Convergence rate of the relaxed CQ algorithm under Hölderian type error bound property

Lufang Zhang* Jinhua Wang† Chong Li‡ Xiaoqi Yang§

Abstract

The relaxed CQ algorithm is one of the most important algorithms for solving the split feasibility problem. We study the issue of strong convergence of the relaxed CQ algorithm in Hilbert spaces together with estimates on the convergence rate. Under a kind of Hölderian type bounded error bound property, strong convergence of the relaxed CQ algorithm is established. Furthermore, qualitative estimates on the convergence rate is presented. In particular, for the case when the involved exponent is equal to 1, the linear convergence of the relaxed CQ algorithm is established. Finally, numerical experiments are performed to show the convergence property of the relaxed CQ algorithm for the compressed sensing problem.

1 Introduction

Let H_1 and H_2 be Hilbert spaces. Let A be a bounded linear operator from H_1 to H_2. Let C and Q be arbitrary nonempty closed convex subsets of H_1 and H_2, respectively. Consider the following split feasibility problem (SFP): find a point $x \in H_1$ such that

$$x \in C \quad \text{and} \quad Ax \in Q. \tag{1.1}$$

Throughout the whole paper, we always assume that the solution set S of (1.1) is nonempty. The SFP (1.1) was introduced by Censor and Elfving in [10], which provides a unified framework for many inverse problems and has been used in various areas such as signal processing, image reconstruction [33] and intensity-modulated radiation therapy [11].

*School of Science, Zhejiang University of Science and Technology, Hangzhou 310023, P. R. China (lufangzhang@zust.edu.cn).

†Department of Mathematics, Hangzhou Normal University, Hangzhou 311121, P. R. China (wangjh@hznu.edu.cn). This author's work was supported in part by the National Natural Science Foundation of China (grant 12171131).

‡School of Mathematical Sciences, Zhejiang University, Hangzhou 310027, P. R. China (cli@zju.edu.cn). This author's work was supported in part by the National Natural Science Foundation of China (grant 11971429).

§Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong (mayangxq@polyu.edu.hk). This author's work was supported in part by the Research Grants Council of Hong Kong (PolyU 152182/19E).

The CQ algorithm introduced in [8] is one of the most popular and practical algorithms to solve the SFP (1.1) which, with the stepsizes $\{\beta_n\} \subseteq \mathbb{R}_+$, is formulated as follows:

$$x_{n+1} = P_C(x_n - \beta_n A^*(\mathbb{I} - P_Q)Ax_n),$$

where $\mathbb{I}$ is the identity in H_2, A^* is the adjoint of A, while P_Ω is the metric projection onto Ω (a subset of a Hilbert space H). Convergence properties of the CQ algorithm with different types of stepsizes have been extensively explored; see [8, 26, 35, 37, 39] and references therein. In particular, the weak convergence property of the CQ algorithm in Hilbert spaces with constant stepsize and with dynamic stepsizes were established in [37] and [26], respectively. Recently, Wang et al [35] established the linear convergence result for the CQ algorithm with the constant stepsize or with the dynamic stepsizes in Hilbert spaces. The main computational cost of the CQ algorithm at each step depends on the projections $P_C(\cdot)$ and $P_Q(\cdot)$, which in general are expensive (see [2]). To reduce the computational cost, Yang [38] proposed a relaxed CQ algorithm, i.e., Algorithm 1, for solving the SFP (1.1), when the involved convex sets C and Q are respectively the level sets of the continuous convex functions $c : H_1 \to \mathbb{R}$ and $q : H_2 \to \mathbb{R}$:

$$C := \{x \in H_1 : c(x) \le 0\} \quad \text{and} \quad Q := \{y \in H_2 : q(y) \le 0\}. \tag{1.2}$$

Recall that, for a convex function $f : H \to \mathbb{R}$, the *subdifferential* of f is denoted by ∂f and defined by

$$\partial f(x) := \{u \in H : \langle u, y - x\rangle \le f(y) - f(x), \ \ \forall y \in H\} \quad \text{for each } x \in H.$$

f is said to be subdifferentiable at $x \in H$ if $\partial f(x) \ne \emptyset$.

Algorithm 1. Let $x_0 \in H_1$ be given. Having $x_0, x_1, \cdots, x_n$, choose $\xi_n \in \partial c(x_n)$, $\eta_n \in \partial q(Ax_n)$, and stepsize $\beta_n \ge 0$, and determine x_{n+1} by

$$x_{n+1} := P_{C_n}(x_n - \beta_n A^*(\mathbb{I} - P_{Q_n})Ax_n),$$

where

$$C_n := \{x \in H_1 : c(x_n) + \langle \xi_n, x - x_n\rangle \le 0\} \tag{1.3}$$

and

$$Q_n := \{y \in H_2 : q(Ax_n) + \langle \eta_n, y - Ax_n\rangle \le 0\}. \tag{1.4}$$

Note that C_n and Q_n are half-spaces and so the projections P_{C_n} and P_{Q_n} have closed form expressions; see [13, 38]. Moreover, Algorithm 1 also covers the CQ algorithm by taking $c(\cdot) := \mathrm{d}_C(\cdot)$ and $q(\cdot) := \mathrm{d}_Q(\cdot)$, the distance functions of C and Q, respectively. Weak convergence results of Algorithm 1 were established in [37] and [26] for the case when the stepsizes $\{\beta_n\}$ are constants satisfying

$$\sigma \le \beta_n \le \frac{2}{\|A\|^2} - \sigma, \tag{1.5}$$

or stepsizes $\{\beta_n\}$ are dynamic given by

$$\beta_n := \frac{\rho_n\|(\mathbb{I} - P_{Q_n})Ax_n\|^2}{\|A^*(\mathbb{I} - P_{Q_n})Ax_n\|^2} \quad \text{with} \quad \sigma < \rho_n < 2 - \sigma, \tag{1.6}$$

for some $0 < \sigma \le \frac{1}{\|A\|^2}$ for (1.5) or $0 < \sigma < 1$ for (1.6), where the convention that $\frac{0}{0} = 0$ is adopted.

To the best of our knowledge, the result regarding the linear convergence for Algorithm 1, similar to the one studied in [35] for the CQ algorithm, has not been explored. It seems to us that the bounded linear regularity property, which plays an important role in the analysis for the CQ algorithm in [35], is not enough here and the techniques used in [35] do not work for us to establish linear convergence of Algorithm 1. Recall that the bounded error bound property has been studied extensively and is a powerful tool in convergence analysis of optimization algorithms; see [1, 6, 16, 22] and references therein. A natural extension is the error bound with fractional exponent (which is usually called the Hölderian type error bound) where the exponent may have a close relationship with the convergence rate of some algorithms; see, e.g. [5, 18–20, 23, 24, 27] and references therein. Inspired by this and the work of [35], in the present paper, we will study the issue of strong convergence of Algorithm 1 in Hilbert spaces together with estimates on the convergence rate. Under Hölderian type bounded error bound property, strong convergence of Algorithm 1 is established. Furthermore, qualitative estimates on the convergence rate is presented. In particular, for the case when the involved exponent is equal to 1, the linear convergence of Algorithm 1 is established.

The paper is organized as follows. As usual, some auxiliary results are presented in the next section. Strong convergence together with estimates on the convergence rate of Algorithm 1 is established in Section 3. Numerical experiments to show the convergence property of Algorithm 1 are given in Section 4.

2 Preliminaries

For simplicity, let $\mathbb{N}$ be the set of all positive integers and let $\mathbb{N}^* = \mathbb{N} \cup \{0\}$. For $x \in H$ and $r > 0$, we use $\mathbb{B}(0, r)$ to denote the closed metric ball at x with radius r, that is,

$$\mathbb{B}(x, r) := \{y \in H : \|x - y\| \le r\}.$$

The following propositions are useful. Proposition 2.1 is about some well-known properties of the projection operator, which is taken from [3, Corollary 4.10 and Theoren 3.14]. Proposition 2.2 is known in [32, p46, Lemma 6].

Proposition 2.1. *Let Ω be a nonempty closed convex set in H and let $x, y \in H$. Then the following assertions hold:*

(i) $\langle (\mathbb{I} - P_\Omega)x - (\mathbb{I} - P_\Omega)y, x - y\rangle \ge \|(\mathbb{I} - P_\Omega)x - (\mathbb{I} - P_\Omega)y\|^2$;

(ii) $[z = P_\Omega(x)] \Longleftrightarrow [z \in \Omega$ *and* $\langle x - z, y - z\rangle \le 0,\ \forall y \in \Omega]$.

Proposition 2.2. *Let $\alpha > 0$, $p > 0$, and let $\{\beta_k\}$ be a sequence of nonnegative numbers satisfying*

$$\beta_{k+1} \le \beta_k(1 - \alpha\beta_k^p) \quad \textit{for each} \quad k \in \mathbb{N}^*.$$

Then,

$$\beta_{k+1} \le \Big(\beta_0^{-p} + p\alpha(k+1)\Big)^{-\frac{1}{p}} \quad \textit{for each} \quad k \in \mathbb{N}^*.$$

Recall that the solution set of (1.1) is nonempty, that is, $S := C \bigcap A^{-1}Q \neq \emptyset$. We need the following lemma about useful properties for the sequences generated by Algorithm 1. Set $[a]_+ := \max\{a, 0\}$ for each $a \in \mathbb{R}$.

Lemma 2.3. *Let $\{x_n\}$ be a sequence generated by Algorithm 1 (together with associated sequences $\{C_n\}, \{Q_n\}, \{\xi_n\}, \{\eta_n\}$). Then, for each $n \in \mathbb{N}^*$, we have the following assertions:*

(i) $C \subseteq C_n$ *and* $Q \subseteq Q_n$.

(ii) $[0 \in \partial c(x_n)] \Longrightarrow [x_n \in C]$ *and* $[0 \in \partial q(Ax_n)] \Longrightarrow [Ax_n \in Q]$.

(iii) $[c(x_n)]_+ = \|\xi_n\| \mathrm{d}_{C_n}(x_n)$ *and* $[q(Ax_n)]_+ = \|\eta_n\| \mathrm{d}_{Q_n}(Ax_n)$.

(iv) $[x_n \in C_n \cap A^{-1}Q_n] \Longleftrightarrow [x_n \in S]$.

Proof. (i) Let $x \in C$. Then $c(x) \leq 0$. Since $\xi_n \in \partial c(x_n)$, one has by definition that

$$\langle \xi_n, x - x_n \rangle \leq c(x) - c(x_n),$$

and so

$$\langle \xi_n, x - x_n \rangle + c(x_n) \leq c(x) \leq 0$$

which gives that $x \in C_n$. Similarly, let $y \in Q$. Then $q(y) \leq 0$. Since $\eta_n \in \partial q(Ax_n)$, one has by definition that

$$\langle \eta_n, y - Ax_n \rangle \leq q(y) - q(Ax_n)$$

and so

$$\langle \eta_n, y - Ax_n \rangle + q(Ax_n) \leq q(y) \leq 0,$$

which implies that $y \in Q_n$.

(ii) Assume $0 \in \partial c(x_n)$. Then we have that

$$\langle 0, x - x_n \rangle \leq c(x) - c(x_n) \quad \text{ for each } x \in H_1.$$

This implies that $c(x_n) \leq 0$ and so $x_n \in C$. Similarly, assume $0 \in \partial q(Ax_n)$. Then we have that

$$\langle 0, y - Ax_n \rangle \leq q(y) - q(Ax_n) \quad \text{ for each } y \in H_2.$$

This gives that $q(Ax_n) \leq 0$ and so $Ax_n \in Q$.

(iii) For the case when $x_n \in C_n$, it follows from (1.3) that $c(x_n) \leq 0$ and so $[c(x_n)]_+ = 0 = \|\xi_n\| \mathrm{d}_{C_n}(x_n)$. Below, we consider the case when $x_n \notin C_n$. Note by assertions (i)-(ii) that $0 \notin \partial c(x_n)$ (so $\xi_n \neq 0$) and $c(x_n) > 0$. Thus, for any $x \in C_n$, one has that

$$\langle x_n - (x_n - \frac{c(x_n)}{\|\xi_n\|^2}\xi_n), x - (x_n - \frac{c(x_n)}{\|\xi_n\|^2}\xi_n) \rangle = \frac{c(x_n)}{\|\xi_n\|^2}(c(x_n) + \langle \xi_n, x - x_n \rangle) \leq 0,$$

which together with Proposition 2.1 (ii) gives that $P_{C_n}(x_n) = x_n - \frac{c(x_n)}{\|\xi_n\|^2}\xi_n$ and so

$$\mathrm{d}_{C_n}(x_n) = \|x_n - P_{C_n}(x_n)\| = \frac{|c(x_n)|}{\|\xi_n\|} = \frac{c(x_n)}{\|\xi_n\|},$$

which implies that $[c(x_n)]_+ = \|\xi_n\| \mathrm{d}_{C_n}(x_n)$.

Similarly, we can obtain that $[q(Ax_n)]_+ = \|\eta_n\| \mathrm{d}_{Q_n}(Ax_n)$.

(iv) Assume $x_n \in C_n \cap A^{-1}Q_n$. Then we have by (1.3) and (1.4) that $c(x_n) \leq 0$ and $q(Ax_n) \leq 0$, respectively. It means that $x_n \in C \cap A^{-1}Q = S$. The converse can be obtained by (i).

□

3 Convergence rate of the relaxed CQ algorithm

Recalling that the error bound property is found to have many important applications in various areas such as instance sensitivity analysis, convergence analysis of many algorithms for optimization problems, and so on; see, e.g. [28,31] and references therein. The notion of the (Lipschitz-type) error bound was originally introduced by Hoffman in [15] for systems of affine functions in finite dimensional spaces. Since then, various extensions of this result to convex and/or non-convex inequalities and/or equality systems, have been explored and well developed; see [30,31,36] and references therein. Another natural extension is the error bound with fractional exponent (i.e., Hölderian type error bound) where the exponent may have a close relationship with the convergence rate of some algorithms; see, e.g. [5, 18–20, 23, 24, 27] and references therein. In order to establish strong convergence together with estimates on the convergence rate of Algorithm 1, we need the following notion of Hölderian type bounded error bound property for the SFP (1.1)-(1.2).

Definition 3.1. Let $\tau \in (0,1]$. The SFP (1.1)-(1.2) is said to satisfy the bounded error bound property with exponent τ, if for any bounded subset W of H_1 with $W \cap S \neq \varnothing$, there exists $\gamma_W > 0$ such that

$$\mathrm{d}_S(x) \leq \gamma_W (\max\{[c(x)]_+, [q(Ax)]_+\})^\tau \quad \text{for each } x \in W. \tag{3.1}$$

For the remainder of this paper, we always assume that $\tau \in (0,1]$, and that c and q, or equivalently, their subdifferentials, are bounded on bounded sets.

Remark 3.2. (a) The study of the Hölderian type error bound properties for convex polynomial systems, nonnegative convex polynomial systems and general polynomial systems was done in [5, Corollary 3.4], [21, Theorem 4.3] and [25], respectively.

(b) In the special case when $c(\cdot) := \mathrm{d}_C(\cdot)$ and $q(\cdot) := \mathrm{d}_Q(\cdot)$, the bounded error bound property with exponent $\tau = 1$ for the SFP (1.1)-(1.2) is equivalent to the bounded linear regularity property for the SFP (1.1) considered in [35]. In fact, note that, for each $x \in H$,

$$\mathrm{d}_S(x) \leq \|x - P_C x\| + \mathrm{d}_S(P_C x) \quad \text{and} \quad \mathrm{d}_Q(AP_C x) \leq \mathrm{d}_Q(Ax) + \|A\| \mathrm{d}_C(x);$$

then the equivalence can be checked directly by definition; see also [14] for more details.

For the general case, using the Lipschitz property of the functions c and q on bounded subsets, one checks directly by definition that the bounded error bound property with the exponent $\tau = 1$ implies the bounded linear regularity property. However, the converse is not true in general; indeed, for a nontrivial SFP (1.1) satisfying the bounded linear regularity property, the SFP (1.1)-(1.2) with $c(\cdot) := \mathrm{d}_C^2(\cdot)$ and $q(\cdot) := \mathrm{d}_Q^2(\cdot)$, fails to satisfy the bounded error bound property with exponent $\tau = 1$.

Our main theorem in the present paper is as follows, which extends the corresponding one in [35, Theorem 2.3].

Theorem 3.3. *Suppose that the SFP* (1.1)-(1.2) *satisfies the bounded error bound property with exponent τ and let $\{x_n\}$ be a sequence generated by Algorithm 1 with each stepsize β_n satisfying* (1.5) *or* (1.6). *Then, $\{x_n\}$ converges strongly to a solution x^* of the SFP*

(1.1)-(1.2). *Furthermore, there exists $\alpha > 0$ such that for each $n \in \mathbb{N}^*$,*

$$\|x_n - x^*\| \le \begin{cases} 2(1-\alpha)^{\frac{1}{2}n} \mathrm{d}_S(x_0), & \tau = 1, \\ 2\left(\mathrm{d}_S^{2(1-\frac{1}{\tau})}(x_0) + \alpha(\frac{1}{\tau} - 1)n\right)^{-\frac{\tau}{2(1-\tau)}}, & \tau \in (0,1). \end{cases} \tag{3.2}$$

In particular, if $\tau = 1$, then $\{x_n\}$ converges linearly.

Proof. For simplicity, write

$$I_0 := \{n \in \mathbb{N}^* : Ax_n \in Q_n\}.$$

Without loss of generality, we may assume that

$$0 \notin \partial c(x_n) \quad \text{for each } n \in I_0$$

because, if there is $n \in I_0$ such that $0 \in \partial c(x_n)$, then $x_n \in C_n$ by Lemma 2.3(i)-(ii), and so $x_n \in S$ by Lemma 2.3(iv). Let $z \in S$. Fix $n \in \mathbb{N}^*$, and set

$$\nabla_{x_n} := A^*(\mathbb{I} - P_{Q_n})Ax_n.$$

Then, one checks that

$$\|\nabla_{x_n}\| \le \|A\| \mathrm{d}_{Q_n}(Ax_n) \quad \text{and} \quad \langle x_n - z, \nabla_{x_n}\rangle \ge \mathrm{d}^2_{Q_n}(Ax_n). \tag{3.3}$$

In fact, the first inequality follows directly from the definition of ∇_{x_n}, while for the second one, we check from the fact $(\mathbb{I} - P_{Q_n})Az = 0$ that

$$\begin{aligned} \langle x_n - z, \nabla_{x_n}\rangle &= \langle Ax_n - Az, (\mathbb{I} - P_{Q_n})Ax_n - (\mathbb{I} - P_{Q_n})Az\rangle \\ &\ge \| (\mathbb{I} - P_{Q_n})Ax_n - (\mathbb{I} - P_{Q_n})Az \|^2 \\ &= \mathrm{d}^2_{Q_n}(Ax_n), \end{aligned}$$

where the inequality holds by Proposition 2.1(i). Below, we show by the choice of β_n (see (1.5) and (1.6)), that for each $n \notin I_0$,

$$\beta_n \ge \min\left\{\sigma, \frac{\sigma \mathrm{d}^2_{Q_n}(Ax_n)}{\|\nabla_{x_n}\|^2}\right\} \ge \sigma \min\{1, \|A\|^{-2}\} \tag{3.4}$$

and

$$2 - \frac{\beta_n \|\nabla_{x_n}\|^2}{\mathrm{d}^2_{Q_n}(Ax_n)} \ge \min\{2 - \beta_n\|A\|^2, 2 - \rho_n\} \ge \sigma \min\{\|A\|^2, 1\}. \tag{3.5}$$

Indeed, for the case when β_n is given by (1.5), (3.4) follows directly from (3.3), while for the case when β_n is given by (1.6), we obtain from (3.3) that

$$\beta_n = \frac{\rho_n \|(\mathbb{I} - P_{Q_n})Ax_n\|^2}{\|A^*(\mathbb{I} - P_{Q_n})Ax_n\|^2} > \frac{\sigma \mathrm{d}^2_{Q_n}(Ax_n)}{\|\nabla_{x_n}\|^2} \ge \frac{\sigma}{\|A\|^2};$$

hence (3.4) is seen to hold. Additionally, for the case when β_n is given by (1.5), we have by (3.3) that

$$2 - \frac{\beta_n \|\nabla_{x_n}\|^2}{\mathrm{d}^2_{Q_n}(Ax_n)} \geq 2 - \beta_n \|A\|^2 \geq \sigma \|A\|^2,$$

while for the case when β_n is given by (1.6), it follows from the fact $\beta_n = \frac{\rho_n \mathrm{d}^2_{Q_n}(Ax_n)}{\|\nabla_{x_n}\|^2}$ and $\rho_n < 2 - \sigma$ that

$$2 - \frac{\beta_n \|\nabla_{x_n}\|^2}{\mathrm{d}^2_{Q_n}(Ax_n)} = 2 - \rho_n > \sigma;$$

hence (3.5) is checked.

To proceed, for simplicity, we write

$$u_n := x_n - \beta_n \nabla_{x_n}. \tag{3.6}$$

Noting $(\mathbb{I} - P_{C_n})z = 0$, one has from Proposition 2.1(i) that

$$\langle u_n - P_{C_n}(u_n), u_n - z\rangle = \langle u_n - P_{C_n}(u_n) - (\mathbb{I} - P_{C_n})z, u_n - z\rangle \geq \|P_{C_n}(u_n) - u_n\|^2.$$

Thus, it follows that

$$\|x_{n+1} - z\|^2 = \|P_{C_n}(u_n) - u_n + u_n - z\|^2 \leq \|u_n - z\|^2 - \mathrm{d}^2_{C_n}(u_n). \tag{3.7}$$

Now set $\zeta := \sigma \min\{\|A\|^2, 1\}$. Below we shall show that

$$\|x_{n+1} - z\|^2 \leq \|x_n - z\|^2 - \mathrm{d}^2_{C_n}(u_n) - \beta_n \zeta \mathrm{d}^2_{Q_n}(Ax_n) \quad \text{for each } n \in \mathbb{N}^*. \tag{3.8}$$

This follows trivially from (3.7) in the case when $n \in I_0$ (noting that $\mathrm{d}^2_{Q_n}(Ax_n) = 0, \nabla_{x_n} = 0$ and so $u_n = x_n$ by (3.6) in this case). For the case when $n \notin I_0$, it follows from (3.6), (3.3) and (3.5) that

$$\begin{aligned}
\|u_n - z\|^2 &= \|x_n - z\|^2 - 2\beta_n \langle x_n - z, \nabla_{x_n}\rangle + \beta_n^2 \|\nabla_{x_n}\|^2 \\
&\leq \|x_n - z\|^2 - \beta_n \left(2 - \beta_n \frac{\|\nabla_{x_n}\|^2}{\mathrm{d}^2_{Q_n}(Ax_n)}\right) \mathrm{d}^2_{Q_n}(Ax_n) \\
&\leq \|x_n - z\|^2 - \beta_n \zeta \mathrm{d}^2_{Q_n}(Ax_n).
\end{aligned}$$

Combining this with (3.7) yields (3.8). Thus, $\{\|x_n - z\|\}$ is monotone decreasing, and so $\{\|x_n - z\|\}$ is bounded. Let W be a bounded subset of H_1 containing z and $\{x_n\}$. Then by the assumed bounded error bound property, there exists $\gamma_W > 0$ such that

$$\mathrm{d}_S(x_n) \leq \gamma_W (\max\{[c(x_n)]_+, [q(Ax_n)]_+\})^\tau \quad \text{for each } n \in \mathbb{N}^*. \tag{3.9}$$

Set

$$\alpha := \gamma_W^{\frac{-2}{\tau}} \min\left\{\inf_{n \notin I_0} \frac{\beta_n \zeta}{\|\eta_n\|^2}, \inf_n \frac{\zeta}{\|\xi_n\|^2(\zeta + 2)}\right\}. \tag{3.10}$$

Then α is well defined and $\alpha > 0$ by (3.4) and the fact that $\{\|\xi_n\|\}, \{\|\eta_n\|\}$ are bounded by assumption. Below, we show that

$$\mathrm{d}^2_{C_n}(u_n) + \beta_n \zeta \mathrm{d}^2_{Q_n}(Ax_n) \geq \alpha \mathrm{d}_S^{\frac{2}{\tau}}(x_n) \quad \text{for each } n \in \mathbb{N}^*. \tag{3.11}$$

This is clear for the case when $[c(x_n)]_+ < [q(Ax_n)]_+$ because, by (3.9) and Lemma 2.3(iii), one has that

$$\gamma_W^{\frac{-1}{\tau}} \mathrm{d}_S^{\frac{1}{\tau}}(x_n) \le [q(Ax_n)]_+ = \|\eta_n\| \mathrm{d}_{Q_n}(Ax_n),$$

and so

$$\mathrm{d}_{C_n}^2(u_n) + \beta_n \zeta \mathrm{d}_{Q_n}^2(Ax_n) \ge \beta_n \zeta \mathrm{d}_{Q_n}^2(Ax_n) \ge \frac{\gamma_W^{\frac{-2}{\tau}} \beta_n \zeta}{\|\eta_n\|^2} \mathrm{d}_S^{\frac{2}{\tau}}(x_n)$$

(noting that $n \notin I_0$ in this case). Thus, we only consider the case when $[c(x_n)]_+ \ge [q(Ax_n)]_+$. To do this, we note first by (3.6) that $\|x_n - P_{C_n}(u_n)\| = \|u_n - P_{C_n}(u_n) + \beta_n \nabla_{x_n}\| \le \mathrm{d}_{C_n}(u_n) + \|\beta_n \nabla_{x_n}\|$, and so we have that

$$\begin{aligned} \|x_n - P_{C_n}(u_n)\|^2 &\le \mathrm{d}_{C_n}^2(u_n) + 2\mathrm{d}_{C_n}(u_n)\, \|\beta_n \nabla_{x_n}\| + \|\beta_n \nabla_{x_n}\|^2 \\ &\le \tfrac{2+\zeta}{\zeta} \mathrm{d}_{C_n}^2(u_n) + \left(1 + \tfrac{\zeta}{2}\right) \|\beta_n \nabla_{x_n}\|^2, \end{aligned} \tag{3.12}$$

where the last inequality holds by the following basic inequality:

$$2\mathrm{d}_{C_n}(u_n)\, \|\beta_n \nabla_{x_n}\| \le \frac{2}{\zeta} \mathrm{d}_{C_n}^2(u_n) + \frac{\zeta}{2} \|\beta_n \nabla_{x_n}\|^2.$$

Furthermore, from (3.5), one sees that $2 - \frac{\beta_n \|\nabla_{x_n}\|^2}{\mathrm{d}_{Q_n}^2(Ax_n)} > 0$ and so $\beta_n \|\nabla_{x_n}\|^2 \le 2\mathrm{d}_{Q_n}^2(Ax_n)$; then it follows from (3.12) that

$$\mathrm{d}_{C_n}^2(x_n) \le \|x_n - P_{C_n}(u_n)\|^2 \le \frac{2+\zeta}{\zeta} \mathrm{d}_{C_n}^2(u_n) + 2(1 + \frac{\zeta}{2}) \beta_n \mathrm{d}_{Q_n}^2(Ax_n).$$

This implies that

$$\frac{\zeta}{2+\zeta} \mathrm{d}_{C_n}^2(x_n) \le \mathrm{d}_{C_n}^2(u_n) + \zeta \beta_n \mathrm{d}_{Q_n}^2(Ax_n). \tag{3.13}$$

Recalling that $[c(x_n)]_+ \ge [q(Ax_n)]_+$, we have by (3.9) and Lemma 2.3(iii) again that $\gamma_W^{\frac{-1}{\tau}} \mathrm{d}_S^{\frac{1}{\tau}}(x_n) \le [c(x_n)]_+ = \|\xi_n\| \mathrm{d}_{C_n}(x_n)$. Combining this with (3.13) gives that

$$\frac{\gamma_W^{\frac{-2}{\tau}} \zeta}{(2+\zeta)\|\xi_n\|^2} \mathrm{d}_S^{\frac{2}{\tau}}(x_n) \le \mathrm{d}_{C_n}^2(u_n) + \zeta \beta_n \mathrm{d}_{Q_n}^2(Ax_n);$$

then (3.11) is seen to hold by the definition of α in (3.10).

Now combining (3.8) and (3.11), we arrive at

$$\|x_{n+1} - z\|^2 \le \|x_n - z\|^2 - \alpha \mathrm{d}_S^{\frac{2}{\tau}}(x_n) \quad \text{for each} \quad n \in \mathbb{N}^*. \tag{3.14}$$

Since $z \in S$ is arbitrary, it follows from (3.14) that

$$\mathrm{d}_S^2(x_{n+1}) \le (1 - \alpha \mathrm{d}_S^{2(\frac{1}{\tau}-1)}(x_n)) \mathrm{d}_S^2(x_n) \quad \text{for each} \quad n \in \mathbb{N}^*. \tag{3.15}$$

Then, we have that for each $n \in \mathbb{N}^*$,

$$\mathrm{d}_S^2(x_{n+1}) \le \begin{cases} (1-\alpha)^{n+1} \mathrm{d}_S^2(x_0), & \tau = 1, \\ \left(\mathrm{d}_S^{2(1-\frac{1}{\tau})}(x_0) + \alpha(\frac{1}{\tau} - 1)(n+1) \right)^{-\frac{\tau}{1-\tau}}, & \tau \in (0,1). \end{cases} \tag{3.16}$$

In fact, for the case when $\tau = 1$, (3.16) follows directly from (3.15), while for the case when $\tau \in (0,1)$, (3.16) holds by (3.15) and applying Proposition 2.2 to $\{\mathrm{d}_S^2(x_k)\}, \frac{1}{\tau} - 1$ in place of $\{\beta_k\}, p$. Fix $n \in \mathbb{N}^*$. Recalling that the sequence $\{\|x_m - P_S(x_n)\| : m \in \mathbb{N}^*\}$ is monotonically decreasing, we have that for each $m > n$

$$\|x_m - x_n\| \leq \|x_m - P_S(x_n)\| + \|x_n - P_S(x_n)\| \leq 2\|x_n - P_S(x_n)\| = 2\mathrm{d}_S(x_n).$$

This, together with (3.16), implies that

$$\|x_m - x_n\| \leq \begin{cases} 2(1-\alpha)^{\frac{1}{2}n}\mathrm{d}_S(x_0), & \tau = 1, \\ 2\left(\mathrm{d}_S^{2(1-\frac{1}{\tau})}(x_0) + \alpha(\frac{1}{\tau} - 1)n\right)^{-\frac{\tau}{2(1-\tau)}}, & \tau \in (0,1). \end{cases}$$

Hence, it follows that $\{x_n\}$ is a Cauchy sequence and converges to a point $x^* \in H_1$. Passing to the limit of $m \to \infty$, we arrive at (3.2). Furthermore, letting $n \to \infty$ in (3.16), one has that $x^* \in S$ because S is closed. Hence x^* is a solution of the SFP (1.1)-(1.2). The proof is complete. □

We end this section with a remark to discuss some comparisons related to algorithms appeared in [4, 7, 9, 34].

Remark 3.4. (a) In [34], the authors proposed a relaxed gradient projection algorithm (RGPA for short), which is an extension of the relaxed simultanneous iterative algorithm (RSSEA for short) introduced in [9], for solving the split equality problem (SEQ for short), and studied the linear convergence property of the RGPA under a bounded linear regularity assumption. In the case when the involved linear operator B is the identity, the SEQ is reduced to the SFP, and it is easy to check that Algorithm 1 is different from the RGPA and the RSSEA for the SFP; see also the the numerical experiments for comparisons of Algorithm 1 and the RSSEA in Section 4.

(b) Under a bounded Hölderian regularity assumption, the convergence rate of the quasi-cyclic algorithm (which was proposed in [4]) for finding a common fixed point of a finite family of averaged nonexpansive operators in a Hilbert space was analyzed in [7], which, in particular, extends the corresponding linear convergence results in [4]. As explained in [7], this algorithm covers many important iterative methods including Krasnoselskii-Man iterations, the cyclic projection algorithm, forward-backward splitting methods and the Douglas-Rachford feasibility algorithm along with some variants. Note that the SFP (1.1) is equivalent to the problem: finding a point $x \in H_1$ such that $x \in C \cap A^{-1}Q$. Thus, in the case when A is the identity, the results in [7] can apply directly. However, in the general case, it seems unclear how to construct an averaged nonexpansive operator T_Q avoiding the "inverse" A^{-1} such that $A^{-1}Q$ is the set of fixed points of T_Q. Furthermore, even in the case when A is the identity, Algorithm 1 is, in general, different from the quasi-cyclic algorithm (note that the involved projection operators in Algorithm 1 are adaptive).

4 Numerical Experiments

In this section, we give a numerical experiment to show the convergence property of Algorithm 1. Here we consider the compressed sensing problem known in [12] which is represented approximately by a linear system of the form $b = Ax + e$, where $A \in \mathbb{R}^{m \times n}$ and

$b \in \mathbb{R}^m$ are known, and $e \in \mathbb{R}^m$ is an arbitrary and unknown vector of errors, and $x \in \mathbb{R}^n$ is the variable to be estimated. The sparsity of x is measured by the ℓ_1 norm, which is defined by $\|x\|_1 := \sum_{i=1}^n |x_i|$ (cf. [12]). Let $t \geq 0$ be a constant and $\varepsilon := \|e\|$. Write

$$C := \{x \in \mathbb{R}^n \mid \|x\|_1 \leq t\} \text{ and } Q := \{y \in \mathbb{R}^m \mid \|y - b\|_2 \leq \varepsilon\}. \tag{4.1}$$

Thus, the compressed sensing problem can be regarded as the SFP (1.1)-(1.2) with $H_1 = \mathbb{R}^n$, $H_2 = \mathbb{R}^m$,

$$c(x) := \max_{\alpha \in \Lambda}\{\alpha^T x - t\} \quad \text{for each} \quad x \in \mathbb{R}^n, \tag{4.2}$$

where $\Lambda := \{\alpha \in \mathbb{R}^n : \alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)^T, \alpha_i \in \{1, -1\}, i = 1, 2, \cdots, n\}$, and

$$q(y) := \|y - b\|_2^2 - \varepsilon^2 \quad \text{for each} \quad y \in \mathbb{R}^m. \tag{4.3}$$

Furthermore, the SFP (1.1)-(1.2) satisfies the bounded error bound property with exponent $\frac{1}{2}$; see Remark 4.1 for details.

Remark 4.1. The SFP (1.1)-(1.2) (with $C, Q, c(\cdot), q(\cdot)$ given by (4.1), (4.2), (4.3), respectively) satisfies the bounded error bound property with exponent $\frac{1}{2}$. In fact, in the case when the minimum value of $q \circ A$ on C is negative, the conclusion is clear by the known result for quadratic convex systems satisfying the Slater condition; see, e.g., [27, Theorem 3.1]. Hence, we only need to consider the case when

$$q(Ax) \geq 0 \text{ for each } x \in C. \tag{4.4}$$

Recall that S is the solution set of the SFP (1.1)-(1.2). Let W be a bounded subset of $\mathbb{R}^n$ with $W \cap S \neq \varnothing$ and $\delta > 0$. Since $q \circ A$ is a convex quadratic function and its minimum value on C is 0 (due to (4.4)), it follows from [29, Lemma 2.3] (with $q \circ A$, C, S in place of f, X, X^*) that there exists a constant $\kappa_1 > 0$ such that

$$\mathrm{d}_S(x) \leq \kappa_1 \sqrt{q(Ax)} \qquad \text{for all} \quad x \in C \quad \text{with} \quad q(Ax) \leq \delta. \tag{4.5}$$

Noting that $q \circ A$ is Lipschtiz continuous on any bounded subset, there exists a constant $\kappa_2 > 0$ such that

$$q(A(y)) - q(A(x)) \leq \kappa_2 \|y - x\| \quad \text{for each } x, y \in C \cup W. \tag{4.6}$$

Since C is a polyhedron, it follows from Hoffman theorem [15, Theorem] that there exists a constant $\kappa_3 > 0$ such that

$$\mathrm{d}_C(x) \leq \kappa_3 [c(x)]_+ \quad \text{for any} \quad x \in \mathbb{R}^n. \tag{4.7}$$

Let $\bar{x} \in W \cap S$ and set

$$\rho_1 := \sup_{x \in W} \|x - \bar{x}\|, \quad \kappa := \max\{\kappa_1, \frac{\rho_1}{\sqrt{\delta}}\}, \quad \rho_2 := \sup_{x \in W} \sqrt{\mathrm{d}_C(x)} \tag{4.8}$$

and

$$\gamma_W := (\frac{\rho_2}{\sqrt{\kappa_2}} + \kappa)\sqrt{1 + \kappa_2 \kappa_3}. \tag{4.9}$$

To complete the proof, it suffices to show that (3.1) holds with $\tau = \frac{1}{2}$ and γ_W given by (4.9). To do this, let $x \in W$. By the triangle inequality, one has that

$$\mathrm{d}_S(x) \leq \mathrm{d}_C(x) + \mathrm{d}_S(P_C(x)). \tag{4.10}$$

If $q(AP_C(x)) > \delta$, then it follows from the definition of ρ_1 in (4.8) that

$$\mathrm{d}_S(P_C(x)) \leq \| P_C(x) - P_C(\bar{x}) \| \leq \| x - \bar{x} \| \leq \frac{\| x - \bar{x} \|}{\sqrt{\delta}} \sqrt{q(AP_C(x))} \leq \frac{\rho_1}{\sqrt{\delta}} \sqrt{q(AP_C(x))};$$

otherwise, it follows from (4.5) that $\mathrm{d}_S(P_C(x)) \leq \kappa_1 \sqrt{q(AP_C(x))}$. This, together with the definition of κ in (4.8), gives that

$$\mathrm{d}_S(P_C(x)) \leq \kappa \sqrt{q(AP_C(x))}. \tag{4.11}$$

Note by (4.6) that

$$q(AP_C(x)) \leq q(A(x)) + \kappa_2 \mathrm{d}_C(x) \leq [q(A(x))]_+ + \kappa_2 \mathrm{d}_C(x).$$

Combining this with (4.11), (4.10) and (4.7) yields that

$$\begin{aligned} \mathrm{d}_S(x) &\leq \mathrm{d}_C(x) + \kappa \sqrt{[q(A(x))]_+ + \kappa_2 \mathrm{d}_C(x)} \\ &\leq (\sqrt{\mathrm{d}_C(x)/\kappa_2} + \kappa) \sqrt{[q(A(x))]_+ + \kappa_2 \kappa_3 [c(x)]_+}. \end{aligned}$$

Thus by the definition of γ_W in (4.9), (3.1) is seen to hold with $\tau = \frac{1}{2}$.

Four experiments are performed to show the comparison of the convergence results between Algorithm 1 and the RSSEA in [34] for the compressed sensing problem. In each experiment, we consider the synthetic data described in [17] for compressed sensing problem. Each entry in matrix $A \in \mathbb{R}^{m \times n}$ is randomly and independently generated from a Gaussian distribution, where $A^\top A = I$ is satisfied. The true sparse solution $\bar{x} \in \mathbb{R}^n$ has s nonzero elements drawn independently from a Gaussian distribution, and t is obtained by $t = \|\bar{x}\|_1$. The observation b is generated by $b = Ax$.

The dimension of problem is set as $m = 256$ and $n = 1024$, and we select initial point $x_0 = \mathbf{0}$ and set $\varepsilon = 10^{-6}$. To evalute the performance of algorithms, we compute the total violation by

$$\text{Total violation} := [\|x\|_1 - t]_+ + [\|Ax - b\|_2 - \epsilon]_+.$$

All numerical experiments are implemented in R (3.5.2) on personal desktop (Intel Core i7-10510U, 2.30GHz, 16GB of RAM).

The first experiment shows the convergence results of Algorithm 1 with constant stepsizes $\beta_n \equiv 0.9$. We conduct 100 trials with randomly simulated data to show the convergence property of Algorithm 1. See Figure 1.

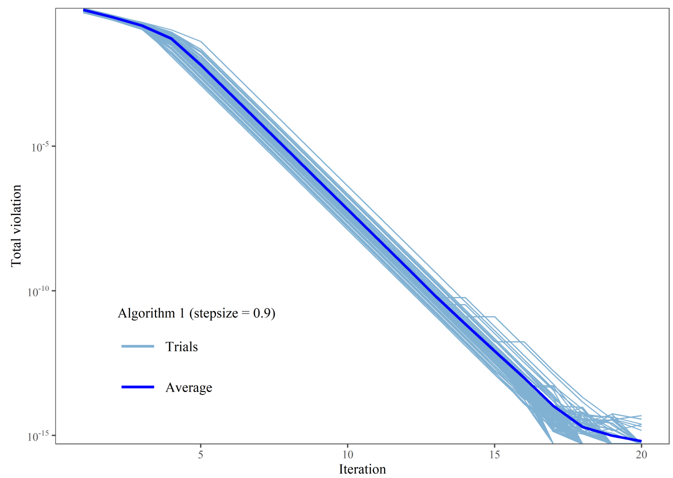
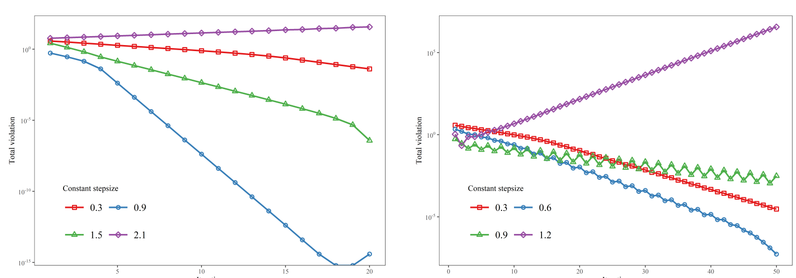


Figure 1: Convergence results of Algorithm 1 after 100 trials.

The second experiment illustrates the choice of stepsize for Algorithm 1 and the RSSEA. We consider different constant stepsizes $\beta_n \in \{0.3, 0.9, 1.5, 2.1\}$ for Algorithm 1 and $\beta_n \in \{0.3, 0.6, 0.9, 1.2\}$ for the RSSEA. Part (a) of Figure 2 indicates that Algorithm 1 converges when $0 < \beta_n \leq 2$ and stepsize $\beta_n \equiv 0.9$ provides a faster convergence than others. Moreover, part (b) of Figure 2 shows that the RSSEA converges when $0 < \beta_n \leq 1$ and stepsize $\beta_n \equiv 0.6$ provides a better convergence than others.



(a) Algorithm 1 (b) RSSEA

Figure 2: Convergence results of Algorithm 1 and the RSSEA with different stepsizes.

The third experiment compares the convergence results of Algorithm 1 and the RSSEA. We consider the constant stepsize $\beta_n \equiv 0.9$ for Algorithm 1 and $\beta_n \equiv 0.6$ for the RSSEA. The results show that Algorithm 1 owns faster convergence than the RSSEA based on total violation. See Figure 3.

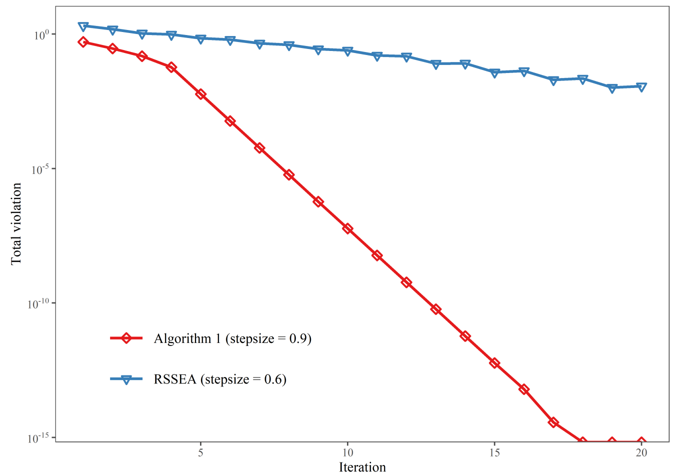


Figure 3: Convergence results of Algorithm 1 and the RSSEA.

The last experiment is implemented to compare the performance of Algorithm 1 and the RSSEA with different sparsity. We consider the constant stepsize $\beta_n \equiv 0.9$ for Algorithm 1 and $\beta_n \equiv 0.6$ for the RSSEA in this experiment, and we set the maximum number of iterations as 50. Figure 4 shows the performance of Algorithm 1 and the RSSEA based on 100 random trials, and it indicated that Algorithm 1 performs better than the RSSEA with different sparsity levels.

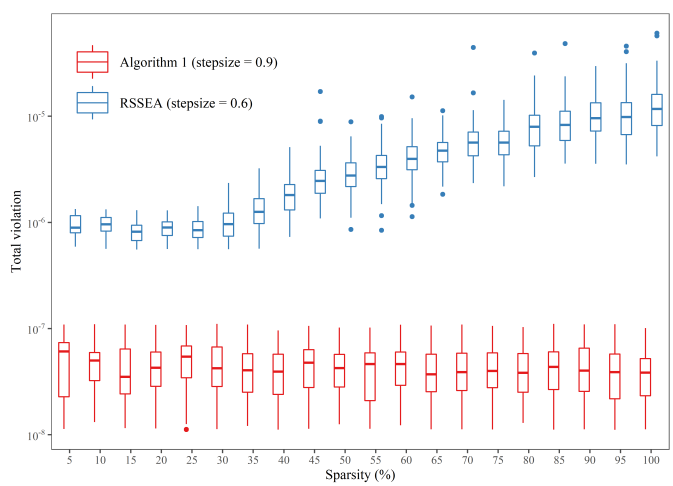


Figure 4: Performance of Algorithm 1 and the RSSEA with different sparsity.

5 Conclusion

Under the Hölderian type bounded error bound property, we established strong convergence with estimates on the convergence rate of the relaxed CQ algorithm in Hilbert spaces. In particular, for the case when the involved exponent is equal to 1, the linear convergence of the relaxed CQ algorithm was established. Numerical experiments were performed to show the convergence property of the relaxed CQ algorithm for the compressed sensing problem.

References

[1] T. Aspelmeier, C. Charitha and D. R. Luke. Local linear convergence of the ADMM/Douglas-Rachford algorithms without strong convexity and application to statistical imaging. *SIAM J. Imaging Sci.*, 9(2): 842–868, 2016.

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

[3] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer Science + Business Media, 2011.

[4] H. H. Bauschke, D. Noll and H. M. Phan. Linear and strong convergence of algorithms involving averaged nonexpansive operators. *J. Math. Anal. Appl.*, 421: 1–20, 2015.

[5] J. M. Borwein, G. Li and L. Yao. Analysis of the convergence rate for the cyclic projection algorithm applied to basic semialgebraic convex sets. *SIAM J. Optim.*, 24(1): 498–527, 2014.

[6] J. V. Burke and S. Deng. Weak sharp minima revisited. I. Basic theory. *Control Cybernet.*, 31(3): 439–469, 2002. Well-posedness in optimization and related topics (Warsaw, 2001).

[7] J. M. Borwein, G. Li and M. K. Tam. Convergence rate analysis for averaged fixed point iterations in common fixed point problems. *SIAM J. Optim.*, 27(1): 1-33, 2017.

[8] C. L. Byrne. Iterative oblique projection onto convex sets and the split feasibility problem. *Inverse Problems*, 18(2): 441–453, 2002.

[9] C. L. Byrne and A. Moudafi. Extensions of the CQ algorithm for the split feasibility and split equality problems. *J. Nonlinear Convex Anal.*, 18(8): 1485–1496, 2017.

[10] Y. Censor and T. Elfving. A multiprojection algorithm using Bregman projections in a product space. *Numer. Algorithms*, 8(2-4): 221–239, 1994.

[11] Y. Censor, T. Elfving, N. Kopf and T. Bortfeld. The multiple-sets split feasibility problem and its applications for inverse problems. *Inverse Problems*, 21(6): 2071–2084, 2005.

[12] E. J. Candes and T. Tao Decoding by linear programming. *IEEE Trans. Inform. Theory*, 51(12): 4203–4215, 2005.

[13] M. Fukushima. A relaxed projection method for variational inequalities. *Math. Program.*, 35(1): 58–70, 1986.

[14] T. Gao, C. Li, J. Wang and X. Yang. Relative regularity conditions and bounded linear regularity properties for split feasibility problems in normed linear spaces. *submitted.*

[15] A. J. Hoffman. On approximate solutions of systems of linear inequalities. *J. Research Nat. Bur. Standards*, 49: 263–265, 1952.

[16] Y. Hu, C. Li and X. Yang. On convergence rates of linearized proximal algorithms for convex composite optimization with applications. *SIAM J. Optim.*, 26(2): 1207–1235, 2016.

[17] Y. Hu, C. Li, K. Meng, J. Qin and X. Yang. Group sparse optimization via $\ell p, q$ regularization. *The Journal of Machine Learning Research*, 18(1):960–1011, 2017.

[18] A. Kruger, M. Lopez, X. Yang and J. Zhu. Hölder error bounds and Hölder calmness with applications to convex semi-infinite optimization. *Set-Valued Var. Anal.*, 27(4): 995–1023, 2019.

[19] D. Klatte. Hoffman's error bound for systems of convex inequalities. In *Mathematical programming with data perturbations*, volume 195 of *Lecture Notes in Pure and Appl. Math.*, pages 185–199. Dekker, New York, 1998.

[20] A. S. Lewis and J.-S. Pang. Error bounds for convex inequality systems. In *Generalized convexity, generalized monotonicity: recent results (Luminy, 1996)*, volume 27 of *Nonconvex Optim. Appl.*, pages 75–110. Kluwer Acad. Publ., Dordrecht, 1998.

[21] G. Li. On the asymptotically well behaved functions and global error bound for convex polynomials. *SIAM J. Optim.*, 20(4): 1923–1943, 2010.

[22] C. Li, L. Meng, L. Peng, Y. Hu and J.-C. Yao. Weak sharp minima for convex infinite optimization problems in normed linear spaces. *SIAM J. Optim.*, 28(3): 1999–2021, 2018.

[23] G. Li and B. S. Mordukhovich. Hölder metric subregularity with applications to proximal point method. *SIAM J. Optim.*, 22(4): 1655–1684, 2012.

[24] G. Li, B. S. Mordukhovich, T. T. A. Nghia and T. S. Phạm. Error bounds for parametric polynomial systems with applications to higher-order stability analysis and convergence rates. *Math. Program.*, 168(1-2, Ser. B): 313–346, 2018.

[25] G. Li, B. S. Mordukhovich and T. S. Phạm. New fractional error bounds for polynomial systems with applications to Hölderian stability in optimization and spectral theory of tensors. *Math. Program.*, 153(2, Ser. A): 333–362, 2015.

[26] G. López, V. Martín-Márquez, F. Wang and H.-K. Xu. Solving the split feasibility problem without prior knowledge of matrix norms. *Inverse Problems*, 28(8): 085004, 18, 2012.

[27] X.-D. Luo and Z.-Q. Luo. Extension of Hoffman's error bound to polynomial systems. *SIAM J. Optim.*, 4(2): 383–392, 1994.

[28] Z.-Q. Luo and P. Tseng. Error bounds and convergence analysis of feasible descent methods: a general approach. *Ann. Oper. Res.*, 46/47: 157–178, 1993. Degeneracy in optimization problems.

[29] W. Li. Error bounds for piecewise convex quadratic programs and applications. *SIAM J. Control Optim.*, 33(5):1510–1529, 1995.

[30] H. V. Ngai and M. Théra. Error bounds for systems of lower semicontinuous functions in Asplund spaces. *Math. Program.*, 116(1-2, Ser. B): 397–427, 2009.

[31] J.-S. Pang. Error bounds in mathematical programming. *Math. Program.*, 79(1-3, Ser. B): 299–332, 1997. Lectures on mathematical programming (ismp97) (Lausanne, 1997).

[32] B. T. Polyak. *Introduction to Optimization.* Nauka, Moscow, 1983.

[33] H. Stark, editor. *Image recovery: theory and application.* Academic Press, Inc., Orlando, FL, 1987.

[34] T. Tian, L. Shi and R. Chen. Linear convergence of the relaxed gradient projection algorithm for solving the split equality problems in Hilbert spaces. *J. Inequal. Appl.*, 2019(1): 1–12, 2019.

[35] J. Wang, Y. Hu, C. Li and J.-C. Yao. Linear convergence of *cq* algorithms and applications in gene regulatory network inference. *Inverse Problems*, 33(5): 055017, 25, 2017.

[36] Z. Wu and J. J. Ye. On error bounds for lower semicontinuous functions. *Math. Program.*, 92(2, Ser. A): 301–314, 2002.

[37] H.-K. Xu. Iterative methods for the split feasibility problem in infinite-dimensional Hilbert spaces. *Inverse Problems*, 26(10): 105018, 17, 2010.

[38] Q. Yang. The relaxed CQ algorithm solving the split feasibility problem. *Inverse Problems*, 20(4): 1261–1266, 2004.

[39] Q. Yang. On variable-step relaxed projection algorithm for variational inequalities. *J. Math. Anal. Appl.*, 302(1): 166–179, 2005.