

A strongly convergent Krasnosel'skiĭ-Mann-type algorithm for finding a common fixed point of a countably infinite family of nonexpansive operators in Hilbert spaces

Radu Ioan Bot^{*} Dennis Meier[†]

March 2, 2021

Abstract. In this article, we propose a Krasnosel'skiĭ-Mann-type algorithm for finding a common fixed point of a countably infinite family of nonexpansive operators $(T_n)_{n \geq 0}$ in Hilbert spaces. We formulate an asymptotic property which the family $(T_n)_{n \geq 0}$ has to fulfill such that the sequence generated by the algorithm converges strongly to the element in $\bigcap_{n \geq 0} \text{Fix } T_n$ with minimum norm. Based on this, we derive a forward-backward algorithm that allows variable step sizes and generates a sequence of iterates that converge strongly to the zero with minimum norm of the sum of a maximally monotone operator and a cocoercive one. We demonstrate the superiority of the forward-backward algorithm with variable step sizes over the one with constant step size by means of numerical experiments on variational image reconstruction and split feasibility problems in infinite dimensional Hilbert spaces.

Key Words. fixed points of families of nonexpansive mappings, Tikhonov regularization, splitting methods, forward-backward algorithm

AMS subject classification. 47J25, 47H09, 47H05, 90C25

1 Introduction

Let \mathcal{H} be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\|\cdot\| := \sqrt{\langle \cdot, \cdot \rangle}$. For a given nonexpansive (i.e. 1-Lipschitz continuous) mapping $T : \mathcal{H} \rightarrow \mathcal{H}$, one of the most prominent iterative methods for finding a fixed point of T is the so-called *Krasnosel'skiĭ-Mann algorithm*, which reads as (see [1])

$$x_{n+1} = x_n + \lambda_n(Tx_n - x_n) \quad \forall n \geq 0,$$

where $x_0 \in \mathcal{H}$ is an arbitrary starting point and $(\lambda_n)_{n \geq 0}$ is a sequence of nonnegative real numbers. Under the assumption that the set of fixed points of T is nonempty, one can show under mild conditions imposed on $(\lambda_n)_{n \geq 0}$ that the sequence $(x_n)_{n \geq 0}$ converges weakly to a fixed point of T .

Since the solving of many monotone inclusion and convex optimization problems can be reduced to the solving of a fixed point problem, there is a huge interest in designing corresponding efficient and stable algorithms. For instance, the *forward-backward*

^{*}University of Vienna, Faculty of Mathematics, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria, email: radu.bot@univie.ac.at. Research partially supported by the Austrian Science Fund (FWF), project I 2419-N32.

[†]University of Vienna, Faculty of Mathematics, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria, email: meierd61@univie.ac.at. Research supported by the Austrian Science Fund (FWF), project I 2419-N32.

algorithm, for determining a zero of the sum of a set-valued maximally monotone operator and a single-valued and cocoercive one, and the *Douglas-Rachford algorithm*, for determining a zero of the sum of two set-valued maximally monotone operators, can be embedded in the framework of the Krasnosel'skiĭ-Mann algorithm. A shortcoming which all of the above mentioned algorithms share is that the convergence of the generated iterates take place only with respect to the weak topology. In order to achieve strong convergence, one has to assume that the involved operators satisfy a stronger notion of monotonicity, like strong monotonicity ([1]).

Since, however, many interesting problems to be solved are formulated in infinite dimensional function spaces, where weak and strong convergence do not coincide, and do not involve strong monotone operators, there is a strong interest in developing algorithms, which generate iterates that strongly converge under minimal assumptions. A variant of the Krasnosel'skiĭ-Mann algorithm which overcomes the drawback of weak convergence has been proposed in [2] and reads as follows:

$$x_{n+1} = \beta_n x_n + \lambda_n (T(\beta_n x_n) - \beta_n x_n) \quad \forall n \geq 0, \quad (1.1)$$

where $x_0 \in \mathcal{H}$ is an arbitrary starting point and $(\lambda_n)_{n \geq 0}$ and $(\beta_n)_{n \geq 0}$ are suitably chosen sequences of positive numbers. The formulation of the iterative scheme has its roots in a standard Tikhonov regularization approach. The sequence $(\beta_n)_{n \geq 0}$ is called *Tikhonov regularization sequence* and has the role to enforce the strong convergence of $(x_n)_{n \geq 0}$ to the fixed point of T with minimum norm. The iterative scheme (1.1) was the starting point for deriving in [2] strongly convergent forward-backward, Douglas-Rachford as well as primal-dual algorithms for monotone inclusion problems. Tikhonov regularization techniques has been used also in [3, 4, 5] in order to enforce strong convergence of numerical algorithms.

In this article we will address the problem of finding a common fixed point of a family of nonexpansive operators $T_n : \mathcal{H} \rightarrow \mathcal{H}$, for $n \geq 0$. We will formulate a Krasnosel'skiĭ-Mann-type algorithm endowed with Tikhonov regularization terms, which evaluates in the iteration $n \geq 0$ the operator T_n . Under the hypothesis that the sequence $(T_n)_{n \geq 0}$ fulfils an appropriate asymptotic condition, and provided that the intersection of the sets of fixed points of $(T_n)_{n \geq 0}$ is nonempty, we show that the sequence of generated iterates converges strongly to the common fixed point with minimum norm of the mappings $(T_n)_{n \geq 0}$. Based on this, we derive a strongly convergent forward-backward method with variable step sizes for finding the zeros of the sum of a maximally monotone operator and a cocoercive one. This method is much more flexible than the variant with constant step size, a fact which we also emphasize by means of numerical experiments on variational image reconstruction and split feasibility problems in infinite dimensional Hilbert spaces.

2 A strongly convergent Krasnosel'skiĭ-Mann-type algorithm

The symbols \rightharpoonup and \rightarrow denote weak and strong convergence, respectively. We recall that a mapping $T : \mathcal{H} \rightarrow \mathcal{H}$ is called nonexpansive, if $\|Tx - Ty\| \leq \|x - y\|$ for all $x, y \in \mathcal{H}$. The set of fixed points of T is denoted by $\text{Fix } T := \{x \in \mathcal{H} : T(x) = x\}$. For a nonempty convex and closed set $C \subseteq \mathcal{H}$, the projection operator onto C , $P_C : \mathcal{H} \rightarrow \mathcal{H}$, is defined as $P_C(x) := \operatorname{argmin}_{c \in C} \|x - c\|$.

The following result, which is a direct consequence of [5, Lemma 2.5], will play a

key role in the proof of the main result of this paper, which we formulate and prove subsequently.

Lemma 2.1. *Let $(a_n)_{n \geq 0}$ be a sequence of non-negative real numbers satisfying the inequality*

$$a_{n+1} \leq (1 - \theta_n)a_n + \theta_n b_n + \epsilon_n \quad \forall n \geq 0,$$

where

- (i) $0 \leq \theta_n \leq 1$ for all $n \geq 0$ and $\sum_{n \geq 0} \theta_n = +\infty$;
- (ii) $\limsup_{n \rightarrow +\infty} b_n \leq 0$;
- (iii) $\epsilon_n \geq 0$ for all $n \geq 0$ and $\sum_{n \geq 0} \epsilon_n < +\infty$.

Then the sequence $(a_n)_{n \geq 0}$ converges to 0.

Theorem 2.1. *Let $(T_n)_{n \geq 0}$ be a sequence of nonexpansive operators $T_n : \mathcal{H} \rightarrow \mathcal{H}$, for $n \geq 0$, with $S := \bigcap_{n \geq 0} \text{Fix } T_n \neq \emptyset$. Consider the iterative scheme*

$$x_{n+1} = \beta_n x_n + \lambda_n (T_n(\beta_n x_n) - \beta_n x_n) \quad \forall n \geq 0,$$

with starting point $x_0 \in \mathcal{H}$ and $(\lambda_n)_{n \geq 0}$, $(\beta_n)_{n \geq 0}$ real sequences satisfying the conditions:

- (i) $0 < \beta_n \leq 1$ for any $n \geq 0$, $\lim_{n \rightarrow +\infty} \beta_n = 1$, $\sum_{n \geq 0} (1 - \beta_n) = +\infty$ and $\sum_{n \geq 1} |\beta_n - \beta_{n-1}| < +\infty$;
- (ii) $0 < \lambda_n \leq 1$ for any $n \geq 0$, $\liminf_{n \rightarrow +\infty} \lambda_n > 0$ and $\sum_{n \geq 1} |\lambda_n - \lambda_{n-1}| < +\infty$;
- (iii)

$$\sum_{n \geq 1} \|T_n(\beta_{n-1} x_{n-1}) - T_{n-1}(\beta_{n-1} x_{n-1})\| < +\infty.$$

Then

$$\|x_n - T_n x_n\| \rightarrow 0 \text{ as } n \rightarrow +\infty.$$

In addition, if we suppose that the sequence $(T_n)_{n \geq 0}$ satisfies the asymptotic condition

for any subsequences $(T_{n_k})_{k \geq 0}$ of $(T_n)_{n \geq 0}$ and $(x_{n_k})_{k \geq 0}$ of $(x_n)_{n \geq 0}$

with $x_{n_k} \rightharpoonup x \in \mathcal{H}$ and $x_{n_k} - T_{n_k} x_{n_k} \rightarrow 0$ as $k \rightarrow +\infty$ it holds $x \in \bigcap_{n \geq 0} \text{Fix } T_n$, (2.1)

then $(x_n)_{n \geq 0}$ converges strongly to $P_S(0)$.

Proof. First, we show that the sequence $(x_n)_{n \geq 0}$ is bounded. Let $x \in S$. Due to the nonexpansiveness of T_n , we have for any $n \geq 0$

$$\begin{aligned} \|x_{n+1} - x\| &= \|(1 - \lambda_n)(\beta_n x_n - x) + \lambda_n(T_n(\beta_n x_n) - T_n x)\| \\ &\leq (1 - \lambda_n)\|(\beta_n x_n - x)\| + \lambda_n\|(T_n(\beta_n x_n) - T_n x)\| \\ &\leq \|\beta_n x_n - x\| = \|\beta_n(x_n - x) + (\beta_n - 1)x\| \\ &\leq \beta_n\|x_n - x\| + (1 - \beta_n)\|x\|. \end{aligned}$$

From here it follows that

$$\|x_n - x\| \leq \max\{\|x_0 - x\|, \|x\|\} \quad \forall n \geq 0,$$

what shows that $(x_n)_{n \geq 0}$ is bounded. Since T_n is nonexpansive, we also have for any $n \geq 0$

$$\|T_n(\beta_n x_n) - x\| = \|T_n(\beta_n x_n) - T_n x\| \leq \|\beta_n x_n - x\|.$$

This shows that the sequence $(T_n(\beta_n x_n))_{n \geq 0}$ is also bounded.

Next, we prove that

$$\|x_{n+1} - x_n\| \rightarrow 0 \text{ as } n \rightarrow +\infty. \quad (2.2)$$

Indeed, by taking into account that the operators $(T_n)_{n \geq 0}$ are nonexpansive, we can derive for any $n \geq 1$ the following estimate

$$\begin{aligned} & \|x_{n+1} - x_n\| \\ &= \|(1 - \lambda_n)\beta_n x_n - (1 - \lambda_{n-1})\beta_{n-1}x_{n-1} + \lambda_n T_n(\beta_n x_n) - \lambda_{n-1} T_{n-1}(\beta_{n-1}x_{n-1})\| \\ &\leq \|(1 - \lambda_n)(\beta_n x_n - \beta_{n-1}x_{n-1}) + (\lambda_{n-1} - \lambda_n)\beta_{n-1}x_{n-1}\| \\ &\quad + \|\lambda_n(T_n(\beta_n x_n) - T_{n-1}(\beta_{n-1}x_{n-1})) + \lambda_n T_{n-1}(\beta_{n-1}x_{n-1}) - \lambda_{n-1} T_{n-1}(\beta_{n-1}x_{n-1})\| \\ &\leq \|\beta_n x_n - \beta_{n-1}x_{n-1}\| + |\lambda_{n-1} - \lambda_n| \|\beta_{n-1}x_{n-1}\| \\ &\quad + \|\lambda_n(T_n(\beta_{n-1}x_{n-1}) - T_{n-1}(\beta_{n-1}x_{n-1})) + (\lambda_n - \lambda_{n-1})T_{n-1}(\beta_{n-1}x_{n-1})\| \\ &\leq \|\beta_n x_n - \beta_{n-1}x_{n-1}\| + |\lambda_{n-1} - \lambda_n| (\|\beta_{n-1}x_{n-1}\| + \|T_{n-1}(\beta_{n-1}x_{n-1})\|) \\ &\quad + \lambda_n \|T_n(\beta_{n-1}x_{n-1}) - T_{n-1}(\beta_{n-1}x_{n-1})\|. \end{aligned}$$

Thanks to the boundedness of the sequences $(\beta_n x_n)_{n \geq 0}$ and $(T_n(\beta_n x_n))_{n \geq 0}$, there exists a constant $C_2 > 0$ such that

$$\begin{aligned} \|x_{n+1} - x_n\| &\leq \|\beta_n x_n - \beta_{n-1}x_{n-1}\| + |\lambda_{n-1} - \lambda_n| C_2 \\ &\quad + \lambda_n \|T_n(\beta_{n-1}x_{n-1}) - T_{n-1}(\beta_{n-1}x_{n-1})\| \quad \forall n \geq 1. \end{aligned}$$

Since $(x_n)_{n \geq 0}$ is bounded, there exists a constant $C_1 > 0$ such that

$$\begin{aligned} \|x_{n+1} - x_n\| &\leq \|\beta_n(x_n - x_{n-1}) + (\beta_n - \beta_{n-1})x_{n-1}\| \\ &\quad + |\lambda_{n-1} - \lambda_n| C_2 + \lambda_n \|T_n(\beta_{n-1}x_{n-1}) - T_{n-1}(\beta_{n-1}x_{n-1})\| \\ &\leq \beta_n \|x_n - x_{n-1}\| + |\beta_n - \beta_{n-1}| C_1 \\ &\quad + |\lambda_{n-1} - \lambda_n| C_2 + \lambda_n \|T_n(\beta_{n-1}x_{n-1}) - T_{n-1}(\beta_{n-1}x_{n-1})\|. \end{aligned}$$

Now statement (2.2) is a consequence of Lemma 2.1, by taking for $n \geq 1$ the choices $a_n := \|x_n - x_{n-1}\|$, $b_n := 0$, $\theta_n := 1 - \beta_n$ and

$$\epsilon_n := |\beta_n - \beta_{n-1}| C_1 + |\lambda_{n-1} - \lambda_n| C_2 + \lambda_n \|T_n(\beta_{n-1}x_{n-1}) - T_{n-1}(\beta_{n-1}x_{n-1})\|.$$

Next we prove that

$$\|x_n - T_n x_n\| \rightarrow 0 \text{ as } n \rightarrow +\infty. \quad (2.3)$$

For any $n \geq 0$ we have the following estimate

$$\begin{aligned} \|x_n - T_n x_n\| &\leq \|x_{n+1} - x_n\| + \|x_{n+1} - T_n x_n\| \\ &= \|x_{n+1} - x_n\| + \|(1 - \lambda_n)(\beta_n x_n - T_n x_n) + \lambda_n(T_n(\beta_n x_n) - T_n x_n)\| \\ &\leq \|x_{n+1} - x_n\| + (1 - \lambda_n) \|\beta_n x_n - T_n x_n\| + \lambda_n \|\beta_n x_n - x_n\| \\ &\leq \|x_{n+1} - x_n\| + (1 - \lambda_n) \|\beta_n x_n - \beta_n T_n x_n\| \\ &\quad + (1 - \lambda_n) \|\beta_n T_n x_n - T_n x_n\| + \lambda_n (1 - \beta_n) \|x_n\| \\ &= \|x_{n+1} - x_n\| + (1 - \lambda_n) \|x_n - T_n x_n\| \\ &\quad + (1 - \lambda_n)(1 - \beta_n) \|T_n x_n\| + \lambda_n (1 - \beta_n) \|x_n\|. \end{aligned}$$

From here it follows that

$$\lambda_n \|x_n - T_n x_n\| \leq \|x_{n+1} - x_n\| + (1 - \lambda_n)(1 - \beta_n) \|T_n x_n\| + \lambda_n(1 - \beta_n) \|x_n\| \quad \forall n \geq 0.$$

Since

$$\|T_n x_n - x\| = \|T_n x_n - T_n x\| \leq \|x_n - x\| \leq \max\{\|x_0 - x\|, \|x\|\},$$

it follows that the sequence $(T_n x_n)_{n \geq 0}$ is also bounded. Taking into account that $(x_n)_{n \geq 0}$ is bounded, (2.2) and the assumptions (i) and (ii), it follows from the last inequality that (2.3) holds.

In order to prove the last statement of the theorem, we suppose that the sequence $(T_n)_{n \geq 0}$ fulfills the asymptotic condition (2.1). We denote by $\bar{x} := P_S(0)$ the element in S with minimum norm. Using again that T_n is nonexpansive, we have for any $n \geq 0$

$$\begin{aligned} \|x_{n+1} - \bar{x}\| &= \|(1 - \lambda_n)(\beta_n x_n - \bar{x}) + \lambda_n(T_n(\beta_n x_n) - T_n \bar{x})\| \\ &\leq (1 - \lambda_n) \|\beta_n x_n - \bar{x}\| + \lambda_n \|T_n(\beta_n x_n) - T_n \bar{x}\| \\ &\leq \|\beta_n x_n - \bar{x}\|. \end{aligned}$$

Hence,

$$\begin{aligned} \|x_{n+1} - \bar{x}\|^2 &\leq \|\beta_n x_n - \bar{x}\|^2 = \|\beta_n(x_n - \bar{x}) + (\beta_n - 1)\bar{x}\|^2 \\ &= \beta_n^2 \|x_n - \bar{x}\|^2 + 2\beta_n(1 - \beta_n) \langle -\bar{x}, x_n - \bar{x} \rangle + (1 - \beta_n)^2 \|\bar{x}\|^2 \\ &\leq \beta_n \|x_n - \bar{x}\|^2 + (1 - \beta_n)(2\beta_n \langle -\bar{x}, x_n - \bar{x} \rangle + (1 - \beta_n) \|\bar{x}\|^2) \quad \forall n \geq 0. \end{aligned} \tag{2.4}$$

Next we show that

$$\limsup_{n \rightarrow +\infty} \langle -\bar{x}, x_n - \bar{x} \rangle \leq 0. \tag{2.5}$$

Assuming the contrary, there would exist a positive real number l and a subsequence $(x_{n_k})_{k \geq 0}$ such that

$$\langle -\bar{x}, x_{n_k} - \bar{x} \rangle \geq l > 0 \quad \forall k \geq 0.$$

Due to the boundedness of the sequence $(x_n)_{n \geq 0}$, we can assume without losing the generality that $(x_{n_k})_{k \geq 0}$ weakly converges to an element $y \in \mathcal{H}$. Taking into account (2.3), it follows from the asymptotic condition (2.1) that y lies in S . On the other hand, from the variational characterization of the projection we have

$$l \leq \lim_{k \rightarrow +\infty} \langle -\bar{x}, x_{n_k} - \bar{x} \rangle = \langle -\bar{x}, y - \bar{x} \rangle \leq 0,$$

which leads to a contradiction. This shows that (2.5) holds. Thus

$$\limsup_{n \rightarrow +\infty} (2\beta_n \langle -\bar{x}, x_n - \bar{x} \rangle + (1 - \beta_n) \|\bar{x}\|^2) \leq 0.$$

A direct application of Lemma 2.1 to (2.4), by taking for $n \geq 0$ the choices $a_n := \|x_n - \bar{x}\|^2$,

$$b_n := 2\beta_n \langle -\bar{x}, x_n - \bar{x} \rangle + (1 - \beta_n) \|\bar{x}\|^2,$$

$\epsilon_n := 0$ and $\theta_n := 1 - \beta_n$, delivers the desired conclusion. \square

Remark 2.1. (i) The asymptotic condition (2.1) was introduced in [7] and used in the convergence analysis of inertial Krasnosel'skiĭ-Mann algorithms designed for finding a common fixed point of a countably infinite family of nonexpansive operators.

(ii) In the particular case when $T_n = T$ for any $n \geq 0$, where $T : \mathcal{H} \rightarrow \mathcal{H}$ is a nonexpansive operator, the asymptotic condition (2.1) becomes the so-called *demi-closedness principle*, which is known to hold for any nonexpansive operator (see [1, Corollary 4.18]). In this setting, Theorem 2.1 reduces to Theorem 3 in [2].

(iii) The assumptions imposed on the sequence $(\lambda_n)_{n \geq 0}$ are met by every monotonically increasing or decreasing (and hence convergent) sequence with a positive limit. For the Tikhonov regularization sequence $(\beta_n)_{n \geq 0}$ one can choose, for instance, $\beta_0 \in (0, \frac{1}{2})$ and $\beta_n = 1 - \frac{1}{1+n}$ for any $n \geq 1$.

The following result is a consequence of Theorem 2.1 and addresses the problem of finding a common fixed point of a countably infinite family of averaged operators. It will play a determinant role in the convergence analysis of the forward-backward method with variable step sizes which we propose in the next section. We recall that a mapping $T : \mathcal{H} \rightarrow \mathcal{H}$ is called firmly nonexpansive, if $\|Tx - Ty\|^2 + \|(\text{Id} - T)x - (\text{Id} - T)y\|^2 \leq \|x - y\|^2$ for all $x, y \in \mathcal{H}$. Every firmly nonexpansive mapping is also nonexpansive. Let $\alpha \in (0, 1)$ be fixed. We say that $R : \mathcal{H} \rightarrow \mathcal{H}$ is an α -averaged operator, if there exists a nonexpansive operator $T : \mathcal{H} \rightarrow \mathcal{H}$ such that $R = (1 - \alpha)\text{Id} + \alpha T$. Here, Id denotes the identity operator on \mathcal{H} .

Corollary 2.1. *Let $(R_n)_{n \geq 0}$ be a sequence of α_n -averaged operators $R_n : \mathcal{H} \rightarrow \mathcal{H}$ with $S := \bigcap_{n \geq 0} \text{Fix } R_n \neq \emptyset$ and $(\alpha_n)_{n \geq 0} \subseteq (0, 1)$ fulfilling the conditions*

$$\liminf_{n \rightarrow +\infty} \alpha_n > 0 \text{ and } \sum_{n \geq 1} \left| \frac{1}{\alpha_n} - \frac{1}{\alpha_{n-1}} \right| < +\infty. \quad (2.6)$$

Consider the iterative scheme

$$x_{n+1} = \beta_n x_n + \lambda_n (R_n(\beta_n x_n) - \beta_n x_n) \quad \forall n \geq 0, \quad (2.7)$$

with starting point $x_0 \in \mathcal{H}$ and $(\lambda_n)_{n \geq 0}, (\beta_n)_{n \geq 0}$ real sequences satisfying the conditions:

- (i) $0 < \beta_n \leq 1$ for any $n \geq 0$, $\lim_{n \rightarrow +\infty} \beta_n = 1$, $\sum_{n \geq 0} (1 - \beta_n) = +\infty$ and $\sum_{n \geq 1} |\beta_n - \beta_{n-1}| < +\infty$;*
- (ii) $0 < \lambda_n \leq \frac{1}{\alpha_n}$ for any $n \geq 0$, $\liminf_{n \rightarrow +\infty} \lambda_n > 0$ and $\sum_{n \geq 1} |\lambda_n - \lambda_{n-1}| < +\infty$;*
- (iii)*

$$\sum_{n \geq 1} \|R_n(\beta_{n-1} x_{n-1}) - R_{n-1}(\beta_{n-1} x_{n-1})\| < +\infty.$$

Then

$$\|x_n - R_n x_n\| \rightarrow 0 \text{ as } n \rightarrow +\infty.$$

In addition, if we suppose that the sequence $(R_n)_{n \geq 0}$ satisfies the asymptotic condition

for any subsequences $(R_{n_k})_{k \geq 0}$ of $(R_n)_{n \geq 0}$ and $(x_{n_k})_{k \geq 0}$ of $(x_n)_{n \geq 0}$,

$$\text{with } x_{n_k} \rightharpoonup x \in \mathcal{H} \text{ and } x_{n_k} - R_{n_k} x_{n_k} \rightarrow 0 \text{ as } k \rightarrow +\infty \text{ it holds } x \in \bigcap_{n \geq 0} \text{Fix } R_n, \quad (2.8)$$

then $(x_n)_{n \geq 0}$ converges strongly to $P_S(0)$.

Proof. Fix $n \geq 0$. Let $T_n : \mathcal{H} \rightarrow \mathcal{H}$ be the nonexpansive operator such that $R_n = (1 - \alpha_n) \text{Id} + \alpha_n T_n$. The iterative scheme in (2.7) can be rewritten as

$$x_{n+1} = \beta_n x_n + \alpha_n \lambda_n (T_n(\beta_n x_n) - \beta_n x_n) \quad \forall n \geq 0.$$

We have $T_n = \frac{1}{\alpha_n} R_n + \left(1 - \frac{1}{\alpha_n}\right) \text{Id}$ and therefore

$$\begin{aligned} & \|T_n(\beta_{n-1} x_{n-1}) - T_{n-1}(\beta_{n-1} x_{n-1})\| \\ & \leq \frac{1}{\alpha_n} \|R_n(\beta_{n-1} x_{n-1}) - R_{n-1}(\beta_{n-1} x_{n-1})\| + \left| \frac{1}{\alpha_n} - \frac{1}{\alpha_{n-1}} \right| \|R_{n-1}(\beta_{n-1} x_{n-1})\| \\ & \quad + \left| \frac{1}{\alpha_n} - \frac{1}{\alpha_{n-1}} \right| \|\beta_{n-1} x_{n-1}\|. \end{aligned}$$

From the proof of Theorem 2.1 it follows that the sequences $(R_n(\beta_n x_n))_{n \geq 0}$ and $(\beta_n x_n)_{n \geq 0}$ are bounded. Combining this with (2.6) and assumption (iii), it follows that

$$\sum_{n \geq 1} \|T_n(\beta_{n-1} x_{n-1}) - T_{n-1}(\beta_{n-1} x_{n-1})\| < +\infty.$$

Finally, since $\text{Fix } T_n = \text{Fix } R_n$,

$$\|x_n - T_n x_n\| = \frac{1}{\alpha_n} \|x_n - R_n x_n\|,$$

and $\liminf_{n \rightarrow +\infty} \alpha_n > 0$, it is easy to see that, if $(R_n)_{n \geq 0}$ satisfies the asymptotic condition (2.8), then $(T_n)_{n \geq 0}$ satisfies the asymptotic condition (2.1). The conclusion follows by applying Theorem 2.1. \square

3 A strongly convergent forward-backward algorithm with variable step sizes

Based on the general scheme proposed in Theorem 2.1, we will formulate in this section a strongly convergent forward-backward algorithm with variable step sizes for solving the monotone inclusion problem

$$\text{find } x \in \mathcal{H} \text{ such that } 0 \in Ax + Bx, \quad (3.1)$$

where $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is a maximally monotone operator and $B : \mathcal{H} \rightarrow \mathcal{H}$ is a cocoercive operator.

Having a set-valued operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$, we denote by $\text{zer } A := \{x \in \mathcal{H} : 0 \in Ax\}$ its set of zeros, by $\text{Gr } A := \{(x, u) \in \mathcal{H} \times \mathcal{H} : u \in Ax\}$ its graph and by $A^{-1} : \mathcal{H} \rightrightarrows \mathcal{H}$ the inverse operator of A , which is the operator having as graph $\text{Gr } A^{-1} := \{(x, u) \in \mathcal{H} \times \mathcal{H} : x \in Au\}$. The operator A is said to be monotone, if $\langle x - y, u - v \rangle \geq 0$ for all $(x, u), (y, v) \in \text{Gr } A$, and maximally monotone, if it is monotone and there exists no proper monotone extension of the graph of A on $\mathcal{H} \times \mathcal{H}$. The resolvent of A , $J_A : \mathcal{H} \rightarrow \mathcal{H}$, is defined by

$$J_A := (\text{Id} + A)^{-1}.$$

If A is maximally monotone, then $J_A : \mathcal{H} \rightarrow \mathcal{H}$ is single-valued, maximally monotone and firmly nonexpansive (see [1, Proposition 23.7 and Corollary 23.10]). Further, we denote by $R_A : \mathcal{H} \rightarrow \mathcal{H}$, $R_A := 2J_A - \text{Id}$, the reflected resolvent of A .

We recall that, if $f : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ is a proper, convex and lower semicontinuous function, then the (convex) subdifferential $\partial f : \mathcal{H} \rightrightarrows \mathcal{H}$ of f ,

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

for $x \in \mathcal{H}$ with $f(x) \neq +\infty$ and as $\partial f(x) = \emptyset$, otherwise, is a maximally monotone operator (see [8]). Its resolvent is given by $J_{\partial f} = \text{prox}_f$ (see [1]), where

$$\text{prox}_f : \mathcal{H} \rightarrow \mathcal{H}, \quad \text{prox}_f(x) = \operatorname{argmin}_{y \in \mathcal{H}} \left\{ f(y) + \frac{1}{2} \|y - x\|^2 \right\},$$

denotes the proximal operator of f . For a nonempty closed and convex set $C \subseteq \mathcal{H}$, one has that $P_C = \text{prox}_{\delta_C}$, where

$$\delta_C(x) = \begin{cases} 0, & x \in C \\ +\infty, & x \notin C \end{cases},$$

denotes the indicator function of C .

A single-valued operator $B : \mathcal{H} \rightarrow \mathcal{H}$ is called β -cocoercive, for $\beta > 0$, if $\langle x - y, Bx - By \rangle \geq \beta \|Bx - By\|^2$ for all $x, y \in \mathcal{H}$. According to the Baillon-Haddad Theorem, if $g : \mathcal{H} \rightarrow \mathbb{R}$ is a convex and Fréchet differentiable function with $\frac{1}{\beta}$ -Lipschitz gradient, then $\nabla g : \mathcal{H} \rightarrow \mathcal{H}$ is a β -cocoercive operator.

The strongly convergent forward-backward algorithm with variable step sizes for solving (3.1) that we propose in this section will be formulated as a particular case of (2.7) for $R_n = J_{\gamma_n A} \circ (\text{Id} - \gamma_n B)$, for $n \geq 0$, where $\inf_{n \geq 0} \gamma_n > 0$. To this end we will prove that $(R_n)_{n \geq 0}$ fulfills the asymptotic condition (2.8) (see [9, Corollary 17]). The proof relies on the following lemma, which is a special instance of [1, Corollary 25.5].

Lemma 3.1. *Let $A, B : \mathcal{H} \rightrightarrows \mathcal{H}$ be maximally monotone operators, and $(x_n, u_n)_{n \geq 0} \in \text{Gr } A$, $(y_n, v_n)_{n \geq 0} \in \text{Gr } B$ such that $x_n \rightarrow x$, $y_n \rightarrow y$, $u_n \rightarrow u$, $v_n \rightarrow v$, $u_n + v_n \rightarrow 0$, $x_n - y_n \rightarrow 0$ as $n \rightarrow +\infty$. Then $x = y \in \text{zer}(A + B)$, $(x, u) \in \text{Gr } A$ and $(y, v) \in \text{Gr } B$.*

Proposition 3.1. *Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be maximally monotone, $B : \mathcal{H} \rightarrow \mathcal{H}$ be β -cocoercive, for $\beta > 0$, and $\inf_{n \geq 0} \gamma_n > 0$. Suppose that $\text{zer}(A + B) \neq \emptyset$, and set*

$$R_n := J_{\gamma_n A} \circ (\text{Id} - \gamma_n B) \quad \forall n \geq 0.$$

Then $(R_n)_{n \geq 0}$ fulfills the asymptotic condition (2.8).

Proof. By [1, Proposition 25.1(iv)] we have $\text{Fix } R_n = \text{zer}(A + B)$, hence $\bigcap_{n \geq 0} \text{Fix } R_n = \text{zer}(A + B)$.

Let $(x_n)_{n \geq 0}$ be the sequence generated by the iterative scheme (2.7). Further, let $(x_{n_k})_{k \geq 0}$ be a subsequence of $(x_n)_{n \geq 0}$ such that $x_{n_k} \rightarrow x \in \mathcal{H}$ and $x_{n_k} - R_{n_k} x_{n_k} \rightarrow 0$ as $k \rightarrow +\infty$. We set $y_k := R_{n_k} x_{n_k}$ for any $k \geq 0$. It holds $x_{n_k} - y_k \rightarrow 0$, therefore $y_k \rightarrow x$ as $k \rightarrow +\infty$. Since

$$x_{n_k} - y_k \in (\text{Id} + \gamma_{n_k} A)y_k + \gamma_{n_k} Bx_{n_k} - y_k = \gamma_{n_k} (Ay_k + Bx_{n_k}) \quad \forall k \geq 0,$$

it follows that for any $k \geq 0$ there exist $(x_{n_k}, u_k) \in \text{Gr } B$, $(y_k, v_k) \in \text{Gr } A$ such that

$$\gamma_{n_k} (v_k + u_k) = x_{n_k} - y_k \rightarrow 0 \quad \text{as } k \rightarrow +\infty.$$

Since $\inf_{n \geq 0} \gamma_n > 0$, we obtain $v_k + u_k \rightarrow 0$ as $k \rightarrow +\infty$.

On the other hand, since B is β -cocoercive, we have $\|Bx_{n_k}\| \leq \beta^{-1}\|x_{n_k}\| + \|B0\|$ for any $k \geq 0$. Further, as $(x_{n_k})_{k \geq 0}$ is weakly convergent, the uniform boundedness principle implies that $(x_{n_k})_{k \geq 0}$ is bounded, hence $(Bx_{n_k})_{k \geq 0}$ is also bounded. Consequently, there exists a convergent subsequence $(Bx_{n_{k_l}})_{l \geq 0}$ of $(Bx_{n_k})_{k \geq 0}$ such that $Bx_{n_{k_l}} = u_{k_l} \rightharpoonup u$ as $l \rightarrow +\infty$. Since $v_k + u_k \rightarrow 0$ as $k \rightarrow +\infty$ it follows that $v_{k_l} \rightharpoonup -u$ as $l \rightarrow +\infty$.

Finally, Lemma 3.1 applied to the sequences $(x_{n_{k_l}})_{l \geq 0}$, $(y_{k_l})_{l \geq 0}$, $(u_{k_l})_{l \geq 0}$ and $(v_{k_l})_{l \geq 0}$ gives $x = y \in \text{zer}(A + B)$. \square

The following lemma will be useful in the proof of the main result of this section.

Lemma 3.2. *Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximally monotone operator and $(\gamma_n)_{n \geq 0} \subseteq (0, +\infty)$. Then for any $x \in \mathcal{H}$ and any $n, m \geq 0$ we have*

$$\|J_{\gamma_n A}(x - \gamma_n Bx) - J_{\gamma_m A}(x - \gamma_m Bx)\| \leq \left|1 - \frac{\gamma_m}{\gamma_n}\right| \|J_{\gamma_n A}(x - \gamma_n Bx) - x\|.$$

Proof. Let $x \in \mathcal{H}$ and $n, m \geq 0$. According to [1, Proposition 23.28(i)], we have that

$$J_{\gamma_n A}x = J_{\gamma_m A} \left(\frac{\gamma_m}{\gamma_n}x + \left(1 - \frac{\gamma_m}{\gamma_n}\right) J_{\gamma_n A}x \right).$$

Further, using the nonexpansivity of the resolvent $J_{\gamma_m A}$, we get

$$\begin{aligned} & \|J_{\gamma_n A}(x - \gamma_n Bx) - J_{\gamma_m A}(x - \gamma_m Bx)\| \\ & \leq \left\| \frac{\gamma_m}{\gamma_n}(x - \gamma_n Bx) + \left(1 - \frac{\gamma_m}{\gamma_n}\right) J_{\gamma_n A}(x - \gamma_n Bx) - (x - \gamma_m Bx) \right\| \\ & = \left\| \frac{\gamma_m}{\gamma_n}x + \left(1 - \frac{\gamma_m}{\gamma_n}\right) J_{\gamma_n A}(x - \gamma_n Bx) - x \right\| \\ & = \left|1 - \frac{\gamma_m}{\gamma_n}\right| \|J_{\gamma_n A}(x - \gamma_n Bx) - x\|. \end{aligned} \quad \square$$

Theorem 3.1. *Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximally monotone operator and $B : \mathcal{H} \rightarrow \mathcal{H}$ a β -cocoercive operator, for $\beta > 0$, such that $\text{zer}(A + B) \neq \emptyset$. Consider the iterative scheme*

$$x_{n+1} = (1 - \lambda_n)\beta_n x_n + \lambda_n J_{\gamma_n A}(\beta_n x_n - \gamma_n B(\beta_n x_n)) \quad \forall n \geq 0, \quad (3.2)$$

with starting point $x_0 \in \mathcal{H}$ and $(\lambda_n)_{n \geq 0}$, $(\beta_n)_{n \geq 0}$ and $(\gamma_n)_{n \geq 0}$ real sequences satisfying the conditions:

(i) $0 < \beta_n \leq 1$ for any $n \geq 0$, $\lim_{n \rightarrow +\infty} \beta_n = 1$, $\sum_{n \geq 0} (1 - \beta_n) = +\infty$ and $\sum_{n \geq 1} |\beta_n - \beta_{n-1}| < +\infty$.

(ii) $0 < \lambda_n \leq \frac{4\beta - \gamma_n}{2\beta}$ for any $n \geq 0$, $\liminf_{n \rightarrow +\infty} \lambda_n > 0$ and $\sum_{n \geq 1} |\lambda_n - \lambda_{n-1}| < +\infty$.

(iii) $0 < \gamma_n < 2\beta$ for any $n \geq 0$, $\liminf_{n \rightarrow +\infty} \gamma_n > 0$, and $\sum_{n \geq 1} |\gamma_n - \gamma_{n-1}| < +\infty$.

Then $(x_n)_{n \geq 0}$ converges strongly to $P_{\text{zer}(A+B)}(0)$.

Proof. The iterative scheme in (3.2) can be rewritten as

$$x_{n+1} = \beta_n x_n + \lambda_n (R_n(\beta_n x_n) - \beta_n x_n) \quad \forall n \geq 0,$$

for $R_n := J_{\gamma_n A} \circ (\text{Id} - \gamma_n B)$. For any $n \geq 0$ we have that $\text{Fix } R_n = \text{zer}(A + B)$ (see [1, Proposition 25.1]), thus $\bigcap_{n \geq 0} \text{Fix } R_n = \text{zer}(A + B) \neq \emptyset$.

We set $\alpha_n := \frac{2\beta}{4\beta - \gamma_n}$ for any $n \geq 0$. According to (iii) we have that $\liminf_{n \rightarrow +\infty} \alpha_n > 0$ and

$$\sum_{n \geq 1} \left| \frac{1}{\alpha_n} - \frac{1}{\alpha_{n-1}} \right| = \sum_{n \geq 1} \left| \frac{4\beta - \gamma_n}{2\beta} - \frac{4\beta - \gamma_{n-1}}{2\beta} \right| = \frac{1}{2\beta} \sum_{n \geq 1} |\gamma_n - \gamma_{n-1}| < +\infty,$$

hence (2.6) holds.

Thanks to Lemma 3.2, we have that for any $n \geq 1$

$$\|R_n(\beta_{n-1}x_{n-1}) - R_{n-1}(\beta_{n-1}x_{n-1})\| \leq \left| 1 - \frac{\gamma_{n-1}}{\gamma_n} \right| \|R_n(\beta_{n-1}x_{n-1}) - \beta_{n-1}x_{n-1}\|.$$

We have seen in the proof of Corollary 2.1 that $(\beta_n x_n)_{n \geq 0}$ and $(T_n(\beta_{n-1}x_{n-1}))_{n \geq 1}$ are bounded, too. By (iii) it follows that $\sum_{n \geq 0} \left| 1 - \frac{\gamma_{n-1}}{\gamma_n} \right| < +\infty$ and we can conclude that

$$\sum_{n \geq 1} \|R_n(\beta_{n-1}x_{n-1}) - R_{n-1}(\beta_{n-1}x_{n-1})\| < +\infty,$$

which means that the conditions (i)-(iii) in Corollary 2.1 are fulfilled.

By Proposition 3.1 it follows that the sequence $(R_n)_{n \geq 0}$ fulfils condition (2.8). All assumptions in Corollary 2.1 are fulfilled, which leads to the desired conclusion. \square

Remark 3.1. (i) Let $f : \mathcal{H} \rightarrow (-\infty, +\infty]$ be a proper, convex and lower-semicontinuous function and $g : \mathcal{H} \rightarrow \mathbb{R}$ a convex and Fréchet differentiable function with $\frac{1}{\beta}$ -Lipschitz continuous gradient, for $\beta > 0$, such that $\operatorname{argmin}(f + g) \neq \emptyset$. The iterative scheme

$$x_{n+1} = (1 - \lambda_n)\beta_n x_n + \lambda_n \operatorname{prox}_{\gamma_n f}(\beta_n x_n - \gamma_n \nabla g(\beta_n x_n)) \quad \forall n \geq 0, \quad (3.3)$$

with starting point $x_0 \in \mathcal{H}$ and $(\lambda_n)_{n \geq 0}$, $(\beta_n)_{n \geq 0}$ and $(\gamma_n)_{n \geq 0}$ real sequences satisfying the conditions (i)-(iii) in Theorem 3.1 generates a sequence $(x_n)_{n \geq 0}$ which converges strongly to $\operatorname{P}_{\operatorname{argmin}(f+g)}(0)$.

(ii) Based on Theorem 3.1, one can derive strongly convergent primal-dual splitting algorithms with variable step sizes ([10, 11, 12, 13, 14, 15]) of forward-backward type. These methods are known for their high efficiency when solving highly structured monotone inclusion problems involving mixtures of linearly composed maximally monotone operators and parallel sums of maximally monotone operators. The derivation of strongly convergent primal-dual splitting methods with variable step sizes of forward-backward type can be done in an analogous way as in [2, Section 5.1].

(iii) For $A, B : \mathcal{H} \rightrightarrows \mathcal{H}$ two maximally monotone operators, the classical Douglas-Rachford algorithm operates according to the iterative scheme

$$(\forall n \geq 0) \begin{cases} y_n = J_{\gamma B} x_n \\ z_n = J_{\gamma A}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n) \end{cases},$$

with starting point $x_0 \in \mathcal{H}$. Under mild conditions imposed on the sequence $(\lambda_n)_{n \geq 0}$ and under the assumption that $\operatorname{zer}(A+B) \neq \emptyset$, there exists an element $x \in \operatorname{Fix} R_{\gamma A} R_{\gamma B}$ such that the sequence $(x_n)_{n \geq 0}$ converges weakly to x and $(y_n)_{n \geq 0}$, and $(z_n)_{n \geq 0}$ converge weakly to $J_{\gamma B}$ (see [1, Theorem 25.6]). One can design a strongly convergent Douglas-Rachford algorithm with variable step sizes from the setting of Theorem 2.1, by considering $T_n : \mathcal{H} \rightarrow \mathcal{H}$, $T_n := R_{\gamma_n A} \circ R_{\gamma_n B}$ for $n \geq 0$. However, it is not clear if the family of operators $(T_n)_{n \geq 0}$ satisfies the asymptotical condition (2.1). On the other hand, for a constant step size, i.e. $\gamma_n = \gamma$ for any $n \geq 0$ and some $\gamma \in (0, 2\beta)$, a strongly convergent Douglas-Rachford method can be found in [2].

4 Numerical experiments

The numerical experiments presented in this section were implemented in Mathematica on a 4 × Intel® Core™ i5-4670S CPU @ 3.10 GHz computer with 8GB of RAM. It is worth to mention that the implementations were carried out by only using the symbolic computation packages of Mathematica. In other words, no discretization was performed. By doing so, we could fully exploit the spirit of the Tikhonov regularization technique which is employed in the algorithms we propose in this paper.

4.1 A variational minimization problem

For the first numerical experiment we considered an optimization problem which occurs in variational image reconstruction in infinite dimensional Hilbert spaces. For more details concerning this topic, we refer the reader to [16] and the references therein. Let $\Omega \subset \mathbb{R}^n$, $n \geq 1$, be a domain, $\mathcal{H} := L^2(\Omega) := \{u : \Omega \rightarrow \mathbb{R} : \int_{\Omega} |u(x)|^2 dx < +\infty\}$ be equipped with the scalar product $\langle u, v \rangle := \int_{\Omega} u(x)v(x)dx$ and the associated norm $\|u\| := (\int_{\Omega} |u(x)|^2 dx)^{1/2}$ for all $u, v \in L^2(\Omega)$. The optimization problem we solved has the following formulation

$$\min_{u \in \mathcal{H}} \left\{ \frac{\lambda}{2} \int_{\Omega} ((Ku)(x) - b(x))^2 dx + \frac{1}{2} \int_{\Omega} F(x, u(x)) dx \right\}, \quad (4.1)$$

where $\lambda > 0$, $b \in L^2(\Omega)$,

$$K : \mathcal{H} \rightarrow \mathcal{H}, \quad (Ku)(x) := \int_{\Omega} k(x, y)u(y)dy,$$

with $k \in L^2(\Omega \times \Omega)$, and $F : \Omega \times \mathcal{H} \rightarrow \mathbb{R}$ a measurable function. The parameter λ controls the trade-off between a good fit of u and a smoothness requirement due to the regularization term, while the integral operator K models the deblurring of the original image u . The linear operator K is bounded with $\|K\| \leq \|k\|_{L^2(\Omega \times \Omega)}$ and its adjoint operator $K^* : \mathcal{H} \rightarrow \mathcal{H}$ is given by $(K^*u)(x) := \int_{\Omega} k(y, x)u(y)dy$ for any $x \in \Omega$ (see, for instance, [6, Chapter 1]).

For our numerical example, we considered $\Omega := (0, 1)$, $F := u^2$ and

$$k(x, y) := \begin{cases} 1, & y \leq x \\ 0, & y > x \end{cases}.$$

For this choice, K is the so-called *Volterra operator* given by

$$(Ku)(x) = \int_0^x u(y) dy.$$

We have that

$$\|K\|^2 \leq \|k\|_{L^2((0,1) \times (0,1))}^2 = \int_0^1 \int_0^1 k(x, y)^2 dx dy = \int_0^1 \int_y^1 1 dx dy = \int_0^1 (1 - y) dy = \frac{1}{2},$$

while the adjoint operator of K is given by

$$K^* : \mathcal{H} \rightarrow \mathcal{H}, \quad (K^*u)(x) = \int_x^1 u(y) dy.$$

In this setting, problem (4.1) becomes

$$\min_{u \in \mathcal{H}} \left\{ \frac{\lambda}{2} \int_{\Omega} ((Ku)(x) - b(x))^2 dx + \frac{1}{2} \int_{(0,1)} u^2(x) dx, \right\},$$

Since both operator summands are differentiable with Lipschitz continuous gradient, we had the options: (1) to use the iterative scheme (3.3) as a gradient method (i.e. f is identical zero and g is the function in the objective); (2) to divide the objective into two parts and evaluate one of the two smooth functions via its proximal operator, hence, end up with a proximal-gradient scheme. We pursued both approaches, by taking also into account [17], which suggests that the evaluation of a smooth objective activated via its proximal operator may be advantageous in terms of computational performance compared to evaluating the whole objective through its gradient.

When choosing f identical zero and $g(u) = \frac{\lambda}{2} \int_{\Omega} ((Ku)(x) - b(x))^2 dx + \frac{1}{2} \int_{(0,1)} u^2(x) dx$, the gradient of g reads

$$\nabla g(u) = \lambda K^*(Ku - b) + u \quad \forall u \in \mathcal{H}.$$

For $u_1, u_2 \in \mathcal{H}$ it holds

$$\begin{aligned} \|\nabla g(u_1) - \nabla g(u_2)\| &= \|\lambda K^*K(u_1 - u_2) + (u_1 - u_2)\| \leq (\lambda \|K\|^2 + 1) \|u_1 - u_2\| \\ &\leq \left(\frac{\lambda}{4} + 1\right) \|u_1 - u_2\|, \end{aligned}$$

thus ∇g is $\left(\frac{\lambda}{4} + 1\right)$ -Lipschitz continuous.

In this setting, the iterative scheme (3.3) reads

$$u_{n+1} = \beta_n u_n - \gamma_n [\lambda K^*(K(\beta_n u_n) - b) + \beta_n u_n] \quad \forall n \geq 0, \quad (4.2)$$

with starting point $u_0 \in L^2((0, 1))$ and $(\beta_n)_{n \geq 0} \subseteq (0, 1]$, $(\lambda_n)_{n \geq 0} \subseteq (0, 2 - \gamma_n(\lambda + 4)/8]$ and $(\gamma_n)_{n \geq 0} \subseteq (0, 8/(\lambda + 4))$ are real sequences fulfilling the assumptions (i), (ii) and (iii), respectively, in Theorem 3.1. In Table 2 we report some numerical results obtained when running the iterative scheme (4.2) for different starting points.

When choosing $f(u) = \frac{1}{2} \int_{(0,1)} u^2(x) dx$ and $g(u) = \frac{\lambda}{2} \int_{\Omega} ((Ku)(x) - b(x))^2 dx$, the gradient of g reads

$$\nabla g(u) = \lambda K^*(Ku - b) \quad \forall u \in \mathcal{H},$$

while the proximal operator of f is given by

$$\text{prox}_{\gamma f}(u) = \frac{u}{1 + \gamma},$$

where $\gamma > 0$. For $u_1, u_2 \in \mathcal{H}$ it holds

$$\|\nabla g(u_1) - \nabla g(u_2)\| = \|\lambda K^*K(u_1 - u_2)\| \leq \lambda \|K\|^2 \|u_1 - u_2\| \leq \frac{\lambda}{4} \|u_1 - u_2\|,$$

hence ∇g is $\frac{\lambda}{4}$ -Lipschitz continuous. In this setting, the iterative scheme (3.3) reads

$$u_{n+1} = (1 - \lambda_n) \beta_n u_n + \frac{\lambda_n}{1 + \gamma_n} [\beta_n u_n - \gamma_n \lambda K^*(Ku_n - f)(\beta_n u_n)] \quad \forall n \geq 0, \quad (4.3)$$

with starting point $u_0 \in L^2((0, 1))$ and $(\beta_n)_{n \geq 0} \subseteq (0, 1]$, $(\lambda_n)_{n \geq 0} \subseteq (0, 2 - \gamma_n \lambda / 8]$, and $(\gamma_n)_{n \geq 0} \subseteq (0, 8\lambda)$ are real sequences fulfilling the assumptions (i), (ii) and (iii), respectively, in Theorem 3.1. In Table 1 we report some numerical results obtained when running the iterative scheme (4.3) for different starting points.

u_0	f	$\gamma = 1.3$		$\gamma_n = 1.3 - 0.1 \cdot (-1)^n$	
		Number of it.	CPU time in sec.	Number of it.	CPU time in sec.
$x^2/10$	x	13	4.09281	-	> 600
$2^x/16$	x	13	248.686	-	> 600
$\sin(x)$	x	7	12.9818	-	> 600
$\cos(x)$	x	12	10.2537	-	> 600
$x^2/10$	x^2	13	24.5828	-	> 600
$2^x/16$	x^2	11	25.3676	-	> 600
$\sin(x)$	x^2	9	18.5916	-	> 600
$\cos(x)$	x^2	14	89.7744	-	> 600
$x^2/10$	$\sin(x)$	13	16.3641	-	> 600
$2^x/16$	$\sin(x)$	12	30.4143	-	> 600
$\sin(x)$	$\sin(x)$	5	8.41192	-	> 600
$\cos(x)$	$\sin(x)$	14	17.4393	-	> 600

Table 1: Numerical performances of the Tikhonov regularized gradient method with constant and variable step sizes and different starting points u_0 .

u_0	f	$\gamma = 1.3$		$\gamma_n = 1.3 - 0.1 \cdot (-1)^n$	
		Number of it.	CPU time in sec.	Number of it.	CPU time in sec.
$x^2/10$	x	11	1.333	7	0.291091
$2^x/16$	x	11	25.5198	7	13.7977
$\sin(x)$	x	11	25.3024	7	12.8907
$\cos(x)$	x	11	7.72126	7	2.83167
$x^2/10$	x^2	10	12.9192	7	5.38062
$2^x/16$	x^2	10	22.3199	7	13.649
$\sin(x)$	x^2	10	21.8628	7	13.3971
$\cos(x)$	x^2	10	20.5305	7	12.5219
$x^2/10$	$\sin(x)$	11	9.79028	7	3.4632
$2^x/16$	$\sin(x)$	11	27.1717	7	14.137
$\sin(x)$	$\sin(x)$	11	27.5599	7	13.958
$\cos(x)$	$\sin(x)$	11	6.65283	7	2.18238

Table 2: Numerical performances of the Tikhonov regularized proximal-gradient method with constant and variable step sizes and different starting points u_0 .

Interpretation. In the numerical experiments we carried out for both approaches described above we considered as regularization parameter $\lambda := 1$, as Tikhonov regularization sequence $\beta_0 := \frac{1}{4}$, $\beta_n := 1 - \frac{1}{1+n}$ for $n \geq 1$, and the constant sequence of relaxation parameters $\lambda_n = 0.9$ for any $n \geq 0$. As stopping criterion we used for both iterative schemes $\|u_{n+1} - u_n\| \leq 10^{-4}$.

Table 1 shows that the Tikhonov regularized gradient method with constant step sizes do not reach the demanded accuracy in a reasonable period of time (in our case 600 sec.), while the variant with variable step size does. Table 2 shows that the Tikhonov regularized proximal-gradient method with variable step sizes outperforms both the variant with constant step size as well as the Tikhonov regularized gradient method (4.2) from the point of view of the number of iterations and of the CPU time.

4.2 A split feasibility problem

Let \mathcal{H} and \mathcal{G} be real Hilbert spaces and $L : \mathcal{H} \rightarrow \mathcal{G}$ a bounded linear operator. Let C and Q be nonempty, convex and closed subsets of \mathcal{H} and \mathcal{G} , respectively. The *split feasibility problem* (SFP) searches for a point with the property

$$x \in C \text{ and } Lx \in Q. \quad (4.4)$$

It was originally introduced by Censor and Elfving [18] for solving inverse problems in the context of phase retrieval, medical image reconstruction and intensity modulated radiation therapy.

We will use the strongly convergent proximal-gradient algorithm stated in (3.3) for solving the (SFP). For this purpose, we note that, provided it has a solution, the problem (4.4) can be equivalently written as

$$\min_{x \in \mathcal{H}} \left\{ \delta_C(x) + \frac{1}{2} \|Lx - P_Q(Lx)\|^2 \right\}. \quad (4.5)$$

We choose in the framework of Remark 3.1 (i) $f = \delta_C$ and $g(x) = \frac{1}{2} \|Lx - P_Q(Lx)\|^2$. The function g is Fréchet differentiable with gradient $\nabla g = L^* \circ (\text{Id} - P_Q) \circ L$ and it holds for $x_1, x_2 \in \mathcal{H}$

$$\|\nabla g(x_1) - \nabla g(x_2)\| \leq \|L\|^2 \|x_1 - x_2\|$$

hence ∇g is Lipschitz continuous with constant $\|L\|^2$. The iterative scheme in (3.3) applied to problem (4.5) reads

$$(\forall n \geq 0) \begin{cases} v_n &= (L^* \circ (\text{Id} - P_Q) \circ L)(\beta_n x_n) \\ x_{n+1} &= (1 - \lambda_n) \beta_n x_n + \lambda_n P_C(\beta_n x_n - \gamma_n v_n). \end{cases} \quad (4.6)$$

For the numerical experiments we considered

$$\mathcal{H} = \mathcal{G} = L^2([0, 2\pi]) := \left\{ f : [0, 2\pi] \rightarrow \mathbb{R} : \int_0^{2\pi} |f(s)|^2 ds < +\infty \right\}$$

equipped with the scalar product $\langle f, g \rangle := \int_0^{2\pi} f(s)g(s)ds$ and the associated norm $\|f\| := \left(\int_0^{2\pi} |f(s)|^2 ds \right)^{1/2}$ for all $f, g \in L^2([0, 2\pi])$, and the nonempty, convex and closed sets

$$C := \left\{ x \in L^2([0, 2\pi]) : \int_0^{2\pi} x(s)ds \leq 1 \right\},$$

and

$$Q := \mathbb{R}_+ v, \text{ with } v(t) := t^2.$$

Further, we considered the bounded self-adjoint linear operator $L : L^2([0, 2\pi]) \rightarrow L^2([0, 2\pi])$

$$(Lx)(t) := \frac{3t}{8\pi^3} \int_0^{2\pi} s \cdot x(s) ds.$$

For the norm of L we have the following estimate

$$\|L\|^2 \leq \frac{9}{64\pi^6} \int_0^{2\pi} \int_0^{2\pi} (s \cdot t)^2 ds dt = 1.$$

The projection onto the set C and can be calculated for any $x \in \mathcal{H}$ as ([1, Example 28.16])

$$P_C(x) = \begin{cases} \frac{1 - \int_0^{2\pi} x(s) ds}{2\pi} + x, & \text{if } \int_0^{2\pi} x(s) ds > 1, \\ x, & \text{else} \end{cases}.$$

On the other hand P_Q is given for any $x \in \mathcal{H}$ by ([1, Example 28.24])

$$(P_Q(x))(t) = \begin{cases} \frac{5 \int_0^{2\pi} s^2 \cdot x(s) ds}{32\pi^5} \cdot t^2, & \text{if } \int_0^{2\pi} s^2 \cdot x(s) ds > 0 \\ 0, & \text{else} \end{cases}.$$

Note that, since $x = 0$ is a solution to the (SFP), we have $C \cap L^{-1}(Q) \neq \emptyset$.

x_0	$\gamma_n = 0.5$		$\gamma_n = 1 - \frac{0.5}{1+n}$	
	Number of it.	CPU time in sec.	Number of it.	CPU time in sec.
t	8	0.292133	6	0.197763
t^2	12	1.74567	8	0.3915235
t^3	17	842.058	10	0.655565
$\sin(t)$	3	0.473435	2	0.28365
$\cos(t)$	1	0.050345	1	0.050911
$\exp(t)$	19	836.925	11	2.90818
$\log(t)$	5	16.0949	4	12.8893
\sqrt{t}	6	2.72155	5	2.24045

Table 3: Numerical performances of the iterative scheme (4.6) with constant and variable step sizes, different starting points and the constant sequence of relaxation parameters $\lambda_n = 0.4$ for any $n \geq 0$.

x_0	$\gamma_n = 0.5$		$\gamma_n = 1 - \frac{0.5}{1+n}$	
	Number of it.	CPU time in sec.	Number of it.	CPU time in sec.
t	4	0.123398	3	0.088829
t^2	6	0.251096	4	0.158801
t^3	9	0.481185	5	0.204154
$\sin(t)$	4	0.561261	3	0.412813
$\cos(t)$	1	0.048766	1	0.048814
$\exp(t)$	10	2.14205	6	1.11041
$\log(t)$	3	10.0328	3	10.2823
\sqrt{t}	3	1.4994	3	1.49998

Table 4: Numerical performances of the iterative scheme (4.6) with constant and variable step sizes, different starting points and the sequence of relaxation parameters $\lambda_n = \frac{1}{2} + \frac{1}{2+n}$ for any $n \geq 0$.

Interpretation. We implemented the iterative scheme (4.6) in Mathematica using symbolic computation. We considered as Tikhonov regularization sequence $\beta_0 := \frac{1}{4}$, $\beta_n := 1 - \frac{1}{1+n}$ for $n \geq 1$ and different choices for the sequences of relaxation variables and for the step sizes. As stopping criterion we used

$$\frac{1}{2} \|P_C(x_n) - x_n\|^2 + \frac{1}{2} \|P_Q(Lx_n) - Lx_n\|^2 \leq 10^{-3}.$$

In Table 3 and Table 4 we compare the numerical performances of (4.6) with constant and variable step sizes for different starting points and for two different settings of relaxation variables. One can notice that, in both settings, the variant of the iterative scheme (4.6) with variable step sizes outperforms the one with constant step sizes from the point of view of the number of iterations and of the CPU time.

5 Conclusions

The problem of finding a common fixed point of a countably infinite family of operators is of particular importance since it covers many monotone inclusion and optimization problems that appear in applications and it provides an unifying setting for algorithms with variable step sizes and regularization parameters. We propose in this paper an algorithm which, by making use of Tikhonov regularization techniques, generates a sequence that converges strongly to such a common fixed point of minimum norm. We also show that, when applied to the problem of finding the zeros of the sum of a maximally monotone operator and a cocoercive operator, its convergence properties are inherited by the resulting forward-backward algorithm with variable step sizes and illustrate this by numerical examples. An interesting question of future research is to investigate to what extent the algorithm can be applied to the solving of monotone inclusion problems of other type, of course, without losing its convergence properties.

References

- [1] H.H. Bauschke, P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics, Springer, New York, 2011.
- [2] R.I. Boş, E.R. Csetnek, D. Meier, Inducing strong convergence into the asymptotic behaviour of proximal splitting algorithms in Hilbert spaces, *Optim. Methods Softw.* 34 (3) (2019) 489–514.
- [3] H. Attouch, Viscosity solutions of minimization problems, *SIAM J. Optim.* 6 (3) (1996) 769–806.
- [4] N. Lehdili, A. Moudafi, Combining the proximal algorithm and Tikhonov regularization, *Optimization* 37 (3) (1996) 239–252.
- [5] H.-K. Xu, Iterative algorithms for nonlinear operators, *J. London Math. Soc.* 66 (1) (2002), 240–256.
- [6] H. W. Alt, *Linear Functional Analysis: An Application-Oriented Introduction*, Springer, London, 2016.
- [7] P.-E. Maingé, Convergence theorems for inertial KM-type algorithms, *J. Comput. Appl. Math.* 219 (2008) 223–236.
- [8] R.T. Rockafellar, On the maximal monotonicity of subdifferential mappings, *Pac. J. Math.* 33 (1) (1970) 209–216.
- [9] J.V. Dias, *Inertial Krasnosel'skii-Mann algorithms*, Master thesis, Faculty of Mathematics, University of Vienna, 2017.
- [10] H. Attouch, L.M. Briceño-Arias, P.L. Combettes, A parallel splitting method for coupled monotone inclusions, *SIAM J. Control. Optim.* 48 (2010) 3246–3270.
- [11] R.I. Boş, E.R. Csetnek, A. Heinrich, A primal-dual splitting algorithm for finding zeros of sums of maximal monotone operators, *SIAM J. Optim.* 23(4) (2013) 2011–2036.
- [12] R.I. Boş, C. Hendrich, A Douglas-Rachford type primal-dual method for solving inclusions with mixtures of composite and parallel-sum type monotone operators, *SIAM J. Optim.* 23 (4) (2013) 2541–2565.
- [13] L. M. Briceño-Arias, P.L. Combettes, A monotone + skew splitting model for composite monotone inclusions in duality, *SIAM J. Optim.* 21 (4) (2011) 1230–1250.
- [14] P.L. Combettes, J-C. Pesquet, Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators, *Set-Valued. Var. Anal.* 20 (2) (2012) 307–330.
- [15] Băng C. Vũ, A splitting algorithm for dual monotone inclusions involving cocoercive operators, *Adv Comput. Math.* 38 (3) (2013) 667–681.
- [16] K. Chen, Introduction to variational image-processing models and applications, *Int. J. Comput. Math.* 90 (1) (2013) 1–8.

- [17] P.L. Combettes, L.E. Glaudin, Proximal activation of smooth functions in splitting algorithms for convex minimization, *SIAM J. Imaging Sci.* 12 (4) (2019) 1905–1935.
- [18] Y. Censor, T. Elfving, A multiprojection algorithm using Bregman projections in a product space, *Numer. Algorithms* 8 (1994), 221–239.