Journal of Global Optimization
https://doi.org/10.1007/s10898-024-01425-w

Novel algorithms based on forward-backward splitting technique: effective methods for regression and classification

Yunus Atalan[1] · **Emirhan Hacıoğlu**[2] · **Müzeyyen Ertürk**[3] · **Faik Gürsoy**[3] · **Gradimir V. Milovanović**[4,5]

Received: 15 May 2023 / Accepted: 29 July 2024

Abstract
In this paper, we introduce two novel forward-backward splitting algorithms (FBSAs) for nonsmooth convex minimization. We provide a thorough convergence analysis, emphasizing the new algorithms and contrasting them with existing ones. Our findings are validated through a numerical example. The practical utility of these algorithms in real-world applications, including machine learning for tasks such as classification, regression, and image deblurring reveal that these algorithms consistently approach optimal solutions with fewer iterations, highlighting their efficiency in real-world scenarios.

Keywords Iterative algorithm · Variational inequalities · Relaxed (κ,ω)-cocoercive mappings · Nonexpansive mappings

1 Introduction

A large number of problems encountered in distinct disciplines like variational inequalities, mini-max, optimization problems, etc. can be modelled as the finding zero points of sum of

✉ Emirhan Hacıoğlu
emirhanhacioglu@hotmail.com

Yunus Atalan
yunusatalan@aksaray.edu.tr

Müzeyyen Ertürk
merturk3263@gmail.com

Faik Gürsoy
faikgursoy02@hotmail.com

Gradimir V. Milovanović
gvm@mi.sanu.ac.rs

1 Department of Mathematics, Aksaray University, 68100 Aksaray, Turkey

2 Department of Mathematics, Trakya University, 22030 Edirne, Turkey

3 Department of Mathematics, Adiyaman University, 02040 Adiyaman, Turkey

4 Serbian Academy of Sciences and Arts, 11000 Beograd, Serbia

5 Faculty of Science and Mathematics, University of Niš, Niš 18000, Serbia

Published online: 07 August 2024

two nonlinear operators in a Hilbert space $\mathcal{H}$ as follows (see [1–4]):

$$0 \in (\Gamma_1 + \Gamma_2)x, \tag{1}$$

in which $\Gamma_1 : \mathcal{H} \to 2^{\mathcal{H}}$ is a multivalued mapping and $\Gamma_2 : \mathcal{H} \to \mathcal{H}$ is a mapping. Since most of the problems encountered in applied and computational fields such as signal processing, machine learning, and image recovery can be modeled as the inclusion composed of the sum of two nonlinear operators, splitting methods have attracted the attention of many researchers (see [5–7]). These methods are very useful as they process each operator separately instead of processing the sum of the operators. Here, it is possible to divide the handled process into two parts as the case in which the value of the operator is determined by *forward* calculations and the case in which the operator's resolvent is determined by *backward* calculations.

For the identity operator $I : \mathcal{H} \to \mathcal{H}$, a mapping $J^{\Gamma_1} = (I + \Gamma_1)^{-1}$ is named the resolvent of Γ_1.

The forward-backward algorithm (FBA) is one of the well-known splitting methods when Γ_2 in (1) is single-valued. Each step of this algorithm combines forward calculation of Γ_2 and backward calculation of Γ_1 to reach the solution of (1). A standard forward-backward splitting algorithm (FBSA) is formulated as follows (see [8]):

$$\textbf{Algorithm FBSA}: \; u_{n+1} = (I + \varrho\Gamma_1)^{-1}(I - \varrho\Gamma_2)u_n, \quad n \geq 0,$$

with an appropriate $u_0 \in \mathcal{H}$ and $\varrho > 0$. This algorithm involves both the proximal point algorithm and the gradient methods (see [9, 10]). Lions and Mercier [8] proposed two splitting algorithms for the evolutionary and stationary problems composed of the sum of two multivalued monotone mappings:

$$\textbf{Splitting algorithm 1}: \; v_{n+1} = (2J_\varrho^{\Gamma_1} - I)(2J_\varrho^{\Gamma_2} - I)v_n, \;\; n \geq 0;$$

$$\textbf{Splitting algorithm 2}: \; v_{n+1} = J_\varrho^{\Gamma_1}(2J_\varrho^{\Gamma_2} - I)v_n + (I - J_\varrho^{\Gamma_2})v_n, \;\; n \geq 0,$$

in which $C = \Gamma_1$ or $C = \Gamma_2$ such that $J_\varrho^C = (I + \varrho C)^{-1}$. They proved various convergence results for these algorithms and implemented these results to the minimization problems and to the obstacle problem.

Due to their diverse range of applications, these algorithms have proven to be highly valuable. Consequently, a substantial number of FBSA have been introduced and thoroughly investigated within the context of monotone operators (see [11–19]) and references therein.

In the next section, we will recall some recently defined algorithms in this context and present some key facts which will be used to derive the main results of this exposition.

2 Preliminaries and formulations

Throughout this exposition, unless otherwise stated, $(\mathcal{H}, \|\cdot\|)$ stands for a real Hilbert space in which $\|\cdot\|$ is defined by the inner product $\langle \cdot, \cdot \rangle$, $\emptyset \neq C \subseteq \mathcal{H}$ denotes convex and closed set and $T : C \to C$ be a mapping.

Definition 1 A mapping $\Gamma_2 : \mathcal{H} \to \mathcal{H}$ is called: (i) *monotone* if

$$(\forall \nu_1, \nu_2 \in \mathcal{H}) \quad \langle \Gamma_2\nu_1 - \Gamma_2\nu_2, \nu_1 - \nu_2 \rangle \geq 0;$$

(ii) *ϕ-strongly monotone* if there exists a constant $\phi > 0$ such that

$$(\forall \nu_1, \nu_2 \in \mathcal{H}) \quad \langle \Gamma_2\nu_1 - \Gamma_2\nu_2, \nu_1 - \nu_2 \rangle \geq \phi \|\nu_1 - \nu_2\|^2;$$

(iii) ϕ-*inverse strongly monotone* (ϕ-ism) if there exists a constant $\phi > 0$ such that

$$(\forall \nu_1, \nu_2 \in \mathcal{H}) \quad \langle \Gamma_2 \nu_1 - \Gamma_2 \nu_2, \nu_1 - \nu_2 \rangle \geq \phi \left\| \Gamma_2 \nu_1 - \Gamma_2 \nu_2 \right\|^2 ;$$

(iv) *firmly nonexpansive* if

$$(\forall \nu_1, \nu_2 \in \mathcal{H}) \quad \langle \nu_1 - \nu_2, \Gamma_2 \nu_1 - \Gamma_2 \nu_2 \rangle \geq \left\| \Gamma_2 \nu_1 - \Gamma_2 \nu_2 \right\|^2 ;$$

(v) *expanding* if there exists a constant $h > 1$ such that

$$(\forall \nu_1, \nu_2 \in \mathcal{H}) \qquad \left\| \Gamma_2 \nu_1 - \Gamma_2 \nu_2 \right\| \geq h \left\| \nu_1 - \nu_2 \right\| .$$

Definition 2 A multivalued mapping $\Gamma_1 : \mathcal{H} \to 2^{\mathcal{H}}$ is called (i) *monotone* if for all $\varkappa_1, \varkappa_2 \in \mathcal{H}$, $\nu_1 \in \Gamma_1 \varkappa_1,\ \nu_2 \in \Gamma_1 \varkappa_2$, $\langle \nu_1 - \nu_2, \varkappa_1 - \varkappa_2 \rangle \geq 0$; (ii) *strongly monotone* if there exists a constant $\phi > 0$ such that for all $\varkappa_1, \varkappa_2 \in \mathcal{H}$, $\nu_1 \in \Gamma_1 \varkappa_1,\ \nu_2 \in \Gamma_1 \varkappa_2$,

$$\langle \nu_1 - \nu_2, \varkappa_1 - \varkappa_2 \rangle \geq \phi \left\| \varkappa_1 - \varkappa_2 \right\|^2 ;$$

(iii) *maximal monotone* if Γ_1 monotone and $(I + \varrho \Gamma_1)\mathcal{H} = \mathcal{H}$ holds for all $\varrho > 0$ in which I is the identity mapping in $\mathcal{H}$.

Let Γ_1 be a multivalued monotone mapping with the graph $\mathrm{Gr}(\Gamma_1)$. If a monotone mapping Γ_2 with graph $\mathrm{Gr}(\Gamma_2)$ such that $\mathrm{Gr}(\Gamma_1) \subset \mathrm{Gr}(\Gamma_2)$ cannot be found, then Γ_1 is called a maximal monotone mapping.

Remark 1 (i) The sum $\Gamma_1 + \Gamma_2$ is monotone, if Γ_1 and Γ_2 are monotone. (ii) A mapping $\kappa \Gamma_1$ for $\kappa \geq 0$ is monotone, if Γ_1 is monotone. (iii) A mapping Γ_1^{-1} is monotone, if Γ_1 is monotone.

Remark 2 [6] If $\Gamma_1 : \mathcal{H} \to 2^{\mathcal{H}}$ is a multivalued maximal monotone operator, then there exists a unique $p \in \mathcal{H}$ such that $x \in (I + \varrho \Gamma_1) p$ for each $x \in \mathcal{H}$ and $\varrho > 0$.

Definition 3 [20, p. 182] Let Γ_1 be a maximal monotone operator with resolvent $J_\varrho^{\Gamma_1} = (I + \varrho \Gamma_1)^{-1}$. For every $\varrho > 0$, the Yosida approximation of Γ_1 defined by

$$\Gamma_{1_\varrho} = \frac{1}{\varrho} \left(I - J_\varrho^{\Gamma_1} \right) .$$

Definition 4 [6] $P_C : \mathcal{H} \to C$ is called metric projection for $u \in \mathcal{H}$ such that $d(u, C) = \| u - P_C(u) \| = \inf \{ \| u - p \| : p \in C \}$, in which $P_C(u) \in C$ is the singleton set.

Remark 3 [21] P_C is firmly-nonexpansive of $\mathcal{H}$ and hence a nonexpansive mapping of $\mathcal{H}$.

We also use the following notation $J_\varrho^{\Gamma_1, \Gamma_2} = J_\varrho^{\Gamma_1}(I - \varrho \Gamma_2)$ (see [6]).

To advance research in applied and computational domains, it is crucial to develop new algorithms that exhibit improved convergence rates and to thoroughly investigate their qualitative properties. One such algorithm, known as the Normal-S algorithm, was introduced by Sahu [22]. Research has demonstrated that the Normal-S algorithm achieves a higher convergence rate compared to other well-known iteration algorithms such as Picard [23], Mann [24], Ishikawa [25], Noor [26], and S [27] for the class of contraction mappings (as detailed in [22]). Building on the successes of the Normal-S algorithm and the ongoing research in this area, Gursoy proposed the Picard-S algorithm (as described in [28]). The Picard-S algorithm is particularly intriguing and merits further exploration due to its ability to converge more rapidly. Importantly, it operates independently of the aforementioned algorithms and exhibits favorable behavior when applied to both contractive and nonexpansive mappings.

These properties make the Picard-S algorithm a promising solution for operator equations in various scientific and computational research scenarios. As a special case of Picard-S algorithm, Karakaya et al. proposed a PMP algorithm [29].

Sahu et al. [6] proposed below FBSAs based on several well-known algorithms to solve the problem (1):

$$\begin{aligned}
&\textbf{AlgorithmM}-\textbf{FBSA}:\\
&\quad v_{n+1}=(1-\eta_n^{(1)})v_n+\eta_n^{(1)}(J_\varrho^{\Gamma_1,\Gamma_2}v_n),\quad \eta_{n+1}^{(1)}\in[0,1];\\
&\textbf{AlgorithmS}-\textbf{FBSA}:\quad z_n=(1-\eta_n^{(3)})w_n+\eta_n^{(3)}(J_\varrho^{\Gamma_1,\Gamma_2}w_n),\\
&\quad w_{n+1}=(1-\eta_n^{(2)})J_\varrho^{\Gamma_1,\Gamma_2}w_n+\eta_n^{(2)}(J_\varrho^{\Gamma_1,\Gamma_2}z_n),\quad \eta_{n+1}^{(2)},\eta_{n+1}^{(3)}\in[0,1];
\end{aligned}$$

with appropriate $v_0, w_0 \in \mathcal{H}$, $\varrho > 0$ and the corresponding $\eta_0^{(i)} \in [0, 1]$ $(i = 1, 2, 3)$. Notice that in the case $\eta_n^{(1)} = 1$ for all n $(n \geq 0)$, M-FSBA transforms into P-FSBA [6]. Likewise, when $\eta_n^{(2)} = 1$ for all n $(n \geq 0)$, S-FSBA transforms into NS-FSBA [6].

Sahu et al. [6] obtained some convergence results; they also demonstrated theoretically and experimentally that these algorithms are faster than the classical FBA (forward-backward algorithm).

Drawing inspiration from the mentioned algorithms and their successes, we introduce two new FBSAs: PS-FBSA and PMP-FBSA, respectively:

$$\begin{aligned}
&\textbf{AlgorithmPS}-\textbf{FBSA}:\quad e_n=(1-\eta_n^{(6)})c_n+\eta_n^{(6)}(J_\varrho^{\Gamma_1,\Gamma_2}c_n),\\
&d_n=(1-\eta_n^{(5)})J_\varrho^{\Gamma_1,\Gamma_2}c_n+\eta_n^{(5)}(J_\varrho^{\Gamma_1,\Gamma_2}e_n),\quad c_{n+1}=J_\varrho^{\Gamma_1,\Gamma_2}d_n,\ \eta_n^{(5)},\eta_n^{(6)}\in[0,1];\\
&\textbf{AlgorithmPMP}-\textbf{FBSA}:\quad e_n=(J_\varrho^{\Gamma_1,\Gamma_2}c_n),\\
&d_n=(1-\eta_n^{(7)})e_n+\eta_n^{(7)}(J_\varrho^{\Gamma_1,\Gamma_2}e_n),\quad c_{n+1}=J_\varrho^{\Gamma_1,\Gamma_2}d_n,\quad \eta_{n+1}^{(7)}\in[0,1],
\end{aligned}$$

with appropriate $c_0 \in \mathcal{H}$, $\varrho > 0$ and the corresponding $\eta_0^{(i)} \in [0, 1]$ $(i = 5, 6, 7)$. Observe that if $\eta_n^{(6)} = 1$ for all $n \in \mathbb{N}$, then PS-FBSA reduces to PMP-FBSA.

In the present work, we establish strong convergence results for three algorithms: PS-FBSA, NS-FBSA, and S-FBSA. Specifically, we demonstrate the equivalence of convergence between PS-FBSA and NS-FBSA, and we provide a comparison of their convergence rates. We also applied them to solve the convex minimization problem. We show the validity of our findings through a nontrivial numerical example. As a practical application, in classification/regression of the machine learning problem and in image deblurring tasks, the performances of the newly proposed algorithms with that of their predecessors are compared. By conducting some numerical experiments, we confirm that these new algorithms converge to the optimum value with fewer steps, and we show that the newly proposed algorithms exhibit much better approximations to the optimum value even when fewer steps are used. We need the following facts to derive the main results of this exposition:

Lemma 1 *[30, Corollary 2.14] For $a_1, a_2 \in H$ and $\Lambda \in [0, 1]$, it holds that*

$$\|\Lambda a_1-(\Lambda-1)a_2\|^2=\Lambda\|a_1\|^2-(\Lambda-1)\|a_2\|^2+\Lambda(\Lambda-1)\|a_1-a_2\|^2.$$

Lemma 2 *[31] Let $\{\rho_n^{(\nu)}\}_{n=0}^{\infty}$, $\nu = 1, 2, 3$, be three nonnegative sequences. Assume that $\rho_n^{(2)} = o(\rho_n^{(3)})$, $\sum_{n=1}^{\infty} \rho_n^{(3)} = \infty$, and $\rho_n^{(3)} \in (0, 1)$ for all $n \geq n_0$. If $\rho_{n+1}^{(1)} \leq (1 - \rho_n^{(3)})\rho_n^{(1)} + \rho_n^{(2)}$, then $\lim_{n\to\infty} \rho_n^{(1)} = 0$.*

Lemma 3 *[32] Let* $\{\rho_n^{(\nu)}\}_{n=0}^{\infty}$, $\nu = 1, 2$, *be two nonnegative sequences. Assume that* $\lim_{n\to\infty} \rho_n^{(2)} = 0$ *and* $\mu \in (0, 1)$. *If* $\rho_{n+1}^{(1)} \le \mu\rho_n^{(1)} + \rho_n^{(2)}$, *then* $\lim_{n\to\infty} \rho_n^{(1)} = 0$.

Definition 5 (see [33]) If $\lim_{n\to\infty} \|\Theta_n^{(1)} - \Theta_1\|/\|\Theta_n^{(2)} - \Theta_2\| = 0$, where $\{\Theta_n^{(\nu)}\}_{n=0}^{\infty}$, $\nu = 1, 2$, are two sequences with $\lim_{n\to\infty} \Theta_n^{(\nu)} = \Theta_\nu$, $\nu = 1, 2$, then it is said that $\{\Theta_n^{(1)}\}_{n=0}^{\infty}$ converges faster than $\{\Theta_n^{(2)}\}_{n=0}^{\infty}$.

Definition 6 (see [34]) Let $\{\Theta_n^{(\nu)}\}_{n=0}^{\infty}$ and $\{\Pi_n^{(\nu)}\}_{n=0}^{\infty}$ ($\nu = 1, 2$) be four sequences, such that $\Pi_n^{(\nu)} \ge 0$ for each $n \in \mathbb{N}$, $\lim_{n\to\infty} \Theta_n^{(\nu)} = \Theta^*$, and $n \lim_{n\to\infty} \Pi_n^{(\nu)} = 0$, $\nu = 1, 2$. Suppose that for each $n \in \mathbb{N}$ the following error estimates are available (and the best possible [35]) $\|\Theta_n^{(\nu)} - \Theta^*\| \le \Pi_n^{(\nu)}$ ($\nu = 1, 2$). If $\{\Pi_n^{(1)}\}_{n=0}^{\infty}$ converges faster than $\{\Pi_n^{(2)}\}_{n=0}^{\infty}$ (in the sense of Definition 5), then we say that $\{\Theta_n^{(1)}\}_{n=0}^{\infty}$ converges to Θ^* faster than $\{\Theta_n^{(2)}\}_{n=0}^{\infty}$.

3 Convergence Analysis of forward–backward splitting algorithms

Let $\emptyset \ne C$ be a subset of $\mathcal{H}$, $\Gamma_2 : C \to \mathcal{H}$ a ϕ-ism and expanding operator, and $\Gamma_1 : C \to 2^{\mathcal{H}}$ a multivalued maximal monotone operator. Our purpose is to reach the solution $p \in \mathcal{H}$ such that $0 \in \Gamma_1 p + \Gamma_2 p$. For this purpose, we use the following fixed-point characterization:

$$0 \in (\Gamma_1 + \Gamma_2)p \Longleftrightarrow p = J_\varrho^{\Gamma_1}(I - \varrho\Gamma_2)p. \tag{2}$$

The following two propositions play a key role in the proof of the subsequent theorems.

Proposition 4 *(see Proposition 3.1 in [6]) For all* $u_1, u_2 \in \mathcal{H}$ *and* $\varrho > 0$, *we have*

$$\left\| J_\varrho^{\Gamma_1} u_1 - J_\varrho^{\Gamma_1} u_2 \right\|^2 \le \|u_1 - u_2\|^2,$$

in which $\Gamma_1 : \mathcal{H} \to 2^{\mathcal{H}}$ *is a multivalued maximal monotone operator.*

Proposition 5 *For* $\varrho > 0$ *and for all* $u_1, u_2 \in C$, *we have*

$$\|(I - \varrho\Gamma_2)u_1 - (I - \varrho\Gamma_2)u_2\| \le \theta\|u_1 - u_2\|,$$

in which $\theta = \sqrt{1 + \varrho^2/\phi^2 - 2\varrho\phi}$. *If* $\varrho < \phi\sqrt{2\varrho\phi} < \sqrt{\varrho^2 + \phi^2}$, *then* $\theta < 1$ *which implies that* $(I - \varrho\Gamma_2) : C \to H$ *is a contraction mapping.*

Proof Let $u_1, u_2 \in C$. For $\varrho > 0$, we have

$$\left\|(I - \varrho\Gamma_2)u_1 - (I - \varrho\Gamma_2)u_2\right\|^2 = \|u_1 - u_2\|^2 + \varrho^2 \|\Gamma_2 u_1 - \Gamma_2 u_2\|^2 - 2\varrho \langle \Gamma_2 u_1 - \Gamma_2 u_2, u_1 - u_2\rangle.$$

If (i) Γ_2 is ϕ-ism then, Γ_2 is $\frac{1}{\phi}$-Lipschitzian. Hence, we have

$$\varrho^2\|\Gamma_2 u_1 - \Gamma_2 u_2\|^2 \le \frac{\varrho^2}{\phi^2}\|u_1 - u_2\|^2;$$

If (ii) Γ_2 is ϕ-ism, then $-2\varrho \langle \Gamma_2 u_1 - \Gamma_2 u_2, u_1 - u_2\rangle \le -2\varrho\phi\|\Gamma_2 u_1 - \Gamma_2 u_2\|^2$;
If (iii) Γ_2 is expanding, then

$$-2\varrho\phi\|\Gamma_2 u_1 - \Gamma_2 u_2\|^2 \le -2\varrho\phi h^2\|u_1 - u_2\|^2 < -2\varrho\phi\|u_1 - u_2\|^2.$$

Exploiting (i)–(iii), we get $\left\|(I - \varrho\Gamma_2)u_1 - (I - \varrho\Gamma_2)u_2\right\|^2 \le \theta^2\|u_1 - u_2\|^2$. ☐

Now, we proceed to demonstrate the convergence theorem for the algorithm S-FBSA.

Theorem 6 *Let* $\{w_n\}_{n=0}^{\infty}$ *be a sequence generated by the algorithm* S-FBSA, *with* $\varrho > 0$ *such that* $\varrho < \phi\sqrt{2\varrho\phi} < \sqrt{\varrho^2 + \phi^2}$ *and* $P_{(\Gamma_1+\Gamma_2)^{-1}(0)}w_0 = p_*$. *Then, for any initial point* $w_0 \in C$, *the sequence* $\{w_n\}_{n=0}^{\infty}$ *converges strongly to* p_* *so that*

$$(\forall n \in \mathbb{N}) \quad \|w_{n+1} - p_*\| \leq \theta^{n+1} \|w_0 - p_*\|,$$

where θ *is defined in Proposition 5.*

Proof From the algorithm S-FBSA, Lemma 1 and Proposition 4, we have

$$\begin{aligned}
\|z_n - p_*\|^2 &\leq (1-\eta_n^{(3)})\|w_n - p_*\|^2 + \eta_n^{(3)}\left\|J_\varrho^{\Gamma_1,\Gamma_2}w_n - J_\varrho^{\Gamma_1,\Gamma_2}p_*\right\|^2 \\
&= (1-\eta_n^{(3)})\|w_n - p_*\|^2 + \eta_n^{(3)}\left\|J_\varrho^{\Gamma_1}(I-\varrho\Gamma_2)w_n - J_\varrho^{\Gamma_1}(I-\varrho\Gamma_2)p_*\right\|^2 \\
&\leq (1-\eta_n^{(3)})\|w_n - p_*\|^2 + \eta_n^{(3)}\|(I-\varrho\Gamma_2)w_n - (I-\varrho\Gamma_2)p_*\|^2.
\end{aligned}$$

By using Proposition 5, we attain

$$\begin{aligned}
\|z_n - p_*\|^2 &\leq (1-\eta_n^{(3)})\|w_n - p_*\|^2 + \eta_n^{(3)}\left[1 + \frac{\varrho^2}{\phi^2} - 2\varrho\phi\right]\|w_n - p_*\|^2 \\
&= \left[1 - \eta_n^{(3)}(1-\theta^2)\right]\|w_n - p_*\|^2,
\end{aligned}$$

which implies

$$\|z_n - p_*\|^2 \leq \|w_n - p_*\|^2, \tag{3}$$

as $[1-\eta_n^{(3)}(1-\theta^2)] \leq 1$. Also, we have

$$\begin{aligned}
\|w_{n+1} - p_*\|^2 &\leq (1-\eta_n^{(2)})\left\|(J_\varrho^{\Gamma_1}(I-\varrho\Gamma_2)w_n - J_\varrho^{\Gamma_1}(I-\varrho\Gamma_2)p_*)\right\|^2 \\
&\quad + \eta_n^{(2)}\left\|J_\varrho^{\Gamma_1}(I-\varrho\Gamma_2)z_n - J_\varrho^{\Gamma_1}(I-\varrho\Gamma_2)p_*\right\|^2 \\
&\leq (1-\eta_n^{(2)})\|(I-\varrho\Gamma_2)w_n - (I-\varrho\Gamma_2)p_*)\|^2 \\
&\quad + \eta_n^{(2)}\|(I-\varrho\Gamma_2)z_n - (I-\varrho\Gamma_2)p_*\|^2.
\end{aligned}$$

By using Proposition 5 and the inequality (3), we attain $\|w_{n+1} - p_*\| \leq \theta\|w_n - p_*\| \leq \theta^{n+1}\|w_0 - p_*\|$, which implies that $\lim_{n\to\infty}\|w_n - p_*\| = 0$ as $\theta < 1$. □

Remark 4 By imposing an additional condition $\sum_{k=1}^{n}\eta_k^{(1)} = \infty$ on $\eta_n^{(1)}$, one can obtain the convergence of M-FSBA to p_* similar to the proof of Theorem 6.

The following theorem is a direct consequence of Theorem 6 by setting $\eta_n^{(2)} = 1$ for all $n \in \mathbb{N}$.

Theorem 7 *Let* $\{a_n\}_{n=0}^{\infty}$ *be a sequence generated by the algorithm* NS-FBSA [6] *with* $\varrho > 0$, *such that* $\varrho < \phi\sqrt{2\varrho\phi} < \sqrt{\varrho^2 + \phi^2}$ *and* $P_{(\Gamma_1+\Gamma_2)^{-1}(0)}a_0 = p_*$. *Then, for any initial point* $a_0 \in C$, $\{a_n\}_{n=0}^{\infty}$ *converges strongly to* p_*, *so that*

$$(\forall n \in \mathbb{N}) \quad \|a_{n+1} - p_*\| \leq \theta^{n+1}\|a_0 - p_*\|, \tag{4}$$

where θ *is defined in Proposition 5.*

The following theorem provides the convergence result for the algorithm PS-FBSA.

Theorem 8 *Let $\{c_n\}_{n=0}^{\infty}$ be a sequence generated by the algorithm* PS-FBSA, *with $\varrho > 0$ such that $\varrho < \phi\sqrt{2\varrho\phi} < \sqrt{\varrho^2 + \phi^2}$ and $P_{(\Gamma_1+\Gamma_2)^{-1}(0)}c_0 = p_*$. Then, for any initial point $c_0 \in C$, $\{x_n\}_{n=0}^{\infty}$ converges strongly to p_*, so that*

$$(\forall n \in \mathbb{N}) \quad \|c_{n+1} - p_*\| \leq \theta^{n+1} \|c_0 - p_*\|,$$

where θ is defined in Proposition 5.

Proof From the algorithm PS-FBSA, Lemma 1, and Proposition 4, we have

$$\begin{aligned}\|e_n - p_*\|^2 &\leq (1 - \eta_n^{(6)}) \|c_n - p_*\|^2 + \eta_n^{(6)} \left\| J_\varrho^{\Gamma_1}(I - \varrho\Gamma_2)c_n - J_\varrho^{\Gamma_1}(I - \varrho\Gamma_2)p_* \right\|^2 \\ &\leq (1 - \eta_n^{(6)}) \|c_n - p_*\|^2 + \eta_n^{(6)} \left\| (I - \varrho\Gamma_2)c_n - (I - \varrho\Gamma_2)p_* \right\|^2.\end{aligned}$$

By using Proposition 5, we obtain

$$\begin{aligned}\|e_n - p_*\|^2 &\leq (1 - \eta_n^{(6)}) \|c_n - p_*\|^2 + \eta_n^{(6)} \left[1 + \frac{\varrho^2}{\phi^2} - 2\varrho\phi\right] \|c_n - p_*\|^2 \\ &= [1 - \eta_n^{(6)}(1 - \theta^2)] \|c_n - p_*\|^2,\end{aligned}$$

which yields $\|e_n - p_*\|^2 \leq \|c_n - p_*\|^2$ as $[1 - \eta_n^{(6)}(1 - \theta^2)] \leq 1$. Also,

$$\begin{aligned}\|d_n - p_*\|^2 &\leq (1 - \eta_n^{(5)}) \left\| J_\varrho^{\Gamma_1}(I - \varrho\Gamma_2)c_n - J_\varrho^{\Gamma_1}(I - \varrho\Gamma_2)p_* \right\|^2 \\ &\quad + \eta_n^{(5)} \left\| J_\varrho^{\Gamma_1}(I - \varrho\Gamma_2)e_n - J_\varrho^{\Gamma_1}(I - \varrho\Gamma_2)p_* \right\|^2 \\ &\leq (1 - \eta_n^{(5)}) \|(I - \varrho\Gamma_2)c_n - (I - \varrho\Gamma_2)p_*\|^2 \\ &\quad + \eta_n^{(5)} \|(I - \varrho\Gamma_2)e_n - (I - \varrho\Gamma_2)p_*\|^2.\end{aligned}$$

By using again Proposition 5, we get

$$\begin{aligned}\|d_n - p_*\|^2 &\leq (1 - \eta_n^{(5)}) \left[1 + \frac{\varrho^2}{\phi^2} - 2\varrho\phi\right] \|c_n - p_*\|^2 + \eta_n^{(5)} \left[1 + \frac{\varrho^2}{\phi^2} - 2\varrho\phi\right] \|e_n - p_*\|^2 \\ &\leq \left[1 - \eta_n^{(5)}(1 - \theta^2)\right] \|c_n - p_*\|^2,\end{aligned}$$

which yields $\|d_n - p_*\|^2 < \|c_n - p_*\|^2$ as $[1 - \eta_n^{(5)}(1 - \theta^2)] \leq 1$. Finally,

$$\|c_{n+1} - p_*\|^2 \leq \|(I - \varrho\Gamma_2)d_n - (I - \varrho\Gamma_2)p_*)\|^2.$$

As in the proof of Theorem 6, we get $\|c_{n+1} - p_*\| \leq \theta^{n+1} \|c_0 - p_*\|$, which gives $\lim_{n\to\infty} \|c_n - p_*\| = 0$ as $\theta < 1$. □

The convergence result for PMP-FBSA can be derived by applying Theorem 8 with the specific choice of $\eta_n^{(6)} = 1$ for all $n \in \mathbb{N}$.

Theorem 9 *Let $\{c_n\}_{n=0}^{\infty}$ be a sequence generated by the algorithm* PMP-FBSA, *with $\varrho > 0$ such that $\varrho < \phi\sqrt{2\varrho\phi} < \sqrt{\varrho^2 + \phi^2}$ and $P_{(\Gamma_1+\Gamma_2)^{-1}(0)}c_0 = p_*$. Then, for any initial point $c_0 \in C$, the sequence $\{c_n\}_{n=0}^{\infty}$ converges strongly to p_*, so that*

$$(\forall n \in \mathbb{N}) \quad \|c_{n+1} - p_*\| \leq \theta^{2(n+1)} \|c_0 - p_*\|, \tag{5}$$

where θ is defined in Proposition 5.

The following theorem asserts the equivalence of the convergences of PS-FBSA and NS-FBSA, meaning that if PS-FBSA is convergent, then NS-FBSA is convergent, and vice versa.

Theorem 10 *Let* $\{a_n\}_{n=0}^{\infty}$ *and* $\{c_n\}_{n=0}^{\infty}$ *be sequences generated by the algorithms* NS-FBSA [6] *and* PS-FBSA*, respectively, in which* $\{\eta_n^{(\nu)}\}_{n=0}^{\infty} \subset [0,1]$ *for* $\nu = 4,5,6$. *Let* $\varrho < \phi\sqrt{2\varrho\phi} < \sqrt{\varrho^2+\phi^2}$ *for* $\varrho > 0$, θ *be as in Proposition 5, and let*

$$\widehat{B}_n = (2\theta^2+\theta)\max_{\nu\in\{4,5,6\}}\left\{1-\eta_n^{(\nu)}(1-\theta)\right\},\quad g_n = \frac{\widehat{B}_n}{\eta_n^{(5)}(1-\theta)}\quad (n \geq 0). \tag{6}$$

Then the following claims hold: (i) *If* $\sum_{n=0}^{\infty}\eta_n^{(5)} = \infty$ *and the sequence* $\{g_n\}_{n=0}^{\infty}$ *is bounded, then* $\lim_{n\to\infty}(a_n - c_n) = 0$, *with*

$$\|a_{n+1}-c_{n+1}\| \leq [1-\eta_n^{(5)}(1-\theta)]\|a_n - c_n\| + \widehat{B}_n\|a_n - p_*\|,$$

for each $n \in \mathbb{N}$, *and* $\lim_{n\to\infty} c_n = p_* = P_{(\Gamma_1+\Gamma_2)^{-1}(0)}c_0$.

(ii) *If* $\lim_{n\to\infty} c_n = p_* = P_{(\Gamma_1+\Gamma_2)^{-1}(0)}c_0$, *then* $\lim_{n\to\infty}(c_n - a_n) = 0$ *with*

$$\|c_{n+1}-a_{n+1}\| \leq \theta\|c_n - a_n\| + \widehat{B}_n\|c_n - p_*\|,$$

for each $n \in \mathbb{N}$, *and* $\lim_{n\to\infty} a_n = p_* = P_{(\Gamma_1+\Gamma_2)^{-1}(0)}a_0$.

Proof (i) It is known from Theorem 7 that $\lim_{n\to\infty}\|a_n - p_*\| = 0$. We prove now that $\lim_{n\to\infty}\|a_n - c_n\| = 0$ and $\lim_{n\to\infty}\|c_n - p_*\| = 0$. By Propositions 4, 5, and the algorithms NS-FBSA [6] and PS-FBSA, we conclude that

$$\begin{aligned}\|a_{n+1}-c_{n+1}\| &\leq \|(I-\varrho\Gamma_2)b_n - (I-\varrho\Gamma_2)p_*)\| + \|(I-\varrho\Gamma_2)d_n - (I-\varrho\Gamma_2)p_*)\| \\ &\leq \theta\,\|b_n - p_*\| + \theta\,\|d_n - p_*\|\,. \end{aligned} \tag{7}$$

Similarly, we have $\|b_n - p_*\| \leq \left[1-\eta_n^{(4)}(1-\theta)\right]\|a_n - p_*\|$,

$$\|d_n - p_*\| \leq (1-\eta_n^{(5)})\theta\,\|c_n - p_*\| + \eta_n^{(5)}\theta\,\|e_n - p_*\|\,,$$

and $\|e_n - p_*\| \leq \left[1-\eta_n^{(6)}(1-\theta)\right]\|c_n - p_*\|$, so that

$$\|d_n - p_*\| \leq (1-\eta_n^{(5)})\theta\,\|c_n - p_*\| + \eta_n^{(5)}\theta[1-\eta_n^{(6)}(1-\theta)]\,\|c_n - p_*\|\,.$$

Using these inequalities, we get

$$\begin{aligned}\|a_{n+1}-c_{n+1}\| \leq \theta[1-\eta_n^{(4)}(1-\theta)]\|a_n - p_*\| + \theta^2(1-\eta_n^{(5)})\|c_n - p_*\| \\ +\theta^2\eta_n^{(5)}[1-\eta_n^{(6)}(1-\theta)]\|c_n - p_*\|,\end{aligned}$$

or equivalently

$$\|a_{n+1}-c_{n+1}\| \leq A_n\|a_n - c_n\| + B_n\|a_n - p_*\|, \tag{8}$$

where

$$\begin{aligned}A_n &= \theta^2(1-\eta_n^{(5)}) + \theta^2\eta_n^{(5)}[1-\eta_n^{(6)}(1-\theta)],\\ B_n &= \theta[1-\eta_n^{(4)}(1-\theta)] + \theta^2(1-\eta_n^{(5)}) + \theta^2\eta_n^{(5)}[1-\eta_n^{(6)}(1-\theta)].\end{aligned}$$

Since $\theta < 1$ and $\left\{\eta_n^{(\nu)}\right\} \subset [0,1]$ $(\nu = 4,5,6)$, we have $A_n \leq 1-\eta_n^{(5)}(1-\theta)$,

$$\theta^2(1-\eta_n^{(5)}) \leq \theta^2[1-\eta_n^{(5)}(1-\theta)],\quad \theta^2\eta_n^{(5)}[1-\eta_n^{(6)}(1-\theta)] \leq \theta^2[1-\eta_n^{(6)}(1-\theta)],$$

as well as

$$B_n \le \widehat{B}_n = (2\theta^2 + \theta) \max_{\nu \in \{4,5,6\}} \left\{1 - \eta_n^{(\nu)}(1-\theta)\right\}.$$

Using these inequalities, (8) becomes

$$\|a_{n+1} - c_{n+1}\| \le [1 - \eta_n^{(5)}(1-\theta)]\|a_n - c_n\| + \widehat{B}_n\|a_n - p_*\|. \tag{9}$$

For each $n \in \mathbb{N}$, we set $\rho_n^{(1)} = \|a_n - c_n\| \ge 0$, $\rho_n^{(3)} = \eta_n^{(5)}(1-\theta) \in (0,1)$, and $\rho_n^{(2)} = \widehat{B}_n\|a_n - p_*\|$.

Note that $\sum_{n=0}^{\infty} \eta_n^{(5)} = \infty$. Since the sequence $\{g_n\}_{n=0}^{\infty}$ is bounded, there exists a constant $M > 0$ such that for each $n \in \mathbb{N}$, $|g_n| < M$.

Let $\varepsilon > 0$. Since $\lim_{n\to\infty} \xi_n = 0$ in which $\xi_n = \|a_n - p_*\|$ and $\varepsilon/M > 0$, there exists $n_0 \in \mathbb{N}$ such that for each $n \ge n_0$, $|\xi_n| < \varepsilon/M$. Hence for each $n \ge n_0$, $|g_n\xi_n| < \varepsilon$ and, therefore, $\rho_n^{(2)} = o(\rho_n^{(3)})$. As (9) satisfies all the conditions of Lemma 2, $\lim_{n\to\infty} \|a_n - c_n\| = 0$. Since $\lim_{n\to\infty} \|a_n - p_*\| = 0$ and $\|c_n - p_*\| \le \|a_n - c_n\| + \|a_n - p_*\|$, we conclude that $\lim_{n\to\infty} \|c_n - p_*\| = 0$.

(ii) Now, we prove that $\lim_{n\to\infty} \|c_n - a_n\| = \lim_{n\to\infty} \|a_n - p_*\| = 0$. By Propositions 4, 5, NS-FBSA [6] and PS-FBSA, we have

$$\begin{aligned} \|c_{n+1} - a_{n+1}\| &\le \|(I - \varrho\Gamma_2)d_n - (I - \varrho\Gamma_2)p_*)\| + \|(I - \varrho\Gamma_2)b_n - (I - \varrho\Gamma_2)p_*)\| \\ &\le \theta\,\|d_n - p_*\| + \theta\,\|b_n - p_*\|\,. \end{aligned}$$

As in (i), we find

$$\|c_{n+1} - a_{n+1}\| \le \theta\left[1 - \eta_n^{(4)}(1-\theta)\right]\|c_n - a_n\| + B_n\|c_n - p_*\|, \tag{10}$$

where B_n is defined as in (i). Using inequalities from (i), as well as the inequality $\theta[1 - \eta_n^{(4)}(1-\theta)] \le \theta$, we conclude that the inequality (10) reduces to

$$\|c_{n+1} - a_{n+1}\| \le \theta\|c_n - a_n\| + \widehat{B}_n\|c_n - p_*\|, \tag{11}$$

where $\widehat{B}_n$ is defined in (6).

Similarly as in (i), for each $n \in \mathbb{N}$, we set $\rho_n^{(1)} = \|c_n - a_n\| \ge 0$, $\theta \in (0,1)$, and $\rho_n^{(2)} = \widehat{B}_n\|c_n - p_*\|$. Since $\left\{\widehat{B}_n\right\}_{n=0}^{\infty}$ is a bounded sequence, there exists $R > 0$ such that $|\widehat{B}_n| < R$ for each $n \in \mathbb{N}$. Let $\varepsilon > 0$. Since $\lim_{n\to\infty} \eta_n = 0$ in which $\eta_n = \|c_n - p_*\|$ and $\varepsilon/R > 0$, there exists $n_0 \in \mathbb{N}$ such that $|\eta_n| < \varepsilon/R$, for all $n \ge n_0$. Therefore, for each $n \in \mathbb{N}$, we have $|\widehat{B}_n| < \varepsilon$. Thus, $\lim_{n\to\infty} \rho_n^{(2)} = 0$.

As (11) satisfies all the conditions of Lemma 3, we have $\lim_{n\to\infty} \|c_n - a_n\| = 0$. According to $\lim_{n\to\infty} \|c_n - p_*\| = 0$ and $\|a_n - p_*\| \le \|c_n - a_n\| + \|c_n - p_*\|$, we conclude that $\lim_{n\to\infty} \|a_n - p_*\| = 0$. □

In the next result, we compare the convergence rates of PS-FBSA and NS-FBSA as they approach the solution of problem (1). More precisely, we demonstrate that PS-FBSA converges faster to the solution of this problem compared to NS-FBSA.

Theorem 11 *Let* $\{a_n\}_{n=0}^{\infty}$ *and* $\{c_n\}_{n=0}^{\infty}$ *be sequences generated by the algorithms* NS-FBSA [6] *and* PS-FBSA, *respectively, with* $\{\eta_n^{(\nu)}\}_{n=0}^{\infty} \subset [0,1]$ *for* $\nu = 4, 5, 6$, *in which* $\lim_{n\to\infty} \eta_n^{(5)}\eta_n^{(6)} = 1$ *and* $\lim_{n\to\infty} \eta_n^{(4)} = 1$. *Let* $\varrho < \phi\sqrt{2\varrho\phi} < \sqrt{\varrho^2 + \phi^2}$ *for* $\varrho > 0$ *and* θ *be as in Proposition 5. Then,* $\{c_n\}_{n=0}^{\infty}$ *converges faster than* $\{a_n\}_{n=0}^{\infty}$ *to the solution* $p_* = P_{(\Gamma_1+\Gamma_2)^{-1}(0)}c_0$ *provided that* $a_0 = c_0 \ne p_*$.

Proof According to Theorems 7 and 8, we have the following estimates (which are the best possible) for each $n \in \mathbb{N}$,

$$\|a_{n+1} - p_*\| \leq \theta^{n+1} \|a_0 - p_*\| \prod_{i=0}^{n} \sqrt{[1 - \eta_i^{(4)} \left(1 - \theta^2\right)]} = \vartheta_n^{(2)}$$

and

$$\|c_{n+1} - p_*\| \leq \theta^{2(n+1)} \|c_0 - p_*\| \prod_{i=0}^{n} \sqrt{[1 - \eta_i^{(5)} \eta_i^{(6)} \left(1 - \theta^2\right)]} = \vartheta_n^{(1)}.$$

Observe that $\lim_{n\to\infty} \vartheta_n^{(1)} = \lim_{n\to\infty} \vartheta_n^{(2)} = 0$. Thus, all the conditions of Definition 6 are met. For each $n \in \mathbb{N}$ we define

$$\Phi_n = \frac{\|\vartheta_n^{(1)} - 0\|}{\|\vartheta_n^{(2)} - 0\|} = \theta^{n+1} \Delta_n,$$

in which

$$\Delta_n = \prod_{k=0}^{n} \sqrt{\frac{1 - \eta_k^{(5)} \eta_k^{(6)} (1 - \theta^2)}{1 - \eta_k^{(4)} (1 - \theta^2)}} \cdot \frac{\|c_0 - p_*\|}{\|a_0 - p_*\|}.$$

Then

$$\frac{\Phi_{n+1}}{\Phi_n} = \theta \frac{\Delta_{n+1}}{\Delta_n} = \theta \sqrt{\frac{1 - \eta_{n+1}^{(5)} \eta_{n+1}^{(6)} \left(1 - \theta^2\right)}{1 - \eta_{n+1}^{(4)} \left(1 - \theta^2\right)}}.$$

By applying the assumptions $\lim_{n\to\infty} \eta_n^{(5)} \eta_n^{(6)} = 1$ and $\lim_{n\to\infty} \eta_n^{(4)} = 1$, we get $\lim_{n\to\infty} \Phi_{n+1}/\Phi_n = \lim_{n\to\infty} \theta < 1$, and by ratio test $\sum_{n=0}^{\infty} \Phi_n$ converges. Hence,

$$\lim_{n\to\infty} \Phi_n = \lim_{n\to\infty} \frac{\|\vartheta_n^{(1)} - 0\|}{\|\vartheta_n^{(2)} - 0\|} = 0.$$

By Definition 5, $\{\vartheta_n^{(1)}\}_{n=0}^{\infty}$ converges faster than $\{\vartheta_n^{(2)}\}_{n=0}^{\infty}$ and hence $\{c_n\}_{n=0}^{\infty}$ converges faster than $\{a_n\}_{n=0}^{\infty}$ to $p_* = P_{(\Gamma_1+\Gamma_2)^{-1}(0)} c_0$. □

3.1 An application to convex minimization problem

In this section, the following convex minimization problem will be tackled:

$$\min\{f(u) : u \in C\}. \tag{12}$$

in which $f : C \to \mathbb{R}$ is a convex mapping. Assume that $f : C \to \mathbb{R}$ is a Fréchet differentiable convex function. It is well-known that p_* is a solution of (12) if and only if $p_* = P_C(p_* - \varrho \nabla f(p_*))$ in which $\nabla f : C \to \mathcal{H}$ is the gradient of f and $\varrho > 0$.

Let us consider $J_\varrho^{\Gamma_1,\Gamma_2} = J_\varrho^{\Gamma_1}(I - \varrho\Gamma_2)$ again. Let $\Gamma_1 = \varrho^{-1}(P_C^{-1} - I)$ for any $\varrho > 0$. It follows from Remarks 1 and 3 that Γ_1 is a monotone operator. Also, $(I + \varrho\Gamma_1)C = P_C^{-1}C = \mathcal{H}$ holds for any $\varrho > 0$. That is, Γ_1 is maximal monotone and its resolvent is $J_\varrho^{\Gamma_1} = (I + \varrho\Gamma_1)^{-1} = P_C$. However, if we take ∇f instead of $\Gamma_2 : C \to \mathcal{H}$ in $J_\varrho^{\Gamma_1,\Gamma_2}$, we get $J_\varrho^{\Gamma_1}(I - \varrho\Gamma_2) = P_C(I - \varrho\nabla f)$ for any $\varrho > 0$. As a result, if p holds (2), it is also a solution of minimization problem (12) when $\Gamma_1 = \varrho^{-1}(P_C^{-1} - I)$ and $\Gamma_2 = \nabla f$. The next result is a straight consequence of Theorems 6 and 8 for these choices of Γ_1 and Γ_2.

Corollary 12 *Assume that p_* is a solution of (12). Let $f : C \to \mathbb{R}$ be a Fréchet differentiable convex function such that its gradient ∇f is ϕ-ism and expanding operator for $\varrho > 0$ such that $\varrho < \phi\sqrt{2\varrho\phi} < \sqrt{\varrho^2 + \phi^2}$. Let $\Gamma_1 = \varrho^{-1}(P_C^{-1} - I)$ and $\Gamma_2 = \nabla f$ be in the all algorithms* M-FBSA, S-FBSA, PS-FBSA, PMP-FBSA. *Then, the algorithm* M-FBSA, *with $\sum_{n=0}^{\infty} \eta_n^{(1)} = \infty$, and the algorithms* S-FBSA, PS-FBSA, PMP-FBSA *converge strongly to p_*.*

Example 1 Let $\mathcal{H} = \ell_2$ be a standard Hilbert space of the real sequences with the inner product $\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{k=0}^{\infty} u_k v_k$ and the norm $\|\mathbf{u}\| = \sqrt{\sum_{k=0}^{\infty} |u_k|^2}$. Then, the set $C = \left\{ \mathbf{u} = \{u_k\}_{k=0}^{\infty} : \|\mathbf{u}\| \le 1 \right\} \subset \mathcal{H}$ is convex and closed. Let $f : C \to \mathbb{R}$ be defined by $f(\mathbf{u}) = \|\mathbf{u}\|^2$. For all $\mathbf{u}, \mathbf{v} \in C$ and $\eta \in (0, 1)$, we have

$$\begin{aligned} f\left(\eta \mathbf{u} + (1-\eta)\,\mathbf{v}\right) &= \|\eta \mathbf{u} + (1-\eta)\mathbf{v}\|^2 = \sum_{k=0}^{\infty} |\eta u_k + (1-\eta)\, v_k|^2 \\ &\le \eta \sum_{k=0}^{\infty} |u_k|^2 + (1-\eta) \sum_{k=0}^{\infty} |v_k|^2 = \eta f(\mathbf{u}) + (1-\eta) f(\mathbf{v}), \end{aligned}$$

which implies f is a convex function. Observe that the set of solution of (12) for f is $S = \{\mathbf{0}\} = \{(0, 0, \ldots)\}$. By definition (cf. [36, p. 169]) we conclude that f is Fréchet differentiable at $\mathbf{u}$, with the Fréchet derivative $\nabla f \mathbf{u} = \left(2u_0, 2u_1, \ldots\right) = 2\mathbf{u}$. Since for each $\mathbf{u}, \mathbf{v} \in \mathcal{H}$,

$$\langle \nabla f \mathbf{u} - \nabla f \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle = \langle 2\mathbf{u} - 2\mathbf{v}, \mathbf{u} - \mathbf{v} \rangle = 2 \sum_{k=0}^{\infty} |u_k - v_k|^2 = \frac{1}{2} \|\nabla f \mathbf{u} - \nabla f \mathbf{v}\|^2,$$

we conclude that ∇f is $\frac{1}{2}$-ism. Also, for each $\mathbf{u}, \mathbf{v} \in \mathcal{H}$ we have

$$\|\nabla f \mathbf{u} - \nabla f \mathbf{v}\|^2 = \|2\mathbf{u} - 2\mathbf{v}\|^2 = 4 \sum_{k=0}^{\infty} |u_k - v_k|^2 = 4 \|\mathbf{u} - \mathbf{v}\|^2,$$

which implies that ∇f is 2-expanding mapping. Moreover, $P_C : \mathcal{H} \to C$ is defined as

$$P_C(\mathbf{u}) = \begin{cases} \mathbf{u}, & \mathbf{u} \in C, \\ \mathbf{u}/\|\mathbf{u}\|, & \mathbf{u} \notin C. \end{cases}$$

Also, it is easy to see that $\varrho < \phi\sqrt{2\varrho\phi} < \sqrt{\varrho^2 + \phi^2}$ holds for $\varrho = 1/9$ and $\phi = 1/2$.

By $\mathbf{u}_n = \left\{u_k^{(n)}\right\}_{k=0}^{\infty} = \left(u_0^{(n)}, u_1^{(n)}, \ldots\right)$ we denote the n^{th} iteration obtained by any of the previous algorithms. We take $\eta_n^{(\nu)} = 1/(n+2)$ for $\nu = 1, 3, 6, 7$ and $\eta_n^{(\nu)} = 1/(n+2)^2$ for $\nu = 2, 5$.

The convergence behaviors of the algorithms P-FBSA, M-FBSA, S-FBSA, NS-FBSA, PS-FBSA, PMP-FBSA are listed in Tables 1 and 2 and illustrated in Fig. 1. In these tables we give the obtained values after the first iteration ($n = 1$), as well as ones after $n = 250, 500, 750$ and 1000 iterations. Numbers in parentheses indicate decimal exponents, e.g., 1.2937(−28) means 1.2937×10^{-28}.

Table 1 Convergence behavior of the algorithms P-FBSA, M-FBSA and S-FBSA

n	P-FBSA	M-FBSA	S-FBSA
1	$(0.388889, 0.194444, \ldots)$	$(0.444444, 0.222222, \ldots)$	$(0.378086, 0.189043, \ldots)$
250	$(2.58733(-28), 1.2937(-28), \ldots)$	$(0.158162, 0.079081, \ldots)$	$(2.473(-28), 1.236(-28), \ldots)$
500	$(1.33886(-55), 6.6943(-56), \ldots)$	$(0.135667, 0.067833, \ldots)$	$(1.280(-55), 6.398(-56), \ldots)$
750	$(6.92816(-83), 3.46408(-83), \ldots)$	$(0.124003, 0.062001, \ldots)$	$(6.621(-83), 3.311(-83), \ldots)$
1000	$(3.5851(-110), 1.7925(-110), \ldots)$	$(0.116335, 0.058168, \ldots)$	$(3.426(-110), 1.713(-110), \ldots)$

Table 2 Convergence behavior of the algorithms NS-FBSA, PS-FBSA and PMP-FBSA

n	NS-FBSA	PS-FBSA	PMP-FBSA
1	$(0.345679012, 0.172839506, \ldots)$	$(0.294067215, 0.147033608, \ldots)$	$(0.268861, 0.1344307, \ldots)$
250	$(8.18436(-29), 4.09218(-29), \ldots)$	$(1.2795(-55), 6.3976(-56), \ldots)$	$(4.235(-56), 2.118(-56), \ldots)$
500	$(3.63278(-56), 1.81639(-56), \ldots)$	$(3.426(-110), 1.713(-110), \ldots)$	$(9.728(-111), 4.86(-111), \ldots)$
750	$(1.71822(-83), 8.59111(-84), \ldots)$	$(9.174(-165), 4.587(-165), \ldots)$	$(2.381(-165), 1.19(-165), \ldots)$
1000	$(8.3415(-111), 4.1707(-111), \ldots)$	$(2.457(-219), 1.228(-219), \ldots)$	$(5.98(-220), 2.991(-220), \ldots)$

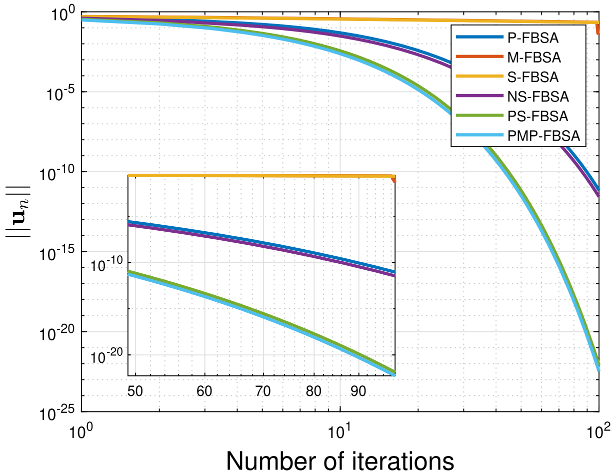

Fig. 1 Convergence behaviour of $\|\mathbf{u}_n\|$

Tables 1 and 2 show that the sequences $\{\mathbf{u}_n\}_{n=0}^{\infty}$, generated by the algorithms P-FBSA, M-FBSA, S-FBSA, NS-FBSA, PS-FBSA, and PMP-FBSA, with the initial guess $\mathbf{u}_0 = \left\{1/2^{k+1}\right\}_{k=0}^{\infty}$, converge to $\mathbf{0} = \{0\}_{k=0}^{\infty}$. Figure 1 shows that the sequence $\|\mathbf{u}_n - \mathbf{0}\| = \|\mathbf{u}_n\|$ converges to 0.

4 Applications

In this section, we inquire about the application of the newly defined FBAs on convex minimization. We define new iterative shrinkage algorithms corresponding to PS-FBSA and PMP-FBSA and apply them to the image deblurring problem.

Let $g : \mathcal{H} \to (-\infty, \infty]$ and $f : \mathcal{H} \to \mathbb{R}$ be proper lower semi-continuous convex functions. Assume that g is a non-smooth function and f is differentiable on $\mathcal{H}$ and has an L-Lipschitz continuous gradient for some $L > 0$. Here we consider the following problem:

$$\min \{F(x) = f(x) + g(x) : x \in \mathcal{H}\},$$

which is equivalent to $0 \in \nabla f(x^*) + \partial g(x^*)$. The set of all solutions of this problem is denoted by X^*.

4.1 Application to supervised learning

In this section, the Least Absolute Shrinkage and Selection Operator-LASSO problem is taken as the basis to perform this process, along with the algorithms mentioned in this paper. We apply these adapted algorithms to four real data sets and, we present a detailed and comparative analysis among them. The outcome of this experiment sets forth that the algorithms PS-FBSA and PMP-FBSA have better computation time, lower cost function value, and higher estimation accuracy than algorithms M-FBSA and S-FBSA in general.

Let us consider a dataset $X \in \mathbb{R}^{m \times d}$ in which every row is a sample and every coloumns are attributes of the samples. $Y \in \mathbb{R}^m$ denotes the set of outcomes, that is, labels of the samples. Now, we can employ the following minimization problem:

$$\min\left\{F(w) = \frac{1}{2}\|Xw - Y\|_2^2 + \delta\|w\|_1 : w \in \mathbb{R}^d\right\}. \tag{13}$$

Let $\Gamma_1(w) = \mathrm{prox}_{\delta\kappa_n}(w) = \left(\left|w^i\right| - \delta\kappa_n\right)_+ \mathrm{sgn}\left(w^i\right)$ in which sgn is the signum function and $\Gamma_2(w) = \left(w - \delta\kappa_n X^t\,(Xw - Y)\right)$. Then, we have

$$J_{\delta\kappa_n}^{\Gamma_1,\Gamma_2}(w) = \mathrm{prox}_{\delta\kappa_n\|\cdot\|_1}(w - \delta\kappa_n X^t(Xw - Y)). \tag{14}$$

The P-FSBA algorithm associated to $J_{\delta\kappa_n}^{\Gamma_1,\Gamma_2}$ in (14) is

$$\textbf{Algorithm ISTA}: \quad w_{n+1} = T(w_n), \quad n \geq 0, \quad w_0 \in \mathbb{R}^d \quad (\kappa_0, \delta > 0),$$

and it can be used to approximate the solution of (13).

In a similar manner, M-ISTA [6], S-ISTA [6], NS-ISTA [6], PS-ISTA, and PMP-ISTA can be defined by utilizing M-FSBA, S-FSBA, NS-FSBA, PS-FSBA, and PMP-FSBA associated with $J_{\delta\kappa_n}^{\Gamma_1,\Gamma_2}$ in (14), respectively (Table 3).

The datasets employed in numerical experiments are listed as follows:

+ ARCENE: It is derived from features showing the abundance of proteins in human serum of a given mass value and includes a set of distracting features called 'probes' which are not predictive. A total of 900 samples with 10000 attributes are contained in this dataset.[1]
+ Heart Disease (Cleveland): With this dataset, researchers aim to distinguish individuals who have heart disease from healthy individuals. A total of 303 samples with 14 attributes are contained in this dataset.[1]

[1] https://archive.ics.uci.edu/ml/index.php.

Table 3 Comparing the effectiveness of algorithms ISTA-PMP ISTA

		ISTA	M-ISTA	S-ISTA
	# of iterations	1585	1591	1447
	$F\left(w^*\right)$	13.81401722	13.81411474	13.72394014
ARCENE	The Best rMse	0.35145778	0.351461535	0.348954979
	rMse (Test)	0.27687903	0.27688189	0.27475253
	Time (s)	361.743	361.342	419.633
	# of iterations	25066	25076	15511
	$F\left(w^*\right)$	120.09814552	120.0981276	119, 6709659
CLEVELAND	The Best rMse	0.88424808	0.88424801	0.8824663
	rMse (Test)	0.67817932	0.67817932	0.67794317
	Time (s)	15.148	16.673	5.840
	# of iterations	16244	16253	10781
	$F\left(w^*\right)$	466.63645322	466.63650778	466.24032396
CNAE9	The Best rMse	0.77800836593	0.77800853	0.7767118
	rMse (Test)	0.7588156	0.75881596	0.75711847
	Time (s)	416.185	416.465	197.668
	# of iterations	100000	100000	76557
	$F\left(w^*\right)$	5551.16531732	5551.17630421	5527.54820185
ISOLET	The Best rMse	2.22942	2.22943697	2.17014111
	rMse (Test)	3.76784608	3.76783122	3.8192689
	Time (s)	9816.789	9879.204	5784.561
		NS-ISTA	PS-ISTA	PMP-ISTA
	# of iterations	1445	1278	1277
	$F\left(w^*\right)$	13.72382997	13.68401091	13.68390731
ARCENE	The Best rMse	0.34895153	0.34767863	0.34767474
	rMse (Test)	0.27475036	0.27382289	0.27382058
	Time (s)	418.531	449.972	449.944
	# of iterations	15506	11536	11534
	$F\left(w^*\right)$	119.67103542	119.52367232	119.52360218
CLEVELAND	The Best rMse	0.88246659	0.88184886	0.88184857
	rMse (Test)	0.67794319	0.67789085	0.67789085
	Time (s)	5.476	2.890	2.514
	# of iterations	10777	8215	8213
	$F\left(w^*\right)$	466.24031897	466.1140877	466.1140287
CNAE9	The Best rMse	0.77671178	0.77627494	0.77627473
	rMse (Test)	0.75711848	0.7564296	0.7564292
	Time (s)	197.037	122.615	122.230
	# of iterations	76552	52244	52240
	$F\left(w^*\right)$	5527.54822498	5527.39823113	5527.39830966
ISOLET	The Best rMse	2.17014097	2.16721879	2.16722057
	rMse (Test)	3.81926902	3.82214176	3.82214007
	Time (s)	5966.230	2826.811	2805.983

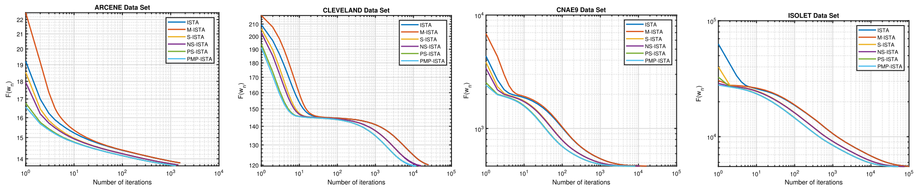

Fig. 2 Comparing the effectiveness of algorithms ISTA-PMP ISTA based on reduction in function values $F(w_n)$ in each step

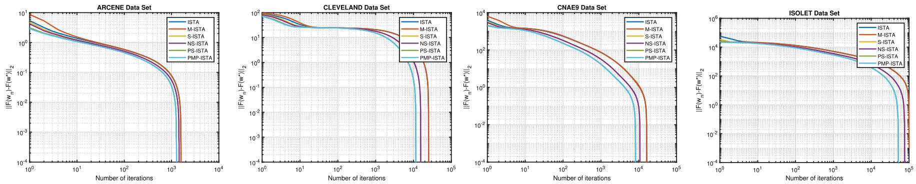

Fig. 3 Comparing the effectiveness of algorithms ISTA-PMP ISTA based on $\|F(w_n) - F(w^*)\|$ in each step

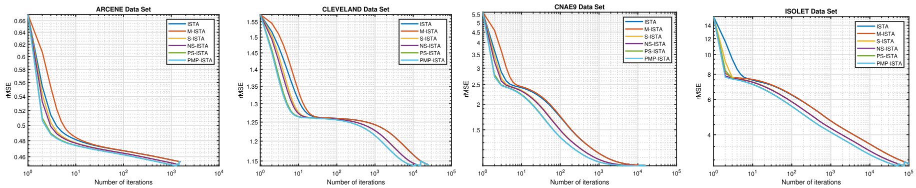

Fig. 4 Comparing the effectiveness of algorithms ISTA-PMP ISTA based on rMSE in each step

+ CNAE: This is a dataset consisting of 9 categories that belong to free-text job description documents of Brazilian companies and is cataloged in a table called the National Classification of Economic Activities (CNAE). A total of 1080 samples with 586 attributes are contained in this dataset.[1]
+ ISOLET: In this data set, whose purpose is to capture letters from sounds and 150 people pronounce each letter in the alphabet twice, a total of 1559 samples with 617 attributes are contained.[1]

The preprocessing of the datasets, application of the algorithms, and preparation of the result analyzes have been carried out in the MATLAB®. The optimal values of $\kappa = (\kappa_n)$ have been found by using a backtracking algorithm.

As a preprocess, all datasets were split into 60% training and 40% test data with a bias added. The stopping criteria for every algorithm is either $|F(w_n) - F(w_{n-1})| < 10^{-5}$ or the maximum number of iteration steps $n = 10^5$.

We compare the performance of algorithms on each dataset in terms of the calculation times, $F(w_n)$ (Fig. 2) and $\|F(w_n) - F(w^*)\|$ (Fig. 3), the accuracy of prediction (rMSE) for training samples (Fig. 4), as well as the accuracy of prediction (rMSE) for testing samples. The graphs presented in Figs. 2-4 are plotted on the log scale.

Fig. 5 (Top) Orginal images; (bottom) Blurred images

4.2 Application to image deblurring problem

The wavelet based deblurring problem defined by

$$\min_x F_{\text{deb}} = \|Ax - b\|_2^2 + \lambda\|x\|_1, \tag{15}$$

in which $A : \mathbb{R}^{n\times d} \to \mathbb{R}^{n\times d}$ is Haar wavelet transform of a blurring matrix, $x \in \mathbb{R}^{n\times d}$ is the original image, and $b \in \mathbb{R}^{n\times d}$ is the blurred image and $\lambda > 0$ is a control parameter. In [37], problem (15) has been argued detailed and the interested reader can consult [38–40]. Gradient projection algorithms are widely used numerical tools to solve this problem.

Let $\Gamma_1(u) = \text{prox}_{\lambda t_n}(u) = \left(\left|u^i\right| - \lambda t_n\right)_+ \text{sgn}\left(u^i\right)$ in which t_n is an appropriate stepsize and $\Gamma_2(u) = \left(u - 2t_n A^t\,(Au - b)\right)$. Then

$$T(u) := J_{\lambda t_n}^{\Gamma_1,\Gamma_2}(u) = \text{prox}_{\lambda t_n\|\cdot\|_1}\left(u - 2t_n A^t\,(Au - b)\right).$$

In this case, the iterative shrinkage-thresholding (ISTA) defined as $u_{n+1} = T(u_n)$. M-ISTA, S-ISTA, NS-ISTA, PS-ISTA and PMP-ISTA algorithms can be defined in a similar manner as above. Taking advantage of the supporting MATLAB® library files provided by Beck and Teboulle, we use MATLAB® functions to produce blurred images from classical test images of Cameraman, Lena, Peppers, and Goldhill. Original and blurred images are shown in Figs. 5.

We present the comparison results regarding the performances of PS-ISTA, PMP-ISTA, ISTA, M-ISTA, S-ISTA, and NS-ISTA in image deblurring tasks. The deblurred Cameraman, Lena, Peppers, Goldhill images are shown in Fig. 6, including PSNR ratio.

We present the deblurring function value at n^{th} iteration with $n = 100$ and $n = 20$ for the Cameraman, Lena, Peppers, and Goldhill in Figs. 7, 8, 9, 10, respectively. As it is understood from these figures, the function values for PS-ISTA and PMP-ISTA are decreasing rapidly during the first 20 steps and are lower than the others even at $n = 100$.

We also present the Frobenius norm of the difference of two successive iterations, that is, $\|u_n - u_{n-1}\|_{\text{fro}}$ value at n^{th} iteration with $n = 100$ and $n = 20$ for Cameraman, Lena, Peppers, and Goldhill in Figs. 11, 12, 13, 14, respectively.

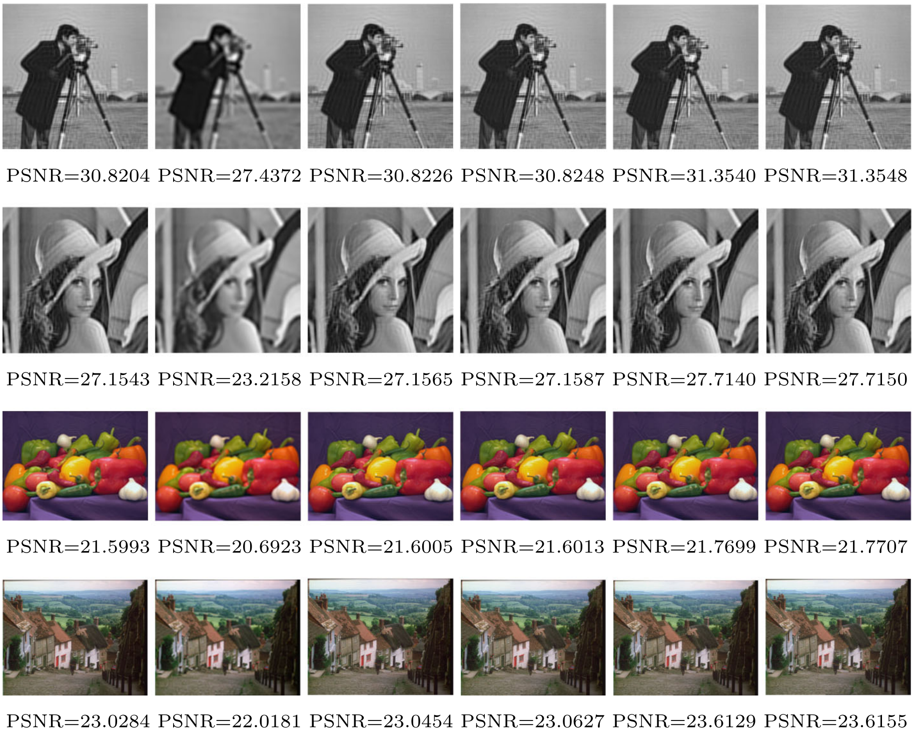

Fig. 6 Comparison results for deblurred Cameraman, Lena, Peppers, Goldhill images (from top to bottom) by ISTA, M-ISTA, S-ISTA, NS-ISTA, PS-ISTA, PMP-ISTA

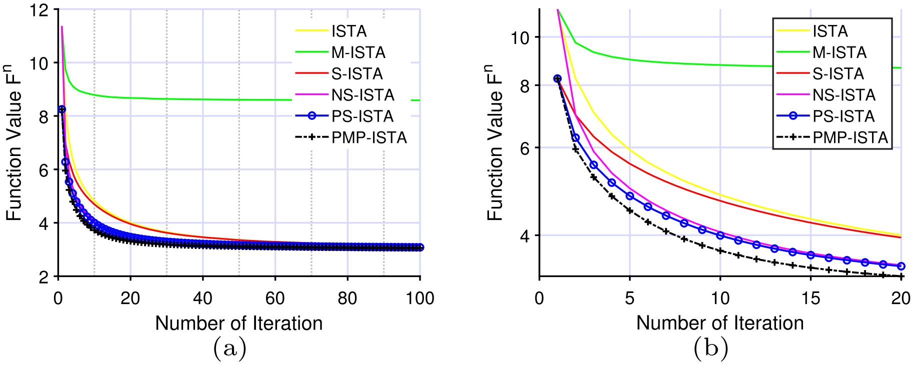

Fig. 7 Function value F^n in each step for Cameraman: (a) $n = 100$ and (b) $n = 20$

5 Conclusion

In this study, we proved some strong convergence theorems through Picard-S and PMP forward-backward algorithms originated from Picard-S [28] and PMP [29] fixed point algorithms. In addition to showing there is an equivalency between convergence of NS-FBSA [6] and PS-FBSA, we compared the rate of convergence of these algorithms. We modified all the algorithms handled in this paper and applied them to the convex minimization problem. We

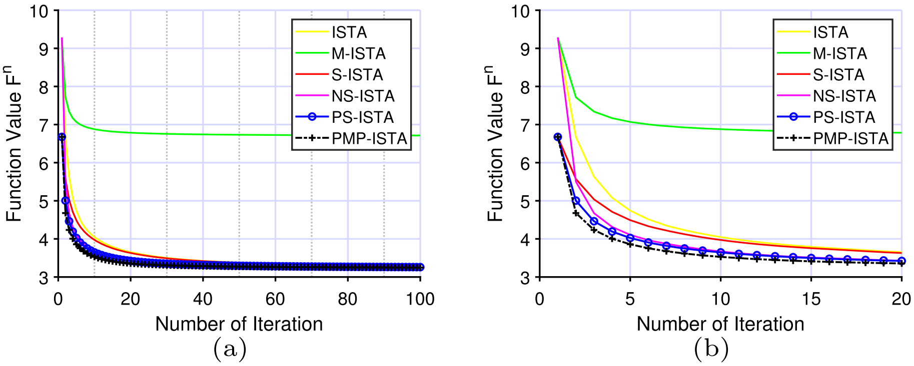

Fig. 8 Function value F^n in each step for Lena: (a) $n = 100$ and (b) $n = 20$

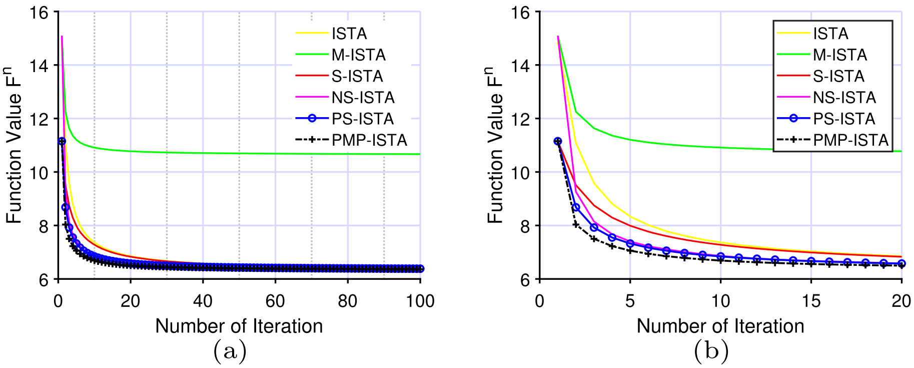

Fig. 9 Function value F^n in each step for Peppers: (a) $n = 100$ and (b) $n = 20$

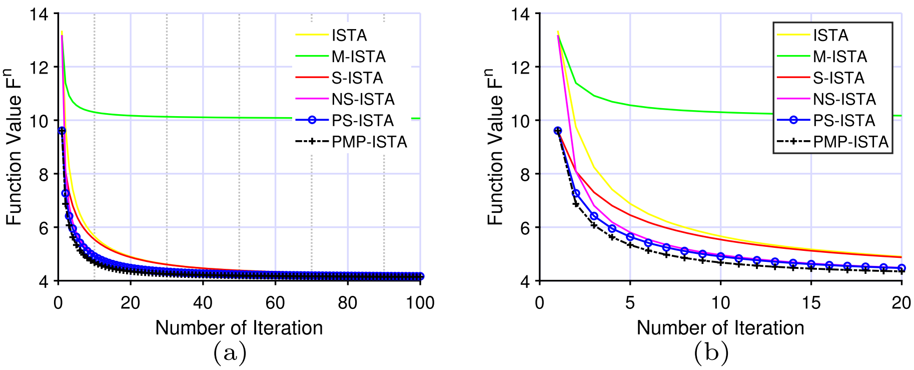

Fig. 10 Function value F^n in each step for Goldhill: (a) $n = 100$ and (b) $n = 20$

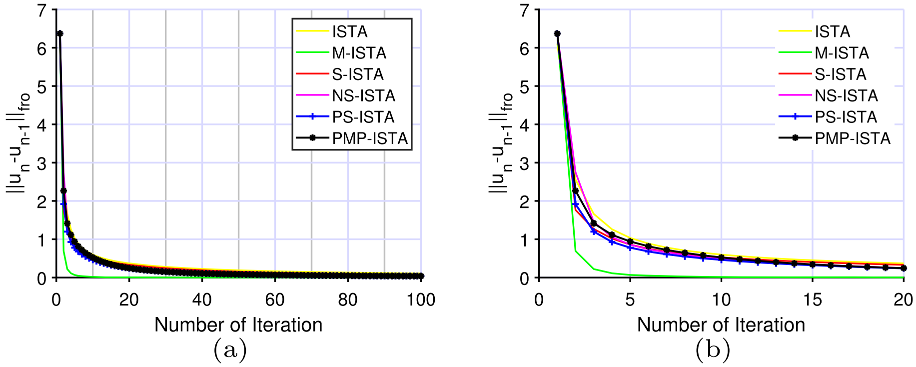

Fig. 11 $\|u_n - u_{n-1}\|_{\text{fro}}$ for Cameraman: (a) $n = 100$ (b) $n = 20$

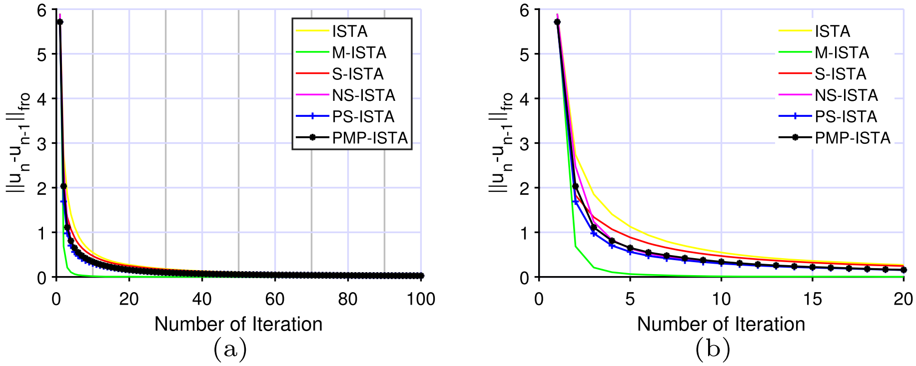

Fig. 12 $\|u_n - u_{n-1}\|_{\text{fro}}$ for Lena: (a) $n = 100$ (b) $n = 20$

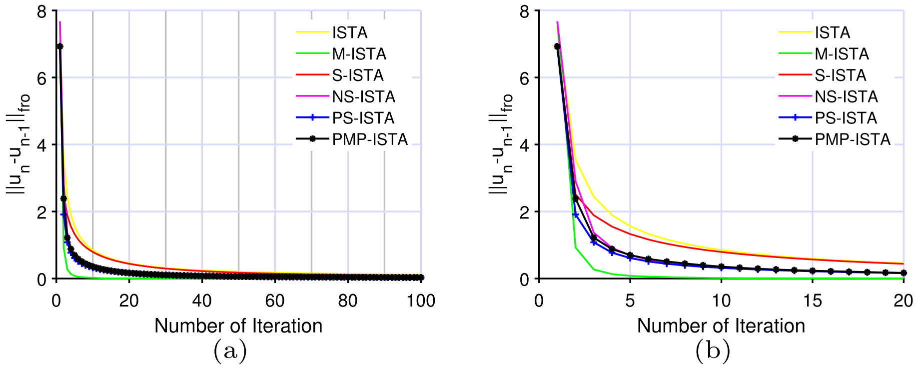

Fig. 13 $\|u_n - u_{n-1}\|_{\text{fro}}$ for Peppers: (a) $n = 100$ (b) $n = 20$

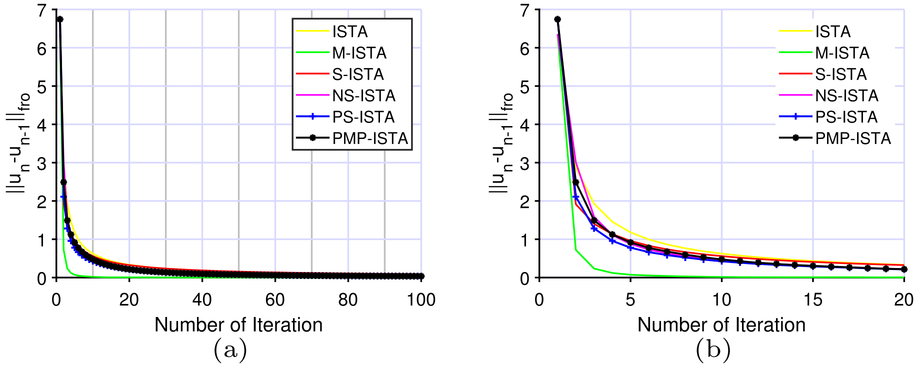

Fig. 14 $\|u_n - u_{n-1}\|_{\text{fro}}$ for Goldhill: (a) $n = 100$ (b) $n = 20$

furnished an academic example in support of Corollary 12 and to illustrate the convergence behaviors of the algorithms M-FBSA, S-FBSA, PS-FBSA, PMP-FBSA. We applied these algorithms to the image deblurring problem and machine learning (classification/regression) with datasets derived from real-world problems. The numerical experiments presented in Sect. 4 reveal that the newly defined algorithms PS-FBSA and PMP-FBSA outperform the algorithms M-FBSA and S-FBSA in solving the problems tackled herein. Our results refine and correct the corresponding results in [41, 42].

Acknowledgements The authors are deeply grateful to the referees for their valuable comments and constructive suggestions for improvements of this paper and its better presentation. This work was supported in part by the Aksaray University, # 2020-009 (Y. Atalan and F. Gürsoy) and by the Serbian Academy of Sciences and Arts, # Φ-96 (G.V. Milovanović).

Funding Open access funding provided by the Scientific and Technological Research Council of Türkiye (TÜBİTAK).

Declarations

Conflict of interest/ Conflict of interest The authors have no Conflict of interest to declare that are relevant to the content of this article.

Availability of Data and Material The datasets are available at https://archive.ics.uci.edu/ml/index.php

References

1. Kinderlehrer, D., Stampacchia, G.: An Introduction to Variational Inequalities and Their Applications. Academic Press, New York (1980)

2. Cho, S.Y., Kang, S.M.: Approximation of common solutions of variational inequalities via strict pseudo-contractions. Acta Math. Sci. **32**, 1607–1618 (2012)
3. Ansari, Q.H.: Vector equilibrium problems and vector variational inequalities. In Vector Variational Inequalities and Vector Equilibria, Springer, Boston, MA (2000)
4. Gürsoy, F., Sahu, D.R., Ansari, Q.H.: S-iteration process for variational inclusions and its rate of convergence. J. Nonlinear Convex Anal. **17**, 1753–1767 (2016)
5. Combettes, P.L., Wajs, R.: Signal recovery by proximal forward-backward splitting. Multiscale Model. Simul. **4**, 1168–1200 (2005)
6. Sahu, D.R., Yao, J.C., Verma, M., Shukla, K.K.: Convergence rate analysis of proximal gradient methods with applications to composite minimization problems. Optimization **70**(1), 75–100 (2021)
7. Osei-Bryson, K.M., Giles, K.: Splitting methods for decision tree induction: An exploration of the relative performance of two entropy-based families. Inf. Syst. Front. **8**(3), 195–209 (2006)
8. Lions, P.L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. **16**, 964–979 (1979)
9. Rockafellar, R.T.: Monotone operators and the proximal point algorithm. SIAM J. Control. Optim. **14**, 877–898 (1976)
10. Dunn, J.C.: Convexity, monotonicity, and gradient processes in Hilbert space. J. Math. Anal. Appl. **53**, 145–158 (1976)
11. Combettes, P.L., Vu, B.C.: Variable metric forward-backward splitting with applications to monotone inclusions in duality. Optimization **63**(9), 1289–1318 (2014)
12. Salzo, S.: The variable metric forward-backward splitting algorithm under mild differentiability assumptions. SIAM J. Optim. **27**(4), 2153–2181 (2017)
13. Latafat, P., Patrinos, P.: Asymmetric forward-backward-adjoint splitting for solving monotone inclusions involving three operators. Comput. Optim. Appl. **68**(1), 57–93 (2017)
14. Pholasa, N., Cholamjiak, P., Cho, Y.J.: Modified forward-backward splitting methods for accretive operators in Banach spaces. J. Nonlinear Sci. Appl. **9**(5), 2766–2778 (2016)
15. Cui, F., Tang, Y., Zhu, C.: Convergence analysis of a variable metric forward-backward splitting algorithm with applications. J. Inequal. Appl. **2019**(1), 141, 27 pp. (2019)
16. Cholamjiak, P.: A generalized forward–backward splitting method for solving quasi inclusion problems in Banach spaces. Numer. Algorithms **71**(4), 915–932 (2016)
17. Dadashi, V., Postolache, M.: Forward-backward splitting algorithm for fixed point problems and zeros of the sum of monotone operators. Arab. J. Math. **9**(1), 89–99 (2020)
18. Zong, C., Tang, Y., Cho, Y.J.: Convergence analysis of an inexact three-operator splitting algorithm. Symmetry **10**(11), 563 (2018)
19. Wang, Y., Wang, F.: Strong convergence of the forward-backward splitting method with multiple parameters in Hilbert spaces. Optimization **67**(4), 493–505 (2018)
20. Brezis, H.: Functional Analysis. Sobolev Spaces and Partial Differential Equations. Springer, New York (2011)
21. Browder, F.E.: Convergence theorems for sequences of nonlinear operators in Banach spaces. Math. Z. **100**(3), 201–225 (1967)
22. Sahu, D.R.: Applications of the S-iteration process to constrained minimization problems and split feasibility problems. Fixed Point Theory **12**(1), 187–204 (2011)
23. Picard, E.: Mémoire sur la théorie des équations aux dérivées partielles et la méthode des approximations successives. J. de Math. Pur. et App. **4**(6), 145–210 (1890)
24. Mann, W.R.: Mean value methods in iteration. Proc. Am. Math. Soc. **4**, 506–510 (1953)
25. Ishikawa, S.: Fixed points by a new iteration method. Proc. Am. Math. Soc. **44**, 147–150 (1974)
26. Noor, M.A.: New approximation schemes for general variational inequalities. J. Math. Anal. Appl. **251**, 217–229 (2000)
27. Agarwal, R.P., O'Regan, D., Sahu, D.R.: Iterative construction of fixed points of nearly asymptotically nonexpansive mappings. J. Nonlinear Convex Anal. **8**(1), 61–79 (2007)
28. Gürsoy, F.: A Picard-S iterative method for approximating fixed point of weak-contraction mappings. Filomat **30**(10), 2829–2845 (2016)
29. Karakaya, V., Atalan, Y., Dogan, K., Bouzara, N.E.H.: Some fixed point results for a new three steps iteration process in Banach spaces. Fixed Point Theory **18**(2), 625–640 (2017)
30. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces, 2nd edn. Springer, Berlin (2017)
31. Weng, X.: Fixed point iteration for local strictly pseudocontractive mapping. Proc. Am. Math. Soc. **113**(3), 727–731 (1991)
32. Ostrowski, A.: The round-off stability of iterations. ZAMM-J. Appl. Math. Mech. **47**(2), 77–81 (1967)

33. Brezinski, C.: Convergence acceleration during the 20th century. J. Comput. Appl. Math. **122**(1–2), 1–21 (2000)
34. Berinde, V.: Picard iteration converges faster than Mann iteration for a class of quasi-contractive operators. Fixed Point Theory Appl. **2**, 97–105 (2004)
35. Berinde, V.: On a notion of rapidity of convergence used in the study of fixed point iterative methods. Creat. Math. Inform. **25**, 29–40 (2016)
36. Milovanović, G.V.: Numerical Analysis, Part I, Naučna Knjiga, Beograd, 1991 (Serbian)
37. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imaging Sci. **2**(1), 183–202 (2009)
38. Chambolle, A., Dossal, C.: On the convergence of the iterates of the "fast iterative shrinkage/thresholding algorithm." J. Optim. Theory Appl. **166**(3), 968–982 (2015)
39. Bioucas-Dias, J.M., Figueiredo, M.A.: An iterative algorithm for linear inverse problems with compound regularizers. Proceedings – International Conference on Image Processing, ICIP. (pp. 685–688) (2008). https://doi.org/10.1109/ICIP.2008.4711847
40. Voronin, S., Woerdeman, H.J.: A new iterative firm-thresholding algorithm for inverse problems with sparsity constraints. Appl. Comput. Harmon. Anal. **35**(1), 151–164 (2013)
41. Atalan, Y.: On a new fixed point iterative algorithm for general variational inequalities. J. Nonlinear Convex Anal. **20**(11), 2371–2386 (2019)
42. Gürsoy, F., Ertürk, M., Abbas, M.: A Picard-type iterative algorithm for general variational inequalities and nonexpansive mappings. Numer. Algorithms **83**(3), 867–883 (2020)