J. Math. Math. Sci. (2026) nn–nnn

An Inertial and Relaxed Projection Algorithm for the Split Feasibility Problem in Hilbert Spaces

Nguyen Van Hung (Hanoi, Vietnam)

(Received Mar. 23, 2026; accepted May 8, 2026)

Abstract. In this paper, we propose a new self-adaptive inertial and relaxed CQ algorithm for solving the split feasibility problem in real Hilbert spaces. The method incorporates alternated inertial extrapolation, a conjugate-gradient-inspired search direction, and a self-adaptive step-size strategy, which eliminates the need for operator norm evaluations or line search procedures. Under suitable assumptions, we establish strong convergence of the generated sequence to the minimum-norm solution of the problem. To illustrate the practical applicability of the proposed method, we apply it to an elastic net regularization model. The numerical experiment shows that the algorithm converges effectively and is suitable for high-dimensional regression problems.

1. Introduction

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be real Hilbert spaces, $\mathcal{C} \subseteq \mathcal{H}_1$ and $\mathcal{Q} \subseteq \mathcal{H}_2$ be nonempty closed convex sets, and let $\mathcal{F} : \mathcal{H}_1 \to \mathcal{H}_2$ be a bounded linear operator. The *Split Feasibility Problem* (SFP) consists of finding

$$\text{Find } u^* \in \mathcal{C} \text{ such that } \mathcal{F}u^* \in \mathcal{Q}. \tag{1.1}$$

That is, one seeks a point in $\mathcal{C}$ whose image under $\mathcal{F}$ lies in $\mathcal{Q}$. The solution set of (1.1) is denoted by

$$\Omega = \{u^* \in \mathcal{C} \mid \mathcal{F}u^* \in \mathcal{Q}\} = \mathcal{C} \cap \mathcal{F}^{-1}(\mathcal{Q}).$$

Key words and phrases: Split Feasibility Problem; Hilbert Space; Inertial Method; Conjugate-Gradient-Type Direction; Adaptive Step Size; Strong Convergence
2020 Mathematics Subject Classification: 47J25, 47J05, 65K10, 90C25

Since $\mathcal{C}$ and $\mathcal{Q}$ are nonempty closed convex sets and $\mathcal{F}$ is a bounded linear operator, the set Ω is a nonempty closed convex subset of $\mathcal{H}_1$. The SFP was introduced by Censor and Elfving [4] and has attracted considerable attention due to its broad range of applications in science and engineering, including X-ray computed tomography [5], machine learning [15], medical image reconstruction, signal processing, and game theory, particularly Nash equilibrium models [7, 16].

A classical approach for solving (1.1) is Byrne's CQ algorithm [2, 3], defined by

$$x^{k+1} = P_{\mathcal{C}}\Big(x^k - \gamma \mathcal{F}^*(I - P_{\mathcal{Q}})\mathcal{F}x^k\Big), \tag{1.2}$$

where I is the identity operator on $\mathcal{H}_2$, $P_{\mathcal{C}}$ and $P_{\mathcal{Q}}$ are the metric projections onto $\mathcal{C}$ and $\mathcal{Q}$, respectively, $\mathcal{F}^*$ is the adjoint of $\mathcal{F}$, and $\gamma \in (0, 2/\|\mathcal{F}\|^2)$. This algorithm can be interpreted as a gradient–projection method applied to a constrained convex minimization reformulation of the SFP.

Xu [20] established weak convergence of (1.2) in infinite-dimensional Hilbert spaces. To ensure strong convergence, Wang and Xu [21] proposed the modified iteration

$$x^{k+1} = P_{\mathcal{C}}\Big[(1 - \alpha_k)\Big(x^k - \gamma \mathcal{F}^*(I - P_{\mathcal{Q}})\mathcal{F}x^k\Big)\Big], \tag{1.3}$$

where $\gamma \in (0, 2/\|\mathcal{F}\|^2)$ and the sequence $\{\alpha_k\}$ satisfies

$$\alpha_k \in (0,1), \quad \lim_{k\to\infty} \alpha_k = 0, \quad \sum_{k=1}^{\infty} \alpha_k = \infty, \tag{1.4}$$

$$\text{either } \sum_{k=1}^{\infty} |\alpha_{k+1} - \alpha_k| < \infty \text{ or } \lim_{k\to\infty} \frac{|\alpha_{k+1} - \alpha_k|}{\alpha_k} = 0. \tag{1.5}$$

They proved that $\{x^k\}$ converges strongly to the minimum-norm solution of (1.1). Later, Yu et al. [24] showed that condition (1.5) can be removed.

In many applications, however, computing the projections onto $\mathcal{C}$ and $\mathcal{Q}$ may be computationally demanding. To address this difficulty, Fukushima [8] introduced an approximation technique based on supporting half-spaces. Assume that

$$\mathcal{C} = \{x \in \mathcal{H}_1 \mid g(x) \le 0\}, \tag{1.6}$$

$$\mathcal{Q} = \{y \in \mathcal{H}_2 \mid h(y) \le 0\}, \tag{1.7}$$

where g and h are convex and weakly lower semicontinuous functions. At each iteration k, the sets $\mathcal{C}$ and $\mathcal{Q}$ are replaced by the half-spaces

$$\mathcal{C}^k = \left\{x \in \mathcal{H}_1 \mid g(x^k) + \langle \xi^k, x - x^k\rangle \le 0\right\}, \quad \xi^k \in \partial g(x^k), \tag{1.8}$$

$$\mathcal{Q}^k = \left\{y \in \mathcal{H}_2 \mid h(\mathcal{F}x^k) + \langle \zeta^k, y - \mathcal{F}x^k\rangle \le 0\right\}, \quad \zeta^k \in \partial h(\mathcal{F}x^k). \tag{1.9}$$

Based on this idea, Yang [22] proposed the relaxed CQ algorithm

$$x^{k+1} = P_{\mathcal{C}^k}\Big(x^k - \gamma \mathcal{F}^*(I - P_{\mathcal{Q}^k})\mathcal{F}x^k\Big), \tag{1.10}$$

where $\gamma \in (0, 2/\|\mathcal{F}\|^2)$. Since projections onto half-spaces admit closed-form expressions, this scheme is computationally attractive. To avoid estimating $\|\mathcal{F}\|$, López et al. [10] introduced the adaptive step size

$$\gamma_k = \rho_k \frac{\|(I - P_{\mathcal{Q}^k})\mathcal{F}x^k\|^2}{\|\mathcal{F}^*(I - P_{\mathcal{Q}^k})\mathcal{F}x^k\|^2}, \tag{1.11}$$

where

$$\rho_k \in (0, 2) \text{ and } \inf_{k \geq 1} \rho_k(2 - \rho_k) > 0. \tag{1.12}$$

To further accelerate convergence, Polyak's inertial extrapolation technique [14] has been incorporated into SFP algorithms (see, e.g., [11, 12, 17, 18, 19]). Although inertial methods often enhance practical performance, they may generate oscillatory behavior. To mitigate this drawback, Mu and Peng [13] proposed an alternated inertial strategy that restores partial monotonicity along a suitable subsequence, thereby improving stability and convergence speed.

Another effective acceleration mechanism consists in replacing the steepest descent direction $-\mathcal{F}^*(I - P_{\mathcal{Q}^k})\mathcal{F}x^k$ with a conjugate gradient-type direction

$$d^k = -\mathcal{F}^*(I - P_{\mathcal{Q}^k})\mathcal{F}x^k + \beta_k d^{k-1}, \tag{1.13}$$

where $\beta_k \geq 0$. Such techniques have been successfully applied to the SFP; see, for instance, Che et al. [6].

Motivated by the above developments, the present paper aims to design a new algorithm for solving (1.1) in infinite-dimensional Hilbert spaces that simultaneously incorporates:

- an alternated inertial extrapolation scheme,
- a conjugate gradient-type search direction,
- an adaptive step size rule, and
- a relaxation mechanism based on supporting half-spaces.

We prove strong convergence of the proposed method under mild assumptions and introduce an improved choice of inertial parameters that simplifies implementation.

The remainder of the paper is organized as follows. Section 2 recalls basic preliminaries. Section 3 presents the proposed algorithm and establishes its strong convergence. Section 4 presents a numerical experiment illustrating the effectiveness of the proposed algorithm in elastic net regularization. Finally, Section 5 concludes the paper.

2. Preliminaries

In this section, we recall several notions and auxiliary results that will be used in the convergence analysis of the proposed algorithm. Throughout the paper, let $\mathcal{H}$ be a real Hilbert space endowed with the inner product $\langle \cdot, \cdot \rangle$ and the induced norm $\| \cdot \|$. For a sequence $\{x^k\} \subset \mathcal{H}$, the notation $x^k \rightharpoonup x$ means that $\{x^k\}$ converges weakly to x, while $x^k \to x$ denotes strong convergence. The weak ω-limit set of $\{x^k\}$, denoted by $\omega(x^k)$, is defined as the collection of all points $x \in \mathcal{H}$ for which there exists a subsequence $\{x^{k_\ell}\}$ satisfying $x^{k_\ell} \rightharpoonup x$ as $\ell \to \infty$.

For any $x, y \in \mathcal{H}$ and $\lambda \in \mathbb{R}$, the following identities hold:

$$\|x + y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2, \tag{2.1}$$

$$\|\lambda x + (1-\lambda)y\|^2 = \lambda\|x\|^2 + (1-\lambda)\|y\|^2 - \lambda(1-\lambda)\|x - y\|^2. \tag{2.2}$$

Let $\mathcal{D}$ be a nonempty closed convex subset of $\mathcal{H}$. An operator $T : \mathcal{D} \to \mathcal{H}$ is said to be *nonexpansive* if

$$\|Tx - Ty\| \le \|x - y\| \quad \text{for all } x, y \in \mathcal{D};$$

η-inverse strongly monotone (for some $\eta > 0$) if

$$\langle Tx - Ty, x - y\rangle \ge \eta\|Tx - Ty\|^2 \quad \text{for all } x, y \in \mathcal{D}.$$

For any $x \in \mathcal{H}$, the metric projection of x onto $\mathcal{D}$ is defined by

$$P_{\mathcal{D}}x = \arg\min\left\{\|u - x\| \mid u \in \mathcal{D}\right\}.$$

It is well known that the projection operator enjoys the following properties.

Lemma 2.1 ([1]). *Let $\mathcal{D}$ be a nonempty closed convex subset of $\mathcal{H}$. Then, for any $x \in \mathcal{H}$ and $u \in \mathcal{D}$, the following statements hold:*

(i) $\langle x - P_{\mathcal{D}}x,\, u - P_{\mathcal{D}}x\rangle \le 0$*;*

(ii) *both $P_{\mathcal{D}}$ and $I - P_{\mathcal{D}}$ are nonexpansive;*

(iii) *both $P_{\mathcal{D}}$ and $I - P_{\mathcal{D}}$ are* 1*-inverse strongly monotone.*

We next recall some basic facts from convex analysis.

A function $f : \mathcal{H} \to (-\infty, +\infty]$ is *convex* if

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y) \quad \text{for all } x, y \in \mathcal{H},\ \lambda \in (0,1).$$

It is said to be *weakly lower semicontinuous* at $x \in \mathcal{H}$ if

$$f(x) \le \liminf_{k\to\infty} f(x^k) \quad \text{whenever } x^k \rightharpoonup x.$$

The *subdifferential inequality* associated with a convex function f reads

$$f(x) \ge f(y) + \langle \xi, x - y\rangle, \tag{2.3}$$

where $\xi \in \mathcal{H}$. Any ξ satisfying (2.3) is called a subgradient of f at y, and the set of all such subgradients is denoted by $\partial f(y)$.

Finally, we recall a technical lemma that will be used in the convergence proof.

Lemma 2.2 ([9, Lemma 8])**.** *Let $\{\Phi_k\}$ be a sequence of nonnegative real numbers satisfying*

$$\Phi_{k+1} \le (1 - p_k)\Phi_k + p_k q_k, \qquad \Phi_{k+1} \le \Phi_k - r_k + s_k, \quad k \ge 0,$$

where $\{p_k\} \subset (0, 1)$, $\{r_k\}$ is a sequence of nonnegative real numbers, and $\{q_k\}, \{s_k\} \subset \mathbb{R}$. Assume that

(i) $\sum_{k=0}^{\infty} p_k = \infty$*;*

(ii) $\lim_{k\to\infty} s_k = 0$*;*

(iii) *for any subsequence $\{k_\ell\}$, the condition $r_{k_\ell} \to 0$ implies* $\limsup_{\ell\to\infty} q_{k_\ell} \le 0$.

Then $\lim_{k\to\infty} \Phi_k = 0$.

3. Main Result

In this section, we introduce a modified self-adaptive relaxed CQ algorithm for solving the Split Feasibility Problem (1.1). The proposed method combines an alternated inertial extrapolation technique with a two-term conjugate gradient-type direction. We then establish its strong convergence in real Hilbert spaces. To proceed, we first state the standing assumptions.

Assumptions.

(A1) The constraint sets $\mathcal{C} \subseteq \mathcal{H}_1$ and $\mathcal{Q} \subseteq \mathcal{H}_2$ are defined by (1.6)–(1.7).

(A2) The subdifferentials ∂g and ∂h are bounded on bounded subsets of $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively.

(A3) The operator $\mathcal{F} : \mathcal{H}_1 \to \mathcal{H}_2$ is bounded and linear.

(A4) The solution set Ω of (1.1) is nonempty.

In addition, the control parameters are assumed to satisfy the following conditions:

(C1) The sequence $\{\alpha_k\} \subset (0,1)$ satisfies (1.4).

(C2) The sequence $\{\eta_k\} \subset (0,1)$ satisfies $\lim_{k\to\infty} \frac{\eta_k}{\alpha_k} = 0$.

(C3) The sequence $\{\delta_k\} \subset (0,1)$ satisfies $\liminf_{k\to\infty} \delta_k > 0$.

(C4) The inertial parameter θ_k satisfies $0 < \theta_k < \frac{1-\delta_{k-1}}{\delta_{k-1}(1+\delta_k)}$.

(C5) The sequence $\{\rho_k\} \subset \left(0, \frac{2}{\varepsilon}\right)$, for some $\varepsilon > 0$, satisfies $\inf_{k\geq 1} \rho_k \left(\frac{2}{\varepsilon} - \rho_k\right) > 0$.

(C6) The sequence $\{\widetilde{\rho}_k\} \subset (0,4)$ satisfies $\liminf_{k\to\infty} \widetilde{\rho}_k(4-\widetilde{\rho}_k) > 0$.

(C7) The parameter $\beta_k \in [0, 1/2)$ satisfies $\beta_k \leq \alpha_k^2$.

We next introduce the relaxed sets associated with a given sequence $\{x^k\} \subset \mathcal{H}_1$. For each $k \geq 0$, let $\mathcal{C}^k$ and $\mathcal{Q}^k$ be the supporting half-spaces defined in (1.8)–(1.9). These sets serve as outer approximations of the original constraint sets $\mathcal{C}$ and $\mathcal{Q}$, respectively. It is straightforward to verify that

$$\mathcal{C} \subseteq \mathcal{C}^k, \qquad \mathcal{Q} \subseteq \mathcal{Q}^k \quad \text{for all } k \geq 0.$$

For each $x \in \mathcal{H}_1$, define

$$f_k(x) := \frac{1}{2} \left\| (I - P_{\mathcal{Q}^k}) \mathcal{F} x \right\|^2. \tag{3.1}$$

The function f_k is convex and Fréchet differentiable on $\mathcal{H}_1$, and its gradient is given by $\nabla f_k(x) = \mathcal{F}^*(I - P_{\mathcal{Q}^k})\mathcal{F}x$, which is Lipschitz continuous.

Algorithm 3.1 presents the proposed self-adaptive relaxed CQ method combining alternated inertial extrapolation with a two-term conjugate gradient-like direction.

Algorithm 3.1. Initialization: Choose constants $\varepsilon > 0$ and $\beta \in (0,1]$, together with parameter sequences $\{\alpha_k\}$, $\{\eta_k\}$, $\{\delta_k\}$, $\{\theta_k\}$, $\{\rho_k\}$, $\{\widetilde{\rho}_k\}$, and $\{\beta_k\}$ satisfying conditions (C1)–(C7).

Select initial points $x^0, x^1 \in \mathcal{H}_1$ and set $d^0 := -\widetilde{\gamma}_0 \nabla f_0(x^0)$, where

$$\widetilde{\gamma}_0 := \begin{cases} \dfrac{\widetilde{\rho}_0 f_0(x^0)}{\|\nabla f_0(x^0)\|^2}, & \text{if } \nabla f_0(x^0) \neq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Set $k = 1$.

Iterative Step: Given x^{k-1} and x^k for $k \geq 1$, perform:

Step 1. Compute

$$w^k := \begin{cases} x^k, & \text{if } k \text{ is even,} \\ x^k + \theta_k(x^k - x^{k-1}), & \text{if } k \text{ is odd.} \end{cases}$$

Step 2. Compute

$$\gamma_k := \begin{cases} \dfrac{\rho_k f_k(x^k)}{\|\nabla f_k(x^k)\|^2}, & \text{if } \nabla f_k(x^k) \neq 0, \\ 0, & \text{otherwise,} \end{cases}$$

and set

$$y^k := (1 - \eta_k)\Big(w^k - \varepsilon\gamma_k \nabla f_k(x^k)\Big).$$

Step 3. Compute

$$\widetilde{\gamma}_k := \begin{cases} \dfrac{\widetilde{\rho}_k f_k(y^k)}{\|\nabla f_k(y^k)\|^2}, & \text{if } \nabla f_k(y^k) \neq 0, \\ 0, & \text{otherwise,} \end{cases}$$

$$d^k := -\widetilde{\gamma}_k \nabla f_k(y^k) + \beta\beta_k d^{k-1}, \qquad v^k := y^k + d^k.$$

Step 4. Compute

$$z^k := P_{\mathcal{C}^k}\big[(1 - \alpha_k)v^k\big], \qquad x^{k+1} := (1 - \delta_k)w^k + \delta_k z^k.$$

Set $k := k + 1$ and return to Step 1.

Lemma 3.1. *Assume that the sequence $\{(I - P_{\mathcal{Q}^k})\mathcal{F}y^k\}$ is bounded. Let $\{d^k\}$ be the sequence generated by Algorithm 3.1. Then $\{d^k\}$ is bounded.*

Proof. We prove the boundedness of $\{d^k\}$ by induction.

Since

$$\nabla f_k(y^k) = \mathcal{F}^*(I - P_{\mathcal{Q}^k})\mathcal{F}y^k,$$

the boundedness of $\{(I - P_{\mathcal{Q}^k})\mathcal{F}y^k\}$ and the boundedness of the linear operator $\mathcal{F}^*$ imply that $\{\nabla f_k(y^k)\}$ is bounded.

Moreover, from the definition

$$\widetilde{\gamma}_k := \begin{cases} \dfrac{\widetilde{\rho}_k f_k(y^k)}{\|\nabla f_k(y^k)\|^2}, & \text{if } \nabla f_k(y^k) \neq 0, \\ 0, & \text{otherwise,} \end{cases}$$

and the condition $\widetilde{\rho}_k \in (0,4)$, it follows that $\{\widetilde{\gamma}_k \nabla f_k(y^k)\}$ is bounded.

Define

$$M := \frac{2}{2-\beta} \sup_{k\geq 0} \|\widetilde{\gamma}_k \nabla f_k(y^k)\|.$$

Clearly $M < \infty$.

For $k = 0$, by definition $d^0 = -\tilde{\gamma}_0 \nabla f_0(x^0)$, so

$$\|d^0\| = \|\tilde{\gamma}_0 \nabla f_0(x^0)\| \leq \sup_{k\geq 0} \|\tilde{\gamma}_k \nabla f_k(y^k)\| \leq \frac{2-\beta}{2} M \leq M.$$

Assume that $\|d^k\| \leq M$ for some $k \geq 0$. From Step 3 of Algorithm 3.1,

$$d^{k+1} = -\widetilde{\gamma}_{k+1} \nabla f_{k+1}(y^{k+1}) + \beta\beta_{k+1} d^k.$$

Thus,

$$\|d^{k+1}\| \leq \|\widetilde{\gamma}_{k+1} \nabla f_{k+1}(y^{k+1})\| + \beta\beta_{k+1}\|d^k\|.$$

By condition (C7), $\beta_{k+1} < \frac{1}{2}$ for all $k \geq 0$. Using the inductive hypothesis and the definition of M, we obtain

$$\|d^{k+1}\| \leq \frac{2-\beta}{2} M + \frac{\beta}{2} M = M.$$

By induction, $\|d^k\| \leq M$ for all $k \geq 0$, and hence $\{d^k\}$ is bounded. ■

Lemma 3.2. *Let $\{y^k\}$ and $\{z^k\}$ be the sequences generated by Algorithm 3.1. Then for all $k \geq 1$ and $u \in \Omega$, the following estimates hold:*

$$\begin{aligned}\|y^k - u\|^2 \leq{}& \|w^k - u\|^2 + (1-\eta_k)\|x^k - w^k\|^2 + \eta_k\|u\|^2 \\ &- 2\varepsilon\rho_k(2-\varepsilon\rho_k)(1-\eta_k)\frac{f_k^2(x^k)}{\|\nabla f_k(x^k)\|^2},\end{aligned} \tag{3.2}$$

and

$$\begin{aligned}\|z^k - u\|^2 \leq{}& (1-\alpha_k)\|w^k - u\|^2 + [\alpha_k + (1-\alpha_k)\eta_k]\|u\|^2 \\ &- 2\varepsilon\rho_k(2-\varepsilon\rho_k)(1-\eta_k)(1-\alpha_k)\frac{f_k^2(x^k)}{\|\nabla f_k(x^k)\|^2} \\ &+ \beta^2\beta_k^2(1-\alpha_k)\|d^{k-1}\|^2 + 2\beta\beta_k(1-\alpha_k)\big\langle (I - \widetilde{\gamma}_k \nabla f_k)y^k - u, d^{k-1}\big\rangle \\ &+ (1-\eta_k)(1-\alpha_k)\|x^k - w^k\|^2 - \widetilde{\rho}_k(4-\widetilde{\rho}_k)(1-\alpha_k)\frac{f_k^2(y^k)}{\|\nabla f_k(y^k)\|^2}.\end{aligned} \tag{3.3}$$

Proof. Let $u \in \Omega$. Then $u \in \mathcal{C}$ and $\mathcal{F}u \in \mathcal{Q}$. Since $\mathcal{Q} \subseteq \mathcal{Q}^k$ for all $k \geq 1$, we have $\mathcal{F}u = P_{\mathcal{Q}^k}\mathcal{F}u$. Using the 1-inverse strongly monotone property of $I - P_{\mathcal{Q}^k}$, we obtain

$$\begin{aligned} \langle \nabla f_k(x^k), x^k - u \rangle &= \big\langle (I - P_{\mathcal{Q}^k})\mathcal{F}x^k - (I - P_{\mathcal{Q}^k})\mathcal{F}u, \mathcal{F}x^k - \mathcal{F}u \big\rangle \\ &\geq \|(I - P_{\mathcal{Q}^k})\mathcal{F}x^k\|^2 = 2f_k(x^k). \end{aligned} \tag{3.4}$$

Estimate for y^k. Using identity (2.1), we compute

$$\begin{aligned} \|w^k - \varepsilon\gamma_k \nabla f_k(x^k) - u\|^2 &= \|w^k - u\|^2 + \varepsilon^2\gamma_k^2\|\nabla f_k(x^k)\|^2 - 2\varepsilon\gamma_k\langle \nabla f_k(x^k), x^k - u\rangle \\ &\quad + 2\varepsilon\gamma_k\langle \nabla f_k(x^k), x^k - w^k\rangle. \end{aligned}$$

By the Cauchy-Schwarz inequality and Young's inequality, we have

$$2\varepsilon\gamma_k\langle \nabla f_k(x^k), x^k - w^k\rangle \leq \|x^k - w^k\|^2 + \varepsilon^2\gamma_k^2\|\nabla f_k(x^k)\|^2.$$

Combining the above two inequalities and applying (3.4), we obtain

$$\begin{aligned} &\|w^k - \varepsilon\gamma_k \nabla f_k(x^k) - u\|^2 \\ &\quad \leq \|w^k - u\|^2 + \|x^k - w^k\|^2 + 2\varepsilon^2\gamma_k^2\|\nabla f_k(x^k)\|^2 - 4\varepsilon\gamma_k f_k(x^k). \end{aligned} \tag{3.5}$$

Now assume first that $\nabla f_k(x^k) \neq 0$. Substituting $\gamma_k = \frac{\rho_k f_k(x^k)}{\|\nabla f_k(x^k)\|^2}$ into (3.5), using condition (C5), we obtain

$$\begin{aligned} \|w^k - \varepsilon\gamma_k \nabla f_k(x^k) - u\|^2 &\leq \|w^k - u\|^2 + \|x^k - w^k\|^2 - 2\varepsilon\rho_k(2 - \varepsilon\rho_k)\frac{f_k^2(x^k)}{\|\nabla f_k(x^k)\|^2} \\ &\leq \|w^k - u\|^2 + \|x^k - w^k\|^2. \end{aligned} \tag{3.6}$$

If $\nabla f_k(x^k) = 0$, then $\gamma_k = 0$ and the same inequality holds.

Using the convexity of $\|\cdot\|^2$, (C2), we deduce

$$\begin{aligned} \|y^k - u\|^2 &= \|(1 - \eta_k)(w^k - \varepsilon\gamma_k\nabla f_k(x^k) - u) + \eta_k(-u)\|^2 \\ &\leq (1 - \eta_k)\|w^k - \varepsilon\gamma_k\nabla f_k(x^k) - u\|^2 + \eta_k\|u\|^2. \end{aligned}$$

Combining the above estimates yields (3.2).

Estimate for z^k. From Step 3, $v^k = (I - \widetilde{\gamma}_k\nabla f_k)y^k + \beta\beta_k d^{k-1}$. A computation similar to the previous one gives

$$\|(I - \widetilde{\gamma}_k\nabla f_k)y^k - u\|^2 \leq \|y^k - u\|^2 - \widetilde{\rho}_k(4 - \widetilde{\rho}_k)\frac{f_k^2(y^k)}{\|\nabla f_k(y^k)\|^2}. \tag{3.7}$$

Using identity (2.1), we obtain

$$\|v^k - u\|^2 = \|(I - \widetilde{\gamma}_k \nabla f_k)y^k - u\|^2 + \beta^2\beta_k^2\|d^{k-1}\|^2 + 2\beta\beta_k\langle (I - \widetilde{\gamma}_k \nabla f_k)y^k - u, d^{k-1}\rangle. \tag{3.8}$$

Finally, from Step 4, the nonexpansiveness of $P_{\mathcal{C}^k}$ and $u \in \mathcal{C}^k$ imply

$$\|z^k - u\|^2 \leq (1 - \alpha_k)\|v^k - u\|^2 + \alpha_k\|u\|^2.$$

Combining this inequality with (3.7), (3.8) and the previous estimate for $\|y^k - u\|^2$ yields (3.3). ■

Lemma 3.3. *Assume that the sequences $\{(I - P_{\mathcal{Q}^k})\mathcal{F}y^k\}$ and $\{(I - \widetilde{\gamma}_k\nabla f_k)y^k - u\}$ are bounded for every $u \in \Omega$. Let $\{x^k\}$ be the sequence generated by Algorithm 3.1. Then, for any $u \in \Omega$, the even subsequence $\{\|x^{2k} - u\|\}$ is bounded.*

Proof. Let $u \in \Omega$. It follows from Step 4, (2.2), and (3.3) that

$$\begin{aligned}
\|x^{k+1} - u\|^2 &= \|(1 - \delta_k)(w^k - u) + \delta_k(z^k - u)\|^2 \\
&= (1 - \delta_k)\|w^k - u\|^2 + \delta_k\|z^k - u\|^2 - \delta_k(1 - \delta_k)\|w^k - z^k\|^2 \\
&\leq (1 - \alpha_k\delta_k)\|w^k - u\|^2 + \delta_k[\alpha_k + (1 - \alpha_k)\eta_k]\|u\|^2 \\
&\quad - 2\varepsilon\delta_k\rho_k(2 - \varepsilon\rho_k)(1 - \eta_k)(1 - \alpha_k)\frac{f_k^2(x^k)}{\|\nabla f_k(x^k)\|^2} \\
&\quad + \beta^2\beta_k^2\delta_k(1 - \alpha_k)\|d^{k-1}\|^2 + 2\beta\beta_k\delta_k(1 - \alpha_k)\big\langle (I - \widetilde{\gamma}_k\nabla f_k)y^k - u, d^{k-1}\big\rangle \\
&\quad + \delta_k(1 - \alpha_k)(1 - \eta_k)\|w^k - x^k\|^2 - \delta_k\widetilde{\rho}_k\big(4 - \widetilde{\rho}_k\big)(1 - \alpha_k)\frac{f_k^2(y^k)}{\|\nabla f_k(y^k)\|^2} \\
&\quad - \delta_k(1 - \delta_k)\|w^k - z^k\|^2.
\end{aligned} \tag{3.9}$$

It follows from Step 1 and (3.9) that

$$\begin{aligned}
\|x^{2k+1} - u\|^2 \leq\ & (1 - \alpha_{2k}\delta_{2k})\|x^{2k} - u\|^2 + \delta_{2k}[\alpha_{2k} + (1 - \alpha_{2k})\eta_{2k}]\|u\|^2 \\
& + \beta^2\beta_{2k}^2\delta_{2k}(1 - \alpha_{2k})\|d^{2k-1}\|^2 \\
& + 2\beta\beta_{2k}\delta_{2k}(1 - \alpha_{2k})\big\langle (I - \widetilde{\gamma}_{2k}\nabla f_{2k})y^{2k} - u, d^{2k-1}\big\rangle \\
& - 2\varepsilon\delta_{2k}\rho_{2k}(2 - \varepsilon\rho_{2k})(1 - \eta_{2k})(1 - \alpha_{2k})\frac{f_{2k}^2(x^{2k})}{\|\nabla f_{2k}(x^{2k})\|^2} \\
& - \delta_{2k}\widetilde{\rho}_{2k}\big(4 - \widetilde{\rho}_{2k}\big)(1 - \alpha_{2k})\frac{f_{2k}^2(y^{2k})}{\|\nabla f_{2k}(y^{2k})\|^2} \\
& - \delta_{2k}(1 - \delta_{2k})\|w^{2k} - z^{2k}\|^2.
\end{aligned} \tag{3.10}$$

From Steps 1 and 4, we have

$$\begin{aligned}\|x^{2k+1}-x^{2k}\|^2 &= \|(1-\delta_{2k})(w^{2k}-x^{2k})+\delta_{2k}(z^{2k}-x^{2k})\|^2\\ &= \delta_{2k}^2\big\|z^{2k}-x^{2k}\big\|^2.\end{aligned} \tag{3.11}$$

It follows from Step 1, (2.2), (3.10), and (3.11) that

$$\begin{aligned}\|w^{2k+1}-u\|^2 &= \|(1+\theta_{2k+1})(x^{2k+1}-u)-\theta_{2k+1}(x^{2k}-u)\|^2\\ &= (1+\theta_{2k+1})\|x^{2k+1}-u\|^2-\theta_{2k+1}\|x^{2k}-u\|^2\\ &\quad+\theta_{2k+1}(1+\theta_{2k+1})\|x^{2k+1}-x^{2k}\|^2\\ &\le (1-\alpha_{2k}\delta_{2k})\|x^{2k}-u\|^2\\ &\quad+\delta_{2k}(1+\theta_{2k+1})\Big\{[\alpha_{2k}+(1-\alpha_{2k})\eta_{2k}]\|u\|^2\\ &\qquad+\beta^2\beta_{2k}^2(1-\alpha_{2k})\|d^{2k-1}\|^2\\ &\qquad+2\beta\beta_{2k}(1-\alpha_{2k})\big\langle(I-\widetilde{\gamma}_{2k}\nabla f_{2k})y^{2k}-u,d^{2k-1}\big\rangle\\ &\qquad+(\delta_{2k}\theta_{2k+1}-1+\delta_{2k})\|z^{2k}-x^{2k}\|^2\\ &\qquad-\widetilde{\rho}_{2k}(4-\widetilde{\rho}_{2k})(1-\alpha_{2k})\frac{f_{2k}^2(y^{2k})}{\|\nabla f_{2k}(y^{2k})\|^2}\\ &\qquad-2\varepsilon\rho_{2k}(2-\varepsilon\rho_{2k})(1-\eta_{2k})(1-\alpha_{2k})\frac{f_{2k}^2(x^{2k})}{\|\nabla f_{2k}(x^{2k})\|^2}\Big\}.\end{aligned} \tag{3.12}$$

It follows from Step 1 and (3.11) that

$$\|w^{2k+1}-x^{2k+1}\|^2 \le \delta_{2k}^2\theta_{2k+1}(1+\theta_{2k+1})\|z^{2k}-x^{2k}\|^2. \tag{3.13}$$

From (3.9), (3.12), and (3.13), we have

$$\begin{aligned}\|x^{2k+2}-u\|^2 &\le (1-\alpha_{2k}\delta_{2k})\|x^{2k}-u\|^2\\ &\quad+\delta_{2k}(1+\theta_{2k+1})\big[(\alpha_{2k}+\eta_{2k})\|u\|^2+\phi_{2k}\big]\\ &\quad+\delta_{2k+1}\big[(\alpha_{2k+1}+\eta_{2k+1})\|u\|^2+\phi_{2k+1}\big]\\ &\quad-\delta_{2k}(1+\theta_{2k+1})\big[(1-\delta_{2k})-\delta_{2k}\theta_{2k+1}(1+\delta_{2k+1})\big]\|z^{2k}-x^{2k}\|^2\\ &\quad-\delta_{2k}(1+\theta_{2k+1})\widetilde{\rho}_{2k}(4-\widetilde{\rho}_{2k})(1-\alpha_{2k})\frac{f_{2k}^2(y^{2k})}{\|\nabla f_{2k}(y^{2k})\|^2}\\ &\quad-2\delta_{2k}(1+\theta_{2k+1})\varepsilon\rho_{2k}(2-\varepsilon\rho_{2k})(1-\eta_{2k})(1-\alpha_{2k})\frac{f_{2k}^2(x^{2k})}{\|\nabla f_{2k}(x^{2k})\|^2}\\ &\quad-2\varepsilon\delta_{2k+1}\rho_{2k+1}(2-\varepsilon\rho_{2k+1})(1-\eta_{2k+1})(1-\alpha_{2k+1})\frac{f_{2k+1}^2(x^{2k+1})}{\|\nabla f_{2k+1}(x^{2k+1})\|^2}\end{aligned}$$

$$- \delta_{2k+1}\widetilde{\rho}_{2k+1}(4-\widetilde{\rho}_{2k+1})(1-\alpha_{2k+1})\frac{f_{2k+1}^2(y^{2k+1})}{\|\nabla f_{2k+1}(y^{2k+1})\|^2}$$

$$- \delta_{2k+1}(1-\delta_{2k+1})\|w^{2k+1}-z^{2k+1}\|^2. \tag{3.14}$$

where

$$\phi_{2k} = \beta^2\beta_{2k}^2\|d^{2k-1}\|^2 + 2\beta\beta_{2k}\big\langle (I-\widetilde{\gamma}_{2k}\nabla f_{2k})y^{2k}-u, d^{2k-1}\big\rangle.$$

Taking

$$M_1 = \sup_{k\geq 1}\Big[(1+\theta_{2k+1})\Big(\big(1+\frac{\eta_{2k}}{\alpha_{2k}}\big)\|u\|^2+\frac{\phi_{2k}}{\alpha_{2k}}\Big)+\frac{\delta_{2k+1}}{\delta_{2k}}\Big(\frac{\alpha_{2k+1}}{\alpha_{2k}}+\frac{\eta_{2k+1}}{\alpha_{2k}}\Big)\|u\|^2+\frac{\phi_{2k+1}}{\alpha_{2k}}\Big],$$

together with conditions (C1)–(C7) and (3.14), we obtain

$$\begin{aligned}\|x^{2k+2}-u\|^2 &\leq (1-\alpha_{2k}\delta_{2k})\|x^{2k}-u\|^2+\alpha_{2k}\delta_{2k}M_1\\ &\leq \max\{\|x^{2k}-u\|^2, M_1\}\\ &\leq \cdots \leq \max\{\|x^0-u\|^2, M_1\}.\end{aligned} \tag{3.15}$$

Hence, the sequence $\{\|x^{2k+2}-u\|\}$ is bounded. Therefore, the even subsequence $\{x^{2k}\}$ of the sequence $\{x^k\}$ generated by Algorithm 3.1 is bounded.

Consequently, from inequality (3.10), it follows directly that the odd subsequence $\{x^{2k+1}\}$ is also bounded. ■

Proposition 3.1. *Let $\{x^k\}$ be the sequence generated by Algorithm 3.1. Then, for any $u \in \Omega$, there exists a constant $M_2 > 0$ such that, for all $k \geq 1$, the following inequalities hold:*

$$\|x^{2k+2}-u\|^2 \leq (1-\alpha_{2k}\delta_{2k})\|x^{2k}-u\|^2+\alpha_{2k}\delta_{2k}q_{2k}, \tag{3.16}$$

$$\|x^{2k+2}-u\|^2 \leq \|x^{2k}-u\|^2 - r_{2k} + s_{2k}, \tag{3.17}$$

where $\{r_{2k}\}$ is a sequence of nonnegative real numbers, $\{q_{2k}\}$ and $\{s_{2k}\}$ are real sequences, and $\sup_{k\geq 1}|q_{2k}| \leq M_2$.

Proof. Let $u \in \Omega$. From Step 2, (C2), (2.1), and (3.6),

$$\begin{aligned}\big\|y^k-u\big\|^2 &= \big\|(1-\eta_k)\big[w^k-\varepsilon\gamma_k\nabla f_k(x^k)-u\big]-\eta_k u\big\|^2\\ &= (1-\eta_k)^2\big\|w^k-\varepsilon\gamma_k\nabla f_k(x^k)-u\big\|^2+\eta_k^2\big\|u\big\|^2\\ &\quad + 2\eta_k(1-\eta_k)\langle w^k-\varepsilon\gamma_k\nabla f_k(x^k)-u, -u\rangle\\ &\leq (1-\eta_k)^2\big[\|w^k-u\|^2+\|x^k-w^k\|^2\big]+\eta_k^2\|u\|^2\\ &\quad + 2\eta_k(1-\eta_k)\langle w^k-u, -u\rangle\end{aligned}$$

$$
\begin{aligned}
&\quad + 2\eta_k(1-\eta_k)\varepsilon\gamma_k\langle \nabla f_k(x^k), u\rangle \\
&\le \|w^k-u\|^2 + \eta_k\Big[\eta_k\|u\|^2 \\
&\quad + 2(1-\eta_k)\big(\langle w^k-u,-u\rangle + \varepsilon\gamma_k\langle \nabla f_k(x^k), u\rangle\big)\Big] \\
&\quad + (1-\eta_k)\|x^k-w^k\|^2. \qquad (3.18)
\end{aligned}
$$

It follows from Step 3, (2.1), (3.7), and (3.18) that

$$
\begin{aligned}
\|v^k-u\|^2 &= \|(I-\widetilde{\gamma}_k\nabla f_k)y^k - u + \beta\beta_k d^{k-1}\|^2 \\
&= \|(I-\widetilde{\gamma}_k\nabla f_k)y^k - u\|^2 + \beta^2\beta_k^2\|d^{k-1}\|^2 \\
&\quad + 2\beta\beta_k\langle (I-\widetilde{\gamma}_k\nabla f_k)y^k - u, d^{k-1}\rangle \\
&\le \|w^k-u\|^2 + \eta_k^2\|u\|^2 \\
&\quad + \eta_k\Big[2(1-\eta_k)\big(\langle w^k-u,-u\rangle + \varepsilon\gamma_k\langle \nabla f_k(x^k), u\rangle\big)\Big] \\
&\quad + (1-\eta_k)\|x^k-w^k\|^2 + \beta^2\beta_k^2\|d^{k-1}\|^2 \\
&\quad + 2\beta\beta_k\langle (I-\widetilde{\gamma}_k\nabla f_k)y^k - u, d^{k-1}\rangle. \qquad (3.19)
\end{aligned}
$$

Now, Step 4, using the convexity of the function $\|\cdot\|^2$, the non-expansiveness of the projection operator $P_{\mathcal{C}^k}$, $u \in \mathcal{C}^k$, (2.1), Lemma 2.1(ii), and (3.19), we obtain that

$$
\begin{aligned}
\|z^k-u\|^2 &\le (1-\alpha_k)^2\|v^k-u\|^2 + \alpha_k^2\|u\|^2 + 2\alpha_k(1-\alpha_k)\langle v^k-u,-u\rangle \\
&\le (1-\alpha_k)\big\|w^k-u\big\|^2 + [\alpha_k^2 + (1-\alpha_k)\eta_k^2]\|u\|^2 \\
&\quad + 2\alpha_k(1-\alpha_k)\langle v^k-u,-u\rangle \\
&\quad + 2\eta_k(1-\eta_k)(1-\alpha_k)\big[\langle w^k-u,-u\rangle \\
&\qquad + \varepsilon\gamma_k\langle \nabla f_k(x^k), u\rangle\big] \\
&\quad + (1-\alpha_k)\Big[(1-\eta_k)\|x^k-w^k\|^2 \\
&\qquad + \beta^2\beta_k^2\|d^{k-1}\|^2 \\
&\qquad + 2\beta\beta_k\langle (I-\widetilde{\gamma}_k\nabla f_k)y^k - u, d^{k-1}\rangle\Big]. \qquad (3.20)
\end{aligned}
$$

It follows from Step 4, (2.2), and (3.20) that

$$
\begin{aligned}
\|x^{k+1}-u\|^2 &\le (1-\alpha_k\delta_k)\big\|w^k-u\big\|^2 \\
&\quad + \delta_k\Big\{[\alpha_k^2 + (1-\alpha_k)\eta_k^2]\|u\|^2 \\
&\qquad + 2\alpha_k(1-\alpha_k)\langle v^k-u,-u\rangle \\
&\qquad + 2\eta_k(1-\eta_k)(1-\alpha_k)\big[\langle w^k-u,-u\rangle
\end{aligned}
$$

$$
\begin{aligned}
&\qquad + \varepsilon\gamma_k\langle\nabla f_k(x^k), u\rangle\big]\\
&+ (1-\alpha_k)\Big[(1-\eta_k)\|x^k - w^k\|^2\\
&\qquad + \beta^2\beta_k^2\|d^{k-1}\|^2\\
&\qquad + 2\beta\beta_k\langle(I - \widetilde{\gamma}_k\nabla f_k)y^k - u, d^{k-1}\rangle\Big]\Big\}\\
&- \delta_k(1-\delta_k)\|w^k - z^k\|^2. \qquad (3.21)
\end{aligned}
$$

It follows from Step 1 and (3.21) that

$$
\begin{aligned}
\|x^{2k+1} - u\|^2 \le{}& (1-\alpha_{2k}\delta_{2k})\big\|w^{2k} - u\big\|^2\\
&+ \delta_{2k}\Big\{[\alpha_{2k}^2 + (1-\alpha_{2k})\eta_{2k}^2]\|u\|^2\\
&\quad + 2\alpha_{2k}(1-\alpha_{2k})\langle v^{2k} - u, -u\rangle\\
&\quad + 2\eta_{2k}(1-\eta_{2k})(1-\alpha_{2k})\big[\langle w^{2k} - u, -u\rangle\\
&\qquad + \varepsilon\gamma_{2k}\langle\nabla f_{2k}(x^{2k}), u\rangle\big]\\
&\quad + \beta^2\beta_{2k}^2\|d^{2k-1}\|^2\\
&\quad + 2\beta\beta_{2k}\langle(I - \widetilde{\gamma}_{2k}\nabla f_{2k})y^{2k} - u, d^{2k-1}\rangle\Big\}\\
&\quad - \delta_{2k}(1-\delta_{2k})\|w^{2k} - z^{2k}\|^2. \qquad (3.22)
\end{aligned}
$$

From Step 1, (2.2), (3.11), and (3.22), we get

$$
\begin{aligned}
\|w^{2k+1} - u\|^2 \le{}& (1-\alpha_{2k}\delta_{2k})\|x^{2k} - u\|^2\\
&+ \delta_{2k}(1+\theta_{2k+1})\Big\{[\alpha_{2k}^2 + (1-\alpha_{2k})\eta_{2k}^2]\|u\|^2\\
&\quad + 2\alpha_{2k}(1-\alpha_{2k})\langle v^{2k} - u, -u\rangle\\
&\quad + 2\eta_{2k}(1-\eta_{2k})(1-\alpha_{2k})\big[\langle w^{2k} - u, -u\rangle\\
&\qquad + \varepsilon\gamma_{2k}\langle\nabla f_{2k}(x^{2k}), u\rangle\big]\\
&\quad + \beta^2\beta_{2k}^2\|d^{2k-1}\|^2\\
&\quad + 2\beta\beta_{2k}\langle(I - \widetilde{\gamma}_{2k}\nabla f_{2k})y^{2k} - u, d^{2k-1}\rangle\Big\}\\
&- \delta_{2k}(1+\theta_{2k+1})\big[(1-\delta_{2k}) - \delta_{2k}\theta_{2k+1}\big]\|z^{2k} - x^{2k}\|^2. \qquad (3.23)
\end{aligned}
$$

It follows from (3.21), (3.23), (3.13), and (C4) that

$$
\begin{aligned}
\|x^{2k+2}-u\|^2 \le{}& (1-\alpha_{2k}\delta_{2k})\|x^{2k}-u\|^2 \\
&+ \delta_{2k+1}\Big\{[\alpha_{2k+1}^2+(1-\alpha_{2k+1})\eta_{2k+1}^2]\|u\|^2 \\
&\quad + 2\alpha_{2k+1}(1-\alpha_{2k+1})\langle v^{2k+1}-u,-u\rangle \\
&\quad + 2\eta_{2k+1}(1-\eta_{2k+1})(1-\alpha_{2k+1}) \\
&\qquad \times \big[\langle w^{2k+1}-u,-u\rangle \\
&\qquad\quad + \varepsilon\gamma_{2k+1}\langle \nabla f_{2k+1}(x^{2k+1}),u\rangle\big] \\
&\quad + \beta^2\beta_{2k+1}^2\|d^{2k}\|^2 \\
&\quad + 2\beta\beta_{2k+1}\langle (I-\widetilde{\gamma}_{2k+1}\nabla f_{2k+1})y^{2k+1}-u,d^{2k}\rangle\Big\} \\
&+ \delta_{2k}(1+\theta_{2k+1})\Big\{[\alpha_{2k}^2+(1-\alpha_{2k})\eta_{2k}^2]\|u\|^2 \\
&\quad + 2\alpha_{2k}(1-\alpha_{2k})\langle v^{2k}-u,-u\rangle \\
&\quad + 2\eta_{2k}(1-\eta_{2k})(1-\alpha_{2k}) \\
&\qquad \times \big[\langle w^{2k}-u,-u\rangle \\
&\qquad\quad + \varepsilon\gamma_{2k}\langle \nabla f_{2k}(x^{2k}),u\rangle\big] \\
&\quad + \beta^2\beta_{2k}^2\|d^{2k-1}\|^2 \\
&\quad + 2\beta\beta_{2k}\langle (I-\widetilde{\gamma}_{2k}\nabla f_{2k})y^{2k}-u,d^{2k-1}\rangle\Big\} \\
&- \delta_{2k+1}(1-\delta_{2k+1})\|w^{2k+1}-z^{2k+1}\|^2 \\
&- \delta_{2k}(1+\theta_{2k+1})\big[(1-\delta_{2k})-\delta_{2k}\theta_{2k+1}(1+\delta_{2k+1})\big]\|z^{2k}-x^{2k}\|^2 \\
\le{}& (1-\alpha_{2k}\delta_{2k})\|x^{2k}-u\|^2+\delta_{2k}\alpha_{2k}q_{2k}.
\end{aligned}
$$

Therefore, we have (3.16) with

$$
\begin{aligned}
q_{2k} = \frac{1}{\alpha_{2k}\delta_{2k}}\Big\{&\delta_{2k+1}\Big[(\alpha_{2k+1}^2+(1-\alpha_{2k+1})\eta_{2k+1}^2)\|u\|^2 \\
&+ 2\alpha_{2k+1}(1-\alpha_{2k+1})\langle v^{2k+1}-u,-u\rangle \\
&+ 2\eta_{2k+1}(1-\eta_{2k+1})(1-\alpha_{2k+1})\Big[\langle w^{2k+1}-u,-u\rangle \\
&\qquad + \varepsilon\gamma_{2k+1}\langle \nabla f_{2k+1}(x^{2k+1}),u\rangle\Big] \\
&+ \beta^2\beta_{2k+1}^2\|d^{2k}\|^2 \\
&+ 2\beta\beta_{2k+1}\langle (I-\widetilde{\gamma}_{2k+1}\nabla f_{2k+1})y^{2k+1}-u,d^{2k}\rangle\Big] \\
&+ \delta_{2k}(1+\theta_{2k+1})\Big[(\alpha_{2k}^2+(1-\alpha_{2k})\eta_{2k}^2)\|u\|^2 \\
&+ 2\alpha_{2k}(1-\alpha_{2k})\langle v^{2k}-u,-u\rangle
\end{aligned}
$$

$$
\begin{aligned}
&+ 2\eta_{2k}(1-\eta_{2k})(1-\alpha_{2k})\Big[\langle w^{2k}-u, -u\rangle \\
&\qquad + \varepsilon\gamma_{2k}\langle \nabla f_{2k}(x^{2k}), u\rangle\Big] \\
&+ \beta^2\beta_{2k}^2\|d^{2k-1}\|^2 \\
&+ 2\beta\beta_{2k}\langle (I-\widetilde{\gamma}_{2k}\nabla f_{2k})y^{2k}-u, d^{2k-1}\rangle\Big]\Big\}.
\end{aligned}
$$

Since $\{x^{2k}\}$ is bounded (Lemma 3.3), the sequences $\{w^k\}$, $\{y^k\}$, $\{v^k\}$, and $\{d^k\}$ (Lemma 3.1) are all bounded. By the Cauchy–Schwarz inequality and conditions (C2), (C7), all terms in q_{2k} are bounded, so there exists a constant $M_2 > 0$ such that

$$\sup_{k\geq 1} |q_{2k}| \leq M_2.$$

It follows from conditions (C1)–(C6) that there exist constants $m_1, m_2, m_3 > 0$ such that for all $k \geq 1$

$$
\begin{aligned}
&\delta_k(1+\theta_{k+1})\big[(1-\delta_k)-\delta_k\theta_{k+1}(1+\delta_{k+1})\big] \geq m_1, \\
&2\varepsilon\delta_k\rho_k(2-\varepsilon\rho_k)(1-\eta_k)(1-\alpha_k) \geq m_2, \\
&\delta_k\widetilde{\rho}_k(4-\widetilde{\rho}_k)(1-\alpha_k) \geq m_3.
\end{aligned}
$$

Combining this with (3.14), we obtain (3.17) with

$$
\begin{aligned}
r_{2k} &= m_1\|x^{2k}-z^{2k}\|^2 \\
&\quad + m_2\left[\frac{f_{2k}^2(x^{2k})}{\|\nabla f_{2k}(x^{2k})\|^2} + \frac{f_{2k+1}^2(x^{2k+1})}{\|\nabla f_{2k+1}(x^{2k+1})\|^2}\right] \\
&\quad + m_3\left[\frac{f_{2k}^2(y^{2k})}{\|\nabla f_{2k}(y^{2k})\|^2} + \frac{f_{2k+1}^2(y^{2k+1})}{\|\nabla f_{2k+1}(y^{2k+1})\|^2}\right],
\end{aligned}
$$

$$
\begin{aligned}
s_{2k} &= \delta_{2k}(1+\theta_{2k+1})\big[(\alpha_{2k}+\eta_{2k})\|u\|^2+\phi_{2k}\big] \\
&\quad + \delta_{2k+1}\big[(\alpha_{2k+1}+\eta_{2k+1})\|u\|^2+\phi_{2k+1}\big].
\end{aligned}
$$

■

Next, we apply Lemma 2.2 to establish the strong convergence of the even subsequence $\{x^{2k}\}$ generated by Algorithm 3.1 to the minimum-norm solution of the split feasibility problem, namely

$$u^* = P_\Omega(0).$$

Theorem 3.1. *Suppose that conditions* (A1)–(A4) *and* (C1)–(C7) *hold. Let* $\{x^k\}$ *be the sequence generated by Algorithm 3.1. Then the even subsequence* $\{x^{2k}\}$ *converges strongly to a point* $u^* \in \Omega$*, where*

$$u^* = P_\Omega(0).$$

Proof. Since the solution set Ω of (1.1) is nonempty, closed and convex, there exists a unique $u^* = P_\Omega(0)$. In particular, $u^* \in \Omega$. Now, it follows from Proposition 3.1 with u replaced by u^* that

$$\|x^{2k+2} - u^*\|^2 \leq (1 - \alpha_{2k}\delta_{2k})\|x^{2k} - u^*\|^2 + \alpha_{2k}\delta_{2k}q^*_{2k},$$
$$\|x^{2k+2} - u^*\|^2 \leq \|x^{2k} - u^*\|^2 - r_{2k} + s^*_{2k},$$

where

$$\begin{aligned} q^*_{2k} = \frac{1}{\alpha_{2k}\delta_{2k}} \Big\{ &\delta_{2k+1}\Big[(\alpha^2_{2k+1} + (1 - \alpha_{2k+1})\eta^2_{2k+1})\|u^*\|^2 \\ &+ 2\alpha_{2k+1}(1 - \alpha_{2k+1})\langle v^{2k+1} - u^*, -u^*\rangle \\ &+ 2\eta_{2k+1}(1 - \eta_{2k+1})(1 - \alpha_{2k+1})\Big[\langle w^{2k+1} - u^*, -u^*\rangle \\ &\qquad + \varepsilon\gamma_{2k+1}\langle \nabla f_{2k+1}(x^{2k+1}), u^*\rangle\Big] \\ &+ \beta^2\beta^2_{2k+1}\|d^{2k}\|^2 \\ &+ 2\beta\beta_{2k+1}\langle (I - \widetilde{\gamma}_{2k+1}\nabla f_{2k+1})y^{2k+1} - u^*, d^{2k}\rangle\Big] \\ &+ \delta_{2k}(1 + \theta_{2k+1})\Big[(\alpha^2_{2k} + (1 - \alpha_{2k})\eta^2_{2k})\|u^*\|^2 \\ &+ 2\alpha_{2k}(1 - \alpha_{2k})\langle v^{2k} - u^*, -u^*\rangle \\ &+ 2\eta_{2k}(1 - \eta_{2k})(1 - \alpha_{2k})\Big[\langle w^{2k} - u^*, -u^*\rangle \\ &\qquad + \varepsilon\gamma_{2k}\langle \nabla f_{2k}(x^{2k}), u^*\rangle\Big] \\ &+ \beta^2\beta^2_{2k}\|d^{2k-1}\|^2 \\ &+ 2\beta\beta_{2k}\langle (I - \widetilde{\gamma}_{2k}\nabla f_{2k})y^{2k} - u^*, d^{2k-1}\rangle\Big]\Big\}, \end{aligned}$$

$$\begin{aligned} s^*_{2k} = &\,\delta_{2k}(1 + \theta_{2k+1})\big[(\alpha_{2k} + \eta_{2k})\|u^*\|^2 + \phi_{2k}\big] \\ &+ \delta_{2k+1}\big[(\alpha_{2k+1} + \eta_{2k+1})\|u^*\|^2 + \phi_{2k+1}\big]. \end{aligned}$$

We apply Lemma 2.2 with $\Phi_{2k} = \|x^{2k} - u^*\|^2$, $p_k = \alpha_{2k}\delta_{2k}$, and verify the three conditions.

(i) Since $\liminf_{k\to\infty} \delta_k > 0$ by (C3), there exist $c > 0$ and $k_0 \geq 1$ such that $\delta_{2k} \geq c$ for all $k \geq k_0$. Together with $\sum_{k=0}^{\infty} \alpha_k = \infty$ from (C1), we obtain

$$\sum_{k=1}^{\infty} \alpha_{2k}\delta_{2k} \geq c \sum_{k=k_0}^{\infty} \alpha_{2k} = \infty.$$

(ii) From (C1),(C2),(C7), and the boundedness of $\{d^k\}$ (Lemma 3.1), we have $\phi_{2k}, \phi_{2k+1} \to 0$, and hence $s^*_{2k} \to 0$ as $k \to \infty$.

(iii) We now show that for every subsequence $\{k_\ell\}$ of $\{k\}$,

$$\lim_{\ell\to\infty} r_{2k_\ell} = 0 \quad \text{implies} \quad \limsup_{\ell\to\infty} q^*_{2k_\ell} \le 0.$$

Suppose that $\lim_{\ell\to\infty} r_{2k_\ell} = 0$. Then we obtain

$$\lim_{\ell\to\infty} \|x^{2k_\ell} - z^{2k_\ell}\| = 0, \tag{3.24}$$

$$\lim_{\ell\to\infty} \frac{f^2_{2k_\ell}(x^{2k_\ell})}{\|\nabla f_{2k_\ell}(x^{2k_\ell})\|^2} = 0, \quad \lim_{\ell\to\infty} \frac{f^2_{2k_\ell+1}(x^{2k_\ell+1})}{\|\nabla f_{2k_\ell+1}(x^{2k_\ell+1})\|^2} = 0, \tag{3.25}$$

$$\lim_{\ell\to\infty} \frac{f^2_{2k_\ell}(y^{2k_\ell})}{\|\nabla f_{2k_\ell}(y^{2k_\ell})\|^2} = 0, \quad \lim_{\ell\to\infty} \frac{f^2_{2k_\ell+1}(y^{2k_\ell+1})}{\|\nabla f_{2k_\ell+1}(y^{2k_\ell+1})\|^2} = 0. \tag{3.26}$$

Since $I - P_{\mathcal{Q}^{2k_\ell}}$ is nonexpansive by Lemma 2.1(ii), $\mathcal{F}$ is a bounded linear operator, we have

$$\begin{aligned}
\|\nabla f_{2k_\ell}(x^{2k_\ell})\|^2 &= \left\|\mathcal{F}^*(I - P_{\mathcal{Q}^{2k_\ell}})\mathcal{F}x^{2k_\ell}\right\|^2 \\
&= \left\|\mathcal{F}^*(I - P_{\mathcal{Q}^{2k_\ell}})\big(\mathcal{F}x^{2k_\ell} - \mathcal{F}u^*\big)\right\|^2 \\
&\le \|\mathcal{F}\|^2\left\|(I - P_{\mathcal{Q}^{2k_\ell}})\big(\mathcal{F}x^{2k_\ell} - \mathcal{F}u^*\big)\right\|^2 \\
&\le \|\mathcal{F}\|^2\|\mathcal{F}x^{2k_\ell} - \mathcal{F}u^*\|^2 \\
&\le \|\mathcal{F}\|^4\|x^{2k_\ell} - u^*\|^2.
\end{aligned}$$

Given that the sequence $\{x^{2k_\ell}\}$ is bounded and $\mathcal{F}$ is a bounded linear operator, it follows from the previous estimate that $\{\nabla f_{2k_\ell}(x^{2k_\ell})\}$ is bounded. Similarly, the sequence $\{\nabla f_{2k_\ell+1}(x^{2k_\ell+1})\}$ is bounded. This, together with (3.25) and (3.26), implies that

$$\begin{aligned}
&\lim_{\ell\to\infty} f_{2k_\ell}(x^{2k_\ell}) = 0, \quad \lim_{\ell\to\infty} f_{2k_\ell+1}(x^{2k_\ell+1}) = 0, \\
&\lim_{\ell\to\infty} f_{2k_\ell}(y^{2k_\ell}) = 0, \quad \lim_{\ell\to\infty} f_{2k_\ell+1}(y^{2k_\ell+1}) = 0.
\end{aligned}$$

Hence,

$$\lim_{\ell\to\infty} \|x^{2k_\ell} - z^{2k_\ell}\| = 0, \tag{3.27}$$

$$\lim_{\ell\to\infty} \|(I - P_{\mathcal{Q}^{2k_\ell}})\mathcal{F}x^{2k_\ell}\| = 0, \quad \lim_{\ell\to\infty} \|(I - P_{\mathcal{Q}^{2k_\ell+1}})\mathcal{F}x^{2k_\ell+1}\| = 0, \tag{3.28}$$

$$\lim_{\ell\to\infty} \|(I - P_{\mathcal{Q}^{2k_\ell}})\mathcal{F}y^{2k_\ell}\| = 0, \quad \lim_{\ell\to\infty} \|(I - P_{\mathcal{Q}^{2k_\ell+1}})\mathcal{F}y^{2k_\ell+1}\| = 0. \tag{3.29}$$

Now, we show that $\limsup_{\ell\to\infty} q^*_{2k_\ell} \le 0$. Indeed, take a subsequence $\{x^{2k_{\ell_m}}\}$ of $\{x^{2k_\ell}\}$ such that

$$\limsup_{\ell\to\infty}\langle x^{2k_\ell} - u^*, -u^*\rangle = \lim_{m\to\infty} \langle x^{2k_{\ell_m}} - u^*, -u^*\rangle. \tag{3.30}$$

Since the sequence $\{x^{2k_{\ell_m}}\}$ is bounded, there exists a subsequence of $\{x^{2k_{\ell_m}}\}$ which converges weakly to some $\widehat{u}$. Without loss of generality, we assume that

$$x^{2k_{\ell_m}} \rightharpoonup \widehat{u}.$$

Since $P_{\mathcal{Q}^{2k_{\ell_m}}} \mathcal{F} x^{2k_{\ell_m}} \in \mathcal{Q}^{2k_{\ell_m}}$, it follows from the definition of $\mathcal{Q}^{2k_{\ell_m}}$ that

$$h\big(\mathcal{F}x^{2k_{\ell_m}}\big) \leq \big\langle \zeta^{2k_{\ell_m}}, \mathcal{F}x^{2k_{\ell_m}} - P_{\mathcal{Q}^{2k_{\ell_m}}} \mathcal{F}x^{2k_{\ell_m}} \big\rangle,$$

where $\zeta^{2k_{\ell_m}} \in \partial h(\mathcal{F}x^{2k_{\ell_m}})$. By Assumption (A2), the sequence $\{\zeta^{2k_{\ell_m}}\}$ is bounded.

Combining this with (3.28), we obtain

$$h\big(\mathcal{F}x^{2k_{\ell_m}}\big) \leq \|\zeta^{2k_{\ell_m}}\| \, \|(I - P_{\mathcal{Q}^{2k_{\ell_m}}})\mathcal{F}x^{2k_{\ell_m}}\| \to 0 \quad \text{as } m \to \infty. \tag{3.31}$$

Since $\mathcal{F}$ is a bounded linear operator and $x^{2k_{\ell_m}} \rightharpoonup \widehat{u}$, we have

$$\mathcal{F}x^{2k_{\ell_m}} \rightharpoonup \mathcal{F}\widehat{u}.$$

Moreover, since h is convex and subdifferentiable, it is weakly lower semicontinuous. Therefore, it follows from (3.31) that

$$h(\mathcal{F}\widehat{u}) \leq \liminf_{m\to\infty} h(\mathcal{F}x^{2k_{\ell_m}}) \leq 0,$$

which implies that

$$\mathcal{F}\widehat{u} \in \mathcal{Q}.$$

On the other hand, from the definition of $\mathcal{C}^{2k_{\ell_m}}$ and the fact that $z^{2k_{\ell_m}} \in \mathcal{C}^{2k_{\ell_m}}$, we have

$$g(x^{2k_{\ell_m}}) \leq \big\langle \xi^{2k_{\ell_m}}, x^{2k_{\ell_m}} - z^{2k_{\ell_m}} \big\rangle, \tag{3.32}$$

where $\xi^{2k_{\ell_m}} \in \partial g(x^{2k_{\ell_m}})$. By Assumption (A2), the sequence $\{\xi^{2k_{\ell_m}}\}$ is bounded. Combining this with (3.27), we obtain

$$g(x^{2k_{\ell_m}}) \leq \|\xi^{2k_{\ell_m}}\| \, \|x^{2k_{\ell_m}} - z^{2k_{\ell_m}}\| \to 0 \quad \text{as } m \to \infty. \tag{3.33}$$

Since $x^{2k_{\ell_m}} \rightharpoonup \widehat{u}$ and g is convex and subdifferentiable, hence weakly lower semicontinuous, it follows from (3.33) that

$$g(\widehat{u}) \leq \liminf_{m\to\infty} g(x^{2k_{\ell_m}}) \leq 0,$$

which implies that $\widehat{u} \in \mathcal{C}$. Consequently, we conclude that $\widehat{u} \in \Omega$.

From Lemma 2.1(i) and (3.30), we obtain

$$\limsup_{\ell\to\infty}\langle x^{2k_\ell}-u^*,-u^*\rangle=\langle \widehat{u}-u^*,-u^*\rangle\le 0. \tag{3.34}$$

The following is also obtained by combining (3.11), (3.27), and condition (C3):

$$\|x^{2k_\ell+1}-x^{2k_\ell}\|^2=\delta_{2k_\ell}^2\|x^{2k_\ell}-z^{2k_\ell}\|^2\to 0 \quad \text{as } \ell\to\infty. \tag{3.35}$$

Combining Step 1, Step 2, the convexity of $\|\cdot\|^2$, condition (C2), (3.28), and (3.35), we obtain

$$\begin{aligned}\|y^{2k_\ell+1}-x^{2k_\ell}\|^2 &\le (1-\eta_{2k_\ell+1})\|w^{2k_\ell+1}-\varepsilon\gamma_{2k_\ell+1}\nabla f_{2k_\ell+1}(x^{2k_\ell+1})-x^{2k_\ell}\|^2\\ &\quad+\eta_{2k_\ell+1}\|x^{2k_\ell}\|^2\to 0 \quad \text{as } \ell\to\infty.\end{aligned} \tag{3.36}$$

It follows from (3.35) and (3.36) that

$$\|y^{2k_\ell+1}-x^{2k_\ell+1}\|^2\le\|y^{2k_\ell+1}-x^{2k_\ell}\|^2+\|x^{2k_\ell}-x^{2k_\ell+1}\|^2\to 0 \quad \text{as } \ell\to\infty. \tag{3.37}$$

From (3.34), (3.35), (3.36), and (3.37), we deduce that

$$\limsup_{\ell\to\infty}\langle y^{2k_\ell}-u^*,-u^*\rangle\le 0 \quad \text{and} \quad \limsup_{\ell\to\infty}\langle y^{2k_\ell+1}-u^*,-u^*\rangle\le 0. \tag{3.38}$$

From conditions (C1)–(C5) and (3.38), we finally obtain

$$\limsup_{\ell\to\infty} q^*_{2k_\ell}=2\limsup_{\ell\to\infty}\left[\langle y^{2k_\ell}-u^*,-u^*\rangle+\langle y^{2k_\ell+1}-u^*,-u^*\rangle\right]\le 0.$$

By Lemma 2.2, we conclude that

$$x^{2k}\to u^*=P_\Omega(0).$$

The proof is complete. ■

Remark 3.1. Since $x^{2k}\to u^*$ and, by (3.35), $\|x^{2k+1}-x^{2k}\|\to 0$, we have

$$\|x^{2k+1}-u^*\|\le\|x^{2k+1}-x^{2k}\|+\|x^{2k}-u^*\|\to 0 \quad \text{as } k\to\infty.$$

Hence, the odd-indexed subsequence $\{x^{2k+1}\}$ also converges strongly to u^*.

Consequently, the whole sequence $\{x^k\}$ generated by Algorithm 3.1 converges strongly to $u^*\in\Omega$, where $u^*=P_\Omega(0)$ is the minimum-norm solution of the split feasibility problem.

4. Numerical Example and Application

In this section, we present a numerical example to demonstrate the effectiveness and practical applicability of the proposed algorithm. Specifically, we consider an application to the Elastic Net regularization problem.

All simulations were implemented in `Python 3.12.13`, using standard scientific computing libraries such as `NumPy` and `SciPy`. The experiments were conducted on a personal computer equipped with an `AMD Ryzen 5 5600H` with `Radeon Graphics CPU @ 3.30GHz, 16 GB of RAM`, running `Windows 11`.

Example 4.1 (Elastic Net Regularization, [23]). In this example, we consider the elastic net regression problem

$$\min_{x\in\mathbb{R}^L} \left\{ \|y - \mathcal{F}x\|^2 + \lambda_1\|x\|_1 + \lambda_2\|x\|_2^2 \right\}, \tag{4.1}$$

where $\mathcal{F} \in \mathbb{R}^{K\times L}$ is a data matrix, $y \in \mathbb{R}^K$ is the observation vector, $\lambda_1, \lambda_2 > 0$ are regularization parameters, and $x \in \mathbb{R}^L$ is the unknown parameter vector. The elastic net combines the ℓ_1-penalty (promoting sparsity) and the ℓ_2-penalty (improving numerical stability), and is particularly effective when the predictors are correlated.

To relate (4.1) to the split feasibility framework, we consider its constrained formulation:

$$\min_{x\in\mathbb{R}^L} \|y - \mathcal{F}x\|^2 \quad \text{subject to} \quad (1-\lambda)\|x\|_1 + \lambda\|x\|_2^2 \le t, \tag{4.2}$$

where $\lambda = \frac{\lambda_2}{\lambda_1+\lambda_2} \in (0, 1)$ and $t > 0$ is chosen appropriately.

Define

$$\mathcal{C} = \{x \in \mathbb{R}^L \mid (1-\lambda)\|x\|_1 + \lambda\|x\|_2^2 \le t\},$$

and

$$\mathcal{Q} = \{z \in \mathbb{R}^K \mid \|z - y\|^2 \le \varphi\},$$

where $\varphi > 0$ is a tolerance parameter. Then problem (4.2) can be formulated as the split feasibility problem:

$$\text{find } x^* \in \mathcal{C} \quad \text{such that} \quad \mathcal{F}x^* \in \mathcal{Q}.$$

Since both $\mathcal{C}$ and $\mathcal{Q}$ are closed and convex sets, and $\mathcal{F}$ is a bounded linear operator, the problem fits naturally into the framework of Algorithm 3.1. Hence, the inertial and relaxed method developed in this paper can be applied to compute an approximate elastic net solution.

The split feasibility reformulation (4.2) is particularly well-suited to sparse regression for several reasons. First, the constraint set $\mathcal{C}$ encodes a combined sparsity-and-smoothness prior via the elastic net penalty $(1-\lambda)\|x\|_1 + \lambda\|x\|_2^2$: the ℓ_1 component promotes sparsity by driving small coefficients to zero, while the ℓ_2 component induces grouping of correlated predictors and ensures numerical stability when the data matrix $\mathcal{F}$ is ill-conditioned or the predictors are collinear. Second, the set $\mathcal{Q} = \{z \in \mathbb{R}^K \mid \|z - y\|^2 \leq \varphi\}$ encodes a noise-aware data-fidelity constraint: instead of requiring exact reproduction of the observations y, it permits a tolerance φ calibrated to the noise level, which is natural in statistical regression settings. Third, the splitting structure separates the signal-space constraint $\mathcal{C}$ from the measurement-space constraint $\mathcal{Q}$, enabling the relaxed CQ framework to work with half-space approximations of both sets without ever computing projections onto the original elastic net ball directly. This is computationally advantageous since the projection onto $\mathcal{C}$ does not admit a simple closed form in general, whereas the projection onto a supporting half-space is explicit. Consequently, the SFP framework simultaneously exploits the structural properties of the elastic net regularizer and avoids the computational bottleneck of evaluating exact projections, making it an efficient approach for high-dimensional sparse recovery.

Problem setup. For the numerical experiment, we set $K = 1500, L = 2000$. The matrix $\mathcal{F}$ was generated with entries independently sampled from the standard normal distribution and then normalized column-wise. The true signal $x^\dagger \in \mathbb{R}^L$ was constructed with $n = 50$ nonzero components, whose values were drawn uniformly from $[-2, 2]$. The observation vector was generated by

$$y = \mathcal{F}x^\dagger + \varepsilon,$$

where $\varepsilon \sim \mathcal{N}(0, \sigma^2 I_K)$ is a Gaussian noise vector with $\sigma = 10^{-3}$.

The tolerance parameter φ for the set $\mathcal{Q}$ is chosen based on the noise statistics. We set $\varphi = \tau K \sigma^2$ with $\tau = 1$ to ensure that the ground-truth signal $x^\dagger$ remains a feasible solution with high probability. The regularization parameters are chosen as $\lambda = 0.4$ and $t = 1.05 g(x^\dagger)$, where $g(x) = (1-\lambda)\|x\|_1 + \lambda\|x\|_2^2$, providing a 5% buffer so that the ground-truth signal lies in the interior of $\mathcal{C}$.

Parameter selection. The control parameters satisfying conditions (C1)–(C7) are chosen as follows:

$$\alpha_k = \frac{1}{100k+1}, \quad \eta_k = \frac{1}{(k+1)^2}, \quad \rho_k = 1.9, \quad \varepsilon = 1,$$

$$\widetilde{\rho}_k = 3.999, \quad \delta_k = 0.5, \quad \theta_k = 0.5, \quad \beta_k = \frac{1}{(100k+1)^2}, \quad \beta = 0.5.$$

The algorithm is initialized at

$$x^0 = x^1 = (1, 1, \dots, 1)^\top.$$

Stopping criterion. The iterative process is terminated whenever

$$\|x^{k+1} - x^k\|_2 \le 10^{-4}, \tag{4.3}$$

or when a maximum iteration limit $k_{\max} = 2000$ is reached.

To monitor convergence, we track the feasibility residual

$$\mathrm{Res}(k) := d(\mathcal{F}x^k, \mathcal{Q}) = \|(I - P_{\mathcal{Q}})(\mathcal{F}x^k)\|_2, \tag{4.4}$$

the mean squared error

$$\mathrm{MSE}(k) := \frac{1}{L}\|x^k - x^\dagger\|_2^2, \tag{4.5}$$

and the combined feasibility residual

$$\mathrm{TOL}(k) := \frac{1}{2}\Big[\|(I - P_{\mathcal{Q}^k})\mathcal{F}x^k\|_2^2 + \|(I - P_{\mathcal{C}^k})x^k\|_2^2\Big], \tag{4.6}$$

which simultaneously measures the degree of violation of both relaxed constraint sets $\mathcal{C}^k$ and $\mathcal{Q}^k$.

These three metrics serve complementary roles in assessing convergence. The mean squared error $\mathrm{MSE}(k)$ measures the approximation quality of the iterate x^k relative to the ground-truth signal $x^\dagger$. It reflects how well the recovered solution matches the true sparse vector and is relevant when $x^\dagger$ is available. The feasibility residual $\mathrm{Res}(k)$ quantifies the distance of $\mathcal{F}x^k$ from the original constraint set $\mathcal{Q}$. It provides a direct measure of feasibility with respect to the data-fitting constraint and is small when the output iterate satisfies $\mathcal{F}x^k \approx y$ to within the prescribed noise tolerance. The combined feasibility residual $\mathrm{TOL}(k)$, by contrast, measures the violation of the relaxed half-space approximations $\mathcal{C}^k$ and $\mathcal{Q}^k$ used in the algorithm. Since $\mathcal{C} \subseteq \mathcal{C}^k$ and $\mathcal{Q} \subseteq \mathcal{Q}^k$, this quantity is typically smaller than $\mathrm{Res}(k)$ and reflects the internal consistency of the iterative updates. Together, these indicators provide a comprehensive picture of both the numerical performance and the practical accuracy of the proposed method.

Numerical results. The algorithm terminated after **218 iterations** with a total CPU time of **4.124 seconds**. The final values of the monitored quantities are summarized in Table 1.

Quantity	**Final value**
Iterations	218
CPU time (s)	4.124
MSE(k)	6.48×10^{-5}
Res(k)	4.02×10^{-4}
$\|x^{k+1} - x^k\|_2$	9.87×10^{-5}
TOL(k)	$\approx 10^{-7}$

Table 1: Summary of numerical results for Algorithm 3.1 applied to the Elastic Net problem.

The numerical results are reported in Figures 1, 2, 3, and 4, which display the convergence behavior of Algorithm 3.1 in terms of four quantities: the mean squared error MSE(k), the successive difference $\|x^{k+1} - x^k\|_2$, the feasibility residual Res(k), and the combined feasibility residual TOL(k), each plotted against both the number of iterations and CPU time (in seconds). All four quantities decrease monotonically on a logarithmic scale. In particular, MSE(k) drops rapidly from $\mathcal{O}(1)$ to below $\mathcal{O}(10^{-4})$ within the first 100 iterations, while TOL(k) converges to $\mathcal{O}(10^{-7})$, significantly faster than Res(k). This suggests that the iterates satisfy the relaxed constraints $\mathcal{C}^k$ and $\mathcal{Q}^k$ with high precision, while the distance to the true feasible set $\mathcal{Q}$ decreases more gradually, as expected from the relaxation mechanism.

We note that the algorithm recovers a feasible approximate solution to the elastic net problem rather than the exact sparse signal $x^\dagger$: the SFP framework seeks a point satisfying the feasibility constraints, and the final iterate achieves Res(k) $\approx 4.02\times10^{-4}$, confirming feasibility with respect to $\mathcal{Q}$ up to a prescribed tolerance. These results confirm that Algorithm 3.1 converges effectively and is well-suited for high-dimensional regression problems of this type.

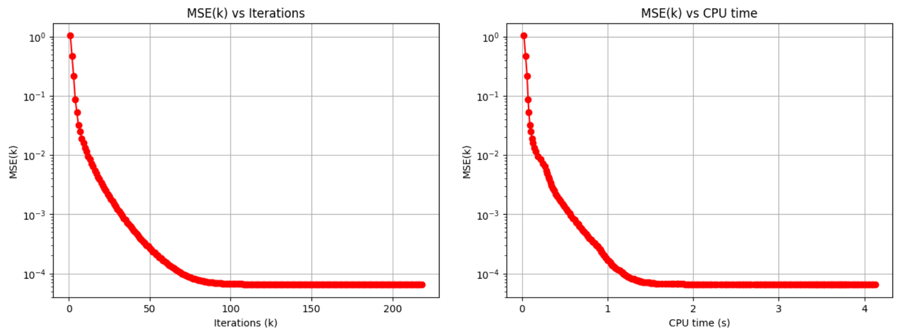

Figure 1: Convergence of MSE(k) for Algorithm 3.1: as a function of iterations (left) and CPU time in seconds (right).

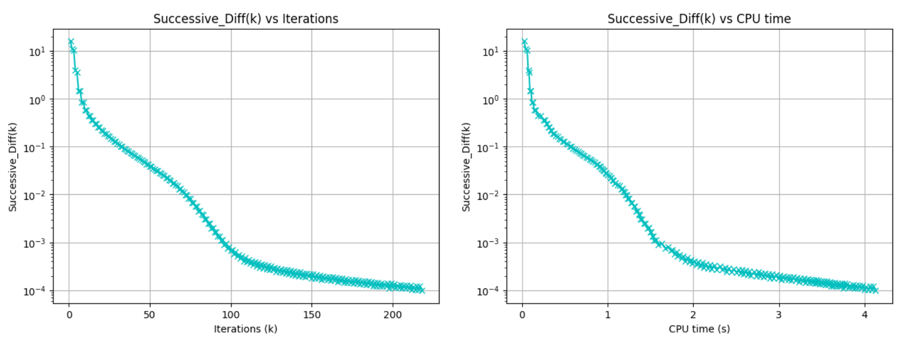

Figure 2: Convergence of the successive difference $\|x^{k+1} - x^k\|_2$ for Algorithm 3.1: as a function of iterations (left) and CPU time in seconds (right).

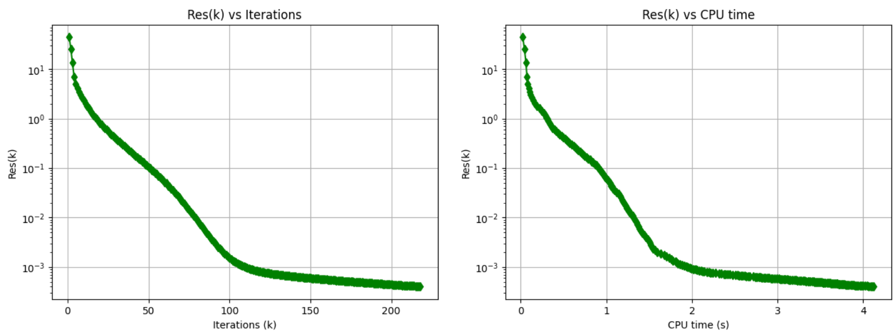

Figure 3: Convergence of the feasibility residual Res(k) for Algorithm 3.1: as a function of iterations (left) and CPU time in seconds (right).

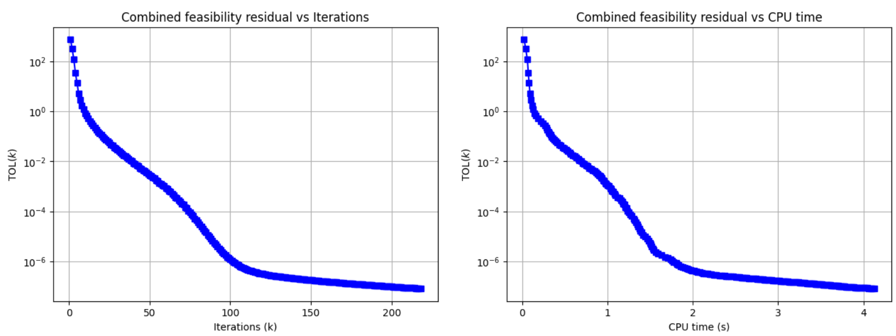

Figure 4: Convergence of the combined feasibility residual TOL(k) for Algorithm 3.1: as a function of iterations (left) and CPU time in seconds (right).

5. Conclusion

In this paper, we have proposed a new iterative algorithm for solving the split feasibility problem (Algorithm 3.1). The method combines alternated inertial extrapolation, a self-adaptive Polyak-type step size, a conjugate-gradient-inspired search direction, and a relaxed projection technique, aiming to improve convergence behavior and numerical stability.

Under suitable assumptions, we established the strong convergence of the proposed algorithm to the minimum-norm solution (Theorem 3.1).

To illustrate its practical applicability, we applied the method to an elastic net regularization problem. The numerical experiment demonstrates that the algorithm converges effectively and is suitable for handling high-dimensional regression models.

These results indicate that the proposed framework provides a feasible and flexible approach for solving structured feasibility problems arising in optimization and inverse problems.

Conflict of interest. We have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] **Bauschke H.H., Combettes P.L.,** Convex Analysis and Monotone Operator Theory in Hilbert Spaces, Springer, Berlin (2011).

[2] **Byrne C.,** *Iterative oblique projection onto convex sets and the split feasibility problem*, Inverse Probl. **18** (2002), no. 2, 441–453.

[3] **Byrne C.,** *A unified treatment of some iterative algorithms in signal processing and image reconstruction*, Inverse Probl. **20** (2004), no. 1, 103–120.

[4] **Censor Y., Elfving T.,** *A multi-projection algorithm using Bregman projections in a product space*, Numer. Algorithms **8** (1994), no. 2–4, 221–239.

[5] **Censor Y., Bortfeld T., Martin B., Trofimov A.,** *A unified approach for inversion problems in intensity-modulated radiation therapy*, Phys. Med. Biol. **51** (2006), no. 10, 2353–2365.

[6] **Che H., Zhuang Y., Wang Y., Chen H.,** *A relaxed inertial and viscosity method for split feasibility problem and applications to image recovery*, J. Global Optim. **87** (2023), no. 2–4, 619–639.

[7] **Dong Q.L., Li X.H., Rassias T.M.,** *Two projection algorithms for a class of split feasibility problems with jointly constrained Nash equilibrium models*, Optimization **70** (2021), no. 4, 871–897.

[8] **Fukushima M.,** *A relaxed projection method for variational inequalities*, Math. Program. **35** (1986), no. 1, 58–70.

[9] **He S., Yang C.,** *Solving the variational inequality problem defined on the intersection of finite level sets*, Abstr. Appl. Anal. **2013** (2013), Article ID 942315.

[10] **López G., Martín-Márquez V., Wang F., Xu H.K.,** *Solving the split feasibility problem without prior knowledge of matrix norms*, Inverse Probl. **28** (2012), no. 8, 085004.

[11] **Ly L.X., Thuy N.T.T., Anh N.Q.,** *Relaxed two-step inertial method for solving a class of split variational inequality problems with application to traffic analysis*, Optimization (2025).

[12] **Ly L.X., Thuy N.T.T., Tung T.T.,** *Two novel two-step inertial algorithms for a class of bilevel variational inequalities*, Optimization (2024).

[13] **Mu Z., Peng Y.,** *A note on the inertial proximal point method*, Stat. Optim. Inf. Comput. **3** (2015), no. 3, 241–248.

[14] **Polyak B.T.,** *Some methods of speeding up the convergence of iteration methods*, U.S.S.R. Comput. Math. Math. Phys. **4** (1964), no. 5, 1–17.

[15] **Suantai S., Peeyada P., Fulga A., Cholamjiak W.,** *Heart disease detection using inertial Mann relaxed CQ algorithms for split feasibility problems*, AIMS Math. **8** (2023), no. 8, 18898–18918.

[16] **Tan B., Qin X., Wang X.,** *Alternated inertial algorithms for split feasibility problems*, Numer. Algorithms **95** (2023), no. 2, 773–812.

[17] **Thuy N.T.T., Nghia N.T.,** *An inertial-type algorithm for a class of bilevel variational inequalities with the split variational inequality problem constraints*, Optimization (2023).

[18] **Thuy N.T.T., Tung T.T.,** *A self adaptive inertial algorithm for solving variational inequalities over the solution set of the split variational inequality problem*, Optim. Lett. **18** (2024), no. 6, 1619–1645.

[19] **Thuy N.T.T., Tung T.T.,** *Two relaxed CQ methods for the split feasibility problem with multiple output sets*, Bull. Malays. Math. Sci. Soc. **47** (2024), no. 68.

[20] **Xu H.K.,** *Iterative methods for the split feasibility problem in infinite-dimensional Hilbert spaces*, Inverse Probl. **26** (2010), no. 10, 105018.

[21] **Wang F., Xu H.K.,** *Approximating curve and strong convergence of the CQ algorithm for the split feasibility problem*, J. Inequal. Appl. **2010** (2010), Article ID 102085.

[22] **Yang Q.,** *The relaxed CQ algorithm solving the split feasibility problem*, Inverse Probl. **20** (2004), no. 4, 1261–1266.

[23] **Yu H., Zhan W.R., Wang F.H.,** *The ball-relaxed CQ algorithms for the split feasibility problem*, Optimization **67** (2018), no. 10, 1687–1699.

[24] **Yu X., Shahzad N., Yao Y.,** *Implicit and explicit algorithms for solving the split feasibility problem*, Optim. Lett. **6** (2012), no. 6, 1447–1462.

Nguyen Van Hung
Faculty of Mathematics and Informatics,
Hanoi University of Science and Technology (HUST)
Hanoi
Vietnam
hungalberto2805@gmail.com