

Article

A Hybrid Forward–Backward Algorithm and Its Optimization Application

Liya Liu ¹, Xiaolong Qin ² and Jen-Chih Yao ^{3,*}

¹ School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China; lya42@std.uestc.edu.cn

² Department of Mathematics, Hangzhou Normal University, Hangzhou 31121, China; qxlxajh@hznu.edu.cn

³ Research Center for Interneural Computing, China Medical University Hospital, Taichung 40447, Taiwan

* Correspondence: yaojc@mail.cmu.edu.tw

Received: 25 February 2020; Accepted: 18 March 2020; Published: 19 March 2020



Abstract: In this paper, we study a hybrid forward–backward algorithm for sparse reconstruction. Our algorithm involves descent, splitting and inertial ideas. Under suitable conditions on the algorithm parameters, we establish a strong convergence solution theorem in the framework of Hilbert spaces. Numerical experiments are also provided to illustrate the application in the field of signal processing.

Keywords: forward–backward method; hybrid steepest decent method; inertial extrapolation; maximally monotone; strong convergence

1. Introduction

Let H be a real Hilbert space. $\langle \cdot, \cdot \rangle$ denotes the associated scalar product and $\| \cdot \|$ stands for the induced norm. Recall that a set-valued operator $G : H \rightarrow 2^H$ is said to be maximally monotone iff $\langle x - x', y - y' \rangle \geq 0, \forall y \in G(x), y' \in G(x')$, and its graph, denoted by $\text{gph } G := \{(x, y) \in H \times H \mid y \in G(x)\}$, is not properly contained in the graph of any other monotone operator on H .

A fundamental and classical problem is to find a zero of a maximally monotone operator G in a real Hilbert space H , namely

$$\text{find } x \in H \text{ such that } 0 \in G(x). \quad (1)$$

This problem includes, as special cases, bifunction equilibrium problems, convex-concave saddle-point problems, minimization problems, non-smooth variational inequalities, etc. Due to its diverse applications in economics, medicine, and engineering, the techniques and methods for solving (1) have received much attention in the optimization community, see [1–4]. Indeed, the interdisciplinary nature of Problem (1) is evident from the viewpoint of algorithmic developments; see, for instance, [5–8] and the references therein. Among them, a celebrated method for solving (1) is the following proximal point algorithm. It can be traced back to the early results obtained in [6–9]. Given the current iterate p_n , calculate the next iterate p_{n+1} via

$$p_{n+1} = (Id + \gamma_n G)^{-1} p_n, \quad \forall n \in \mathbb{N}, \quad (2)$$

where Id stands for the identity operator and γ_n is a positive real number. The operator $(Id + \gamma_n G)^{-1}$ is the so-called resolvent operator, which was introduced by Moreau in [8]. In the context of algorithms, the resolvent operator is often referred as the backward operator. In [7], Rockafeller studied the operator and proved that the sequence generated by algorithm (2) is weakly convergent via the inexact evaluation of the resolvent operator in the framework of infinite-dimensional real Hilbert spaces. However, the evaluation error has to satisfy a summability restriction. Indeed, this restriction

essentially implies that the resolvent operator is computed with the increasing accuracy. In fact, due to the computation of the resolvent operator is often hard to control in practice situation, this is still somewhat limiting. Quite often, computing the resolvent of a maximally monotone operator, which is an inverse problem, is not easy in many cases. Thus, this limits the practicability of the proximal point algorithm in its plain form. Aiming at this, a favorable situation occurs when the operator G can be written as the sum of two operators. Precisely, let us define $G = A + B$ such that the resolvent $(Id + \alpha A)^{-1}$ (implicit, backward step), and the evaluation of B (explicit, forward step) are much easier to compute than the full resolvent $(Id + \alpha G)^{-1}$. In so doing, we only consider the following inclusion problem: find a zero of an additively structured operator $A + B$, acting on space H , namely

$$\text{find } x \in H \text{ such that } 0 \in Ax + Bx, \tag{3}$$

where B is a smooth operator for which we can use direct evaluations, and A is a proxfriendly operator for which we can compute the resolvent. For solving (3), Lions and Mercier [10] proposed the forward–backward splitting method. It is based on the recursive application of a forward step with respect to B , followed by a backward step with respect to A . For any initial data $p_0 \in H$,

$$p_{n+1} = (Id + \gamma_n A)^{-1}(p_n - \gamma_n B p_n), \quad \forall n \in \mathbb{N},$$

where Id stands for the identity operator and $\gamma_n > 0$. Basically, the operator $A : H \rightarrow 2^H$ is maximally monotone, and the operator B is κ -cocoercive (i.e., $\exists \kappa > 0$ such that $\langle Bx - By, x - y \rangle \geq \rho \|Bx - By\|^2, \forall x, y \in H$) and ι -Lipschitz continuous (i.e., $\exists \iota > 0$ such that $\|Bx - By\| \leq \iota \|x - y\|, \forall x, y \in H$). It was proven that the generated sequence $(p_n)_{n \geq 0}$ converges weakly to a solution of (3). Of course, the problem decomposition is not the only consideration, the convergence rate is another. Accelerating first-order method is a subject of active research, which has been extensively studied. Since Polyak [6] introduced the so-called heavy ball method for minimizing a smooth convex function, much has been done on the development of first-order accelerated methods; see [11–15]. The inertial nature of the first-order accelerated method can be exploited in numerical computations to accelerate the trajectories and speed up the rate of convergence. In the context of algorithms, it is often referred as the inertial extrapolation, which involves two iterative steps and the second iterative step is defined based on the previous two iterates. In [16], Alvarez and Attouch employed the first-order accelerated method to study an initial proximal point algorithm for solving the problem of finding zero of a maximally monotone operator. This iteration can be written as the following form: for any initial data $p_0, p_1 \in H$,

$$p_{n+1} = (Id + \lambda_n G)^{-1}(p_n + \alpha_n(p_n - p_{n-1})), \tag{4}$$

where Id stands for the identity operator, G is a maximally monotone operator, the parameters $(\lambda_n)_{n \geq 0} \subset (0, \infty), (\alpha_n)_{n \geq 0} \subset [0, 1], \sum_{n=1}^{\infty} \alpha_n \|p_n - p_{n-1}\|^2 < \infty$. It was showed that the iterative sequence $(p_n)_{n \geq 0}$ generated by (4) converges to a solution of inclusion Problem (1) weakly. An alternative modification to the inertial forward–backward splitting algorithm or its modifications is the following algorithm proposed by Lorenz and Pock [17]. The algorithm weakly converges to a zero of the sum of two maximally monotone operators, with one of two operators being Lipschitz continuous and single valued. Starting with any data $p_0, p_1 \in H$, we define a sequence $(p_k)_{k \geq 0}$ as

$$\begin{cases} w_n = p_n + \alpha_n(p_n - p_{n-1}), \\ p_{n+1} = (M + \gamma_n A)^{-1}(M - \gamma_n B)w_n, \end{cases}$$

where M is a positive definite and linear self-adjoint operator that can be used as a preconditioner for the scheme, γ_n is a step-size parameter and $\alpha_n \in [0, 1)$ is an extrapolation factor. The algorithm keeps

the weak convergence property of the iterates. To motivate the so-called the inertial forward–backward splitting algorithm, we consider the following differential equation

$$\ddot{p}(t) + \iota(t)\dot{p}(t) + M_{A,B,\gamma}(p(t)) = 0, \tag{5}$$

where $A : H \rightarrow H$ and $B : H \rightarrow H$ are operators, $M_{A,B,\gamma} : H \rightarrow H$ is the operator defined by

$$M_{A,B,\gamma}(p) = \frac{1}{\gamma}(Id + \gamma A)^{-1}(p - \gamma B(p)).$$

It comes naturally by discretizing the continuous dynamics (5) explicitly with a time step $h_n > 0$. Set $t_n = \sum_{i=1}^n h_i$, $\iota_n = \iota(t_n)$, $\gamma_n = \gamma(t_n)$ and $p_n = p(t_n)$. An explicit finite-difference scheme for (5) with centered second-order variation gives that

$$\frac{1}{h_n^2}(p_{n+1} - 2p_n + p_{n-1}) + \frac{\iota_n}{h_n}(p_n - p_{n-1}) + M_{A,B,\gamma_n}(w_n) = 0, \tag{6}$$

where w_n is a point in the line passing through p_{n-1} and p_n (we have some flexibility for the choice of w_n). The above equality (6) can be rewritten as

$$p_{n+1} = \left(1 - \frac{h_n^2}{\gamma_n}\right) (p_n + (1 - \iota_n h_n)(p_n - p_{n-1})) + \frac{h_n^2}{\gamma_n}(I + \gamma_n A)^{-1}(y_n - \gamma_n B(w_n)). \tag{7}$$

Set $\gamma_n = h_n^2$ and $\alpha_n = 1 - \iota_n h_n$ in (7). With the aid of the classical Nesterov extrapolation choice for w_n , continuous dynamics (5) leads to a special case of the algorithm as

$$\begin{cases} w_n = p_n + \alpha_n(p_n - p_{n-1}), \\ p_{n+1} = (I + \gamma_n A)^{-1}(w_n - \gamma_n B(w_n)), \end{cases} \tag{8}$$

where γ_n is a step-size parameter and α_n is an extrapolation factor, the extrapolation term $\alpha_n(p_n - p_{n-1})$ is intended to speed up the rate of convergence. In so doing, this dynamical approach leads to a special case of the forward–backward algorithm of inertial type.

Consider the following monotone variational inequality problem (VIP, in short), which consists of finding $z \in \Omega$ such that

$$\langle S(z), x - z \rangle \geq 0, \quad \forall x \in \Omega, \tag{9}$$

where $S : \Omega \rightarrow H$ is a monotone operator. We denote the solution set of the variational inequality problem by $VI(\Omega, S)$.

Remark 1. It is known that x solves the $VI(\Omega, S)$ iff x is an equilibrium point of the following dynamical system, i.e.,

$$x = Proj_{\Omega}(x - \mu Sx), \quad \mu > 0,$$

where $Proj_{\Omega}$ is the nearest point (or metric) projection from H onto Ω .

Variational inequality problems, which serve as a powerful mathematical model, unify several important concepts in applied mathematics, such as systems of nonlinear equations, complementarity problems, and equilibrium problems under a general and unified framework [18–20]. Recently, spotlight has been shed on developing efficient and implementable iterative schemes for solving monotone variational inequality problems, see [21–24]. A significant body of the work on iteration methods for VIPs has accumulated in the literature recently. In 2001, Yamada [25] investigated a so-called hybrid steepest descent method. Given the current iterate p_n , calculate the next iterate p_{n+1} via

$$p_{n+1} = (I - \alpha \chi_n S)T p_n, \quad \forall n \geq 0,$$

where the operator $T : \Omega \rightarrow H$ is nonexpansive, the operator S is ι -Lipschitz continuous for some $\iota \geq 0$ and κ -strongly monotone (i.e., $\exists \kappa > 0$ such that $\langle Sx - Sy, x - y \rangle \geq \rho \|x - y\|^2, \forall x, y \in H$), while the parameters $\alpha \in (0, 2\kappa/\iota^2)$, $(\chi_n)_{n \geq 0} \in (0, 1]$, $\lim_{n \rightarrow \infty} \chi_n = 0$ and $\sum_{n=1}^{\infty} \chi_n = \infty$. In this paper, the set Ω can be regarded as the solution set of the inclusion problem. In continuous case, Attouch and Mainé [26] developed a second-order autonomous system with a Tikhonov-like regularizing term ∇F , which reads

$$\begin{cases} \ddot{p}(t) + \alpha \dot{p}(t) + \nabla \phi(p(t)) + B(p(t)) + \beta(t) \nabla F(p(t)) = 0, \\ p(0) = p_0, \dot{p}(0) = q_0, \end{cases} \tag{10}$$

where $\alpha > 0$ and $(p_0, q_0) \in H^2$ is an arbitrarily given initial data. With the assumptions: (i) $F : H \rightarrow \mathbb{R}$ is a convex, differentiable operator with ∇F strongly monotone and Lipschitz continuous; (ii) $\phi : H \rightarrow \mathbb{R}$ is a convex, differentiable operator with $\nabla \phi : H \rightarrow H$ Lipschitz continuous; (iii) $B : H \rightarrow H$ is a maximally monotone and cocoercive operator; (iv) $\beta : [0, +\infty) \rightarrow (0, +\infty)$ is a positive and decreasing function of class C^1 such that $\int_0^{+\infty} \beta(t) dt = +\infty$ and $\beta(t) \rightarrow 0$ as $t \rightarrow +\infty$, with $\dot{\beta}$ Lipschitz continuous and bounded, they proved that each trajectory of (10) strongly converges to p^* as $t \rightarrow \infty$, which solves the following variational inequality problem:

$$\text{find } p^* \in (B + \nabla \phi)^{-1}(0) \text{ such that } \langle \nabla F(p^*), q - p^* \rangle \geq 0, \quad q \in (B + \nabla \phi)^{-1}(0).$$

In the spirit of the splitting forward–backward method and the hybrid steepest descent method, we present an iterative scheme as a new strategy, parallel to that of the autonomous system (10). We analyze the convergence with the underlying operator B cocoercive and the extension from condition (ii) to the general maximally monotone case, which is considered in Section 3. From this perspective, our study is the natural extension of the convergence results obtained by Attouch and Mainé [26] in the case of continuous dynamical systems.

2. Preliminaries

Lemma 1 ([25]). *Suppose that $W : H \rightarrow \mathbb{R}$ is Fréchet differentiable with $\nabla W : H \rightarrow H$ being κ -strongly monotone and ι -Lipschitz continuous. Define $S := Id - \chi \alpha \nabla W$, where $\chi \in [0, 1]$ and $\alpha \in (0, 2\kappa/\iota^2)$. Then $\|S(x) - S(y)\| \leq (1 - \chi \vartheta) \|x - y\| (x, y \in H)$, where $\vartheta := 1 - \sqrt{1 - \alpha(2\kappa - \alpha \iota^2)} \in (0, 1]$.*

Lemma 2 ([27]). *If $B : H \rightarrow H$ is a ρ -cocoercive operator, then*

- (i) *B is a $\frac{1}{\rho}$ -Lipschitz continuous and monotone operator;*
- (ii) *if v is any constant in $(0, 2\rho]$, then $Id - vB$ is nonexpansive, where Id stands for the identity operator on H .*

Lemma 3. *Let H be a Hilbert space, $A : H \rightarrow 2^H$ be a maximally monotone operator and $B : E \rightarrow E$ be a κ -cocoercive on H . Then*

(i)

$$\|(Id + sA)^{-1}(Id - sB)x - x\| \leq 2\|(Id + tA)^{-1}(Id - tB)x - x\|, \quad \forall x \in H,$$

where t and s are any positive real numbers with $t \geq s$.

(ii) $(A + B)^{-1}(0) = \text{Fix}((Id + \lambda A)^{-1}(Id - \lambda B)), \forall \lambda > 0$.

Proof. Since A is maximally monotone, one has

$$\left\langle \frac{x - (Id + sA)^{-1}(Id - sB)x}{s} - \frac{x - (Id + tA)^{-1}(Id - tB)x}{t} \right\rangle \geq 0.$$

□

Lemma 4 ([28]). Let S be a nonexpansive mapping defined on a closed convex subset C of a Hilbert space H . Then $Id - S$ is demi-closed, i.e., whenever $(x_n)_{n \geq 0}$ is a sequence in C weakly converging to some $x \in C$ and the sequence $((Id - S)x_n)_{n \geq 0}$ strongly converges to some $y \in H$, it follows that $y = (Id - S)x$.

Lemma 5 ([29]). Let $(\mu_n)_{n \geq 0}$ be a sequence of nonnegative real numbers such that

$$\mu_{n+1} \leq (1 - \lambda_n)\mu_n + \lambda_n v_n + \gamma_n,$$

where $(\lambda_n)_{n \geq 0} \subset (0, 1)$ and $(v_n)_{n \geq 0}, (\gamma_n)_{n \geq 0}$ satisfy the following conditions (i) $\lim_{n \rightarrow \infty} \lambda_n = 0, \sum_{n=0}^{\infty} \lambda_n = \infty$; (ii) $\limsup_{n \rightarrow \infty} v_n \leq 0$; (iii) $\gamma_n \geq 0, \sum_{n=0}^{\infty} \gamma_n < \infty$. Then $\lim_{n \rightarrow \infty} \mu_n = 0$.

Lemma 6 ([30]). Let $(\zeta_n)_{n \geq 0}$ be a sequence of nonnegative real numbers such that there exists a subsequence (ζ_{n_i}) of $(\zeta_n)_{n \geq 0}$ such that $\zeta_{n_{i+1}} > \zeta_{n_i}$ for all $i \in \mathbb{N}$. Then there exists a nondecreasing sequence (m_j) of \mathbb{N} such that $\lim_{j \rightarrow \infty} m_j = \infty$ and the following properties are satisfied by all (sufficiently large) number of $j \in \mathbb{N}$, $\zeta_{m_j+1} \geq \zeta_{m_j}, \zeta_{m_j+1} \geq \zeta_k$. In fact, m_j is the largest number n in the set $\{1, 2, \dots, j\}$ such that $\zeta_{n+1} \geq \zeta_n$.

3. Main Results

Throughout this section, we make the following standing assumptions

- Assumption 1.** (a) Id stands for the identity operator. The operator $S : H \rightarrow H$ is supposed to be κ -strongly monotone and ι -Lipschitz continuous for some $\kappa, \iota > 0$;
 (b) Let $A : H \rightarrow 2^H$ be a maximally monotone operator, and $B : H \rightarrow H$ be a ρ -cocoercive operator for some $\rho > 0$, with the solution set $\Omega = (A + B)^{-1}(0) \neq \emptyset$.

Our Algorithm 1 is formally designed as follows.

Algorithm 1: The hybrid forward–backward algorithm

Input: Input the algorithm parameters $(\eta_i)_{i \geq 0}, (v_i)_{i \geq 0}$ and δ ;
Output: Output \hat{p} ;
 1 Initialize the data $p_0, p_1 \in H$;
 2 Set $n \leftarrow 1$;
 3 **while** not converged **do**
 4 **if** $\|p_n - p_{n-1}\| \neq 0$ **then**
 5 Update μ_n such that $\mu_n = o\left(\frac{\eta_n}{\|p_n - p_{n-1}\|}\right)$;
 6 **end**
 7 **else**
 8 Choose μ_n as any positive number;
 9 **end**
 10 Update $w_n = p_n + \mu_n(p_n - p_{n-1})$;
 11 Update $q_n = (Id + v_n A)^{-1}(Id - v_n B)w_n$;
 12 Update $p_{n+1} = (Id - \delta \eta_n S)q_n$;
 13 Set $n \leftarrow n + 1$;
 14 **end**
 15 **return** $\hat{p} = p_n$

We make the following assumption with respect to the algorithm parameters.

- Assumption 2.** (C₁) $0 < \liminf_{n \rightarrow \infty} v_n \leq \limsup_{n \rightarrow \infty} v_n < 2\rho$; (C₂) $(\eta_n)_{n \geq 0} \subset (0, 1), \lim_{n \rightarrow \infty} \eta_n = 0, \eta_{n+1} = O(\eta_n), \sum_{n=0}^{\infty} \eta_n = \infty$; (C₃) $\lim_{n \rightarrow \infty} \frac{\mu_n}{\eta_n} \|p_n - p_{n-1}\| = 0$; (C₄) $\delta \in (0, 2\kappa/\iota^2)$.

Remark 2. Please note that the condition (C₃) of Assumption 2 can be easily implemented since the value of $\|p_n - p_{n-1}\|$ is known before choosing μ_n . Indeed, the parameter μ_n can be chosen such that

$$\mu_n = \begin{cases} \omega, & \text{if } p_n = p_{n-1}; \\ \mu_n = \frac{\tau_n}{\|p_n - p_{n-1}\|}, & \text{if otherwise,} \end{cases}$$

where $0 \leq \omega$ and $(\tau_n)_{n \geq 0}$ is a positive sequence such that $\tau_n = o(\eta_n)$.

Now we are in a position to state and prove the main result of this section.

Theorem 1. Suppose that Assumptions 1, 2 hold. Then for any initial data $p_0, p_1 \in H$, the weak sequential cluster point of sequences $(p_n)_{n \geq 0}, (w_n)_{n \geq 0}$ and $(q_n)_{n \geq 0}$ generated by Algorithm 1 belongs to the solution set of Problem 3. In addition, the three sequences converge strongly to the unique element of $VI(\Omega, S)$.

Proof. First, we show that $I - \delta R$ is a contraction operator. According to Lemma 1, we have

$$\begin{aligned} \|(Id - \delta S)x - (Id - \delta S)y\|^2 &= \|x - y\|^2 + \delta^2 \|Sx - Sy\|^2 - 2\delta \langle x - y, Sx - Sy \rangle \\ &\leq \|x - y\|^2 + \delta^2 l^2 \|x - y\|^2 - 2\delta \kappa \|x - y\|^2 \\ &= (1 - \alpha)^2 \|x - y\|^2, \end{aligned} \tag{11}$$

where $\alpha = \frac{1}{2} \delta (2\kappa - \delta l^2)$. Thus, we find that $Id - \delta S$ is a contraction operator with the constant $1 - \alpha$. In light of the nonexpansivity of $Proj_\Omega$, we further obtain that $Proj_\Omega (Id - \delta S)$ is a contraction operator. From the Banach contraction principle, there exists a unique point $a \in \Lambda$ such that $a = Proj_\Omega (Id - \delta S)a$.

Now, it remains to prove that $(p_n)_{n \geq 0}$ is bounded. Since $B : H \rightarrow H$ is ρ -cocoercive and $(Id + v_n A)^{-1}$ is firmly nonexpansive, one concludes from Lemma 2 that $Id - v_n B$ is nonexpansive. Let $z \in \Lambda$ be arbitrarily chosen. Invoking Assumption 2 (C₁), one infers that

$$\begin{aligned} \|q_n - z\|^2 &= \|(Id + v_n A)^{-1} (Id - v_n B)w_n - (Id + v_n A)^{-1} (Id - v_n B)z\|^2 \\ &\leq \|w_n - z + v_n Bz - v_n Bw_n\|^2 \\ &\leq \|w_n - z\|^2 - v_n (2\rho - v_n) \|Bz - Bw_n\|^2 \\ &\leq \|w_n - z\|^2, \end{aligned} \tag{12}$$

which yields

$$v_n (2\rho - v_n) \|Bz - Bw_n\|^2 \leq \|w_n - z\|^2 - \|q_n - z\|^2, \tag{13}$$

and

$$\|q_n - z\|^2 \leq \|w_n - z\|^2. \tag{14}$$

From the definition of w_n , one reaches

$$\|w_n - z\| \leq \|p_n - z\| + \mu_n \|p_n - p_{n-1}\|. \tag{15}$$

Invoking Assumption 2 (C₃), there exists a positive constant $M_1 \in \mathbb{R}$ such that

$$\frac{\mu_n}{\eta_n} \|p_n - p_{n-1}\| \leq M_1, \quad \forall n \geq 0. \tag{16}$$

From the definition of p_n , which together with (14)–(16), one obtains

$$\begin{aligned}
 \|p_{n+1} - z\| &= \|(Id - \delta\eta_n S)q_n - (Id - \delta\eta_n S)z - \delta\eta_n Sz\| \\
 &\leq (1 - \xi\eta_n)\|w_n - z\| + \delta\eta_n\|Sz\| \\
 &\leq (1 - \xi\eta_n)(\|p_n - z\| + \eta_n M_1) + \delta\eta_n\|Sz\| \\
 &\leq \max\left\{\|p_n - z\|, \frac{1}{\xi}(M_1 + \delta\|Sz\|)\right\} \\
 &\dots \\
 &\leq \max\left\{\|p_1 - z\|, \frac{1}{\xi}(M_1 + \delta\|Sz\|)\right\},
 \end{aligned}
 \tag{17}$$

where $\xi = 1 - \sqrt{1 - \delta(2\kappa - \delta\iota^2)} \in (0, 1]$, due to Assumption 2 (C_4). This implies that sequence $(p_n)_{n \geq 0}$ is bounded. At the same time, by putting together (14) and (15), one concludes that $(q_n)_{n \geq 0}$ and $(w_n)_{n \geq 0}$ are bounded. Once again, using the definition of w_n , one concludes that

$$\begin{aligned}
 \|w_n - z\|^2 &\leq \|p_n - z\|^2 + \mu_n\|p_n - p_{n-1}\|(\mu_n\|p_n - p_{n-1}\| + 2\|p_n - z\|) \\
 &\leq \|p_n - z\|^2 + \frac{\mu_n}{\eta_n}\|p_n - p_{n-1}\|M_2,
 \end{aligned}
 \tag{18}$$

where

$$M_2 = \sup_{n \geq 0} \eta_n(\mu_n\|p_n - p_{n-1}\| + 2\|p_n - z\|) < \infty.
 \tag{19}$$

By combining (12) with (18), one immediately concludes that

$$\|q_n - z\|^2 \leq \|p_n - z\|^2 + \frac{\mu_n}{\eta_n}\|p_n - p_{n-1}\|M_2 - \nu_n(2\kappa - \nu_n)\|Bz - Bw_n\|^2.
 \tag{20}$$

Invoking (20), which together with the definition of p_n , one deduces that

$$\begin{aligned}
 \|p_{n+1} - z\|^2 &= (Id - \xi\eta_n)\|q_n - z\|^2 - 2\delta\eta_n\langle p_{n+1} - z, Sz \rangle \\
 &\leq \|p_n - z\|^2 + \frac{\mu_n}{\eta_n}\|p_n - p_{n-1}\|M_2 - \nu_n(2\kappa - \nu_n)\|Bz - Bw_n\|^2 + \eta_n M_3,
 \end{aligned}
 \tag{21}$$

where $M_3 = \sup_{n \geq 0} 2\delta\langle z - p_{n+1}, Sz \rangle < \infty$, due to the boundedness of $(p_n)_{n \geq 0}$. Let us rewrite (21) as

$$\nu_n(2\kappa - \nu_n)\|Bz - Bw_n\|^2 \leq \|p_n - z\|^2 - \|p_{n+1} - z\|^2 + \frac{\mu_n}{\eta_n}\|p_n - p_{n-1}\|M_2 + \eta_n M_3.
 \tag{22}$$

By using the firmly nonexpansive property of $(Id + \nu_n A)^{-1}$, one arrives at

$$\begin{aligned}
 \|q_n - z\|^2 &\leq \langle q_n - z, (Id - \nu_n B)w_n - (Id - \nu_n B)z \rangle \\
 &= \frac{1}{2}\|q_n - z\|^2 + \frac{1}{2}\|(Id - \nu_n B)w_n - (Id - \nu_n B)z\|^2 \\
 &\quad - \frac{1}{2}\|(q_n - z) - (Id - \nu_n B)w_n + (Id - \nu_n B)z\|^2 \\
 &\leq \frac{1}{2}\|q_n - z\|^2 + \frac{1}{2}\|w_n - z\|^2 - \frac{1}{2}\|q_n - w_n\|^2 - \frac{\nu_n^2}{2}\|Bw_n - Bz\|^2 \\
 &\quad + \nu_n\|q_n - w_n\|\|Bw_n - Bz\|,
 \end{aligned}$$

which can be equivalently rewritten as

$$\|q_n - z\|^2 \leq \|w_n - z\|^2 - \|q_n - w_n\|^2 - \nu_n^2\|Bw_n - Bz\|^2 + 2\nu_n\|q_n - w_n\|\|Bw_n - Bz\|.
 \tag{23}$$

Returning to (14), (18) and (23), one concludes

$$\begin{aligned} \|p_{n+1} - z\|^2 &= \|\eta_n(Id - \delta S)q_n + (1 - \eta_n)q_n - z\|^2 \\ &= \eta_n\|(Id - \delta S)q_n - z\|^2 + (1 - \eta_n)\|q_n - z\|^2 - \eta_n(1 - \eta_n)\|\delta S q_n\|^2 \\ &\leq \eta_n\|(Id - \delta S)q_n - z\|^2 + \|q_n - z\|^2 \\ &\leq \eta_n\|(Id - \delta S)q_n - z\|^2 + \|p_n - z\|^2 + \frac{\mu_n}{\eta_n}\|p_n - p_{n-1}\|M_2 \\ &\quad - \|q_n - w_n\|^2 - \frac{\nu_n^2}{2}\|Bw_n - Bz\|^2 + 2\nu_n\|q_n - w_n\|\|Bw_n - Bz\|, \end{aligned}$$

that is,

$$\begin{aligned} \|q_n - w_n\|^2 &\leq \eta_n\|(Id - \delta S)q_n - z\|^2 + \frac{\mu_n}{\eta_n}\|p_n - p_{n-1}\|M_2 \\ &\quad + 2\nu_n\|q_n - w_n\|\|Bw_n - Bz\| + \|p_n - z\|^2 - \|p_{n+1} - z\|^2. \end{aligned} \tag{24}$$

Next, we show that $(\|p_n - z\|^2)_{n \geq 0}$ converges to zero by considering two possible cases on the sequence $(\|p_n - z\|^2)_{n \geq 0}$.

Case 1. There exists $N \in \mathbb{N}$ such that $\|p_{n+1} - z\|^2 \leq \|p_n - z\|^2, \forall n \geq N$. Recalling that $\|p_n - z\|^2$ is lower bounded, one deduces that $\lim_{n \rightarrow \infty} \|p_n - z\|^2$ exists. Using Assumption 2 (C₂), and letting n tend to infinity in (22), one finds that

$$\lim_{n \rightarrow \infty} \nu_n \|Bz - Bw_n\| = 0. \tag{25}$$

Taking account of Assumption 2 (C₂), (C₃), (25), and letting n tend to infinity in (24), one concludes that

$$\lim_{n \rightarrow \infty} \|q_n - w_n\|^2 = 0. \tag{26}$$

It follows from Assumption 2 (C₂), (C₃) the definitions of w_n and p_{n+1} that

$$\lim_{n \rightarrow \infty} \|w_n - p_n\| = \lim_{n \rightarrow \infty} \mu_n \|p_n - p_{n-1}\| = 0, \tag{27}$$

and

$$\lim_{n \rightarrow \infty} \|p_{n+1} - q_n\| = \lim_{n \rightarrow \infty} \delta \eta_n \|Rq_n\| = 0. \tag{28}$$

Resorting to (26)–(28), one finds that

$$\lim_{n \rightarrow \infty} \|p_{n+1} - p_n\| \leq \lim_{n \rightarrow \infty} (\|p_{n+1} - q_n\| + \|q_n - w_n\| + \|w_n - p_n\|) = 0. \tag{29}$$

Denote $G_{\nu_n} = (Id + \nu_n A)^{-1}(Id - \nu_n B)$. It is then immediately that

$$\lim_{n \rightarrow \infty} \|G_{\nu_n} w_n - w_n\| = \lim_{n \rightarrow \infty} \|q_n - w_n\| = 0. \tag{30}$$

For any $\gamma \geq 0$ such that $\gamma \leq \nu_n, \forall n \geq 0$, it implies from Lemma 3 that

$$\lim_{n \rightarrow \infty} \|G_\gamma w_n - w_n\| \leq 2 \lim_{n \rightarrow \infty} \|G_{\nu_n} w_n - w_n\| = 0. \tag{31}$$

Since $(p_n)_{n \geq 0}$ is a bounded sequence, there exists a subsequence $(p_{n_i})_{n_i \geq 0}$ of $(p_n)_{n \geq 0}$ and a weak sequential cluster point $s \in H$ such that $p_{n_i} \rightharpoonup s$ as $i \rightarrow \infty$. Due to (26) and (27), one finds that $w_{n_i} \rightharpoonup s$ and $q_{n_i} \rightharpoonup s$ as $i \rightarrow \infty$. Thus,

$$\limsup_{n \rightarrow \infty} \langle Sz, z - w_n \rangle = \lim_{i \rightarrow \infty} \langle Sz, z - w_{n_i} \rangle. \tag{32}$$

From the demiclosedness of $I - G_\gamma$ and Lemma 4, one obtains that $s \in \text{Fix}(G_\gamma)$. In view of $\text{Fix}(G_\gamma) = (A + B)^{-1}(0)$, one concludes $s \in \Lambda$. From the fact that $a = \text{Proj}_\Lambda(I - \delta S)a$, one sees $\langle a - (I - \delta S)a, s - a \rangle \geq 0$. As a straightforward consequence, one finds

$$\langle Sa, a - s \rangle \leq 0. \tag{33}$$

Coming back to (29), (32) and (33), one has

$$\limsup_{n \rightarrow \infty} \langle Sa, a - p_{n+1} \rangle = \limsup_{n \rightarrow \infty} \langle Sa, a - p_n \rangle = \lim_{i \rightarrow \infty} \langle Sa, a - p_{n_i} \rangle = \langle Sa, a - s \rangle \leq 0. \tag{34}$$

Returning to (11) and owing to the definitions of p_n and q_n , one finds that

$$\begin{aligned} \|p_{n+1} - a\|^2 &= \|\eta_n(Id - \delta S)q_n + (1 - \eta_n)q_n - a\|^2 \\ &\leq (1 - \eta_n)^2 \|q_n - a\|^2 + 2\eta_n \langle (Id - \delta S)q_n - a, p_{n+1} - a \rangle \\ &= (1 - \eta_n)^2 \|q_n - a\|^2 + 2\eta_n \langle (Id - \delta S)q_n - (Id - \delta S)a + (Id - \delta S)a - a, p_{n+1} - a \rangle \\ &\leq (1 - \eta_n)^2 \|q_n - a\|^2 + 2\eta_n(1 - \alpha) \|q_n - a\| \|p_{n+1} - a\| + 2\delta\eta_n \langle Sa, a - p_{n+1} \rangle \\ &\leq (1 - \eta_n)^2 (\|p_n - a\| + \mu_n \|p_n - p_{n-1}\|)^2 \\ &\quad + 2\eta_n(1 - \alpha) (\|p_n - a\|^2 + \mu_n \|p_n - p_{n-1}\| \|p_n - a\|) + 2\delta\eta_n \langle Sa, a - p_{n+1} \rangle \\ &\leq (1 - 2\eta_n\alpha) (\|p_n - a\|^2 + 2\mu_n \|p_n - p_{n-1}\| \|p_n - a\|) \\ &\quad + \mu_n^2 \|p_n - p_{n-1}\|^2 + 2\delta\eta_n \langle Sa, a - p_{n+1} \rangle, \end{aligned} \tag{35}$$

where $\alpha = \frac{1}{2}\delta(2\kappa - \delta t^2) \in (0, 1)$. Let

$$\Gamma_n = \|p_n - a\|^2 + 2\mu_n \|p_n - p_{n-1}\| \|p_n - a\|$$

and

$$M_4 = \sup_{n \geq 0} \frac{\eta_{n+1}}{\eta_n\alpha} \|p_{n+1} - a\|.$$

Furthermore, due to Assumption 2 (C₂) and the boundedness of $(p_n)_{n \geq 0}$, we find that $M_4 < \infty$. In so doing, it asserts that

$$\Gamma_{n+1} \leq (1 - 2\eta_n\alpha)\Gamma_n + 2\eta_n\alpha \left(\frac{\delta}{\alpha} \langle Sa, p_{n+1} - a \rangle + \left(\frac{\mu_{n+1}}{\eta_{n+1}} \|p_{n+1} - p_n\| \right) M_4 + \frac{\eta_n}{2\alpha} \left(\frac{\mu_n}{\eta_n} \|p_n - p_{n-1}\| \right)^2 \right).$$

Invoking (32), we infer that

$$\limsup_{n \rightarrow \infty} \left(\frac{\delta}{\alpha} \langle Sa, p_{n+1} - a \rangle + \left(\frac{\mu_{n+1}}{\eta_{n+1}} \|p_{n+1} - p_n\| \right) M_4 + \frac{\eta_n}{2\alpha} \left(\frac{\mu_n}{\eta_n} \|p_n - p_{n-1}\| \right)^2 \right) \leq 0.$$

According to Assumption 2 (C₂), one finds that $2\eta_n\alpha \in (0, 1)$ and $\lim_{n \rightarrow \infty} 2\eta_n\alpha = 0$. So doing, Lemma 5 asserts that $\lim_{n \rightarrow \infty} \Gamma_n = 0$, which further implies that $\lim_{n \rightarrow \infty} \|p_n - a\| = 0$. Coming back to (26) and (27), one concludes that $(p_n)_{n \geq 0}$, $(q_n)_{n \geq 0}$, $(w_n)_{n \geq 0}$ converge strongly to a .

Case 2. There exists a subsequence $(\|p_{n_j} - z\|^2)_{j \geq 0}$ of $(\|p_n - z\|^2)_{n \geq 0}$ such that

$$\|p_{n_j} - z\|^2 < \|p_{n_{j+1}} - z\|^2, \quad \forall j \in \mathbb{N}.$$

In this case, it is obvious that there exists a nondecreasing sequence $(m_k)_{k \geq 0} \in \mathbb{N}$ such that $m_k \rightarrow \infty$ as $k \rightarrow \infty$. Lemma 6 asserts that the following inequalities hold,

$$\|p_{m_k} - z\|^2 \leq \|p_{m_{k+1}} - z\|^2, \quad \|p_k - z\|^2 \leq \|p_{m_k} - z\|^2, \quad \forall k \in \mathbb{N}. \tag{36}$$

Thanks to (21) and (36), one finds

$$\nu_n(2\kappa - \nu_n)\|Bz - Bw_n\|^2 \leq \mu_n M_1 \|p_n - p_{n-1}\| + M_2 \eta_n.$$

By letting n tend to infinity in the above inequality, one infers that $\lim_{n \rightarrow \infty} \nu_n \|Bz - Bw_n\| = 0$. According to (24) and (36), one finds that

$$\|q_n - w_n\|^2 \leq \eta_n \|(Id - \delta S)q_n - z\|^2 + \mu_n M_1 \|p_n - p_{n-1}\| + 2\nu_n \|q_n - w_n\| \|Bw_n - Bz\|.$$

Accordingly, one finds that $\lim_{n \rightarrow \infty} \|q_n - w_n\| = 0$. A calculation similar to the proof in Case 1. guarantees that

$$\lim_{n \rightarrow \infty} \|p_{m_k+1} - p_{m_k}\| = 0,$$

and

$$\limsup_{k \rightarrow \infty} \langle Sa, a - p_{m_k+1} \rangle \leq 0. \tag{37}$$

Invoking (35), one finds

$$\begin{aligned} \|p_{m_k+1} - a\|^2 &\leq (1 - \eta_{m_k})^2 (\|p_{m_k} - a\| + \mu_{m_k} \|p_{m_k} - p_{m_k-1}\|)^2 + 2\eta_{m_k} (1 - \alpha) (\|p_{m_k} - a\| \\ &\quad + \mu_{m_k} \|p_{m_k} - p_{m_k-1}\|) \|p_{m_k+1} - a\| + 2\delta\eta_{m_k} \langle Sa, a - p_{m_k+1} \rangle \\ &\leq (1 - \eta_{m_k})^2 (\|p_{m_k+1} - a\| + \mu_{m_k} \|p_{m_k} - p_{m_k-1}\|)^2 + 2\eta_{m_k} (1 - \alpha) (\|p_{m_k+1} - a\|^2 \\ &\quad + \mu_{m_k} \|p_{m_k} - p_{m_k-1}\| \|p_{m_k+1} - a\|) + 2\delta\eta_{m_k} \langle Sa, a - p_{m_k+1} \rangle \\ &\leq (1 - 2\eta_{m_k} \alpha) (\|p_{m_k+1} - a\|^2 + 2\mu_{m_k} \|p_{m_k} - p_{m_k-1}\| \|p_{m_k+1} - a\|) \\ &\quad + \mu_{m_k}^2 \|p_{m_k} - p_{m_k-1}\|^2 + 2\delta\eta_{m_k} \langle Sa, a - p_{m_k+1} \rangle \\ &\leq (1 - 2\eta_{m_k} \alpha) \|p_{m_k+1} - a\|^2 + 2\mu_{m_k} \|p_{m_k} - p_{m_k-1}\| \|p_{m_k+1} - a\| \\ &\quad + \mu_{m_k}^2 \|p_{m_k} - p_{m_k-1}\|^2 + 2\delta\eta_{m_k} \langle Sa, a - p_{m_k+1} \rangle, \end{aligned} \tag{38}$$

which sends us to

$$\|p_{m_k+1} - a\|^2 \leq \frac{\mu_{m_k}}{\eta_{m_k}} \|p_{m_k} - p_{m_k-1}\| M_5 + \left(\frac{\mu_{m_k}}{\eta_{m_k}} \|p_{m_k} - p_{m_k-1}\| \right)^2 M_6 + \langle Sa, a - p_{m_k+1} \rangle M_7, \tag{39}$$

where $M_5 = \sup_{n \geq 0} \frac{\|p_{n+1} - a\|}{\alpha} < \infty$, $M_6 = \sup_{n \geq 0} \frac{\eta_n}{2\alpha} < \infty$ and $M_7 = \frac{\delta}{\alpha}$. By applying Assumption 2 (C₃), which together with (39), we additionally derive that

$$\limsup_{k \rightarrow \infty} \|p_{m_k+1} - a\|^2 \leq 0. \tag{40}$$

From $\|p_k - a\| \leq \|p_{m_k+1} - a\|$, we obtain that $\limsup_{k \rightarrow \infty} \|p_k - a\| = 0$. This further implies that $(p_k)_{k \geq 0}$ converges strongly to a . Furthermore, one has

$$\lim_{n \rightarrow \infty} \|q_n - w_n\|^2 = 0, \quad \lim_{n \rightarrow \infty} \|w_n - p_n\| = 0. \tag{41}$$

Recalling that the sequence $(p_k)_{k \geq 0}$ converges strongly to a , which together with (41), one finds that the sequences $(q_n)_{n \geq 0}$ and $(w_n)_{n \geq 0}$ also converge strongly to a . From the above, one can conclude that the sequences generated by Algorithm 1 converge strongly to the unique solution $a \in \Omega$ such that $a = Proj_{\Omega}(Id - \delta S)a$. In view of Remark 1, one sees that $a \in VI(\Omega, S)$. Furthermore, the solution of such variational inequality is unique due to the properties of S . This completes the proof. \square

If $(\mu_n)_{n \in \mathbb{N}} := 0$, we obtain the following Algorithm 2 without the inertial extrapolation as a special case of Algorithm 1.

Algorithm 2: The hybrid forward–backward algorithm without the inertial term

Input: Input the algorithm parameters $(\nu_i)_{i \geq 0}, (\eta_i)_{i \geq 0}$ and δ ;
Output: Output \hat{p} ;
 1 Initialize the data $p_0 \in H$;
 2 Set $n \leftarrow 1$;
 3 **while** not converged **do**
 4 Choose μ_n as any positive number;
 5 Update $q_n = (Id + \nu_n A)^{-1}(I - \nu_n B)p_n$;
 6 Update $p_{n+1} = (Id - \delta \eta_n S)q_n$;
 7 Set $n \leftarrow n + 1$;
 8 **end**
 9 **return** $\hat{p} = p_n$

Recently, much attention has been paid to the relationship between continuous-time systems and neural network models, see [31,32]. One can views an iterative algorithm as a time discretized version of a continuous-time dynamic system. We propose a recurrent neural network to Problem (1). By taking constant parameters, we obtain the following consequence of the dynamical with respect to the time variable t

$$\begin{cases} w(t) = p(t) + \alpha \dot{p}(t), \\ u(t) = (Id + \gamma A)^{-1}(w(t) - \gamma Bw(t)), \\ \dot{p}(t) + \dot{\check{p}}(t) = (Id - \delta \eta(t)S)u(t), \\ p(0) = p_0, \dot{p}(0) = z_0. \end{cases} \tag{42}$$

Algorithm 1 is strongly connected with the damped inertial system (42). Indeed, Algorithm 1 comes naturally into play by performing an implicit discretization of the inertial system with respect to the time variable t . Without ambiguity, we omit to write the variable t to get simplified notations. In so doing, p stands for $p(t)$, and so on. By setting $S := Id$ and $k = 1 - \delta \eta$, the circuit architecture diagram of the model (42) is shown in Figure 1.

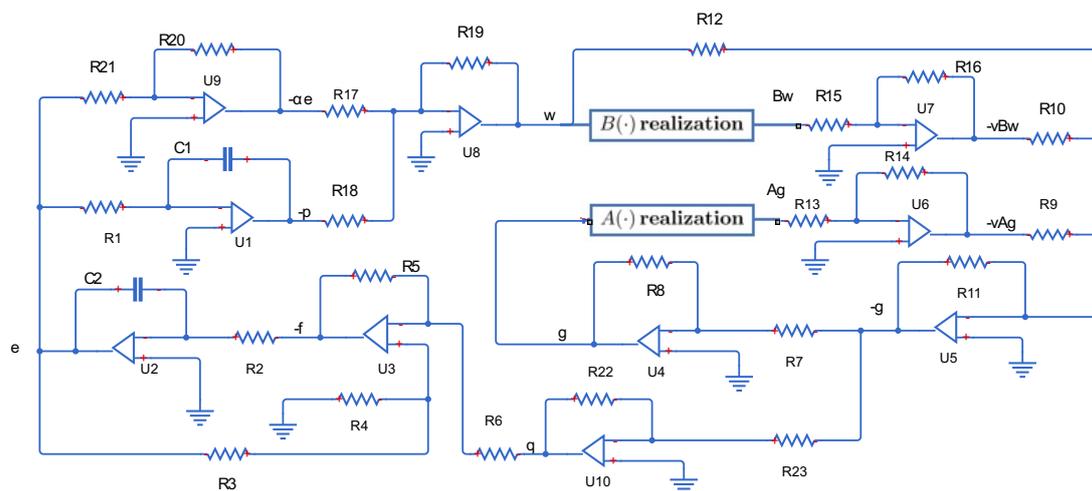


Figure 1. Circuit architecture diagram of model (42), where $e = \dot{p}$, $f = \dot{\check{p}}$, $w = p + \alpha \dot{p}$, $q = \dot{p} + \dot{\check{p}}$, $g = q/k$.

The operational amplifiers $U1, U2$ are combined with capacitors $C1, C2$, resistors $R1, R2$ to work as integrators to realize the transformations between $p, \dot{p}, \dot{\check{p}}$. The amplifier $U9$ cooperated with resistors $R20, R21$ brings about effectiveness in amplification and opposition. Hence \dot{p} is translated into $-\alpha \dot{p}$. The amplifier $U3$ and resistors $R3, R4, R5, R6$ are united as the subtractor block to achieve $-\dot{\check{p}}$. Also, in the same way, the amplifier $U8$ and resistors $R17, R18, R19$ are united as another subtractor block to achieve $p + \alpha \dot{p}$. One sees that $A((\dot{p} + \dot{\check{p}})/k)$ and $B(p + \alpha \dot{p})$

are obtained, respectively, by “ $A(\cdot)$ realization” block and “ $B(\cdot)$ realization” block. The amplifier $U6$ cooperated with resistors $R13, R14$ brings about effectiveness in amplification and opposition, therefore, $A((\dot{p} + \ddot{p})/k)$ is translated into $-vA((\dot{p} + \ddot{p})/k)$. The amplifier $U7$ cooperated with resistors $R15, R16$ brings about effectiveness in amplification and opposition, and hence $B(p + \alpha\dot{p})$ is translated into $-vB(p + \alpha\dot{p})$. The amplifier $U5$ and resistors $R9, R10, R11, R12$ are united as the subtractor block to achieve $vA((\dot{p} + \ddot{p})/k) + vB(p + \alpha\dot{p}) - (p + \alpha\dot{p})$, that is $-(\dot{p} + \ddot{p})/k$. On the other hand, $-(\dot{p} + \ddot{p})/k$ is translated into $(\dot{p} + \ddot{p})/k$ through the phase inverter composed by the operational amplifier $U4$ and resistors $R7, R8$.

4. Numerical Experiment

In this section, we consider a computational experiment to illustrate the convergence properties of the proposed method. The experiment is performed on a PC with Intel (R) Core (TM) i5-8250U CPU @1.60GHz under the MATLAB computing environment.

Digital signal reconstruction is one of the earliest and most classical problem in the file restoration, the video and image coding, the medical, the astronomical imaging and some other applications. Many problems in signal processing can be formulated as inverting the linear system, which are modeled as

$$b = Qx + v, \tag{43}$$

where $x \in R^i$ is the original signal to be reconstructed, $v \in R^j$ is the noise, $b \in R^j$ is the noisy measurement, $Q \in R^{j \times i}$ is a bounded linear observation operator, often ill conditioned because it models a process with loss of information.

In our experiment, we consider a general compressed sensing scenario, where the goal is to reconstruct an i -length sparse signal x with exactly k nonzero components from j ($k \ll j < i$) observations, i.e., the number of measurements is much larger than the sparsity level of x and at the same time smaller than the number of the signal length. Considering the storage limitation of the PC, we test a small size signal with $i = 2^{12}$ and the original signal contains $k = 180$ randomly nonzero elements. We reconstruct this signal from $j = 2^{10}$ observations. More precisely, the observation $b = Qx + v$, where $Q \in R^{j \times i}$ is the Gaussian matrix whose elements are randomly obtained from the standard normal distribution $\mathbb{N}(0, 1)$ and v is the Gaussian noise distributed as $\mathbb{N}(0, \sigma^2 I)$ with $\sigma^2 = 10^{-4}$.

A classical and significant approach to the problems of the signal processing is the regularization method, which has attracted a considerable amount of attention and revived much interest in the compressed sensing literature. We restrict our attention to the l_1 -regularized least squares model (43). Lasso framework is a particular instant of the linear problems of type (43) with the non-smooth l_1 -norm as a regularizer, in which minimizes a squared loss function, and seeks to find the sparse solutions of

$$\text{minimize}_{x \in \mathbb{R}^i} \left\{ \frac{1}{2} \|Qx - b\|_2^2 + \alpha \|x\|_1 \right\}, \tag{44}$$

where the regularization parameter $\alpha (\alpha \geq 0)$ provides a tradeoff between the noise sensitivity and the fidelity to the measurements. One knows that Problem (44) always has a solution. However, it needs not to be unique. Please note that the optimal solution tends to zero as $\alpha \rightarrow \infty$. As $\alpha \rightarrow 0$, the limiting point has the minimum l_1 norm among all points that satisfy $Q^T(Qx - b) = 0$, i.e., $x = \arg \min_{Q^T(Qx - b) = 0} \|x\|_1$. Since the objective function of Problem (44) is convex but not differentiable, one uses a first-order optimality condition based on the subdifferential calculus. For $m = 1, 2, \dots$, one can obtain the necessary and sufficient condition for the optimal solution as follows

$$(Q^T(Qx - b))_m \in \begin{cases} \{-\alpha\}, & x_m > 0; \\ \{\alpha\}, & x_m < 0; \\ [-\alpha, \alpha], & x_m = 0. \end{cases}$$

From above, one sees that the optimal solution of Problem (44) is 0, for $(Q^T b)_m \in [-\alpha, \alpha]$, $(m = 1, 2, \dots)$, i.e., $\alpha \geq \|Q^T b\|_\infty$. Thus, one can now derive the formula $\alpha_{max} = \|Q^T b\|_\infty$. For l_1 -regularized least squares; however, the convergence occurs for a finite value of α ($\alpha \geq \alpha_{max}$). To avoid the optimal sparse solution is a zero vector, in this experiment, the regularization parameter was denoted by $\alpha := 0.01\alpha_{max}$, where the value of α_{max} is computed by the above formula.

The proximal methods give a better modeling of the sparsity structure in the dataset. The major step in proximal methods is to find a solution of $\arg \min_x (g(x) + \frac{1}{2\alpha} \|x - v\|_2^2)$ with respect to the function g and the parameter $\alpha > \|Q^T b\|_\infty$. On the other hand, the proximity operator is defined as the resolvent operator of the subgradient $prox_{\alpha g} = (I + \alpha \partial g)^{-1}$. Furthermore, the proximity operator for l_1 -norm is described as the shrinkage operator, which is defined as $prox_{\alpha \|\cdot\|_1}(x)_m = (|\cdot| - \alpha)_+ sgn((x)_m)$. Now one considers the deblurring Problem (43) via the definition of the iterative shrinkage algorithm. Resorting to $g(x) = \alpha \|x\|_1$ and $f(x) = \frac{1}{2} \|Qx - b\|_2^2$, one can see that Problem (44) is a special instance of the problem

$$\text{find } x \in H \text{ such that } 0 \in (\partial g + \nabla f)x.$$

Under the class of regularized loss minimization problems, the function g is considered to be a non-smooth regularization function and the function f is viewed as a smooth loss function with gradient being $\|Q^T Q\|$ -cocoercive. Furthermore,

$$\begin{cases} prox_{v_n g} x = sgn((x)_m) \max\{|(x)_m| - \alpha v_n, 0\}, \\ (Id - v_n \nabla f)x = x - v_n Q^T(Qx - b), \quad \forall v_n > 0. \end{cases}$$

One solves the deblurring problem by using the definition of the iterative shrinkage algorithm. Setting $A := \partial g$, $B := \nabla f$ and $S := Id$ in Algorithm 1, one can obtain the proximal gradient type algorithm given as follows

$$\begin{cases} q_n = (Id + v_n A)^{-1}(Id - v_n B)(p_n + \mu_n(p_n - p_{n-1})), \\ p_{n+1} = (Id - \delta \eta_n S)q_n, \end{cases}$$

where $(\mu_n)_{n \in \mathbb{N}} \in \mathbb{R}^+$. Meanwhile, we set $\delta := \frac{1}{50}$, $\eta_n := \frac{1}{n}$ and $v_n = v := 10^{-8} \times \|Q^T Q\|$, where $n \in \mathbb{N}$. Next, we randomly choose the starting signal in the range of $(0, 1)^{4096}$ and take the number of iterations $n = 2000$ as the stopping criterion.

In this experiment, the minimum norm solution is the point in the set $\{x \in \mathbb{R}^{4096} \mid Q^T Qx = Q^T b\}$, which is closest to the original sparse signal. Thus, we can see that the result of this experiment for a signal sparse reconstruction is showed in Figure 2. By comparing the last two plots in Figure 2 and the top plot in Figure 2, one finds that the original sparse signal is recovered almost exactly. We hence conclude that our algorithm is efficient for dealing with the signal deblurring problem.

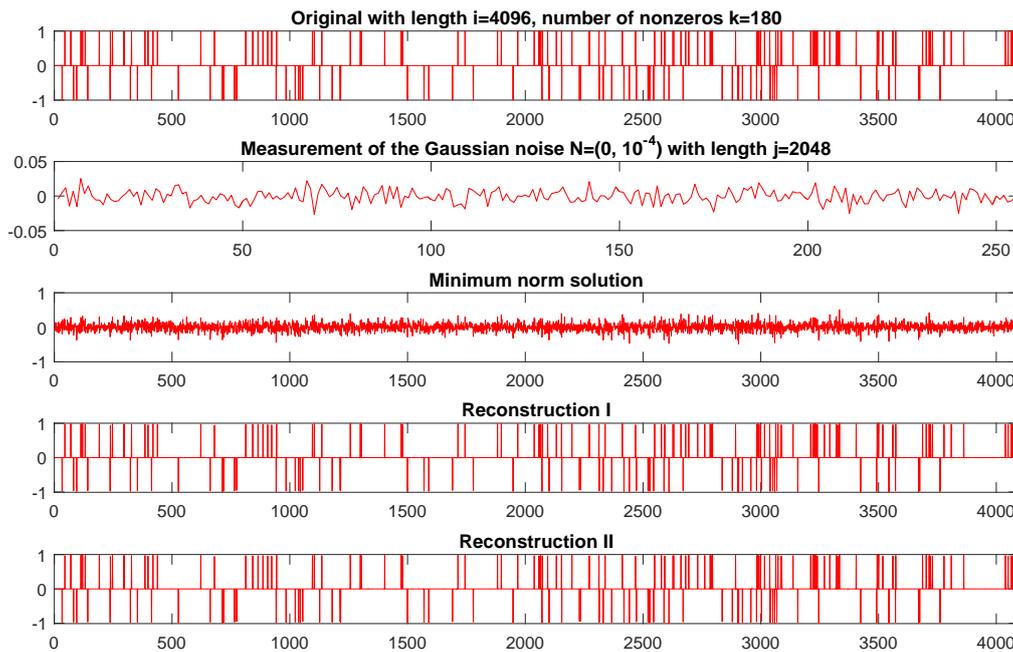


Figure 2. From top to bottom: the original signal, the noisy measurement, the minimum norm solution, the reconstruction signals respectively by Algorithm 1 ($\mu_n = 0.5$) and Algorithm 2.

To illustrate that Algorithm 1 has a competitive performance compared with Algorithm 2, we describe the following numerical results shown in Figure 3.

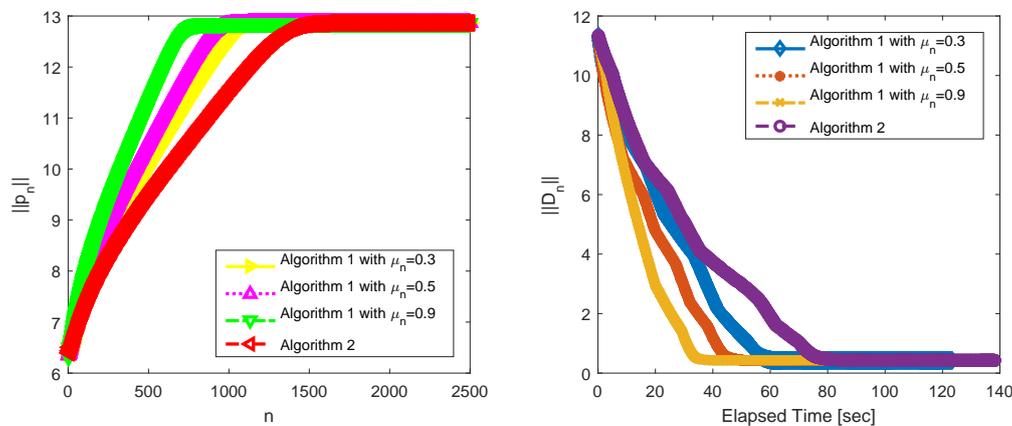


Figure 3. The behavior of $\|p_n\|$ with the number of iterations (left), the behavior of $\|D_n\|$ with the running time (right). The number of iterations is 2000. The first CPU time to compute $\|p_{2000}\|$ is about 123.41 s; the second CPU time to compute $\|p_{2000}\|$ is about 116.64 s; the third CPU time to compute $\|p_{2000}\|$ is about 116.59 s; the fourth CPU time to compute $\|p_{2000}\|$ is about 138.18 s (from top to bottom in legend).

The left plot in Figure 3 shows that the behaviors of the term $(\|p_n\|)_{n \in \mathbb{N}}$ (the y -axis) with respect to the number of iterations (x -axis), where p_n is generated by Algorithms 1 and 2. It can be observed from the test result reported in Figure 3 that $(\|p_n\|)_{n \in \mathbb{N}}$ converges to 12.8603. Thus, we can use $(\|p_n\|)_{n \in \mathbb{N}}$ to study the convergence and the computational performance of Algorithms 1 and 2. We denote $D_n = p_n - x$, where x is the original signal to be reconstructed, and p_n is an estimated signal of x . The right plot in Figure 3 shows that the value of $(\|D_n\|)_{n \in \mathbb{N}}$ (the y -axis) when the execution time in second elapses (x -axis). The sequence $(\|D_n\|)_{n \in \mathbb{N}}$ converges to 0.4152 with the running time. The restoration accuracy can be measured by means of the mean squared error. We find that the

mean squared error $MSE = \|D_n\|^2/i$ converges to $4.2088e^{-5}$, which further implies that the iterative sequence converges to the original signal x in this experiment.

We make a comparison of the behaviors of $(\|p_n\|)_{n \in \mathbb{N}}$, $(\|D_n\|)_{n \in \mathbb{N}}$ generated respectively by Algorithms 1 and 2. It shows that the bigger μ_n is, the fewer the required number of iterations becomes, the faster the convergence rate becomes. Furthermore, we find that all cases in Algorithm 1 need less computer time and enjoy a faster rate of the convergence than Algorithm 2. It can be observed from the plots that the changing process in all cases of Algorithm 1 outperforms Algorithm 2. The inertial extrapolation of Algorithm 1 plays a key role in the acceleration. For μ_k judiciously chosen, this inertial term improves the convergence speed of this algorithm.

5. Conclusions

In this paper, we introduced a variational inequality problem over the zero solution set of the inclusion problems. Since this problem has a double structure, it can be considered as a double-hierarchical optimization problem. The proposed algorithm uses the extrapolation term $\alpha_n(p_n - p_{n-1})$, which can be viewed as the procedure of speeding up the rate of convergence. As an application, the algorithm was used to solve the signal deblurring problem to support our convergence result.

Author Contributions: Conceptualization, L.L., X.Q. and J.-C.Y.; methodology, L.L., X.Q. and J.-C.Y.; data curation, L.L., X.Q. and J.-C.Y.; visualization, L.L.; supervision, X.Q. and J.-C.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

References

1. Moudafi, A. Split monotone variational inclusions. *J. Optim. Theory Appl.* **2011**, *150*, 275–283. [[CrossRef](#)]
2. Iiduka, H. Two stochastic optimization algorithms for convex optimization with fixed point constraints. *Optim. Methods Softw.* **2019**, *34*, 731–757. [[CrossRef](#)]
3. Ansari, Q.H.; Bao, T.Q. A limiting subdifferential version of Ekeland's variational principle in set optimization. *Optim. Lett.* **2019**, 1–5. 10.1007/s11590-019-01489-8. [[CrossRef](#)]
4. Al-Homidan, S.; Ansari, Q.H.; Kassay, G. Vectorial form of Ekeland variational principle with applications to vector equilibrium problems. *Optimization* **2020**, *69*, 415–436. [[CrossRef](#)]
5. Li, X.; Huang, N.; Ansari, Q.H.; Yao, J.C. Convergence rate of descent method with new inexact line-search on Riemannian manifolds. *J. Optim. Theory Appl.* **2019**, *180*, 830–854. [[CrossRef](#)]
6. Polyak, B.T. Some methods of speeding up the convergence of iteration methods. *USSR Comput. Math. Math. Phys.* **1964**, *4*, 1–17. [[CrossRef](#)]
7. Rockafellar, R.T. Monotone operators and the proximal point algorithm. *SIAM J. Control Optim.* **1976**, *14*, 877–898. [[CrossRef](#)]
8. Moreau, J.J. Proximité et dualité dans un espace hilbertien. *Bull. De La Société Mathématique De Fr.* **1965**, *93*, 273–299. [[CrossRef](#)]
9. Bruck, R.E.; Reich, S. Nonexpansive projections and resolvents of accretive operators in Banach spaces. *Houst. J. Math.* **1977**, *3*, 459–470.
10. Lions, P.L.; Mercier, B. Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.* **1979**, *16*, 964–979. [[CrossRef](#)]
11. Cho, S.Y.; Kang, S.M. Approximation of fixed points of pseudocontraction semigroups based on a viscosity iterative process. *Appl. Math. Lett.* **2011**, *24*, 224–228. [[CrossRef](#)]
12. Bot, R.I.; Csetnek, E.R. Second order forward-backward dynamical systems for monotone inclusion problems. *SIAM J. Control Optim.* **2016**, *54*, 1423–1443. [[CrossRef](#)]
13. Cho, S.Y.; Kang, S.M. Approximation of common solutions of variational inequalities via strict pseudocontractions. *Acta Math. Sci.* **2012**, *32*, 1607–1618. [[CrossRef](#)]

14. Cho, S.Y.; Li, W.; Kang, S.M. Convergence analysis of an iterative algorithm for monotone operators. *J. Inequal. Appl.* **2013**, *2013*, 199. [[CrossRef](#)]
15. Luo, Y.; Shang, M.; Tan, B. A general inertial viscosity type method for nonexpansive mappings and its applications in signal processing. *Mathematics* **2020**, *8*, 288. [[CrossRef](#)]
16. Alvarez, F.; Attouch, H. An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping. *Set-Valued Anal.* **2001**, *9*, 3–11. [[CrossRef](#)]
17. Lorenz, D.A.; Pock, T. An inertial forward-backward algorithm for monotone inclusions. *J. Math. Imaging Vis.* **2015**, *51*, 311–325. [[CrossRef](#)]
18. He, S.; Liu, L.; Xu, H.K. Optimal selections of stepsizes and blocks for the block-iterative ART. *Appl. Anal.* **2019**, 1–14. [[CrossRef](#)]
19. Cho, S.Y. Generalized mixed equilibrium and fixed point problems in a Banach space. *J. Nonlinear Sci. Appl.* **2016**, *9*, 1083–1092. [[CrossRef](#)]
20. Liu, L. A hybrid steepest descent method for solving split feasibility problems involving nonexpansive mappings. *J. Nonlinear Convex Anal.* **2019**, *20*, 471–488.
21. Kassay, G.; Reich, S.; Sabach, S. Iterative methods for solving systems of variational inequalities in reflexive Banach spaces. *SIAM J. Optim.* **2011**, *21*, 1319–1344. [[CrossRef](#)]
22. Censor, Y.; Gibali, A.; Reich, S.; Sabach, S. Common solutions to variational inequalities. *Set-Valued Var. Anal.* **2012**, *20*, 229–247. [[CrossRef](#)]
23. Gibali, A.; Liu, L.W.; Tang, Y.C. Note on the modified relaxation CQ algorithm for the split feasibility problem. *Optim. Lett.* **2018**, *12*, 817–830. [[CrossRef](#)]
24. Cho, S.Y. Strong convergence analysis of a hybrid algorithm for nonlinear operators in a Banach space. *J. Appl. Anal. Comput.* **2018**, *8*, 19–31.
25. Yamada, I. The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings. *Inherently Parallel Algorithms Feasibility Optim. Their Appl.* **2001**, *8*, 473–504.
26. Attouch, H.; Maingé, P.E. Asymptotic behavior of second-order dissipative evolution equations combining potential with non-potential effects. *ESAIM Control Optim. Calc. Var.* **2011**, *17*, 836–857. [[CrossRef](#)]
27. Iiduka, H.; Takahashi, W. Strong convergence theorems for nonexpansive mappings and inverse-strongly monotone mappings. *Nonlinear Anal.* **2005**, *61*, 341–350. [[CrossRef](#)]
28. Opial, Z. Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bull. Am. Math. Soc.* **1967**, *73*, 591–597. [[CrossRef](#)]
29. Maingé, P.E. Inertial iterative process for fixed points of certain quasi-nonexpansive mappings. *Set-Valued Anal.* **2007**, *15*, 67–79. [[CrossRef](#)]
30. Maingé, P.E. The viscosity approximation process for quasi-nonexpansive mappings in Hilbert spaces. *Comput. Math. Appl.* **2010**, *59*, 74–79. [[CrossRef](#)]
31. Peypouquet, J.; Sorin, S. Evolution equations for maximal monotone operators: Asymptotic analysis in continuous and discrete time. *arXiv* **2009**, arXiv:0905.1270.
32. He, X.; Huang, T.; Yu, J.; Li, C.; Li, C. An inertial projection neural network for solving variational inequalities. *IEEE Trans. Cybern.* **2016**, *47*, 809–814. [[CrossRef](#)] [[PubMed](#)]

