Nonlinear Functional Analysis and Applications
Vol. 23, No. 1 (2018), pp. 1-11
ISSN: 1229-1595(print), 2466-0973(online)

http://nfaa.kyungnam.ac.kr/journal-nfaa

CONVERGENCE OF GENERALIZED CONTRACTION-PROXIMAL POINT ALGORITHMS FOR SOLVING UNCONSTRAINED CONVEX OPTIMIZATION PROBLEMS

Pei-Chao Duan[1] and Yi Gao[2]

[1]Department of Mathematics
Civil Aviation University of China, Tianjin 300300, China
e-mail: pcduancauc@126.com

[2]Department of Mathematics
Civil Aviation University of China, Tianjin 300300, China
e-mail: yigaocauc@126.com

Abstract. In this paper, we get new contraction-proximal point algorithms for solving the unconstrained convex optimization problems which have the following iterative form:

$$\begin{cases} x_{n+1} = \alpha_n h(x_n) + \beta_n x_n + \mu_n V_{\lambda_n} x_n, \\ V_{\lambda_n} = prox_{\lambda_n g}(I - \lambda_n \nabla f) x_n. \end{cases}$$

Furthermore, we present the algorithm with bounded error: $x_{n+1} = \alpha_n h(x_n) + \beta_n x_n + \mu_n V_{\lambda_n} x_n + e_n$, where h is a τ-contractive mapping with $0 \leq \tau < 1$, V_{λ_n} is an averaged operator and e_n is the sequence error generated by itself of iteration. We also get the relative strong convergence under some conditions. It also extends the use of already existing algorithms.

1. Introduction and preliminaries

Throughout this paper, suppose that H is a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. $\Gamma_0(H)$ is the space of convex functions in H that are proper, lower semicontinuous and convex.

[0]Received January 9, 2017. Revised June 1, 2017.
[0]2010 Mathematics Subject Classification: 58E35, 47H09, 65J15.
[0]Keywords: Nonexpansive mapping, general iterative algorithm, proximal point operators, fixed point, variational inequality.
[0]Corresponding author: P.C. Duan(pcduancauc@126.com).

Consider the following unconstrained convex optimization problem:

$$\min_{x\in H} f(x)+g(x), \tag{1.1}$$

where $f,g\in\Gamma_0(H)$. It is often the case where f is differential and g is subdifferential. It was firstly proposed in 1978 by Mangasarian, Meyer and Robinson [8].

As we know, there already exists a lot of methods solving the unstrained convex minimization problem [8] which is a quite popular sector. Because some corresponding practical problems are arisen from image or signal processing, machine learning can be transferred into the machine form. And, in 2005, there comes a new iterative form which involves the proximal operator [2]. Combettes and Wajs proposed a classical method for solving such problem. For any initial guess $x_0\in H$, it generates the following iterative sequence as

$$x_{n+1}=(prox_{\lambda_n g}(I-\lambda_n\nabla f))x_n, \tag{1.2}$$

where $\lambda_n\in(0,\frac{2}{L})$ and it is well known that the algorithm converges weakly. Subsequently, Xu [11] proposed the relaxed proximal point algorithm:

$$x_{n+1}=(1-\alpha_n)x_n+\alpha_n(prox_{\lambda_n g}(I-\lambda_n\nabla f))x_n, \tag{1.3}$$

where $\lambda_n\in(0,\frac{2}{L})$, $\alpha_n\in(0,\frac{4}{2+L\limsup_{n\to\infty}\lambda_n})$ and obtained weak convergence under appropriate conditions. However, it is well known that strongly convergent algorithms are very important for solving the problem in infinite dimensional spaces. In 2015, Duan and Song in [3] proposed a generalized viscosity approximately algorithm:

$$x_{n+1}=\alpha_n h(x_n)+(1-\alpha_n)(prox_{\lambda_n g}(I-\lambda_n\nabla f))x_n, \tag{1.4}$$

where $\lambda_n\in(0,\frac{2}{L})$, $\alpha_n\in(0,\frac{2+\lambda_n L}{4})$ and prove its strong convergence under appropriate assumptions.

Now, in this paper, motivated by works of [2], [3], [12], and many other articles, we extend the above algorithms as:

$$x_{n+1}=\alpha_n h(x_n)+\beta_n x_n+\mu_n(prox_{\lambda_n g}(I-\lambda_n\nabla f))x_n, \tag{1.5}$$

where $\lambda_n\in(0,\frac{2}{L})$, $\alpha_n\in(0,\frac{2+\lambda_n L}{4})$, $\beta_n\in[0,1)$ and $\mu_n\in(0,1)$ satisfy with $\alpha_n+\beta_n+\mu_n=1$. It is worth noting that this algorithm may be regarded as the combination of algorithms (1.4) and (1.5). As a matter of fact, if we let $\rho_n=\frac{\mu_n}{1-\alpha_n}$, then we have

$$x_{n+1}=\alpha_n h(x_n)+(1-\alpha_n)[(1-\rho_n)x_n+\rho_n(prox_{\lambda_n g}(I-\lambda_n\nabla f))x_n]. \tag{1.6}$$

Next, we also prove the convergence of such following iterative form:

$$x_{n+1}=\alpha_n h(x_n)+\beta_n x_n+\mu_n(prox_{\lambda_n g}(I-\lambda_n\nabla f))x_n+e_n, \tag{1.7}$$

where e_n is the error sequence generated by (1.7). Namely, (1.6) is the exact version of algorithm (1.7).

Especially, taking $h(x_n) = u$, where $u \in H$ is fixed, then, under some conditions, the corresponding algorithm holds as following:

$$x_{n+1} = \alpha_n u + \beta_n x_n + \mu_n (prox_{\lambda_n g}(I - \lambda_n \nabla f))x_n + e_n. \tag{1.8}$$

Now, we recall some results which will be useful in proving our main results. Firstly, we use "$\rightarrow$" stands for strong convergence and "$\rightharpoonup$" stands for weak convergence. And the weak ω-limit set of a sequence $\{x_n\}$ will be denoted by $\omega_w(x_n)$.

A mapping $h : H \to H$ is called τ-contractive if there exists a contraction constant $\tau \in [0, 1)$ such that $\|h(x) - h(y)\| \leq \tau\|x - y\|$, for all $x, y \in H$. If we take $\tau = 1$, namely, $\|Tx - Ty\| \leq \|x - y\|$, for all $x, y \in H$, then T is called nonexpansive. And, a mapping is called Lipschitzian if there exists a positive constant L such that $\|Fx - Fy\| \leq L|x - y\|$, for all $x, y \in D(A)$. We denote a mapping F is η-inverse strongly monotone(η-ism), if there exists a constant $\eta > 0$ satisfies the following inequality $\langle Fx - Fy, x - y\rangle \geq \eta\|Fx - Fy\|^2$, for all $x, y \in H$. Also, a mapping $V : H \to H$ is called α-averaged(α-av for short) if $V = (1 - \alpha)I + \alpha T$, where $\alpha \in (0, 1)$, $T : H \to H$ is nonexpansive.

Let C be a nonempty closed convex subset of H. We use P_C to denote the projection form H onto C; namely, for $x \in H$, P_C is the unique point in C with the property:

$$\|x - P_C x\| = \min_{y \in C} \|x - y\|.$$

It is well known that $P_C x$ is characterized by:

$$\langle x - P_C x, z - P_C x\rangle \leq 0, \quad \forall z \in C.$$

Then, we will give some elementary properties of norms and proximal-operator in Hilbert spaces.

Lemma 1.1. *For all $x, y \in H$, there holds the following relation:*

$$\|x + y\|^2 \leq \|x\|^2 + 2\langle y, x + y\rangle.$$

Proposition 1.2. ([1]) *If $T_1, T_2, \cdots, T_n$ are averaged mappings, we can get that $T_n T_{n-1} \cdots T_1$ is averaged. In particular, if T_i is α_i-av, $i = 1, 2$, then $T_2 T_1$ is $(\alpha_2 + \alpha_1 - \alpha_2\alpha_1)$-av.*

Proposition 1.3. ([11]) *Let $T : H \to H$ be an operator from H to H.*

(i) *T is nonexpansive if and only if the complement $I - T$ is $\frac{1}{2}$-ism.*
(ii) *If T is v-ism, then for $\gamma > 0$, γT is $\frac{v}{\gamma}$-ism.*
(iii) *T is averaged if and only if the complement $I - T$ is v-ism for some $v > \frac{1}{2}$. Indeed, for $\alpha \in (0, 1)$, T is α-averaged if and only if the complement $I - T$ is $\frac{1}{2\alpha}$-ism.*

Lemma 1.4. ([4]) (**Demiclosedness Principle**) *Let H be a real Hilbert space, and let $T : H \to H$ be a nonexpensive mapping with $Fix(T) \neq \emptyset$. If $\{x_n\}$ is a sequence in H converges weakly to x and if $\{(I-T)x_n\}$ converges strongly to y, then $(I-T)x = y$; in particulary, if $y = 0$, then $x \in Fix(T)$.*

Lemma 1.5. ([5]) *Assume $\{s_n\}$ is a sequence of nonnegative real numbers such that*

$$s_{n+1} \leq (1-\gamma_n)s_n + \gamma_n\delta_n, \ n \geq 0,$$
$$s_{n+1} \leq s_n - \eta_n + \varphi_n, \ n \geq 0,$$

where $\{\gamma_n\}$ is a sequence in $(0,1)$, $\{\eta_n\}$ is a sequence of nonnegative real numbers and $\{\delta_n\}$ and $\{\varphi_n\}$ are two sequences in $\mathbb{R}$ such that

(i) $\Sigma_{n=0}^{\infty}\gamma_n = \infty$;
(ii) $\lim_{n\to\infty} \varphi_n = 0$;
(iii) $\lim_{k\to\infty} \eta_{n_k} = 0$ *implies* $\limsup_{k\to\infty} \delta_{n_k} \leq 0$ *for any subsequence* $\{n_k\} \subset \{n\}$.

Then $\lim_{n\to\infty} s_n = 0$.

Lemma 1.6. ([11]) *Let $\{s_n\}$ be a nonnegative real sequence satisfying*

$$s_{n+1} \leq (1-\gamma_n)s_n + \epsilon_n.$$

where the sequences $\{\gamma_n\} \subset (0,1)$ and $\{\epsilon_n\}$ are real sequences. Then $s_n \to 0$ as $n \to \infty$ provided that

(i) $\Sigma_{n=0}^{\infty}\gamma_n = \infty$;
(ii) $\limsup_{n\to\infty} \frac{\epsilon_n}{\gamma_n} \leq 0$ *or* $\Sigma_{n=0}^{\infty}|\epsilon_n| < \infty$.

Definition 1.7. ([6],[7]) The proximal operator of $\varphi \in \Gamma_0(H)$ is defined by

$$prox_{\varphi}(x) = \arg\min_{v\in H}\left\{\varphi(v) + \frac{1}{2}\|v-x\|^2\right\}, \quad x \in H.$$

The proximal operator of φ of order $\lambda > 0$ is defined as the proximal operator of $\lambda\varphi$, that is,

$$prox_{\lambda\varphi}(x) = arg\min_{v\in H}\left\{\varphi(v) + \frac{1}{2\lambda}\|v-x\|^2\right\}, \quad x \in H.$$

Follows this, some properties of proximal operator will also be listed.

Lemma 1.8. ([2],[9]) *Let $\varphi \in \Gamma_0(H)$ and $\lambda \in (0,\infty)$. Then we have the following statements.*

(i) *If C is a nonempty closed convex subset of H and $\varphi = I_C$ is the indicator function of C, then the proximal operator $prox_{\lambda\varphi} = P_C$ for all $\lambda > 0$, where P_C is the metric projection from C onto H.*

(ii) *The operator* $prox_{\lambda\varphi} = (I + \lambda\partial\varphi)^{-1} = J_\lambda^{\partial\varphi}$*, the resolvent of the subdifferential* $\partial\varphi$ *of* φ*.*

(iii) *If* $f : H \to \mathbb{R}$ *is a differentiable functional, then we denote by* ∇f *the gradient of* f*. Assume that* ∇f *is Lipschitz continuous on* H*. Then the operator* $V_\lambda = prox_{\lambda g}(I - \lambda\nabla f)$ *is* $\frac{2+\lambda L}{4}$*-av for each* $0 < \lambda < \frac{2}{L}$*.*

The proximal operator can be used to minimize the sum of two convex functions.

Lemma 1.9. ([12]) *The proximal identity*

$$prox_{\lambda\varphi}x = prox_{\mu\varphi}\left(\frac{\mu}{\lambda}x + (1 - \frac{\mu}{\lambda})prox_{\lambda\varphi}x\right) \tag{1.9}$$

holds for $\varphi \in \Gamma_0(H), \lambda > 0$ *and* $\mu > 0$*.*

It is useful in inducing some important equalities.

Proposition 1.10. ([12]) *Let* $f, g \in \Gamma_0(H)$*. Let* $x^* \in H$ *and* $\lambda > 0$*. Assume that* f *is finite-valued and differential on* H*. Then* x^* *is a solution to* (1.10) *if and only if* x^* *solves the fixed point equation*

$$x^* = (prox_{\lambda g}(I - \lambda\nabla f))x^*. \tag{1.10}$$

2. Main results

In what follows, we assume that $f, g \in \Gamma_0(H)$, $h(x)$ is a τ-contractive mapping of H with $0 \le \tau < 1$ and (1.1) is consist. Let S be the nonempty solution set of (1.1), $\{\alpha_n\} \subset (0, \frac{2+\lambda_n L}{4})$, $\{\beta_n\} \subset [0, 1)$, $0 < \liminf_{n\to\infty} \mu_n \le \limsup_{n\to\infty} \mu_n < 1$, and $\alpha_n + \beta_n + \mu_n = 1$. We first show the convergence of algorithm (1.6), and then extend to the algorithm (1.7).

Theorem 2.1. *Let the following conditions hold:*

(*a*) $\lim_{n\to\infty} \alpha_n = 0$;

(*b*) $\Sigma_{n=1}^{\infty}\alpha_n = \infty$;

(*c*) $0 < \liminf_{n\to\infty} \lambda_n \le \limsup_{n\to\infty} \lambda_n < \frac{2}{L}$.

Then, for any initial guess $x_0 \in H$*, define* $V_{\lambda_n} = prox_{\lambda_n g}(I - \lambda_n\nabla f)$*, the sequence* $\{x_n\}$ *generated by*

$$x_{n+1} = \alpha_n h(x_n) + \beta_n x_n + \mu_n V_{\lambda_n} x_n \tag{2.1}$$

converges strongly to z*. Where* z *is a solution of* (1.1) *and it is also the unique solution of variational inequality* $\langle h(z) - z, \tilde{x} - z\rangle \le 0$ *for all* $\tilde{x} \in S$*.*

Proof. The proof is divided into several steps.

Step 1. We will show that $\{x_n\}$ is bounded. Let $y_n = (1-\rho_n)x_n + \rho_n V_{\lambda_n}(x_n)$, where $\rho_n = \frac{\mu_n}{1-\alpha_n}$, $0 < \liminf_{n\to\infty} \rho_n \le \limsup_{n\to\infty} \rho_n < 1$. Then, It follows that

$$x_{n+1} = \alpha_n h(x_n) + (1-\alpha_n) y_n. \tag{2.2}$$

So, for any $z \in S$, we get

$$\begin{aligned} \|x_{n+1} - z\| &= \|\alpha_n h(x_n) + (1-\alpha_n)y_n - z\| \\ &\le \alpha_n \|h(x_n) - z\| + (1-\alpha_n)\|y_n - z\|. \end{aligned} \tag{2.3}$$

We also know

$$\begin{aligned} \|y_n - z\| &= \|(1-\rho_n)x_n + \rho_n V_{\lambda_n} x_n - z\| \\ &= \|(1-\rho_n)(x_n - z) + \rho_n (V_{\lambda_n} x_n - z)\| \\ &\le (1-\rho_n)\|x_n - z\| + \rho_n \|x_n - z\| \\ &= \|x_n - z\|. \end{aligned} \tag{2.4}$$

So,

$$\begin{aligned} \|x_{n+1} - z\| &\le \alpha_n \|h(x_n) - z\| + (1-\alpha_n)\|x_n - z\| \\ &\le \alpha_n \|h(x_n) - h(z)\| + \alpha_n \|h(z) - z\| + (1-\alpha_n)\|x_n - z\| \\ &\le (1 - \alpha_n(1-\tau))\|x_n - z\| + \alpha_n \|h(z) - z\| \\ &= (1 - \alpha_n(1-\tau))\|x_n - z\| + \alpha_n (1-\tau) \frac{\|h(z) - z\|}{1-\tau}. \end{aligned} \tag{2.5}$$

By induction, we obtain

$$\|x_{n+1} - z\| \le \max\left\{ \|x_0 - z\|, \frac{\|h(z) - z\|}{1-\tau} \right\}.$$

Hence, the sequence $\{x_n\}$ is bounded.

Step 2. Show that

$$\lim_{k\to\infty} \|x_{n_k} - V_{\lambda_{n_k}} x_{n_k}\| \to 0, \quad \forall \{n_k\} \subset \{n\}. \tag{2.6}$$

$$\begin{aligned} \|x_{n+1} - z\|^2 &= \|\alpha_n h(x_n) + \beta_n x_n + \mu_n V_{\lambda_n} x_n - z\|^2 \\ &= \|\alpha_n h(x_n) + (1-\alpha_n) y_n - z\|^2 \\ &= \|(1-\alpha_n)(y_n - z) + \alpha_n (h(x_n) - z)\|^2 \\ &\le (1-\alpha_n)^2 \|y_n - z\|^2 + 2\alpha_n \langle h(x_n) - z, x_{n+1} - z \rangle \\ &\le (1-\alpha_n)^2 \|y_n - z\|^2 + 2\alpha_n \tau \|x_n - z\| \|x_{n+1} - z\| \qquad (2.7) \\ &\quad + 2\alpha_n \langle h(z) - z, x_{n+1} - z \rangle \\ &\le (1-\alpha_n)^2 \|y_n - z\|^2 + \alpha_n \tau \|x_n - z\|^2 + \alpha_n \tau \|x_{n+1} - z\|^2 \\ &\quad + 2\alpha_n \langle h(z) - z, x_{n+1} - z \rangle. \end{aligned}$$

So, combine (2.4) and (2.7), we get that

$$\begin{aligned}\|x_{n+1}-z\|^2 &\le (1-\alpha_n)^2\|x_n-z\|^2+\alpha_n\tau\|x_n-z\|^2\\ &\quad +\alpha_n\tau\|x_{n+1}-z\|^2+2\alpha_n\langle h(z)-z,x_{n+1}-z\rangle. \end{aligned}\tag{2.8}$$

Namely,

$$\begin{aligned}(1-\alpha_n\tau)\|x_{n+1}-z\|^2 &\le ((1-\alpha_n)^2+\alpha_n\tau)\|x_n-z\|^2\\ &\quad +2\alpha_n\langle h(z)-z,x_{n+1}-z\rangle. \end{aligned}\tag{2.9}$$

Thus we have

$$\begin{aligned}&\|x_{n+1}-z\|^2\\ &\le (1-\frac{2\alpha_n(1-\tau)}{1-\alpha_n\tau})\|x_n-z\|^2+\frac{2\alpha_n(1-\tau)}{1-\alpha_n\tau}\frac{1}{1-\tau}\langle h(z)-z,x_{n+1}-z\rangle\\ &\quad +\frac{2\alpha_n(1-\tau)}{1-\alpha_n\tau}\frac{\alpha_n}{2(1-\tau)}M,\end{aligned}\tag{2.10}$$

where $M=\sup\|x_n-z\|^2$. Furthermore, we also have

$$\begin{aligned}&\|x_{n+1}-z\|^2\\ &=\|\alpha_nh(x_n)+(1-\alpha_n)y_n-z\|^2\\ &=\|\alpha_nh(x_n)-\alpha_ny_n+y_n-z\|^2\\ &=\alpha_n^2\|h(x_n)-y_n\|^2+\|y_n-z\|^2+2\alpha_n\langle h(x_n)-y_n,y_n-z\rangle\\ &=\|(1-\rho_n)x_n+\rho_nV_{\lambda_n}x_n-z\|^2+\alpha_n^2\|h(x_n)-y_n\|^2\\ &\quad +2\alpha_n\langle h(x_n)-y_n,y_n-z\rangle\\ &\le (1-\rho_n)\|x_n-z\|^2+\rho_n\|V_{\lambda_n}x_n-z\|^2-\rho_n(1-\rho_n)\|V_{\lambda_n}x_n-x_n\|^2\\ &\quad +\alpha_n^2\|h(x_n)-y_n\|^2+2\alpha_n\langle h(x_n)-y_n,y_n-z\rangle\\ &=\|x_n-z\|^2-\rho_n(1-\rho_n)\|V_{\lambda_n}x_n-x_n\|^2+\alpha_n^2\|h(x_n)-y_n\|^2\\ &\quad +2\alpha_n\langle h(x_n)-y_n,y_n-z\rangle.\end{aligned}\tag{2.11}$$

Set $\gamma_n=\frac{2\alpha_n(1-\tau)}{1-\alpha_n\tau}$, $\delta_n=\frac{1}{1-\tau}\langle h(z)-z,x_{n+1}-z\rangle+\frac{\alpha_n}{2(1-\tau)}M$, $\eta_n=\rho_n(1-\rho_n)\|V_{\lambda_n}x_n-x_n\|^2$, $\varphi_n=\alpha_n^2\|h(x_n)-y_n\|^2+2\alpha_n\langle h(x_n)-y_n,y_n-z\rangle$. Then $\Sigma_{n=0}^{\infty}\gamma_n=\infty$ and $\varphi_n\to 0$ hold obviously. In order to complete the proof by using Lemma 1.5, we have

$$\limsup_{k\to\infty}\langle h(z)-z,x_{n_k}-z\rangle=\langle h(z)-z,x\ -z\rangle\le 0,\ \ \forall\, x\in S.\tag{2.12}$$

It suffices to verify that $\eta_{n_k}\to 0(k\to\infty)$ implies that $\limsup_{k\to\infty}\delta_{n_k}\le 0$ for any subsequence $(n_k)\subset(n)$.

Step 3. Show that

$$\omega_w(x_n)\subset S.\tag{2.13}$$

Here, $\omega_w(x_n)$ is the set of all weak cluster points of $\{x_n\}$. Note that $\{x_n\}$ is bounded and (2.17) together guarantee that $\{x_n\}$ converges weakly to a point in S and then the proof is consist. To see (2.13), we prove as follows. Take $\tilde{x} \in \omega_w(x_n)$ and assume that $\{x_{n_j}\}$ is a subsequence of $\{x_n\}$ weakly converging to $\tilde{x}$. Hence by (2.6), $x_{n_j+1} \rightharpoonup \tilde{x}$ as well. Without loss of generality, we may assume $\lambda_{n_j} \to \lambda$, hence we get that $0 < \lambda < \frac{2}{L}$. Setting $V_\lambda = prox_{\lambda g}(I - \lambda \nabla f)$, then V_λ is nonexpansive. Set

$$y_j = x_{n_j} - \lambda_{n_j} \nabla f(x_{n_j}), \quad z_j = x_{n_j} - \lambda \nabla f(x_{n_j}).$$

Using the proximal identify of Lemma 1.9, we deduce that

$$\begin{aligned}
&\|V_{\lambda_{n_j}} x_{n_j} - V_\lambda x_{n_j}\| \\
&= \|prox_{\lambda_{n_j} g} y_j - prox_{\lambda g} z_j\| \\
&= \|prox_{\lambda g}(\frac{\lambda}{\lambda_{n_j}} y_j + (1 - \frac{\lambda}{\lambda_{n_j}}) prox_{\lambda_{n_j} g} y_j) - prox_{\lambda g} z_j\| \\
&\leq \|\frac{\lambda}{\lambda_{n_j}} y_j + (1 - \frac{\lambda}{\lambda_{n_j}}) prox_{\lambda_{n_j} g} y_j - z_j\| \\
&\leq \frac{\lambda}{\lambda_{n_j}} \|y_j - z_j\| + (1 - \frac{\lambda}{\lambda_{n_j}}) \|prox_{\lambda_{n_j} g} y_j - z_j\| \\
&= \frac{\lambda}{\lambda_{n_j}} |\lambda_{n_j} - \lambda| \|\nabla f(x_{n_j})\| + (1 - \frac{\lambda}{\lambda_{n_j}}) \|prox_{\lambda_{n_j} g} y_j - z_j\|. \quad (2.14)
\end{aligned}$$

Since $\{x_n\}$ is bounded, ∇f is Lipschitz continuous, and $\lambda_{n_j} \to \lambda$, we immediately derive from the last relation that $\|V_{\lambda_{n_j}} x_{n_j} - V_\lambda x_{n_j}\| \to 0$. As a result, we find

$$\|x_{n_j} - V_\lambda x_{n_j}\| \leq \|x_{n_j} - V_{\lambda_{n_j}} x_{n_j}\| + \|V_{\lambda_{n_j}} x_{n_j} - V_\lambda x_{n_j}\| \to 0. \quad (2.15)$$

Now the demiclosedness of the nonexpansive mapping $I - V_\lambda$ implies that $(I - V_\lambda)\tilde{x} = 0$. Namely, $\tilde{x} \in Fix(V_\lambda) = S$.
Indeed, since z is a solution of the variational inequality of $\langle h(z) - z, \tilde{x} - z \rangle \leq 0$, it is easy to see $\limsup_{k\to\infty} \delta_{n_k} \leq 0$. This shows condition (iii) in Lemma 1.5 holds and therefore the desired result at once follows. Thus we can conclude that $x_n \to z$. □

Remark 2.2. Note that $\beta_n = 0$ holds in Theorem 2.1, we can obtain the result of the reference [3]. However, the proof of Theorem 2.1 in this paper is simpler and the conditions are weaker than Theorem in reference [3].

Now, we are now in a position to state and prove a strong convergence result for sequences generated from algorithm (1.7) under several alternative conditions on the error sequences e_n.

Still, we assume that $f, g \in \Gamma_0(H)$, $h(x)$ is a τ-contractive operator of H to H with $0 \leq \tau < 1$ and (1.1) is consist. Let S be the nonempty solution set of (1.1), $\alpha_n \subset (0,1)$, $\beta_n \subset [0,1)$, $\mu_n \subset (0,1)$, and $\alpha_n + \beta_n + \mu_n = 1$.

Theorem 2.3. *Let the following conditions hold:*

(*a*) $\lim_{n\to\infty} \alpha_n = 0$;
(*b*) $\Sigma_{n=1}^{\infty} \alpha_n = \infty$;
(*c*) $0 < \liminf_{n\to\infty} \lambda_n \leq \limsup_{n\to\infty} \lambda_n < \frac{2}{L}$;
(*d*) $\Sigma_{n=1}^{\infty} \|e_n\| < \infty$.

Then, for any initial guess $x_0 \in H$, the sequence $\{x_n\}$ generated by

$$x_{n+1} = \alpha_n h(x_n) + \beta_n x_n + \mu_n V_{\lambda_n} x_n + e_n \tag{2.16}$$

converges strongly to z. where z is a solution of (1.1) *and it is also the unique solution of variational inequality $\langle h(z) - z, \tilde{x} - z\rangle \leq 0$ for all $\tilde{x} \in S$.*

Proof. Taking Theorem 2.1 into account, it is enough to prove that $\|x_n - v_n\| \to 0$ as $n \to \infty$, for $v_{n+1} = \alpha_n h(v_n) + \beta_n v_n + \mu_n V_{\lambda_n} v_n$. Since V_{λ_n} is nonexpansive and $h(x)$ is contractive, we have

$$\begin{aligned}
&\|x_{n+1} - v_{n+1}\| \\
&\leq \alpha_n \tau \|h(x_n) - h(v_n)\| + \beta_n \|x_n - v_n\| + \mu_n \|x_n - v_n\| + \|e_n\| \\
&= \alpha_n \tau \|x_n - v_n\| + (1 - \alpha_n)\|x_n - v_n\| + \|e_n\| \\
&= (1 - \alpha_n(1-\tau))\|x_n - v_n\| + \|e_n\| \\
&= (1 - \alpha_n(1-\tau))\|x_n - v_n\| + \alpha_n(1-\tau)\frac{\|e_n\|}{\alpha_n(1-\tau)}.
\end{aligned} \tag{2.17}$$

Since the sequence of errors satisfies the condition (d), it readily follows from Lemma 1.5 that $\|x_n - v_n\| \to 0$. This completes the proof. □

Especially, if we fix $h(x_n)$, namely, take $h(x_n) = u$, where $u \in H$ is fixed. Then, we have the following results. And the proof is similar to Theorem 2.1.

Theorem 2.4. *Let $f, g \in \Gamma_0(H)$, and assume that* (1.1) *is consist. Let $V_{\lambda_n} = prox_{\lambda_n g}(I - \lambda_n \nabla f)$, where ∇f is L-Lipschitizian. Given $x_0 \in H$ and define the sequence $\{x_n\}$ by the following iterative algorithm:*

$$x_{n+1} = \alpha_n u + \beta_n x_n + \mu_n V_{\lambda_n} x_n, \tag{2.18}$$

where $\alpha_n \subset (0,1)$, $\beta_n \subset [0,1)$, $\mu_n \subset (0,1)$, and $\lambda_n \in (0, \frac{2}{L})$, Suppose that

(a) $\lim_{n\to\infty} \alpha_n = 0$;
(b) $\Sigma_{n=1}^{\infty} \alpha_n = \infty$;
(c) $0 < \liminf_{n\to\infty} \lambda_n \leq \limsup_{n\to\infty} \lambda_n < \frac{2}{L}$;
(d) $\Sigma_{n=1}^{\infty} \|e_n\| < \infty$.

Then, $\{x_n\}$ *converges strongly to* $P_S(u)$ *which is the nearest point of* S *to* u *and is also a solution of* (1.1)*.*

Theorem 2.5. *Let* $f, g \in \Gamma_0(H)$*, and assume that* (1.1) *is consist. Let* $V_{\lambda_n} = prox_{\lambda_n g}(I - \lambda_n \nabla f)$*, where* ∇f *is L-Lipschitzian. Given* $x_0 \in H$ *and define the sequence* $\{x_n\}$ *by the following iterative algorithm:*

$$x_{n+1} = \alpha_n u + \beta_n x_n + \mu_n V_{\lambda_n} x_n + e_n, \tag{2.19}$$

where $\alpha_n \subset (0, 1)$*,* $\beta_n \subset [0, 1)$*,* $\mu_n \subset (0, 1)$*, and* $\lambda_n \in (0, \frac{2}{L})$*. Suppose that*

(a) $\lim_{n\to\infty} \alpha_n = 0$*;*
(b) $\Sigma_{n=1}^{\infty} \alpha_n = \infty$*;*
(c) $0 < \liminf_{n\to\infty} \lambda_n \leq \limsup_{n\to\infty} \lambda_n < \frac{2}{L}$*.*

Then, $\{x_n\}$ *converges strongly to the nearest point of* S *to* $u(P_S u)$ *which is also the solution of* (1.1)*.*

Proof. For given initial point v_0, we define $\{v_n\}$ recursively by

$$v_{n+1} = \alpha_n u + \beta_n v_n + \mu_n V_{\lambda_n} v_n, \tag{2.20}$$

By Theorem 2.4, $\{v_n\}$ converges strongly to $P_S u$. So it remains to show that $\|x_n - v_n\| \to 0$. Since

$$\begin{aligned} \|x_{n+1} - v_{n+1}\| &\leq \|\beta_n(x_n - v_n) + \mu_n(V_{\lambda_n} x_n - V_{\lambda_n} v_n)\| + \|e_n\| \\ &\leq \beta_n \|(x_n - v_n)\| + \mu_n \|V_{\lambda_n} x_n - V_{\lambda_n} v_n)\| + \|e_n\| \\ &\leq (\beta_n + \delta_n)\|x_n - v_n\| + \|e_n\| \\ &= (1 - \alpha_n)\|x_n - v_n\| + \alpha_n \frac{\|e_n\|}{\alpha_n}, \end{aligned} \tag{2.21}$$

it follows from Lemma 1.6 that $\|x_n - v_n\| \to 0$. Therefore we have the desired result. □

Acknowledgments: The authors thank the referees for their helping comments, which notably improved the presentation of this paper. This work was supported by Fundamental Research Funds for the Central Univ.(3122017072).

References

[1] P.L. Combettes, *Solxing monotone inclusions via compositions of nonexpansive averaged operators*, Opitim., **53** (2004), 475–504.
[2] P.L. Combettes and R. Wajs, *Signal recovery by proximal forward-backward spilitting*, Multiscale Model. Simul., **4** (2005), 1168–1200.
[3] P.C. Duan and M.M. Song, *Strong convergence of an iterative method for the lasso*, Nonlinear Funct. Anal. and Appl., **20**(2) (2015), 247–256.
[4] K. Geobel and W.A. Kirk, *Topics in metric fixed point theory*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 1990.

[5] S.N. He and C.P. Yang, *Solving the variational inequality problem defined on intersection of finite level sets*, Abstract Appl. Anal., **2013** (2013), Article ID 942315.
[6] J.J. Moreau, *Proximites des applications "prox"*, C.R. Acad. Sci. Paris, Sér.A Math., **256** (1963), 1069–1071.
[7] J.J. Moreau, *Proximite et dualite dans un espace hulbertien*, Bull. Soc. Math. France, **93** (1965), 279–299.
[8] O.L. Mangasarian, R.R. Meyer, S.M. Robinson and A. Auslender, *Mininisation de fonctions localement Lipschitiziwnnes: appligramming a la programmation mi-convexe, mi-differential*, Nonlnear Programming 3, Eds.,Academic Press, New York, 1978, 429–460.
[9] C.A. Micchellil, L. Shen and Y. Xu, *Proximity algorithms for image models*, Denoising, Inverse Probl., **27**(4) (2011), Article ID 045009.
[10] F.H. Wang and H.H. Cui, *Convergence of the generalized contraction-proximal point algorithm in a Hilbert space*, Optim., **64**(4) (2015), 709–715.
[11] H.K. Xu, *Averaged mapping and the gradient-projection algorithm*, J. of Optimi. Theory and Appl., **150**(3) (2011), 360–378.
[12] H.K. Xu, *Properties and iterative methods for the lasso and ite variants*, Chin. Ann. Math., **35B**(3) (2014), 1–18.