axioms

Article

Bounded Perturbation Resilience of Two Modified Relaxed CQ Algorithms for the Multiple-Sets Split Feasibility Problem

Yingying Li and Yaxuan Zhang *

College of Science, Civil Aviation University of China, Tianjin 300300, China; yingyl99@163.com
* Correspondence: bunnyxuan@tju.edu.cn

Abstract: In this paper, we present some modified relaxed CQ algorithms with different kinds of step size and perturbation to solve the Multiple-sets Split Feasibility Problem (MSSFP). Under mild assumptions, we establish weak convergence and prove the bounded perturbation resilience of the proposed algorithms in Hilbert spaces. Treating appropriate inertial terms as bounded perturbations, we construct the inertial acceleration versions of the corresponding algorithms. Finally, for the LASSO problem and three experimental examples, numerical computations are given to demonstrate the efficiency of the proposed algorithms and the validity of the inertial perturbation.

Keywords: multiple-sets split feasibility problem; CQ algorithm; bounded perturbation; armijo-line search; self-adaptive step size

MSC: 47J20; 47J25; 49J40

check for updates

Citation: Li, Y.; Zhang, Y. Bounded Perturbation Resilience of Two Modified Relaxed CQ Algorithms for the Multiple-Sets Split Feasibility Problem. *Axioms* **2021**, *10*, 197. https://doi.org/10.3390/axioms10030197

Academic Editor: Chris Goodrich

Received: 16 July 2021
Accepted: 20 August 2021
Published: 23 August 2021

1. Introduction

In this paper, we focus on the Multiple-sets Split Feasibility Problem (MSSFP), which is formulated as follows.

$$\text{Find a point } x^* \in C = \bigcap_{i=1}^{t} C_i \text{ such that } Ax^* \in Q = \bigcap_{j=1}^{r} Q_j, \tag{1}$$

where $A : \mathcal{H}_1 \to \mathcal{H}_2$ is a bounded and linear operator, $C_i \subset \mathcal{H}_1$, $i = 1, \cdots, t$, and $Q_j \subset \mathcal{H}_2$, $j = 1, \cdots, r$ are nonempty closed and convex sets, and $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces. When $t = 1$, $r = 1$, it is the Split Feasibility Problem (SFP). Byrne in [1,2] introduced the following CQ algorithm to solve the SFP,

$$x^{k+1} = P_C(x^k - \alpha_k A^*(I - P_Q)Ax^k), \tag{2}$$

where $\alpha_k \in (0, \frac{2}{\|A\|^2})$. It is proven that the iterates $\{x^k\}$ converge to a solution of the SFP. When P_C and P_Q have explicit expressions, the CQ algorithm is easy to carry out. However, P_C and P_Q have no explicit formulas in general; thus the computation of P_C and P_Q is itself an optimization problem.

To avoid the computation of P_C and P_Q, Yang [3] proposed the relaxed CQ algorithm in finite dimensional spaces. The algorithm is

$$x^{k+1} = P_{C^k}(x^k - \alpha_k A^*(I - P_{Q^k})Ax^k), \tag{3}$$

where $\alpha_k \in (0, \frac{2}{\|A\|^2})$, C^k and Q^k are sequences of closed half spaces containing C and Q, respectively.

As for the MSSFP (1), Censor et al. in [4] proposed the following algorithm,

$$x^{k+1} = P_\Omega(x^k - \alpha \nabla p(x^k)), \tag{4}$$

where Ω is an auxiliary closed subset, and $p(x)$ is a function to measure the distance from a point to all the sets C_i and Q_j,

$$p(x) = \frac{1}{2}\sum_{i=1}^{t} \lambda_i \|x - P_{C_i}(x)\|^2 + \frac{1}{2}\sum_{j=1}^{r} \beta_j \|Ax - P_{Q_j}(Ax)\|^2, \tag{5}$$

where $\lambda_i > 0$, $\beta_j > 0$ for every i and j, and $\sum_{i=1}^{t} \lambda_i + \sum_{j=1}^{r} \beta_j = 1$, $0 < \alpha < 2/L$, $L = \sum_{i=1}^{t} \lambda_i + \|A\|^2 \sum_{j=1}^{r} \beta_j$. The convergence of the algorithm (4) is proved in finite dimensional spaces.

Later, He et al. [5] introduced a relaxed self-adaptive CQ algorithm,

$$x^{k+1} = \tau_k \mu + (1 - \tau_k)(x^k - \alpha_k \nabla p_k(x^k)), \tag{6}$$

where the sequence $\{\tau_k\} \subset (0,1)$, $\mu \in \mathcal{H}$, $p_k(x) = \frac{1}{2}\sum_{i=1}^{t} \lambda_i \|x - P_{C_i^k}(x)\|^2 + \frac{1}{2}\sum_{j=1}^{r} \beta_j \|Ax - P_{Q_j^k}(Ax)\|^2$, where the closed convex sets C_i^k and Q_j^k are level sets of some convex functions containing C_i and Q_j, and self-adaptive step size $\alpha_k = \frac{\rho_k p_k(x^k)}{\|\nabla p_k(x^k)\|^2}$, $0 < \rho_k < 4$. They proved that the sequence $\{x^k\}$ generated by algorithm (6) converges in norm to $P_S(\mu)$, where S is the solution set of the MSSFP.

In order to improve the rate of convergence, many scholars have investigated the choice of the step size of the algorithms. Based on the CQ algorithm (2), Yang [6] proposed the step size

$$\alpha_k = \frac{\rho_k}{\|\nabla f(x^k)\|},$$

where $\{\rho_k\}$ is a sequence of positive real numbers satisfying $\sum_{n=0}^{\infty} \rho_k = \infty$ and $\sum_{n=0}^{\infty} \rho_k^2 < +\infty$, and $f(x) = \frac{1}{2}\|(I - P_Q)Ax\|^2$. Assuming that Q is bounded and A is a matrix with full column rank, Yang proved the convergence of the underlying algorithm in finite dimensional spaces. In 2012, López et al. [7] introduced another choice of the step size sequence $\{\alpha_k\}$ in the algorithm (3) as follows

$$\alpha_k = \frac{\rho_k f_k(x^k)}{\|\nabla f_k(x^k)\|^2},$$

where $0 < \rho_k < 4$, $f_k(x) = \frac{1}{2}\|(I - P_{Q^k})Ax\|^2$, and they proved the weak convergence of the iteration sequence in Hilbert spaces. The advantage of this choice of the step size lies in the fact that neither prior information about the matrix norm A nor any other conditions on Q and A are required. Recently, Gibali et al. [8] and Chen et al. [9] used step size determined by Armijo-line search and proved the convergence of the algorithm. For more information on the relaxed CQ algorithm and the selection of step size, please refer to references [10–12].

On the other hand, in order to make the algorithms converge faster, specific perturbations have been introduced into the iterative format, since the perturbations guide the iteration to a lower objective function value without losing the overall convergence. So far, bounded perturbation recovery has been used in many problems.

Consider the usage of the bounded perturbation for the non-smooth optimization problems, $\min_{x \in H} \phi(x) = f(x) + g(x)$, where f and g are proper lower semicontinuous convex functions in real Hilbert spaces, f is differentiable, g is not necessarily differentiable, and ∇f is L-Lipschitz continuous. One of the classic algorithms is the proximal gradient (PG) algorithm, based on which Guo et al. [13] proposed the following PG algorithm with perturbations,

$$x^{k+1} = \mathrm{prox}_{\lambda_k g}(I - \lambda_k D \nabla f + e)(x^k). \tag{7}$$

Assume that (i) D is a bounded linear operator, (ii) $0 < \inf \lambda_k \leq \lambda_k \leq \sup \lambda_k < \frac{2}{L}$, (iii) $e(x^k)$ satisfies $\sum_{k=0}^{\infty} \|e(x^k)\| < +\infty$, and (iv) $\theta_k = \nabla f(x^k) - D(x^k)\nabla f(x^k)$ satisfies $\sum_{k=0}^{\infty} \|\theta_k\| < +\infty$. They asserted that the generated sequence $\{x^k\}$ converges weakly to a solution. Later, Guo and Cui [14] proposed the modified PG algorithm for solving this problem,

$$x^{k+1} = \tau_k h(x^k) + (1-\tau_k)\text{prox}_{\lambda_k g}(I - \lambda_k \nabla f)(x^k) + e(x^k), \tag{8}$$

where $\tau_k \subset [0,1]$, h is a $\rho \in (0,1)$-contractive operator. They proved that the sequence $\{x^k\}$ generated by the algorithm (8) converges strongly to a solution x^*. In 2020, Pakkaranang et al. [15] considered PG algorithm combined with inertial technique

$$\begin{cases} y^k = x^k + \theta_k(x^k - x^{k-1}), \\ x^{k+1} = \tau_k h(y^k) + (1-\tau_k)\text{prox}_{\lambda_k g}(I - \lambda_k \nabla f)(y^k) + e(y^k), \end{cases} \tag{9}$$

and they proved its strong convergence under suitable conditions.

For the convex minimization problem, $\min_{x\in\Omega} f(x)$, where Ω is a nonempty closed convex subset in finite dimensional space and the objective function f is convex, Jin et al. [16] presented the following projected scaled gradient (PSG) algorithm with errors

$$x^{k+1} = P_\Omega(x^k - \lambda_k D(x^k)\nabla f(x^k) + e(x^k)). \tag{10}$$

Assume that (i) $\{D(x^k)\}_{k=0}^{\infty}$ is a sequence of diagonal scaling matrices, and that (ii) (iii) (iv) are the same as the conditions in algorithm (7); then the generated sequence $\{x^k\}$ converges weakly to a solution.

In 2017, Xu [17] applied the superiorization techniques to the relaxed PSG. The iterative form is

$$x^{k+1} = (1-\tau_k)x^k + \tau_k P_\Omega(x^k - \lambda_k D(x^k)\nabla f(x^k) + e(x^k)), \tag{11}$$

where τ_k is a sequence in $[0,1]$, and $D(x^k)$ is a diagonal scaling matrix. He established weak convergence of the above algorithm under appropriate conditions imposed on $\{\tau_k\}$ and $\{\lambda_k\}$.

For the variational inequality problem (VIP for short) $\langle F(x^*), x - x^* \rangle \geq 0$, $\forall x \in C$, where F is a nonlinear operator, Dong et al. [18] considered the external gradient algorithm with perturbations

$$\begin{cases} \bar{x}^k = P_C(x^k - \alpha_k F(x^k) + e_1(x^k)), \\ x^{k+1} = P_C(x^k - \alpha_k F(\bar{x}^k) + e_2(x^k)). \end{cases} \tag{12}$$

where $\alpha_k = \gamma l^{m_k}$ with m_k the smallest non-negative integer such that

$$\alpha_k \|F(x^k) - F(\bar{x}^k)\| \leq \mu \|x^k - \bar{x}^k\|.$$

Assume that F is monotonous and L-Lipschitz is continuous and that the error sequence is summable; the sequence $\{x^k\}$ generated by the algorithm converges weakly to a solution of the VIP.

For the split variational inclusion problem, Duan and Zheng [19] in 2020 proposed the following algorithm

$$x^{k+1} = \tau_k h(x^k) + (1-\tau_k)J_\gamma^{B_1}(I - \lambda_k A^*(I - J_\gamma^{B_2})A)(x^k) + e(x^k), \tag{13}$$

where A is a bounded linear operator, B_1 and B_2 are maximal monotone operators. Assuming that $\lim_{k\to\infty} \tau_k = 0$, $\sum_{k=0}^{\infty} \tau_k = \infty$, $0 < \inf_{k\to\infty} \lambda_k \leq \sup_{k\to\infty} \lambda_k < \frac{2}{L}$, $L = \|A\|^2$ and $\sum_{k=0}^{\infty} \|e(x^k)\| < +\infty$, they proved that the sequence $\{x^k\}$ strongly converges to a solution of the split variational inclusion problem, which is also the unique solution of some variational inequality problem.

For the convex feasibility problem, Censor and Zaslavski [20] considered the perturbation resilience and convergence of dynamic string-averaging projection method.

Adding an inertial term can improve the convergence rate, which is also a perturbation. Recently, for a common solution of the split minimization problem and the fixed point problem, Kaewyong and Sitthithakerngkiet [21] combined the proximal algorithm and a modified Mann's iterative method with the inertial extrapolation and improved related results. Shehu et al. [22] and Li et al. [23] added alternated inertial perturbation to the algorithms for solving the SFP and improved the convergence rate.

At present, the (multiple-sets) split feasibility problem is widely used in application fields, such as CT tomography, image restoration, and image reconstruction, etc. There are many related literatures on the iterative algorithms for solving the (multiple-sets) split feasibility problem. However, there are relatively fewer documents studying the algorithms of the (multiple-sets) split feasibility problem with perturbations, especially with self-adaptive step size. In fact, the latter also has a bounded disturbance recovery property. Motivated by [9,18], we focus on the modified relaxed CQ algorithms to solve the MSSFP (1) in real Hilbert spaces and assert that the proposed algorithms are also bounded-perturbation-resilient.

The rest of the paper is arranged as follows. In Section 2, definitions and notions that will be useful for our analysis are presented. In Section 3, we present our algorithms and prove their weak convergence. In Section 4, we prove that the proposed algorithms have bounded perturbation resilience and construct the inertial modification of the algorithms. Furthermore, finally, in Section 5, we present some numerical simulations to show the validity of the proposed algorithms.

2. Preliminaries

In this section, we first define some symbols and then review some definitions and basic results that will be used in this paper.

Throughout this paper, $\mathcal{H}$ denotes a real Hilbert space endowed with an inner product $\langle \cdot, \cdot \rangle$ and its deduced norm $\| \cdot \|$, and I is the identity operator on $\mathcal{H}$. We denote by S the solution set of the MSSFP (1). Moreover, $x^k \to x$ ($x^k \rightharpoonup x$) represents that the sequence $\{x^k\}$ converges strongly (weakly) to x. Finally, we denote by $\omega_\omega(x^k)$ all the weak cluster points of $\{x^k\}$.

An operator $T : \mathcal{H} \to \mathcal{H}$ is said to be nonexpansive if for all $x, y \in \mathcal{H}$,

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

$T : \mathcal{H} \to \mathcal{H}$ is said to be firmly nonexpansive if for all $x, y \in \mathcal{H}$,

$$||Tx - Ty||^2 \leq ||x - y||^2 - ||(I - T)x - (I - T)y||^2,$$

or equivalently

$$||Tx - Ty||^2 \leq \langle Tx - Ty, x - y \rangle.$$

It is well known that T is firmly nonexpansive if and only if $I - T$ is firmly nonexpansive.

Let C be a nonempty closed convex subset of $\mathcal{H}$. Then the metric projection P_C from $\mathcal{H}$ onto C is defined as

$$P_C(x) = \operatorname*{argmin}_{y \in C} ||x - y||^2, \quad x \in \mathcal{H}.$$

The metric projection P_C is a firmly nonexpansive operator.

Definition 1 ([24]). *A function $f : \mathcal{H} \to \mathbf{R}$ is said to be weakly lower semicontinuous at $\hat{x}$ if x^k converges weakly to $\hat{x}$ implies*

$$f(\hat{x}) \le \liminf_{k\to\infty} f(x^k).$$

Definition 2. *If $\varphi : \mathcal{H} \to \mathbf{R}$ is a convex function, the subdifferential of φ at x is defined as*

$$\partial\varphi(x) = \{\xi \in \mathcal{H} \mid \varphi(y) \ge \varphi(x) + \langle \xi, y - x\rangle,\ \forall y \in \mathcal{H}\}.$$

Lemma 1 ([24]). *Let C be a nonempty closed and convex subset of $\mathcal{H}$; then for any x, $y \in \mathcal{H}$, $z \in C$, the following assertions hold:*
(i) $\langle x - P_C x, z - P_C x\rangle \le 0$;
(ii) $\|P_C x - z\|^2 \le \|x - z\|^2 - \|P_C x - x\|^2$;
(iii) $2\langle x, y\rangle \le \|x\| + \|x\|\|y\|^2$;
(iv) $2\langle x, y\rangle \le \|x\|^2 + \|y\|^2$.

Lemma 2 ([25]). *Assume that $\{a^k\}_{k=0}^{\infty}$ is a sequence of nonnegative real numbers such that*

$$a^{k+1} \le (1 + \sigma_k)a^k + \delta_k, \forall k \ge 0,$$

where the nonnegative sequences $\{\sigma_k\}_{k=0}^{\infty}$ and $\{\delta_k\}_{k=0}^{\infty}$ satisfies $\sum_{k=0}^{\infty} \sigma_k < +\infty$ and $\sum_{k=0}^{\infty} \delta_k < +\infty$, respectively. Then $\lim_{k\to\infty} a^k$ exists.

Lemma 3 ([25]). *Let S be a nonempty closed and convex subset of $\mathcal{H}$ and $\{x^k\}$ be a sequence in $\mathcal{H}$ that satisfies the following properties:*
(i) $\lim_{k\to\infty} \|x^k - x\|$ *exists for each* $x \in S$;
(ii) $\omega_\omega(x^k) \subset S$.
Then $\{x^k\}$ converges weakly to a point in S.

Definition 3. *An algorithmic operator P is said to be bounded perturbations resilient if the iteration $x^{k+1} = P(x^k)$ and $x^{k+1} = P(x^k + \lambda_k v_k)$ all converge, where $\{\lambda_k\}$ is a sequence of nonnegative real numbers, $\{v_k\}$ is a sequence in $\mathcal{H}$, and $M \in \mathbf{R}$ and satisfies*

$$\sum_{k=0}^{\infty} \lambda_k < +\infty, \|v_k\| \le M.$$

3. Algorithms and Their Convergence

In this section, we introduce two algorithms of the MSSFP (1) and prove their weak convergence. First assume that the following four assumptions hold.

(A1) The solution set S of the MSSFP (1) is nonempty.

(A2) The level sets of convex functions can be expressed by

$$C_i = \{x \in \mathcal{H}_1 \mid c_i(x) \le 0\} \quad \text{and} \quad Q_j = \{y \in \mathcal{H}_2 \mid q_j(y) \le 0\},$$

where $c_i : \mathcal{H}_1 \to \mathbf{R}$ $(i = 1, \cdots, t)$ and $q_j : \mathcal{H}_2 \to \mathbf{R}$ $(j = 1, \cdots, r)$ are weakly lower semicontinuous and convex functions.

(A3) For any $x \in \mathcal{H}_1$ and $y \in \mathcal{H}_2$, at least one subgradient $\xi_i \in \partial c_i(x)$ and $\eta_j \in \partial q_j(y)$ can be calculated. The subdifferential ∂c_i and ∂q_j are bounded on the bounded sets.

(A4) The sequences of perturbations $\{e_i(x^k)\}_{k=0}^{\infty}$ $(i = 1, 2, 3)$ is summable, i.e.,

$$\sum_{k=0}^{\infty} \|e_i(x^k)\| < +\infty.$$

Define two sets at point x^k by

$$C_i^k = \{x \in \mathcal{H}_1 \mid c_i(x^k) + \langle \xi_i^k, x - x^k \rangle \leq 0\},$$

and

$$Q_j^k = \{y \in \mathcal{H}_2 \mid q_j(Ax^k) + \langle \eta_j^k, y - Ax^k \rangle \leq 0\},$$

where $\xi_i^k \in \partial c_i(x^k)$ and $\eta_j^k \in \partial q_j(Ax^k)$. Define the function f_k by

$$f_k(x) = \frac{1}{2} \sum_{j=1}^{r} \beta_j \|(I - P_{Q_j^k})Ax\|^2, \tag{14}$$

where $\beta_j > 0$. Then it is easy to verify that the function $f_k(x)$ is convex and differentiable with gradient

$$\nabla f_k(x) = \sum_{j=1}^{r} \beta_j A^*(I - P_{Q_j^k})Ax, \tag{15}$$

and the L-Lipschitz constant of $\nabla f_k(x)$ is $L = \|A\|^2 \sum_{j=1}^{r} \beta_j$.

We see that C_i^k $(i = 1, \cdots, t)$ and Q_j^k $(j = 1, \cdots, r)$ are half spaces such that $C_i \subset C_i^k$, $Q_j \subset Q_j^k$, for all $k \geq 1$. We now present Algorithm 1 with Armijo-line search step size.

Algorithm 1 (The relaxed CQ algorithm with Armijo-line search and perturbation)

Given constant $\gamma > 0$, $l \in (0,1)$, $\mu \in (0,1)$. Let x^0 be arbitrarily chosen, for $k = 0, 1, \cdots$, compute

$$\bar{x}^k = P_{C_{[k]}^k}(x^k - \alpha_k \nabla f_k(x^k) + e_1(x^k)), \tag{16}$$

where $[k] = k \bmod t$ and $\alpha_k = \gamma l^{m_k}$ with m_k the smallest non-negative integer such that

$$\alpha_k \|\nabla f_k(x^k) - \nabla f_k(\bar{x}^k)\| \leq \mu \|x^k - \bar{x}^k\|. \tag{17}$$

Construct the next iterate x^{k+1} by

$$x^{k+1} = P_{C_{[k]}^k}(x^k - \alpha_k \nabla f_k(\bar{x}^k) + e_2(x^k)). \tag{18}$$

Lemma 4 ([6])**.** *The Armijo-line search terminates after a finite number of steps. In addition,*

$$\frac{\mu l}{L} < \alpha_k \leq \gamma, \ for\ all\ \ k \geq 0. \tag{19}$$

where $L = \|A\|^2 \sum_{j=1}^{r} \beta_j$.

The weak convergence of Algorithm 1 is established below.

Theorem 1. *Let $\{x^k\}$ be the sequence generated by Algorithm 1, and the assumptions (A1)∼(A4) hold. Then $\{x^k\}$ converges weakly to a solution of the MSSFP (1).*

Proof. Let x^* be a solution of the MSSFP. Note that $C \subset C_i \subset C_i^k$, $Q \subset Q_j \subset Q_j^k$, $i = 1, \cdots, t$, $j = 1, \cdots, r$, $k = 0, 1, \cdots$, so $x^* = P_C(x^*) = P_{C_i}(x^*) = P_{C_i^k}(x^*)$ and $Ax^* = P_Q(Ax^*) = P_{Q_j}(Ax^*) = P_{Q_j^k}(Ax^*)$, and thus $f_k(x^*) = 0$ and $\nabla f_k(x^*) = 0$.

First, we prove that $\{x^k\}$ is bounded. Following Lemma 1 (ii), we have

$$\begin{aligned}
& \|x^{k+1} - x^*\|^2 \\
= & \|P_{C^k_{[k]}}(x^k - \alpha_k \nabla f_k(\bar{x}^k) + e_2(x^k)) - x^*\|^2 \\
\leq & \|x^k - \alpha_k \nabla f_k(\bar{x}^k) + e_2(x^k) - x^*\|^2 - \|x^{k+1} - x^k + \alpha_k \nabla f_k(\bar{x}^k) - e_2(x^k)\|^2 \\
= & \|x^k - x^*\|^2 - \|x^{k+1} - x^k\|^2 - 2\langle \alpha_k \nabla f_k(\bar{x}^k) - e_2(x^k), x^k - x^* \rangle \\
& -2\langle \alpha_k \nabla f_k(\bar{x}^k) - e_2(x^k), x^{k+1} - x^k \rangle \\
= & \|x^k - x^*\|^2 - \|x^{k+1} - x^k\|^2 - 2\langle \alpha_k \nabla f_k(\bar{x}^k) - e_2(x^k), x^{k+1} - x^* \rangle \\
= & \|x^k - x^*\|^2 - \|x^{k+1} - x^k\|^2 - 2\langle \alpha_k \nabla f_k(\bar{x}^k), x^{k+1} - x^* \rangle \\
& +2\langle e_2(x^k), x^{k+1} - x^* \rangle \\
= & \|x^k - x^*\|^2 - \|x^{k+1} - \bar{x}^k\|^2 - \|\bar{x}^k - x^k\|^2 - 2\langle x^{k+1} - \bar{x}^k, \bar{x}^k - x^k \rangle \\
& -2\alpha_k \langle \nabla f_k(\bar{x}^k), x^{k+1} - x^* \rangle + 2\langle e_2(x^k), x^{k+1} - x^* \rangle \\
= & \|x^k - x^*\|^2 - \|x^{k+1} - \bar{x}^k\|^2 - \|\bar{x}^k - x^k\|^2 - 2\langle x^{k+1} - \bar{x}^k, \bar{x}^k - x^k \rangle \\
& -2\alpha_k \langle \nabla f_k(\bar{x}^k), x^{k+1} - \bar{x}^k \rangle - 2\alpha_k \langle \nabla f_k(\bar{x}^k), \bar{x}^k - x^* \rangle + 2\langle e_2(x^k), x^{k+1} - x^* \rangle \\
= & \|x^k - x^*\|^2 - \|x^{k+1} - \bar{x}^k\|^2 - \|\bar{x}^k - x^k\|^2 - 2\alpha_k \langle \nabla f_k(\bar{x}^k), \bar{x}^k - x^* \rangle \\
& +2\langle x^k - \bar{x}^k - \alpha_k \nabla f_k(\bar{x}^k), x^{k+1} - \bar{x}^k \rangle + 2\langle e_2(x^k), x^{k+1} - x^* \rangle. \qquad (20)
\end{aligned}$$

From Lemma 1 (iii), we have that

$$2\langle e_2(x^k), x^{k+1} - x^* \rangle \leq \|e_2(x^k)\| + \|e_2(x^k)\| \|x^{k+1} - x^*\|^2. \qquad (21)$$

Since $I - P_C$ is firmly nonexpensive, $\nabla f_k(x^*) = 0$, and Lemma 4, we get that

$$\begin{aligned}
& 2\alpha_k \langle \nabla f_k(\bar{x}^k), \bar{x}^k - x^* \rangle \\
= & 2\alpha_k \langle \sum_{j=1}^{r} \beta_j A^*(I - P_{Q^k_j}) A\bar{x}^k - \sum_{j=1}^{r} \beta_j A^*(I - P_{Q^k_j}) Ax^*, \bar{x}^k - x^* \rangle \\
= & 2\alpha_k \sum_{j=1}^{r} \beta_j \langle (I - P_{Q^k_j}) A\bar{x}^k - (I - P_{Q^k_j}) Ax^*, A\bar{x}^k - Ax^* \rangle \\
\geq & 2\frac{\mu l}{L} \sum_{j=1}^{r} \beta_j \|(I - P_{Q^k_j}) A\bar{x}^k\|^2. \qquad (22)
\end{aligned}$$

Based on the definition of $\bar{x}^k$ and Lemma 1 (i), we know that

$$\langle \bar{x}^k - x^k + \alpha_k \nabla f_k(x^k) - e_1(x^k), x^{k+1} - \bar{x}^k \rangle \geq 0. \qquad (23)$$

Note that (17), (23), and Lemma 1 (iii) yield that

$$\begin{aligned}
& 2\langle x^k - \bar{x}^k - \alpha_k \nabla f_k(\bar{x}^k), x^{k+1} - \bar{x}^k \rangle \\
\leq & 2\langle -e_1(x^k) + \alpha_k \nabla f_k(x^k) - \alpha_k \nabla f_k(\bar{x}^k), x^{k+1} - \bar{x}^k \rangle \\
= & 2\alpha_k \langle \nabla f_k(x^k) - \nabla f_k(\bar{x}^k), x^{k+1} - \bar{x}^k \rangle - 2\langle e_1(x^k), x^{k+1} - \bar{x}^k \rangle \\
\leq & 2\alpha_k \|\nabla f_k(x^k) - \nabla f_k(\bar{x}^k)\| \|x^{k+1} - \bar{x}^k\| + 2\|e_1(x^k)\| \|x^{k+1} - \bar{x}^k\| \\
\leq & 2\mu \|x^k - \bar{x}^k\| \|x^{k+1} - \bar{x}^k\| + \|e_1(x^k)\| + \|e_1(x^k)\| \|x^{k+1} - \bar{x}^k\|^2 \\
\leq & \mu \|x^k - \bar{x}^k\|^2 + \mu \|x^{k+1} - \bar{x}^k\|^2 + \|e_1(x^k)\| + \|e_1(x^k)\| \|x^{k+1} - \bar{x}^k\|^2 \\
= & \mu \|x^k - \bar{x}^k\|^2 + (\mu + \|e_1(x^k)\|) \|x^{k+1} - \bar{x}^k\|^2 + \|e_1(x^k)\|. \qquad (24)
\end{aligned}$$

From assumption (A4), we know that $\lim_{k\to\infty}\|e_i(x^k)\| = 0$, $i = 1, 2$, and thus $\forall\varepsilon > 0$, $\exists K$, it holds that $\|e_i(x^k)\| < \varepsilon$ for $k > K$. We can therefore assume $\|e_1(x^k)\| \in [0, 1-\mu-\tau)$ and $\|e_2(x^k)\| \in [0, 1/2)$ for $k \geq K$, where $\tau \in (0, 1-\mu)$. Hence, from (24), we get that

$$2\langle x^k - \bar{x}^k - \alpha_k \nabla f_k(\bar{x}^k), x^{k+1} - \bar{x}^k\rangle \leq \mu\|x^k - \bar{x}^k\|^2 + (1-\tau)\|x^{k+1} - \bar{x}^k\|^2 + \|e_1(x^k)\|. \quad (25)$$

Substituting (21), (22), and (25) into (20) yields

$$\begin{aligned}\|x^{k+1} - x^*\|^2 \leq\ & \|x^k - x^*\|^2 - (1-\mu)\|x^k - \bar{x}^k\|^2 - \tau\|x^{k+1} - \bar{x}^k\|^2 \\ & + \|e_1(x^k)\| + \|e_2(x^k)\| + \|e_2(x^k)\|\|x^{k+1} - x^*\|^2 \\ & - 2\frac{\mu l}{L}\sum_{j=1}^{r}\beta_j\|(I - P_{Q_j^k})A\bar{x}^k\|^2. \end{aligned} \quad (26)$$

Organizing the above formula we know that

$$\begin{aligned}\|x^{k+1} - x^*\|^2 \leq\ & \frac{1}{1-\|e_2(x^k)\|}\|x^k - x^*\|^2 - \frac{1-\mu}{1-\|e_2(x^k)\|}\|x^k - \bar{x}^k\|^2 \\ & - \frac{\tau}{1-\|e_2(x^k)\|}\|x^{k+1} - \bar{x}^k\|^2 + \frac{\|e_1(x^k)\| + \|e_2(x^k)\|}{1-\|e_2(x^k)\|} \\ & - \frac{2\mu l}{(1-\|e_2(x^k)\|)L}\sum_{j=1}^{r}\beta_j\|(I - P_{Q_j^k})A\bar{x}^k\|^2. \end{aligned} \quad (27)$$

Since $\|e_2(x^k)\| \in [0, 1/2)$ for $k \geq K$, we get

$$1 \leq \frac{1}{1-\|e_2(x^k)\|} \leq 1 + 2\|e_2(x^k)\| < 2. \quad (28)$$

This together with (27) shows that

$$\begin{aligned}\|x^{k+1} - x^*\|^2 \leq\ & (1 + 2\|e_2(x^k)\|)\|x^k - x^*\|^2 - (1-\mu)\|x^k - \bar{x}^k\|^2 + 2\|e_1(x^k)\| \\ & + 2\|e_2(x^k)\| - \tau\|x^{k+1} - \bar{x}^k\|^2 - 2\frac{\mu l}{L}\sum_{j=1}^{r}\beta_j\|(I - P_{Q_j^k})A\bar{x}^k\|^2 \\ \leq\ & (1 + 2\|e_2(x^k)\|)\|x^k - x^*\|^2 + 2\|e_1(x^k)\| + 2\|e_2(x^k)\|. \end{aligned} \quad (29)$$

Using Lemma 2 and assumption (A4), we know the existence of $\lim_{k\to\infty}\|x^k - x^*\|^2$ and the boundedness of $\{x^k\}_{k=0}^{\infty}$.

From (29), it follows

$$\begin{aligned} & (1-\mu)\|x^k - \bar{x}^k\|^2 + \tau\|x^{k+1} - \bar{x}^k\|^2 + 2\frac{\mu l}{L}\sum_{j=1}^{r}\beta_j\|(I - P_{Q_j^k})A\bar{x}^k\|^2 \\ \leq\ & (1 + 2\|e_2(x^k)\|)\|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2 + 2\|e_1(x^k)\| + 2\|e_2(x^k)\|, \end{aligned} \quad (30)$$

which means that

$$\sum_{k=0}^{\infty}\|x^k - \bar{x}^k\| < +\infty, \sum_{k=0}^{\infty}\|x^{k+1} - \bar{x}^k\| < +\infty.$$

We therefore have

$$\lim_{k\to\infty}\|x^k - \bar{x}^k\| = 0, \lim_{k\to\infty}\|x^{k+1} - \bar{x}^k\| = 0. \quad (31)$$

Thus, by taking $k \to \infty$ in the inequality $\|x^{k+1} - x^k\| \le \|x^{k+1} - \bar{x}^k\| + \|\bar{x}^k - x^k\|$, we have

$$\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0. \tag{32}$$

From (30), we also know

$$\lim_{k\to\infty} \sum_{j=1}^{r} \beta_j \|(I - P_{Q_j^k})A\bar{x}^k\| = 0. \tag{33}$$

Hence for every $j = 1, 2, \cdots, r$, we have

$$\lim_{k\to\infty} \|(I - P_{Q_j^k})A\bar{x}^k\| = 0. \tag{34}$$

Since $\{x^k\}$ is bounded, the set $\omega_\omega(x^k)$ is nonempty. Let $\widehat{x} \in \omega_\omega(x^k)$; then there exists a subsequence $\{x^{k_n}\}$ of $\{x^k\}$ such that $x^{k_n} \rightharpoonup \widehat{x}$. Next, we show that $\widehat{x}$ is a solution of the MSSFP (1), which will show that $\omega_\omega(x^k) \subset S$. In fact, since $x^{k_n+1} \in C_{[k_n]}^{k_n}$, then by the definition of $C_{[k_n]}^{k_n}$, we have

$$c_{[k_n]}(x^{k_n}) + \langle \xi_{[k_n]}^{k_n}, x^{k_n+1} - x^{k_n} \rangle \le 0, \tag{35}$$

where $\xi_{[k_n]}^{k_n} \in \partial c_{[k_n]}(x^{k_n})$. For every $i = 1, 2, \cdots, t$, choose a subsequence $\{k_{n_s}\} \subset \{k_n\}$ such that $[k_{n_s}] = i$, then

$$c_i(x^{k_{n_s}}) + \langle \xi_i^{k_{n_s}}, x^{k_{n_s}+1} - x^{k_{n_s}} \rangle \le 0. \tag{36}$$

Following the assumption (A3) on the boundedness of ∂c_i and (32), there exists M_1 such that

$$\begin{aligned} c_i(x^{k_{n_s}}) &\le \langle \xi_i^{k_{n_s}}, x^{k_{n_s}} - x^{k_{n_s}+1} \rangle \\ &\le \|\xi_i^{k_{n_s}}\| \|x^{k_{n_s}} - x^{k_{n_s}+1}\| \\ &\le M_1 \|x^{k_{n_s}} - x^{k_{n_s}+1}\| \to 0,\ s \to \infty. \end{aligned} \tag{37}$$

From the weak lower semicontinuity of the convex function c_i, we deduce from (37) that $c_i(\widehat{x}) \le \liminf_{s\to\infty} c_i(x^{k_{n_s}}) \le 0$, i.e., $\widehat{x} \in C = \bigcap_{i=1}^{t} C_i$.

Noting the fact that $I - P_{Q_j^{k_n}}$ is nonexpansive, together with (31), (34), and A being a bounded and linear operator, we get that

$$\begin{aligned} \|(I - P_{Q_j^{k_n}})Ax^{k_n}\| &\le \|(I - P_{Q_j^{k_n}})Ax^{k_n} - (I - P_{Q_j^{k_n}})A\bar{x}^{k_n}\| + \|(I - P_{Q_j^{k_n}})A\bar{x}^{k_n}\| \\ &\le \|Ax^{k_n} - A\bar{x}^{k_n}\| + \|(I - P_{Q_j^{k_n}})A\bar{x}^{k_n}\| \\ &\le \|A\| \|x^{k_n} - \bar{x}^{k_n}\| + \|(I - P_{Q_j^{k_n}})A\bar{x}^{k_n}\| \to 0,\ n \to \infty. \end{aligned} \tag{38}$$

Since $P_{Q_j^{k_n}}(Ax^{k_n}) \in Q_j^{k_n}$, we have

$$q_j(Ax^{k_n}) + \langle \eta_j^{k_n}, P_{Q_j^{k_n}}(Ax^{k_n}) - Ax^{k_n} \rangle \le 0, \tag{39}$$

where $\eta_j^{k_n} \in \partial q_j(Ax^{k_n})$. From the boundedness assumption (A3), (38), and (39), there exists M_2 such that

$$\begin{aligned} q_j(Ax^{k_n}) &\le \|\eta_j^{k_n}\|\|Ax^{k_n} - P_{Q_j^{k_n}}(Ax^{k_n})\| \\ &\le M_2\|(I - P_{Q_j^{k_n}})Ax^{k_n}\| \to 0,\ n \to \infty. \end{aligned} \tag{40}$$

Then $q_j(A\hat{x}) \le \liminf_{n\to\infty} q_j(Ax^{k_n}) \le 0$, thus $A\hat{x} \in Q = \bigcap_{j=1}^r Q_j$, and therefore $\hat{x} \in S$. Using Lemma 3, we conclude that the sequence $\{x^k\}$ converges weakly to a solution of the MSSFP (1). □

Now, we present Algorithm 2 in which the step size is given by the self-adaptive method and prove its weak convergence.

Algorithm 2 (The relaxed CQ algorithm with self-adaptive step size and perturbation)

Take arbitrarily the initial guess x^0, and calculate

$$x^{k+1} = P_{C_{[k]}^k}(x^k - \alpha_k \nabla f_k(x^k) + e_3(x^k)), \tag{41}$$

where $\alpha_k = \frac{\rho_k f_k(x^k)}{\|\nabla f_k(x^k)\|^2}$, $0 < \rho_k < 4$, and C_i, Q_j, C_i^k, Q_j^k and $\nabla f_k(x)$ were defined at the beginning of this section.

The convergence result of Algorithm 2 is stated in the next theorem.

Theorem 2. *Let $\{x^k\}$ be the sequence generated by Algorithm 2. Assumptions (A1)∼(A4) hold and ρ_k satisfies $\inf_k \rho_k(4-\rho_k) > 0$. Then $\{x^k\}$ converges weakly to a solution of the MSSFP (1).*

Proof. First, we prove $\{x^k\}$ is bounded. Let $x^* \in S$. Following Lemma 1 (ii), we have

$$\begin{aligned} & \|x^{k+1} - x^*\|^2 \\ = & \|P_{C_{[k]}^k}(x^k - \alpha_k \nabla f_k(x^k) + e_3(x^k)) - x^*\|^2 \\ \le & \|x^k - \alpha_k \nabla f_k(x^k) + e_3(x^k) - x^*\|^2 - \|x^{k+1} - x^k + \alpha_k \nabla f_k(x^k) - e_3(x^k)\|^2 \\ = & \|x^k - x^*\|^2 - \|x^{k+1} - x^k\|^2 - 2\langle \alpha_k \nabla f_k(x^k) - e_3(x^k), x^k - x^* \rangle \\ & -2\langle \alpha_k \nabla f_k(x^k) - e_3(x^k), x^{k+1} - x^k \rangle \\ = & \|x^k - x^*\|^2 - \|x^{k+1} - x^k\|^2 - 2\alpha_k \langle \nabla f_k(x^k), x^k - x^* \rangle \\ & -2\langle \alpha_k \nabla f_k(x^k), x^{k+1} - x^k \rangle + 2\langle e_3(x^k), x^{k+1} - x^* \rangle. \end{aligned} \tag{42}$$

From Lemma 1 (iii), it follows

$$2\langle e_3(x^k), x^{k+1} - x^* \rangle \le \|e_3(x^k)\| + \|e_3(x^k)\|\|x^{k+1} - x^*\|^2. \tag{43}$$

Similar with (22), it holds that

$$2\alpha_k \langle \nabla f_k(x^k), x^k - x^* \rangle \ge 2\alpha_k \sum_{j=1}^r \beta_j \|(I - P_{Q_j^k})Ax^k\|^2 = 4\alpha_k f_k(x^k). \tag{44}$$

From Lemma 1 (iv), one has

$$-2\langle \alpha_k \nabla f_k(x^k), x^{k+1} - x^k \rangle \le \alpha_k^2 \|\nabla f_k(x^k)\|^2 + \|x^{k+1} - x^k\|^2. \tag{45}$$

Substituting (43)–(45) into (42), we get that

$$\begin{aligned}\|x^{k+1}-x^*\|^2 &\le \|x^k-x^*\|^2+\alpha_k^2\|\nabla f_k(x^k)\|^2-4\alpha_k f_k(x^k)+\|e_3(x^k)\| \\ &\quad +\|e_3(x^k)\|\|x^{k+1}-x^*\|^2, \\ &= \|x^k-x^*\|^2+\frac{\rho_k^2 f_k^2(x^k)}{\|\nabla f_k(x^k)\|^4}\|\nabla f_k(x^k)\|^2-4\frac{\rho_k f_k(x^k)}{\|\nabla f_k(x^k)\|^2}f_k(x^k) \\ &\quad +\|e_3(x^k)\|+\|e_3(x^k)\|\|x^{k+1}-x^*\|^2, \\ &= \|x^k-x^*\|^2-\rho_k(4-\rho_k)\frac{f_k^2(x^k)}{\|\nabla f_k(x^k)\|^2}+\|e_3(x^k)\|\|x^{k+1}-x^*\|^2 \\ &\quad +\|e_3(x^k)\|. \end{aligned} \tag{46}$$

Organizing the above formula, we obtain that

$$\begin{aligned}&\|x^{k+1}-x^*\|^2 \\ \le\; & \frac{1}{1-\|e_3(x^k)\|}\|x^k-x^*\|^2-\frac{\rho_k(4-\rho_k)}{1-\|e_3(x^k)\|}\frac{f_k^2(x^k)}{\|\nabla f_k(x^k)\|^2}+\frac{\|e_3(x^k)\|}{1-\|e_3(x^k)\|}.\end{aligned} \tag{47}$$

From assumption (A4), we know that $\lim_{k\to\infty} e_3(x^k)=0$, so we can assume without loss of generality that $\|e_3(x^k)\|\in[0,1/2)$, $k\ge 0$, then

$$1\le \frac{1}{1-\|e_3(x^k)\|}\le 1+2\|e_3(x^k)\|<2. \tag{48}$$

So (47) can be reduced as

$$\|x^{k+1}-x^*\|^2\le (1+2\|e_3(x^k)\|)\|x^k-x^*\|^2+2\|e_3(x^k)\|. \tag{49}$$

Using Lemma 2, we get the existence of $\lim_{k\to\infty}\|x^k-x^*\|^2$ and the boundedness of $\{x^k\}_{k=0}^{\infty}$.

From (47), we know

$$\begin{aligned}&\frac{\rho_k(4-\rho_k)}{1-\|e_3(x^k)\|}\frac{f_k^2(x^k)}{\|\nabla f_k(x^k)\|^2} \\ \le\; & \frac{1}{1-\|e_3(x^k)\|}\|x^k-x^*\|^2-\|x^{k+1}-x^*\|^2+\frac{\|e_3(x^k)\|}{1-\|e_3(x^k)\|}\to 0,\end{aligned} \tag{50}$$

then the fact that $\inf_k \rho_k(4-\rho_k)>0$ asserts that $\frac{f_k^2(x^k)}{\|\nabla f_k(x^k)\|^2}\to 0$. Since ∇f_k is Lipschitz continuity and $\nabla f_k(x^*)=0$, we get that

$$\|\nabla f_k(x^k)\|^2=\|\nabla f_k(x^k)-\nabla f_k(x^*)\|^2\le L^2\|x^k-x^*\|^2. \tag{51}$$

This implies that $\nabla f_k(x^k)$ is bounded, and thus (50) yields $f_k(x^k)\to 0$. Hence for every $j=1,2,\cdots,r$, we have

$$\|(I-P_{Q_j^k})Ax^k\|\to 0,\; k\to\infty. \tag{52}$$

Let $\{x^{k_n}\}$ be a subsequence of $\{x^k\}$ such that $x^{k_n}\rightharpoonup \hat{x}\in\omega_\omega(x^k)$, and $\{k_{n_s}\}$ are a subsequence of $\{k_n\}$ such that $[k_{n_s}]=i$. Similar to the proof of Theorem 1, we know that $c_i(\hat{x})\le \liminf_{s\to\infty} c_i(x^{k_{n_s}})\le 0$, i.e., $\hat{x}\in C=\bigcap_{i=1}^t C_i$. Since (52) indicates that $q_j(A\hat{x})\le\liminf_{n\to\infty} q_j(Ax^{k_n})\le 0$, $A\hat{x}\in Q=\bigcap_{j=1}^r Q_j$. Therefore $\hat{x}\in S$. Using Lemma 3, we conclude that the sequence $\{x^k\}$ converges weakly to a solution of the MSSFP (1). □

4. The Bounded Perturbation Resilience

4.1. Bounded Perturbation Resilience of the Algorithms

In this subsection, we consider the bounded perturbation algorithms of Algorithms 1 and 2. Based on Definition 3, in Algorithm 1, let $e_i(x^k) = 0$, $i = 1, 2$. The original algorithm is

$$\begin{cases} \bar{x}^k = P_{C^k_{[k]}}(x^k - \alpha_k \nabla f_k(x^k)), \\ x^{k+1} = P_{C^k_{[k]}}(x^k - \alpha_k \nabla f_k(\bar{x}^k)), \end{cases} \tag{53}$$

where α_k is obtained by Armijo-line search step size such that $\alpha_k \|\nabla f_k(x^k) - \nabla f_k(\bar{x}^k)\| \leq \mu \|x^k - \bar{x}^k\|$, where $\mu \in (0,1)$. The generated iteration sequence is weakly convergent, which is proved as a special case in Section 3. The algorithm with the bounded perturbation of (53) is that

$$\begin{cases} \bar{x}^k = P_{C^k_{[k]}}(x^k + \lambda_k \nu_k - \alpha_k \nabla f_k(x^k + \lambda_k \nu_k)), \\ x^{k+1} = P_{C^k_{[k]}}(x^k + \lambda_k \nu_k - \alpha_k \nabla f_k(\bar{x}^k)). \end{cases} \tag{54}$$

where $[k] = k \bmod t$ and $\alpha_k = \gamma l^{m_k}$ with m_k the smallest non-negative integer such that

$$\begin{aligned} \alpha_k \|\nabla f_k(x^k + \lambda_k \nu_k) - \nabla f_k(\bar{x}^k)\| &\leq \mu \|x^k + \lambda_k \nu_k - \bar{x}^k\| \\ &\leq \mu(\|x^k - \bar{x}^k\| + \lambda_k \|\nu_k\|). \end{aligned} \tag{55}$$

The following theorem shows that the algorithm (53) is bounded perturbation-resilient.

Theorem 3. *Assume that (A1)∼(A3) are true; the sequence $\{\nu_k\}_{k=0}^{\infty}$ is bounded and the scalar sequence $\{\lambda_k\}_{k=0}^{\infty}$ satisfies $\lambda_k \geq 0$ and $\Sigma_{k=0}^{\infty} \lambda_k < +\infty$. Then the sequence $\{x^k\}_{k=0}^{\infty}$ generated by iterative scheme (54) converges weakly to a solution of the MSSFP (1). Thus, the algorithm (53) is bounded perturbation-resilient.*

Proof. Let $x^* \in S$. Since $\Sigma_{k=0}^{\infty} \lambda_k < +\infty$ and the sequence $\{\nu_k\}_{k=0}^{\infty}$ are bounded, we have

$$\sum_{k=0}^{\infty} \lambda_k \|\nu_k\| < +\infty, \tag{56}$$

thus

$$\lim_{k \to \infty} \lambda_k \|\nu_k\| = 0. \tag{57}$$

So we can assume that $\lambda_k \|\nu_k\| \in [0, (1 - \mu - \tau)/2)$, where $\tau \in (0, 1 - \mu)$, without loss of generality. Replacing $e_2(x^k)$ with $\lambda_k \nu_k$ in (20) and using Lemma 1 (iii) show

$$\begin{aligned} & \|x^{k+1} - x^*\|^2 \\ \leq\ & \|x^k - x^*\|^2 - \|\bar{x}^k - x^k\|^2 - \|x^{k+1} - \bar{x}^k\|^2 - 2\alpha_k \langle \nabla f_k(\bar{x}^k), \bar{x}^k - x^* \rangle \\ & + 2\langle x^k - \bar{x}^k - \alpha_k \nabla f_k(\bar{x}^k), x^{k+1} - \bar{x}^k \rangle + 2\langle \lambda_k \nu_k, x^{k+1} - x^* \rangle \\ \leq\ & \|x^k - x^*\|^2 - \|\bar{x}^k - x^k\|^2 - \|x^{k+1} - \bar{x}^k\|^2 - 2\alpha_k \langle \nabla f_k(\bar{x}^k), \bar{x}^k - x^* \rangle \\ & + 2\langle x^k - \bar{x}^k - \alpha_k \nabla f_k(\bar{x}^k), x^{k+1} - \bar{x}^k \rangle + \lambda_k \|\nu_k\| + \lambda_k \|\nu_k\| \|x^{k+1} - x^*\|^2. \end{aligned} \tag{58}$$

Since $I - P_C$ is firmly nonexpensive, $\nabla f_k(x^*) = 0$ and Lemma 4, we get that

$$2\alpha_k \langle \nabla f_k(\bar{x}^k), \bar{x}^k - x^* \rangle \geq 2\frac{\mu l}{L} \sum_{j=1}^{r} \beta_j \|(I - P_{Q^k_j}) A \bar{x}^k\|^2. \tag{59}$$

Based on the definition of $\bar{x}^k$ and Lemma 1 (i), we know that

$$\langle \bar{x}^k - x^k + \alpha_k \nabla f_k(x^k + \lambda_k \nu_k) - \lambda_k \nu_k, x^{k+1} - \bar{x}^k \rangle \geq 0. \tag{60}$$

Based on (55), the following formulas holds

$$\begin{aligned}
& 2\langle \alpha_k \nabla f_k(x^k + \lambda_k \nu_k) - \alpha_k \nabla f_k(\bar{x}^k), x^{k+1} - \bar{x}^k \rangle \\
\leq \;& 2\alpha_k \|\nabla f_k(x^k + \lambda_k \nu_k) - \nabla f_k(\bar{x}^k)\| \|x^{k+1} - \bar{x}^k\| \\
\leq \;& 2\mu \|x^k - \bar{x}^k\| \|x^{k+1} - \bar{x}^k\| + 2\mu\lambda_k \|\nu_k\| \|x^{k+1} - \bar{x}^k\| \\
= \;& \mu \|x^k - \bar{x}^k\|^2 + (\mu + \lambda_k \|\nu_k\|) \|x^{k+1} - \bar{x}^k\|^2 + \mu^2 \lambda_k \|\nu_k\|.
\end{aligned} \tag{61}$$

Lemma 1 (iii) reads that

$$-2\lambda_k \langle \nu_k, x^{k+1} - \bar{x}^k \rangle \leq \lambda_k \|\nu_k\| + \lambda_k \|\nu_k\| \|x^{k+1} - \bar{x}^k\|^2. \tag{62}$$

Substituting (60)–(62) into the fifth item of (58), we get

$$\begin{aligned}
& 2\langle x^k - \bar{x}^k - \alpha_k \nabla f_k(\bar{x}^k), x^{k+1} - \bar{x}^k \rangle \\
\leq \;& 2\langle x^k - \bar{x}^k - \alpha_k \nabla f_k(\bar{x}^k), x^{k+1} - \bar{x}^k \rangle \\
& +2\langle \bar{x}^k - x^k + \alpha_k \nabla f_k(x^k + \lambda_k \nu_k) - \lambda_k \nu_k, x^{k+1} - \bar{x}^k \rangle \\
= \;& 2\langle \alpha_k \nabla f_k(x^k + \lambda_k \nu_k) - \alpha_k \nabla f_k(\bar{x}^k), x^{k+1} - \bar{x}^k \rangle - 2\lambda_k \langle \nu_k, x^{k+1} - \bar{x}^k \rangle \\
\leq \;& \mu \|x^k - \bar{x}^k\|^2 + (\mu + 2\lambda_k \|\nu_k\|) \|x^{k+1} - \bar{x}^k\|^2 + (1 + \mu^2) \lambda_k \|\nu_k\| \\
\leq \;& \mu \|x^k - \bar{x}^k\|^2 + (1 - \tau) \|x^{k+1} - \bar{x}^k\|^2 + 2\lambda_k \|\nu_k\|.
\end{aligned} \tag{63}$$

Substituting (59) and (63) into (58) we get

$$\begin{aligned}
\|x^{k+1} - x^*\|^2 \leq \;& \frac{1}{1 - \lambda_k \|\nu_k\|} \Big[\|x^k - x^*\|^2 + 3\lambda_k \|\nu_k\| - (1 - \mu) \|\bar{x}^k - x^k\|^2 \\
& -\tau \|x^{k+1} - \bar{x}^k\|^2 - 2\frac{\mu l}{L} \sum_{j=1}^{r} \beta_j \|(I - P_{Q_j^k}) A\bar{x}^k\|^2 \Big].
\end{aligned} \tag{64}$$

Since $\lambda_k \|\nu_k\| \in [0, (1 - \mu - \tau)/2)$, we get

$$1 \leq \frac{1}{1 - \lambda_k \|\nu_k\|} \leq 1 + 2\lambda_k \|\nu_k\| < 2. \tag{65}$$

This, together with (64), shows that

$$\begin{aligned}
\|x^{k+1} - x^*\|^2 \leq \;& (1 + 2\lambda_k \|\nu_k\|) \Big[\|x^k - x^*\|^2 - (1 - \mu) \|x^k - \bar{x}^k\|^2 - \tau \|x^{k+1} - \bar{x}^k\|^2 \\
& -2\frac{\mu l}{L} \sum_{j=1}^{r} \beta_j \|(I - P_{Q_j^k}) A\bar{x}^k\|^2 \Big] + 6\lambda_k \|\nu_k\| \\
\leq \;& (1 + 2\lambda_k \|\nu_k\|) \|x^k - x^*\|^2 + 6\lambda_k \|\nu_k\|.
\end{aligned} \tag{66}$$

Using Lemma 2, we know the existence of $\lim_{k\to\infty} \|x^k - x^*\|^2$ and the boundedness of $\{x^k\}_{k=0}^{\infty}$.

From (64), it follows that

$$\begin{aligned}
& (1 - \mu) \|\bar{x}^k - x^k\|^2 + \tau \|x^{k+1} - \bar{x}^k\|^2 + 2\frac{\mu l}{L} \sum_{j=1}^{r} \beta_j \|(I - P_{Q_j^k}) A\bar{x}^k\|^2 \\
\leq \;& \|x^k - x^*\|^2 - (1 - \lambda_k \|\nu_k\|) \|x^{k+1} - x^*\|^2 + 3\lambda_k \|\nu_k\|.
\end{aligned} \tag{67}$$

Thus, we have $\lim_{k\to\infty}\|x^k-\bar{x}^k\|=0$, $\lim_{k\to\infty}\|x^{k+1}-\bar{x}^k\|=0$ and $\lim_{k\to\infty}\sum_{j=1}^r\beta_j\|(I-P_{Q_j^k})A\bar{x}^k\|^2=0$. Hence,

$$\lim_{k\to\infty}\|x^{k+1}-x^k\|=0, \tag{68}$$

and for every $j=1,2,\cdots,r$,

$$\lim_{k\to\infty}\|(I-P_{Q_j^k})A\bar{x}^k\|=0. \tag{69}$$

Similarly to with Theorem 1, we conclude that the sequence $\{x^k\}$ converges weakly to a solution of the MSSFP (1). □

Remark 1. *When $t=1$, $r=1$, the MSSFP reduces to the SFP; thus Theorems 1 and 3 guarantee that algorithm (53) is bounded perturbation-resilient with Armijo-line search step size for the SFP.*

Remark 2. *Replace $f_k(x)$ in algorithm (53) by $g_k(x)$, and $\nabla f_k(x)$ by $\nabla g_k(x)$, where $g_k(x)=\frac{1}{2}\|(I-P_{Q_{[k]}^k})Ax\|^2$, and $\nabla g_k(x)=A^*(I-P_{Q_{[k]}^k})Ax$, $[k]=k \bmod r$. The corresponding algorithm is also bounded perturbation-resilient.*

Next, we will prove that Algorithm 2 with self-adaptive step size is bounded perturbation-resilient. Based on Definition 3, let $e_3(x^k)=0$ in Algorithm 2. The original algorithm is

$$x^{k+1}=P_{C_{[k]}^k}(x^k-\alpha_k\nabla f_k(x^k)), \tag{70}$$

where $\alpha_k=\frac{\rho_k f_k(x^k)}{\|\nabla f_k(x^k)\|^2}$, $0<\rho_k<4$. The iterative sequence converges weakly to a solution of the MSSFP (1); see [26]. Consider the algorithm with the bounded perturbation

$$x^{k+1}=P_{C_{[k]}^k}(x^k+\lambda_k\nu_k-\tilde{\alpha}_k\nabla f_k(x^k+\lambda_k\nu_k)), \tag{71}$$

where $\tilde{\alpha}_k=\frac{\rho_k f_k(x^k+\lambda_k\nu_k)}{\|\nabla f_k(x^k+\lambda_k\nu_k)\|^2}$, $0<\rho_k<4$. The following theorem shows that the algorithm (70) is bounded-perturbation-resilient.

Theorem 4. *Suppose that (A1)∼(A3) are true; the sequence $\{\nu_k\}_{k=0}^\infty$ is bounded and the scalar sequence $\{\lambda_k\}_{k=0}^\infty$ satisfies $\lambda_k\geq 0$, $\Sigma_{k=0}^\infty\lambda_k<+\infty$, and ρ_k satisfies $\inf_k\rho_k(4-\rho_k)>0$. Then the sequence $\{x^k\}_{k=0}^\infty$ generated by iterative scheme (71) converges weakly to a solution of the MSSFP (1). Thus, the algorithm (70) is bounded-perturbation-resilient.*

Proof. Set $e_3(x^k)=\lambda_k\nu_k+\alpha_k\nabla f_k(x^k)-\tilde{\alpha}_k\nabla f_k(x^k+\lambda_k\nu_k)$, then (71) can be rewritten as $x^{k+1}=P_{C_{[k]}^k}(x^k-\alpha_k\nabla f_k(x^k)+e_3(x^k))$, which is the form of Algorithm 2. According to Theorem 2, it suffices to prove that $\sum_{k=0}^\infty e_3(x^k)<+\infty$. Since $\frac{\rho_k f_k(x^k+\lambda_k\nu_k)}{\|\nabla f_k(x^k+\lambda_k\nu_k)\|^2}\nabla f_k(x^k+\lambda_k\nu_k)$ is continuous, we write

$$\frac{\rho_k f_k(x^k+\lambda_k\nu_k)}{\|\nabla f_k(x^k+\lambda_k\nu_k)\|^2}\nabla f_k(x^k+\lambda_k\nu_k)=\frac{\rho_k f_k(x^k)}{\|\nabla f_k(x^k)\|^2}\nabla f_k(x^k)+O(\lambda_k\nu_k), \tag{72}$$

where $O(\lambda_k\nu_k)$ denotes the infinitesimal of the same order of $\lambda_k\nu_k$. From the expression of $e_3(x^k)$, we obtain

$$\begin{aligned}\|e_3(x^k)\| &\le \|\lambda_k v_k\| + \left\|\alpha_k \nabla f_k(x^k) - \tilde{\alpha}_k \nabla f_k(x^k + \lambda_k v_k)\right\| \\ &= \|\lambda_k v_k\| + \left\|\frac{\rho_k f_k(x^k)}{\|\nabla f_k(x^k)\|^2}\nabla f_k(x^k) - \frac{\rho_k f_k(x^k+\lambda_k v_k)}{\|\nabla f_k(x^k+\lambda_k v_k)\|^2}\nabla f_k(x^k+\lambda_k v_k)\right\| \\ &= \|\lambda_k v_k\| + \left\|\frac{\rho_k f_k(x^k)}{\|\nabla f_k(x^k)\|^2}\nabla f_k(x^k) - \left(\frac{\rho_k f_k(x^k)}{\|\nabla f_k(x^k)\|^2}\nabla f_k(x^k) + O(\lambda_k v_k)\right)\right\| \\ &= \|\lambda_k v_k\| + \|O(\lambda_k v_k)\|. \end{aligned} \tag{73}$$

Since $\{\lambda_k v_k\}$ is summable, we know that $\{e_3(x^k)\}$ is summable, i.e., $\sum_{k=0}^{\infty}\|e_3(x^k)\| \le +\infty$. Thus, we conclude that the sequence $\{x^k\}$ converges weakly to a solution of the MSSFP (1); i.e., the algorithm (70) is the bounded-perturbation-resilient. □

Remark 3. *When $t = 1$, $r = 1$, the MSSFP reduces to the SFP; thus Theorems 2 and 4 guarantee that algorithm (70) is bounded-perturbation-resilient with the self-adaptive step size for the SFP.*

4.2. Construction of the Inertial Algorithms by Bounded Perturbation Resilience

In this subsection, we consider algorithms with inertial terms as a special case of Algorithms 1 and 2. In Algorithm 1, letting $e_i(x^k) = \theta_k^{(i)}(x^k - x^{k-1})$, $i = 1, 2$, we obtain

$$\begin{cases} \bar{x}^k = P_{C^k_{[k]}}(x^k - \alpha_k \nabla f_k(x^k) + \theta_k^{(1)}(x^k - x^{k-1})), \\ x^{k+1} = P_{C^k_{[k]}}(x^k - \alpha_k \nabla f_k(\bar{x}^k) + \theta_k^{(2)}(x^k - x^{k-1})), \end{cases} \tag{74}$$

where the step size α_k is obtained by Armijo-line search and

$$\theta_k^{(i)} = \begin{cases} \dfrac{\lambda_k^{(i)}}{\|x^k - x^{k-1}\|}, & \|x^k - x^{k-1}\| > 1, \\ \lambda_k^{(i)}, & \|x^k - x^{k-1}\| \le 1, \end{cases} \quad i = 1, 2. \tag{75}$$

Theorem 5. *Assume that the assumptions (A1)∼(A3) are true, and the sequence $\{\lambda_k\}_{k=0}^{\infty}$ satisfies $\lambda_k \ge 0$, and $\Sigma_{k=0}^{\infty}\lambda_k^{(i)} < +\infty$, $i = 1, 2$. Then, the sequence $\{x^k\}_{k=0}^{\infty}$ generated by iterative scheme (74) converges weakly to a solution of the MSSFP (1).*

Proof. Let $e_i(x^k) = \lambda_k^{(i)} v_k$, $i = 1, 2$, where

$$v_k = \begin{cases} \dfrac{x^k - x^{k-1}}{\|x^k - x^{k-1}\|}, & \|x^k - x^{k-1}\| > 1, \\ x^k - x^{k-1}, & \|x^k - x^{k-1}\| \le 1. \end{cases} \tag{76}$$

Thus, we know that $\|v_k\| \le 1$ and $\{e_i(x^k)\}_{k=0}^{\infty}$ satisfies assumption (A4). According to Theorem 1, we conclude that the sequence $\{x^k\}$ converges weakly to a solution of the MSSFP (1). □

Considering the algorithm with inertial bounded perturbation

$$\begin{cases} \bar{x}^k = P_{C^k_{[k]}}(x^k + \theta_k(x^k - x^{k-1}) - \alpha_k \nabla f_k(x^k + \theta_k(x^k - x^{k-1}))), \\ x^{k+1} = P_{C^k_{[k]}}(x^k + \theta_k(x^k - x^{k-1}) - \alpha_k \nabla f_k(\bar{x}^k)). \end{cases} \tag{77}$$

where

$$\theta_k = \begin{cases} \dfrac{\lambda_k}{\|x^k - x^{k-1}\|}, & \|x^k - x^{k-1}\| > 1, \\ \lambda_k, & \|x^k - x^{k-1}\| \le 1. \end{cases} \tag{78}$$

According to Theorem 3, it is easy to know that the sequence $\{x^k\}$ converges weakly to a solution of the MSSFP (1). More relevant evidence can be found in reference [27].

Similarly, we can get Theorem 6, which asserts that Algorithm 2 with the inertial perturbation is weakly convergent.

Theorem 6. *Assume that (A1)∼(A3) are true; the scalar sequence $\{\lambda_k\}_{k=0}^{\infty}$ satisfies $\lambda_k \geq 0$, and $\Sigma_{k=0}^{\infty}\lambda_k < +\infty$, and ρ_k satisfies $\inf_k \rho_k(4-\rho_k) > 0$. Then the sequence $\{x^k\}_{k=0}^{\infty}$ is generated by each of the following iterative scheme,*

$$x^{k+1} = P_{C_{[k]}^k}(x^k - \alpha_k \nabla f_k(x^k) + \theta_k(x^k - x^{k-1})), \tag{79}$$

$$x^{k+1} = P_{C_{[k]}^k}(x^k - \alpha_k \nabla f_k(x^k + \theta_k(x^k - x^{k-1})) + \theta_k(x^k - x^{k-1})), \tag{80}$$

where θ_k is the same as (78) and α_k is self-adaptive step size which is the same as in Algorithm 2, converges weakly to a solution of the MSSFP (1).

5. Numerical Experiments

In this section, we compare the asymptotic behavior of algorithms (53) (Chen et al. [9]), (77) (Algorithm 1), (70) (Wen et al. [26]) and (80) (Algorithm 2), denoted by NP1, HP1, NP2, and HP2, respectively. For the sake of convenience, we denote $\mathbf{e}_0 = (0, 0, \cdots, 0)^T$ and $\mathbf{e}_1 = (1, 1, \cdots, 1)^T$, respectively. The codes are written in Matlab 2016a and run on Inter(R) Core(TM) i7-8550U CPU @ 1.80 GHz 2.00 GHz, RAM 8.00 GB. We present two kinds of experiments. One is a real-life problem called LASSO problem, the other kind is some numerical simulation including three examples of the MSSFP.

5.1. LASSO Problem

Let us consider the following LASSO problem [28]

$$\min \{\frac{1}{2}\|Ax-b\|_2^2 \mid x \in \mathbf{R^n}, \|x\|_1 \leq \varepsilon\}$$

where $A \in \mathbf{R^{m\times n}}$, $m < n$, $b \in \mathbf{R^m}$, and $\varepsilon > 0$. The matrix A is generated from a standard normal distribution with mean zero and unit variance. The true sparse signal x^* is generated from uniformly distribution in the interval $[-2, 2]$ with random p position nonzero, while the rest is kept zero. The sample data $b = Ax^*$. For the considered MSSFP, let $r = t = 1$ and $C = \{x \mid \|x\|_1 \leq \varepsilon\}$, $Q = \{b\}$. The objective function is defined as $f(x) = \frac{1}{2}\|Ax-b\|_2^2$.

We report the final error between the reconstructed signal and the true signal. Take $\|x^k - x^*\| < 10^{-4}$ as the stopping criterion, where x^* is the true signal. We compare the algorithms NP1, HP1, NP2 and HP2 with Yang's algorithm [3]. Let $\alpha_k = \gamma l^{m_k}$ for all $k \geq 1$, $\gamma = 1$, $l = \frac{1}{2}$, $\mu = \frac{1}{2}$, $\theta_k = \frac{1}{4}$, $\rho_k = 0.1$, and $\alpha_k = 0.1 * \dfrac{1}{\|A\|^2}$ of Yang's algorithm [3].

The results are reported in Table 1. Figure 1 shows the objective function value versus iteration numbers when $m = 240, n = 1024, p = 30$.

From Table 1 and Figure 1, we know that the inertial perturbation can improve the convergence of the algorithms and that the algorithms with Armijo-line search or self-adaptive step size perform better than Yang's algorithm [3].

We also measure the restoration accuracy by means of the mean squared error, i.e., MSE$= (1/k)\|x^* - x^k\|$, where x^* is an estimated signal of x. Figure 2 shows a comparison of the accuracy of the recovered signals when $m = 1440, n = 6144, p = 180$. Given the same number of iterations, the recovered signals generated by algorithms in this paper outperform the one generated by Yang's algorithm; NP1 needs more CPU time and presents lower accuracy; algorithms with self-adaptive step size perform better than the algorithms with step size determined by Armijo-line search in CPU time and imposing inertial perturbation accelerates the convergence rate and accuracy of signal recovery.

Table 1. Comparison of algorithms with different step size.

m	n	p		NP1	HP1	NP2	HP2	Yang's alg.
120	512	15	No. of Iter	1588	1119	10,004	7426	10,944
			cpu(time)	0.8560	0.6906	0.6675	0.4991	0.7011
240	1024	30	No. of Iter	1909	1354	10,726	7969	13,443
			cpu(time)	2.1224	1.4836	1.6236	1.2011	1.9789
480	2048	60	No. of Iter	2972	2117	17,338	12,897	22,118
			cpu(time)	22.5140	14.8782	15.4729	11.1033	19.3376
720	3072	90	No. of Iter	3955	2872	21,853	16,244	28,004
			cpu(time)	134.9243	82.6705	79.1640	57.1230	110.0482

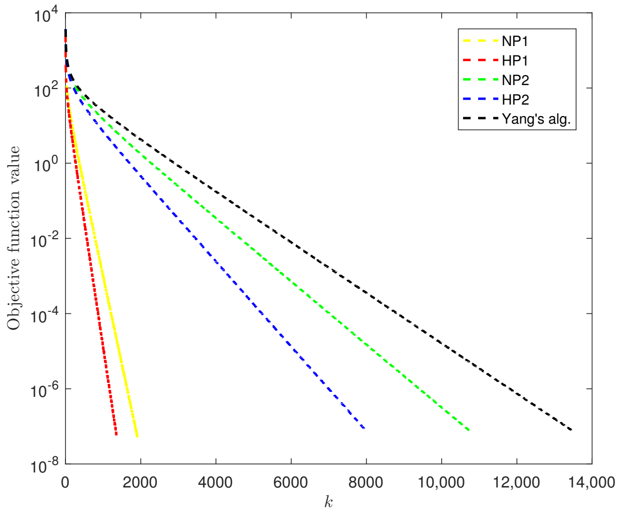

Figure 1. The objective function value versus the iteration number.

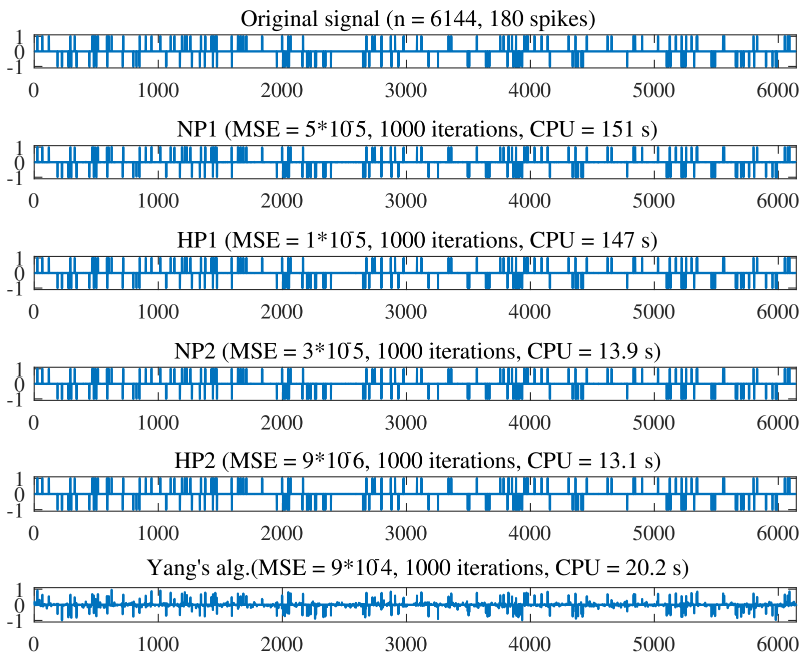

Figure 2. Comparison of signal processing.

5.2. Three MSSFP Problems

Example 1 ([5]). *Take* $\mathcal{H}_1 = \mathcal{H}_2 = \mathbf{R}^3$, $r = t = 2$, $\beta_1 = \beta_2 = \frac{1}{2}$ *and* $\alpha_k = \gamma l^{m_k}$ *for all* $k \geq 1$, $\gamma = 1$, $l = \frac{1}{2}$, $\mu = \frac{1}{2}$, $\theta_k = \frac{1}{4}$, $\rho_k = 0.1$. *Define*

$$
\begin{aligned}
C_1 &= \left\{x = (x_1, x_2, x_3)^T \in \mathbf{R^3} \mid x_1 + x_2^2 + 2x_3 \leq 0\right\}, \\
C_2 &= \left\{x = (x_1, x_2, x_3)^T \in \mathbf{R^3} \mid \frac{x_1^2}{16} + \frac{x_2^2}{9} + \frac{x_3^2}{4} - 1 \leq 0\right\}, \\
Q_1 &= \left\{x = (x_1, x_2, x_3)^T \in \mathbf{R^3} \mid x_1^2 + x_2 - x_3 \leq 0\right\}, \\
Q_2 &= \left\{x = (x_1, x_2, x_3)^T \in \mathbf{R^3} \mid \frac{x_1^2}{4} + \frac{x_2^2}{4} + \frac{x_3^2}{9} - 1 \leq 0\right\},
\end{aligned}
$$

and

$$
A = \begin{pmatrix} 2 & -1 & 3 \\ 4 & 2 & 5 \\ 2 & 0 & 2 \end{pmatrix}.
$$

The underlying MSSFP is to find $x^* \in C_1 \bigcap C_2$ *such that* $Ax^* \in Q_1 \bigcap Q_2$.

We use inertial perturbation to accelerate the convergence of the algorithm. For the convenience of comparison, the initial values of the two inertial algorithms are set to be the same. Let $x^0 = x^1$. We use $E_k = \|x^{k+1} - x^k\| / \|x^k\|$ to measure the error of the k-th iterate. If $E_k < 10^{-5}$, then the iteration process stops. We compare our proposed iteration methods HP1 , HP2 with NP1, NP2 and Liu and Tang's Algorithm 2 in [29]. Algorithm 2 is of the form $x^{k+1} = U_{[k]}(x^k - \alpha_k \sum_{j=1}^r \beta_j A^*(I - T_j)Ax)$, $\alpha_k \in (0, \frac{2}{\|A\|^2})$. We take $U_{[k]} = P_{C^k_{[k]}}$, $T_j = P_{Q^k_j}$ and $\alpha_k = 0.2 * \frac{1}{\|A\|^2}$, and the algorithm is referred to as LT alg.

The convergence results and the CPU time of the five algorithms are shown in Table 2 and Figure 3. The errors are shown in Figure 4.

The results show that (80) (HP2) outperforms (77) (HP1) for certain initial values. The main reason may be that the self-adaptive step size is more efficient than the one determined by the Armijo-line search. Comparison results of five algorithms and the convergence behavior show that in most cases, the convergence rate of the algorithm can be improved by adding an appropriate perturbation.

Table 2. Numerical results of five algorithms for Example 1.

Choice		NP1	HP1	NP2	HP2	LT alg.
1. $x^0 = (0.1, 0.1, 0.1)^T$	No. of Iter	60	43	219	162	420
	cpu(time)	0.0511	0.0450	0.0362	0.0347	0.0879
2. $x^0 = (-0.4, 0.555, 0.888)^T$	No. of Iter	139	85	195	143	178
	cpu(time)	0.0669	0.0509	0.0342	0.0318	0.0552
3. $x^0 = (1, 2, 3)^T$	No. of Iter	142	89	195	141	178
	cpu(time)	0.0694	0.0490	0.0352	0.0339	0.0551
4. $x^0 = (0.123, 0.745, 0.789)^T$	No. of Iter	149	85	108	77	526
	cpu(time)	0.0590	0.0448	0.0295	0.0268	0.1018

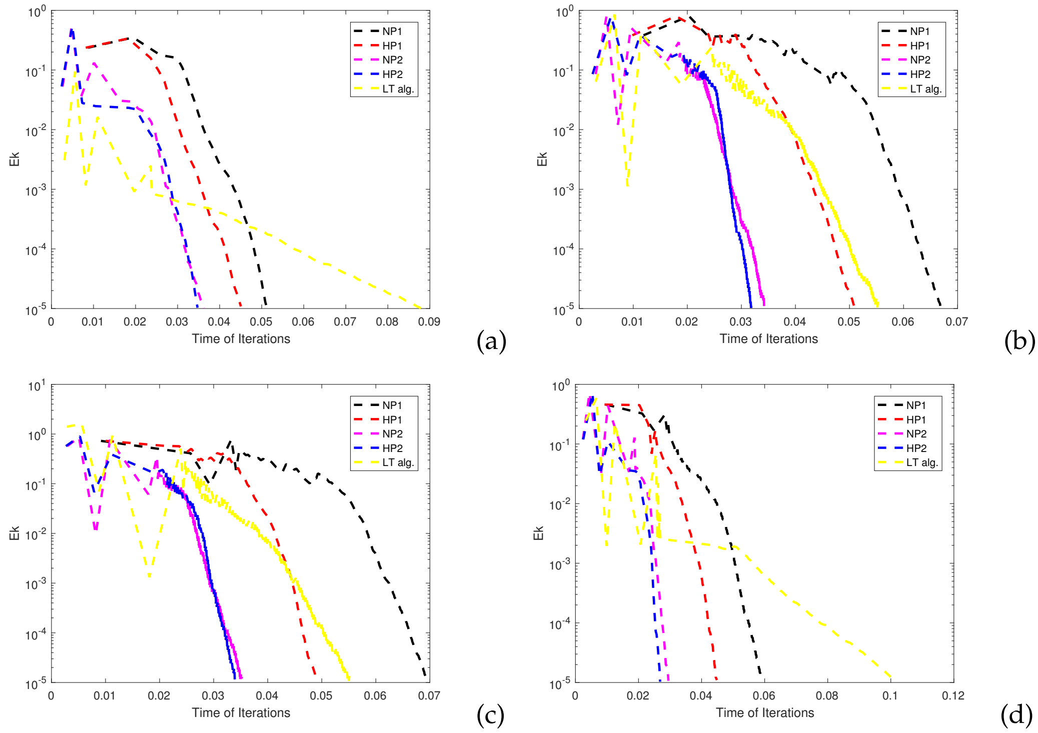

Figure 3. Comparison of CPU times of the algorithms in Example 1: (**a**) Comparison for choice 1. (**b**) Comparison for choice 2. (**c**) Comparison for choice 3. (**d**) Comparison for choice 4.

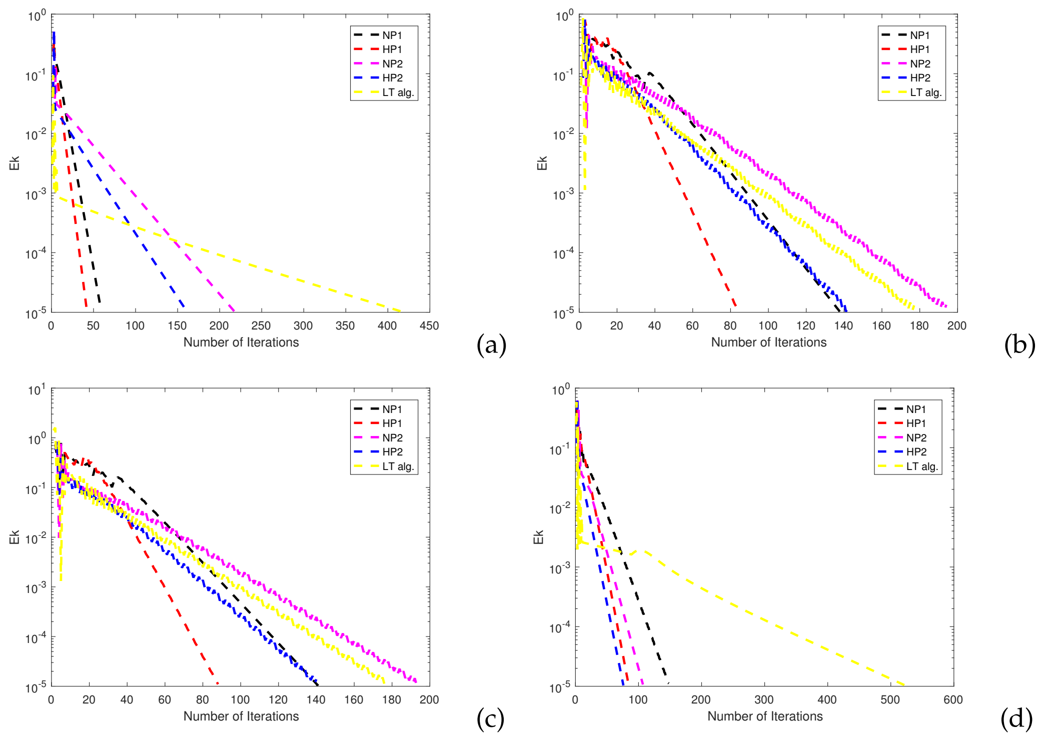

Figure 4. Comparison of iterations of the algorithms in Example 1: (**a**) Comparison for choice 1. (**b**) Comparison for choice 2. (**c**) Comparison for choice 3. (**d**) Comparison for choice 4.

Example 2. *Take* $\mathcal{H}_1 = \mathbf{R}^n$, $\mathcal{H}_2 = \mathbf{R}^m$, $A = (a_{ij})_{m\times n}$ *with* $a_{ij} \in (0,1)$ *generated randomly,* $C_i = \{x \in \mathbf{R}^n \mid \|x - d_i\|_2^2 \le r_i^2\}$, $i = 1,2,\cdots,t$, $Q_j = \{y \in \mathbf{R}^m \mid \|y - l_j\|_1^2 \le h_j^2\}$, $j = 1,2,\cdots,r$, *where* $d_i \in [\mathbf{e}_0, 10\mathbf{e}_1]$, $r_i \in [40,60]$, $l_j \in [\mathbf{e}_0, \mathbf{e}_1]$, $h_j \in [10,20]$ *are all generated* *randomly. Set* $\beta_1 = \beta_2 = \cdots = \beta_r = \frac{1}{r}$ *and* $\alpha_k = \gamma l^{m_k}$ *for all* $k \geq 1$, $\gamma = 1$, $l = \frac{1}{2}$, $\mu = \frac{1}{2}$, $\theta_k = \frac{1}{4}$, $\rho_k = 0.001$.

We consider using inertial perturbation to accelerate the convergence of the algorithm. If $E_k = \|x^{k+1} - x^k\| / \|x^k\| < 10^{-4}$, then the iteration process stops. Let $x^0 = x^1$. We choose arbitrarily three different initial points and consider iterative steps of the four algorithms with m, n, r, t being different values. See Table 3 for details.

Table 3. Numerical results of the algorithms with and without perturbation for Example 2.

Initial Point		NP1	HP1	NP2	HP2
$r = t = 10$, $m = 15$, $n = 20$					
$x^0 = x^1 = 2 * e_1$	No. of Iter	49	36	1281	999
	cpu(time)	0.1031	0.0669	0.1480	1494
$x^0 = x^1 = 50 * e_1$	No. of Iter	187	121	2297	1669
	cpu(time)	0.2485	0.1536	0.2887	0.1868
$x^0 = x^1 = 100 * rand(n, 1)$	No. of Iter	312	225	2357	1811
	cpu(time)	0.4202	0.2908	0.2830	0.2159
$r = t = 10$, $m = n = 40$					
$x^0 = x^1 = 2 * e_1$	No. of Iter	89	66	956	732
	cpu(time)	0.3140	0.1777	0.1534	0.1318
$x^0 = x^1 = 50 * e_1$	No. of Iter	1710	1583	1301	1061
	cpu(time)	4.0390	4.0357	1860	0.1555
$x^0 = x^1 = 100 * rand(n, 1)$	No. of Iter	1674	1658	1487	1219
	cpu(time)	4.6581	3.7752	0.2065	0.1762
$r = t = 30$, $m = n = 40$					
$x^0 = x^1 = 2 * e_1$	No. of Iter	136	103	985	753
	cpu(time)	0.6912	0.5174	0.3312	0.2515
$x^0 = x^1 = 50 * e_1$	No. of Iter	1612	1411	1258	968
	cpu(time)	12.3437	11.7164	0.3991	0.3127
$x^0 = x^1 = 100 * rand(n, 1)$	No. of Iter	1541	1133	1643	1012
	cpu(time)	11.8273	7.4646	1.0363	0.2965

In this example, we found that the algorithm with Armijo-line search needs fewer iteration steps in relatively low-dimensional spaces. In the case of high-dimensional spaces, the algorithm with self-adaptive step size outperforms in time. Generally, the convergence is improved by inertial perturbations for both algorithms in our paper.

Example 3 ([30]). *Take* $\mathcal{H}_1 = \mathbf{R}^n$, $\mathcal{H}_2 = \mathbf{R}^m$, $A = (a_{ij})_{m \times n}$ *with* $a_{ij} \in (0, 1)$ *generated randomly,* $C_i = \{x \in \mathbf{R}^n \mid \|x - d_i\|_2^2 \leq r_i^2\}$, $i = 1, 2, \cdots, t$, $Q_j = \{y \in \mathbf{R}^m \mid y^T B_j y + b_j y + c_j \leq 0\}$, $j = 1, 2, \cdots, r$, *where* $d_i \in (6\mathbf{e}_0, 16\mathbf{e}_1)$, $r_i \in (100, 120)$, $b_j \in (-30\mathbf{e}_1, -20\mathbf{e}_1)$, $c_j \in (-50, -60)$, *and all elements of the matrix* B_j *are all generated randomly in the interval (2,10). Set* $\beta_1 = \beta_2 = \cdots = \beta_r = \frac{1}{r}$ *and* $\alpha_k = \gamma l^{m_k}$ *for all* $k \geq 1$, $\gamma = 1$, $l = \frac{1}{2}$, $\mu = \frac{1}{2}$, $\theta_k = \frac{1}{4}$, $\rho_k = 0.1$.

We consider using inertial perturbation to accelerate the convergence of the algorithm. The stopping criterion is defined by $E_k = \frac{1}{2} \sum_{i=1}^{t} \|x^k - P_{C_i^k} x^k\|^2 + \frac{1}{2} \sum_{j=1}^{r} \|Ax^k - P_{Q_j^k} Ax^k\|^2 < 10^{-4}$. Let $x^0 = x^1$. The details are shown in Table 4.

Table 4. Results of Armijo-line search and self-adaptive algorithms for Example 3.

Initial Point		NP1	HP1	NP2	HP2
$r = t = 10$, $m = n = 20$					
$x^0 = x^1 = e_1$	No. of Iter	477	357	2268	1700
	cpu(time)	1.2453	0.9267	1.0516	0.8038
$x^0 = x^1 = 50 * e_1$	No. of Iter	757	564	3291	2470
	cpu(time)	1.6205	1.2805	1.5623	1.1023
$x^0 = x^1 = 100 * rand(n, 1)$	No. of Iter	996	737	4323	3231
	cpu(time)	1.9087	1.4396	1.9696	1.4493
$r = t = 20$, $m = 40$, $n = 50$					
$x^0 = x^1 = e_1$	No. of Iter	1256	941	5336	4001
	cpu(time)	12.1310	4.0061	5.9165	4.0919
$x^0 = x^1 = 50 * e_1$	No. of Iter	1492	1105	6917	5221
	cpu(time)	12.6430	8.2382	12.9631	9.4880
$x^0 = x^1 = 100 * rand(n, 1)$	No. of Iter	2101	1835	9936	9226
	cpu(time)	16.4070	13.2868	14.9611	12.8079
$r = t = 40$, $m = n = 60$					
$x^0 = x^1 = e_1$	No. of Iter	1758	1317	8328	6245
	cpu(time)	48.2570	38.0668	30.6759	23.4267
$x^0 = x^1 = 50 * e_1$	No. of Iter	2503	1777	12,905	8677
	cpu(time)	59.2127	44.7915	49.5823	32.6868
$x^0 = x^1 = 100 * rand(n, 1)$	No. of Iter	2274	1474	18,781	13,952
	cpu(time)	58.2569	38.1917	72.6622	54.9814

We can see from Table 4 that the convergence rate is improved by inertial perturbations for both algorithms. In most cases, the algorithm with step size determined by Armijo-line search outperforms the one with self-adaptive step size in the number of iterations, whereas the latter outperforms the former in CPU time.

6. Conclusions

In this paper, for the MSSFP, we present two relaxed CQ algorithms with different kinds of self-adaptive step size and discuss their bounded perturbation resilience. Treating appropriate inertial terms as bounded perturbations, we construct the inertial acceleration versions of the corresponding algorithms. For the real-life LASSO problem and three experimental examples, we numerically compare the performance with or without inertial perturbation of the algorithms and also compare the performance of the proposed algorithms with Yang's algorithm [3], and Liu and Tang's algorithm [29]. The results show the efficiency of the proposed algorithms and the validity of the inertial perturbation.

Author Contributions: Both the authors contributed equally to this work. Both authors have read and agreed to the published version of the manuscript.

Funding: The authors was supported by the National Natural Science Foundation under Grant No. 61503385 and No. 11705279, and the Fundamental Research Funds for the Central Universities No. 3122018L004.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would also like to thank the editors and the reviewers for their constructive suggestions and comments, which greatly improved the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Byrne, C. Iterative oblique projection onto convex sets and the split feasibility problem. *Inverse Probl.* **2002**, *18*, 441–453. [CrossRef]
2. Byrne, C. A unified treatment of some iterative algorithms in signal processing and image reconstruction. *Inverse Probl.* **2004**, *20*, 103–120. [CrossRef]
3. Yang, Q.Z. The relaxed CQ algorithm solving the split feasibility problem. *Inverse Probl.* **2004**, *20*, 1261–1266. [CrossRef]
4. Censor, Y.; Elfving, T.; Kopf, N.; Bortfeld, T. The multiple-sets spilt feasibility problem and its applications for inverse problems. *Inverse Probl.* **2005**, *21*, 2071–2084. [CrossRef]
5. He, S.N.; Zhao, Z.Y.; Luo, B. A relaxed self-adaptive CQ algorithm for the multiple-sets split feasibility problem. *Optimization* **2015**, *64*, 1907–1918. [CrossRef]
6. Yang, Q.Z. On variable-step relaxed projection algorithm for variational inequalities. *J. Math. Anal. Appl.* **2005**, *302*, 166–179. [CrossRef]
7. López, G.; Martín-Márquez, V.; Wang, F.H.; Xu, H.K. Solving the split feasibility problem without prior knowledge of matrix norms. *Inverse Probl.* **2012**, *28*, 374–389. [CrossRef]
8. Gibali, A.; Liu, L.W.; Tang, Y.C. Note on the the modified relaxation CQ algorithm for the split feasibility problem. *Optim. Lett.* **2018**, *12*, 817–830. [CrossRef]
9. Chen, Y.; Guo, Y.; Yu, Y.; Chen, R. Self-adaptive and relaxed self-adaptive projection methods for solving the multiple-set split feasibility problem. *Abstr. Appl. Anal.* **2012**, *2012*, 958040. [CrossRef]
10. Xu, H.K. Iterative methods for the split feasibility problem in infinite-dimensional Hilbert spaces. *Inverse Probl.* **2010**, *26*, 105018. [CrossRef]
11. Xu, H.K. A variable Krasnosel'skii-Mann algorithm and the multiple-set split feasibility problem. *Inverse Probl.* **2006**, *22*, 2021–2034. [CrossRef]
12. Yao, Y.H.; Postolache, M.; Liou, Y.C. Strong convergence of a self-adaptive method for the split feasibility problem. *Fixed Point Theory Appl.* **2013**, *2013*, 201. [CrossRef]
13. Guo, Y.N.; Cui, W.; Guo, Y.S. Perturbation resilience of proximal gradient algorithm for composite objectives. *J. Nonlinear Sci. Appl.* **2017**, *10*, 5566–5575. [CrossRef]
14. Guo, Y.N.; Cui, W. Strong convergence and bounded perturbation resilience of a modified proximal gradient algorithm. *J. Inequalities Appl.* **2018**, *2018*, 103. [CrossRef] [PubMed]
15. Pakkaranang, N.; Kumam, P.; Berinde, V.; Suleiman, Y.I. Superiorization methodology and perturbation resilience of inertial proximal gradient algorithm with application to signal recovery. *J. Supercomput.* **2020**, *76*, 9456–9477. [CrossRef]
16. Jin, W.M.; Censor, Y.; Jiang, M. Bounded perturbation resilience of projected scaled gradient methods. *Comput. Optim. Appl.* **2016**, *63*, 365–392. [CrossRef]
17. Xu, H.K. Bound perturbation resilience and superiorization techniques for the projected scaled gradient method. *Inverse Probl.* **2017**, *33*, 044008. [CrossRef]
18. Dong, Q.L.; Gibali, A.; Jiang, D.; Tang, Y. Bounded perturbation resilience of extragradient-type methods and their applications. *J. Inequalities Appl.* **2017**, *2017*, 280. [CrossRef] [PubMed]
19. Duan, P.C.; Zheng, X.B. Bounded perturbation resilience of generalized viscosity iterative algorithm for split variational inclusion problem. *Appl. Set-Valued Anal. Optim.* **2020**, *2*, 49–61.
20. Censor, Y.; Zaslavski, A.J. Convergence and perturbation resilience of dynamic string-averaging projection methods. *Comput. Optim. Appl.* **2013**, *54*, 65–76. [CrossRef]
21. Kaewyong, N.; Sitthithakerngkiet, K. A self-adaptive algorithm for the common solution of the split minimization problem and the fixed point problem. *Axioms* **2021**, *10*, 109. [CrossRef]
22. Shehu, Y.; Dong, Q.L.; Liu, L.L. Global and linear convergence of alternated inertial methods for split feasibility problems. *R. Acad. Sci.* **2021**, *115*, 53.
23. Li, H.Y.; Wu, Y.L.; Wang, F.H. New inertial relaxed CQ algorithms for solving split feasibility problems in Hilbert spaces. *J. Math.* **2021**, *2021*, 6624509. [CrossRef]
24. Bauschke, H.H.; Combettes, P.L. *Convex Analysis and Monotone Operator Theory in Hilbert Space*; Springer: London, UK, 2011.
25. Bauschke, H.H.; Combettes, P.L. A weak-to-strong convergence principle for Fejér-monntone methods in Hilbert spaces. *Math. Oper. Res.* **2001**, *26*, 248–264. [CrossRef]
26. Wen, M.; Peng, J.G.; Tang, Y.C. A cyclic and simultaneous iterative method for solving the multiple-sets split feasibility problem. *J. Optim. Theory Appl.* **2015**, *166*, 844–860. [CrossRef]
27. Sun, J.; Dang, Y.Z.; Xu, H.L. Inertial accelerated algorithms for solving a split feasibility problem. *J. Ind. Manag. Optim.* **2017**, *13*, 1383–1394.
28. Tang, Y.C.; Zhu, C.X.; Yu, H. Iterative methods for solving the multiple-sets split feasibility problem with splitting self-adaptive step size. *Fixed Point Theory Appl.* **2015**, *2015*, 178. [CrossRef]

29. Liu, L.W.; Tang, Y.C. Several iterative algorithms for solving the split common fixed point problem of directed operators with applications. *Optim. A J. Math. Program. Oper. Res.* **2016**, *65*, 53–65.
30. Zhao, J.L.; Yang, Q.Z. Self-adaptive projection methods for the multiple-sets split feasibility problem. *Inverse Probl.* **2011**, *27*, 035009. [CrossRef]