symmetry

Article

Projected-Reflected Subgradient-Extragradient Method and Its Real-World Applications

Aviv Gibali [1,2,*,†], Olaniyi S. Iyiola [3,†], Lanre Akinyemi [4,†] and Yekini Shehu [5,†]

1 Department of Mathematics, ORT Braude College, Karmiel 2161002, Israel
2 The Center for Mathematics and Scientific Computation, U. Haifa, Mt. Carmel, Haifa 3498838, Israel
3 Department of Mathematics and Physical Sciences, California University of Pennsylvania, California, PA 15419, USA; iyiola@calu.edu
4 Department of Mathematics, Prairie View A&M University, Prairie View, TX 77446, USA; laakinyemi@pvamu.edu
5 Department of Mathematics, Zhejiang Normal University, Jinhua 321004, China; yekini.shehu@unn.edu.ng
* Correspondence: avivg@braude.ac.il
† These authors contributed equally to this work.

Abstract: Our main focus in this work is the classical variational inequality problem with Lipschitz continuous and pseudo-monotone mapping in real Hilbert spaces. An adaptive reflected subgradient-extragradient method is presented along with its weak convergence analysis. The novelty of the proposed method lies in the fact that only one projection onto the feasible set in each iteration is required, and there is no need to know/approximate the Lipschitz constant of the cost function a priori. To illustrate and emphasize the potential applicability of the new scheme, several numerical experiments and comparisons in tomography reconstruction, Nash–Cournot oligopolistic equilibrium, and more are presented.

Keywords: subgradient-extragradient; reflected step; variational inequality; pseudo-monotone mapping; Lipschitz mapping

MSC: 47H09; 47J20; 47J05; 47J25

Citation: Gibali, A.; Iyiola, O.S.; Akinyemi, L.; Shehu, Y. Projected-Reflected Subgradient-Extragradient Method and Its Real-World Applications. *Symmetry* **2021**, *13*, 489. https://doi.org10.3390/sym13030489

Academic Editor: Palle E. T. Jorgensen

Received: 5 February 2021
Accepted: 11 March 2021
Published: 16 March 2021

1. Introduction

In this paper, we focus on the classical variational inequality (VI) problem, as can be found in Fichera [1,2], Stampacchia [3], and Kinderlehrer and Stampacchia [4], defined in real Hilbert space H. Given a nonempty, closed, and convex set of $C \subseteq H$ and a continuous mapping $A : H \to H$, the variational inequality (VI) problem consists of finding a point $x^* \in C$ such that:

$$\langle Ax^*, x - x^* \rangle \geq 0 \ \ \forall x \in C. \tag{1}$$

$VI(C, A)$ is used to denote the solution set of VIP(1) for simplicity. A wide range of mathematical and applied sciences rely heavily on variational inequalities in both theory and algorithms. Due to the importance of the variational inequality problem and many of its applications in different fields, several notable researchers have extensively studied this class of problems in the literature, and many more new ideas are emerging in connection with the problems. In the case of finite-dimensional setting, the current state-of-the-art results can be found in [5–7] including the substantial references therein.

Many algorithms (iterative methods) for solving the variational inequality (1) have been developed and well studied; see [5–15] and the references therein. One of the famous methods is the so-called extragradient method (EGM), which was developed by Korpelevich [16] (also by Antipin [17] independently) in the finite-dimensional Euclidean space for a monotone and Lipschitz continuous operator. The extragradient method has been modified in different ways and later was extended to infinite-dimensional spaces. Many of these extensions were well studied in [18–23] and the references therein.

One feature that renders the Korpelevich algorithm less acceptable has to do with the fact that two projections onto the feasible set are required in every iteration. For this reason, there is a need to solve a minimum distance problem twice every iteration. Therefore, the efficiency of this method (Korpelevich algorithm) is affected, which limits its application as well.

A remedy to the second drawback was presented in Censor et al. [18–20]. The authors introduced the subgradient-extragradient method (SEGM). Given $x_1 \in H$,

$$\begin{cases} y_n = P_C(x_n - \lambda A(x_n)), \\ T_n := \{w \in H : \langle x_n - \lambda A(x_n) - y_n, w - y_n \rangle \leq 0\}, \\ x_{n+1} = P_{T_n}(x_n - \lambda A(y_n)) \end{cases} \tag{2}$$

In this algorithm, A is an L-Lipschitz-continuous and monotone mapping and $0 < \lambda < \frac{1}{L}$. One of the novelties in the proposed SEGM (2) is the replacement of the second projection onto a feasible set with a projection onto a half-space. Recently, weak and strong convergence results of SEGM (2) have been obtained in the literature; see [24,25] and the references therein.

Thong and Hieu in [26] came up with inertial subgradient-extragradient method in the following algorithm. Given $x_0, x_1 \in H$,

$$\begin{cases} w_n = x_n + \theta_n(x_n - x_{n-1}), \\ y_n = P_C(w_n - \lambda A(w_n)), \\ T_n := \{w \in H : \langle w_n - \lambda A(w_n) - y_n, w - y_n \rangle \leq 0\}, \\ x_{n+1} = P_{T_n}(w_n - \lambda A(y_n)). \end{cases} \tag{3}$$

The authors proved the weak convergence of the sequence $\{x_n\}$ generated by (3) to a solution of variational inequality (VI), Equation (1), for the case where A is monotone and an L-Lipschitz-continuous mapping. For some:

$$0 < \delta < \frac{1}{2} - 2\theta - \frac{1}{2}\theta^2,$$

the parameter λ is chosen to satisfy:

$$0 < \lambda L \leq \frac{\frac{1}{2} - 2\theta - \frac{1}{2}\theta^2 - \delta}{\frac{1}{2} - \theta + \frac{1}{2}\theta^2}.$$

The sequence $\{\theta_n\}$ is non-decreasing with $0 \leq \theta_n \leq \theta < \sqrt{5} - 2$.

Malitsky in [21] introduced the following projected reflected gradient method, which solves VI (1) when A is Lipschitz continuous and monotone: choose $x_1, x_0 \in C$:

$$\begin{cases} w_n = 2x_n - x_{n-1}, \\ x_{n+1} = P_C(x_n - \lambda A w_n), \end{cases} \tag{4}$$

where $\lambda \in (0, \frac{\sqrt{2}-1}{L})$, and we obtain weak convergence results in real Hilbert spaces.

Recently, Boţ et al. [27] introduced Tseng's forward-backward-forward algorithm with relaxation parameters in Algorithm 1 to solve VI (1).

In [28], the following adaptive golden ratio method in Algorithm 2 for solving VI (1) was proposed.

Algorithm 1: Tseng's forward-backward-forward algorithm with relaxation parameters.

Initialization: Choose $\rho_n \in (0,1]$ with the given parameters $\lambda_0 > 0$ and $0 < \mu < 1$. Let $x_1 \in H$ be arbitrary.
Iterative steps: x_{n+1} is calculated, with the current iterate x_n given as follows:
Step 1. Compute:

$$y_n = P_C(x_n - \lambda_n Ax_n).$$

If $x_n = y_n$ or $Ay_n = 0$, then stop, and y_n is a solution of $VI(C, A)$. Otherwise:
Step 2. Compute:

$$x_{n+1} = (1-\rho_n)x_n + \rho_n(y_n + \lambda_n(Ax_n - Ay_n)),$$

Update:

$$\lambda_{n+1} = \begin{cases} \min\{\frac{\mu\|x_n - y_n\|}{\|Ax_n - Ay_n\|}, \lambda_n\} & if\, Ax_n - Ay_n \neq 0, \\ \lambda_n & otherwise. \end{cases} \tag{5}$$

Set $n := n+1$, and go to Step 1.

Algorithm 2: Adaptive golden ratio method .

Initialization: Choose $x_0, x_1 \in H, \lambda_0 > 0, \phi \in (0, \frac{\sqrt{5}+1}{2}], \bar{\lambda} > 0$. Set $\bar{x}_0 = x_1, \theta_0 = 1, \rho = \frac{1}{\phi} + \frac{1}{\phi^2}$.
Iterative steps: x_{n+1} is calculated, with the current iterate x_n given as follows:
Step 1. Compute:

$$\lambda_n = \min\left\{\rho\lambda_{n-1}, \frac{\phi\theta_{n-1}}{4\lambda_{n-1}} \frac{\|x_n - x_{n-1}\|^2}{\|Ax_n - Ax_{n-1}\|^2}, \bar{\lambda}\right\}.$$

Step 2. Compute:

$$\bar{x}_n = \frac{(\phi - 1)x_n + \bar{x}_{n-1}}{\phi},$$

and:

$$x_{n+1} = P_C(\bar{x}_n - \lambda_n Ax_n). \tag{6}$$

Update:

$$\theta_n = \frac{\lambda_n}{\lambda_{n-1}}\phi \tag{7}$$

Set $n := n+1$, and go to Step 1.

Motivated by the recent works in [18–21,26–28], our aim in this paper is to introduce a reflected subgradient-extragradient method that solves variational inequalities and obtain weak convergence in the case where the cost function is Lipschitz continuous and a pseudo-monotone operator in real Hilbert spaces. This pseudo-monotone operator is in the sense of Karamardian [29]. Our method uses self-adaptive step sizes, and the convergence of the proposed algorithm is proven without any assumption of prior knowledge of the Lipschitz constant of the cost function.

The outline of the paper is as follows. We start with recalling some basic definitions and results in Section 2. Our algorithm and weak convergence analysis are presented in Section 3. In Section 4, we give some numerical experiments to demonstrate the performances of our method compared with other related algorithms.

2. Preliminaries

In this section, we provide necessary definitions and results needed in the sequel.

Definition 1. *An operator $T : H \to H$ is said to be L-*`Lipschitz continuous` *with $L > 0$ if the following inequality is satisfied:*

$$\|Tx - Ty\| \leq L\|x - y\| \quad \forall x, y \in H.$$

Definition 2. *An operator $T : H \to H$ is said to be* `monotone` *if the following inequality is satisfied:*

$$\langle Tx - Ty, x - y \rangle \geq 0 \quad \forall x, y \in H.$$

Definition 3. *An operator $T : H \to H$ is said to be pseudo-monotone if the following inequality implies the other:*

$$\langle Tx, y - x \rangle \geq 0 \Longrightarrow \langle Ty, y - x \rangle \geq 0 \quad \forall x, y \in H.$$

Definition 4. *An operator $T : H \to H$ is said to be* `sequentially weakly continuous` *if for each sequence $\{x_n\}$, we have that x_n converges weakly to x, which implies that $\{Tx_n\}$ converges weakly to Tx.*

Recall that for any given point x chosen in H, $P_C(x)$ denotes the unique nearest point in C. This operator has been shown to be nonexpansive, that is,

$$\|x - P_C(x)\| \leq \|x - y\| \; \forall y \in C.$$

The operator P_C is known as the metric projection of H onto C.

Lemma 1 ([30])**.** *Given $x \in H$ and $z \in C$ with C a nonempty, closed, and convex subset of a real Hilbert space H, then:*

$$z = P_C(x) \Longleftrightarrow \langle x - z, z - y \rangle \geq 0 \; \forall y \in C.$$

Lemma 2 ([30,31])**.** *Given $x \in H$, a real Hilbert space and letting C be a closed and convex subset of H, then the following inequalities are true:*

1. $\|P_C x - P_C(y)\|^2 \leq \langle P_C x - P_C(y), x - y \rangle \; \forall y \in H$
2. $\|P_C(x) - y\|^2 \leq \|x - y\|^2 - \|x - P_C(x)\|^2 \; \forall y \in C.$

Lemma 3 ([32])**.** *Given $x \in H$ and $v \in H$, $v \neq 0$, and letting $T = \{z \in H : \langle v, z - x \rangle \leq 0\}$, then, for all $u \in H$, the projection $P_T(u)$ is defined by:*

$$P_T(u) = u - \max\left\{0, \frac{\langle v, u - x \rangle}{||v||^2}\right\} v.$$

In particular, if $u \notin T$, then:

$$P_T(u) = u - \frac{\langle v, u - x \rangle}{||v||^2} v.$$

The explicit formula provided in Lemma 3 is very important in computing the projection of any point onto a half-space.

Lemma 4 ([33], Lemma 2.1)**.** *Let $A : C \to H$ be continuous and pseudo-monotone where C is a nonempty, closed, and convex subset of a real Hilbert space H. Then, x^* is a solution of $VI(C, A)$ if and only if:*

$$\langle Ax, x - x^* \rangle \geq 0 \; \forall x \in C.$$

Lemma 5 ([34])**.** *Let $\{x_n\}$ be a sequence in H and C a nonempty subset of H with the following conditions satisfied:*

(i) *every sequential weak cluster point of $\{x_n\}$ is in C;*

(ii) $\lim_{n \to \infty} \|x_n - x\|$ *exists for every $x \in C$.*

Then, the sequence $\{x_n\}$ converges weakly to a point in C.

The following lemmas were given in [35].

Lemma 6. *Let h be a real-valued function on a real Hilbert space H, and define $K := \{x \in H : h(x) \leq 0\}$. If h is Lipschitz continuous on H with modulus $\theta > 0$ and K is nonempty, then:*

$$dist(x, K) \geq \frac{1}{\theta} \max\{0, h(x)\} \ \ \forall x \in H,$$

where $dist(x, K)$ is the distance function from x to K.

Lemma 7. *Let H be a real Hilbert space. The following statements are satisfied.*

(i) For all $x, y \in H$, $\|x + y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2$;
(ii) For all $x, y \in H$, $\|x + y\|^2 \leq \|x\|^2 + 2\langle y, x + y\rangle$;
(iii) For all $x, y \in H$, $\|x + y\|^2 = 2\|x\|^2 + 2\|y\|^2 - \|x - y\|^2$.

Lemma 8 (Maingé [36])**.** *Let $\{\delta_n\}$, $\{\varphi_n\}$, and $\{\theta_n\}$ be sequences defined in $[0, +\infty)$ satisfying the following:*

$$\varphi_{n+1} \leq \varphi_n + \theta_n(\varphi_n - \varphi_{n-1}) + \delta_n, \ \ \forall n \geq 1, \ \sum_{n=1}^{+\infty} \delta_n < +\infty,$$

and there exists θ, a real number, with $0 \leq \theta_n \leq \theta < 1$ for all $n \in \mathbb{N}$. Then, the following hold:

(i) $\sum_{n=1}^{+\infty} [\varphi_n - \varphi_{n-1}]_+ < +\infty$, *where* $[t]_+ := \max\{t, 0\}$;
(ii) *there exists* $\varphi^* \in [0, +\infty)$ *such that* $\lim_{n\to\infty} \varphi_n = \varphi^*$.

In the work that follows, $x_n \to x$ as $n \to \infty$ denotes the strong convergence of $\{x_n\}_{n=1}^{\infty}$ to a point x, and $x_n \rightharpoonup x$ as $n \to \infty$ denotes the weak convergence of $\{x_n\}_{n=1}^{\infty}$ to a point x.

3. Main Results

We first provide the following conditions upon which the convergence analysis of our method is based and then present our method in Algorithm 3.

Condition 1. *The feasible set C is a nonempty, closed, and convex subset of H.*

Condition 2. *The VI (1) associated operator $A : H \to H$ is pseudo-monotone, sequentially weakly and Lipschitz continuous on a real Hilbert space H.*

Condition 3. *The solution set of VI (1) is nonempty, that is $VI(C, A) \neq \emptyset$.*

In addition, we also make the following parameter choices. $0 < \alpha \leq \alpha_n \leq \alpha_{n+1} < \frac{1}{2+\delta} := \epsilon, \ \ \delta > 0.$

Remark 1. *We point out here that the proposed Algorithm 3 is different from the method (4) in that the projection step in Algorithm 3 is $P_C(w_n - \lambda_n A w_n)$, while the projection step in (4) is $P_C(x_n - \lambda A w_n)$. Furthermore, A is assumed to be pseudo-monotone in our Algorithm 3, while A is assumed to be monotone in (4).*

Algorithm 3: Adaptive projected reflected subgradient extragradient method.

Initialization: Given $\lambda_0 > 0, \mu \in (0,1)$, let $x_0, x_1 \in H$ be arbitrary
Iterative steps: Given the current iterate x_n, calculate x_{n+1} as follows:
Step 1. Compute:

$$\begin{cases} w_n = 2x_n - x_{n-1} \\ y_n = P_C(w_n - \lambda_n A w_n). \end{cases}$$

If $x_n = w_n = y_n = x_{n+1}$, then stop. Otherwise:
Step 2. Compute:

$$x_{n+1} = (1-\alpha_n)x_n + \alpha_n P_{T_n}(w_n),$$

where:

$$T_n := \{x \in H : h_n(x) \le 0\}$$

and

$$h_n(x) = \langle w_n - y_n - \lambda_n(Aw_n - Ay_n), x - y_n \rangle. \tag{8}$$

Update:

$$\lambda_{n+1} = \begin{cases} \min\left\{ \dfrac{\mu\|w_n - y_n\|}{\|Aw_n - Ay_n\|}, \lambda_n \right\} & if\, Aw_n - Ay_n \neq 0, \\ \lambda_n & otherwise. \end{cases} \tag{9}$$

Set $n := n+1$, and go to Step 1.

The first step towards the convergence proof of Algorithm 3 is to show that the sequence $\{\lambda_n\}$ generated by (9) is well defined. This is done using similar arguments as in [25].

Lemma 9. *The sequence $\{\lambda_n\}$ generated by (9) is a nonincreasing sequence and:*

$$\lim_{n\to\infty} \lambda_n = \lambda \ge \min\Big\{\lambda_0, \frac{\mu}{L}\Big\}.$$

Proof. Clearly, by (9), $\{\lambda_n\}$ is nonincreasing since $\lambda_{n+1} \le \lambda_n$ for all $n \in \mathbb{N}$. Next, using the fact that A is L-Lipschitz continuous, we have:

$$\|Aw_n - Ay_n\| \le L\|w_n - y_n\|.$$

Therefore, we obtain:

$$\mu \frac{\|w_n - y_n\|}{\|Aw_n - Ay_n\|} \ge \frac{\mu}{L} \quad \text{if} \quad Aw_n \neq Ay_n,$$

which together with (9) implies that:

$$\lambda_n \ge \min\Big\{\lambda_0, \frac{\mu}{L}\Big\}.$$

Therefore, the sequence $\{\lambda_n\}$ is nonincreasing and lower bounded. Therefore, there exists $\lim_{n\to\infty} \lambda_n$. $\square$

Lemma 10. *Assume that Conditions 1–3 hold. Let x^* be a solution of Problem (1) and the function h_n be defined by (8). Then, $h_n(x^*) \le 0$, and there exists $n_0 \in \mathbb{N}$ such that:*

$$h_n(w_n) \ge \frac{1-\mu}{2}\|w_n - y_n\|^2 \quad \forall n \ge n_0.$$

In particular, if $w_n \neq y_n$ then $h_n(w_n) > 0$.

Proof. Using Lemma 4 and the fact that x^* denotes a solution of Problem (1), we obtain the following:

$$\langle Ay_n, x^* - y_n \rangle \leq 0. \tag{10}$$

It follows from (10) and $y_n = P_C(w_n - \lambda_n Aw_n)$ that:

$$\begin{aligned} h_n(x^*) &= \langle w_n - y_n - \lambda_n(Aw_n - Ay_n), x^* - y_n \rangle \\ &= \langle w_n - y_n - \lambda_n Aw_n, x^* - y_n \rangle + \lambda_n \langle Ay_n, x^* - y_n \rangle \\ &\leq 0. \end{aligned}$$

Hence, the proof of the first claim of Lemma 10 is achieved. Next, we proceed to the proof of the second claim. Clearly, from the definition of $\{\lambda_n\}$, the following inequality is true:

$$\|Aw_n - Ay_n\| \leq \frac{\mu}{\lambda_{n+1}} \|w_n - y_n\| \ \ \forall n. \tag{11}$$

In fact, Inequality (11) is satisfied if $Aw_n = Ay_n$. Otherwise, it implies from (9) that:

$$\lambda_{n+1} = \min\{\frac{\mu\|w_n - y_n\|}{\|Aw_n - Ay_n\|}, \lambda_n\} \leq \frac{\mu\|w_n - y_n\|}{\|Aw_n - Ay_n\|}.$$

Thus,

$$\|Aw_n - Ay_n\| \leq \frac{\mu}{\lambda_{n+1}} \|w_n - y_n\|.$$

Hence, we can conclude from the above that Inequality (11) is true for $Aw_n = Ay_n$ and $Aw_n \neq Ay_n$.

Using (11), we obtain:

$$\begin{aligned} h_n(w_n) &= \langle w_n - y_n - \lambda_n(Aw_n - Ay_n), w_n - y_n \rangle \\ &= \|w_n - y_n\|^2 - \lambda_n \langle Aw_n - Ay_n, w_n - y_n \rangle \\ &\geq \|w_n - y_n\|^2 - \lambda_n \|Aw_n - Ay_n\| \|w_n - y_n\| \\ &\geq \|w_n - y_n\|^2 - \mu \frac{\lambda_n}{\lambda_{n+1}} \|w_n - y_n\|^2 \\ &= (1 - \mu \frac{\lambda_n}{\lambda_{n+1}}) \|w_n - y_n\|^2. \end{aligned}$$

Since $\lim_{n\to\infty}(1 - \mu \frac{\lambda_n}{\lambda_{n+1}}) = 1 - \mu > \frac{1-\mu}{2} > 0$, there exists $n_0 \in \mathbb{N}$ such that $(1 - \mu \frac{\lambda_n}{\lambda_{n+1}}) > \frac{1-\mu}{2}$ for all $n \geq n_0$. Therefore,

$$h_n(w_n) \geq \frac{1-\mu}{2} \|w_n - y_n\|^2.$$

□

Remark 2. *Lemma 10 implies that $w_n \notin T_n$ with $n \geq n_0$. Based on Lemma 3, we can write x_{n+1} in the form:*

$$x_{n+1} = w_n - \frac{\langle w_n - y_n - \lambda_n(Aw_n - Ay_n), w_n - y_n \rangle}{\|w_n - y_n - \lambda_n(Aw_n - Ay_n)\|^2} (w_n - y_n - \lambda_n(Aw_n - Ay_n)) \ \ \forall n \geq n_0.$$

We present the following result using similar arguments in [14], Theorem 3.1.

Lemma 11. *Let $\{w_n\}$ be a sequence generated by Algorithm 3, and assume that Conditions 1–3 hold. Let $\{w_n\}$ be a sequence generated by Algorithm 3. If there exists $\{w_{n_k}\}$, a subsequence of $\{w_n\}$*

such that $\{w_{n_k}\}$ *converges weakly to* $z \in H$ *and* $\lim_{k\to\infty} \|w_{n_k} - y_{n_k}\| = 0$, *then* $z \in VI(C, A)$.

Proof. From $w_{n_k} \rightharpoonup z, \lim_{k\to\infty} \|w_{n_k} - y_{n_k}\| = 0$ and $\{y_n\} \subset C$, we have $z \in C$. Furthermore, we have:

$$y_{n_k} = P_C(w_{n_k} - \lambda_{n_k} Aw_{n_k}).$$

Thus,

$$\langle w_{n_k} - \lambda_{n_k} Aw_{n_k} - y_{n_k}, x - y_{n_k}\rangle \leq 0 \text{ for all } x \in C.$$

Equivalently, we have:

$$\frac{1}{\lambda_{n_k}}\langle w_{n_k} - y_{n_k}, x - y_{n_k}\rangle \leq \langle Aw_{n_k}, x - y_{n_k}\rangle \text{ for all } x \in C.$$

From this, we obtain:

$$\frac{1}{\lambda_{n_k}}\langle w_{n_k} - y_{n_k}, x - y_{n_k}\rangle + \langle Aw_{n_k}, y_{n_k} - w_{n_k}\rangle \leq \langle Aw_{n_k}, x - w_{n_k}\rangle \ \forall\, x \in C. \tag{12}$$

We have that $\{w_{n_k}\}$ is a bounded sequence and A is Lipschitz continuous on H, and we get that $\{Aw_{n_k}\}$ is bounded and $\lambda_n \geq \min\{\lambda_0, \frac{\mu}{L}\}$. Taking $k \to \infty$ in (12), since $\|w_{n_k} - y_{n_k}\| \to 0$, we get:

$$\liminf_{k\to\infty}\langle Aw_{n_k}, x - w_{n_k}\rangle \geq 0. \tag{13}$$

On the other hand, we have:

$$\langle Ay_{n_k}, x - y_{n_k}\rangle = \langle Ay_{n_k} - Aw_{n_k}, x - w_{n_k}\rangle + \langle Aw_{n_k}, x - w_{n_k}\rangle + \langle Ay_{n_k}, w_{n_k} - y_{n_k}\rangle. \tag{14}$$

Since $\lim_{k\to\infty} \|w_{n_k} - y_{n_k}\| = 0$ and A is Lipschitz continuous on H, we get:

$$\lim_{k\to\infty} \|Aw_{n_k} - Ay_{n_k}\| = 0,$$

which, together with (13) and (14), implies that:

$$\liminf_{k\to\infty}\langle Ay_{n_k}, x - y_{n_k}\rangle \geq 0. \tag{15}$$

Next, we show that $z \in VI(C, A)$.

Next, a decreasing sequence, $\{\epsilon_k\}$, of positive numbers, which tends to zero, is chosen. We denote by N_k, for each k, the smallest positive integer satisfying the inequality:

$$\langle Ay_{n_j}, x - y_{n_j}\rangle + \epsilon_k \geq 0 \text{ for all } j \geq N_k. \tag{16}$$

It should be noted that the existence of N_k is guaranteed by (15). Clearly, the sequence $\{N_k\}$ is increasing from the fact that $\{\epsilon_k\}$ is decreasing. Furthermore, for each k, since $\{y_{N_k}\} \subset C$, we can suppose $Ay_{N_k} \neq 0$. We get:

$$\langle Ay_{N_k}, v_{N_k}\rangle = 1 \text{ for each } k,$$

where:

$$v_{N_k} = \frac{Ay_{N_k}}{\|Ay_{N_k}\|^2}.$$

We can infer from (16) that for each k:

$$\langle Ay_{N_k}, x + \epsilon_k v_{N_k} - y_{N_k}\rangle \geq 0.$$

Using the pseudo-monotonicity of the operator A on H, we obtain:

$$\langle A(x+\epsilon_k v_{N_k}), x+\epsilon_k v_{N_k} - y_{N_k}\rangle \geq 0.$$

Hence, we have:

$$\langle Ax, x-y_{N_k}\rangle \geq \langle Ax - A(x+\epsilon_k v_{N_k}), x+\epsilon_k v_{N_k} - y_{N_k}\rangle - \epsilon_k \langle Ax, v_{N_k}\rangle. \tag{17}$$

Next, we show that:

$$\lim_{k\to\infty} \epsilon_k v_{N_k} = 0.$$

Using the fact that $w_{n_k} \rightharpoonup z$ and $\lim_{k\to\infty} \|w_{n_k} - y_{n_k}\| = 0$, we get $y_{N_k} \rightharpoonup z$ as $k \to \infty$. Furthermore, for A sequentially weakly continuous on C, $\{Ay_{n_k}\}$ converges weakly to Az. We can suppose $Az \neq 0$, since otherwise, z is a solution. Using the fact that the norm mapping is sequentially weakly lower semicontinuous, we obtain:

$$0 < \|Az\| \leq \liminf_{k\to\infty} \|Ay_{n_k}\|.$$

Since $\{y_{N_k}\} \subset \{y_{n_k}\}$ and $\epsilon_k \to 0$ as $k \to \infty$, we get:

$$0 \leq \limsup_{k\to\infty} \|\epsilon_k v_{N_k}\| = \limsup_{k\to\infty}\left(\frac{\epsilon_k}{\|Ay_{n_k}\|}\right) \leq \frac{\limsup_{k\to\infty}\epsilon_k}{\liminf_{k\to\infty}\|Ay_{n_k}\|} = 0.$$

This in fact means $\lim_{k\to\infty} \epsilon_k v_{N_k} = 0$.

Next, letting $k \to \infty$, then the right-hand side of (17) tends to zero by A being Lipschitz continuous, $\{w_{N_k}\}, \{v_{N_k}\}$ are bounded, and:

$$\lim_{k\to\infty} \epsilon_k v_{N_k} = 0.$$

Hence, we obtain:

$$\liminf_{k\to\infty} \langle Ax, x-y_{N_k}\rangle \geq 0.$$

Therefore, for all $x \in C$, we get:

$$\langle Ax, x-z\rangle = \lim_{k\to\infty} \langle Ax, x-y_{N_k}\rangle = \liminf_{k\to\infty}\langle Ax, x-y_{N_k}\rangle \geq 0.$$

Finally, using Lemma 4, we have $z \in VI(C, A)$, which completes the proof. □

Remark 3. *Imposing the sequential weak continuity on A is not necessary when A is a monotone function; see [24].*

Theorem 4. *Any sequence $\{x_n\}$ that is generated using Algorithm 3 converges weakly to an element of $VI(C, A)$ when Conditions 1–3 are satisfied.*

Proof. Claim 1. $\{x_n\}$ is a bounded sequence. Define $u_n := P_{T_n}(w_n)$, and let $p \in VI(C, A)$. Then, we have:

$$\|u_n - p\|^2 = \|P_{T_n} w_n - p\|^2 \leq \|w_n - p\|^2 - \|u_n - w_n\|^2. \tag{18}$$

Furthermore,

$$\begin{aligned} \|u_n - p\|^2 = \|P_{T_n} w_n - p\|^2 &\leq \|w_n - p\|^2 - \|P_{T_n} w_n - w_n\|^2 \\ &= \|w_n - p\|^2 - \mathrm{dist}^2(w_n, T_n). \end{aligned} \tag{19}$$

This implies that:

$$\begin{aligned} \|x_{n+1}-p\|^2 &= \|(1-\alpha_n)(x_n-p)+\alpha_n(u_n-p)\|^2 \\ &= (1-\alpha_n)\|x_n-p\|^2+\alpha_n\|u_n-p\|^2 \\ &\quad -\alpha_n(1-\alpha_n)\|x_n-u_n\|^2, \end{aligned} \tag{20}$$

which in turn implies that:

$$\begin{aligned} \|x_{n+1}-p\|^2 &\leq (1-\alpha_n)\|x_n-p\|^2+\alpha_n\|w_n-p\|^2 \\ &\quad -\alpha_n(1-\alpha_n)\|x_n-u_n\|^2. \end{aligned} \tag{21}$$

Note that:

$$x_{n+1}=(1-\alpha_n)x_n+\alpha_n u_n$$

and this implies:

$$u_n-x_n=\frac{1}{\alpha_n}(x_{n+1}-x_n). \tag{22}$$

Using (22) in (21), we obtain:

$$\begin{aligned} \|x_{n+1}-p\|^2 &\leq (1-\alpha_n)\|x_n-p\|^2+\alpha_n\|w_n-p\|^2 \\ &\quad -\frac{(1-\alpha_n)}{\alpha_n}\|x_{n+1}-x_n\|^2. \end{aligned} \tag{23}$$

Furthermore, by Lemma 7 (iii),

$$\begin{aligned} \|w_n-p\|^2 &= \|2x_n-x_{n-1}-p\|^2 \\ &= \|(x_n-p)+(x_n-x_{n-1})\|^2 \\ &= 2\|x_n-p\|^2-\|x_{n-1}-p\|^2 \\ &\quad +2\|x_n-x_{n-1}\|^2. \end{aligned} \tag{24}$$

Using (24) in (23):

$$\begin{aligned} \|x_{n+1}-p\|^2 &\leq (1-\alpha_n)\|x_n-p\|^2+2\alpha_n\|x_n-p\|^2 \\ &\quad -\alpha_n\|x_{n-1}-p\|^2+2\alpha_n\|x_n-x_{n-1}\|^2 \\ &\quad -\frac{1-\alpha_n}{\alpha_n}\|x_{n+1}-x_n\|^2 \\ &= (1+\alpha_n)\|x_n-p\|^2-\alpha_n\|x_{n-1}-p\|^2 \\ &\quad +2\alpha_n\|x_n-x_{n-1}\|^2 \\ &\quad -\frac{1-\alpha_n}{\alpha_n}\|x_{n+1}-x_n\|^2. \end{aligned} \tag{25}$$

Define:

$$\Gamma_n := \|x_n-p\|^2-\alpha_n\|x_{n-1}-p\|^2+2\alpha_n\|x_n-x_{n-1}\|^2, \;\; n\geq 1.$$

Since $\alpha_n \leq \alpha_{n+1}$, we have:

$$\begin{aligned} \Gamma_{n+1}-\Gamma_n &= \|x_{n+1}-p\|^2-(1+\alpha_{n+1})\|x_n-p\|^2 \\ &\quad +\alpha_n\|x_{n-1}-p\|^2+2\alpha_{n+1}\|x_{n+1}-x_n\|^2 \\ &\quad -2\alpha_n\|x_n-x_{n-1}\|^2 \\ &\leq \|x_{n+1}-p\|^2-(1+\alpha_n)\|x_n-p\|^2+\alpha_n\|x_{n-1}-p\|^2 \\ &\quad +2\alpha_{n+1}\|x_{n+1}-x_n\|^2 \\ &\quad -2\alpha_n\|x_n-x_{n-1}\|^2. \end{aligned} \tag{26}$$

Now, using (25) in (26), one gets:

$$\begin{aligned} \Gamma_{n+1} - \Gamma_n &\leq -\frac{1-\alpha_n}{\alpha_n}\|x_{n+1} - x_n\|^2 + 2\alpha_{n+1}\|x_{n+1} - x_n\|^2 \\ &= -\Big(\frac{1-\alpha_n}{\alpha_n} - 2\alpha_{n+1}\Big)\|x_{n+1} - x_n\|^2. \end{aligned} \tag{27}$$

Observe that:

$$\begin{aligned} \frac{1-\alpha_n}{\alpha_n} - 2\alpha_{n+1} &= \frac{1}{\alpha_n} - 1 - 2\alpha_{n+1} \\ &\geq 2 + \delta - 1 - \frac{2}{2+\delta} \\ &= \delta + \frac{\delta}{2+\delta} \geq \delta. \end{aligned} \tag{28}$$

Using (28) in (27), we get:

$$\Gamma_{n+1} - \Gamma_n \leq -\delta\|x_{n+1} - x_n\|^2. \tag{29}$$

Hence, $\{\Gamma_n\}$ is non-increasing. In a similar way, we obtain:

$$\begin{aligned} \Gamma_n &= \|x_n - p\|^2 - \alpha_n\|x_{n-1} - p\|^2 + 2\alpha_n\|x_n - x_{n-1}\|^2 \\ &\geq \|x_n - p\|^2 - \alpha_n\|x_{n-1} - p\|^2. \end{aligned} \tag{30}$$

Note that:

$$\alpha_n < \frac{1}{2+\delta} = \epsilon < 1.$$

From (30), we have:

$$\begin{aligned} \|x_n - p\|^2 &\leq \alpha_n\|x_{n-1} - p\|^2 + \Gamma_n \\ &\leq \epsilon\|x_{n-1} - p\|^2 + \Gamma_1 \\ &\vdots \\ &\leq \epsilon^n\|x_0 - p\|^2 + (1 + \cdots + \epsilon^{n-1})\Gamma_1 \\ &\leq \epsilon^n\|x_0 - p\|^2 + \frac{\Gamma_1}{1-\epsilon}. \end{aligned} \tag{31}$$

Consequently,

$$\begin{aligned} \Gamma_{n+1} &= \|x_{n+1} - p\|^2 - \alpha_{n+1}\|x_n - p\|^2 \\ &\quad + 2\alpha_{n+1}\|x_{n+1} - x_n\|^2 \\ &\geq -\alpha_{n+1}\|x_n - p\|^2 \end{aligned}$$

and this means from (31) that:

$$\begin{aligned} -\Gamma_{n+1} &\leq \alpha_{n+1}\|x_n - p\|^2 \\ &\leq \epsilon\|x_n - p\|^2 \\ &\vdots \\ &\leq \epsilon^{n+1}\|x_0 - p\|^2 + \frac{\epsilon\Gamma_1}{1-\epsilon}. \end{aligned} \tag{32}$$

By (29) and (32), we get:

$$\begin{aligned}\delta\sum_{n=1}^{k}\|x_{n+1}-x_n\|^2 &\leq \Gamma_1-\Gamma_{k+1}\\ &\leq \epsilon^{k+1}\|x_0-p\|^2+\frac{\Gamma_1}{1-\epsilon}.\end{aligned} \tag{33}$$

This then implies that:

$$\sum_{n=1}^{\infty}\|x_{n+1}-x_n\|^2 \leq \frac{\Gamma_1}{\delta(1-\epsilon)} < +\infty. \tag{34}$$

Hence, $\lim_{n\to\infty}\|x_{n+1}-x_n\| = 0$. We also have from Algorithm 3 that:

$$\|w_n-x_n\| = \|x_n-x_{n-1}\| \to 0, n\to\infty. \tag{35}$$

From (25), we obtain:

$$\begin{aligned}\|x_{n+1}-p\|^2 \leq\ & (1+\alpha_n)\|x_n-p\|^2-\alpha_n\|x_{n-1}-p\|^2\\ & +2\|x_n-x_{n-1}\|^2.\end{aligned} \tag{36}$$

Using Lemma 8 in (36) (noting (34)), we get:

$$\lim_{n\to\infty}\|x_n-p\|^2 = l < \infty. \tag{37}$$

This implies that $\lim_{n\to\infty}\|x_n-p\|$ exists. Therefore, the sequence $\{x_n\}$ is bounded, and so is $\{y_n\}$.

Claim 2. There exists $M>1$ such that:

$$\left[\frac{1}{M}\frac{1-\mu}{2}\|w_n-y_n\|^2\right]^2 \leq \|w_n-p\|^2-\|u_n-p\|^2 \ \ \forall n\geq n_0.$$

We know that $\{Aw_n\}, \{Ay_n\}$ are bounded using the fact that $\{x_n\}, \{y_n\}, \{w_n\}$ are bounded. Hence, there exists $M>1$ such that:

$$\|w_n-y_n-\lambda_n(Aw_n-Ay_n)\| \leq M \ \text{ for all } n.$$

Therefore, for all $u,v\in H$, we obtain:

$$\begin{aligned}\|h_n(u)-h_n(v)\| &= \|\langle w_n-y_n-\lambda_n(Aw_n-Ay_n), u-v\rangle\|\\ &\leq \|w_n-y_n-\lambda_n(Aw_n-Ay_n)\|\|u-v\|\\ &\leq M\|u-v\|.\end{aligned}$$

Then, we have that $h_n(\cdot)$ is M-Lipschitz continuous on H. From Lemma 6, we get:

$$\text{dist}(w_n,T_n) \geq \frac{1}{M}h_n(w_n),$$

from which, together with Lemma 10, we get:

$$\text{dist}(w_n,T_n) \geq \frac{1}{M}\frac{1-\mu}{2}\|w_n-y_n\|^2 \ \ \forall n\geq n_0. \tag{38}$$

Combining (19) and (38), we obtain:

$$\|u_n-p\|^2 \leq \|w_n-p\|^2-\left[\frac{1}{M}\frac{1-\mu}{2}\|w_n-y_n\|^2\right]^2 \ \ \forall n\geq n_0. \tag{39}$$

This complete the proof of Claim 2.

Claim 3. The sequence $\{x_n\}$ converges weakly to an element of $VI(C, A)$. Indeed, since $\{x_n\}$ is a bounded sequence, there exists the subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k} \rightharpoonup z \in C$. Since $\{w_n\}$ and $\{u_n\}$ are bounded, there exists $M^* > 0$ such that $\forall n \geq n_0$, and we have from (39):

$$\begin{aligned} \left[\frac{1}{M}\frac{1-\mu}{2}\|w_n - y_n\|^2\right]^2 &\leq \|w_n - p\|^2 - \|u_n - p\|^2 \\ &= \Big[\|w_n - p\| - \|u_n - p\|\Big]\Big[\|w_n - p\| + \|u_n - p\|\Big] \\ &\leq \|w_n - u_n\|\Big[\|w_n - p\| + \|u_n - p\|\Big] \\ &\leq M^*\|w_n - u_n\|. \end{aligned} \tag{40}$$

From (40), we have:

$$\lim_{n\to\infty} \|w_n - y_n\| = 0. \tag{41}$$

Consequently,

$$\|x_n - y_n\| \leq \|w_n - y_n\| + \|w_n - x_n\| \to 0, n \to \infty.$$

Furthermore, $\{w_{n_k}\}$ of $\{w_n\}$ converges weakly to $z \in C$. This implies from Lemma 11 and (41) that $z \in VI(C, A)$. Therefore, we proved that if $p \in VI(C, A)$, then $\lim_{n\to\infty} \|x_n - p\|$ exists, and each sequential weak cluster point of the sequence $\{x_n\}$ is in $VI(C, A)$. By Lemma 5, the sequence $\{x_n\}$ converges weakly to an element of $VI(C, A)$. □

4. Numerical Illustrations

In this section, we consider many examples in which some are real-life applications for numerical implementations of our proposed Algorithm 3. For a broader overview of the efficiency and accuracy of our proposed algorithm, we investigate and compare the performance of the proposed Algorithm 3 with Algorithm 1 proposed by Boţ et al. in [27] (Bot Alg.), Algorithm 2 proposed by Malitsky in [28] (Malitsky Alg.), the algorithm proposed by Shehu and Iyiola in [37] (Shehu Alg.), the subgradient-extragradient method (SEM) (2), and the extragradient method (EGM) in [16].

Example 1 (Tomography reconstruction model)**.** *In this example, we consider the linear inverse problem:*

$$Bx = \hat{b}, \tag{42}$$

where $x \in \mathbb{R}^k$ is the unknown image, $B \in \mathbb{R}^{m\times k}$ is the projection matrix, and $\hat{b} \in \mathbb{R}^m$ is the given sinogram (set of projections). The aim then is to recover a slice image of an object from a sinogram. To be realistic, we consider noisy $b = \hat{b} + \epsilon$, where $\epsilon \in \mathbb{R}^m$. Problem (42) can be presented as a convex feasibility problem (CFP) with the sets (hyper-planes) $C_i = \{x : \langle a_i, x\rangle = b_i\}$. Since, in practice, the projection matrix B is often rank-deficient, so $b \notin range(B)$; thus, we may assume that the CFP has no solution (also called inconsistent), so we consider the least squares model $\min_x \sum_{i=1}^m \mathrm{dist}(x, C_i)^2$.

Recall that the projection onto the hyper-plane C_i has a closed formula $P_{C_i}x = x - \frac{\langle a_i,x\rangle - b_i}{\|a_i\|^2}a_i$. Therefore, the evaluation of Tx reduces to a matrix-vector multiplication, and this can be realized very efficiently, where $T := \frac{1}{m}(P_{C_1} + \ldots + P_{C_m})$ and $A := I - T$. Note that our approach only exploits feasibility constraints, which is definitely not a state-of-the-art model for tomography reconstruction. More involved methods would solve this problem with the use of some regularization techniques, but we keep such a simple model for illustration purposes only.

As a particular problem, we wish to reconstruct the Shepp–Logan phantom image 128×128 *(thus, $x \in \mathbb{R}^k$ with $k = 2^8$) from the far less measurement $m = 2^7$. We generate the matrix $B \in \mathbb{R}^{m\times k}$ randomly and define $b = Bx + \epsilon$, where $\epsilon \in \mathbb{R}^m$ is a random vector, whose entries are drawn from* $\mathrm{N}(0, 1)$.

Using this example, we give some comparisons of our proposed Algorithm 3, Algorithm 1 *proposed by Boţ et al. in [27] (Bot Alg.), and Algorithm 2 proposed by Malitsky et al. in [28] (Malitsky Alg.) using the residual $\|e_n\|_2 := \|x_{n+1} - x_n\|_2 \leq \epsilon$, where $\epsilon = 10^{-4}$, as our stopping criterion. For our proposed algorithm, the starting point x_0 is chosen randomly with $x_1 = (0, \ldots, 0)$ and $\alpha_n = 0.002$, for Boţ et al., the starting point $x_0 = (0, \ldots, 0)$ and $\rho_n = 0.2$, while for Malitsky et al., the starting point $x_0 = \bar{x}_0 = (0, \ldots, 0)$ with x_1 randomly chosen, $\theta_0 = 1$, and $\bar{\lambda} = 1$. All results are reported in Tables 1 and 2 and Figures 1–6.*

Example 2 (Equilibrium-optimization model)**.** *Here, we consider the Nash–Cournot oligopolistic equilibrium model in electricity markets. Given m companies, such that the i-th company possesses I_i generating units, denote by x the power vector, that is each of its j-th entries x_j corresponds to the power generated by unit j. Assume that the price function p_i is an affine decreasing function of $s := \sum_{j=1}^{N} x_j$ where N is the number of all generating units. Therefore, $p_i(s) := \alpha - \beta_i s$. We can now present the profit of company i as $f_i(x) := p_i(s)\sum_{j \in I_i} x_j - \sum_{j \in I_i} c_j(x_j)$, where $c_j(x_j)$ is the cost for generating x_j by generating unit j. Denote by K_i the strategy set of the i-th company i. Clearly, $\sum_{j \in I_i} x_j \in K_i$ for each i, and the overall strategy set is $C := K_1 \times K_2 \times \ldots \times K_m$.*

The Nash equilibrium concept with regards to the above data is that each company wishes to maximize its profit by choosing the corresponding production level under the presumption that the production of the other companies is a parametric input.

Recall that $x^* \in C = K_1 \times K_2 \times \ldots \times K_m$ is an equilibrium point if:

$$f_i(x^*) \geq f_i(x^*[x_i]) \forall x_i \in K_i, \quad i = 1, 2, \ldots, m,$$

where the vector $x^*[x_i]$ stands for the vector obtained from x^* by replacing x_i^* with x_i. Define:

$$f(x, y) := \psi(x, y) - \psi(x, x)$$

with:

$$\psi(x, y) := -\sum_{i=1}^{n} f_i(x^*[y_i]).$$

Therefore, finding a Nash equilibrium point is formulated as:

$$X^* \in C : f(x^*, x) \geq 0 \ \ \forall x \in C. \tag{43}$$

Suppose for every j, the cost c_j for production and the environmental fee g are increasingly convex functions. This convexity assumption implies that (43) is equivalent to (see [38]):

$$x \in C : \langle Bx - a + \nabla\varphi(x), y - x \rangle \geq 0 \ \ \forall y \in C,$$

where:

$$a := (\alpha, \alpha, \ldots, \alpha)^T$$

$$B_1 = \begin{pmatrix} \beta_1 & 0 & 0 & \ldots & 0 \\ 0 & \beta_2 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \beta_m \end{pmatrix} B = \begin{pmatrix} 0 & \beta_1 & \beta_1 & \ldots & \beta_1 \\ \beta_2 & 0 & \beta_2 & \ldots & \beta_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \beta_m & \beta_m & \beta_m & \ldots & \beta_m \end{pmatrix}$$

$$\varphi(x) := x^T B_1 x + \sum_{j=1}^{N} c_j(x_j).$$

Note that c_j is differentiable convex for every j.

Our proposed scheme is tested with the following cost function:

$$c_j(x_j) = \frac{1}{2} x_j^T D x_j + d^T x_j.$$

The parameters β_j for all $j = 1, \ldots, m$, matrix D, and vector d were generated randomly in the interval $(0, 1]$, $[1, 40]$, *and* $[1, 40]$, *respectively.*

The numerical experiments involve the initial points x_0 and x_1 generated randomly in $[1, 40]$ *and $m = 10$. The stopping role of the algorithm is chosen as $\|e_n\|_2 := \|x_{n+1} - x_n\|_2 \leq \epsilon$, where $\epsilon = 10^{-4}$. Let us assume that each company has the same lower production bound one and upper production bound 40, that is,*

$$K_i := \{x_i : 1 \leq x_i \leq 40\}, \ \ i = 1, \ldots, 10.$$

We compare our proposed Algorithm 3 with Algorithm 1 proposed by Boţ et al. in [27] (Bot Alg.), Algorithm 2 proposed by Malitsky et al. in [28] (Malitsky Alg.), and Shehu and Iyiola's proposed Algorithm 3.2 in [37] (Shehu Alg.). For our proposed algorithm, we choose $\alpha_n = 0.49$, for Boţ et al., $\rho_n = 0.02$, for Malitsky et al., the starting point $x_0 = \bar{x}_0$, $\theta_0 = 1$, and $\bar{\lambda} = 1$, while for Shehu and Iyiola, $\rho = 1$ and $\sigma = 0.5$. All results are reported in Tables 3 and 4 and Figures 7–12.

Example 3. *This example is taken from [39]. First, generate the following matrices randomly B, S, D in $\mathbb{R}^{m\times m}$ where S is skew-symmetric and D is a positive definite diagonal matrix. Then, define the operator A by $A(x) := Mx + q$ with $M = BB^T + S + D$. The symmetric property of S implies that the operator does not arise from an optimization problem, and the positive definiteness of D implies the uniqueness of the solution to the corresponding variational inequality problem.*

We choose here $q = 0$. Choose random matrix $B \in \mathbb{R}^{k\times m}$ and $b \in \mathbb{R}^k$ with nonnegative entries, and define the VI feasible set C by $Bx \leq b$. Clearly, the origin is in C, and it is the unique solution of the corresponding variational inequality. Projections onto C are computed via the MATLAB routine `fmincon`*, and thus, it is costly. We test the algorithm's performances (number of iterations and CPU time in seconds) for different m's and inequality constraints k.*

For this example, the stopping criterion is chosen as $\|e_n\|_2 := \|x_n\|_2 \leq \epsilon$, where $\epsilon = 0.002$. We experiment with different values of k $(30, and 50)$ *and m* $(10, 20, 30, and 40)$*. We randomly generate vector b and matrices B, S, and D. We choose λ_0 and μ appropriately in Algorithm 3. In* (2)*, $L = \|M\|$ is used. Algorithm 3 proposed in this paper is compared with the subgradient-extragradient method (SEM)* (2)*. For our proposed algorithm, we choose $\mu = 0.999$, $\lambda_0 = 0.5$, and $\alpha_n = 0.499$ while for SEM, $\lambda = \frac{0.125}{4L}$. All results are reported in Table 5 and Figures 13–22.*

Example 4. *Consider VI* (1) *with:*

$$A(x) = \begin{pmatrix} 0.5x_1x_2 - 2x_2 - 10^7 \\ -4x_1 + 0.1x_2^2 - 10^7 \end{pmatrix}$$

and:

$$C := \{x \in \mathbb{R}^2 : (x_1 - 2)^2 + (x_2 - 2)^2 \leq 1\}.$$

Then, A is not monotone on C, but pseudo-monotone. Furthermore, VI (1) *has the unique solution $x^* = (2.707, 2.707)^T$. A comparison of our method is made with the extragradient method [16]. We denote the parameter in EGM [16] as λ_n^* to differentiate it from λ_n in our proposed Algorithm 3. We terminate the iterations if:*

$$\|e_n\|_2 := \|x_n - x^*\|_2 \leq \varepsilon$$

with $\varepsilon = 10^{-3}$. In this, our proposed Algorithm 3 is compared with the extragradient method (EGM) in [16]. For our proposed algorithm, we choose $x_0 = (1, 2)^T$ and $\alpha_n = 0.499$, while for EGM, $\lambda_n^* = 0.00000001$. All results are reported in Tables 6–8 and Figures 23–32.

Example 5. *Consider $H := L^2([0, 1])$ and $C := \{x \in H : \|x\| \leq 2\}$. Define $A : L^2([0, 1]) \to L^2([0, 1])$ by:*

$$A(u)(t) := e^{-\|u\|^2} \int_0^t u(s)ds, \ \ \forall u \in L^2([0, 1]), t \in [0, 1].$$

It can also be shown that A is pseudo-monotone, but not monotone on H, Lipschitz continuous with $L = \left(\frac{2}{e} + 1\right)\frac{2}{\pi}$, and sequentially weakly-to-weak;y continuous on H (see Example 2.1 of [27]).

Our proposed Algorithm 3 is compared with Algorithm 1 proposed by Boţ et al. in [27] (Bot Alg.), Algorithm 2 proposed by Malitsky et al. in [28] (Malitsky Alg.), and Shehu and Iyiola's proposed Algorithm 3.2 in [37] (Shehu Alg.). For our proposed algorithm, we choose $x_0 = \frac{1}{9}e^t \sin(t)$ *and* $\alpha_n = 0.49$*, for Boţ et al.,* $\rho_n = 0.02$*, for Malitsky et al.,* $\theta_0 = 1$*, and* $\phi = 1.1$*, while for Shehu and Iyiola,* $\rho = 1$*,* $\sigma = 0.005$*, and* $\gamma = 0.9$*. All algorithms are terminated using the stopping criterion* $\|e_n\|_2 := \|w_n - y_n\|_2 \leq \varepsilon$ *with* $\varepsilon = 10^{-4}$*. All results are reported in Tables 9 and 10 and Figures 33–36.*

Table 1. Example 1 comparison: proposed Algorithm 3, Bot Algorithm 1, and Malitsky Algorithm 2 with $\mu = \phi = 0.9$.

	Proposed Algorithm 3		**Bot Algorithm 1**		**Malitsky Algorithm 2**	
λ_0	**No. of Iter.**	**CPU Time**	**No. of Iter.**	**CPU Time**	**No. of Iter.**	**CPU Time**
0.1	2	5.8318×10^{-3}	98	0.2590	71	0.1533
1	2	4.3773×10^{-3}	61	0.1643	39	0.0826
5	12	4.0209×10^{-2}	38	0.1012	22	0.0491
10	6	1.8319×10^{-2}	207	0.5764	72	0.1359

Table 2. Example 1: proposed Algorithm 3 with $\lambda_0 = 1$ for different μ values.

	$\lambda_0 = 0.1$	$\lambda_0 = 0.3$	$\lambda_0 = 0.7$	$\lambda_0 = 0.9$
No. of Iter.	6	6	3	2
CPU Time	2.6506×10^{-2}	2.7113×10^{-2}	7.1002×10^{-3}	4.7982×10^{-3}

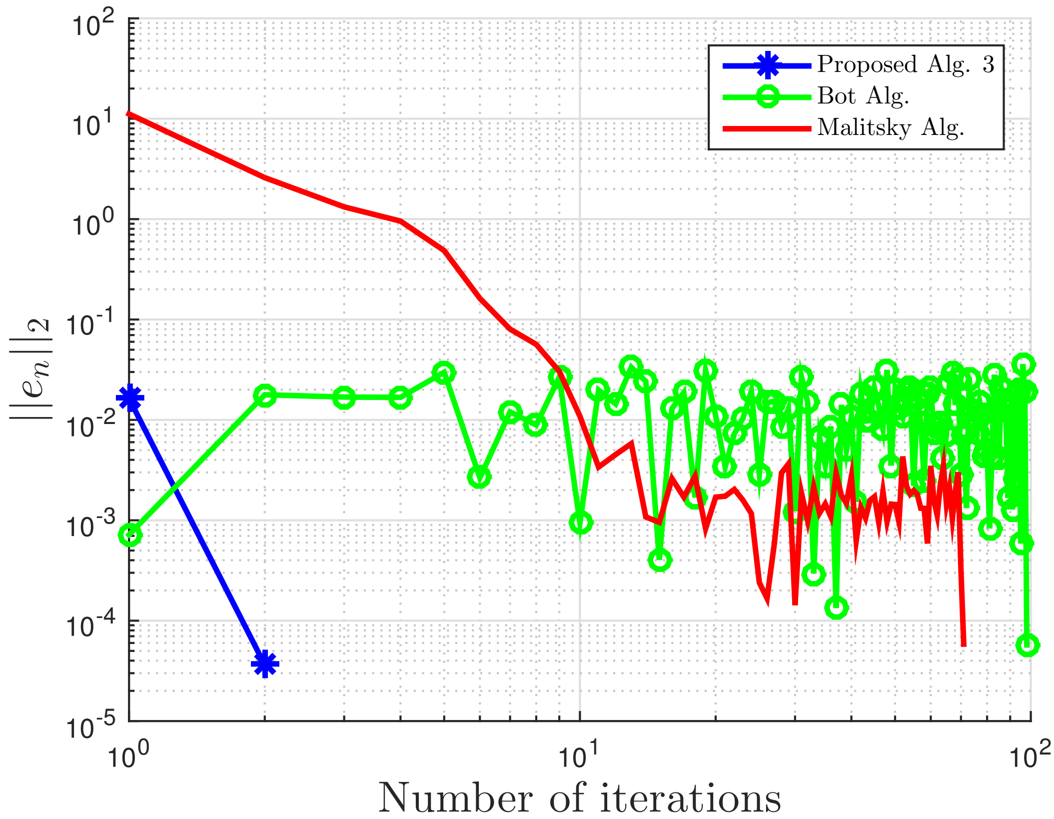

Figure 1. Example 1: $\mu = \phi = 0.9$ and $\lambda_0 = 0.1$. Alg., Algorithm.

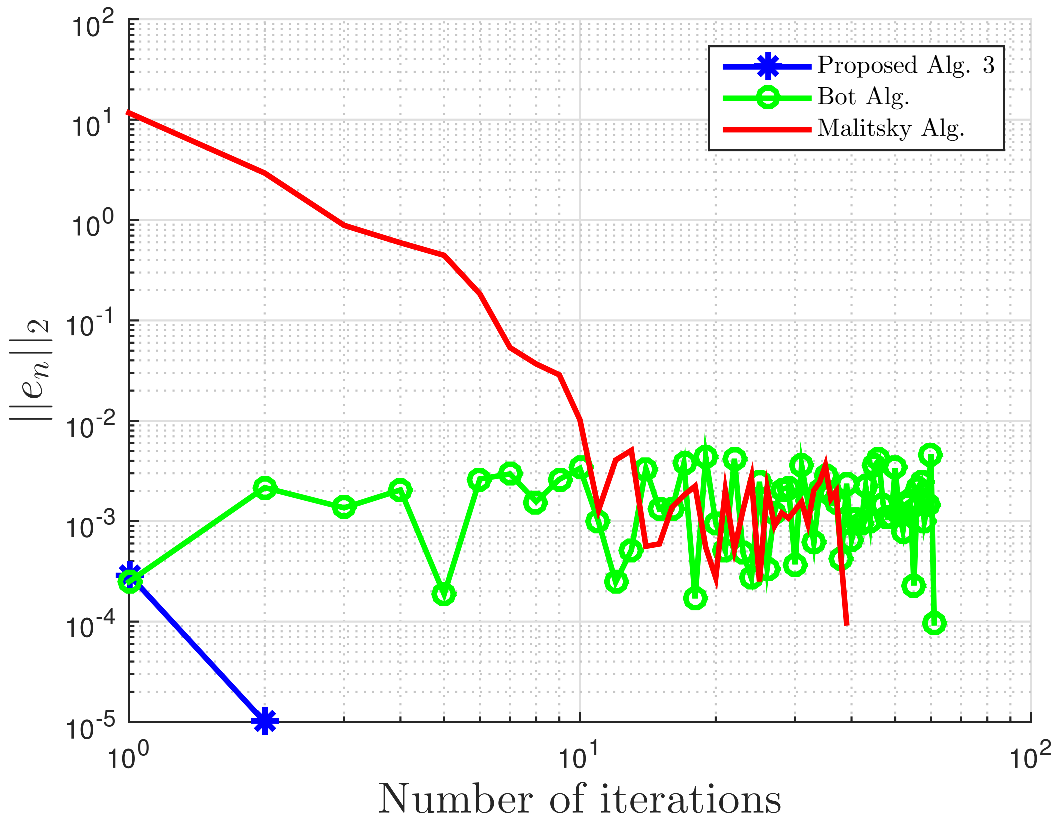

Figure 2. Example 1: $\mu = \phi = 0.9$ and $\lambda_0 = 1$.

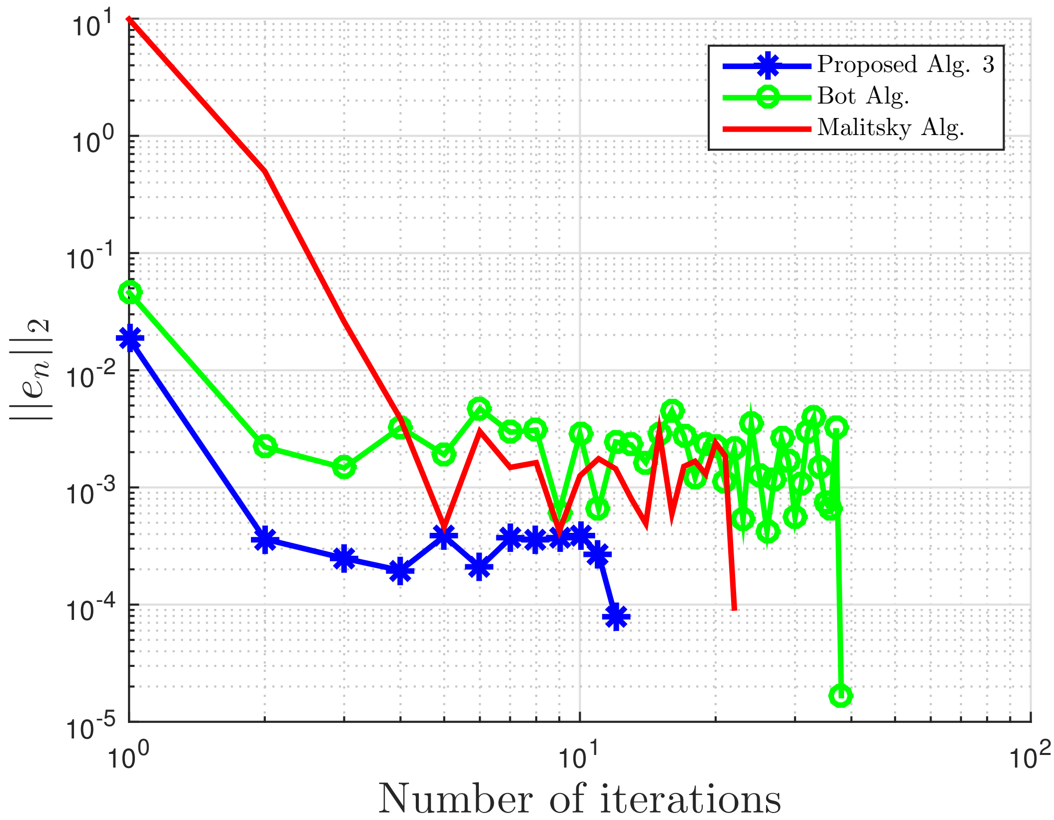

Figure 3. Example 1: $\mu = \phi = 0.9$ and $\lambda_0 = 5$.

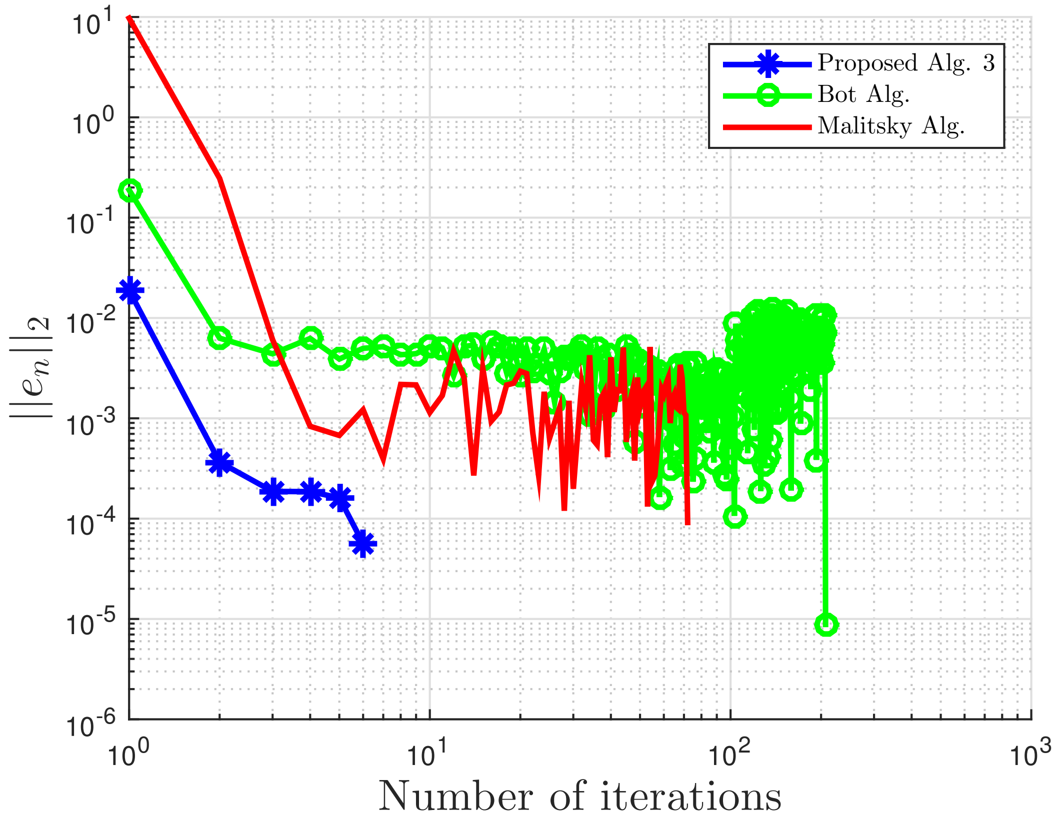

Figure 4. Example 1: $\mu = \phi = 0.9$ and $\lambda_0 = 10$.

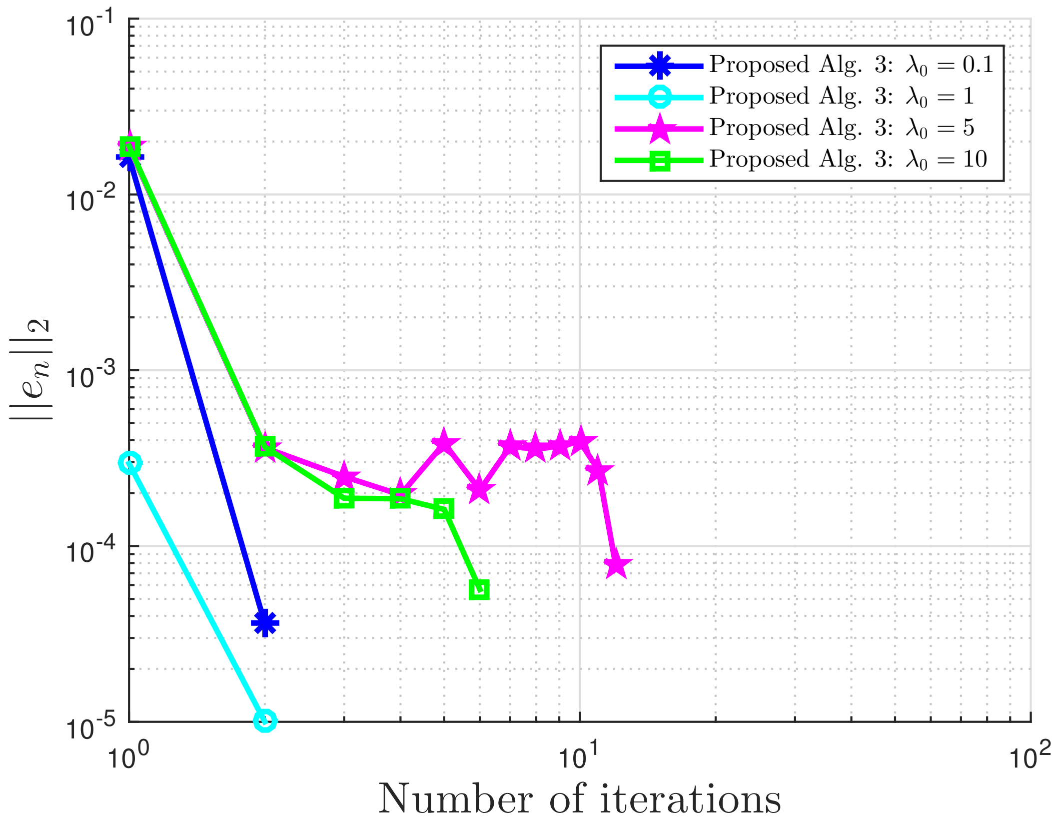

Figure 5. Example 1: $\mu = \phi = 0.9$.

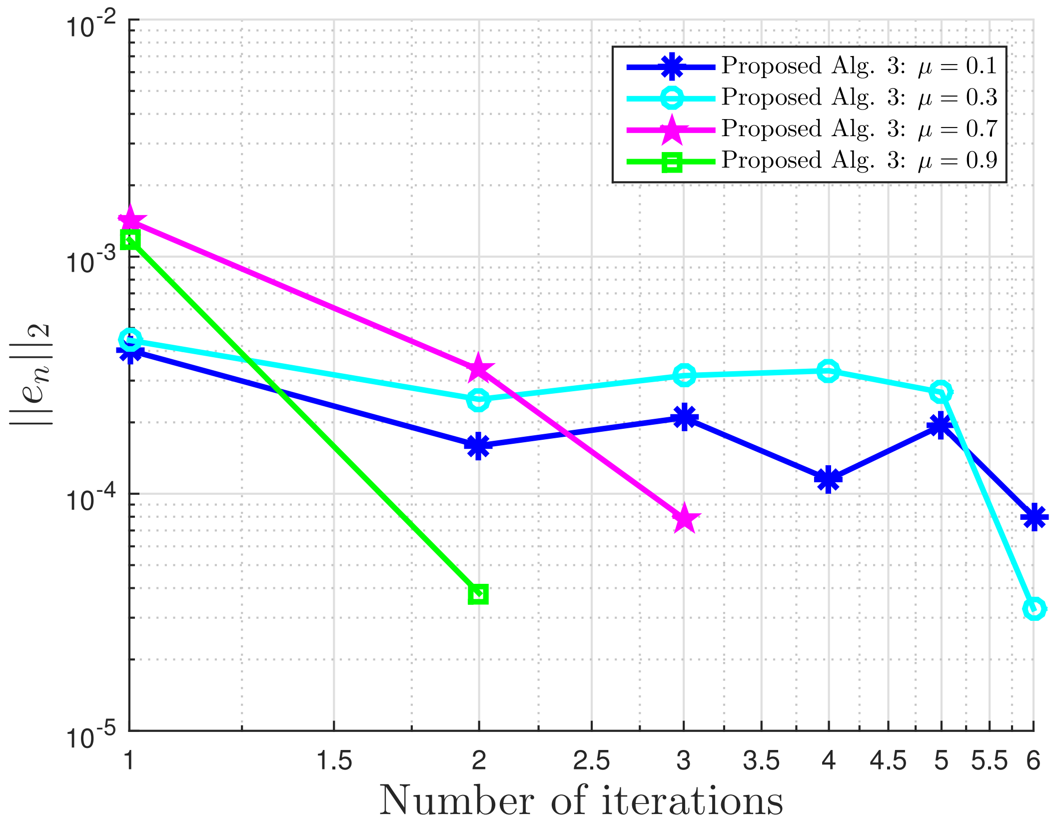

Figure 6. Example 1: $\lambda_0 = 1$.

Table 3. Example 2 comparison: proposed Algorithm 3, Bot Algorithm 1, Malitsky Algorithm 2, and Shehu Alg. [37] with $\lambda_0 = 1$ and $\mu = \phi = \gamma$.

	Proposed Algorithm 3		**Bot Algorithm 1**		**Malitsky Algorithm 2**		**Shehu Alg. [37]**	
μ	**No. of Iter.**	**CPU Time**	**No. of Iter.**	**CPU Time**	**No. of Iter.**	**CPU Time**	**No. of Iter.**	**CPU Time**
0.1	22	8.7749×10^{-3}	582	0.1372	33	8.4833×10^{-3}	15680	8.7073
0.3	24	8.8605×10^{-3}	594	0.1440	47	9.3949×10^{-3}	13047	7.2304
0.7	36	1.8575×10^{-2}	619	0.1914	81	1.9522×10^{-2}	14736	9.8807
0.999	47	3.9262×10^{-2}	581	0.3016	1809	0.7438	7048	5.4446

Table 4. Example 2: proposed Algorithm 3 with $\mu = 0.999$ for different λ_0 values.

	$\boldsymbol{\lambda_0 = 0.1}$	$\boldsymbol{\lambda_0 = 1}$	$\boldsymbol{\lambda_0 = 5}$	$\boldsymbol{\lambda_0 = 10}$
No. of Iter.	45	52	44	50
CPU Time	2.6624×10^{-2}	2.1446×10^{-2}	1.7123×10^{-2}	1.9696×10^{-2}

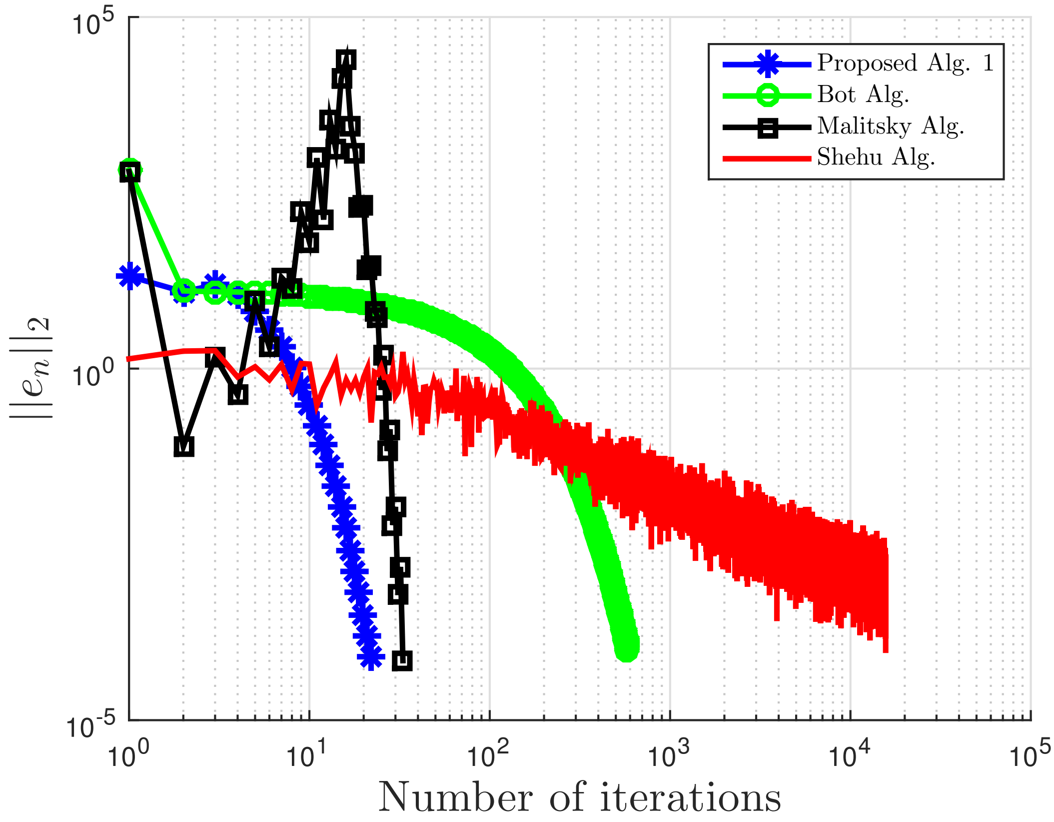

Figure 7. Example 2: $\mu = \phi = \gamma = 0.1$ and $\lambda_0 = 1$.

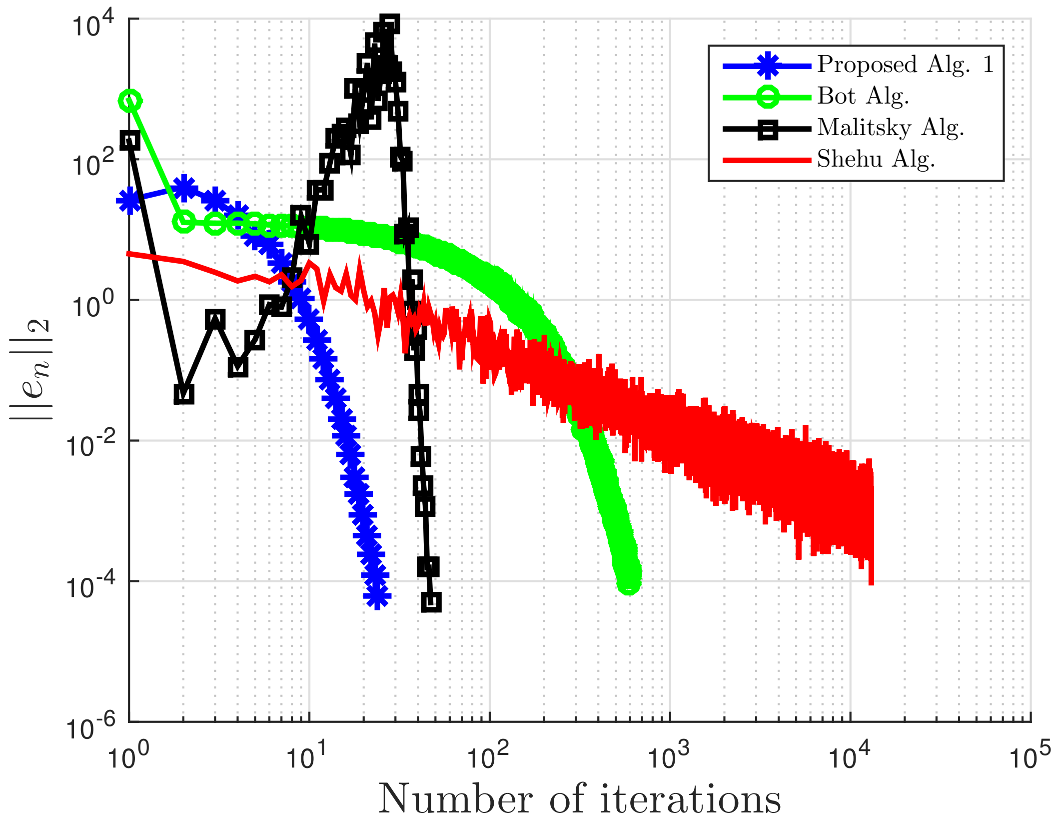

Figure 8. Example 2: $\mu = \phi = \gamma = 0.3$ and $\lambda_0 = 1$.

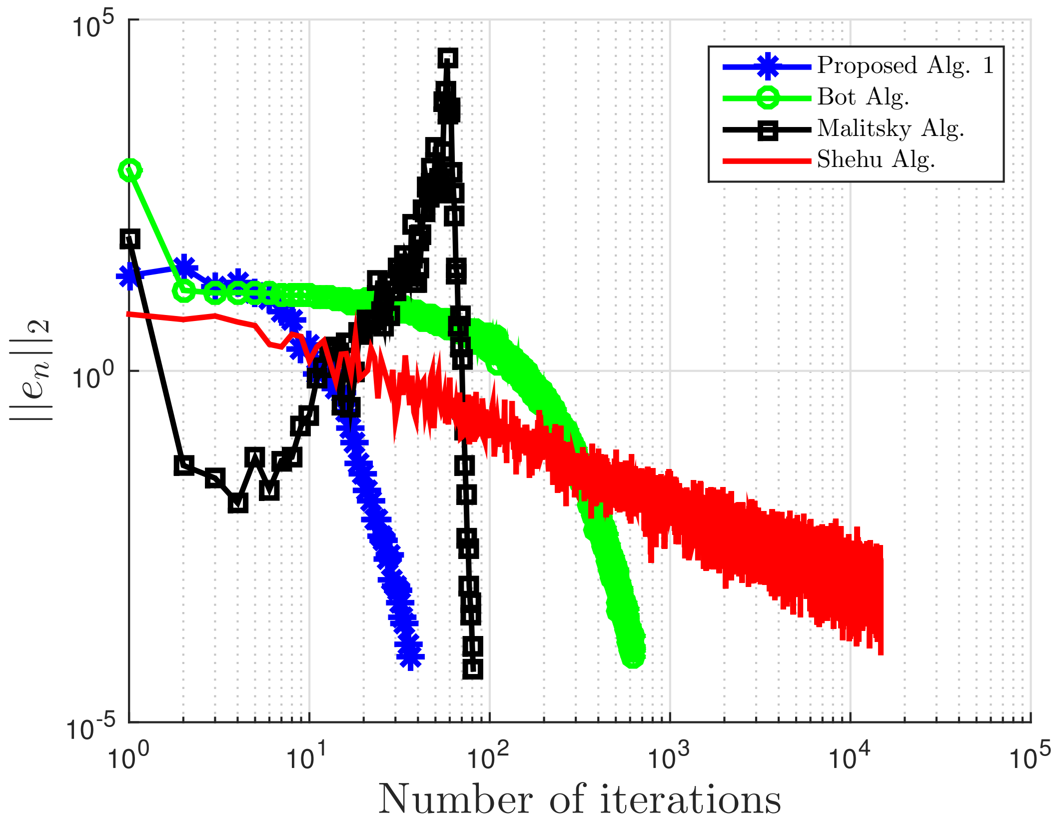

Figure 9. Example 2: $\mu = \phi = \gamma = 0.7$ and $\lambda_0 = 1$.

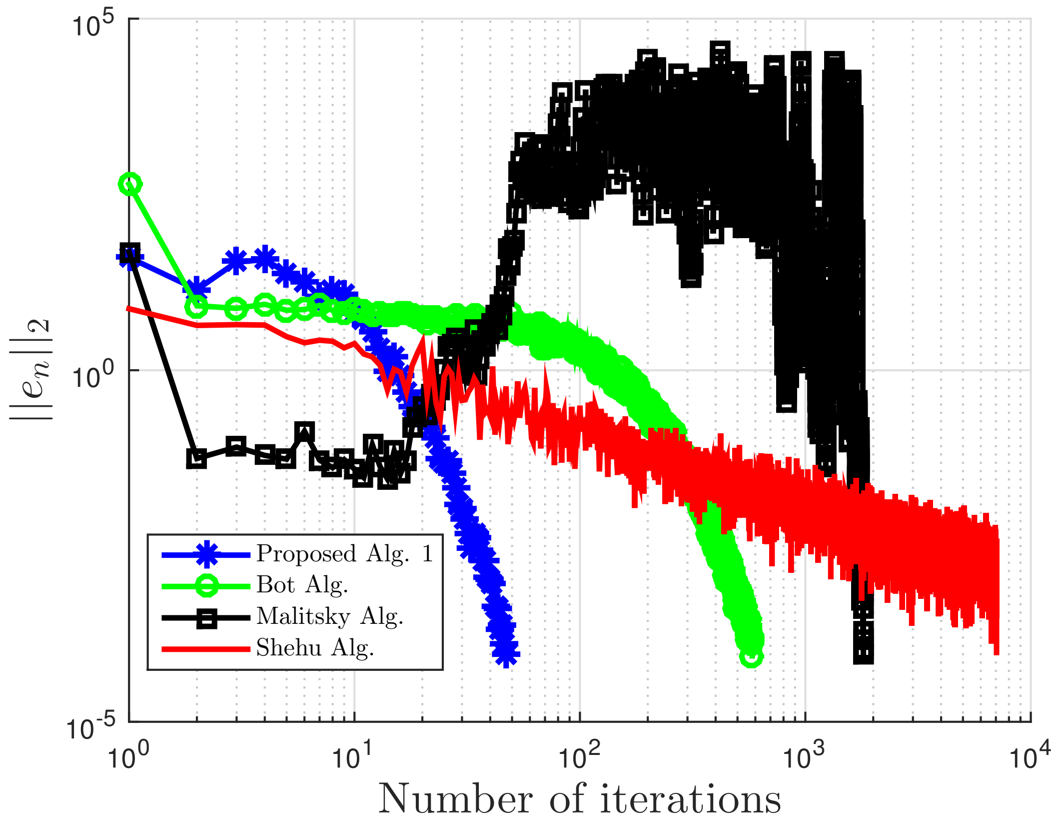

Figure 10. Example 2: $\mu = \phi = \gamma = 0.999$ and $\lambda_0 = 1$.

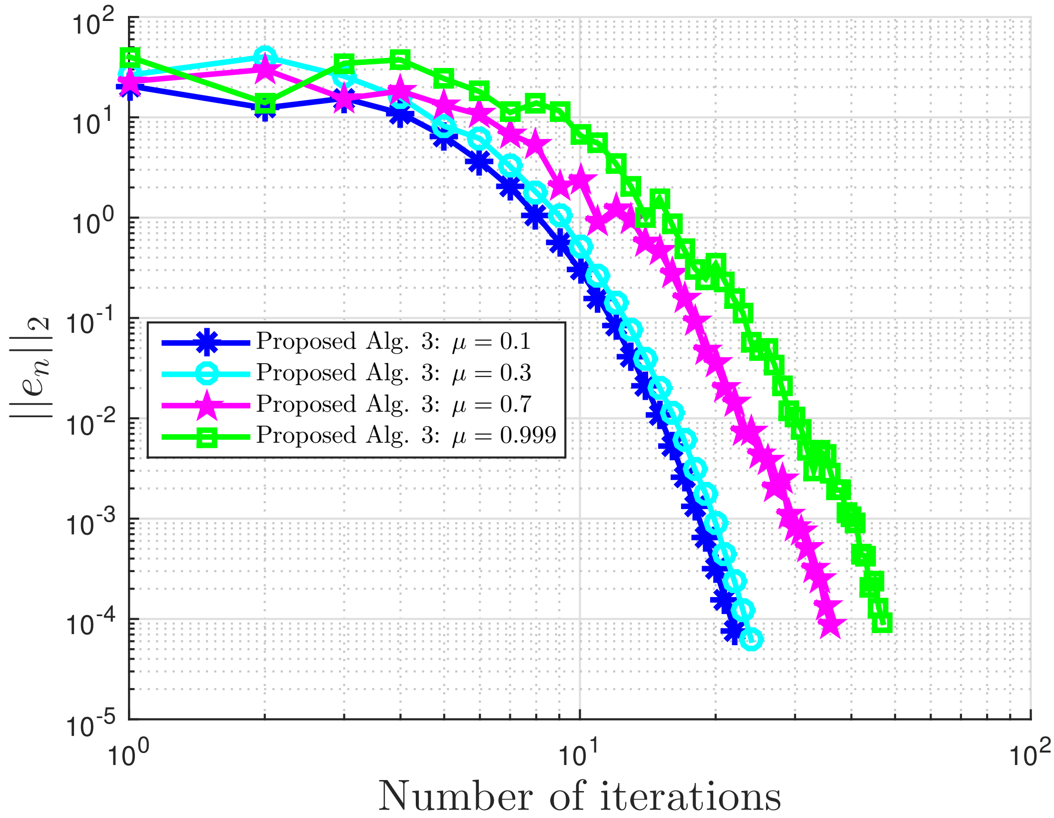

Figure 11. Example 2: $\lambda = 1$.

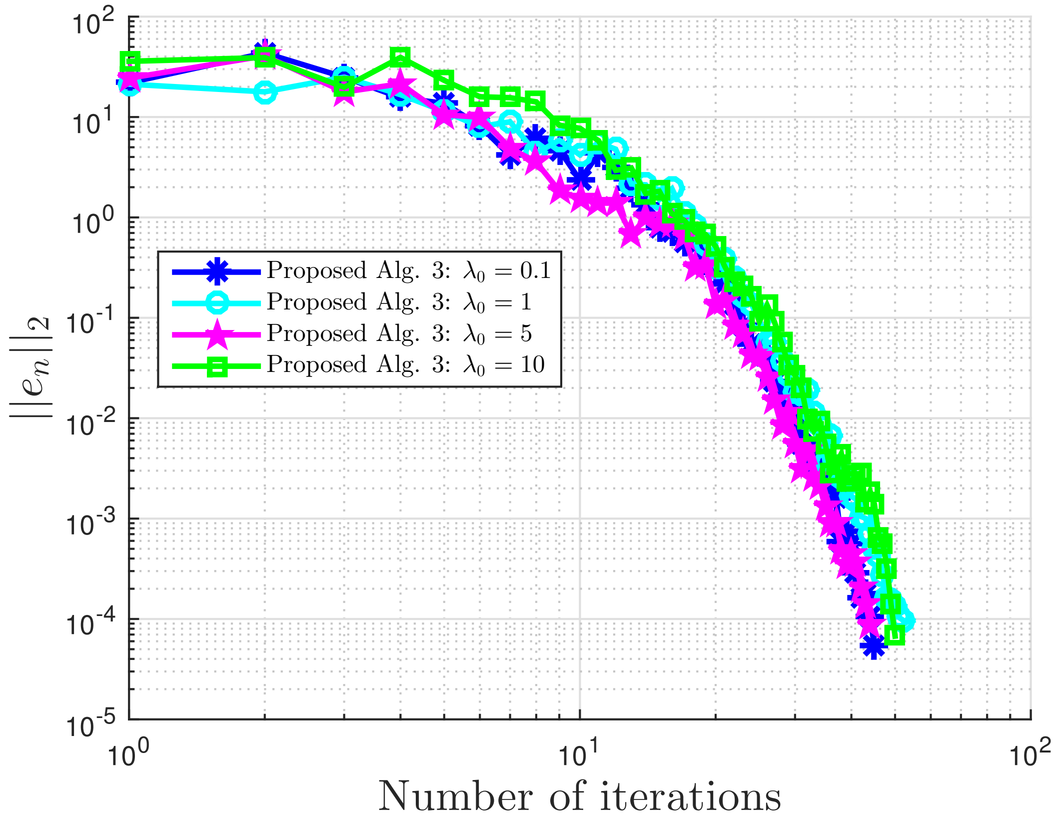

Figure 12. Example 2: $\mu = 0.999$.

Table 5. Comparison of proposed Algorithm 3 and the subgradient-extragradient method (SEM) (2) for Example 3.

	$m = 10$		$m = 20$		$m = 30$		$m = 40$	
$k = 30$	**Iter.**	**CPU Time**	**Iter.**	**CPU Time**	**Iter.**	**CPU Time**	**Iter.**	**CPU Time**
Proposed Algorithm 3	157	2.7327	162	3.9759	144	4.4950	128	4.8193
SEM (2)	3785	64.8752	13,980	243.9019	18,994	345.3686	30,777	567.8440
	$m = 10$		$m = 20$		$m = 30$		$m = 40$	
$k = 50$	**Iter.**	**CPU Time**	**Iter.**	**CPU Time**	**Iter.**	**CPU Time**	**Iter.**	**CPU Time**
Proposed Algorithm 3	185	4.0691	173	4.3798	128	4.8817	130	6.2658
SEM (2)	4176	77.6893	8645	150.4267	21,262	381.0991	30,956	561.4559

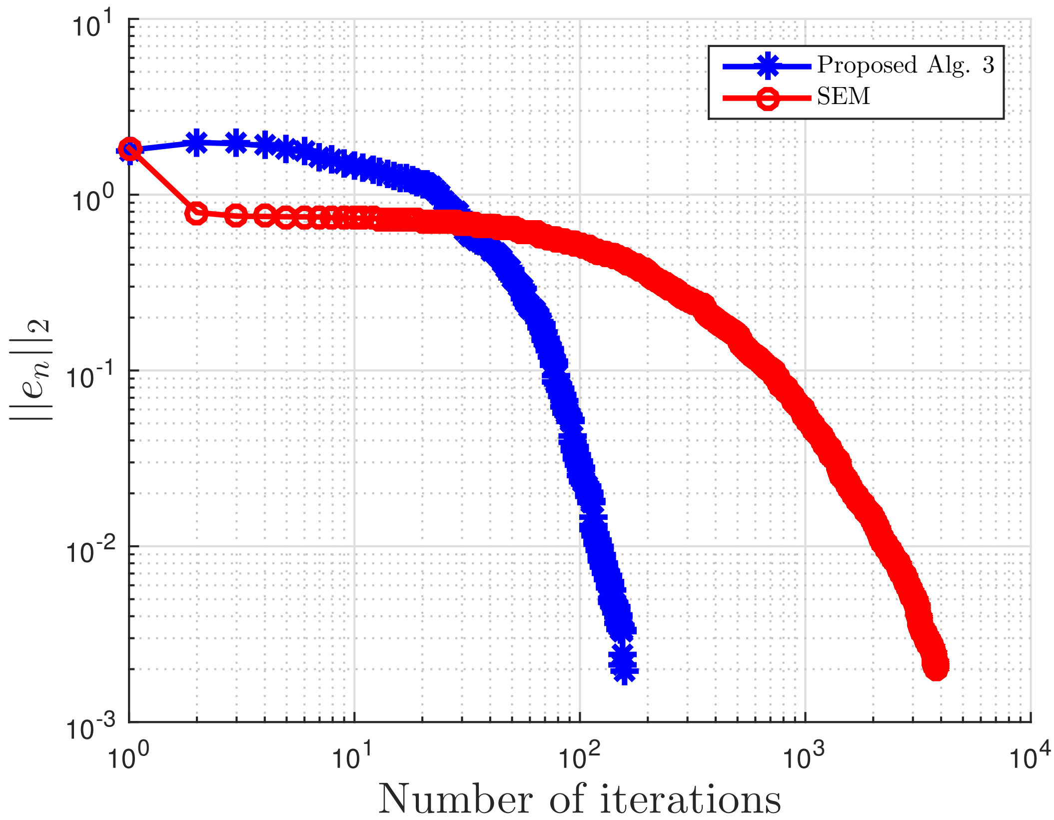

Figure 13. Example 3: $k = 30$ and $m = 10$.

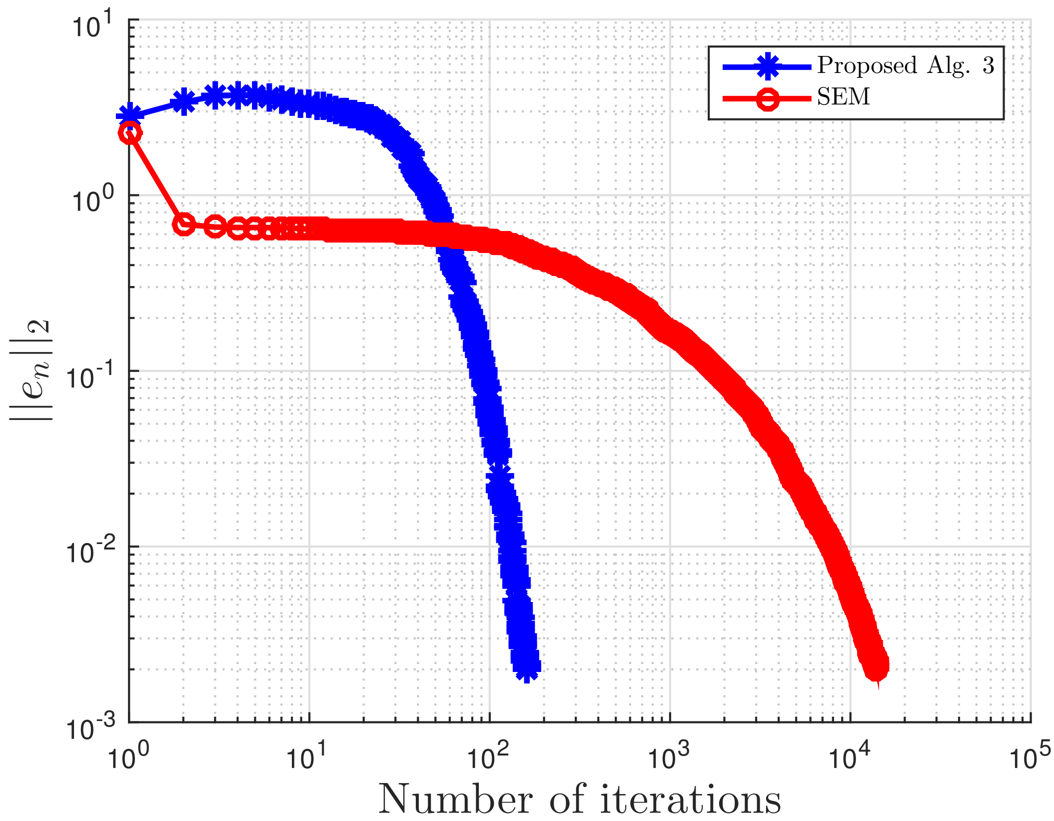

Figure 14. Example 3: $k = 30$ and $m = 20$.

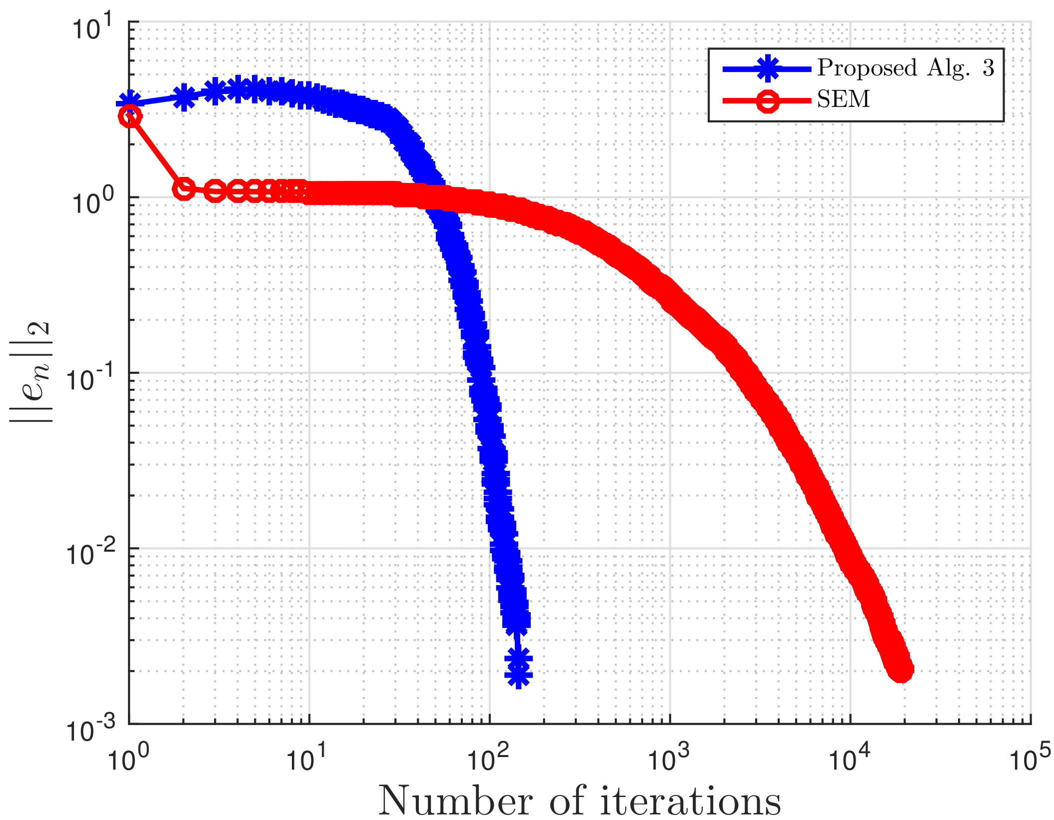

Figure 15. Example 3: $k = 30$ and $m = 30$.

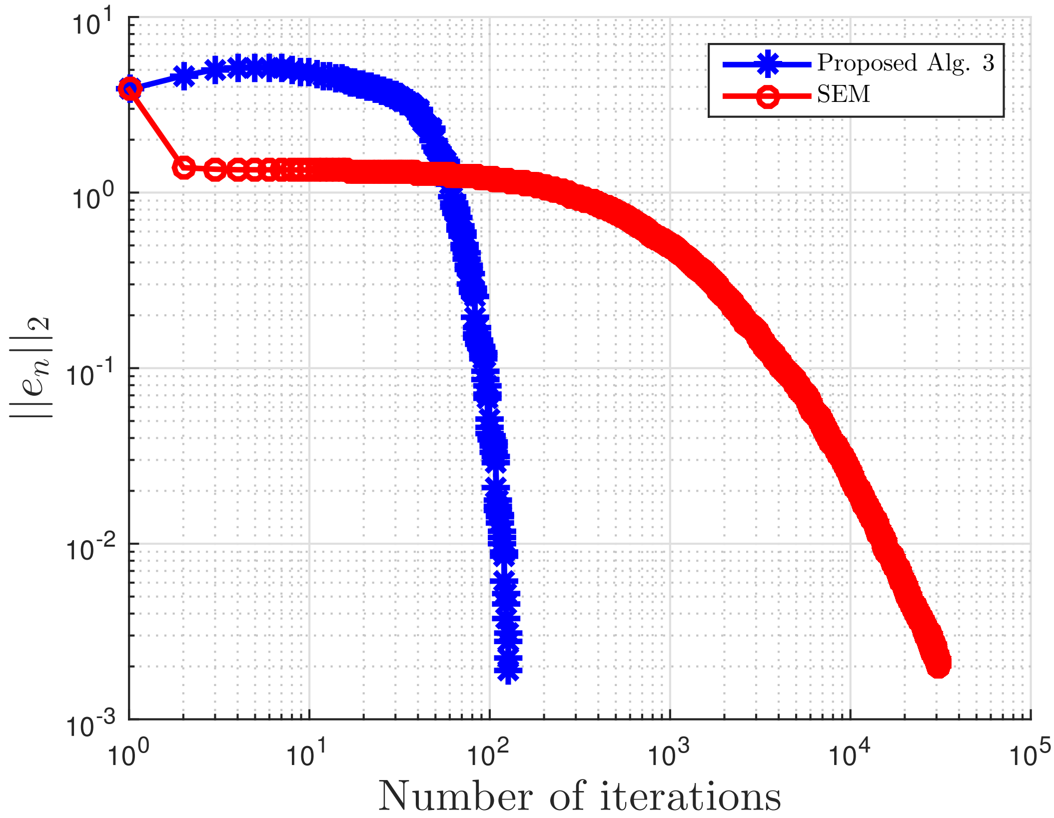

Figure 16. Example 3: $k = 30$ and $m = 40$.

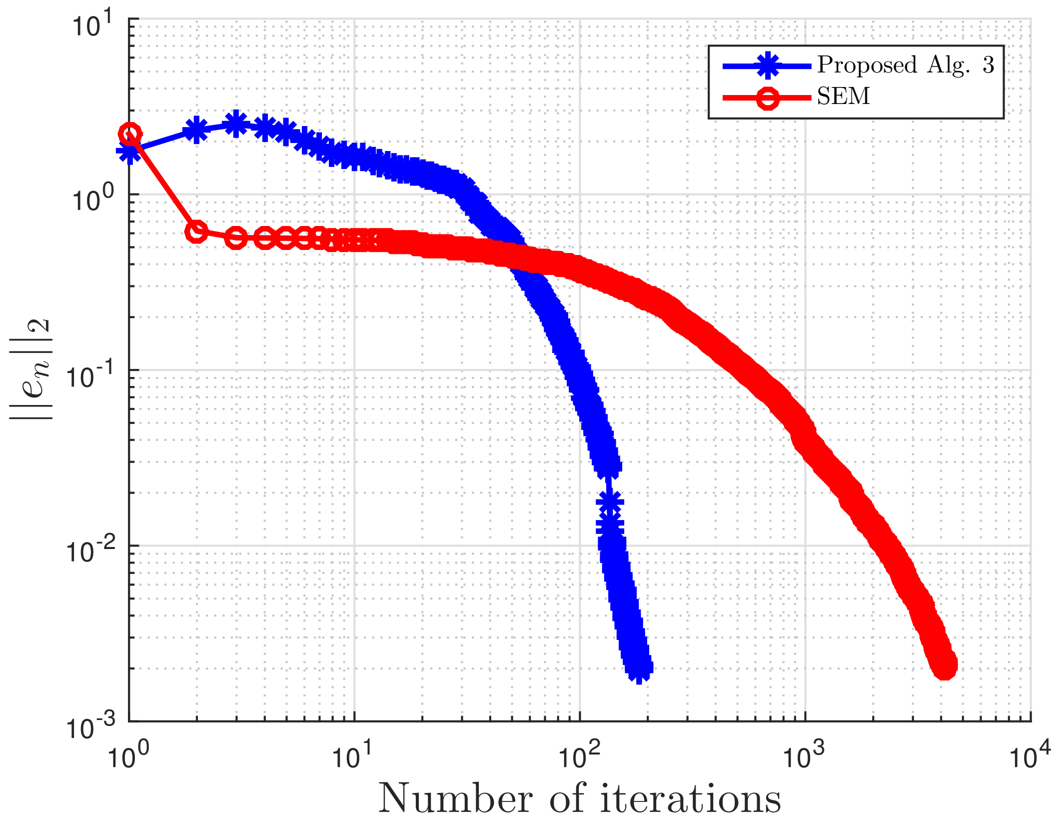

Figure 17. Example 3: $k = 50$ and $m = 10$.

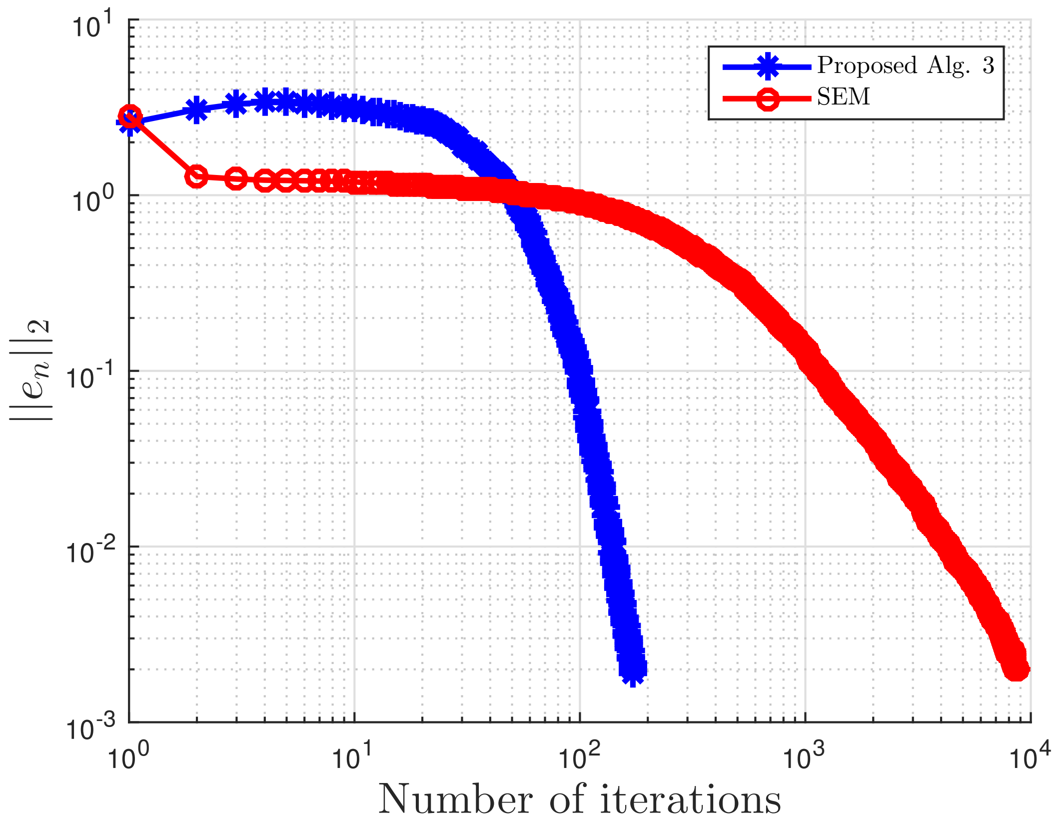

Figure 18. Example 3: $k = 50$ and $m = 20$.

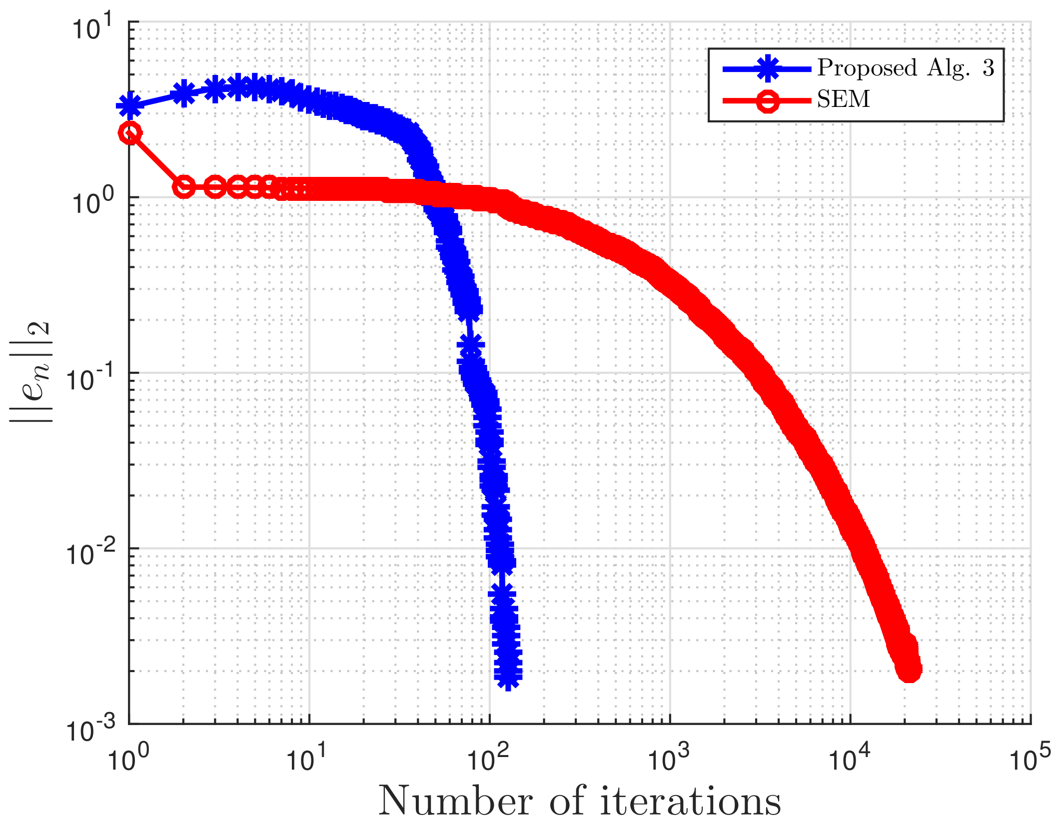

Figure 19. Example 3: $k = 50$ and $m = 30$.

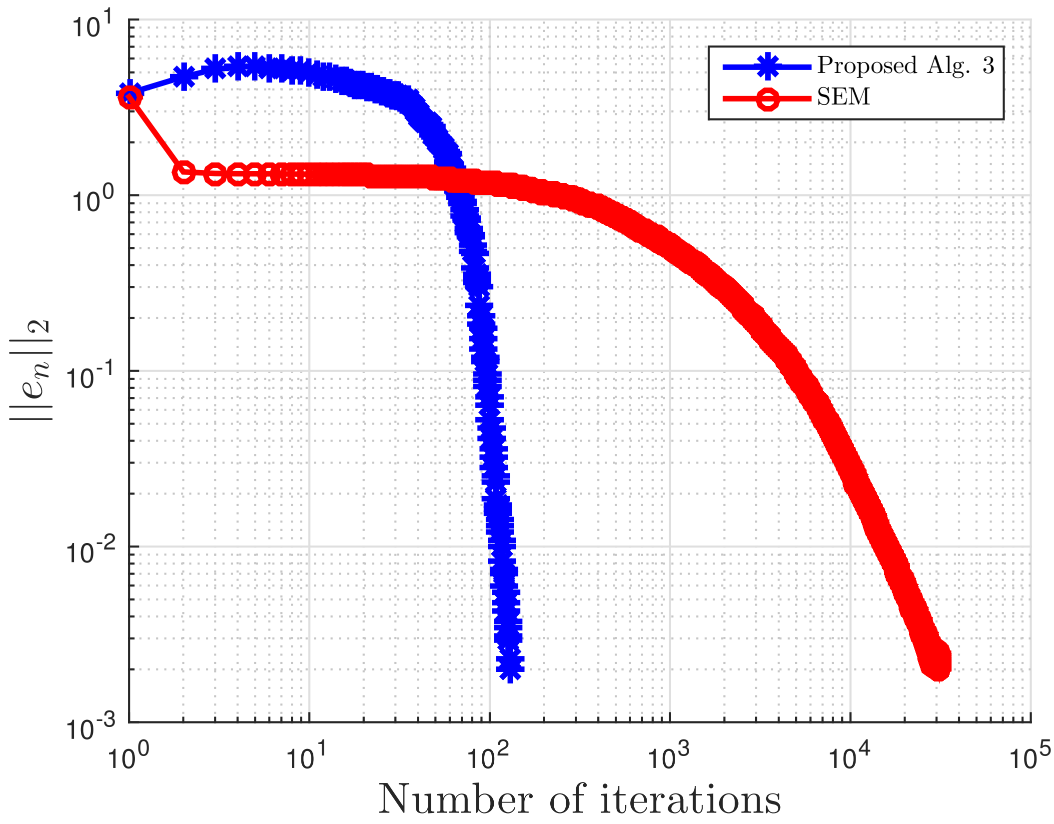

Figure 20. Example 3: $k = 50$ and $m = 40$.

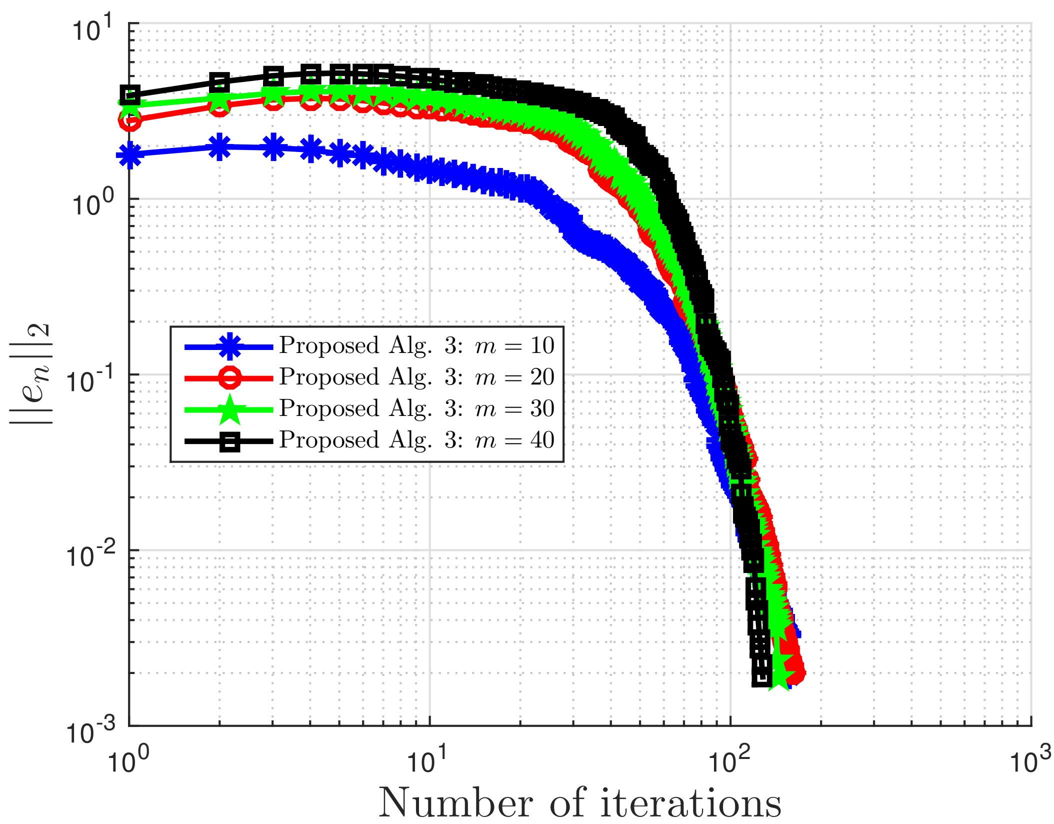

Figure 21. Example 3: $k = 30$.

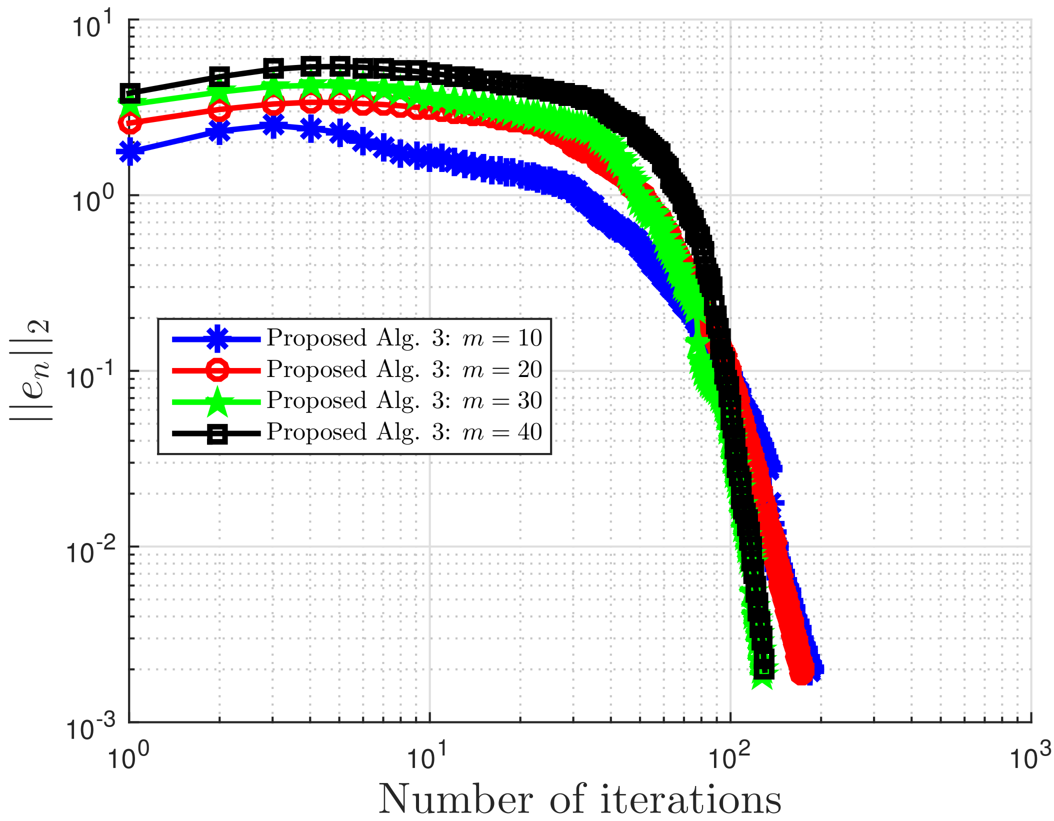

Figure 22. Example 3: $k = 50$.

Table 6. Comparison of proposed Algorithm 3 and the extragradient method (EGM) [16] for Example 4 with $\lambda_0 = 1$ and $\mu = 0.1$.

	Proposed Algorithm 3		**EGM [16]**	
x_1	**No. of Iter.**	**CPU Time**	**No. of Iter.**	**CPU Time**
$(2,1)^T$	13	5.9390×10^{-4}	62	1.8851×10^{-3}
$(1,2)^T$	33	5.2230×10^{-4}	62	1.9666×10^{-3}
$(1.5,1.5)^T$	12	4.7790×10^{-4}	14	3.955×10^{-4}
$(1.25,1.75)^T$	13	6.7980×10^{-4}	58	1.9138×10^{-3}

Table 7. Proposed Algorithm 3 for Example 4 with $\lambda_0 = 5$ and $\mu = 0.999$.

	Proposed Algorithm 3	
x_1	**No. of Iter.**	**CPU Time**
$(2,1)^T$	18	7.5980×10^{-4}
$(1,2)^T$	15	6.0910×10^{-4}
$(1.5,1.5)^T$	13	5.5840×10^{-4}
$(1.25,1.75)^T$	16	6.9110×10^{-4}

Table 8. Example 4: proposed Algorithm 3 with $x_1 = (2,1)^T$ for different μ and λ_0 values.

	$\boldsymbol{\lambda_0 = 5}$			
	$\boldsymbol{\mu = 0.1}$	$\boldsymbol{\mu = 0.3}$	$\boldsymbol{\mu = 0.7}$	$\boldsymbol{\mu = 0.999}$
No. of Iter.	14	14	16	18
CPU Time	6.8600×10^{-4}	7.1170×10^{-4}	8.1080×10^{-4}	8.6850×10^{-4}
	$\boldsymbol{\mu = 0.999}$			
	$\boldsymbol{\lambda_0 = 0.1}$	$\boldsymbol{\lambda_0 = 1}$	$\boldsymbol{\lambda_0 = 5}$	$\boldsymbol{\lambda_0 = 10}$
No. of Iter.	12	18	18	18
CPU Time	5.6360×10^{-4}	9.3920×10^{-4}	9.0230×10^{-4}	8.9310×10^{-4}

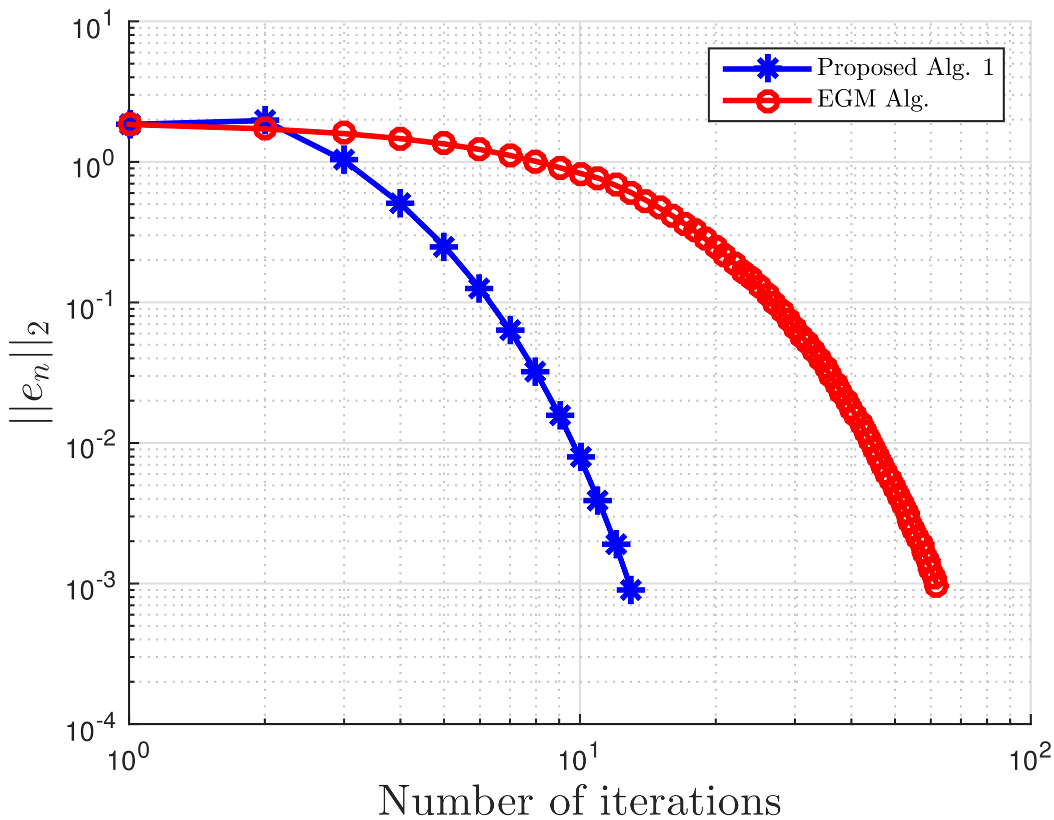

Figure 23. Example 4: $k = 50$ and $m = 10$.

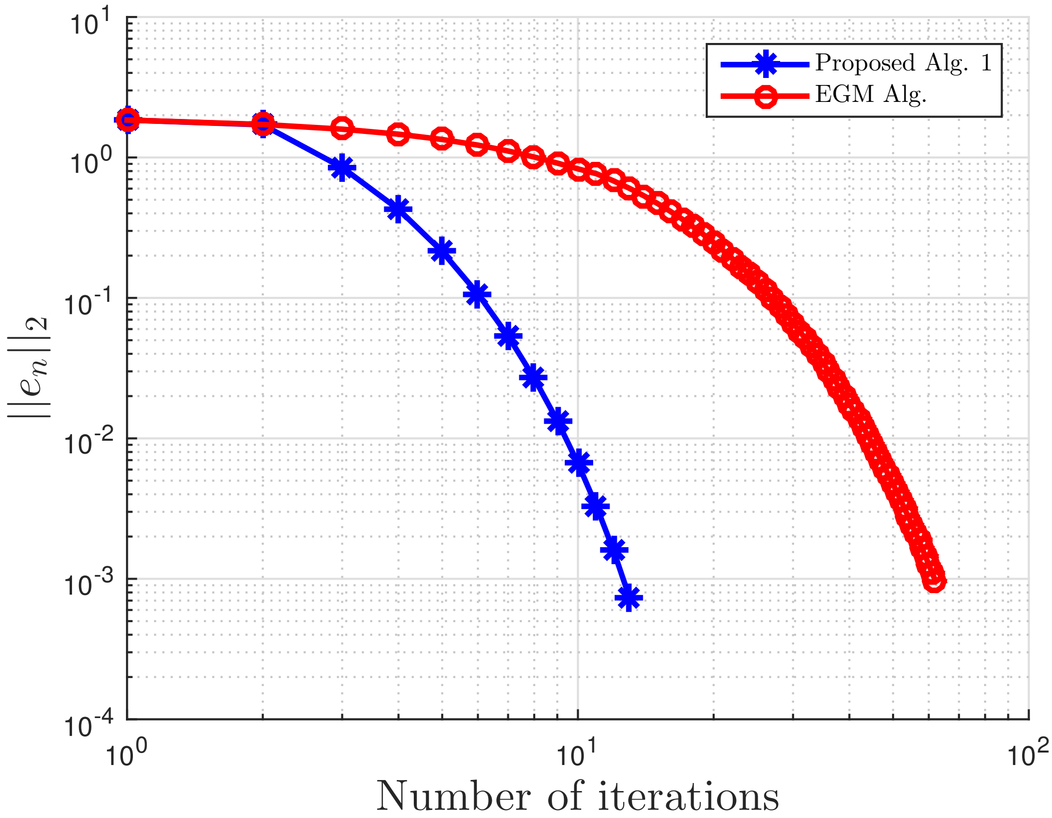

Figure 24. Example 4: $k = 50$ and $m = 20$.

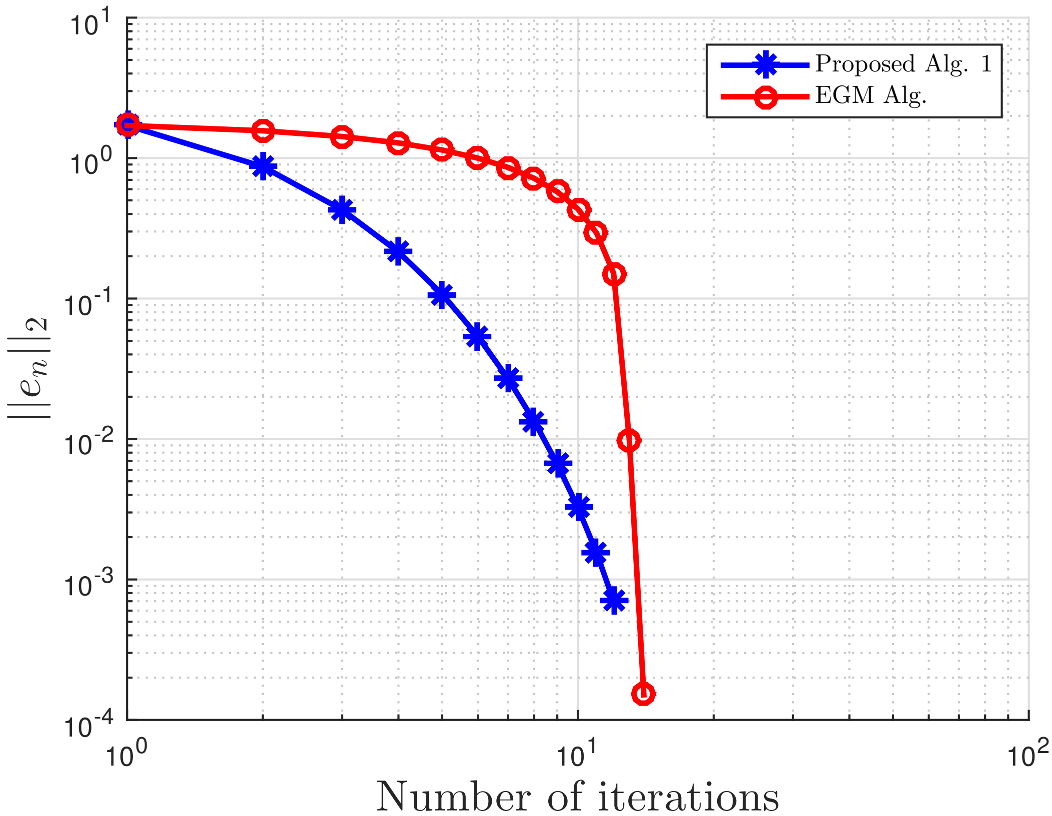

Figure 25. Example 4: $k = 50$ and $m = 30$.

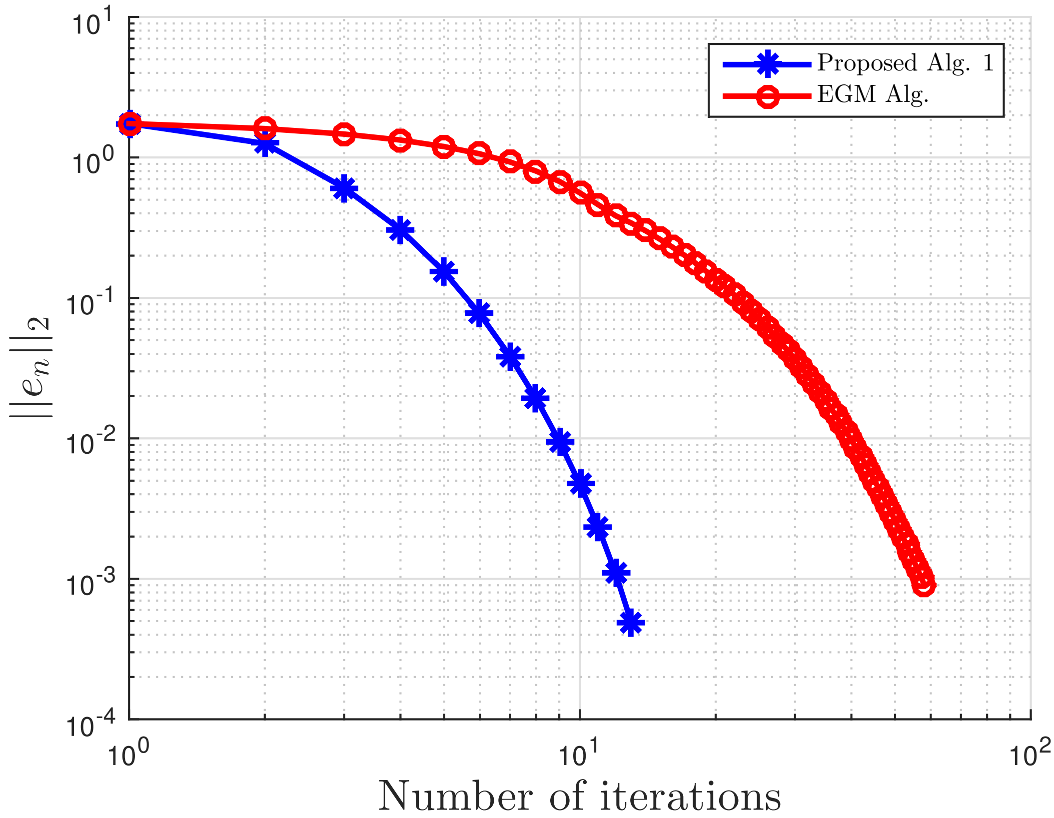

Figure 26. Example 4: $k = 50$ and $m = 40$.

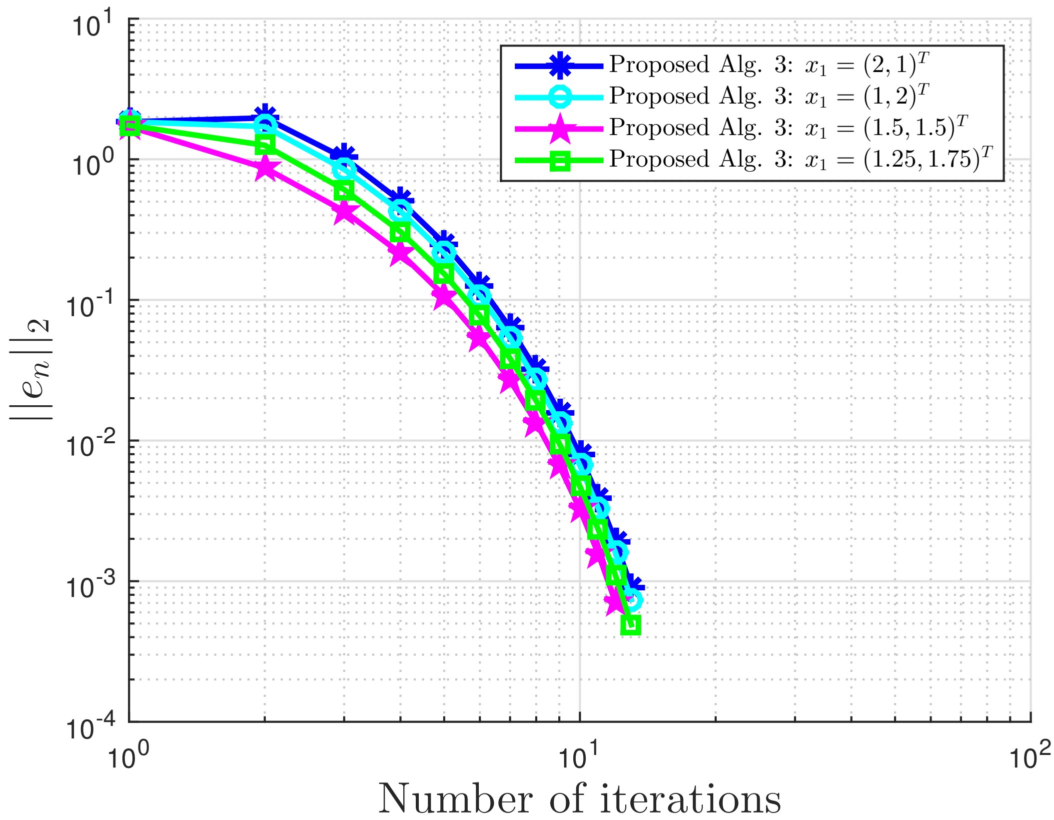

Figure 27. Example 4: $\lambda_0 = 1$ and $\mu = 0.1$.

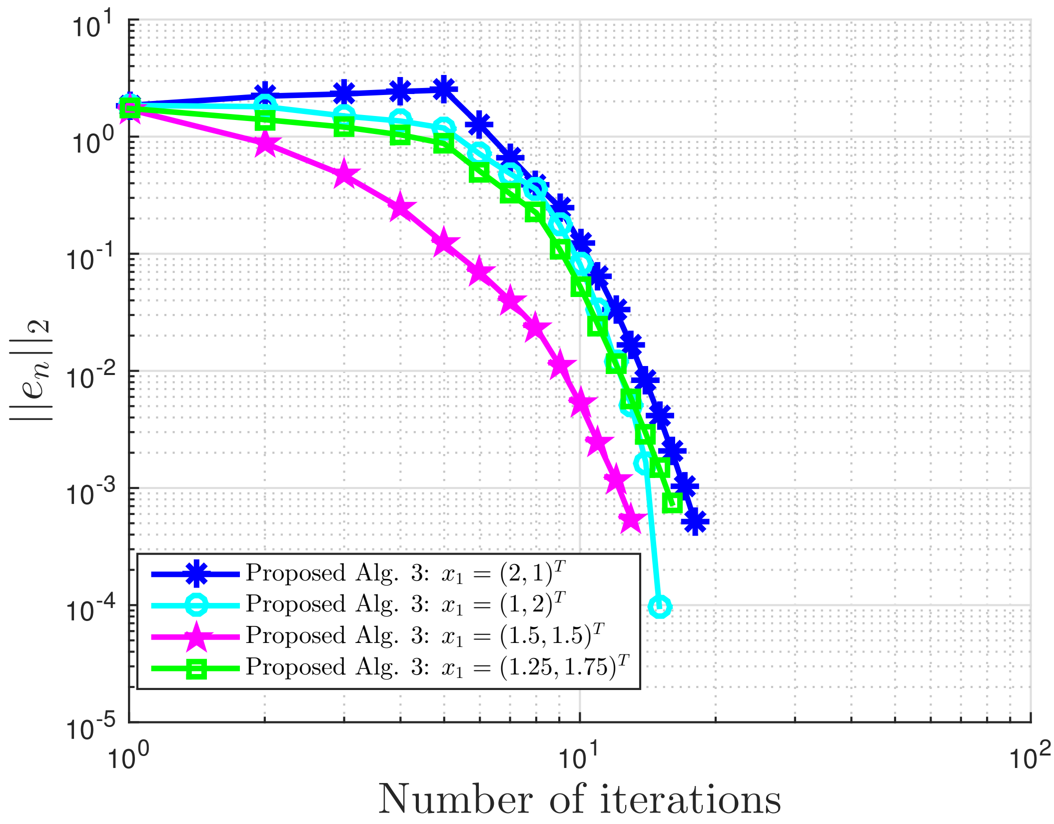

Figure 28. Example 4: $\lambda_0 = 5$ and $\mu = 0.999$.

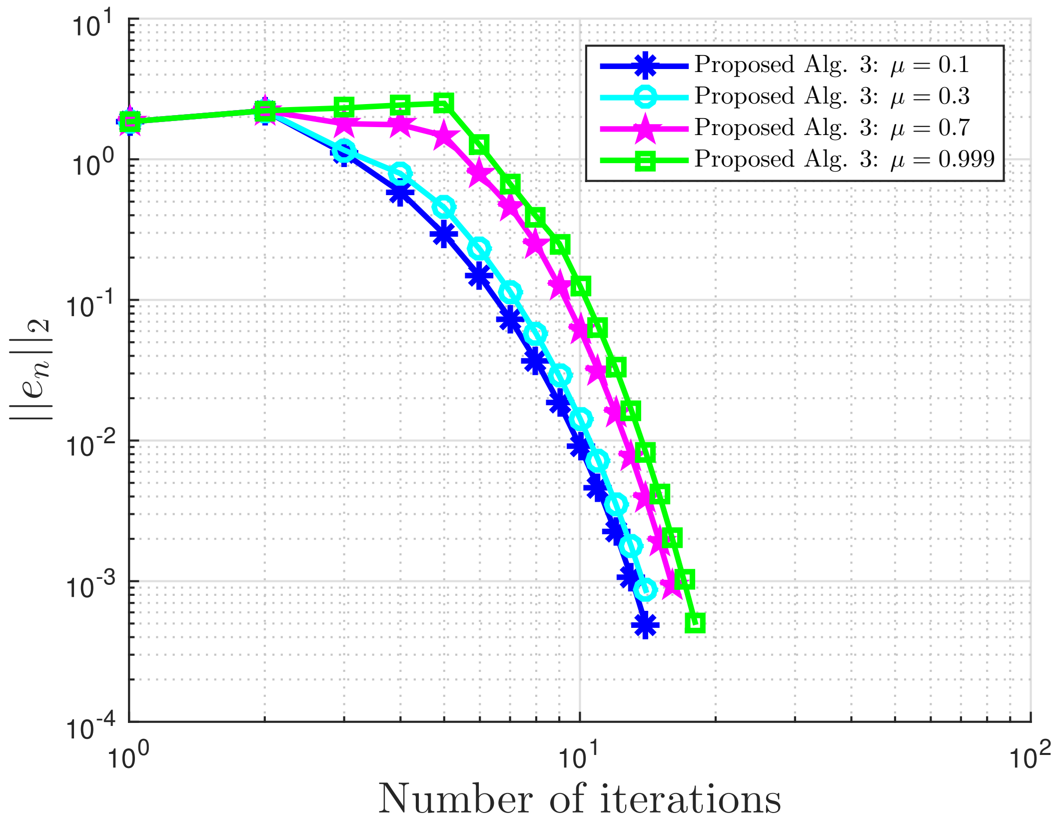

Figure 29. Example 4: $\lambda_0 = 5$ and $x_1 = (2,1)^T$.

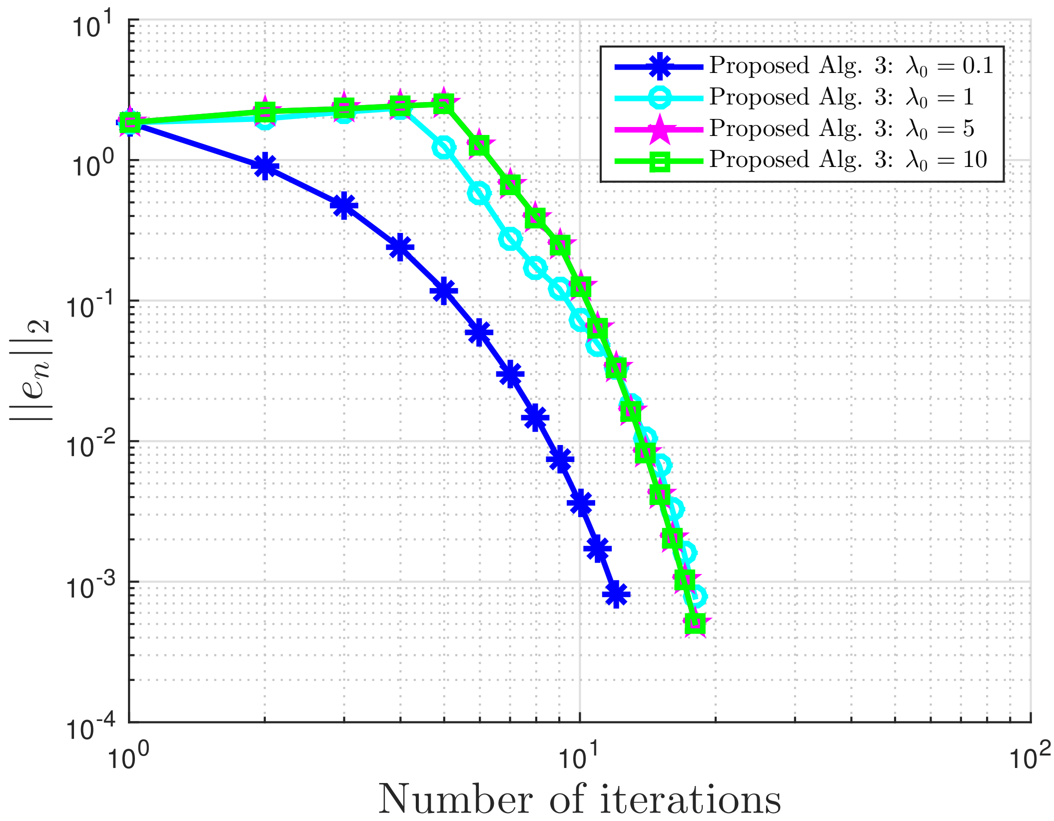

Figure 30. Example 4: $\mu = 0.999$ and $x_1 = (2,1)^T$.

Table 9. Example 5 comparison: proposed Algorithm 3, Bot Algorithm 1, Malitsky Algorithm 2, and Shehu Alg. [37] with $\lambda_0 = 1$ and $\mu = 0.9$.

	Proposed Algorithm 3		**Bot Algorithm 1**		**Malitsky Algorithm 2**		**Shehu Alg. [37]**	
x_1	**Iter.**	**CPU Time**	**Iter.**	**CPU Time**	**Iter.**	**CPU Time**	**Iter.**	**CPU Time**
$\frac{1}{12}(t^2 - 2t + 1)$	23	5.1433×10^{-3}	2159	0.23341	371	3.7625×10^{-2}	37,135	10.8377
$\frac{1}{9}e^t \sin(t)$	18	3.2111×10^{-3}	1681	0.18084	374	3.9159×10^{-2}	70,741	32.5843
$\frac{1}{21}t^2 \cos(t)$	14	2.4545×10^{-3}	4344	0.48573	373	3.8413×10^{-2}	17,741	5.6272
$\frac{1}{7}(3t-2)e^t$	43	8.7538×10^{-3}	2774	0.29515	351	3.7544×10^{-2}	28,758	7.3424

Table 10. Example 5: proposed Algorithm 3 with $x_1 = \frac{t^2-2t+1}{12}$ for different μ and λ_0 values.

	$\boldsymbol{\mu = 0.9}$			
	$\boldsymbol{\lambda_0 = 0.1}$	$\boldsymbol{\lambda_0 = 1}$	$\boldsymbol{\lambda_0 = 2}$	$\boldsymbol{\lambda_0 = 3}$
No. of Iter.	167	23	11	8
CPU Time	3.4828×10^{-2}	4.2089×10^{-3}	2.2288×10^{-3}	1.3899×10^{-3}
	$\boldsymbol{\lambda_0 = 2}$			
	$\boldsymbol{\mu = 0.1}$	$\boldsymbol{\mu = 0.3}$	$\boldsymbol{\mu = 0.7}$	$\boldsymbol{\mu = 0.999}$
No. of Iter.	11	11	11	11
CPU Time	2.0615×10^{-3}	2.1237×10^{-3}	1.9886×10^{-3}	2.0384×10^{-3}

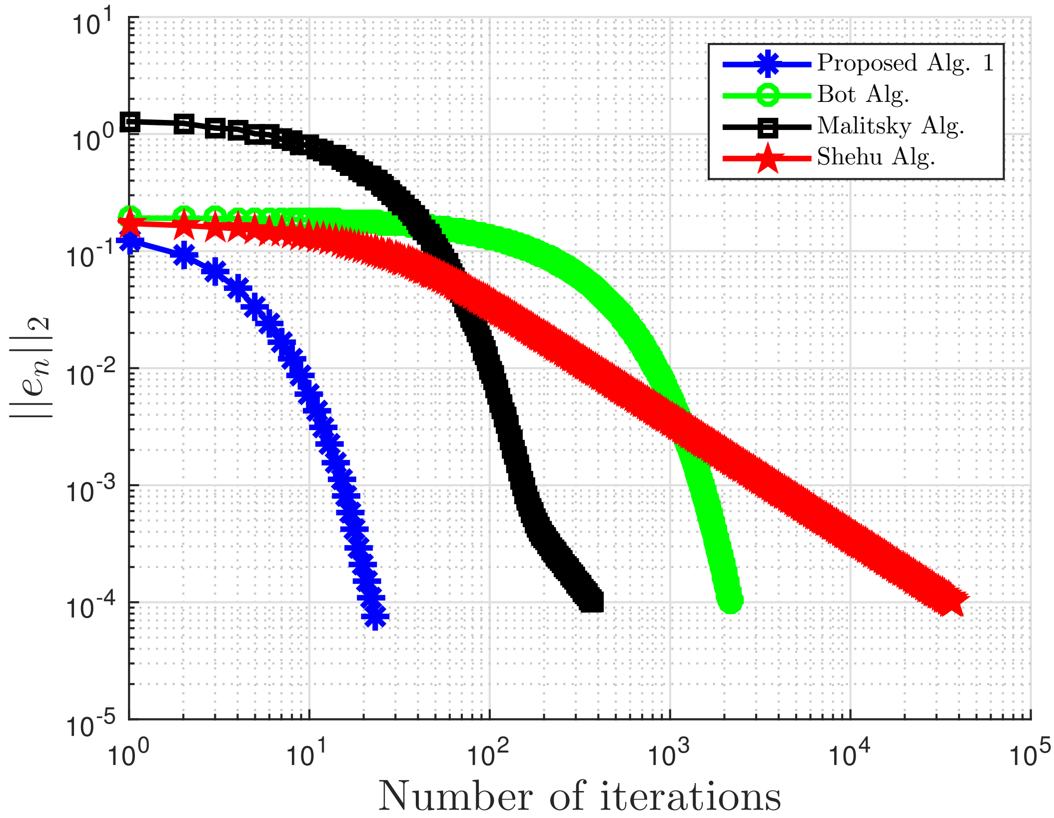

Figure 31. Example 5: $\lambda_0 = 1$, $\mu = 0.9$ and $x_1 = \frac{1}{12}(t^2 - 2t + 1)$.

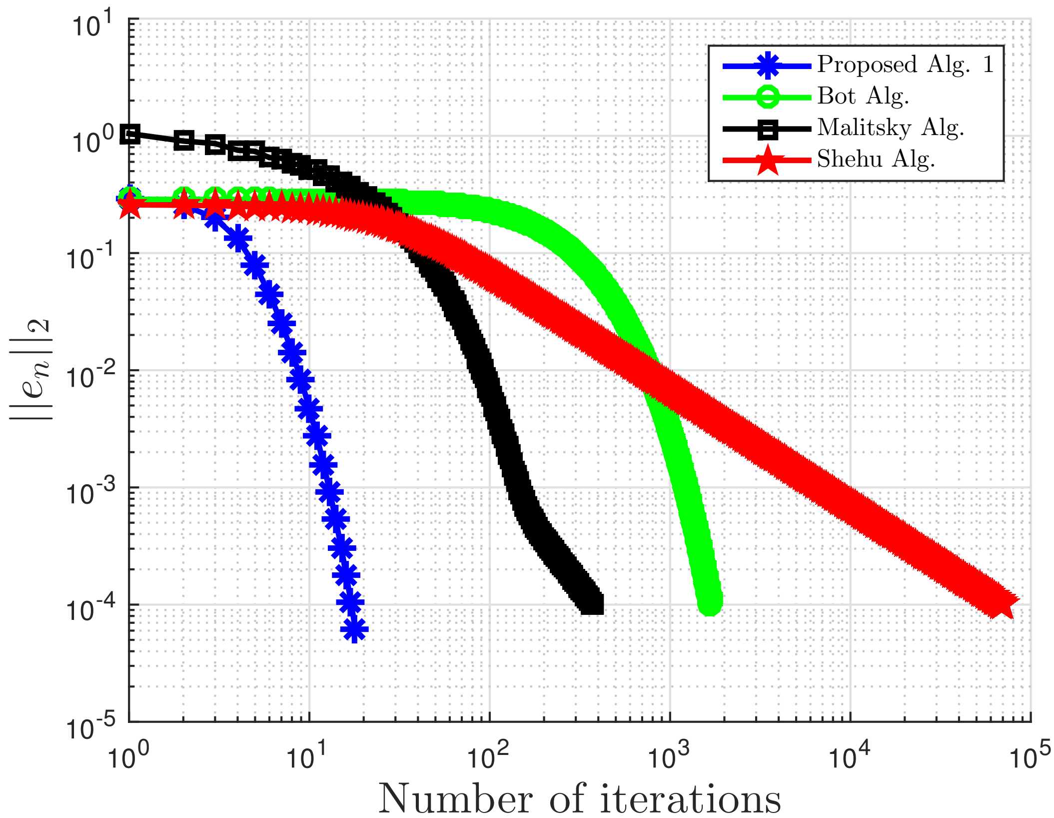

Figure 32. Example 5: $\lambda_0 = 1$, $\mu = 0.9$ and $x_1 = \frac{1}{9}e^t \sin(t)$.

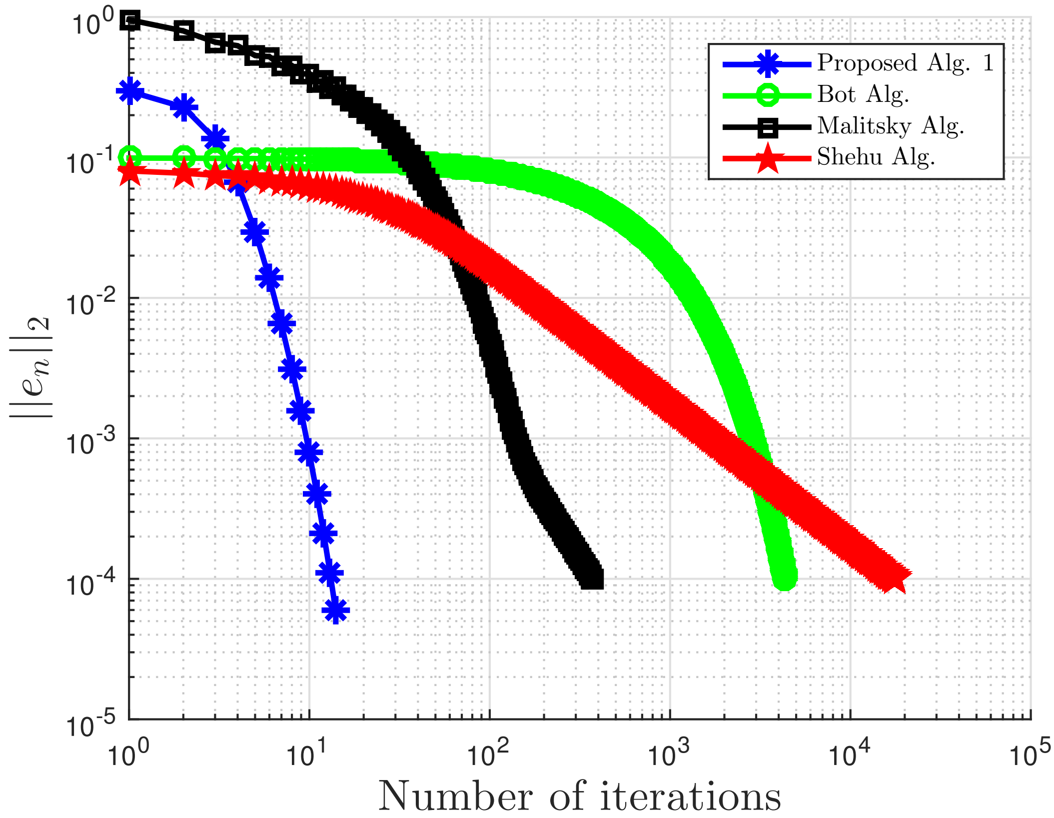

Figure 33. Example 5: $\lambda_0 = 1$, $\mu = 0.9$ and $x_1 = \frac{1}{21}t^2\cos(t)$.

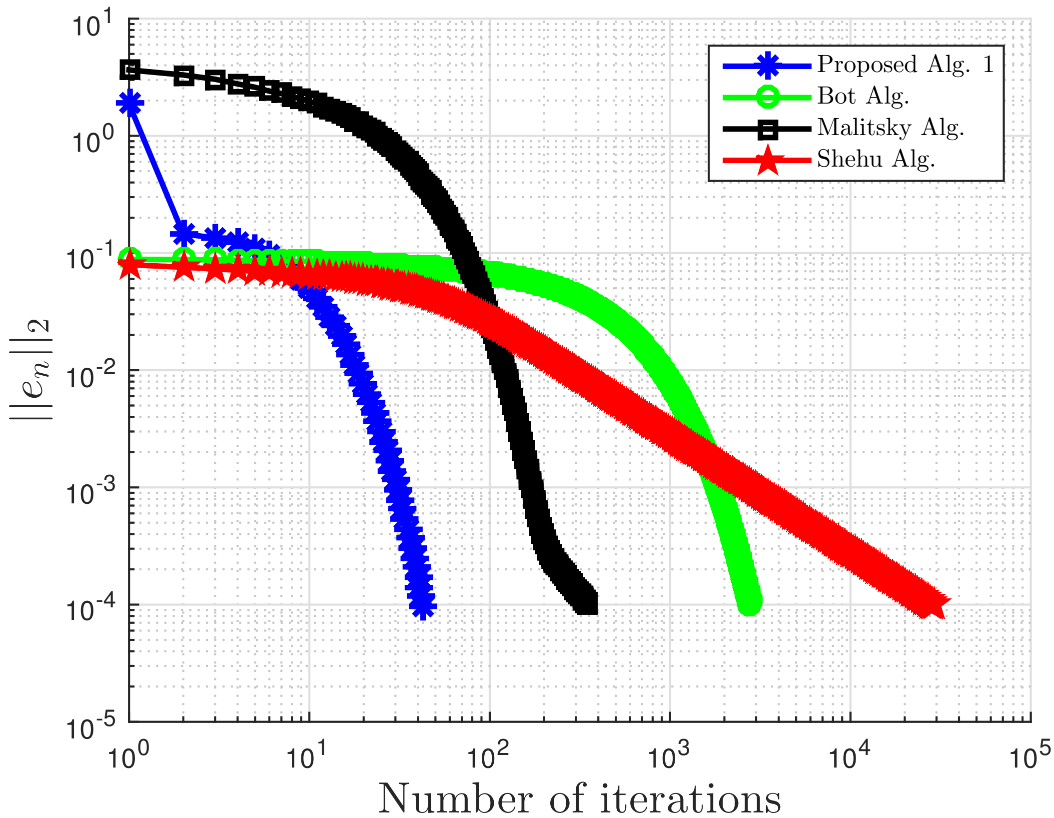

Figure 34. Example 5: $\lambda_0 = 1$, $\mu = 0.9$ and $x_1 = \frac{1}{7}(3t-2)e^t$.

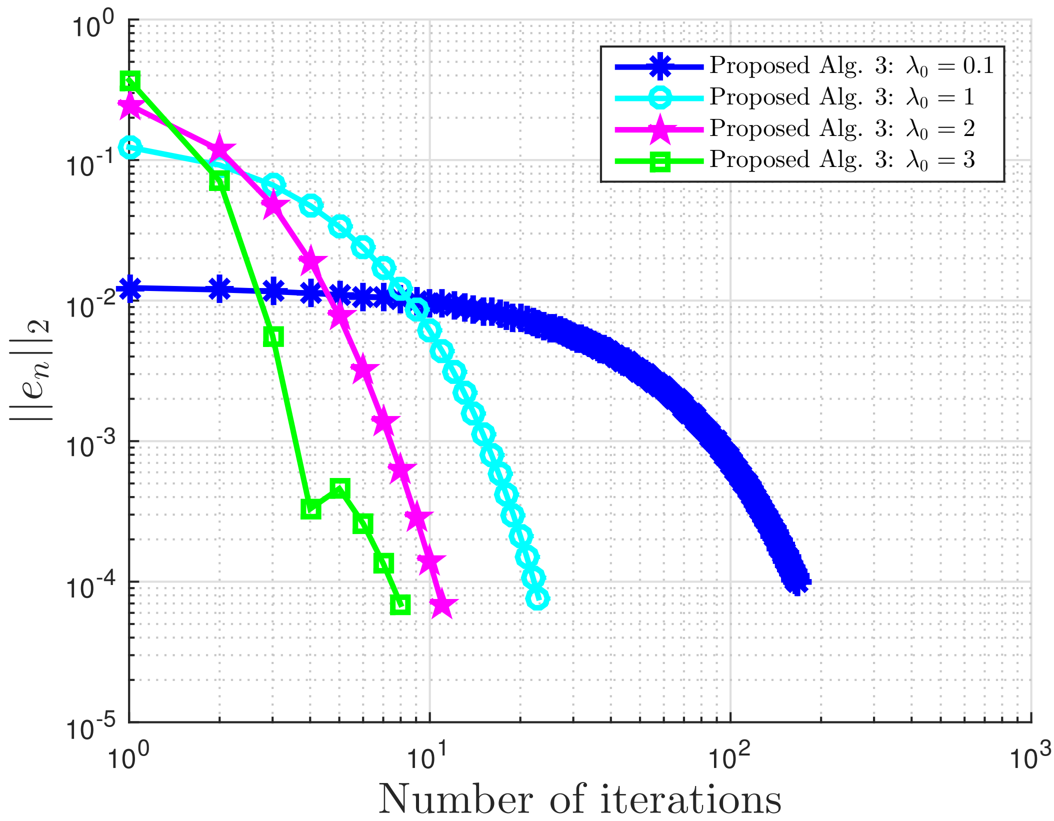

Figure 35. Example 5: $\mu = 0.9$.

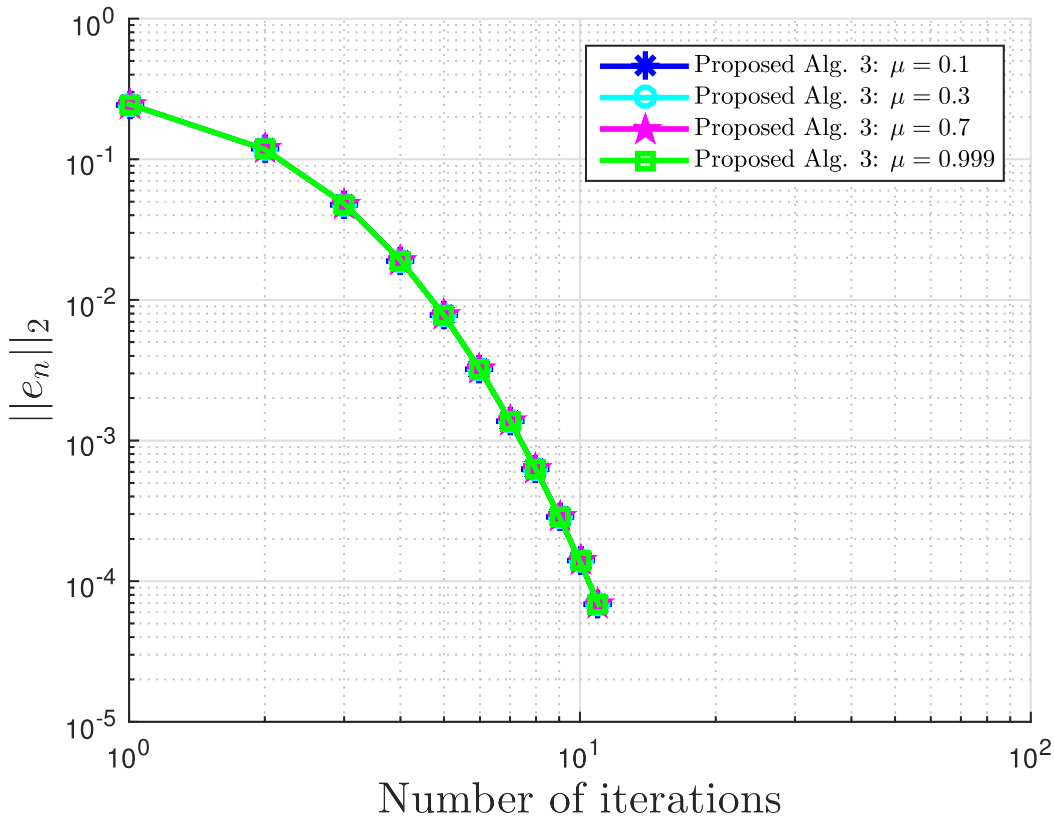

Figure 36. Example 5: $\lambda_0 = 2$.

5. Discussion

The weak convergence analysis of the reflected subgradient-extragradient method for variational inequalities in real Hilbert spaces is obtained in this paper. We provide and intensive numerical illustration and comparison with related works for several applications such as tomography reconstruction and Nash–Cournot oligopolistic equilibrium models. Our result is one of the few results on the subgradient-extragradient method with the reflected step in the literature. Our next project is to modify our results to bilevel variational inequalities.

Author Contributions: Writing—review and editing, A.G., O.S.I., L.A. and Y.S. All authors contributed equally to this work which included mathematical theory and analysis and code implementation. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The study does not report any data.

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

Abbreviations

The following abbreviations are used in this manuscript:

VI	variational inequality problem
EGM	extragradient method
SEGM	subgradient-extragradient method

References

1. Fichera, G. Sul problema elastostatico di Signorini con ambigue condizioni al contorno. *Atti Accad. Naz. Lincei VIII Ser. Rend. Cl. Sci. Fis. Mat. Nat.* **1963**, *34*, 138–142.
2. Fichera, G. Problemi elastostatici con vincoli unilaterali: il problema di Signorini con ambigue condizioni al contorno. *Atti Accad. Naz. Lincei Mem. Cl. Sci. Fis. Mat. Nat. Sez. I VIII Ser.* **1964**, *7*, 91–140.
3. Stampacchia, G. Formes bilineaires coercitives sur les ensembles convexes. *C. R. Acad. Sci.* **1964**, *258*, 4413–4416.
4. Kinderlehrer, D.; Stampacchia, G. *An Introduction to Variational Inequalities and Their Applications*; Academic: New York, NY, USA, 1980.
5. Facchinei, F.; Pang, J.S. *Finite-Dimensional Variational Inequalities and Complementarity Problems. Volume I*; Springer Series in Operations Research; Springer: New York, NY, USA, 2003.
6. Facchinei, F.; Pang, J.S. *Finite-Dimensional Variational Inequalities and Complementarity Problems. Volume II*; Springer Series in Operations Research; Springer: New York, NY, USA, 2003.
7. Konnov, I.V. *Combined Relaxation Methods for Variational Inequalities*; Springer: Berlin/Heidelberg, Germany, 2001.
8. Iusem, A.N. An iterative algorithm for the variational inequality problem. *Comput. Appl. Math.* **1994**, *13*, 103–114.
9. Iusem, A.N.; Svaiter, B.F. A variant of Korpelevich's method for variational inequalities with a new search strategy. *Optimization* **1997**, *42*, 309–321. [CrossRef]
10. Kanzow, C.; Shehu, Y. Strong convergence of a double projection-type method for monotone variational inequalities in Hilbert spaces. *J. Fixed Point Theory Appl.* **2018**, *20*, 51. [CrossRef]
11. Khobotov, E.N. Modifications of the extragradient method for solving variational inequalities and certain optimization problems. *USSR Comput. Math. Math. Phys.* **1987**, *27*, 120–127. [CrossRef]
12. Maingé, P.E. A hybrid extragradient-viscosity method for monotone operators and fixed point problems. *SIAM J. Control Optim.* **2008**, *47*, 1499–1515. [CrossRef]
13. Marcotte, P. Application of Khobotov's algorithm to variational inequalities and network equilibrium problems. *Inf. Syst. Oper. Res.* **1991**, *29*, 258–270. [CrossRef]
14. Vuong, P.T. On the weak convergence of the extragradient method for solving pseudo-monotone variational inequalities. *J. Optim. Theory Appl.* **2018**, *176*, 399–409. [CrossRef]
15. Shehu, Y.; Iyiola, O.S. Weak convergence for variational inequalities with inertial-type method. *Appl. Anal.* **2020**, *2020*, 1–25. [CrossRef]
16. Korpelevich, G.M. The extragradient method for finding saddle points and other problems. *Ekon. Matematicheskie Metody* **1976**, *12*, 747–756.

17. Antipin, A.S. On a method for convex programs using a symmetrical modification of the Lagrange function. *Ekon. Matematicheskie Metody* **1976**, *12*, 1164–1173.
18. Censor, Y.; Gibali, A.; Reich, S. The subgradient-extragradient method for solving variational inequalities in Hilbert space. *J. Optim. Theory Appl.* **2011**, *148*, 318–335. [CrossRef]
19. Censor, Y.; Gibali, A.; Reich, S. Strong convergence of subgradient-extragradient methods for the variational inequality problem in Hilbert space. *Optim. Methods Softw.* **2011**, *26*, 827–845. [CrossRef]
20. Censor, Y.; Gibali, A.; Reich, S. Extensions of Korpelevich's extragradient method for the variational inequality problem in Euclidean space. *Optimization* **2011**, *61*, 1119–1132. [CrossRef]
21. Malitsky, Y.V. Projected reflected gradient methods for monotone variational inequalities. *SIAM J. Optim.* **2015**, *25*, 502–520. [CrossRef]
22. Malitsky, Y.V.; Semenov, V.V. A hybrid method without extrapolation step for solving variational inequality problems. *J. Glob. Optim.* **2015**, *61*, 193–202. [CrossRef]
23. Solodov, M.V.; Svaiter, B.F. A new projection method for variational inequality problems. *SIAM J. Control Optim.* **1999**, *37*, 765–776. [CrossRef]
24. Denisov, S.V.; Semenov, V.V.; Chabak, L.M. Convergence of the modified extragradient method for variational inequalities with non-Lipschitz operators. *Cybern. Syst. Anal.* **2015**, *51*, 757–765. [CrossRef]
25. Yang, J.; Liu, H. Strong convergence result for solving monotone variational inequalities in Hilbert space. *Numer. Algorithms* **2019**, *80*, 741–752. [CrossRef]
26. Thong, D.V.; Hieu, D.V. Modified subgradient-extragradient method for variational inequality problems. *Numer. Algorithms* **2018**, *79*, 597–610. [CrossRef]
27. Boţ, R.I.; Csetnek, E.R.; Vuong, P.T. The forward-backward-forward method from discrete and continuous perspective for pseudomonotone variational inequalities in Hilbert spaces. *Eur. J. Oper. Res.* **2020**, *287*, 49–60. [CrossRef]
28. Malitsky, Y.V. Golden ratio algorithms for variational inequalities. *Math. Program.* **2020**, *184*, 383–410. [CrossRef]
29. Karamardian, S. Complementarity problems over cones with monotone and pseudo-monotone maps. *J. Optim. Theory Appl.* **1976**, *18*, 445–454. [CrossRef]
30. Goebel, K.; Reich, S. *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings*; Marcel Dekker: New York, NY, USA, 1984.
31. Kopecká, E.; Reich, S. A note on alternating projections in Hilbert spaces. *J. Fixed Point Theory Appl.* **2012**, *12*, 41–47. [CrossRef]
32. Cegielski, A. *Iterative Methods for Fixed Point Problems in Hilbert Spaces*; Lecture Notes in Mathematics; Springer: Berlin/Heidelberg, Germany, 2012.
33. Cottle, R.W.; Yao, J.C. Pseudo-monotone complementarity problems in Hilbert space. *J. Optim. Theory Appl.* **1992**, *75*, 281–295. [CrossRef]
34. Opial, Z. Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bull. Am. Math. Soc.* **1967**, *73*, 591–597. [CrossRef]
35. He, Y.R. A new double projection algorithm for variational inequalities. *J. Comput. Appl. Math.* **2006**, *185*, 166–173. [CrossRef]
36. Maingé, P.E. Convergence theorems for inertial KM-type algorithms. *J. Comput. Appl. Math.* **2008**, *219*, 223–236. [CrossRef]
37. Shehu, Y.; Iyiola, O.S. Iterative algorithms for solving fixed point problems and variational inequalities with uniformly continuous monotone operators. *Numer. Algorithms* **2018**, *79*, 529–553. [CrossRef]
38. Yen, L.H.; Muu, L.D.; Huyen, N.T.T. An algorithm for a class of split feasibility problems: Application to a model in electricity production. *Math. Methods Oper. Res.* **2016**, *84*, 549–565 . [CrossRef]
39. Harker, P.T.; Pang, J.-S. A damped-Newton method for the linear complementarity problem. In *Computational Solution of Nonlinear Systems of Equations, Lectures in Applied Mathematics*; Allgower, G., Georg, K., Eds.; AMS: Providence, RI, USA, 1990; Volume 26, pp. 265–284.