for those components associated with the set $\mathcal{B}$ and subscript 2 for those associated with the set $\mathcal{K}_b$, *e.g.* $z^k = (z_1^k, z_2^k)$ where $z_1 \in \mathcal{B}$ and $z_2 \in \mathcal{K}_b$ and the matrix $A = [A_1; A_2]$. Note throughout that $\mathcal{C}^\infty = \{0\} \times \mathcal{K}$ and $(\mathcal{C}^\infty)^\circ = \mathbf{R}^{m_1} \times \mathcal{K}^\circ$.

(i): Divide (15a) by k, take the limit and apply Lemma 1 to get

$$A\delta x = \lim_{k\to\infty} \tfrac{1}{k}\Pi_{\mathcal{C}}(v^k).$$

Due to Lemma A.4 and the compactness of $\mathcal{B}$, we then obtain

$$\begin{aligned} A_1\delta x &= \lim_{k\to\infty} \tfrac{1}{k}\Pi_{\mathcal{B}}(v_1^k) = 0 \\ A_2\delta x &= \lim_{k\to\infty} \tfrac{1}{k}\Pi_{\mathcal{K}_b}(v_2^k) = \Pi_{\mathcal{K}}(\delta v_2). \end{aligned}$$

(ii): Divide (15b) by ρk, take the inner product of both sides with δx and take the limit to obtain

$$\begin{aligned} -\rho^{-1}\langle P\delta x, \delta x\rangle &= \lim_{k\to\infty} \left\langle A\delta x, \tfrac{1}{k}v_k - \tfrac{1}{k}\Pi_{\mathcal{C}}(v^k)\right\rangle \\ &= \left\langle A_1\delta x, \delta v_1 - \lim_{k\to\infty}\tfrac{1}{k}\Pi_{\mathcal{B}}(v_1^k)\right\rangle + \left\langle A_2\delta x, \delta v_2 - \lim_{k\to\infty}\tfrac{1}{k}\Pi_{\mathcal{K}_b}(v_2^k)\right\rangle \\ &= \left\langle \Pi_{\mathcal{K}}(\delta v_2), \delta v_2 - \Pi_{\mathcal{K}}(\delta v_2)\right\rangle \\ &= \langle \Pi_{\mathcal{K}}(\delta v_2), \Pi_{\mathcal{K}^\circ}(\delta v_2)\rangle \\ &= 0, \end{aligned}$$

where we used Lemma 1 in the second equality, $A_1\delta x = 0$, $A_2\delta x = \Pi_{\mathcal{K}}(\delta v_2)$, the compactness of $\mathcal{B}$ and Lemma A.4 in the third, and the Moreau decomposition [BC11, Thm. 6.29] in the fourth and fifth. Then $P\delta x = 0$ since $P \in \mathbf{S}_+^n$.

(iii): Divide (15b) by k, take the limit and use $P\delta x = 0$ to obtain

$$\begin{aligned} 0 &= \lim_{k\to\infty} \tfrac{1}{k}\rho A^T(\mathrm{Id} - \Pi_{\mathcal{C}})(v^k) \\ &= \lim_{k\to\infty} \tfrac{1}{k}\rho\left(A_1^T(\mathrm{Id} - \Pi_{\mathcal{B}})(v_1^k) + A_2^T(\mathrm{Id} - \Pi_{\mathcal{K}_b})(v_2^k)\right) \\ &= \rho A_1^T\left(\delta v_1 - \lim_{k\to\infty}\tfrac{1}{k}\Pi_{\mathcal{B}}(v_1^k)\right) + \rho A_2^T\left(\delta v_2 - \lim_{k\to\infty}\tfrac{1}{k}\Pi_{\mathcal{K}_b}(v_2^k)\right) \\ &= \rho A_1^T\delta v_1 + \rho A_2^T\left(\delta v_2 - \Pi_{\mathcal{K}}(\delta v_2)\right) \\ &= \rho A_1^T\delta v_1 + \rho A_2^T\Pi_{\mathcal{K}^\circ}(\delta v_2) \\ &= \rho \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}^T \begin{bmatrix} \delta v_1 \\ \Pi_{\mathcal{K}^\circ}(\delta v_2) \end{bmatrix} \\ &= A^T\rho\Pi_{(\mathcal{C}^\infty)^\circ}(\delta v) \\ &= A^T\delta y, \end{aligned}$$

where we used Lemma 1 in the third equality, Lemma A.4 and the compactness of $\mathcal{B}$ in the fourth, the Moreau decomposition in the fifth, and $(\mathcal{C}^\infty)^\circ = \mathbf{R}^{m_1} \times \mathcal{K}^\circ$ in the seventh.

(iv): We first show that the sequence $\{\delta z^k\}_{k\in\mathbf{N}}$ converges to δz. From (14) we have

$$-\alpha^{-1}\left(\delta x^{k+1} - \delta x^k\right) = \delta x^k - \delta\tilde{x}^k, \tag{16a}$$

$$-\alpha^{-1}\left(\delta v^{k+1} - \delta v^k\right) = \delta z^k - \delta\tilde{z}^k. \tag{16b}$$

Take the limit of (16a) to obtain

$$\lim_{k\to\infty} \delta\tilde{x}^k = \lim_{k\to\infty} \delta x^k = \delta x.$$

From (13a) we now have $\delta\tilde{z}^k = A\delta\tilde{x}^k \to A\delta x$. Take the limit of (16b) and use the result from (i) to obtain

$$\lim_{k\to\infty} \delta z^k = \lim_{k\to\infty} \delta\tilde{z}^k = A\delta x = \Pi_{\mathcal{C}^\infty}(\delta v).$$

We now show that the sequence $\{\frac{1}{k}z^k\}_{k\in\mathbf{N}}$ also converges to δz. Dividing the expression for z^k in (12) by k and taking the limit, we obtain

$$\lim_{k\to\infty} \tfrac{1}{k}z^k = \lim_{k\to\infty} \tfrac{1}{k}\begin{bmatrix}\Pi_{\mathcal{B}}(v_1^k)\\ \Pi_{\mathcal{K}_b}(v_2^k)\end{bmatrix} = \begin{bmatrix}0\\ \Pi_{\mathcal{K}}(\delta v_2)\end{bmatrix} = \Pi_{\mathcal{C}^\infty}(\delta v).$$

(v): We first show that the sequence $\{\delta y^k\}_{k\in\mathbf{N}}$ converges to δy. Since $\rho^{-1}y^k = v^k - z^k$, we have

$$\begin{aligned}\lim_{k\to\infty} \rho^{-1}\delta y^k &= \lim_{k\to\infty} \delta v^k - \lim_{k\to\infty} \delta z^k\\ &= \delta v - \Pi_{\mathcal{C}^\infty}(\delta v)\\ &= \Pi_{(\mathcal{C}^\infty)^\circ}(\delta v),\end{aligned}$$

where we used the Moreau decomposition in the last equality.

We now show that the sequence $\{\frac{1}{k}y^k\}_{k\in\mathbf{N}}$ also converges to δy. Dividing the expression for y^k in (12) by k and taking the limit, we obtain

$$\lim_{k\to\infty} \tfrac{1}{k}y^k = \rho\lim_{k\to\infty} \tfrac{1}{k}(v^k - z^k) = \rho(\delta v - \delta z) = \rho(\delta v - \Pi_{\mathcal{C}^\infty}(\delta v)) = \rho\Pi_{(\mathcal{C}^\infty)^\circ}(\delta v). \qquad \square$$

Proposition 2 shows that the limits δy and δx will always satisfy the subspace and conic constraints in the primal and dual infeasibility tests (6) and (7), respectively. We next consider the terms appearing in the inequalities in (6) and (7).

Proposition 3. Suppose that Assumption 1 holds. Then the following identities hold for the limits δx and δy:

(i) $\langle q, \delta x\rangle = -\sigma\alpha^{-1}\|\delta x\|^2 - \rho\alpha^{-1}\|A\delta x\|^2$.
(ii) $S_{\mathcal{C}}(\delta y) = -\rho^{-1}\alpha^{-1}\|\delta y\|^2$.

Proof. Take the inner product of both sides of (15b) with δx and use Proposition 2(ii) to obtain

$$\langle \delta x, q\rangle + \rho\left\langle A\delta x, (\mathrm{Id} - \Pi_{\mathcal{C}})(v^k)\right\rangle = -\sigma\alpha^{-1}\left\langle \delta x, \delta x^{k+1}\right\rangle - \rho\alpha^{-1}\left\langle A\delta x, \delta v^{k+1}\right\rangle.$$

Using $A_1\delta x = 0$ from Proposition 2(i) and then taking the limit gives

$$\begin{aligned}\langle q, \delta x\rangle &= -\sigma\alpha^{-1}\|\delta x\|^2 - \rho\alpha^{-1}\langle A\delta x, \delta v\rangle - \rho \lim_{k\to\infty}\left\langle A_2\delta x, \Pi_{\mathcal{K}^\circ}(v_2^k - b)\right\rangle \\ &= -\sigma\alpha^{-1}\|\delta x\|^2 - \rho\alpha^{-1}\langle \Pi_{\mathcal{C}^\infty}(\delta v), \delta v\rangle - \rho \lim_{k\to\infty}\left\langle \Pi_{\mathcal{K}}(\delta v_2), \Pi_{\mathcal{K}^\circ}(v_2^k - b)\right\rangle \\ &= -\sigma\alpha^{-1}\|\delta x\|^2 - \rho\alpha^{-1}\|\Pi_{\mathcal{C}^\infty}(\delta v)\|^2 - \rho \lim_{k\to\infty}\left\langle \Pi_{\mathcal{K}}(\delta v_2), \Pi_{\mathcal{K}^\circ}(v_2^k - b)\right\rangle,\end{aligned} \tag{17}$$

where we used Lemma A.3(ii) in the first equality, Proposition 2(i) in the second, and Lemma A.3(iv) in the third.

Now take the inner product of both sides of (15a) with $\Pi_{(\mathcal{C}^\infty)^\circ}(\delta v)$ to obtain

$$\alpha^{-1}\left\langle \Pi_{(\mathcal{C}^\infty)^\circ}(\delta v), \delta v^{k+1}\right\rangle = \left\langle A^T\Pi_{(\mathcal{C}^\infty)^\circ}(\delta v), x^k + \alpha^{-1}\delta x^{k+1}\right\rangle - \left\langle \Pi_{(\mathcal{C}^\infty)^\circ}(\delta v), \Pi_{\mathcal{C}}(v^k)\right\rangle.$$

According to Proposition 2(iii) the first inner product on the right-hand side is zero, and taking the limit we obtain

$$\begin{aligned}\lim_{k\to\infty}\left\langle \Pi_{(\mathcal{C}^\infty)^\circ}(\delta v), \Pi_{\mathcal{C}}(v^k)\right\rangle &= -\alpha^{-1}\left\langle \Pi_{(\mathcal{C}^\infty)^\circ}(\delta v), \delta v\right\rangle \\ &= -\alpha^{-1}\|\Pi_{(\mathcal{C}^\infty)^\circ}(\delta v)\|^2,\end{aligned}$$

where the second equality follows from Lemma A.3(iv). The limit in the above identity can be expressed as

$$\begin{aligned}\lim_{k\to\infty}\left\langle \Pi_{(\mathcal{C}^\infty)^\circ}(\delta v), \Pi_{\mathcal{C}}(v^k)\right\rangle &= \lim_{k\to\infty}\left\langle \delta v_1, \Pi_{\mathcal{B}}(v_1^k)\right\rangle + \lim_{k\to\infty}\left\langle \Pi_{\mathcal{K}^\circ}(\delta v_2), \Pi_{\mathcal{K}_b}(v_2^k)\right\rangle \\ &= S_{\mathcal{B}}(\delta v_1) + \lim_{k\to\infty}\left\langle \Pi_{\mathcal{K}^\circ}(\delta v_2), b + \Pi_{\mathcal{K}}(v_2^k - b)\right\rangle \\ &= S_{\mathcal{B}}(\delta v_1) + \langle \Pi_{\mathcal{K}^\circ}(\delta v_2), b\rangle + \lim_{k\to\infty}\left\langle \Pi_{\mathcal{K}^\circ}(\delta v_2), \Pi_{\mathcal{K}}(v_2^k - b)\right\rangle,\end{aligned}$$

where the second equality follows from Lemma A.3(i) and Lemma A.5. Since the support function of $\mathcal{K}$ evaluated at any point in $\mathcal{K}^\circ$ is zero, we can write

$$\langle b, \Pi_{\mathcal{K}^\circ}(\delta v_2)\rangle = \langle b, \Pi_{\mathcal{K}^\circ}(\delta v_2)\rangle + \sup_{z\in\mathcal{K}}\langle z, \Pi_{\mathcal{K}^\circ}(\delta v_2)\rangle = S_{\mathcal{K}_b}(\Pi_{\mathcal{K}^\circ}(\delta v_2)).$$

Now we have

$$\begin{aligned}S_{\mathcal{C}}(\Pi_{(\mathcal{C}^\infty)^\circ}(\delta v)) &= S_{\mathcal{B}}(\delta v_1) + S_{\mathcal{K}_b}(\Pi_{\mathcal{K}^\circ}(\delta v_2)) \\ &= -\alpha^{-1}\|\Pi_{(\mathcal{C}^\infty)^\circ}(\delta v)\|^2 - \lim_{k\to\infty}\left\langle \Pi_{\mathcal{K}^\circ}(\delta v_2), \Pi_{\mathcal{K}}(v_2^k - b)\right\rangle,\end{aligned}$$

and due to the positive homogeneity of the support function,

$$S_{\mathcal{C}}(\delta y) = -\rho\alpha^{-1}\|\Pi_{(\mathcal{C}^\infty)^\circ}(\delta v)\|^2 - \rho\lim_{k\to\infty}\left\langle \Pi_{\mathcal{K}^\circ}(\delta v_2), \Pi_{\mathcal{K}}(v_2^k - b)\right\rangle. \tag{18}$$

We will next show that the limits in (17) and (18) are equal to zero. Summing the two equalities, we obtain

$$\begin{aligned}\langle q,\delta x\rangle + S_{\mathcal{C}}(\delta y) + \sigma\alpha^{-1}\|\delta x\|^2 + \rho\alpha^{-1}\|\delta v\|^2 = -\rho\lim_{k\to\infty}&\left\langle \Pi_{\mathcal{K}}(\delta v_2), \Pi_{\mathcal{K}^\circ}(v_2^k - b)\right\rangle \\ -\rho\lim_{k\to\infty}&\left\langle \Pi_{\mathcal{K}^\circ}(\delta v_2), \Pi_{\mathcal{K}}(v_2^k - b)\right\rangle,\end{aligned} \tag{19}$$

where we used $\|\delta v\|^2 = \|\Pi_{\mathcal{C}^\infty}(\delta v)\|^2 + \|\Pi_{(\mathcal{C}^\infty)^\circ}(\delta v)\|^2$ [BC11, Thm. 6.29].

Take the inner product of both sides of (15b) with x^k to obtain

$$\begin{aligned}\left\langle Px^k, x^k\right\rangle + \left\langle q, x^k\right\rangle + \rho\left\langle Ax^k, (\mathrm{Id}-\Pi_{\mathcal{C}})(v^k)\right\rangle = -\alpha^{-1}&\left\langle P\delta x^{k+1}, x^k\right\rangle - \sigma\alpha^{-1}\left\langle \delta x^{k+1}, x^k\right\rangle \\ -\rho\alpha^{-1}&\left\langle Ax^k, \delta v^{k+1}\right\rangle.\end{aligned} \tag{20}$$

We can rewrite the third inner product on the left-hand side of the equality above as

$$\begin{aligned}\left\langle Ax^k, (\mathrm{Id}-\Pi_{\mathcal{C}})(v^k)\right\rangle &= \left\langle \Pi_{\mathcal{C}}(v^k) + \alpha^{-1}\left(\delta v^{k+1} - A\delta x^{k+1}\right), (\mathrm{Id}-\Pi_{\mathcal{C}})(v^k)\right\rangle \\ &= \left\langle \Pi_{\mathcal{B}}(v_1^k), v_1^k\right\rangle - \|\Pi_{\mathcal{B}}(v_1^k)\|^2 + \left\langle \Pi_{\mathcal{K}_b}(v_2^k), (\mathrm{Id}-\Pi_{\mathcal{K}_b})(v_2^k)\right\rangle \\ &\quad + \alpha^{-1}\left\langle \delta v^{k+1} - A\delta x^{k+1}, \rho^{-1}y^k\right\rangle \\ &= \left\langle \Pi_{\mathcal{B}}(v_1^k), v_1^k\right\rangle - \|\Pi_{\mathcal{B}}(v_1^k)\|^2 + \left\langle b, \Pi_{\mathcal{K}^\circ}(v_2^k - b)\right\rangle \\ &\quad + \alpha^{-1}\left\langle \delta v^{k+1} - A\delta x^{k+1}, \rho^{-1}y^k\right\rangle,\end{aligned}$$

where we used (15a) in the first equality, (12) in the second, and Lemma A.3(iii) in the third. Substituting this expression into (20), dividing by k and taking the limit, we then obtain

$$\begin{aligned}&\lim_{k\to\infty}\tfrac{1}{k}\left\langle Px^k, x^k\right\rangle + \langle q,\delta x\rangle + \rho\lim_{k\to\infty}\tfrac{1}{k}\left\langle \Pi_{\mathcal{B}}(v_1^k), v_1^k\right\rangle - \rho\lim_{k\to\infty}\tfrac{1}{k}\|\Pi_{\mathcal{B}}(v_1^k)\|^2 + \rho\left\langle b, \Pi_{\mathcal{K}^\circ}(\delta v_2)\right\rangle \\ &+ \rho\alpha^{-1}\left\langle \delta v - A\delta x, \rho^{-1}\delta y\right\rangle = -\alpha^{-1}\langle P\delta x, \delta x\rangle - \sigma\alpha^{-1}\|\delta x\|^2 - \rho\alpha^{-1}\langle A\delta x, \delta v\rangle.\end{aligned}$$

Due to Lemma A.5, Lemma 2(ii) and the compactness of $\mathcal{B}$, the equality above simplifies to

$$\begin{aligned}\lim_{k\to\infty}\tfrac{1}{k}\left\langle Px^k, x^k\right\rangle + \langle q,\delta x\rangle + S_{\mathcal{C}}(\delta y) + \sigma\alpha^{-1}\|\delta x\|^2 = -\rho\alpha^{-1}&\left\langle \delta v - A\delta x, \rho^{-1}\delta y\right\rangle \\ -\rho\alpha^{-1}&\langle A\delta x, \delta v\rangle.\end{aligned} \tag{21}$$

The sum of inner products appearing on the right-hand side of the equality above can be written as

$$\begin{aligned}\left\langle \delta v - A\delta x, \rho^{-1}\delta y\right\rangle + \langle A\delta x, \delta v\rangle &= \left\langle \delta v - \Pi_{\mathcal{C}^\infty}(\delta v), \Pi_{(\mathcal{C}^\infty)^\circ}(\delta v)\right\rangle + \langle \Pi_{\mathcal{C}^\infty}(\delta v), \delta v\rangle \\ &= \|\Pi_{(\mathcal{C}^\infty)^\circ}(\delta v)\|^2 + \|\Pi_{\mathcal{C}^\infty}(\delta v)\|^2 \\ &= \|\delta v\|^2,\end{aligned}$$

where we used Proposition 2(i) in the first equality, and Lemma A.3(iv) and the Moreau decomposition in the second. Substituting this expression into (21), we obtain

$$\langle q,\delta x\rangle + S_{\mathcal{C}}(\delta y) + \sigma\alpha^{-1}\|\delta x\|^2 + \rho\alpha^{-1}\|\delta v\|^2 = -\lim_{k\to\infty}\tfrac{1}{k}\left\langle Px^k, x^k\right\rangle. \tag{22}$$

Comparing identities in (19) and (22), we get the following relation:

$$\lim_{k\to\infty} \tfrac{1}{k} \left\langle Px^k, x^k \right\rangle = \rho \lim_{k\to\infty} \left\langle \Pi_{\mathcal{K}}(\delta v_2), \Pi_{\mathcal{K}^\circ}(v_2^k - b) \right\rangle + \rho \lim_{k\to\infty} \left\langle \Pi_{\mathcal{K}^\circ}(\delta v_2), \Pi_{\mathcal{K}}(v_2^k - b) \right\rangle .$$

Positive semidefiniteness of P implies that the sequence on the left-hand side is term-wise nonnegative. Since the two sequences on the right-hand side involve inner products of elements in $\mathcal{K}$ and its polar, each sequence is term-wise nonpositive. Consequently, each of these limits must be zero. The claims of the proposition then follow directly from (17) and (18). □

4.1 Optimality and infeasibility certificates

We are now in a position to prove that, in the limit, the iterates of Algorithm 1 either satisfy the optimality conditions (4) or produce a certificate of strong infeasibility. Recall that Fact 1, Lemma 1(ii) and Proposition 2(iv)–(v) ensure convergence of the sequence $\{\delta x^k, \delta z^k, \delta y^k\}_{k\in\mathbf{N}}$.

Proposition 4 (Optimality)**.** If $(\delta x^k, \delta z^k, \delta y^k) \to 0$, then the optimality conditions (4) are satisfied in the limit, *i.e.*

$$\|Px^k + q + A^T y^k\| \to 0 \quad \text{and} \quad \|Ax^k - z^k\| \to 0. \tag{23}$$

Proof. Follows from (12) and (15). □

Lemma 1(iii) is sufficient to prove that if problem (2) is solvable then the sequence of iterates $\{x^k, z^k, y^k\}_{k\in\mathbf{N}}$ converges to its primal-dual solution. However, convergence of $\{\delta x^k, \delta z^k, \delta y^k\}_{k\in\mathbf{N}}$ to zero is not itself sufficient to prove convergence of $\{x^k, z^k, y^k\}_{k\in\mathbf{N}}$; we provide a numerical example in Section 5.3 to show when this scenario can occur. According to Proposition 4, in this case the violation of optimality conditions still goes to zero in the limit.

We next show that if $\{\delta x^k, \delta z^k, \delta y^k\}_{k\in\mathbf{N}}$ converges to a nonzero value, then we can construct a certificate of primal and/or dual infeasibility. Note that due to Proposition 2(i), δz can be nonzero only when δx is nonzero.

Theorem 1. Suppose that Assumption 1 holds.

(i) If $\delta y \neq 0$, then the problem (2) is infeasible and δy satisfies the primal infeasibility conditions (6).

(ii) If $\delta x \neq 0$, then the problem (5) is infeasible and δx satisfies the dual infeasibility conditions (7).

(iii) If $\delta x \neq 0$ and $\delta y \neq 0$, then problems (2) and (5) are simultaneously infeasible.

Proof. (i): Follows from Proposition 2(iii) and Proposition 3(ii).

(ii): Follows from Proposition 2(i)–(ii) and Proposition 3(i).

(iii): Follows from (i) and (ii). □

Remark 1. It is easy to show that δy^k and δx^k would still provide certificates of primal and dual infeasibility if we instead used the norm $\|(x,z)\| = \sqrt{x^T S x + z^T R z}$ in the proximal operators in (8), with R and S being diagonal positive definite matrices.

4.2 Termination criteria

We can define termination criteria for Algorithm 1 so that the iterations stop when either a primal-dual solution or a certificate of primal or dual infeasibility is found with some predefined accuracy.

A reasonable criterion for detecting optimality is that the norms of primal and dual residuals are smaller than some tolerance levels $\varepsilon_{\mathrm{prim}} > 0$ and $\varepsilon_{\mathrm{dual}} > 0$, respectively, *i.e.*

$$\|Ax^k - z^k\| \le \varepsilon_{\mathrm{prim}}, \quad \|Px^k + q + A^T y^k\| \le \varepsilon_{\mathrm{dual}}.$$

Since $(\delta x^k, \delta y^k) \to (\delta x, \delta y)$, a meaningful criterion for detecting primal and dual infeasibility would be to use δy^k and δx^k to check that the conditions (6) and (7) are almost satisfied, *i.e.*

$$\|A^T \delta y^k\| \le \varepsilon_{\mathrm{pinf}}, \quad S_{\mathcal{C}}(\delta y^k) < \varepsilon_{\mathrm{pinf}}, \tag{24}$$

and

$$\|P\delta x^k\| \le \varepsilon_{\mathrm{dinf}}, \quad \langle q, \delta x^k \rangle < \varepsilon_{\mathrm{dinf}}, \quad \mathrm{dist}_{\mathcal{C}^\infty}(A\delta x^k) < \varepsilon_{\mathrm{dinf}}, \tag{25}$$

where $\varepsilon_{\mathrm{pinf}} > 0$ and $\varepsilon_{\mathrm{dinf}} > 0$. Infeasibility detection based on these vectors is used in OSQP [SBG+18], an open-source operator splitting solver for quadratic programming. Note that the tolerance levels are often chosen relative to the scaling of the algorithm's iterates and the problem data; see [SBG+18, Sec. 3.4] for details.

For any positive tolerance levels $\varepsilon_{\mathrm{prim}} > 0$, $\varepsilon_{\mathrm{dual}} > 0$, $\varepsilon_{\mathrm{pinf}} > 0$ and $\varepsilon_{\mathrm{dinf}} > 0$, at least one of the termination criteria will be satisfied after finite time. For weakly infeasible problems both optimality and infeasibility conditions will be satisfied for any given accuracy. This means that an infinitesimally small perturbation to the problem can make it either optimal or strongly infeasible. We provide an example in Section 5.3 illustrating such case.

Remark 2. Even though $(\delta x^k, \delta y^k) \to (\delta x, \delta y)$, termination criteria for detecting infeasibility should not be implemented by simply checking that successive terms in the sequences $\{\delta x^k\}_{k\in\mathbf{N}}$ and $\{\delta y^k\}_{k\in\mathbf{N}}$ are close together. The reason is that these sequences can take values which repeat for many iterations even though they have not reached their limit points, and such repeated values in these sequences will not constitute infeasibility certificates. Instead, we check the infeasibility conditions (24) and (25) directly, with the understanding that these conditions will necessarily be satisfied in the limit for infeasible problems.

5 Numerical examples

In this section we demonstrate via several numerical examples the different asymptotic behaviors of iterates generated by Algorithm 1 for solving optimization problems of the form (2).

5.1 Parametric QP

Consider the QP

$$\begin{array}{ll} \underset{x_1,x_2}{\text{minimize}} & \frac{1}{2}x_1^2 + x_1 - x_2 \\ \text{subject to} & 0 \leq x_1 + ax_2 \leq u_1 \\ & 1 \leq x_1 \leq 3 \\ & 1 \leq x_2 \leq u_3, \end{array} \tag{26}$$

where $a \in \mathbf{R}$, $u_1 \geq 0$ and $u_3 \geq 1$ are parameters. Note that the above problem is an instance of problem (2) with

$$P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad q = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & a \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathcal{C} = [l, u], \quad l = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad u = \begin{bmatrix} u_1 \\ 3 \\ u_3 \end{bmatrix},$$

where $[l, u] \coloneqq \{x \in \mathbf{R}^m \mid l \leq x \leq u\}$. Depending on the values of parameters u_1 and u_3, the constraint set in (26) can be either bounded or unbounded. The projection onto the set $[l, u]$ can be evaluated as

$$\Pi_{[l,u]}(x) = \max\left(\min(x, u), l\right),$$

and the support function of the bounded set $\mathcal{B} = [l, u]$ as

$$S_{\mathcal{B}}(y) = \langle l, \min(y, 0)\rangle + \langle u, \max(y, 0)\rangle,$$

where min and max functions should be taken element-wise. The support function of the translated cone $\mathcal{K}_b$ is

$$S_{\mathcal{K}_b}(y) = \begin{cases} \langle b, y\rangle & y \in \mathcal{K}^\circ \\ +\infty & \text{otherwise.} \end{cases}$$

Note that the condition $S_{\mathcal{K}_b}(y) < 0$ is equivalent to

$$y \in \mathcal{K}^\circ \quad \text{and} \quad \langle b, y\rangle < 0.$$

In the sequel we will discuss four scenarios that can occur depending on the values of the parameters: (i) optimality, (ii) primal infeasibility, (iii) dual infeasibility, (iv) simultaneous primal and dual infeasibility, and will show that Algorithm 1 correctly produces certificates for all four scenarios. In all cases we set the parameters $\alpha = \rho = \sigma = 1$ and set the initial iterate $(x^0, z^0, y^0) = (0, 0, 0)$.

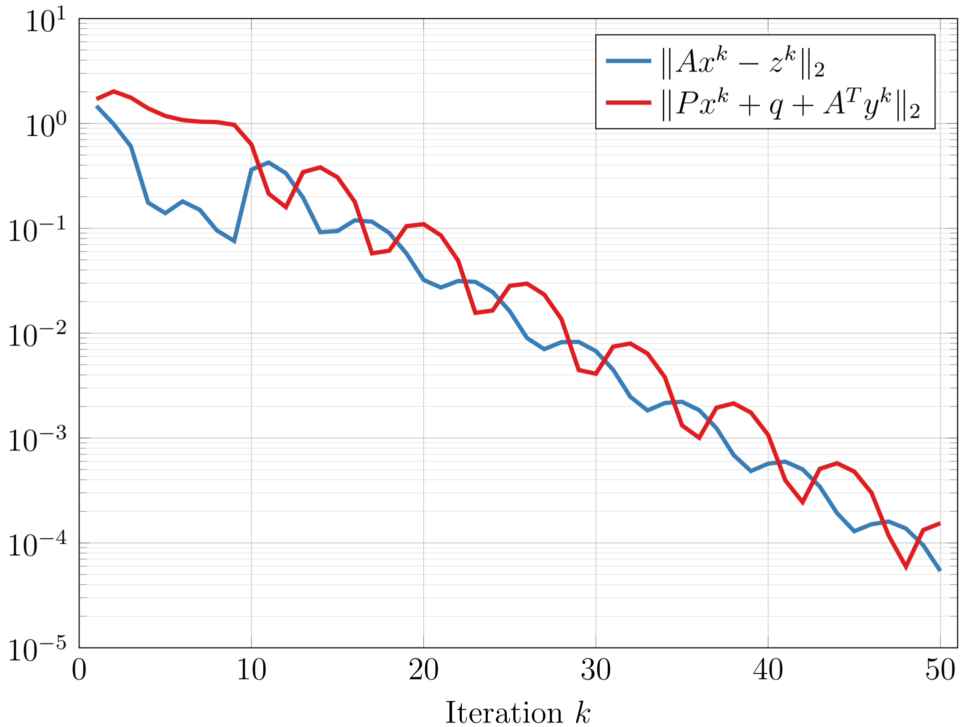

Figure 1: Convergence of $\{x^k, z^k, y^k\}_{k\in\mathbf{N}}$ to a certificate of optimality for problem (26) with $a = 1$, $u_1 = 5$ and $u_3 = 3$.

Optimality. Consider the problem (26) with parameters

$$a = 1, \quad u_1 = 5, \quad u_3 = 3.$$

Algorithm 1 converges to $x^\star = (1, 3)$, $z^\star = (4, 1, 3)$, $y^\star = (0, -2, 1)$, for which the objective value equals -1.5, and we have

$$Ax^\star - z^\star = 0 \quad \text{and} \quad Px^\star + q + A^T y^\star = 0,$$

i.e. the pair $(x^\star, y^\star)$ is a primal-dual solution of problem (26). Figure 1 shows convergence of $\{x^k, z^k, y^k\}_{k\in\mathbf{N}}$ to a certificate of optimality. Recall that the iterates of the algorithm always satisfy the optimality conditions (4c).

Primal infeasibility. We next set the parameters of problem (26) to

$$a = 1, \quad u_1 = 0, \quad u_3 = 3.$$

Note that in this case the constraint set is $\mathcal{C} = \mathcal{B} = \{0\} \times [1, 3] \times [1, 3]$. The sequence $\{\delta y^k\}_{k\in\mathbf{N}}$ generated by Algorithm 1 converges to $\delta y = (2/3, -2/3, -2/3)$, and we have

$$A^T \delta y = 0 \quad \text{and} \quad S_{\mathcal{C}}(\delta y) = -4/3 < 0.$$

According to Proposition 1(i), δy is a certificate of primal infeasibility for the problem. Figure 2 shows convergence of $\{\delta y^k\}_{k\in\mathbf{N}}$ to a certificate of primal infeasibility.

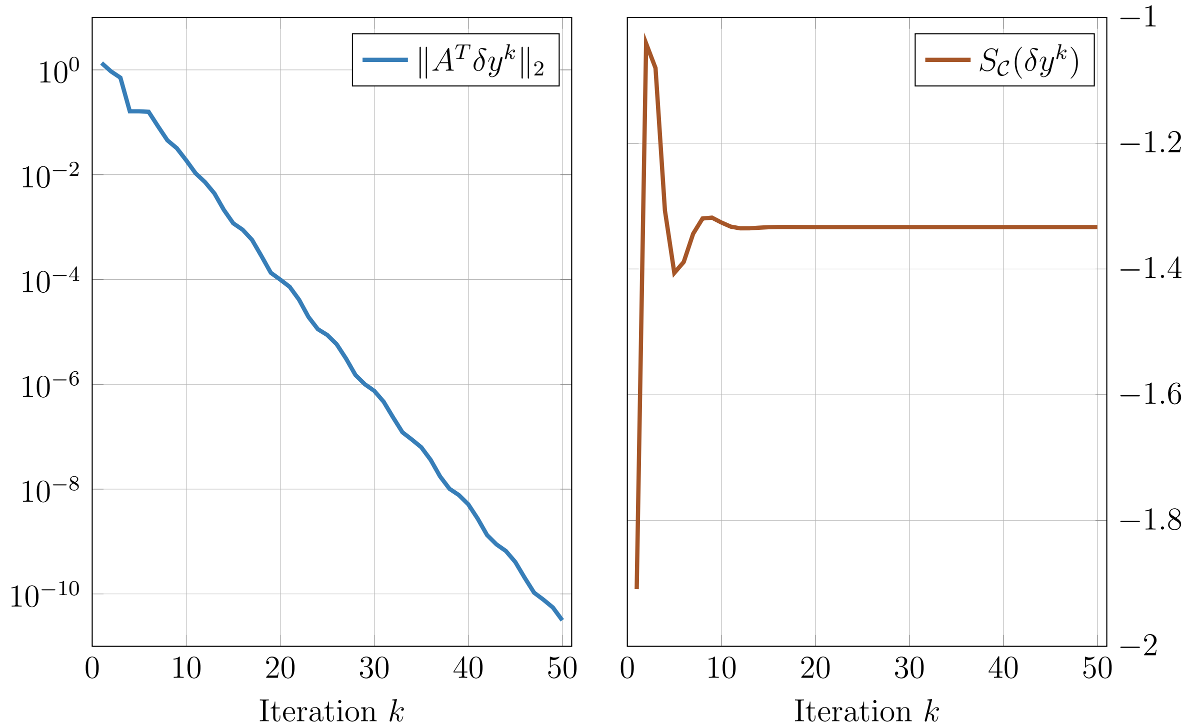

Figure 2: Convergence of $\{\delta y^k\}_{k\in\mathbf{N}}$ to a certificate of primal infeasibility for problem (26) with $a = 1$, $u_1 = 0$ and $u_3 = 3$.

Dual infeasibility. We set the parameters to

$$a = 0, \quad u_1 = 2, \quad u_3 = +\infty.$$

The constraint set has the form $\mathcal{C} = \mathcal{B} \times \mathcal{K}_b$ with

$$\mathcal{B} = [0, 2] \times [1, 3], \quad \mathcal{K} = \mathbf{R}_+, \quad b = 1,$$

and the constraint matrix A can be written as

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \quad \text{with} \quad A_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 0 & 1 \end{bmatrix}. \tag{27}$$

The sequence $\{\delta x^k\}_{k\in\mathbf{N}}$ generated by Algorithm 1 converges to $\delta x = (0, \frac{1}{2})$, and we have

$$P\delta x = 0, \quad A_1\delta x = 0, \quad A_2\delta x = \tfrac{1}{2} \in \mathcal{K}, \quad \langle q, \delta x\rangle = -\tfrac{1}{2} < 0.$$

According to Proposition 1(ii), δx is a certificate of dual infeasibility of the problem. Figure 3 shows convergence of $\{\delta x^k\}_{k\in\mathbf{N}}$ to a certificate of dual infeasibility, where $\mathrm{dist}_{\mathcal{C}^\infty}$ denotes the Euclidean distance to the set $\mathcal{C}^\infty = \{0\} \times \{0\} \times \mathbf{R}_+$.

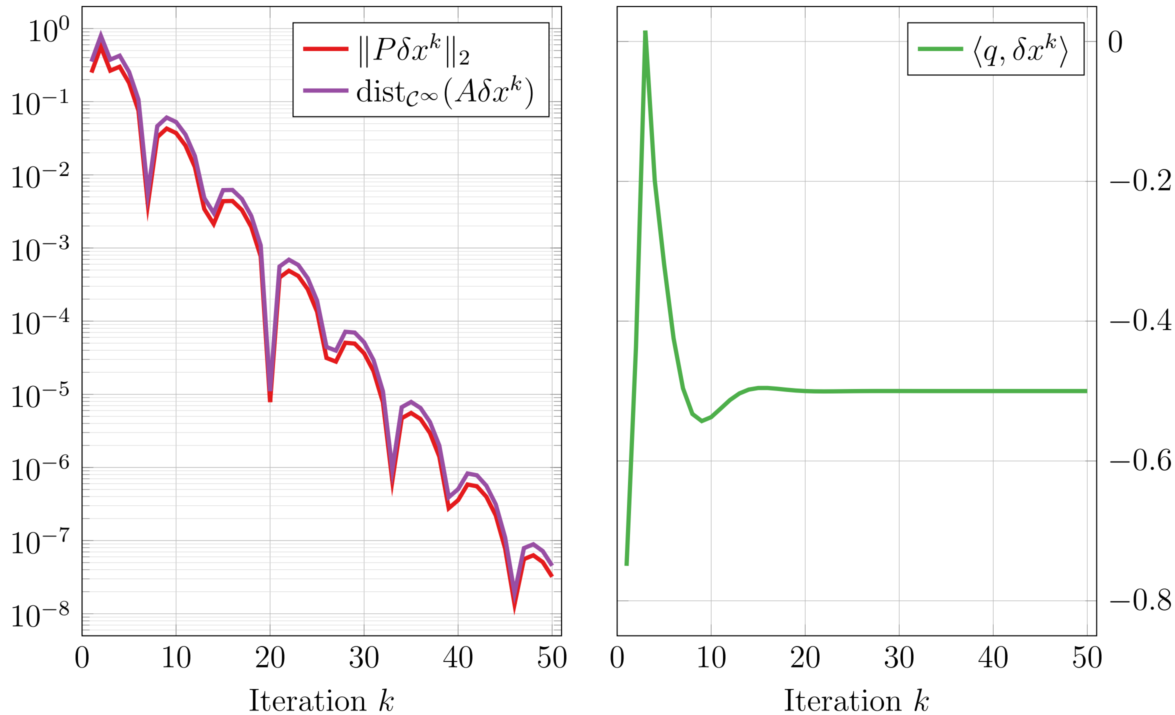

Figure 3: Convergence of $\{\delta x^k\}_{k\in\mathbf{N}}$ to a certificate of dual infeasibility for problem (26) with $a = 0$, $u_1 = 2$ and $u_3 = +\infty$.

Simultaneous primal and dual infeasibility. We set

$$a = 0, \quad u_1 = 0, \quad u_3 = +\infty.$$

The constraint set has the form $\mathcal{C} = \mathcal{B} \times \mathcal{K}_b$ with

$$\mathcal{B} = \{0\} \times [1, 3], \quad \mathcal{K} = \mathbf{R}_+, \quad b = 1,$$

and the constraint matrix A can be written as in (27). The sequences $\{\delta x^k\}_{k\in\mathbf{N}}$ and $\{\delta y^k\}_{k\in\mathbf{N}}$ generated by Algorithm 1 converge to $\delta x = (0, \frac{1}{2})$ and $\delta y = (\frac{1}{2}, -\frac{1}{2}, 0)$, respectively. If we partition δy as $\delta y = (\delta y_1, \delta y_2)$ with $\delta y_1 = (\frac{1}{2}, -\frac{1}{2})$ and $\delta y_2 = 0$, then we have

$$A^T \delta y = 0, \quad S_{\mathcal{C}}(\delta y) = S_{\mathcal{B}}(\delta y_1) + S_{\mathcal{K}_b}(\delta y_2) = -\tfrac{1}{2} < 0,$$

and

$$P\delta x = 0, \quad A_1 \delta x = 0, \quad A_2 \delta x = \tfrac{1}{2} \in \mathcal{K}, \quad \langle q, \delta x \rangle = -\tfrac{1}{2} < 0.$$

Therefore, δx and δy are certificates that the problem is simultaneously primal and dual infeasible. Figure 4 shows convergence of $\{\delta y^k\}_{k\in\mathbf{N}}$ and $\{\delta x^k\}_{k\in\mathbf{N}}$ to certificates of primal and dual infeasibility, respectively.

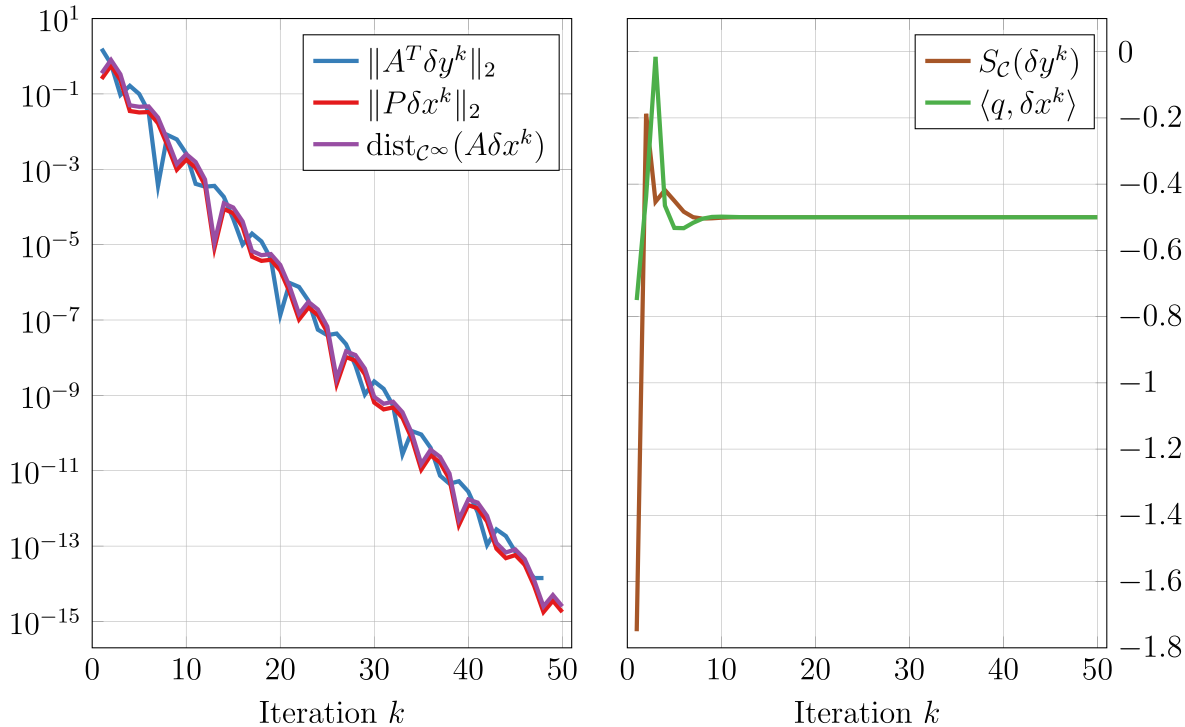

Figure 4: Convergence of $\{\delta y^k\}_{k\in\mathbf{N}}$ and $\{\delta x^k\}_{k\in\mathbf{N}}$ to certificates of primal and dual infeasibility, respectively, for problem (26) with $a=0$, $u_1=0$ and $u_3=+\infty$.

5.2 Infeasible SDPs from SDPLIB

We next demonstrate the asymptotic behavior of Algorithm 1 on two infeasible SDPs from the benchmark library SDPLIB [Bor99]. The problems are given in the following form

$$\begin{array}{ll}\underset{x,z}{\text{minimize}} & q^T x\\ \text{subject to} & Ax = z\\ & z \in \mathcal{S}_b^m,\end{array}$$

where $\mathcal{S}^m$ denotes the vectorized form of $\mathbf{S}_+^m$, *i.e.* $z \in \mathcal{S}^m$ is equivalent to $\mathrm{mat}(z) \in \mathbf{S}_+^m$, and $\mathcal{S}_b^m := \mathcal{S}^m + \{b\}$.

Let $X \in \mathbf{S}^m$ have the following eigenvalue decomposition

$$X = U\,\mathrm{diag}(\lambda_1, \ldots, \lambda_m) U^T.$$

Then the projection of X onto $\mathbf{S}_+^m$ is

$$\Pi_{\mathbf{S}_+^m}(X) = U\,\mathrm{diag}\left(\max(\lambda_1, 0), \ldots, \max(\lambda_m, 0)\right) U^T.$$

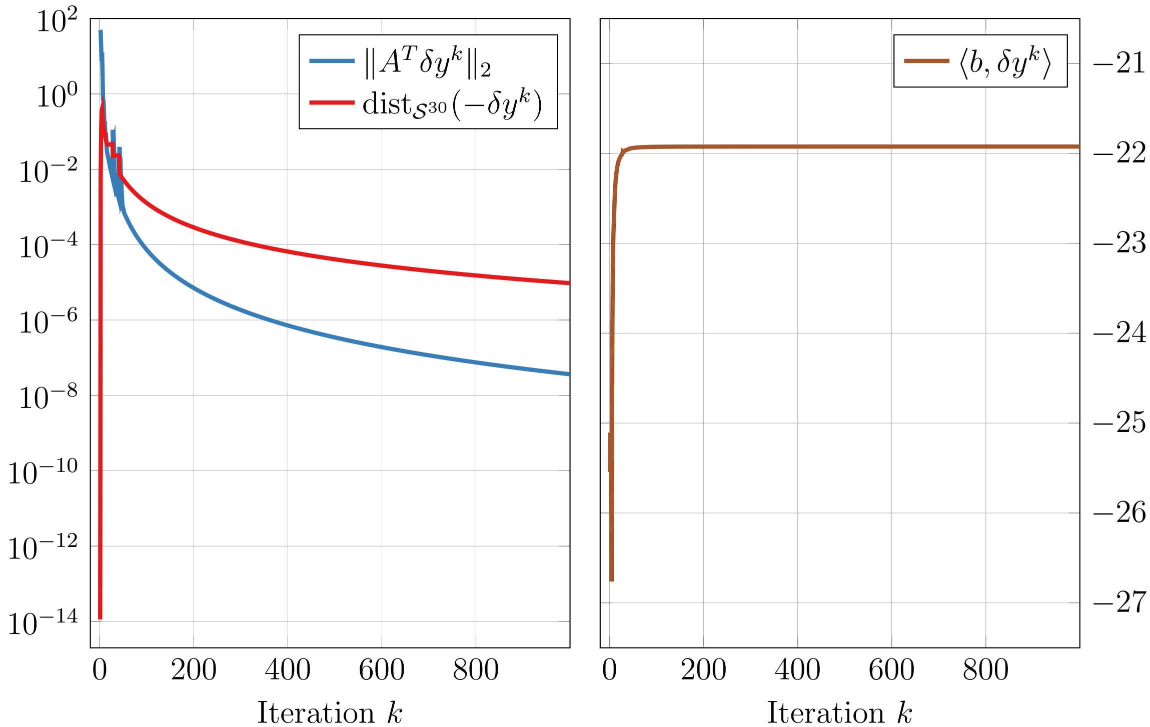

Figure 5: Convergence of $\{\delta y^k\}_{k\in\mathbf{N}}$ to a certificate of primal infeasibility for problem `infp1` from SDPLIB.

Primal infeasible SDPs. The primal infeasible problem `infp1` from SDPLIB has decision variables $x \in \mathbf{R}^{10}$ and $z \in \mathcal{S}^{30}$. We run Algorithm 1 with parameters $\alpha = 1$ and $\rho = \sigma = 0.1$ from the initial iterate $(x^0, z^0, y^0) = (0, 0, 0)$. Figure 5 shows convergence of $\{\delta y^k\}_{k\in\mathbf{N}}$ to a certificate of primal infeasibility, where $\mathrm{dist}_{\mathcal{S}^m}(y)$ denotes the spectral norm distance of $\mathrm{mat}(y)$ to the positive semidefinite cone $\mathbf{S}^m_+$.

Dual infeasible SDPs. Dual infeasible problem `infd1` from SDPLIB has decision variables $x \in \mathbf{R}^{10}$ and $z \in \mathcal{S}^{30}$. We run Algorithm 1 with parameters $\alpha = 1$ and $\rho = \sigma = 0.001$ from the initial iterate $(x^0, z^0, y^0) = (0, 0, 0)$. Figure 6 shows convergence of $\{\delta x^k\}_{k\in\mathbf{N}}$ to a certificate of dual infeasibility.

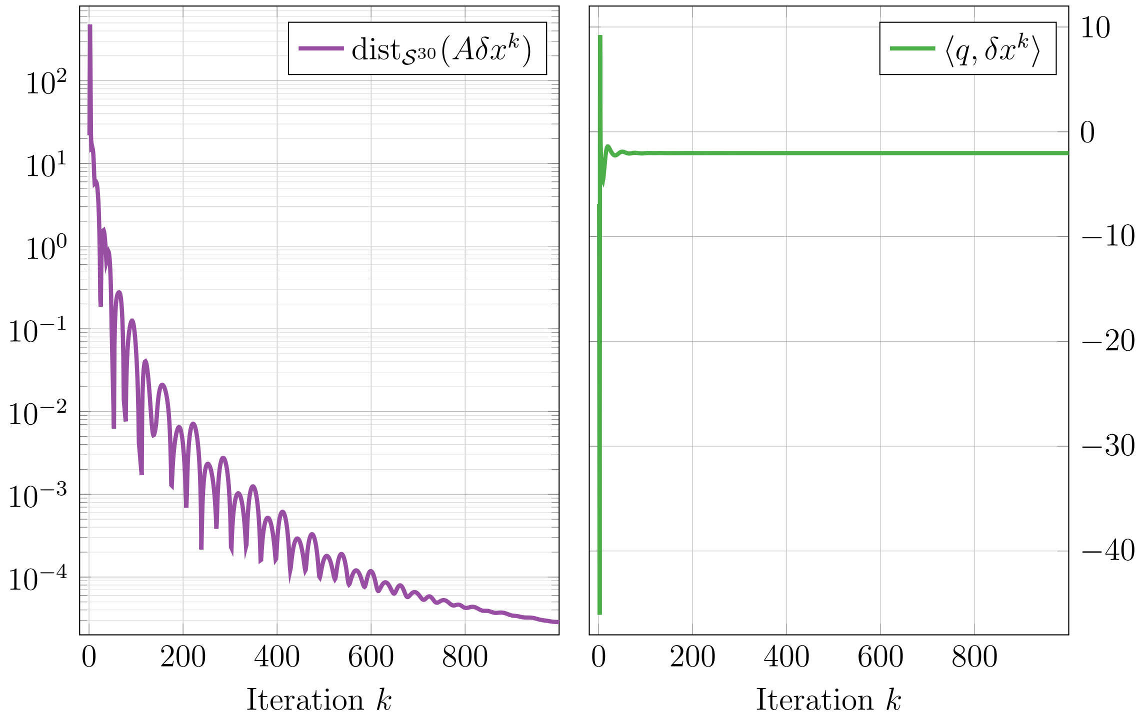

Figure 6: Convergence of $\{\delta x^k\}_{k\in\mathbf{N}}$ to a certificate of dual infeasibility for problem `infd1` from SDPLIB.

5.3 Infeasible SDPs with no certificate

Consider the following feasibility problem [Ram97, Ex. 5]

$$\begin{array}{ll} \underset{x_1,x_2}{\text{minimize}} & 0 \\ \text{subject to} & \begin{bmatrix} x_1 & 1 & 0 \\ 1 & x_2 & 0 \\ 0 & 0 & -x_1 \end{bmatrix} \succeq 0, \end{array} \tag{28}$$

noting that it is primal infeasible by inspection. If we write the constraint set in (28) as

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}}_{A_1} x_1 + \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{A_2} x_2 + \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{A_0} \succeq 0$$

and denote by $A = [\mathrm{vec}(A_1)\ \ \mathrm{vec}(A_2)]$ and $b = -\mathrm{vec}(A_0)$, then the constraint can be written as $Ax \in \mathcal{S}^3_b$, where $\mathcal{S}^3$ denotes the vectorized form of $\mathbf{S}^3_+$. If we define $Y := \mathrm{mat}(y)$, then the primal infeasibility conditions (6) for the above problem amount to

$$Y_{11} - Y_{33} = 0, \quad Y_{22} = 0, \quad Y_{12} < 0, \quad Y \preceq 0,$$

where Y_{ij} denotes the element of $Y \in \mathbf{S}^3$ in the i-th row and j-th column. Given that $Y \preceq 0$ and $Y_{22} = 0$ imply $Y_{12} = 0$, the above system is infeasible as well. Note that $Y = 0$ is a feasible point for the dual of problem (28) and problem (28) is thus not dual infeasible.

We next show that $(\delta x^k, \delta Z^k, \delta Y^k) \to 0$, where $\delta Z^k := \operatorname{mat}(\delta z^k)$ and $\delta Y^k := \operatorname{mat}(\delta y^k)$. Set $x^k = ((1 + \rho\sigma^{-1})\varepsilon, \varepsilon^{-1})$ and $V^k := \operatorname{mat}(v^k) = \operatorname{diag}(\varepsilon, \varepsilon^{-1}, 0)$ where $\varepsilon > 0$. The iteration (11) then produces the following iterates

$$Z^k = V^k, \quad \tilde{x}^k = (\varepsilon, \varepsilon^{-1}), \quad \tilde{Z}^k = \operatorname{diag}(\varepsilon, \varepsilon^{-1}, -\varepsilon),$$

and therefore we have

$$\begin{aligned} \delta x^{k+1} &= \alpha\,(\tilde{x}^k - x^k) = \alpha\,(-\rho\sigma^{-1}\varepsilon, 0), \\ \delta V^{k+1} &= \alpha\,(\tilde{Z}^k - Z^k) = \alpha \operatorname{diag}(0, 0, -\varepsilon). \end{aligned}$$

By taking ε arbitrarily small, we can make $(\delta x^{k+1}, \delta V^{k+1})$ arbitrarily close to zero, which according to Lemma 1 means that $(\delta x^k, \delta V^k) \to (\delta x, \delta V) = 0$, and according to Proposition 4 the optimality conditions (4) are satisfied in the limit. However, the sequence $\{x^k, Z^k, Y^k\}_{k\in\mathbf{N}}$ does not have a limit point; otherwise, such a point would be a certificate for optimality of the problem. Let T denote the fixed-point operator mapping (x^k, V^k) to (x^{k+1}, V^{k+1}). Since $(\delta x, \delta V) \in \overline{\operatorname{ran}(T - \operatorname{Id})}$ by definition, and $(\delta x, \delta V) \notin \operatorname{ran}(T - \operatorname{Id})$, this means that the set $\operatorname{ran}(T - \operatorname{Id})$ is not closed, and the distance from $(\delta x, \delta V)$ to $\operatorname{ran}(T - \operatorname{Id})$ is zero. In other words, the set of matrices in (28) and the semidefinite cone $\mathbf{S}^3_+$ do not intersect, but are not strongly separable.

We run Algorithm 1 with parameters $\alpha = \rho = \sigma = 1$ from the initial iterate $(x^0, Z^0, Y^0) = (0, 0, 0)$. Figure 7 shows convergence of residuals $\|Ax^k - z^k\|_2$ and $\|A^T y^k\|_2$ to zero.

Remark 3. Let $\varepsilon > 0$. Consider the following perturbation to problem (28):

$$\begin{array}{ll} \text{minimize} & 0 \\ \text{subject to} & \begin{bmatrix} x_1 & 1 & 0 \\ 1 & x_2 & 0 \\ 0 & 0 & -x_1 \end{bmatrix} \succeq -\varepsilon I. \end{array}$$

This problem is feasible since the constraint above is satisfied for $x_1 = 0$ and $x_2 = 1/\varepsilon - \varepsilon$.

Consider now the following problem:

$$\begin{array}{ll} \text{minimize} & 0 \\ \text{subject to} & \begin{bmatrix} x_1 & 1 & 0 \\ 1 & x_2 & 0 \\ 0 & 0 & -x_1 \end{bmatrix} \succeq \varepsilon I. \end{array}$$

This problem is strongly infeasible since the vector $\bar{y} = \operatorname{vec}(\operatorname{diag}(-1, 0, -1))$ satisfies primal infeasibility conditions (6).

These two examples show that an infinitesimally small perturbation to problem (28) can make the problem feasible or strongly infeasible.

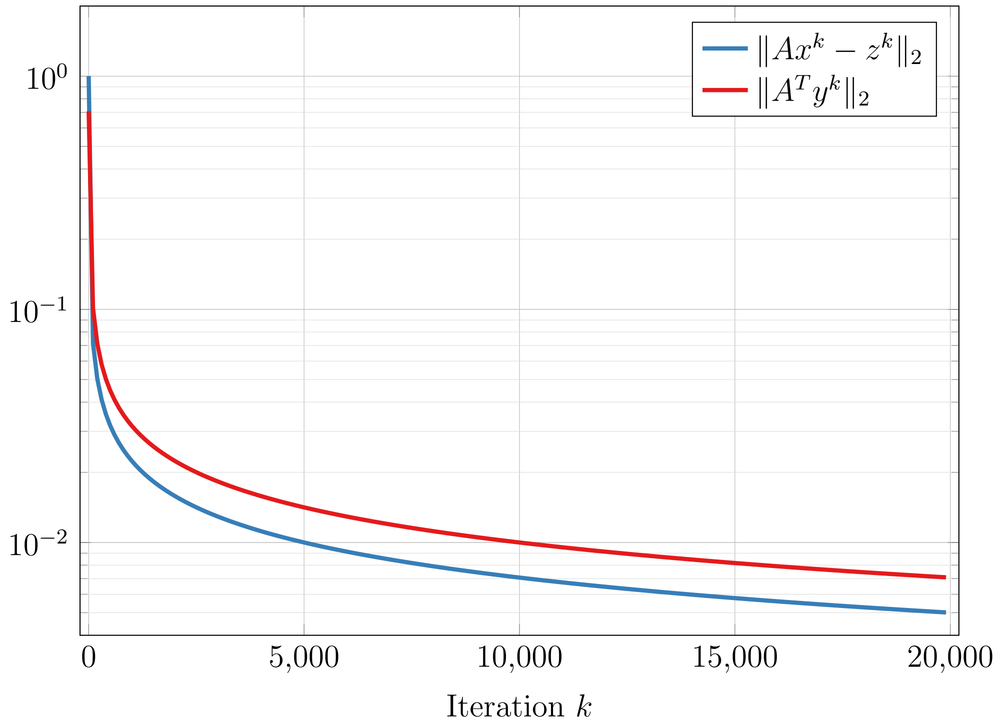

Figure 7: Convergence of residuals $\|Ax^k - z^k\|_2$ and $\|A^T y^k\|_2$ for problem (28).

6 Conclusions

We have analyzed the asymptotic behavior of ADMM for a class of convex optimization problems, and have shown that if the problem is primal and/or dual strongly infeasible, then the sequence of successive differences of the algorithm's iterates converge to a certificate of infeasibility. Based on these results, we have proposed termination criteria for detecting primal and dual infeasibility, providing for the first time a set of reliable and generic stopping criteria for ADMM applicable to infeasible convex problems. We have also provided numerical examples to demonstrate different asymptotic behaviors of the algorithm's iterates.

Acknowledgement

We are grateful to Walaa Moursi for helpful comments and pointing out additional references.

A Supporting results

Lemma A.1. The first-order optimality conditions for problem (3) are the conditions (4).

Proof. The linear constraint in problem (3) can be relaxed by using the following Lagrangian subproblem:

$$
\begin{array}{ll}
\underset{x,y,z}{\text{minimize}} & \frac{1}{2}x^T P x + q^T x + y^T(Ax - z) \\
\text{subject to} & z \in \mathcal{C}.
\end{array}
$$

If we denote the objective function in the above problem by $F(x, y, z)$, then the optimality conditions can be written as [RW98, Thm. 6.12]

$$
\begin{aligned}
&z \in \mathcal{C}, \\
&0 = \nabla_x F(x, y, z) = Px + q + A^T y, \\
&0 = \nabla_y F(x, y, z) = Ax - z, \\
&0 \geq \sup_{z' \in \mathcal{C}} \langle -\nabla_z F(x, y, z), z' - z \rangle = \sup_{z' \in \mathcal{C}} \langle y, z' - z \rangle,
\end{aligned}
$$

where the last condition is equivalent to $y \in N_{\mathcal{C}}(z)$. □

Lemma A.2. The dual of problem (2) is given by problem (5).

Proof. We first rewrite problem (2) in the form

$$
\begin{array}{ll}
\underset{x,z}{\text{minimize}} & \frac{1}{2}x^T P x + q^T x + \mathcal{I}_{\mathcal{C}}(z) \\
\text{subject to} & Ax = z,
\end{array}
$$

then form its Lagrangian,

$$
L(x, z, y) := \tfrac{1}{2}x^T P x + q^T x + \mathcal{I}_{\mathcal{C}}(z) + y^T(Ax - z),
$$

and finally derive the dual function as follows:

$$
\begin{aligned}
g(y) &:= \inf_{x,z} L(x, z, y) \\
&= \inf_x \{\tfrac{1}{2}x^T P x + (A^T y + q)^T x\} + \inf_{z \in \mathcal{C}} \{-y^T z\} \\
&= \inf_x \{\tfrac{1}{2}x^T P x + (A^T y + q)^T x\} - \sup_{z \in \mathcal{C}} \{y^T z\}.
\end{aligned}
$$

Note that the minimum of the Lagrangian over x is obtained when $Px + A^T y + q = 0$, and the second term in the last line is the support function of $\mathcal{C}$. The dual problem, defined as the problem of maximizing the dual function, can then be written in the form (5), where the conic constraint on y is just the restriction of y to the domain of $S_{\mathcal{C}}$ [Roc70, p.112 and Cor. 14.2.1]. □

Lemma A.3. For any vectors $v \in \mathbf{R}^n$, $b \in \mathbf{R}^n$ and a nonempty, closed and convex cone $\mathcal{K} \subseteq \mathbf{R}^n$,

(i) $\Pi_{\mathcal{K}_b}(v) = b + \Pi_{\mathcal{K}}(v-b)$.
(ii) $(\operatorname{Id} - \Pi_{\mathcal{K}_b})(v) = \Pi_{\mathcal{K}^\circ}(v-b)$.
(iii) $\langle \Pi_{\mathcal{K}_b}(v), (\operatorname{Id} - \Pi_{\mathcal{K}_b})(v)\rangle = \langle b, \Pi_{\mathcal{K}^\circ}(v-b)\rangle$.
(iv) $\langle \Pi_{\mathcal{K}}(v), v\rangle = \|\Pi_{\mathcal{K}}(v)\|^2$.

Proof. Part (i) is from [BC11, Prop. 28.1(i)].

(ii): From part (i) we have

$$(\operatorname{Id} - \Pi_{\mathcal{K}_b})(v) = v - b - \Pi_{\mathcal{K}}(v-b) = \Pi_{\mathcal{K}^\circ}(v-b),$$

where the second equality follows from the Moreau decomposition [BC11, Thm. 6.29].

(iii): Follows directly from parts (i) and (ii), and the Moreau decomposition.

(iv): From the Moreau decomposition, we have

$$\langle \Pi_{\mathcal{K}}(v), v\rangle = \langle \Pi_{\mathcal{K}}(v), \Pi_{\mathcal{K}}(v) + \Pi_{\mathcal{K}^\circ}(v)\rangle = \|\Pi_{\mathcal{K}}(v)\|^2. \qquad \square$$

Lemma A.4. Suppose that $\mathcal{K}$ is a nonempty, closed and convex cone and for some sequence $\{v^k\}_{k\in\mathbf{N}}$, where $v^k \in \mathbf{R}^n$, we denote by $\delta v := \lim_{k\to\infty} \frac{1}{k} v^k$, assuming that the limit exists. Then for any $b \in \mathbf{R}^n$,

$$\lim_{k\to\infty} \tfrac{1}{k}\Pi_{\mathcal{K}_b}(v^k) = \Pi_{\mathcal{K}}(\delta v).$$

Proof. Write the limit as

$$\begin{aligned}
\lim_{k\to\infty} \tfrac{1}{k}\Pi_{\mathcal{K}_b}(v^k) &= \lim_{k\to\infty} \tfrac{1}{k}\left(b + \Pi_{\mathcal{K}}(v^k - b)\right) \\
&= \lim_{k\to\infty} \Pi_{\mathcal{K}}\left(\tfrac{1}{k}(v^k - b)\right) \\
&= \Pi_{\mathcal{K}}\left(\lim_{k\to\infty} \tfrac{1}{k} v^k\right),
\end{aligned}$$

where the first equality uses Lemma A.3(i), and the second and third follow from the positive homogeneity [BC11, Prop. 28.22] and continuity [BC11, Prop. 4.8] of $\Pi_{\mathcal{K}}$, respectively. $\square$

Lemma A.5. Suppose that $\mathcal{B} \subseteq \mathbf{R}^n$ is a nonempty, convex and compact set and for some sequence $\{v^k\}_{k\in\mathbf{N}}$, where $v^k \in \mathbf{R}^n$, we denote by $\delta v := \lim_{k\to\infty} \frac{1}{k} v^k$, assuming that the limit exists. Then

$$\lim_{k\to\infty} \tfrac{1}{k}\left\langle v^k, \Pi_{\mathcal{B}}(v^k)\right\rangle = \lim_{k\to\infty} \left\langle \delta v, \Pi_{\mathcal{B}}(v^k)\right\rangle = S_{\mathcal{B}}(\delta v).$$

Proof. Let $z^k := \Pi_{\mathcal{B}}(v^k)$. We have the following inclusion [BC11, Prop. 6.46]

$$v^k - z^k \in N_{\mathcal{B}}(z^k),$$

which is equivalent to [BC11, Thm. 16.23]

$$\left\langle \tfrac{1}{k}(v^k - z^k), z^k \right\rangle = S_{\mathcal{B}}\left(\tfrac{1}{k}(v^k - z^k)\right).$$

Taking the limit of the above identity, we obtain

$$\lim_{k\to\infty} \left\langle \tfrac{1}{k}(v^k - z^k), z^k \right\rangle = \lim_{k\to\infty} S_{\mathcal{B}}\left(\tfrac{1}{k}(v^k - z^k)\right) = S_{\mathcal{B}}\Big(\lim_{k\to\infty} \tfrac{1}{k}(v^k - z^k)\Big) = S_{\mathcal{B}}(\delta v), \tag{29}$$

where the second equality follows from the continuity of $S_{\mathcal{B}}$ [BC11, Ex. 11.2], and the third from the compactness of $\mathcal{B}$. Since $\{z^k\}_{k\in\mathbf{N}}$ remains in the compact set $\mathcal{B}$, we can derive the following relation from (29):

$$\begin{aligned}
\left| S_{\mathcal{B}}(\delta v) - \lim_{k\to\infty} \left\langle \delta v, z^k \right\rangle \right| &= \left| \lim_{k\to\infty} \left\langle \tfrac{1}{k}(v^k - z^k), z^k \right\rangle - \left\langle \delta v, z^k \right\rangle \right| \\
&= \left| \lim_{k\to\infty} \left\langle \tfrac{1}{k}v^k - \delta v, z^k \right\rangle - \tfrac{1}{k}\left\langle z^k, z^k \right\rangle \right| \\
&\le \lim_{k\to\infty} \underbrace{\|\tfrac{1}{k}v^k - \delta v\|}_{\to 0} \|z^k\| + \tfrac{1}{k}\|z^k\|^2 \\
&= 0,
\end{aligned}$$

where the third row follows from the triangle and Cauchy-Schwarz inequalities, and the fourth from the compactness of $\mathcal{B}$. Finally, we can derive the following identity from (29):

$$S_{\mathcal{B}}(\delta v) = \lim_{k\to\infty} \left\langle \tfrac{1}{k}(v^k - z^k), z^k \right\rangle = \lim_{k\to\infty} \left\langle \tfrac{1}{k}v^k, z^k \right\rangle - \underbrace{\tfrac{1}{k}\|z^k\|^2}_{\to 0}.$$

This concludes the proof. $\square$

References

[BB96] H. H. Bauschke and J. M. Borwein. On projection algorithms for solving convex feasibility problems. *SIAM Review*, 38(3):367–426, 1996.

[BBR78] J. B. Baillon, R. E. Bruck, and S. Reich. On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces. *Houston Journal of Mathematics*, 4(1):1–9, 1978.

[BC11] H. H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces.* Springer Science & Business Media, 2011.

[BCL04] H. H. Bauschke, P. L. Combettes, and D. R. Luke. Finding best approximation pairs relative to two closed convex sets in Hilbert spaces. *Journal of Approximation Theory*, 127(2):178–192, 2004.

[BDM16] H. H. Bauschke, M. N. Dao, and W. M. Moursi. The Douglas-Rachford algorithm in the affine-convex case. *Operations Research Letters*, 44(3):379–382, 2016.

[BG18] G. Banjac and P. Goulart. Tight global linear convergence rate bounds for operator splitting methods. *IEEE Transactions on Automatic Control (To appear)*, 2018.

[BM16] H. H. Bauschke and W. M. Moursi. The Douglas-Rachford algorithm for two (not necessarily intersecting) affine subspaces. *SIAM Journal on Optimization*, 26(2):968–985, 2016.

[BM17] H. H. Bauschke and W. M. Moursi. On the Douglas-Rachford algorithm. *Mathematical Programming*, 164(1):263–284, 2017.

[Bol13] D. Boley. Local linear convergence of the alternating direction method of multipliers on quadratic or linear programs. *SIAM Journal on Optimization*, 23(4):2183–2207, 2013.

[Bor99] B. Borchers. SDPLIB 1.2, a library of semidefinite programming test problems. *Optimization Methods and Software*, 11(1):683–690, 1999.

[BPC+11] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1):1–122, 2011.

[BSM+17] G. Banjac, B. Stellato, N. Moehle, P. Goulart, A. Bemporad, and S. Boyd. Embedded code generation using the OSQP solver. In *IEEE Conference on Decision and Control (CDC)*, pages 1906–1911, 2017.

[BT09] A. Beck and M. Teboulle. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Transactions on Image Processing*, 18(11):2419–2434, 2009.

[BV04] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[CP11] P. L. Combettes and J.-C. Pesquet. Proximal splitting methods in signal processing. *Springer Optimization and Its Applications*, 49:185–212, 2011.

[CW05] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Modeling & Simulation*, 4(4):1168–1200, 2005.

[EB92] J. Eckstein and D. P. Bertsekas. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming*, 55(1):293–318, 1992.

[Gab83] D. Gabay. Applications of the method of multipliers to variational inequalities. *Studies in Mathematics and its Applications*, 15(C):299–331, 1983.

[GB17] P. Giselsson and S. Boyd. Linear convergence and metric selection for Douglas-Rachford splitting and ADMM. *IEEE Transactions on Automatic Control*, 62(2):532–544, 2017.

[GFB16] P. Giselsson, M. Fält, and S. Boyd. Line search for averaged operator iteration. In *IEEE Conference on Decision and Control (CDC)*, pages 1015–1022, 2016.

[GTSJ15] E. Ghadimi, A. Teixeira, I. Shames, and M. Johansson. Optimal parameter selection for the alternating direction method of multipliers (ADMM): Quadratic problems. *IEEE Transactions on Automatic Control*, 60(3):644–658, 2015.

[HM12] D. Henrion and J. Malick. *Projection Methods in Conic Optimization*, pages 565–600. Springer US, 2012.

[JGR$^+$14] J. Jerez, P. Goulart, S. Richter, G. Constantinides, E. Kerrigan, and M. Morari. Embedded online optimization for model predictive control at megahertz rates. *IEEE Transactions on Automatic Control*, 59(12):3238–3251, 2014.

[LMT16] B. F. Lourenço, M. Muramatsu, and T. Tsuchiya. Weak infeasibility in second order cone programming. *Optimization Letters*, 10(8):1743–1755, 2016.

[Mou16] W. M. Moursi. *The Douglas-Rachford operator in the possibly inconsistent case: static properties and dynamic behaviour*. PhD thesis, University of British Columbia, 2016.

[NB17] V. V. Naik and A. Bemporad. Embedded mixed-integer quadratic optimization using accelerated dual gradient projection. In *IFAC World Congress*, pages 10723–10728, 2017.

[OCPB16] B. O'Donoghue, E. Chu, N. Parikh, and S. Boyd. Conic optimization via operator splitting and homogeneous self-dual embedding. *Journal of Optimization Theory and Applications*, 169(3):1042–1068, 2016.

[OSB13] B. O'Donoghue, G. Stathopoulos, and S. Boyd. A splitting method for optimal control. *IEEE Transactions on Control Systems Technology*, 21(6):2432–2442, 2013.

[Paz71] A. Pazy. Asymptotic behavior of contractions in Hilbert space. *Israel Journal of Mathematics*, 9(2):235–240, 1971.

[PB13] N. Parikh and S. Boyd. Proximal algorithms. *Foundations and Trends in Optimization*, 1(3):123–231, 2013.

[Ram97] M. V. Ramana. An exact duality theory for semidefinite programming and its complexity implications. *Mathematical Programming*, 77(1):129–162, 1997.

[RDC14] A. U. Raghunathan and S. Di Cairano. Infeasibility detection in alternating direction method of multipliers for convex quadratic programs. In *IEEE Conference on Decision and Control (CDC)*, pages 5819–5824, 2014.

[Roc70] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, USA, 1970.

[RW98] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*. Grundlehren der mathematischen Wissenschaften. Springer, 1998.

[SBG+18] B. Stellato, G. Banjac, P. Goulart, A. Bemporad, and S. Boyd. OSQP: An operator splitting solver for quadratic programs. *arXiv:1711.08013*, 2018.

[SSS+16] G. Stathopoulos, H. Shukla, A. Szucs, Y. Pu, and C. Jones. Operator splitting methods in control. *Foundations and Trends in Systems and Control*, 3(3):249–362, 2016.

[Toh08] K.-C. Toh. An inexact primal-dual path following algorithm for convex quadratic SDP. *Mathematical Programming*, 112(1):221–254, 2008.

[ZFP+17] Y. Zheng, G. Fantuzzi, A. Papachristodoulou, P. Goulart, and A. Wynn. Fast ADMM for semidefinite programs with chordal sparsity. In *American Control Conference (ACC)*, pages 3335–3340, 2017.