Noname manuscript No.
(will be inserted by the editor)

New inertial factors of the Krasnosel'skiĭ-Mann iteration

Yunda Dong

Received: date / Accepted: date

Abstract In this article, we consider the Krasnosel'skiĭ-Mann iteration for approximating a fixed point of any given non-expansive operator in real Hilbert spaces, and we study an inertial version proposed by Maingé recently. As a result, we suggest new conditions on the inertial factors to ensure weak convergence. They are free of iterates and depend on the original coefficient of the Krasnosel'skiĭ-Mann iteration. In particular, in a special case that corresponds to the Douglas-Rachford splitting, the upper bound of the sequence of inertial factors is merely required to strictly less than 1/3. Rudimentary numerical results indicate practical usefulness of our suggested conditions.

Keywords Non-expansive operator · fixed point · Krasnosel'skiĭ-Mann iteration · inertial factor · the Douglas-Rachford splitting

Mathematics Subject Classification (2000) 58E35 · 65K15

1 Introduction

Let $\mathcal{H}$ be an infinite-dimensional real Hilbert space, and let $\mathcal{C} \subseteq \mathcal{H}$ be a nonempty closed convex subset. If $T : \mathcal{C} \to \mathcal{C}$ is non-expansive, then one of its fixed points (if at least one fixed point exists) can be approximated by the following Krasnosel'skiĭ-Mann iteration [28, 24]

$$x_{k+1} = (1 - \alpha_k)x_k + \alpha_k T(x_k), \quad k = 0, 1, ...,$$

Communicated by

Yunda Dong
School of Mathematics and Statistics, Zhengzhou University, Zhengzhou 450001, PR China
E-mail: ydong@zzu.edu.cn

where the coefficient $\alpha_k \in [0,1]$ and the series $\sum \alpha_k(1-\alpha_k)$ diverges [19]. For any given starting point $x_0 \in \mathcal{C}$, the resulting sequence converges weakly to some fixed point of T. Thanks to the celebrated counter-example [17], one has known that the Krasnosel'skiĭ-Mann iteration may fail to converge strongly in some cases. For pertinent discussions, we refer to [33, 23, 30, 22] and the references cited therein.

An appealing feature of the Krasnosel'skiĭ-Mann iteration is that the proximal point algorithm (introduced by Martinet [29] and generalized by Rockafellar [35]; also see [3, 20, 9, 10]), together with its generalized form [18], and the Peaceman/Douglas-Rachford splitting method [34, 14] of Lions and Mercier [25] may well be interpreted as its special cases.

Recently, inspired by the work on the proximal point algorithm [1], [27] suggested adding an inertial term to the Krasnosel'skiĭ-Mann iteration:

$$\begin{aligned} &y_k = x_k + t_k(x_k - x_{k-1}),\\ &x_{k+1} = (1-\alpha_k)y_k + \alpha_k T(y_k), \quad k = 0, 1, ..., \end{aligned}$$

where the inertial factor $t_k > 0$. For any given starting points x_{-1}, x_0 in the set $\mathcal{C}$, the corresponding weak convergence was proved if $\mathcal{C}$ is further assumed to be affine and the following additional condition, which follows the style of [1],

$$\sum_{k=0}^{+\infty} t_k \|x_k - x_{k-1}\|^2 < +\infty \tag{1}$$

is satisfied. See [8, 36] for very recent discussions of adding inertial terms to other iterative schemes.

Although the condition (1) can be verified in practice, it remains desirable to give its alternative free of iterates. A step toward this direction was taken in a recent work [5]. The conditions on the inertial sequence include

$$t_0 = 0, \quad 0 \le t_k \le t_{k+1} \le t < 1$$

and other relations between the coefficient α_k and upper bound of the inertial sequence; also see (3) below.

Unfortunately, as pointed out at the end of Sect. 3 and in the Remark 4.3, these conditions are too complicated to determine upper bound of the inertial sequence in a simple way even if the coefficient α_k has been known. Moreover, in the case of $\alpha_k \equiv 0.5$, they are undesirably restrictive.

To circumvent such difficulty, we mainly consider the case $\alpha_k \in [0.5, 1)$ that covers the Douglas-Rachford splitting method and give a concise relation between the coefficient and (upper bound of) the inertial sequence

$$t_0 = 0, \quad 0 \le t_k \le t_{k+1} \le \frac{1-(1+\varepsilon)\alpha_k}{1+\alpha_k},$$

where ε is any given sufficiently small positive number. In contrast to (3), our suggested conditions above have an appealing property: once the coefficient α_k has been chosen before implementing the algorithm, it becomes very convenient to determine (upper bound of) the inertial sequence dynamically; we refer to [12] for further applications in practice.

Interestingly, in the case of $\alpha_k \equiv 0.5$ that corresponds to the Douglas-Rachford splitting, the conditions become that the inertial sequence is uniformly bounded by 1/3 (see Corollary 1 below) and it coincides with the counterpart for an inertial proximal point algorithm suggested in [1, Proposition 2.1].

The rest of this article is organized as follows. In Sect. 2, we give some useful concepts and preliminary results. In Sect. 3, we formally describe the inertial Krasnosel'skiĭ-Mann iteration, and for comparison, we state the main results in [5]. In Sect. 4, we give our suggested conditions on the inertial sequence and prove weak convergence of the inertial Krasnosel'skiĭ-Mann iteration. Moreover, we compare them with the existing ones [5]. In Sect. 5, we discuss implications to generalized proximal point algorithm and the Peaceman/Douglas-Rachford splitting. In Sect. 6, we did rudimentary numerical experiments to confirm practical usefulness of our suggested conditions, when applied to an inertial Douglas-Rachford splitting. In Sect. 7, we close this article by some concluding remarks.

2 Preliminary Results

In this section, we first give some basic definitions and then provide some auxiliary results for later use.

Let $\mathcal{H}$ be an infinite-dimensional Hilbert space, in which $\langle x, y\rangle$ stands for the usual inner product and $\|x\| := \sqrt{\langle x, x\rangle}$ for the induced norm for any $x, y \in \mathcal{H}$. Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be an operator, which may be multi-valued. To concisely give the following definition, we agree on that the notation $(x, w) \in A$ and $x \in \mathcal{H}$, $w \in A(x)$ have the same meaning. Moreover, $w \in A(x)$ if and only if $x \in A^{-1}w$, where A^{-1} stands for the inverse of A. dom$A$ stands for the effective domain of A.

Definition 1 Let $\mathcal{C} \subseteq \mathcal{H}$ be a nonempty subset. An operator $T : \mathcal{C} \to \mathcal{C}$ is called non-expansive if and only if

$$\|T(x) - T(y)\| \le \|x - y\|, \quad \forall x, y \in \mathcal{C};$$

firmly non-expansive if and only if

$$\|T(x) - T(y)\|^2 \le \langle x - y, T(x) - T(y)\rangle, \quad \forall x, y \in \mathcal{C}.$$

Definition 2 Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be an operator. It is called monotone if and only if

$$\langle x - x',\ w - w' \rangle \geq 0, \quad \forall (x, w) \in A,\ \forall (x', w') \in A;$$

maximal monotone if and only if it is monotone and for given $\hat{x} \in \mathcal{H}$ and $\hat{w} \in \mathcal{H}$ the following implication relation holds

$$\langle x - \hat{x},\ w - \hat{w} \rangle \geq 0, \quad \forall (x, w) \in A \quad \Rightarrow \quad (\hat{x}, \hat{w}) \in A.$$

Definition 3 Let $f : \mathcal{H} \to [-\infty, +\infty]$ be a convex function. Then for any given $x \in \mathcal{H}$ the sub-differential of f at x is defined by

$$\partial f(x) := \{s \in \mathcal{H} : f(y) - f(x) \geq \langle s, y - x \rangle,\ \forall y \in \mathcal{H}\}.$$

Each s is called a sub-gradient of f at x. Moreover, if f is further continuously differentiable, then $\partial f(x) = \{\nabla f(x)\}$, where $\nabla f(x)$ is the gradient of f at x.

As is well known, the sub-differential of any closed proper convex function in an infinite-dimensional Hilbert space is maximal monotone as well. An important example is the sub-differential of the indicator function defined by

$$\delta_{\mathcal{C}}(x) = \begin{cases} 0, & \text{if } x \in \mathcal{C}, \\ +\infty, & \text{if } x \notin \mathcal{C}, \end{cases}$$

where $\mathcal{C}$ is some nonempty closed convex set in $\mathcal{R}^n$. Moreover, for any given positive number $\lambda > 0$, we have $P_{\mathcal{C}} = (I + \lambda \partial \delta_{\mathcal{C}})^{-1}$, where $P_{\mathcal{C}}$ is usual projection onto C.

For any given maximal monotone operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$, it is Minty [31] who proved that there must exist a unique $y \in \mathcal{H}$ such that $(I + \lambda A)y \ni x$ for all $x \in \mathcal{H}$ and $\lambda > 0$, where I stands for the identity operator, i.e., $Ix = x$ for all $x \in \mathcal{H}$. This implies that the corresponding operator $J_{\lambda A} := (I + \lambda A)^{-1}$, also called the resolvent operator of A, is single-valued. Be aware that $J_{\lambda A}$ here is firmly non-expansive, and $2J_{\lambda A} - I$ is non-expansive (cf. [15]).

For any given maximal monotone operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$, there are other related properties. (i) For all $x \in \mathcal{H}$, the set $A(x)$ must be either empty or nonempty closed convex; see [2, Proposition 3, § 6.7]. (ii) The solution set $\{x : 0 \in A(x)\}$ is either empty or nonempty closed convex [32]. Note that if T is non-expansive, then $I - T$ is continuous and monotone, thus maximal monotone. Moreover, the set of its fixed points must be either empty or nonempty closed convex.

3 Inertial Krasnosel'skiĭ-Mann iteration

In this section, we restate and further discuss the above-mentioned inertial Krasnosel'skiĭ-Mann iteration for non-expansive operators on some nonempty, closed and affine subset $\mathcal{C}$ in Hilbert spaces.

First, we restate the inertial Krasnosel'skiĭ-Mann iteration suggested in [27]. For any given starting points x_{-1}, x_0 in the set $\mathcal{C}$, it reads

$$x_{k+1} = ((1-\alpha_k)I + \alpha_k T)\,(x_k + t_k(x_k - x_{k-1})), \quad k = 0, 1, ... \tag{2}$$

Next, we give [5, Theorem 5], which is the basic result there.

Proposition 1 *Let $\mathcal{C} \subseteq \mathcal{H}$ be a nonempty closed affine subset. Assume that $T : \mathcal{C} \to \mathcal{C}$ is non-expansive and there exists at least one fixed point. Consider the inertial Krasnosel'skiĭ-Mann iteration (2) above. If the inertial sequence satisfies*

$$t_0 = 0, \quad 0 \leq t_k \leq t_{k+1} \leq t < 1,$$

and $\{\alpha_k\}$ satisfies

$$0 < \underline{\alpha} \leq \alpha_k \leq \frac{\delta(1-t^2) - t^2 - t^3 - t\sigma}{\delta[1 + t(1+t) + t\delta + \sigma]}, \quad \text{with } \delta > \frac{t^2 + t^3 + t\sigma}{1 - t^2}, \tag{3}$$

where $\delta > 0$ and $\sigma > 0$. Then the sequence $\{x_k\}$ converges weakly to a fixed point of T.

Obviously, these conditions above are complicated. Furthermore, as indicated in Remark 4.3 below, they are restrictive.

4 Convergence

In this section, we mainly analyze weak convergence of the above-mentioned inertial Krasnosel'skiĭ-Mann iteration, and we specially stress that our suggested conditions on the inertial sequence are new and convenient in practice.

This section begins with a well-known lemma [1], which is used for simplifying the proof of our main theorem in this article.

Lemma 4.1 *Let $\{\varphi_k\}$, $\{t_k\}$ and $\{\delta_k\}$ be nonnegative sequences. Assume that*

$$\varphi_{k+1} \leq \varphi_k + t_k(\varphi_k - \varphi_{k-1}) + \delta_k, \quad k = 0, 1, ...,$$

and $0 \leq t_k \leq t < 1$ and $\sum_{k=0}^{+\infty} \delta_k < +\infty$. Then $\lim_{k\to+\infty} \varphi_k$ exists.

The following result can be derived by slightly modifying the proof of [5, Theorem 5] and its proof is in the Appendix.

Lemma 4.2 *Let $\{x_k\}$ be the sequences generated by the inertial Krasnosel'skiĭ-Mann iteration (2). Then, for any given fixed point z, we have*

$$\begin{aligned} \|x_{k+1} - z\|^2 \leq{}& \|x_k - z\|^2 + t_k\left(\|x_k - z\|^2 - \|x_{k-1} - z\|^2\right) \\ & - \left(\frac{1}{\alpha_k} - 1\right)(1 - t_k)\|x_{k+1} - x_k\|^2 + \gamma_k \|x_k - x_{k-1}\|^2 \end{aligned} \tag{4}$$

where γ_k *is given by*

$$\gamma_k := t_k + t_k^2 - \left(\frac{1}{\alpha_k} - 1\right)(t_k - 1)\, t_k.$$

Now we state the main result in this article.

Theorem 4.1 *In the setting of Lemma 4.2, we further assume that*

$$\alpha_k \in [0.5, 1-\varepsilon], \quad t_0 = 0,\ 0 \le t_k \le t_{k+1} \le \frac{1-(1+\varepsilon)\alpha_k}{1+\alpha_k}, \tag{5}$$

where ε *is any given sufficiently small positive number. Then the inertial Krasnosel'skiĭ-Mann iteration (2) has the following properties*

(i) $\sum_{k=0}^{+\infty} \|x_{k+1} - x_k\|^2 < +\infty$.

(ii) The sequence $\{x_k\}$ *converges weakly to a fixed point of* T.

Proof Denote

$$\varphi_k := \|x_k - z\|^2, \quad \mu_k := \varphi_k - t_k\varphi_{k-1} + \gamma_k \|x_k - x_{k-1}\|^2. \tag{6}$$

Since the inertial sequence $\{t_k\}$ is non-decreasing, it follows from Lemma 4.2 that

$$\begin{aligned}
&\mu_{k+1} - \mu_k \\
&\le \varphi_{k+1} - (1+t_k)\varphi_k + t_k\varphi_{k-1} + \gamma_{k+1}\|x_{k+1} - x_k\|^2 - \gamma_k\|x_k - x_{k-1}\|^2 \\
&\le -\left(\left(\frac{1}{\alpha_k} - 1\right)(1 - t_k) - \gamma_{k+1}\right)\|x_{k+1} - x_k\|^2.
\end{aligned} \tag{7}$$

Meanwhile, since $\alpha_k \in [0.5, 1)$ and $t_k \le t_{k+1} < 1$, we have

$$\begin{aligned}
&\left(\frac{1}{\alpha_k} - 1\right)(1 - t_k) - \gamma_{k+1} \\
&= \left(\frac{1}{\alpha_k} - 1\right)(1 - t_k) - t_{k+1} - t_{k+1}^2 + \left(\frac{1}{\alpha_{k+1}} - 1\right)(t_{k+1} - 1)\, t_{k+1} \\
&\ge \left(\frac{1}{\alpha_k} - 1\right)(1 - t_{k+1}) - t_{k+1} - t_{k+1}^2 + \left(\frac{1}{\alpha_{k+1}} - 1\right)(t_{k+1} - 1)\, t_{k+1} \\
&= \frac{1}{\alpha_k} - 1 - \left(\frac{1}{\alpha_k} + \frac{1}{\alpha_{k+1}} - 1\right) t_{k+1} - \left(2 - \frac{1}{\alpha_{k+1}}\right) t_{k+1}^2 \\
&\ge \frac{1}{\alpha_k} - 1 - \left(\frac{1}{\alpha_k} + \frac{1}{\alpha_{k+1}} - 1\right) t_{k+1} - \left(2 - \frac{1}{\alpha_{k+1}}\right) t_{k+1} \\
&= \frac{1}{\alpha_k} - 1 - \left(\frac{1}{\alpha_k} + 1\right) t_{k+1} \\
&\ge \varepsilon,
\end{aligned} \tag{8}$$

where the last inequality follows from (5). Combining this with (7) yields that, in either case, we always have

$$\mu_{k+1} - \mu_k \le -\varepsilon\,\|x_{k+1} - x_k\|^2, \ \ k = 0, 1, \ldots.$$

Consequently, the sequence $\{\mu_k\}$ is non-increasing. Summing up for $j \le k-1$ yields

$$\varepsilon \sum_{j=0}^{k-1} \|x_{j+1} - x_j\|^2 \le \mu_0 - \mu_k \le \mu_0 + t_k \varphi_{k-1},$$

where the last inequality follows from (6). Below, we need to prove that the sequence $\{\varphi_k\}$ has an upper bound. In fact, in view of (6) and that $\{\mu_k\}$ is non-increasing, we have

$$\varphi_k - t_k \varphi_{k-1} \le \mu_k \le \mu_0.$$

So, we further get

$$\begin{aligned} \varphi_k &\le t_k \varphi_{k-1} + \mu_0 \\ &\le t \varphi_{k-1} + \mu_0 \\ &\le t(t\varphi_{k-2} + \mu_0) + \mu_0 \\ &= t^2 \varphi_{k-2} + t\mu_0 + \mu_0 \end{aligned}$$

and eventually

$$\varphi_k \le t^k \varphi_0 + t^{k-1}\mu_0 + \cdots + t\mu_0 + \mu_0 \le t^k \varphi_0 + \frac{\mu_0}{1-t}.$$

Thus, we can say that $\sum_{j=0}^{k-1} \|x_{j+1} - x_j\|^2$ is always bounded for all k and the item (i) of this theorem is proved.

Finally, we prove weak convergence of the sequence $\{x_k\}$. In fact, from Lemma 4.2, we can see that

$$\|x_{k+1} - z\|^2 \le \|x_k - z\|^2 + t_k \left(\|x_k - z\|^2 - \|x_{k-1} - z\|^2 \right) + \gamma_k \|x_k - x_{k-1}\|^2.$$

Combining this with the item (i) of this theorem and Lemma 4.1 yields that $\lim_{k\to+\infty} \|x_k - z\|$ exists. This indicates that the sequence $\{x_k\}$ is bounded in norm, thus it has at least one weak cluster point, say $\bar{z}$, i.e., there exists some subsequence $\{x_{k_j}\}$ such that it converges weakly to $\bar{z}$. Meanwhile, it follows from (2) that

$$(I - T)(x_k + t_k(x_k - x_{k-1})) = (1/\alpha_k)(t_k(x_k - x_{k-1}) - (x_{k+1} - x_k)).$$

Since the item (i) means that $\{x_{k+1} - x_k\}$ converges to zero in norm, taking this into account and taking the limit along k_j yield

$$(I - T)(\bar{z}) = 0 \qquad \Leftrightarrow \qquad T(\bar{z}) = \bar{z},$$

where we have made use of the fact that $I - T$ is continuous and monotone. The proof of uniqueness of weak cluster point is standard [35, 13], thus is omitted. The proof is complete.

Remark 4.1 Impressively, as far as the inner product $2\langle x_{k+1}-x_k, x_k-x_{k-1}\rangle$ in (25) (see the Appendix A below) is concerned, we adopt the following upper bound

$$\|x_{k+1}-x_k\|^2+\|x_k-x_{k-1}\|^2$$

and do not follow [5, Theorem 5] to make use of the upper bound

$$\rho_k\|x_{k+1}-x_k\|^2+(1/\rho_k)\|x_k-x_{k-1}\|^2, \quad \rho_k>0.$$

The reason is that introducing ρ_k makes analysis much more complicated and restrictive as mentioned at the end of Sect. 3. Perhaps, $\rho_k\equiv 1$ itself has been a good choice. As to the relation (8), it is better if we require

$$\frac{1}{\alpha_k}-1-\left(\frac{1}{\alpha_k}+\frac{1}{\alpha_{k+1}}-1\right)t_{k+1}-\left(2-\frac{1}{\alpha_{k+1}}\right)t_{k+1}^2\geq\varepsilon,$$

which implies

$$t_{k+1}\leq\sqrt{p_k^2+q_k}-p_k, \tag{9}$$

where p_k and q_k are given by

$$p_k:=\frac{1}{2}\frac{1}{2-\frac{1}{\alpha_{k+1}}}\left(\frac{1}{\alpha_k}+\frac{1}{\alpha_{k+1}}-1\right), \quad q_k:=\frac{1}{2-\frac{1}{\alpha_{k+1}}}\left(\frac{1}{\alpha_k}-1-\varepsilon\right).$$

In practice, by assuming $\alpha_k\in[0.5+\varepsilon,1-\varepsilon]$ in advance, we may make use of this formula to calculate larger upper bound of the inertial sequence dynamically. Yet, in theory, we prefer (5) to it for simplicity.

Remark 4.2 In Theorem 4.1, the focus of our analysis is on the case of $\alpha_k\in[0.5,1-\varepsilon]$, in which the corresponding Peaceman/Douglas-Rachford splitting method – an instance of the Krasnosel'skiĭ-Mann iteration – typically yields better numerical results [13]. The reader may analyze other cases similarly.

Remark 4.3 Now let us probe into the intrinsic issue behind the aforementioned conditions (3). Consider an inequality there, which can be rewritten as

$$\alpha_k t\delta^2+(\alpha_k t^2+t^2+\alpha_k t+\alpha_k+\alpha_k\sigma-1)\delta+t^2+t^3+t\sigma\leq 0,$$

namely,

$$b_k:=\alpha_k t^2+t^2+\alpha_k t+\alpha_k+\alpha_k\sigma-1, \tag{10}$$
$$\alpha_k t\delta^2+b_k\delta+t^2+t^3+t\sigma\leq 0. \tag{11}$$

Clearly, this corresponds to a quadratic equation involving only one unknown δ. To guarantee that its roots are positive, we require

$$0>b_k, \tag{12}$$
$$0\leq b_k^2-4\alpha_k t(t^2+t^3+t\sigma). \tag{13}$$

Table 1: Numerical demonstration of the conditions (3)

α_k	0.5		0.6		0.7		0.8		0.9	
t	0.30	0.40	0.20	0.30	0.20	0.25	0.10	0.15	0.08	0.09
σ	–	–	–	–	–	–	–	–	–	–

Table 2: Numerical demonstration of our suggested conditions

	α_k	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90
Condition (5)	t	$< 1/3$								
Condition (9)	t		0.303	0.274	0.245	0.216	0.186	0.154	0.121	0.085

For given α_k and t, we first check whether or not there exist some positive number σ such that (12) and (13) hold. If so, we still need to further check whether or not the relation

$$\delta_+(\alpha_k, t, \sigma) > \frac{t^2 + t^3 + t\sigma}{1 - t^2} \tag{14}$$

in (3) holds as well for these given α_k and t, where we have made use of δ_+ to stand for the biggest root of the quadratic equation involved in (11).

For simplicity, we focus on the cases of $\alpha_k \equiv 0.5, 0.6, 0.7, 0.8, 0.9$. So, the condition (12) becomes

$$\begin{aligned}
0 &> 3t^2 + t + \sigma - 1 &\Rightarrow \quad 0 &> 3t^2 + t - 1 &\Rightarrow \quad t &< 0.435,\\
0 &> 16t^2 + 6t + 6\sigma - 4 &\Rightarrow \quad 0 &> 16t^2 + 6t - 4 &\Rightarrow \quad t &< 0.346,\\
0 &> 17t^2 + 7t + 7\sigma - 3 &\Rightarrow \quad 0 &> 17t^2 + 7t - 3 &\Rightarrow \quad t &< 0.261,\\
0 &> 18t^2 + 8t + 8\sigma - 2 &\Rightarrow \quad 0 &> 18t^2 + 8t - 2 &\Rightarrow \quad t &< 0.178,\\
0 &> 19t^2 + 9t + 9\sigma - 1 &\Rightarrow \quad 0 &> 19t^2 + 9t - 1 &\Rightarrow \quad t &< 0.093,
\end{aligned}$$

respectively, and the corresponding results are listed in the following Table 1, where the notation – means that, for given α_k and t, there is no any positive number σ such that (12), (13) and (14) hold simultaneously.

For comparison, we also give numerical demonstration of our suggested conditions – either (5) or (9) (with $\varepsilon = 0.0001$) in the following Table 2.

From Table 1 and Table 2, we can see that, for our suggested conditions, (5) admits the case of $\alpha_k \equiv 0.5$ and $t = 0.33$ and (9) admits the cases of

$$\alpha_k \equiv 0.6, 0.7, 0.8, 0.9 \quad \text{and} \quad t = 0.27, 0.21, 0.15, 0.08,$$

respectively. In contrast, the conditions (12), (13) and (14) rule out all of them and thus seem restrictive.

Remark 4.4 By the way, in the case of $\alpha_k \in (0, 0.5)$, we make use of (8) to consider the following equation

$$\left(\frac{1}{\alpha_{k+1}} - 2\right) t^2 - \left(\frac{1}{\alpha_k} + \frac{1}{\alpha_{k+1}} - 1\right) t + \frac{1}{\alpha_k} - 1 - \varepsilon = 0.$$

Notice that $\frac{1}{\alpha_{k+1}} - 2 > 0$ due to $\alpha_{k+1} < 0.5$. Since

$$\left(\frac{1}{\alpha_k} + \frac{1}{\alpha_{k+1}} - 1\right)^2 - 4\left(\frac{1}{\alpha_{k+1}} - 2\right)\left(\frac{1}{\alpha_k} - 1\right) > 0,$$

we set Δ_k to be the positive square root of

$$\left(\frac{1}{\alpha_k} + \frac{1}{\alpha_{k+1}} - 1\right)^2 - 4\left(\frac{1}{\alpha_{k+1}} - 2\right)\left(\frac{1}{\alpha_k} - 1 - \varepsilon\right).$$

Thus, the equation above has two different positive roots

$$t_{\pm} := \frac{1}{2} \frac{1}{\frac{1}{\alpha_{k+1}} - 2} \left(\frac{1}{\alpha_k} + \frac{1}{\alpha_{k+1}} - 1 \pm \Delta_k\right).$$

If we further assume that

$$\frac{1}{\alpha_k} - \frac{1}{\alpha_{k+1}} + 3 > 0, \tag{15}$$

which is equivalent to saying that

$$\frac{1}{2} \frac{1}{\frac{1}{\alpha_{k+1}} - 2} \left(\frac{1}{\alpha_k} + \frac{1}{\alpha_{k+1}} - 1\right) > 1,$$

then it can be easily seen that the positive roots $t_- < 1$ and $t_+ > 1$. So, such t_- can serve as an upper bound of the inertial factors. Note that the assumption above (15) is mild since it holds automatically for any given constant in the interval $(0, 0.5)$.

As a direct consequence of Theorem 4.1, we give the following result.

Corollary 1 *Let $\mathcal{C} \subseteq \mathcal{H}$ be a nonempty closed affine subset. Assume that $T : \mathcal{C} \to \mathcal{C}$ is non-expansive and there exists at least one fixed point. For any given starting points x_{-1}, x_0 in the set $\mathcal{C}$, consider the recursive relation*

$$x_{k+1} = \frac{1}{2}(I + T)(x_k + t_k(x_k - x_{k-1})), \quad k = 0, 1, ...$$

If the inertial sequence satisfies

$$t_0 = 0, \quad 0 \le t_k \le t_{k+1} \le t < 1/3,$$

then the sequence $\{x_k\}$ converges weakly to a fixed point of T.

5 Implications

In this section, we discuss some implications of our suggested results on inertial Krasnosel'skiĭ-Mann iteration to (generalized) proximal point algorithm and the Peaceman/Douglas-Rachford splitting method of Lions and Mercier.

First, we consider the problem of solving monotone inclusion $0 \in A(x)$ in Hilbert spaces, where A is maximal monotone. Since $2J_{\lambda A} - I$ is non-expansive, taking $T := 2J_{\lambda A} - I$ in the Krasnosel'skiĭ-Mann iteration yields

$$x_{k+1} = (I - 2\alpha_k(I - J_{\lambda A}))\,(x_k), \quad k = 0, 1, ...,$$

where $\alpha_k \in (0, 1)$. This is a generalized proximal point algorithm, and the case $\alpha_k \equiv 0.5$ corresponds to the proximal point algorithm. Thus, we can get an inertial version of this generalized proximal point algorithm

$$x_{k+1} = (I - 2\alpha_k(I - J_{\lambda A}))\,(x_k + t_k(x_k - x_{k-1})), \quad k = 0, 1, ...,$$

where $\alpha_k \in (0, 1)$. If we set $\alpha_k \equiv 0.5$, then the recursive formula above reduces to

$$x_{k+1} = J_{\lambda A}(x_k + t_k(x_k - x_{k-1})), \quad k = 0, 1, ...,$$

which coincides with the one suggested in [1], with the scaling factor being some constant number. Interestingly, the conditions on the inertial sequence (5) becomes

$$0 \le t_k \le t_{k+1} \le t < 1/3,$$

and thus they also coincide with the ones stated in [1, Proposition 2.1]. For other discussions in the setting of convex minimization, the reader may consult a recent work [4]. Notice that that the proximal point algorithm comes from an implicit discretization of the first-order steepest descent method, while its inertial version is a discrete form of a second-order dissipative dynamical system alternatively called "heavy ball with friction" [1].

Next, we consider the problem of solving monotone inclusion $0 \in A(x) + B(x)$ in Hilbert spaces, where A, B are maximal monotone. For any given starting point $u_0 \in \mathcal{H}$, the corresponding Peaceman/Douglas-Rachford splitting method reads

$$u_{k+1} = (1 - w)(2J_{\lambda B} - I)(2J_{\lambda A} - I)(u_k) + w\,[J_{\lambda B}(2J_{\lambda A} - I) + I - J_{\lambda A}]\,(u_k)$$

for $k = 0, 1, ...$, where the co-efficient $w \in [0, 1]$ was introduce by Varga; see [25, Sect. 1.3] for more details. When $w = 0, 1$, it corresponds to the Peaceman-Rachford splitting method and the Douglas-Rachford splitting method, respectively.

Denote $R_{\lambda A} := 2J_{\lambda A} - I$. Then

$$\begin{aligned} &(1-w)(2J_{\lambda B}-I)(2J_{\lambda A}-I) + w\left[J_{\lambda B}(2J_{\lambda A}-I) + I - J_{\lambda A}\right] \\ &= (1-w)R_{\lambda B}R_{\lambda A} + w\left[\frac{I+R_{\lambda B}}{2}R_{\lambda A} + I - \frac{I+R_{\lambda A}}{2}\right] \\ &= \left(1-(1-\frac{w}{2})\right)I + (1-\frac{w}{2})R_{\lambda B}R_{\lambda A} \\ &= (1-\alpha)I + \alpha R_{\lambda B}R_{\lambda A}, \end{aligned}$$

where $\alpha := 1 - \frac{w}{2}$ and thus α here belongs to the interval $[0.5, 1]$. So, we follow [26] to demonstrate that the Peaceman/Douglas-Rachford splitting method can be interpreted as a special case of the Krasnosel'skiĭ-Mann iteration above, in which the co-efficient varies in the interval $[0.5, 1]$ as just mentioned.

From these facts, we may add an inertial term to the Peaceman/Douglas-Rachford splitting method. Thus, we get the following inertial version.

$$u_{k+1} = \left[(1-\alpha_k)I + \alpha_k(2J_{\lambda B}-I)(2J_{\lambda A}-I)\right](u_k + t_k(u_k - u_{k-1})) \tag{16}$$

for $k = 0, 1, ...,$ where $\lambda > 0$ is any given positive number, and the inertial sequence satisfies

$$\alpha_k \in [0.5, 1-\varepsilon], \quad t_0 = 0,\ 0 \le t_k \le t_{k+1} \le \frac{1-(1+\varepsilon)\alpha_k}{1+\alpha_k}, \tag{17}$$

where ε is any given sufficiently small positive number. Note that, when $\alpha_k \equiv 0.5$, the conditions above become

$$t_0 = 0,\ 0 \le t_k \le t_{k+1} \le t < 1/3.$$

This just corresponds to the inertial Douglas-Rachford splitting. Notice that, in this case, the corresponding inertial Douglas-Rachford splitting method is also a special case of the one given in Corollary 1.

As to its convergence behaviours, we have the following results.

Theorem 5.1 *Let $A, B : \mathcal{H} \rightrightarrows \mathcal{H}$ be maximal monotone such that*

$$0 \in A(x) + B(x) \tag{18}$$

has at least one solution. Assume that $domA \supseteq domB \neq \emptyset$. Then for any given starting points $u_{-1}, u_0 \in \mathcal{H}$, the sequence generated by the inertial Peaceman/Douglas-Rachford splitting method (16) is weakly convergent to some point u in $domB$ such that

$$(2J_{\lambda B} - I)(2J_{\lambda A} - I)(u) = u \tag{19}$$

and $x := J_{\lambda A}(u)$ solves the problem (18) above.

Proof In fact, since the inertial Peaceman/Douglas-Rachford splitting method (16) is a special case of the inertial Krasnosel'skiĭ-Mann iteration (2), of course the resulting sequence of iterates converges weakly to a fixed point u of the non-expansive operator $(2J_{\lambda B} - I)(2J_{\lambda A} - I)$. This shows that (19) holds.

It remains to prove that $x := J_{\lambda A}(u)$ solves (18). In views of (19), we have

$$(2J_{\lambda B} - I)(2x - u) = u \quad \Rightarrow \quad J_{\lambda B}(2x - u) = x$$

which implies

$$2x - u \in x + \lambda B(x).$$

Adding this to

$$x := J_{\lambda A}(u) \quad \Rightarrow \quad u \in x + \lambda A(x)$$

yields $0 \in A(x) + B(x)$. The proof is complete.

For further applications of the Peaceman/Douglas-Rachford splitting method, we refer to [16, 11] for more details.

Be care of that the inertial Peaceman/Douglas-Rachford splitting method (16) can be implemented as follows. Choose $u_{-1}, u_0 \in \mathcal{H}$, $x_{-1} \in \mathcal{H}$, $\lambda_0 > 0$. First choose t_k and calculate

$$v_k := u_k + t_k(u_k - u_{k-1}).$$

Then, for known λ_k, solve in order

$$(I + \lambda_k A)(x) \ni v_k, \quad (I + \lambda_k B)(y) \ni 2x_k - v_k \tag{20}$$

to get x_k, y_k, respectively. Finally, choose $\alpha_k \in [0.5, 1)$ and compute

$$u_{k+1} = v_k - 2\alpha_k(x_k - y_k)$$

to return the new iterate. Notice that: (i) The conditions on both α_k and the inertial sequence are given by the Table 2. (ii) At k-th iteration, for known x_{k-1}, x_k, λ_k, we can mimic [13] to self-adaptively get the new scaling factor λ_{k+1}; see (21) below. (iii) The subproblems (20) can be solved approximately with summable errors and it is not difficult to confirm that the method's weak convergence remains valid.

These conditions (17) above are widely different from counterparts stated in [5, Theorem 8], which are obtained from (3) directly. In particular, we are able to require $\alpha_k \geq 0.5$ in a desirable and natural way. This is because our derivation is from a convex combination of the Peaceman-Rachford splitting and the Douglas-Rachford splitting introduced by Varga whereas the authors of [5] did not exclude the undesirable case of $\alpha_k \in (0, 0.5)$ because their derivation is only from the Douglas-Rachford splitting itself. Based on this point, we called (16) inertial Peaceman/Douglas-Rachford splitting method so that one can distinguish such a name from inertial Douglas-Rachford splitting method suggested in [5] recently.

At the end of this section, we shall specially stress that our ultimate goal of studying an inertial version of the standard Krasnosel'skiĭ-Mann iteration is to provide the theoretical basis for a practical Peaceman/Douglas-Rachford splitting method with inertial effects, which is being developed. This explains the reason why our suggested conditions assume $\alpha_k \geq 0.5$, which well covers the case of the Peaceman/Douglas-Rachford splitting. For the inertial Krasnosel'skiĭ-Mann iteration with variable non-expansive operators, we refer to [27] for pertinent discussions. Yet, as far as our ultimate goal is concerned, it does not pay to be complicated.

6 Rudimentary experiments

In this section, our primary goal is to further check inertial effects when the aforementioned inertial Douglas-Rachford splitting method was applied to solving our test problems and to confirm that our suggested conditions are practically useful. The reason why we mainly focus on the Douglas-Rachford splitting is that the corresponding sequence of inertial factors has a relatively large upper bound so that adding inertial terms might significantly affect the method's numerical performance.

All numerical experiments were run in MATLAB R2014a (8.3.0.532) with 64-bit (win64) on a desktop computer with an Intel(R) Core(TM) i5-7400 CPU 3.00 GHz and 8.00 GB of RAM. The operating system is Windows 10.

For the (inertial) Douglas-Rachford splitting method, which corresponds to the case of $\alpha_k = 0.5$, we mimicked [13] to update the involved λ_k in the following way. At k-th iteration, for known λ_k, first calculate

$$\phi_k := \frac{\lambda_k \|A(x_k) - A(x_{k-1})\|}{\|x_k - x_{k-1}\|},$$

where we have further assumed that A is continuous. Then, update λ_k via

$$\lambda_{k+1} = \begin{cases} 1.5\lambda_k, & \text{if } \phi_k \leq 0.5, \\ 0.5\lambda_k, & \text{if } \phi_k \geq 2, \\ \lambda_k, & \text{otherwise.} \end{cases} \tag{21}$$

In the case of A being further linear, we suggested choosing $\lambda_0 \leq 1$ and to be slightly larger than $1/d_{\max}$ (if possible), where $d_{\max}$ is the largest entry in the diagonal of the involved matrix. Here there is no worry about loss of convergence since it can be merely done for the first N (say, $N = 100$) iterations.

Henceforth, we made use of DR and iDR to stand for the Douglas-Rachford splitting method and the inertial version, respectively.

Our first test problem comes from [6, Problem (8.6)] and is about portfolio selection by Markowitz's mean-variance model. Denote by x_1, x_2, x_3 the proportion of the total funds invested in Stocks, Bonds and Money Market, respectively. Consider the problem of minimizing the financial risk (that corresponds to variance):

$$0.02778x_1^2+2\cdot 0.00387x_1x_2+2\cdot 0.00021x_1x_3+0.01112x_2^2-2\cdot 0.00020x_2x_3+0.00115x_3^2,$$

where the constraint set $\mathcal{C}$ is given by

$$\begin{aligned}&0.1073x_1 + 0.0737x_2 + 0.0627x_3 \geq 8\%,\\ &x_1 + x_2 + x_3 = 1,\\ &x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0.\end{aligned}$$

Roughly speaking, the investors wish to bear as little risk as possible for a given minimum rate 8% of expected return.

For this test problem, it can be reformulated into

$$0 \in Qx + \partial\delta_{\mathcal{C}}(x),$$

where $x = (x_1, x_2, x_3)^T$ and

$$Q = 2\begin{pmatrix} 0.02778 & 0.00387 & 0.00021\\ 0.00387 & 0.01112 & -0.00020\\ 0.00021 & -0.00020 & 0.00115\end{pmatrix}.$$

Set

$$A(x) := Qx,\quad B(x) := \partial\delta_{\mathcal{C}}(x).$$

Then, we applied iDR to solving this problem. In practical implementations, we chose

$$u_{-1} = zeros(3,1),\quad u_0 = ones(3,1),\quad x_{-1} = u_{-1},\quad \lambda_0 = 1.$$

Moreover, we adopted (21) as a way of updating λ_k. At each iteration, we had to deal with usual projection$P_{\mathcal{C}} = (I+\lambda_k\partial\delta_{\mathcal{C}})^{-1}$ onto the constraint set and we made use of Matlab solver quadprog to resolve it. The corresponding numerical results were reported in Figure 1, where $e_k := x_k - P_{\mathcal{C}}[x_k - Qx_k]$ stands for residual.

Our second test problem is monotone variational inequality problem, which is from [7] and is a user-optimized traffic pattern for the simple network with only two nodes x, y and five links a_1, a_2, a_3, b_1, b_2, where a_1, a_2, a_3 are directed from x to y and b_1, b_2 are the return of a_1, a_2, respectively.

The travel cost function and the constraint set are given by

$$F(x) = \begin{pmatrix} 10 & 0 & 0 & 5 & 0\\ 0 & 15 & 0 & 0 & 5\\ 0 & 0 & 20 & 0 & 0\\ 2 & 0 & 0 & 20 & 0\\ 0 & 1 & 0 & 0 & 25\end{pmatrix}\begin{pmatrix} x_{a_1}\\ x_{a_2}\\ x_{a_3}\\ x_{b_1}\\ x_{b_2}\end{pmatrix} + \begin{pmatrix} 1000\\ 950\\ 3000\\ 1000\\ 1300\end{pmatrix},$$

Figure 1. Numerical results on the first test problem

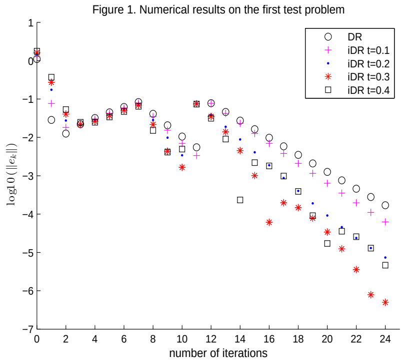
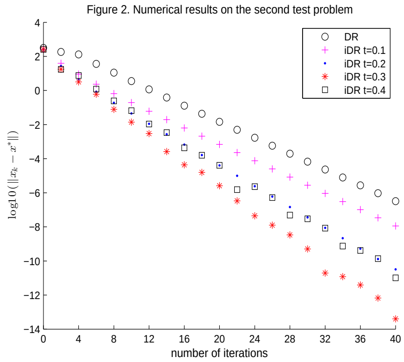


Figure 2. Numerical results on the second test problem



and

$$\mathcal{C} := \{x \in R^5 : x \geq 0, x_{a_1} + x_{a_2} + x_{a_3} = 210, x_{b_1} + x_{b_2} = 120\},$$

respectively. As we know, $x^* = (120, 90, 0, 70, 50)^T$ is its unique solution.

Set

$$A(x) := F(x), \quad B(x) := \partial\delta_{\mathcal{C}}(x).$$

Then, we applied iDR to solving this problem. In practical implementations, we chose

$$u_{-1} = zeros(5,1), \quad u_0 = ones(5,1), \quad x_{-1} = u_{-1}.$$

As to λ_0, we suggested choosing $\lambda_0 \leq 1$ and λ_0 to be slightly larger than $1/25$, where 25 is the largest entry in the diagonal of the involved matrix. Thus, we simply chose $\lambda_0 = 0.1$. Moreover, we adopted (21) as a way of updating λ_k. At each iteration, we had to deal with usual projection $P_{\mathcal{C}} = (I + \lambda_k \partial\delta_{\mathcal{C}})^{-1}$ onto the constraint set and we made use of Matlab solver quadprog to resolve it. The corresponding numerical results were reported in Figure 2.

From Figures 1 and 2, we can see that iDR with $t_k \equiv 0.3$ really outperformed both DR and iDR with $t \in \{0.1, 0.2, 0.4\}$ for our test problems. Specifically speaking, for given numbers of iterations, iDR with $t_k \equiv 0.3$ tends to achieve higher accuracy. Be aware that, as a good choice of the inertial factors, $t_k \equiv 0.3$ is ruled out by existing conditions and admissible for our suggested ones.

7 Conclusions

In this article, we have considered the Krasnosel'skiĭ-Mann iteration for approximating a fixed point of any given non-expansive operator in real Hilbert spaces. For an inertial version proposed by Maingé recently, we have suggested new conditions on the inertial factors to ensure weak convergence. They are free of iterates and depend on the original coefficient of the Krasnosel'skiĭ-Mann iteration. Their appealing property is that, at each iteration, one may choose the inertial factors dynamically. Our analysis of inertial Krasnosel'skiĭ-Mann iteration covers that of an inertial proximal point algorithm, in which case our suggested conditions coincide with the ones in [1, Proposition 2.1] in a desirable way. Furthermore, we have discussed an implication to the Peaceman/Douglas-Rachford splitting and have confirmed via rudimentary numerical experiments that our suggested conditions are practically useful for the inertial Douglas-Rachford splitting.

Very recently, these new techniques of designing inertial factors have been used by X.H. Yu and X. Zhu for splitting methods for monotone inclusions of three operators in their individual thesis, Zhengzhou University, and it would be interesting to report these new findings in the accompanying article.

Acknowledgments The author is grateful to Xiao Zhu for her help in doing numerical experiments.

Disclosure statement

No potential conflict of interest was reported by the author.

References

1. Alvarez, F. and Attouch, H.: An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping, Set-Val Anal. 9(1), 3-11 (2001)
2. Aubin, J.P., Ekeland, I.: Applied Nonlinear Analysis, Interscience Publications: John Wiley, New York, 1984.
3. Brézis, H., Lions, P.L.: Produits infinis de resolvantes, Israel J Math. 29, 329-345 (1978)
4. Cabot, A., Frankel, P.: Asymptotics for some proximal-like method involving inertia and memory aspects, Set-Valued Anal. 19, 59-74 (2011)
5. Bot, R.I., Csetnek, E.R., Hendrich, C.: Inertial Douglas-Rachford splitting for monotone inclusion problems, Appl Math Comput. 256(1), 472-487 (2015)
6. Cornuejols, G., Tütüncü, R.: Optimization Methods in Finance, Cambridge Univeristy, 2007
7. Dafermos, S.: Traffic equilibrium and variational inequalities, Transp. Sci. 14, 42-54 (1980)
8. Dong, Q. L., Cho, Y. J., Zhong, L. L., Rassias, Th. M.: Inertial projection and contraction algorithms for variational inequalities. J. Global Optim. 70(3), 687-704 (2018)

9. Dong, Y.D.: The proximal point algorithm revisited, J Optim Theory Appl. 161, 478-489 (2014)
10. Dong, Y.D.: Comments on "the proximal point algorithm revisited", J Optim Theory Appl. 166, 343-349 (2015)
11. Dong, Y.D.: Douglas-Rachford splitting method for semi-definite programming, J Appl Math Comput. 51, 569-591 (2016)
12. Dong, Y.D., Fang, B.B., Wang, X.Y.: New inertial factors of a splitting method for monotone inclusions, preprint (2018). Avialable at http://www.optimization-online.org/DBFILE/2018/04/6576.pdf
13. Dong, Y.D., Fischer, A.: A family of operator splitting methods revisited, Nonlinear Anal. 72, 4307-4315 (2010)
14. Douglas J., Rachford, H.H.: On the numerical solution of the heat conduction problems in two and three space variables, Trans Am Math Soc. 82, 421-439 (1956)
15. Eckstein, J., Bertsekas, D.P.: On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators, Math Program. 55(3), 293-318 (1992)
16. Fukushima, M.: The primal douglas-rachford splitting algorithm for a class of monotone mappings with application to the traffic equilibrium problem, Math. Program. 72(1), 1-15 (1996)
17. Genel, A., Lindenstrauss, J.: An example concerning fixed points, Isr J Math. 22(1), 81-86 (1975)
18. Gol'shtein, E.G., Tret'yakov, N.V.: Modified Lagrangians in convex programming and their generalizations, Math Program Study. 10, 86–97 (1979)
19. Groetsch, C.W.: A note on segmenting Mann iterates, J Math Anal Appl. 40(2), 369-372 (1972)
20. Güler, O.: On the convergence of the proximal point algorithm for convex minimization, SIAM J Control Optim. 29, 403-419 (1991)
21. Ishikawa, S.: Fixed points by a new iteration, Proc Amer Math Soc. 44, 147-150 (1974)
22. Kanzow, C., Shehu, Y.: Generalized Krasnoselskii-Mann-type iterations for nonexpansive mappings in Hilbert spaces, Comput Optim Appl. 67(3), 595-620 (2017)
23. Kim, T.H., Xu, H.-K.: Strong convergence of modified Mann iterations, Nonlinear Anal. 61, 51-60 (2005)
24. Krasnosel'skiĭ, M.A.: Two remarks on the method of successive approximations, Uspekhi Mat Nauk. 10, 123-127 (1955)
25. Lions, P.L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators, SIAM J Numer Anal. 16, 964-979 (1979)
26. Lawrence J., Spingarn, J.E.: On fixed points of non-expansive piecewise isometric mappings, P Lond Math Soc. 55(1987), pp. 605–624.

27. Maingé, P.: Convergence theorems for inertial KM-type algorithms, J Comput Appl Math. 219(1), 223-236 (2008)
28. Mann, W.R.: Mean value methods in iteration, Bull Am Math Soc. 4, 506-510 (1953)
29. Martinet, B.: Regularisation d'inéquations variationelles par approximations successives, Rev Fr d'Informatique Recherche Opér. 4, 154-158 (1970)
30. Matsushita, S.Y.: On the convergence rate of the Krasnosel'skiĭ-Mann iteration, B Austr Math Soc. 96(1), 162-170 (2017)
31. Minty, G.J.: Monotone (nonlinear) operators in Hilbert space, Duke Math J. 29,341-346 (1962)
32. Minty, G.J.: On the monotonicity of the gradient of a convex function, Pac J Math. 14, 243–247 (1964)
33. Pazy, A.: Asymptotic behavior of contractions in Hilbert space, Israel J Math. 9, 235-240 (1971)
34. Peaceman, D., Rachford, H.H.: The numerical solution of parabolic and elliptic differential equations, J Soc Indust Appl Math. 3, 28-41 (1955)
35. Rockafellar, R.T.: Monotone operators and the proximal point algorithm, SIAM J Control Optim. 14, 877–898 (1976)
36. Thong, D.V., Hieu, D.V.: Inertial extragradient algorithms for strongly pseudomonotone variational inequalities. J. Comput. Appl. Math. 341, 80-98 (2018)

A

Below we give a short proof of Lemma 4.2 for completeness. Note that we no longer follow [5] to introduce the factor ρ_k there.

Proof Rewrite (2) as

$$y_k = x_k + t_k(x_k - x_{k-1}), \tag{22}$$
$$x_{k+1} = (1-\alpha_k)y_k + \alpha_k T(y_k), \quad k = 0, 1, ..., . \tag{23}$$

Since z is a fixed point of T, i.e., $T(z) = z$, it follows from (23) that

$$x_{k+1} - z = \alpha_k(T(y_k) - T(z)) + (1-\alpha_k)(y_k - z).$$

Combing this with the following well-known identity [21]

$$\|\alpha u + (1-\alpha)v\|^2 = \alpha\|u\|^2 + (1-\alpha)\|v\|^2 - \alpha(1-\alpha)\|u-v\|^2, \;\; \forall u, v \in \mathcal{H}$$

for all real number α yields

$$\begin{aligned} &\|x_{k+1} - z\|^2 \\ &= \|\alpha_k(T(y_k) - T(z)) + (1-\alpha_k)(y_k - z)\|^2 \\ &= \alpha_k\|T(y_k) - T(z)\|^2 + (1-\alpha_k)\|y_k - z\|^2 - \alpha_k(1-\alpha_k)\|T(y_k) - y_k\|^2 \\ &\le \alpha_k\|y_k - z\|^2 + (1-\alpha_k)\|y_k - z\|^2 - \alpha_k(1-\alpha_k)\|T(y_k) - y_k\|^2 \\ &= \|y_k - z\|^2 - \alpha_k(1-\alpha_k)\|T(y_k) - y_k\|^2, \end{aligned} \tag{24}$$

where the inequality follows from that T is non-expansive.

In view of the identity above and (22), we have

$$\begin{aligned}\|y_k - z\|^2 &= \|(1+t_k)(x_k - z) - t_k(x_{k-1} - z)\|^2 \\ &= (1+t_k)\|x_k - z\|^2 - t_k\|x_{k-1} - z\|^2 + (1+t_k)t_k\|x_k - x_{k-1}\|^2.\end{aligned}$$

On the other hand, it is easy to see that

$$\begin{aligned}&\|T(y_k) - y_k\|^2 \\ =& \frac{1}{\alpha_k^2}\|(x_{k+1} - x_k) - t_k(x_k - x_{k-1})\|^2 \\ =& \frac{1}{\alpha_k^2}\|x_{k+1} - x_k\|^2 + \frac{t_k^2}{\alpha_k^2}\|x_k - x_{k-1}\|^2 - 2\frac{t_k}{\alpha_k^2}\langle x_{k+1} - x_k, x_k - x_{k-1}\rangle \qquad (25) \\ \geq& \frac{1}{\alpha_k^2}\|x_{k+1} - x_k\|^2 + \frac{t_k^2}{\alpha_k^2}\|x_k - x_{k-1}\|^2 - \frac{t_k}{\alpha_k^2}\left(\|x_{k+1} - x_k\|^2 + \|x_k - x_{k-1}\|^2\right) \\ =& \frac{1}{\alpha_k^2}(1-t_k)\|x_{k+1} - x_k\|^2 + \frac{1}{\alpha_k^2}(t_k^2 - t_k)\|x_k - x_{k-1}\|^2.\end{aligned}$$

Finally, making use of these two relations to bound (24) yields the desired results.