Optimization Letters (2021) 15:2719–2732
https://doi.org/10.1007/s11590-021-01706-3

ORIGINAL PAPER

On the asymptotic behavior of the Douglas–Rachford and proximal-point algorithms for convex optimization

Goran Banjac[1] · **John Lygeros**[1]

Received: 29 April 2020 / Accepted: 15 January 2021 / Published online: 4 February 2021

Abstract
Banjac et al. (J Optim Theory Appl 183(2):490–519, 2019) recently showed that the Douglas–Rachford algorithm provides certificates of infeasibility for a class of convex optimization problems. In particular, they showed that the difference between consecutive iterates generated by the algorithm converges to certificates of primal and dual strong infeasibility. Their result was shown in a finite-dimensional Euclidean setting and for a particular structure of the constraint set. In this paper, we extend the result to real Hilbert spaces and a general nonempty closed convex set. Moreover, we show that the proximal-point algorithm applied to the set of optimality conditions of the problem generates similar infeasibility certificates.

Keywords Douglas–Rachford algorithm · Proximal-point algorithm · Convex optimization · Infeasibility detection

Mathematics Subject Classification 49M27 · 65K10 · 90C25

1 Introduction

Due to its very good practical performance and ability to handle nonsmooth functions, the Douglas–Rachford algorithm has attracted a lot of interest for solving convex optimization problems. Provided that a problem is solvable and satisfies certain constraint qualification, the algorithm converges to an optimal solution [1, Cor. 27.3]. If the problem is infeasible, then some of its iterates diverge [2].

Results on the asymptotic behavior of the Douglas–Rachford algorithm for infeasible problems are very scarce, and most of them study some specific cases such as

✉ Goran Banjac
gbanjac@ethz.ch

John Lygeros
jlygeros@ethz.ch

1 Automatic Control Laboratory, ETH Zurich, Zurich, Switzerland

feasibility problems involving two convex sets that do not intersect [3–5]. Although there have been some recent results studying a more general setting [6,7], they impose some additional assumptions on feasibility of either the primal or the dual problem. The authors in [8] consider a problem of minimizing a convex quadratic function over a particular constraint set, and show that the iterates of the Douglas–Rachford algorithm generate an infeasibility certificate when the problem is primal and/or dual strongly infeasible. A similar analysis was applied in [9] to show that the proximal-point algorithm used for solving a convex quadratic program can also detect infeasibility.

The constraint set of the problem studied in [8] is represented in the form $Ax \in C$, where A is a real matrix and C the Cartesian product of a convex compact set and a translated closed convex cone. This paper extends the result of [8] to real Hilbert spaces and a general nonempty closed convex set C. Moreover, we show that a similar analysis can be used to prove that the proximal-point algorithm for solving the same class of problems generates similar infeasibility certificates.

The paper is organized as follows. We introduce some definitions and notation in the remainder of Sect. 1, and the problem under consideration in Sect. 2. Section 3 presents some supporting results that are essential for generalizing the results in [8]. Finally, Sects. 4 and 5 analyze the asymptotic behavior of the Douglas–Rachford and proximal-point algorithms, respectively, and show that they provide infeasibility certificates for the considered problem.

1.1 Notation

Let $\mathcal{H}, \mathcal{H}_1, \mathcal{H}_2$ be real Hilbert spaces with inner products $\langle \cdot \mid \cdot \rangle$, induced norms $\|\cdot\|$, and identity operators Id. The power set of $\mathcal{H}$ is denoted by $2^{\mathcal{H}}$. Let $\mathbb{N}$ denote the set of positive integers. For a sequence $(s_n)_{n\in\mathbb{N}}$, we denote by $s_n \to s$ ($s_n \rightharpoonup s$) that it converges strongly (weakly) to s and define $\delta s_{n+1} := s_{n+1} - s_n$.

Let D be a nonempty subset of $\mathcal{H}$ with $\overline{D}$ being its *closure*. Then $T\colon D \to \mathcal{H}$ is *nonexpansive* if

$$(\forall x \in D)(\forall y \in D) \quad \|Tx - Ty\| \leq \|x - y\|,$$

and it is α-*averaged* with $\alpha \in]0, 1[$ if there exists a nonexpansive operator $R\colon D \to \mathcal{H}$ such that $T = (1-\alpha)\operatorname{Id} + \alpha R$. We denote the range of T by $\operatorname{ran} T$. A set-valued operator $B\colon \mathcal{H} \to 2^{\mathcal{H}}$, characterized by its *graph*

$$\operatorname{gra} B = \{(x, u) \in \mathcal{H} \times \mathcal{H} \mid u \in Bx\},$$

is *monotone* if

$$(\forall (x, u) \in \operatorname{gra} B)\ (\forall (y, v) \in \operatorname{gra} B) \quad \langle x - y \mid u - v \rangle \geq 0.$$

The *inverse* of B, denoted by B^{-1}, is defined through its graph

$$\operatorname{gra} B^{-1} = \{(u, x) \in \mathcal{H} \times \mathcal{H} \mid (x, u) \in \operatorname{gra} B\}.$$

For a proper lower semicontinuous convex function $f : \mathcal{H} \to]-\infty, +\infty]$, we define its:

$$\begin{aligned}
&\textit{Fenchel conjugate}: && f^* : \mathcal{H} \to]-\infty, +\infty] : u \mapsto \sup_{x \in \mathcal{H}} \left(\langle x \mid u \rangle - f(x)\right),\\
&\textit{proximity operator}: && \operatorname{Prox}_f : \mathcal{H} \to \mathcal{H} : x \mapsto \operatorname*{argmin}_{y \in \mathcal{H}} \left(f(y) + \tfrac{1}{2}\|y - x\|^2\right),\\
&\textit{subdifferential}: && \partial f : \mathcal{H} \to 2^{\mathcal{H}}\\
& && \quad : x \mapsto \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H})\ \langle y - x \mid u \rangle + f(x) \le f(y)\}.
\end{aligned}$$

For a nonempty closed convex set $C \subseteq \mathcal{H}$, we define its:

$$\begin{aligned}
&\textit{polar cone}: && C^{\ominus} = \left\{u \in \mathcal{H} \mid \sup_{x \in C} \langle x \mid u \rangle \le 0\right\},\\
&\textit{recession cone}: && \operatorname{rec} C = \{x \in \mathcal{H} \mid (\forall y \in C)\ x + y \in C\},\\
&\textit{indicator function}: && \iota_C : \mathcal{H} \to [0, +\infty] : x \mapsto \begin{cases} 0 & x \in C \\ +\infty & \text{otherwise}, \end{cases}\\
&\textit{support function}: && \sigma_C : \mathcal{H} \to]-\infty, +\infty] : u \mapsto \sup_{x \in C} \langle x \mid u \rangle,\\
&\textit{projection operator}: && P_C : \mathcal{H} \to \mathcal{H} : x \mapsto \operatorname*{argmin}_{y \in C} \|y - x\|,\\
&\textit{normal cone operator}: && N_C : \mathcal{H} \to 2^{\mathcal{H}}\\
& && \quad : x \mapsto \begin{cases} \left\{u \in \mathcal{H} \mid \sup_{y \in C} \langle y - x \mid u \rangle \le 0\right\} & x \in C \\ \emptyset & x \notin C. \end{cases}
\end{aligned}$$

2 Problem of interest

Consider the following convex optimization problem:

$$\begin{aligned}
&\underset{x \in \mathcal{H}_1}{\text{minimize}}\ \tfrac{1}{2}\langle Qx \mid x \rangle + \langle q \mid x \rangle\\
&\text{subject to } Ax \in C,
\end{aligned} \tag{1}$$

with $Q : \mathcal{H}_1 \to \mathcal{H}_1$ a monotone self-adjoint bounded linear operator, $q \in \mathcal{H}_1$, $A : \mathcal{H}_1 \to \mathcal{H}_2$ a bounded linear operator, and C a nonempty closed convex subset of $\mathcal{H}_2$; we assume that $\operatorname{ran} Q$ and $\operatorname{ran} A$ are closed. The objective function of the problem is convex, continuous, and Fréchet differentiable [1, Prop. 17.36].

When $\mathcal{H}_1$ and $\mathcal{H}_2$ are finite-dimensional Euclidean spaces, problem (1) reduces to the one considered in [8], where the Douglas–Rachford algorithm (which is equivalent to the alternating direction method of multipliers) was shown to generate certificates of primal and dual strong infeasibility. Moreover, the authors proposed termination criteria for infeasibility detection, which are easy to implement and are used in several numerical solvers; see, e.g., [10–12]. To prove the main results, they used the assumption that C can be represented as the Cartesian product of a convex compact set and a translated closed convex cone, which was exploited heavily in their proofs. In this paper we extend these results to the case where $\mathcal{H}_1$ and $\mathcal{H}_2$ are real Hilbert spaces, and C is a general nonempty closed convex set.

2.1 Optimality conditions

We can rewrite problem (1) in the form

$$\underset{x\in\mathcal{H}_1}{\text{minimize}} \quad \tfrac{1}{2}\langle Qx \mid x\rangle + \langle q \mid x\rangle + \iota_C(Ax).$$

Provided that a certain constraint qualification holds, we can characterize its solution by [1, Thm. 27.2]

$$0 \in Qx + q + A^*\partial\iota_C(Ax),$$

and introducing a dual variable $y \in \partial\iota_C(Ax)$, we can rewrite the inclusion as

$$0 \in \begin{pmatrix} Qx + q + A^*y \\ -y + \partial\iota_C(Ax) \end{pmatrix}. \tag{2}$$

Introducing an auxiliary variable $z \in C$ and using $\partial\iota_C = N_C$, we can write the optimality conditions for problem (1) as

$$Ax - z = 0 \tag{3a}$$
$$Qx + q + A^*y = 0 \tag{3b}$$
$$z \in C, \quad y \in N_C z. \tag{3c}$$

2.2 Infeasibility certificates

The authors in [8] derived the following conditions for characterizing strong infeasibility of problem (1) and its dual:

Proposition 2.1 ([8, Prop. 3.1])

(i) *If there exists a* $\bar{y} \in \mathcal{H}_2$ *such that*

$$A^*\bar{y} = 0 \quad \textit{and} \quad \sigma_C(\bar{y}) < 0,$$

then problem (1) is strongly infeasible.

(ii) *If there exists an* $\bar{x} \in \mathcal{H}_1$ *such that*

$$Q\bar{x} = 0, \quad A\bar{x} \in \operatorname{rec} C, \quad \textit{and} \quad \langle q \mid \bar{x}\rangle < 0,$$

then the dual of problem (1) is strongly infeasible.

3 Auxiliary results

Fact 3.1 *Suppose that* $T : \mathcal{H} \to \mathcal{H}$ *is an averaged operator and let* $s_0 \in \mathcal{H}$, $s_n = T^n s_0$, *and* $\delta s := P_{\overline{\operatorname{ran}}(T-\operatorname{Id})}(0)$. *Then*

(i) $\frac{1}{n}s_n \to \delta s$.
(ii) $\delta s_n \to \delta s$.

Proof The first result is [13, Cor. 3] and the second is [14, Cor. 2.3]. □

The following proposition provides essential ingredients for generalizing the results in [8, §5].

Proposition 3.2 *Let* $(s_n)_{n\in\mathbb{N}}$ *be a sequence in* $\mathcal{H}$ *satisfying* $\frac{1}{n}s_n \to \delta s$. *Let* $D \subseteq \mathcal{H}$ *be a nonempty closed convex set and define sequences* $(p_n)_{n\in\mathbb{N}}$ *and* $(r_n)_{n\in\mathbb{N}}$ *by*

$$\begin{aligned} p_n &:= P_D s_n \\ r_n &:= (\operatorname{Id} - P_D)s_n. \end{aligned}$$

Then

(i) $r_n \in (\operatorname{rec} D)^{\ominus}$.
(ii) $\frac{1}{n}p_n \to \delta p := P_{\operatorname{rec} D}(\delta s)$.
(iii) $\frac{1}{n}r_n \to \delta r := P_{(\operatorname{rec} D)^{\ominus}}(\delta s)$.
(iv) $\lim_{n\to\infty} \frac{1}{n}\langle p_n \mid r_n\rangle = \sigma_D(\delta r)$.

Proof (i): Follows from [15, Thm. 3.1].

(ii) and (iii): A related result was shown in [16, Lem. 6.3.13] and [17, Prop. 2.2] in a finite-dimensional setting. Using similar arguments here, together with those in [18, Lem. 4.3], we can only establish the weak convergence, i.e., $\frac{1}{n}p_n \rightharpoonup \delta p$. Using Moreau's decomposition [1, Thm. 6.30], it follows that $\frac{1}{n}r_n \rightharpoonup \delta r$ and $\|\delta s\|^2 = \|\delta p\|^2 + \|\delta r\|^2$. For an arbitrary vector $z \in D$, [1, Thm. 3.16] yields

$$\|s_n - z\|^2 \geq \|p_n - z\|^2 + \|r_n\|^2, \quad \forall n \in \mathbb{N}.$$

Dividing the inequality by n^2 and taking the limit superior, we get

$$\lim \|\tfrac{1}{n}s_n\|^2 \geq \overline{\lim}\,(\|\tfrac{1}{n}p_n\|^2 + \|\tfrac{1}{n}r_n\|^2) \geq \overline{\lim}\,\|\tfrac{1}{n}p_n\|^2 + \underline{\lim}\,\|\tfrac{1}{n}r_n\|^2,$$

and thus

$$\overline{\lim}\,\|\tfrac{1}{n}p_n\|^2 \leq \lim \|\tfrac{1}{n}s_n\|^2 - \underline{\lim}\,\|\tfrac{1}{n}r_n\|^2 \leq \|\delta s\|^2 - \|\delta r\|^2 = \|\delta p\|^2,$$

where the second inequality follows from [1, Lem. 2.42]. The inequality above yields $\overline{\lim}\,\|\frac{1}{n}p_n\| \leq \|\delta p\|$, which due to [1, Lem. 2.51] implies $\frac{1}{n}p_n \to \delta p$. Using Moreau's decomposition, it follows that $\frac{1}{n}r_n \to \delta r$.

(iv): Taking the limit of the inequality

$$(\forall n \in \mathbb{N})(\forall \hat{p} \in D) \quad \left\langle \hat{p} \mid \tfrac{1}{n} r_n \right\rangle \le \sup_{p \in D} \left\langle p \mid \tfrac{1}{n} r_n \right\rangle,$$

we obtain

$$(\forall \hat{p} \in D) \quad \lim_{n\to\infty} \left\langle \hat{p} \mid \tfrac{1}{n} r_n \right\rangle \le \lim_{n\to\infty} \sup_{p \in D} \left\langle p \mid \tfrac{1}{n} r_n \right\rangle,$$

and taking the supremum of the left-hand side over D, we get

$$\sup_{p \in D} \lim_{n\to\infty} \left\langle p \mid \tfrac{1}{n} r_n \right\rangle \le \lim_{n\to\infty} \sup_{p \in D} \left\langle p \mid \tfrac{1}{n} r_n \right\rangle. \tag{4}$$

From [1, Prop. 6.47], we have

$$r_n = s_n - p_n \in N_D p_n,$$

which, due to [1, Thm. 16.29] and the facts that $\iota_D^* = \sigma_D$ and $\partial \iota_D = N_D$, is equivalent to

$$\tfrac{1}{n} \langle p_n \mid r_n \rangle = \sigma_D \left(\tfrac{1}{n} r_n \right). \tag{5}$$

Taking the limit of (5) and using (4), we obtain

$$\lim_{n\to\infty} \tfrac{1}{n} \langle p_n \mid r_n \rangle = \lim_{n\to\infty} \sup_{p \in D} \left\langle p \mid \tfrac{1}{n} r_n \right\rangle \ge \sup_{p \in D} \lim_{n\to\infty} \left\langle p \mid \tfrac{1}{n} r_n \right\rangle = \sigma_D(\delta r).$$

Since $p_n \in D$, we also have

$$\lim_{n\to\infty} \tfrac{1}{n} \langle p_n \mid r_n \rangle \le \sup_{p \in D} \lim_{n\to\infty} \left\langle p \mid \tfrac{1}{n} r_n \right\rangle = \sigma_D(\delta r).$$

The result follows by combining the two inequalities above. □

The results of Prop. 3.2 are straightforward under the additional assumption that D is compact, since then $\operatorname{rec} D = \{0\}$ and $(\operatorname{rec} D)^{\ominus} = \mathcal{H}$, and thus

$$\lim_{n\to\infty} \tfrac{1}{n} p_n = \lim_{n\to\infty} \tfrac{1}{n} P_D s_n = 0 = P_{\operatorname{rec} D}(\delta s)$$
$$\lim_{n\to\infty} \tfrac{1}{n} r_n = \lim_{n\to\infty} \tfrac{1}{n} (s_n - p_n) = \delta s = P_{(\operatorname{rec} D)^{\ominus}}(\delta s).$$

Moreover, the compactness of D implies the continuity of σ_D [1, Example 11.2], and thus taking the limit of (5) yields

$$\lim_{n\to\infty} \tfrac{1}{n} \langle p_n \mid r_n \rangle = \lim_{n\to\infty} \sigma_D \left(\tfrac{1}{n} r_n \right) = \sigma_D \left(\lim_{n\to\infty} \tfrac{1}{n} r_n \right) = \sigma_D(\delta r).$$

When D is a (translated) closed convex cone, its recession cone is the cone itself, and the results of Prop. 3.2 can be shown using Moreau's decomposition and some basic properties of the projection operator; see [8, Lem. A.3 and Lem. A.4] for details.

A result that motivated our generalization of these limits to an arbitrary nonempty closed convex set D is given in [18, Lem. 4.3], where Prop. 3.2(ii) is established in a finite-dimensional setting.

4 Douglas–Rachford algorithm

The Douglas–Rachford algorithm is an operator splitting method, which can be used to solve composite minimization problems of the form

$$\underset{w\in\mathcal{H}}{\text{minimize}} \quad f(w) + g(w), \tag{6}$$

where f and g are proper lower semicontinuous convex functions. An iteration of the algorithm in application to problem (6) can be written as

$$\begin{aligned} w_n &= \mathrm{Prox}_g\, s_n \\ \tilde{w}_n &= \mathrm{Prox}_f(2w_n - s_n) \\ s_{n+1} &= s_n + \alpha(\tilde{w}_n - w_n). \end{aligned}$$

where $\alpha \in]0, 2[$ is the *relaxation parameter*.

If we rewrite problem (1) as

$$\begin{aligned} f(x,z) &= \tfrac{1}{2}\langle Qx \mid x\rangle + \langle q \mid x\rangle + \iota_{Ax=z}(x,z) \\ g(x,z) &= \iota_C(z), \end{aligned}$$

then an iteration of the Douglas–Rachford algorithm takes the following form [8,10]:

$$\tilde{x}_n = \underset{x\in\mathcal{H}_1}{\operatorname{argmin}} \left(\tfrac{1}{2}\langle Qx \mid x\rangle + \langle q \mid x\rangle + \tfrac{1}{2}\|x - x_n\|^2 + \tfrac{1}{2}\|Ax - (2P_C - \mathrm{Id})v_n\|^2\right) \tag{7a}$$

$$x_{n+1} = x_n + \alpha\,(\tilde{x}_n - x_n) \tag{7b}$$

$$v_{n+1} = v_n + \alpha\,(A\tilde{x}_n - P_C v_n) \tag{7c}$$

We will exploit the following well-known result to analyze the asymptotic behavior of the algorithm [19]:

Fact 4.1 *Iteration (7) amounts to*

$$(x_{n+1}, v_{n+1}) = T_{\mathrm{DR}}(x_n, v_n),$$

where $T_{\mathrm{DR}} : (\mathcal{H}_1 \times \mathcal{H}_2) \to (\mathcal{H}_1 \times \mathcal{H}_2)$ *is an* $(\alpha/2)$*-averaged operator.*

The solution to the subproblem in (7a) satisfies the optimality condition

$$Q\tilde{x}_n + q + (\tilde{x}_n - x_n) + A^* \left(A\tilde{x}_n - (2P_C - \mathrm{Id})v_n\right) = 0. \tag{8}$$

If we rearrange (7b) to isolate $\tilde{x}_n$,

$$\tilde{x}_n = x_n + \alpha^{-1}\delta x_{n+1},$$

and substitute it into (7c) and (8), we obtain the following relations between the iterates:

$$Ax_n - P_C v_n = -\alpha^{-1}\left(A\delta x_{n+1} - \delta v_{n+1}\right) \tag{9a}$$

$$Qx_n + q + A^*(\mathrm{Id} - P_C)v_n = -\alpha^{-1}\left((Q + \mathrm{Id})\delta x_{n+1} + A^*\delta v_{n+1}\right). \tag{9b}$$

Let us define the following auxiliary iterates of iteration (7):

$$z_n := P_C v_n \tag{10a}$$

$$y_n := (\mathrm{Id} - P_C)v_n. \tag{10b}$$

Observe that the pair (z_n, y_n) satisfies optimality condition (3c) for all $n \in \mathbb{N}$ [1, Prop. 6.47], and that the right-hand terms in (9) indicate how far the iterates (x_n, z_n, y_n) are from satisfying (3a) and (3b).

The following corollary follows directly from Fact 3.1, Prop. 3.2, Fact 4.1, and Moreau's decomposition [1, Thm. 6.30]:

Corollary 4.2 *Let the sequences* $(x_n)_{n\in\mathbb{N}}$, $(v_n)_{n\in\mathbb{N}}$, $(z_n)_{n\in\mathbb{N}}$, *and* $(y_n)_{n\in\mathbb{N}}$ *be given by (7) and (10), and* $(\delta x, \delta v) := P_{\overline{\mathrm{ran}}(T_{\mathrm{DR}} - \mathrm{Id})}(0)$. *Then*

(i) $\frac{1}{n}(x_n, v_n) \to (\delta x, \delta v)$.
(ii) $(\delta x_n, \delta v_n) \to (\delta x, \delta v)$.
(iii) $y_n \in (\mathrm{rec}\, C)^{\ominus}$.
(iv) $\frac{1}{n} z_n \to \delta z := P_{\mathrm{rec}\, C}(\delta v)$.
(v) $\frac{1}{n} y_n \to \delta y := P_{(\mathrm{rec}\, C)^{\ominus}}(\delta v)$.
(vi) $\lim_{n\to\infty} \frac{1}{n}\langle z_n \mid y_n\rangle = \sigma_C(\delta y)$.
(vii) $\delta z + \delta y = \delta v$.
(viii) $\langle \delta z \mid \delta y\rangle = 0$.
(ix) $\|\delta z\|^2 + \|\delta y\|^2 = \|\delta v\|^2$.

The following two propositions generalize [8, Prop. 5.1 and Prop. 5.2], though the proofs follow very similar arguments.

Proposition 4.3 *The following relations hold between* δx, δz, *and* δy, *which are defined in Cor. 4.2:*

(i) $A\delta x = \delta z$.
(ii) $Q\delta x = 0$.
(iii) $A^*\delta y = 0$.

(iv) $\delta z_n \to \delta z$.
(v) $\delta y_n \to \delta y$.

Proof (i) Divide (9a) by n, take the limit, and use Cor. 4.2(iv) to get

$$A\delta x = \lim_{n\to\infty} \tfrac{1}{n} P_C v_n = \delta z. \tag{11}$$

(ii) Divide (9b) by n, take the inner product of both sides with δx and take the limit to obtain

$$\langle Q\delta x \mid \delta x\rangle = -\lim_{n\to\infty} \left\langle A\delta x, \tfrac{1}{n}(\mathrm{Id} - P_C)v_n \right\rangle = -\langle \delta z \mid \delta y\rangle = 0,$$

where we used (11) and Cor. 4.2(v) in the second equality, and Cor. 4.2(viii) in the third. Due to [1, Cor. 18.18], the equality above implies

$$Q\delta x = 0. \tag{12}$$

(iii) Divide (9b) by n, take the limit, and use (12) to obtain

$$0 = \lim_{n\to\infty} \tfrac{1}{n} A^*(\mathrm{Id} - P_C)v_n = A^*\delta y,$$

where we used Cor. 4.2(v) in the second equality.

(iv) Subtracting (9a) at iterations $n+1$ and n, and taking the limit yield

$$\lim_{n\to\infty} \delta z_n = A\delta x = \delta z,$$

where the second equality follows from (11).

(v) From (10) we have

$$\lim_{n\to\infty} \delta y_n = \lim_{n\to\infty} (\delta v_n - \delta z_n) = \delta v - \delta z = \delta y,$$

where the last equality follows from Cor. 4.2(vii). □

Proposition 4.4 *The following identities hold for δx and δy, which are defined in Cor. 4.2:*

(i) $\langle q \mid \delta x\rangle = -\alpha^{-1}\|\delta x\|^2 - \alpha^{-1}\|A\delta x\|^2$.
(ii) $\sigma_C(\delta y) = -\alpha^{-1}\|\delta y\|^2$.

Proof Take the inner product of both sides of (9b) with δx and use (12) to obtain

$$\langle q \mid \delta x\rangle + \langle A\delta x \mid y_n\rangle = -\alpha^{-1}\langle \delta x \mid \delta x_{n+1}\rangle - \alpha^{-1}\langle A\delta x \mid \delta v_{n+1}\rangle .$$

Taking the limit and using Prop. 4.3(i) and Cor. 4.2(vii) and (viii) give

$$\langle q \mid \delta x\rangle + \alpha^{-1}\|\delta x\|^2 + \alpha^{-1}\|\delta z\|^2 = -\lim_{n\to\infty} \langle \delta z \mid y_n\rangle \geq 0, \tag{13}$$

where the inequality follows from Cor. 4.2(iii) and (iv) as the inner product of terms in rec C and $(\text{rec}\, C)^{\ominus}$ is nonpositive. Now take the inner product of both sides of (9a) with δy to obtain

$$\left\langle A^*\delta y \mid x_n + \alpha^{-1}\delta x_{n+1}\right\rangle - \langle \delta y \mid P_C v_n \rangle = \alpha^{-1} \langle \delta y \mid \delta v_{n+1}\rangle .$$

Due to Prop. 4.3(iii), the first inner product on the left-hand side is zero. Taking the limit and using Cor. 4.2(vii) and (viii), we obtain

$$-\alpha^{-1}\|\delta y\|^2 = \lim_{n\to\infty} \langle \delta y \mid P_C v_n\rangle \le \sup_{z\in C} \langle \delta y \mid z\rangle = \sigma_C(\delta y),$$

or equivalently,

$$\sigma_C(\delta y) + \alpha^{-1}\|\delta y\|^2 \ge 0. \tag{14}$$

Summing (13) and (14) and using Cor. 4.2(ix), we obtain

$$\langle q \mid \delta x\rangle + \sigma_C(\delta y) + \alpha^{-1}\|\delta x\|^2 + \alpha^{-1}\|\delta v\|^2 \ge 0. \tag{15}$$

Now take the inner product of both sides of (9b) with x_n to obtain

$$\begin{aligned}\langle Qx_n \mid x_n\rangle + \langle q \mid x_n\rangle + \langle Ax_n \mid y_n\rangle = &- \alpha^{-1} \langle (Q + \text{Id})\delta x_{n+1} \mid x_n\rangle \\ &- \alpha^{-1} \langle Ax_n \mid \delta v_{n+1}\rangle .\end{aligned}$$

Dividing by n, taking the limit, and using Prop. 4.3(i) and (ii) and Cor. 4.2(vii) and (viii) yield

$$\lim_{n\to\infty} \tfrac{1}{n} \langle Qx_n \mid x_n\rangle + \langle q \mid \delta x\rangle + \lim_{n\to\infty} \tfrac{1}{n} \langle Ax_n \mid y_n\rangle = -\alpha^{-1}\|\delta x\|^2 - \alpha^{-1}\|\delta z\|^2.$$

We can write the last term on the left-hand side as

$$\begin{aligned}\lim_{n\to\infty} \tfrac{1}{n} \langle Ax_n \mid y_n\rangle &= \lim_{n\to\infty} \tfrac{1}{n} \left\langle z_n + \alpha^{-1} (\delta v_{n+1} - A\delta x_{n+1}) \mid y_n\right\rangle \\ &= \lim_{n\to\infty} \tfrac{1}{n} \langle z_n \mid y_n\rangle + \alpha^{-1}\|\delta y\|^2 \\ &= \sigma_C(\delta y) + \alpha^{-1}\|\delta y\|^2,\end{aligned}$$

where the first equality follows from (9a), the second from Prop. 4.3(i) and Cor. 4.2(v) and (vii), and the third from Cor. 4.2(vi). Plugging the equality above in the preceding, we obtain

$$\langle q \mid \delta x\rangle + \sigma_C(\delta y) + \alpha^{-1}\|\delta x\|^2 + \alpha^{-1}\|\delta v\|^2 = - \lim_{n\to\infty} \tfrac{1}{n} \langle Qx_n \mid x_n\rangle \le 0, \tag{16}$$

where the inequality follows from the monotonicity of Q. Comparing inequalities (15) and (16), it follows that they must be satisfied with equality. Consequently, the left-hand sides of (13) and (14) must be zero. This concludes the proof. □

Given the infeasibility conditions in Prop. 2.1, it follows from Prop. 4.3 and Prop. 4.4 that, if the limit δy is nonzero, then problem (1) is strongly infeasible, and similarly, if δx is nonzero, then its dual is strongly infeasible. Thanks to the fact that $(\delta y_n, \delta x_n) \to (\delta y, \delta x)$, we can now extend the termination criteria proposed in [8, §5.2] for the more general case where C is a general nonempty closed convex set. The criteria in [8, §5.2] evaluate conditions given in Prop. 2.1 at δy_n and δx_n, and have already formed the basis for stable numerical implementations [10,11]. Our results pave the way for similar developments in the more general setting considered here.

5 Proximal-point algorithm

The proximal-point algorithm is a method for finding a vector $w \in \mathcal{H}$ that solves the following inclusion problem:

$$0 \in B(w), \tag{17}$$

where $B : \mathcal{H} \to 2^{\mathcal{H}}$ is a maximally monotone operator. An iteration of the algorithm in application to problem (17) can be written as

$$w_{n+1} = (\mathrm{Id} + \gamma B)^{-1} w_n,$$

where $\gamma > 0$ is the *regularization parameter*.

Due to [1, Cor. 16.30], we can rewrite (2) as

$$0 \in \mathcal{M}(x, y) := \begin{pmatrix} Qx + q + A^* y \\ -Ax + \partial \iota_C^*(y) \end{pmatrix},$$

where $\mathcal{M} : (\mathcal{H}_1 \times \mathcal{H}_2) \to 2^{(\mathcal{H}_1 \times \mathcal{H}_2)}$ is a maximally monotone operator [20]. An iteration of the proximal-point algorithm in application to the inclusion above is then

$$(x_{n+1}, y_{n+1}) = (\mathrm{Id} + \gamma \mathcal{M})^{-1} (x_n, y_n), \tag{18}$$

which was also analyzed in [12]. We will exploit the following result [1, Prop. 23.8] to analyze the algorithm:

Fact 5.1 *Operator* $T_{\mathrm{PP}} := (\mathrm{Id} + \gamma \mathcal{M})^{-1}$ *is the resolvent of a maximally monotone operator and is thus* $(1/2)$*-averaged.*

Iteration (18) reads

$$0 = x_{n+1} - x_n + \gamma \left(Qx_{n+1} + q + A^* y_{n+1} \right) \tag{19a}$$

$$0 \in y_{n+1} - y_n + \gamma \left(-Ax_{n+1} + \partial \iota_C^*(y_{n+1}) \right). \tag{19b}$$

Inclusion (19b) can be written as

$$\gamma A x_{n+1} + y_n \in \left(\mathrm{Id} + \gamma \partial \iota_C^* \right) y_{n+1},$$

which is equivalent to [1, Prop. 16.44]

$$y_{n+1} = \mathrm{Prox}_{\gamma \iota_C^*}(\gamma A x_{n+1} + y_n) = \gamma A x_{n+1} + y_n - \gamma P_C(A x_{n+1} + \gamma^{-1} y_n), \quad (20)$$

where the second equality follows from [1, Thm. 14.3]. Let us define the following auxiliary iterates of iteration (18):

$$v_{n+1} := A x_{n+1} + \gamma^{-1} y_n \quad (21a)$$
$$z_{n+1} := P_C v_{n+1}, \quad (21b)$$

and observe from (20) that

$$y_{n+1} = \gamma(\mathrm{Id} - P_C) v_{n+1}.$$

Using (19a) and (20), we now obtain the following relations between the iterates:

$$A x_{n+1} - P_C v_{n+1} = \gamma^{-1} \delta y_{n+1} \quad (22a)$$
$$Q x_{n+1} + q + \gamma A^*(\mathrm{Id} - P_C) v_{n+1} = -\gamma^{-1} \delta x_{n+1}. \quad (22b)$$

Similarly as for the Douglas–Rachford algorithm, the pair (z_{n+1}, y_{n+1}) satisfies optimality condition (3c) for all $n \in \mathbb{N}$. Observe that the optimality residuals, given by the norms of the left-hand terms in (22), can be computed by evaluating the norms of δy_{n+1} and δx_{n+1}.

The following corollary follows directly from Fact 3.1, Prop. 3.2, and Fact 5.1:

Corollary 5.2 *Let the sequences $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$, $(v_n)_{n\in\mathbb{N}}$, and $(z_n)_{n\in\mathbb{N}}$ be given by (18) and (21), and $(\delta x, \delta y) := P_{\overline{\mathrm{ran}}(T_{\mathrm{PP}} - \mathrm{Id})}(0)$. Then*

(i) $\frac{1}{n}(x_n, y_n, v_n) \to (\delta x, \delta y, A\delta x + \gamma^{-1}\delta y)$.
(ii) $(\delta x_n, \delta y_n, \delta v_n) \to (\delta x, \delta y, A\delta x + \gamma^{-1}\delta y)$.
(iii) $y_{n+1} \in (\mathrm{rec}\, C)^{\ominus}$.
(iv) $\frac{1}{n} z_n \to \delta z := P_{\mathrm{rec}\, C}(\delta v)$.
(v) $\delta y = \gamma P_{(\mathrm{rec}\, C)^{\ominus}}(\delta v)$.
(vi) $\lim_{n\to\infty} \frac{1}{n} \langle z_n \mid y_n \rangle = \sigma_C(\delta y)$.

The proofs of the following two propositions follow similar arguments as those in Sect. 4, and are thus omitted.

Proposition 5.3 *The following relations hold between δx, δz, and δy, which are defined in Cor. 5.2:*

(i) $A\delta x = \delta z$.
(ii) $Q\delta x = 0$.
(iii) $A^*\delta y = 0$.

Proposition 5.4 *The following identities hold for δx and δy, which are defined in Cor. 5.2:*

(i) $\langle q \mid \delta x\rangle = -\gamma^{-1}\|\delta x\|^2$.
(ii) $\sigma_C(\delta y) = -\gamma^{-1}\|\delta y\|^2$.

The authors in [12] use similar termination criteria to those given in [8, §5.2] to detect infeasibility of convex quadratic programs using the algorithm given by iteration (18), though they do not prove that δy and δx are indeed infeasibility certificates whenever the problem is strongly infeasible. Identities in (22) show that, when $(\delta y, \delta x) = (0, 0)$, the optimality conditions (3) are satisfied in the limit. Otherwise, Prop. 2.1, Prop. 5.3, and Prop. 5.4 imply that problem (1) and/or its dual is strongly infeasible.

Remark 5.5 *Weak infeasibility of problem (1) means that the sets* ran A *and* C *do not intersect, but the distance between them is zero. In such cases, there exists no* $\bar{y} \in \mathcal{H}_2$ *satisfying the conditions in Prop. 2.1 and the algorithms studied in Sects. 4–5 would yield* $\delta y_n \to \delta y = 0$. *A similar reasoning holds for the weak infeasibility of the dual problem for which the algorithms would yield* $\delta x_n \to \delta x = 0$.

Acknowledgements This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme grant agreement OCAL, No. 787845.

Funding Open Access funding provided by ETH Zurich.

Compliance with ethical standards

Data Availability Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

References

1. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces, 2nd edn. Springer, New York (2017). https://doi.org/10.1007/978-3-319-48311-5
2. Eckstein, J., Bertsekas, D.P.: On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators. Math. Program. **55**(1), 293–318 (1992). https://doi.org/10.1007/BF01581204
3. Bauschke, H.H., Dao, M.N., Moursi, W.M.: The Douglas–Rachford algorithm in the affine-convex case. Oper. Res. Lett. **44**(3), 379–382 (2016). https://doi.org/10.1016/j.orl.2016.03.010
4. Bauschke, H.H., Moursi, W.M.: The Douglas–Rachford algorithm for two (not necessarily intersecting) affine subspaces. SIAM J. Optim. **26**(2), 968–985 (2016). https://doi.org/10.1137/15M1016989
5. Bauschke, H.H., Moursi, W.M.: On the Douglas–Rachford algorithm. Math. Program. **164**(1), 263–284 (2017). https://doi.org/10.1007/s10107-016-1086-3
6. Ryu, E., Liu, Y., Yin, W.: Douglas-Rachford splitting and ADMM for pathological convex optimization. Comput. Optim. Appl. **74**, 747–778 (2019). https://doi.org/10.1007/s10589-019-00130-9

7. Bauschke, H.H., Moursi, W.M.: On the behavior of the Douglas–Rachford algorithm for minimizing a convex function subject to a linear constraint. SIAM J. Optim. **30**(3), 2559–2576 (2020). https://doi.org/10.1137/19M1281538
8. Banjac, G., Goulart, P., Stellato, B., Boyd, S.: Infeasibility detection in the alternating direction method of multipliers for convex optimization. J. Optim. Theory Appl. **183**(2), 490–519 (2019). https://doi.org/10.1007/s10957-019-01575-y
9. Liao-McPherson, D., Kolmanovsky, I.: FBstab: a proximally stabilized semismooth algorithm for convex quadratic programming. Automatica (2020). https://doi.org/10.1016/j.automatica.2019.108801
10. Stellato, B., Banjac, G., Goulart, P., Bemporad, A., Boyd, S.: OSQP: an operator splitting solver for quadratic programs. Math. Program. Comput. **12**(4), 637–672 (2020). https://doi.org/10.1007/s12532-020-00179-2
11. Garstka, M., Cannon, M., Goulart, P.: COSMO: a conic operator splitting method for large convex problems. In: European Control Conference (ECC) (2019). https://doi.org/10.23919/ECC.2019.8796161
12. Hermans, B., Themelis, A., Patrinos, P.: QPALM: a Newton-type proximal augmented Lagrangian method for quadratic programs. In: IEEE Conference on Decision and Control (CDC) (2019). https://doi.org/10.1109/CDC40024.2019.9030211
13. Pazy, A.: Asymptotic behavior of contractions in Hilbert space. Israel J. Math. **9**(2), 235–240 (1971). https://doi.org/10.1007/BF02771588
14. Baillon, J.B., Bruck, R.E., Reich, S.: On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces. Houston J. Math. **4**(1), 1–9 (1978)
15. Zarantonello, E.H.: Projections on convex sets in Hilbert space and spectral theory. In: Zarantonello, E.H. (ed.) Contributions to nonlinear functional analysis, pp. 237–424. Academic Press, Cambridge (1971). https://doi.org/10.1016/B978-0-12-775850-3.50013-3
16. Facchinei, F., Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer Series in Operations Research and Financial Engineering. Springer, New York (2003). https://doi.org/10.1007/b97543
17. Gowda, M.S., Sossa, D.: Weakly homogeneous variational inequalities and solvability of nonlinear equations over cones. Math. Program. **177**, 149–171 (2019). https://doi.org/10.1007/s10107-018-1263-7
18. Shen, J., Lebair, T.M.: Shape restricted smoothing splines via constrained optimal control and nonsmooth Newton's methods. Automatica **53**, 216–224 (2015). https://doi.org/10.1016/j.automatica.2014.12.040
19. Lions, P., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. **16**(6), 964–979 (1979). https://doi.org/10.1137/0716071
20. Rockafellar, R.T.: Augmented Lagrangians and applications of the proximal point algorithm in convex programming. Math. Oper. Res. **1**(2), 97–116 (1976). https://doi.org/10.1287/moor.1.2.97