

On the convergence of algorithms with Tikhonov regularization terms ^{*}

Bruno Dinis^a and Pedro Pinto^b

^a Departamento de Matemática, Faculdade de Ciências da Universidade de Lisboa,
 Campo Grande, Edifício C6, 1749-016 Lisboa, Portugal
 E-mail: bmdinis@fc.ul.pt

^b Department of Mathematics, Technische Universität Darmstadt,
 Schlossgartenstrasse 7, 64289 Darmstadt, Germany
 E-mail: pinto@mathematik.tu-darmstadt.de

May 15, 2020

Abstract

We consider the strongly convergent modified versions of the Krasnosel'skiĭ-Mann, the forward-backward and the Douglas-Rachford algorithms with Tikhonov regularization terms, introduced by Radu Boţ, Ernő Csetnek and Dennis Meier. We obtain quantitative information for these modified iterations, namely rates of asymptotic regularity and metastability. Furthermore, our arguments avoid the use of sequential weak compactness and use only a weak form of the projection argument.

1 Introduction

In nonlinear analysis one is often confronted with the need to find zeros of a sum of monotone operators. Two well-known iterative splitting methods to approximate such zeros are the forward-backward and the Douglas-Rachford algorithms (see [1]). The former, weakly converges to a zero of the sum of a multi-valued maximal monotone operator with a single-valued cocoercive operator, while the latter weakly converges to a zero of the sum of two maximal monotone operators. These convergence results stem from the weak convergence of the Krasnosel'skiĭ-Mann (KM) iteration to a fixed point of a nonexpansive map T in a Hilbert space H :

$$x_{n+1} = x_n + \lambda_n(T(x_n) - x_n), \quad (\text{KM})$$

with $x_0 \in H$ a starting point and $(\lambda_n) \subset [0, 1]$ a sequence of real numbers. One way to obtain strong convergence is to impose stronger conditions on the operators, such as strong monotonicity or strong convexity. However, the application to certain cases may be unfeasible, as the stronger conditions may prove to be too restrictive. The proximal-Tikhonov algorithm (see [10]) was introduced as an alternative way to obtain, under mild conditions, strong convergence to a zero of a maximal monotone operator A . This algorithm is based on the proximal point algorithm [12, 15] and, in a first step, switches from the operator A to the operator $A + \mu_n \text{Id}$, where (μ_n) is a sequence of nonnegative real numbers. As the original proximal point algorithm, in general, is only weakly convergent, the added term $\mu_n \text{Id}$ is crucial in guaranteeing the strong convergence of the iteration. Motivated by this method, in [3] Radu Boţ, Ernő Csetnek and Dennis Meier considered the (KM) iteration with Tikhonov regularization terms as follows:

$$x_{n+1} = \beta_n x_n + \lambda_n(T(\beta_n x_n) - \beta_n x_n), \quad (\text{T-KM})$$

where $x_0 \in H$ is the starting point, $(\lambda_n), (\beta_n)$ are sequences of positive numbers and $T : H \rightarrow H$ is a nonexpansive mapping. Boţ *et al.* obtained a strong convergence result for (T-KM) and for similar modifications of the forward-backward and the Douglas-Rachford algorithms.

In this paper we obtain rates of asymptotic regularity and metastability for these modified iterations through a quantitative analysis of the proofs in [3]. As such we are able to obtain quantitative versions of the convergence results by Boţ *et al.* Similarly to [6, 5, 4, 14], our analysis bypasses the use of sequential weak compactness and only relies on a weak form of the metric projection argument (as explained in [8, 6]). Our results are guided by proof theoretical techniques (in the context of the proof mining program [7, 9]), however no special knowledge of mathematical logic is required to read this paper.

^{*}2010 Mathematics Subject Classification: 47J25, 47H09, 47H05, 03F10. Keywords: Fixed points of nonexpansive mappings, Tikhonov regularization, splitting algorithms, metastability, asymptotic regularity.

2 Preliminaries

Consider H a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$. We start by recalling some notions and properties concerning operators on Hilbert spaces. An operator $A : H \rightrightarrows H$ is said to be *monotone* if and only if whenever (x, y) and (x', y') are elements of the graph of A , it holds that $\langle x - x', y - y' \rangle \geq 0$. A monotone operator A is said to be *maximal monotone* if the graph of A is not properly contained in the graph of any other monotone operator on H . We denote by $\text{zer}(A)$, the set of all zeros of A . Let $A : H \rightrightarrows H$ be a maximal monotone operator and $\gamma > 0$. The resolvent function $J_{\gamma A}$ is the single-valued function defined by $J_{\gamma A} = (I + \gamma A)^{-1}$, and the reflected resolvent function is the function $R_{\gamma A} := 2J_{\gamma A} - \text{Id}$.

A mapping $T : H \rightarrow H$ is called *nonexpansive* if $\|T(x) - T(y)\| \leq \|x - y\|$, for all $x, y \in H$. The set of fixed points of the mapping T will be denoted by $\text{Fix } T$. If T is nonexpansive, then $\text{Fix } T$ is a closed and convex subset of H .

For $\alpha \in (0, 1]$, a functional $T : H \rightarrow H$ is called α -*averaged*¹ if there exists a nonexpansive operator $T' : H \rightarrow H$ such that $T = (1 - \alpha)\text{Id} + \alpha T'$. The α -averaged operators are always nonexpansive. The $\frac{1}{2}$ -averaged operators are also called firmly nonexpansive operators. The resolvent function $J_{\gamma A}$ is firmly nonexpansive and the reflected resolvent function $R_{\gamma A}$ is nonexpansive.

For $\delta > 0$, a functional $B : H \rightarrow H$ is said to be δ -*cocoercive* if for all $x, y \in H$ it holds that $\langle x - y, B(x) - B(y) \rangle \geq \gamma \|B(x) - B(y)\|^2$.

For a comprehensive introduction to convex analysis and the theory of monotone operators in Hilbert spaces we refer to [1].

The main purpose of this paper is to extract quantitative information from the proof of the following result.

Theorem 2.1. ([3, Theorem 3]) *Let $(\beta_n), (\lambda_n) \subset (0, 1]$ be real sequences satisfying:*

$$\begin{aligned} (i) \lim \beta_n = 1, & \quad (ii) \sum_{n \geq 0} (1 - \beta_n) = \infty, & \quad (iii) \sum_{n \geq 1} |\beta_n - \beta_{n-1}| < \infty, \\ (iv) \liminf \lambda_n > 0, & \quad (v) \sum_{n \geq 1} |\lambda_n - \lambda_{n-1}| < \infty. \end{aligned}$$

Consider the iterative scheme (T-KM) with an arbitrary starting point $x_0 \in H$ and a nonexpansive mapping $T : H \rightarrow H$ such that $\text{Fix } T \neq \emptyset$. Then (x_n) converges strongly to $\text{proj}_{\text{Fix } T}(0)$.

Our main result (Theorem 3.5) gives a bound on the *metastability* property of the sequence (x_n) , i.e.

$$\forall k \in \mathbb{N} \forall f : \mathbb{N} \rightarrow \mathbb{N} \exists n \in \mathbb{N} \forall i, j \in [n, f(n)] (\|x_i - x_j\| \leq \frac{1}{k+1}),$$

which is (non-effectively) equivalent to the Cauchy property for (x_n) .

We now present some useful quantitative lemmas which require the notion of monotone functional for two particular cases. This relies on the strong majorizability relation from [2].

Definition 2.2. For functions $f, g : \mathbb{N} \rightarrow \mathbb{N}$ we define

$$g \leq^* f := \forall n, m \in \mathbb{N} (m \leq n \rightarrow (g(m) \leq f(n) \wedge f(m) \leq f(n))).$$

A function $f : \mathbb{N} \rightarrow \mathbb{N}$ is said to be *monotone* if $f \leq^* f$. We say that a functional $\varphi : \mathbb{N} \times \mathbb{N}^{\mathbb{N}} \rightarrow \mathbb{N}$ is *monotone* if for all $m, n \in \mathbb{N}$ and all $f, g : \mathbb{N} \rightarrow \mathbb{N}$,

$$(m \leq n \wedge g \leq^* f) \rightarrow (\varphi(m, g) \leq \varphi(n, f)).$$

Remark 2.3. For $f : \mathbb{N} \rightarrow \mathbb{N}$, the notion of being monotone corresponds to saying that f is a nondecreasing function, i.e. $\forall n \in \mathbb{N} (f(n) \leq f(n+1))$.

In what follows we restrict our metastability results to monotone functions in $\mathbb{N}^{\mathbb{N}}$. However, there is no real restriction in doing so, as for $f : \mathbb{N} \rightarrow \mathbb{N}$, one has $f \leq^* f^{\text{maj}}$, where f^{maj} is the monotone function defined by $f^{\text{maj}}(n) := \max\{f(i) : i \leq n\}$. In this way, we avoid constantly having to switch from f to f^{maj} , and simplify the notation.

Notation 2.4. Consider a function φ on tuples of variables \bar{x}, \bar{y} . If we wish to consider the variables \bar{x} as parameters we write $\varphi[\bar{x}](\bar{y})$. For simplicity of notation we may then even omit the parameters and simply write $\varphi(\bar{y})$.

Lemma 2.5 ([11]). *Let (s_n) be a bounded sequence of non-negative real numbers, with $d \in \mathbb{N} \setminus \{0\}$ an upper bound for (s_n) , such that for any $n \in \mathbb{N}$*

$$s_{n+1} \leq (1 - \alpha_n)s_n + \alpha_n r_n + \gamma_n,$$

¹The standard definition asks for $\alpha \in (0, 1)$. With this extension, 1-averaged is just another way of saying nonexpansive.

where $(\alpha_n) \subset [0, 1]$, (r_n) and $(\gamma_n) \subset [0, +\infty)$ are given sequences of real numbers. Assume that exist functions $A, R, G : \mathbb{N} \rightarrow \mathbb{N}$ such that, for all $k \in \mathbb{N}$

$$(i) \quad \sum_{i=1}^{A(k)} \alpha_i \geq k; \quad (ii) \quad \forall n \geq R(k) \left(r_n \leq \frac{1}{k+1} \right);$$

$$(iii) \quad \forall n \in \mathbb{N} \left(\sum_{i=G(k)+1}^{G(k)+n} \gamma_i \leq \frac{1}{k+1} \right).$$

Then

$$\forall k \in \mathbb{N} \forall n \geq \theta(k) \left(s_n \leq \frac{1}{k+1} \right),$$

with $\theta(k) := \theta[A, R, G, d](k) := A(M - 1 + \lceil \ln(3d(k+1)) \rceil) + 1$, where $M := \max\{R(3k+2), G(3k+2) + 1\}$.

The next result is an easy adaptation of [14, Lemma 14] for the case where $(\gamma_n) \equiv 0$.

Lemma 2.6. *Let (s_n) be a bounded sequence of non-negative real numbers and $d \in \mathbb{N}$ a positive upper bound on (s_n) . Consider sequences of real numbers $(\alpha_n) \subset (0, 1)$, (r_n) and (v_n) and assume the existence of a monotone function A such that $\sum_{i=1}^{A(k)} \alpha_i \geq k$, for all $k \in \mathbb{N}$. For natural numbers k, n and q assume*

$$\forall i \in [n, q] \left(v_i \leq \frac{1}{3(k+1)(q+1)} \wedge r_i \leq \frac{1}{3(k+1)} \right),$$

and $s_{i+1} \leq (1 - \alpha_i)(s_i + v_i) + \alpha_i r_i$, for all $i \in \mathbb{N}$. Then

$$\forall i \in [\sigma(k, n), q] \left(s_i \leq \frac{1}{k+1} \right),$$

with $\sigma(k, n) := \sigma[A, d](k, n) := A(n + \lceil \ln(3d(k+1)) \rceil) + 1$.

The next result corresponds to a quantitative version of a weaker form of the projection argument of zero onto $\text{Fix } T$. Below $B_N := \{x \in H : \|x - p\| \leq N\}$, where $p \in \text{Fix } T$ is made clear by the context and $N \in \mathbb{N}$.

Lemma 2.7 ([14]). *Let $N \in \mathbb{N} \setminus \{0\}$ be such that $N \geq 2\|p\|$ for some point $p \in \text{Fix } T$. For any $k \in \mathbb{N}$ and monotone function $f : \mathbb{N} \rightarrow \mathbb{N}$, there are $n \leq 24N(f^{(R)}(0) + 1)^2$ and $x \in B_N$ such that $\|T(x) - x\| \leq \frac{1}{f(n)+1}$ and*

$$\forall y \in B_N \left(\|T(y) - y\| \leq \frac{1}{n+1} \rightarrow \langle x, x - y \rangle \leq \frac{1}{k+1} \right),$$

with $R := 4N^4(k+1)^2$ and $\check{f} := \max\{f(24N(m+1)^2), 24N(m+1)^2\}$.

3 Quantitative results

In this section we give a quantitative analysis of Theorem 2.1 as well as some of its corollaries. We start by stating the relevant quantitative conditions.

Given positive real sequences $(\beta_n), (\lambda_n)$, a constant $\ell \in \mathbb{N} \setminus \{0\}$ and functions $b, D, B, L : \mathbb{N} \rightarrow \mathbb{N}$ and $h : \mathbb{N} \rightarrow \mathbb{N} \setminus \{0\}$ we consider the following conditions:

$$(Q_1) \quad \forall n \in \mathbb{N} (\beta_n \geq \frac{1}{h(n)})$$

$$(Q_2) \quad \forall k \in \mathbb{N} \forall n \geq b(k) (|1 - \beta_n| \leq \frac{1}{k+1}),$$

$$(Q_3) \quad \forall k \in \mathbb{N} \left(\sum_{i=1}^{D(k)} (1 - \beta_i) \geq k \right),$$

$$(Q_4) \quad \forall k \in \mathbb{N} \forall n \in \mathbb{N} \left(\sum_{i=B(k)+1}^{B(k)+n} |\beta_i - \beta_{i-1}| \leq \frac{1}{k+1} \right),$$

$$(Q_5) \quad \forall n \in \mathbb{N} (\lambda_n \geq \frac{1}{\ell}),$$

$$(Q_6) \quad \forall k \in \mathbb{N} \forall n \in \mathbb{N} \left(\sum_{i=L(k)+1}^{L(k)+n} |\lambda_i - \lambda_{i-1}| \leq \frac{1}{k+1} \right).$$

The first condition states that the function h witnesses the fact that $\beta_n > 0$. The conditions $(Q_2) - (Q_6)$ correspond to quantitative strenghtenings of the conditions $(i) - (v)$ in Theorem 2.1. Indeed, b is a rate of convergence for $\beta_n \rightarrow 1$; the function D is a rate of divergence for $(\sum (1 - \beta_i))$; the functions B and L are rates for the Cauchy property of the convergent series $\sum |\beta_i - \beta_{i-1}|$ and $\sum |\lambda_i - \lambda_{i-1}|$, respectively, and the number ℓ is used to express the fact that (λ_n) is bounded away from 0.

Example 3.1. Consider the sequences defined by $\lambda_n = \beta_n = 1 - \frac{1}{n+2}$, for all $n \in \mathbb{N}$. Then the conditions $(Q_1) - (Q_6)$ are satisfied with $h \equiv 2$, $b = B = L = \text{Id}$, $D(k) := e^{k+2}$ and $\ell = 2$.

In the following lemmas (x_n) always denotes a sequence generated by **(T-KM)**.

Lemma 3.2. Let $N \in \mathbb{N} \setminus \{0\}$ be such that $N \geq \max\{\|x_0 - p\|, \|p\|\}$, for some $p \in \text{Fix}T$. Then, $\|x_n\| \leq 2N$ and $\|T(x_n)\| \leq 2N$, for all $n \in \mathbb{N}$.

Proof. Let $p \in \text{Fix}T$ be such that $N \geq \max\{\|x_0 - p\|, \|p\|\}$. Since T is nonexpansive, we have

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

By induction on n one easily shows that $\|x_n - p\| \leq N$. Since T is nonexpansive, $\|T(x_n) - p\| \leq N$ and the result follows from the fact that $N \geq \|p\|$. \square

The next lemma gives a rate of asymptotic regularity for the sequence (x_n) .

Lemma 3.3. Let $N \in \mathbb{N} \setminus \{0\}$ be such that $N \geq \max\{\|x_0 - p\|, \|p\|\}$, for some fixed point p . Consider $\ell \in \mathbb{N} \setminus \{0\}$ and monotone functions $b, D, B, L : \mathbb{N} \rightarrow \mathbb{N}$ satisfying conditions $(Q_2) - (Q_6)$. Then

$$(i) \quad \forall k \in \mathbb{N} \forall n \geq \nu_1(k) (\|x_{n+1} - x_n\| \leq \frac{1}{k+1}),$$

$$(ii) \quad \forall k \in \mathbb{N} \forall n \geq \nu_2(k) (\|T(x_n) - x_n\| \leq \frac{1}{k+1}),$$

$$\text{where } \nu_1(k) := \nu_1[N, D, B, L](k) := \theta[D, \mathbf{0}, G, 2N](k), \text{ and}$$

$$\nu_2(k) := \nu_2[N, \ell, b, D, B, L](k) := \max\{b(4N\ell(k+1) - 1), \nu_1(2\ell(k+1) - 1)\},$$

with $G(k) := G[N, B, L](k) := \max\{B(4N(k+1) - 1), L(10N(k+1) - 1)\}$, and θ is as in Lemma 2.5 and $\mathbf{0}$ is the zero function.

Proof. For all $n \in \mathbb{N}$ we have

$$\begin{aligned} \|T(\beta_n x_n)\| &\leq \|T(\beta_n x_n) - T(\beta_n p)\| + \|T(\beta_n p) - T(p)\| + \|T(p)\| \\ &\leq \|x_n - p\| + (1 - \beta_n) \|p\| + \|p\| \leq 3N \end{aligned} \tag{1}$$

Using (1) we derive

$$\begin{aligned} \|x_{n+1} - x_n\| &\leq \|(1 - \lambda_n)(\beta_n x_n - \beta_{n-1} x_{n-1}) + (\lambda_{n-1} - \lambda_n)\beta_{n-1} x_{n-1}\| \\ &\quad + \|\lambda_n(T(\beta_n x_n) - T(\beta_{n-1} x_{n-1})) + (\lambda_n - \lambda_{n-1})T(\beta_{n-1} x_{n-1})\| \\ &\leq \|\beta_n x_n - \beta_{n-1} x_{n-1}\| + |\lambda_n - \lambda_{n-1}| (\|\beta_{n-1} x_{n-1}\| + \|T(\beta_{n-1} x_{n-1})\|) \\ &\leq \|\beta_n x_n - \beta_{n-1} x_{n-1}\| + 5N|\lambda_n - \lambda_{n-1}|. \end{aligned}$$

For $n \geq 1$, we have

$$\begin{aligned} \|x_{n+1} - x_n\| &= \|\beta_n(x_n - x_{n-1}) + (\beta_n - \beta_{n-1})x_{n-1}\| + 5N|\lambda_n - \lambda_{n-1}| \\ &\leq \beta_n \|x_n - x_{n-1}\| + 2N|\beta_n - \beta_{n-1}| + 5N|\lambda_n - \lambda_{n-1}|. \end{aligned}$$

Observe that the function G is a Cauchy rate for $(\sum \gamma_n)$, where (γ_n) is given by $\gamma_n := 2N|\beta_n - \beta_{n-1}| + 5N|\lambda_n - \lambda_{n-1}|$. Indeed, for all n, k we have

$$\begin{aligned} \sum_{i=G(k)+1}^{G(k)+n} \gamma_i &= 2N \sum_{i=G(k)+1}^{G(k)+n} |\beta_i - \beta_{i-1}| + 5N \sum_{i=G(k)+1}^{G(k)+n} |\lambda_i - \lambda_{i-1}| \\ &\leq 2N \sum_{i=B(4N(k+1)-1)+1}^{G(k)+n} |\beta_i - \beta_{i-1}| + 5N \sum_{i=L(10N(k+1)-1)+1}^{G(k)+n} |\lambda_i - \lambda_{i-1}| \\ &\leq \frac{2N}{4N(k+1)} + \frac{5N}{10N(k+1)} = \frac{1}{k+1}. \end{aligned}$$

Applying Lemma 2.5 with $d = 2N$, and for all $n \geq 1$, $s_n = \|x_n - x_{n-1}\|$, $r_n = 0$, $\gamma_n = 2N|\beta_n - \beta_{n-1}| + 5N|\lambda_n - \lambda_{n-1}|$ and $\alpha_n = 1 - \beta_n$ we conclude Part (i). As for Part (ii), observe that for all $n \in \mathbb{N}$

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

Hence, for $n \geq \nu_2(k)$

$$\begin{aligned} \|x_n - T(x_n)\| &\leq \frac{1}{\lambda_n} (\|x_{n+1} - x_n\| + (1 - \lambda_n)(1 - \beta_n) \|T(x_n)\| + \lambda_n(1 - \beta_n) \|x_n\|) \\ &\leq \frac{1}{\lambda_n} (\|x_{n+1} - x_n\| + 2N(1 - \lambda_n)(1 - \beta_n) + 2N\lambda_n(1 - \beta_n)) \\ &\leq \ell (\|x_{n+1} - x_n\| + 2N(1 - \beta_n)) \leq \frac{1}{k+1}, \end{aligned}$$

which shows Part (ii). \square

The next result provides quantitative information on the fact that, with $\tilde{x} = \text{proj}_{\text{Fix } T}(0)$, one has $\limsup \langle \tilde{x}, \tilde{x} - x_n \rangle \leq 0$. Notice that, unlike the original proof, the quantitative version below does not require sequential weak compactness – the elimination of this principle is justified in [6].

Lemma 3.4. *Let $N \in \mathbb{N} \setminus \{0\}$ be such that $N \geq \max\{\|x_0 - p\|, \|p\|\}$, for some fixed point p . Consider $\ell \in \mathbb{N} \setminus \{0\}$ and monotone functions $b, D, B, L : \mathbb{N} \rightarrow \mathbb{N}$ satisfying conditions $(Q_2) - (Q_6)$. For any $k \in \mathbb{N}$ and monotone function $f : \mathbb{N} \rightarrow \mathbb{N}$, there are $n \leq \psi(k, f)$ and $x \in B_{2N}$ such that*

$$\|T(x) - x\| \leq \frac{1}{f(n) + 1} \wedge \forall m \geq n (\langle x, x - x_m \rangle \leq \frac{1}{k+1}),$$

where $\psi(k, f) := \psi[N, \ell, b, D, B, L](k, f) := \nu_2(48N(\check{g}^{(R)}(0) + 1)^2)$, $g(m) := f(\nu_2(m))$, $R := 64N^4(k+1)^2$ and $\check{g} := \max\{g(48N(m+1)^2), 48N(m+1)^2\}$, with ν_2 as in Lemma 3.3.

Proof. Given $k \in \mathbb{N}$ and monotone $f : \mathbb{N} \rightarrow \mathbb{N}$, applying Lemma 2.7 to k and the function g one obtains $n_0 \leq 48N(\check{g}^{(R)}(0) + 1)^2$ and $x \in B_{2N}$ such that $\|T(x) - x\| \leq \frac{1}{g(n_0)+1}$ and for all $y \in B_{2N}$

$$\|T(y) - y\| \leq \frac{1}{n_0 + 1} \rightarrow \langle x, x - y \rangle \leq \frac{1}{k+1}. \quad (2)$$

Define $n := \nu_2(n_0)$. By monotonicity, $n \leq \psi(k, f)$ and by the definition of g we conclude that $\|T(x) - x\| \leq \frac{1}{f(n)+1}$. By Part (ii) of Lemma 3.3 we have $\|T(x_m) - x_m\| \leq \frac{1}{n_0+1}$, for all $m \geq n$. The result follows from (2) since $(x_n) \subset B_{2N}$. \square

We are now able to prove our main result.

Theorem 3.5. *Given $(\beta_n), (\lambda_n) \subset (0, 1]$, $x_0 \in H$ and a nonexpansive mapping T , let (x_n) be generated by (T-KM). Assume that there exist $\ell \in \mathbb{N} \setminus \{0\}$ and monotone functions $b, D, B, L : \mathbb{N} \rightarrow \mathbb{N}$ and $h : \mathbb{N} \rightarrow \mathbb{N} \setminus \{0\}$ such that the conditions $(Q_1) - (Q_6)$ hold. Assume that there exists $N \in \mathbb{N} \setminus \{0\}$ such that $N \geq \max\{\|x_0 - p\|, \|p\|\}$, for some $p \in \text{Fix } T$. Then for all $k \in \mathbb{N}$ and monotone function $f : \mathbb{N} \rightarrow \mathbb{N}$*

$$\exists n \leq \mu(k, f) \forall i, j \in [n, f(n)] (\|x_i - x_j\| \leq \frac{1}{k+1}),$$

where $\mu(k, f) := \mu[N, \ell, b, h, D, B, L](k, f) := \sigma(\tilde{f}, \max\{\psi(12\tilde{k} + 1) - 1, \tilde{f}\}, n_1)$, with $\tilde{k} := 4(k+1)^2 - 1$, $n_1 := b(54N^2(\tilde{k} + 1) - 1)$, $\tilde{f}(m) := f[k, N, f, h](m) := 3(10N + 1)(\tilde{k} + 1)(\tilde{f}(m) + 1)h(\tilde{f}(m)) - 1$, for $\tilde{f}(m) := \tilde{f}[k, N, b, f, D](m) := f(\sigma(\tilde{k}, \max\{m, n_1\}))$, ψ as in Lemma 3.4 and $\sigma := \sigma[D, 9N^2]$ as in Lemma 2.6.

Proof. Let $k \in \mathbb{N}$ and $f : \mathbb{N} \rightarrow \mathbb{N}$ monotone be given. With $\tilde{k} := 4(k+1)^2 - 1$ and $n_1 := b(54N^2(\tilde{k} + 1) - 1)$ we define the functions

$$\tilde{f}(m) := f(\sigma(\tilde{k}, \max\{m, n_1\}))$$

and

$$\tilde{f}(m) := 3(10N + 1)(\tilde{k} + 1)(\tilde{f}(m) + 1)h(\tilde{f}(m)) - 1.$$

By Lemma 3.4 there exist $n_0 \leq \psi(12(\tilde{k} + 1) - 1, \tilde{f})$ and $x \in B_{2N}$ such that $\|T(x) - x\| \leq \frac{1}{\tilde{f}(n_0) + 1}$ and

$$\langle x, x - x_m \rangle \leq \frac{1}{12(\tilde{k} + 1)}, \text{ for all } m \geq n_0. \quad (3)$$

We have, for all $n \in \mathbb{N}$

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

Let $w_n := \|T(x) - x\| (2 \|\beta_n x_n - x\| + \|T(x) - x\|)$. Using (Q₁), for all $n \in \mathbb{N}$

$$\begin{aligned} \|x_{n+1} - x\|^2 &\leq (\|\beta_n x_n - x\| + \|T(x) - x\|)^2 \\ &= \|\beta_n x_n - x\|^2 + w_n \\ &= \|\beta_n(x_n - x) + (\beta_n - 1)x\|^2 + w_n \\ &= \beta_n^2 \|x_n - x\|^2 + 2\beta_n(1 - \beta_n)\langle x, x - x_n \rangle + (1 - \beta_n)^2 \|x\|^2 + w_n \\ &\leq \beta_n \|x_n - x\|^2 + (1 - \beta_n)(2\beta_n\langle x, x - x_n \rangle + (1 - \beta_n)\|x\|^2) + w_n \\ &\leq \beta_n(\|x_n - x\|^2 + \frac{w_n}{\beta_n}) + (1 - \beta_n)(2\beta_n\langle x, x - x_n \rangle + (1 - \beta_n)\|x\|^2) \\ &\leq \beta_n(\|x_n - x\|^2 + h(n)w_n) + (1 - \beta_n)(2\beta_n\langle x, x - x_n \rangle + (1 - \beta_n)\|x\|^2) \end{aligned}$$

We will apply Lemma 2.6 with $s_n := \|x_n - x\|^2$, $v_n := h(n)w_n$, $\alpha_n := 1 - \beta_n$, $r_n := 2\beta_n\langle x, x - x_n \rangle + (1 - \beta_n)\|x\|^2$, A is instantiated with the function D and $d := 9N^2$. Let us see that the conditions of the lemma are satisfied. By (Q₃), it is clear that A satisfies the required condition. We have that

$$s_n = \|x_n - x\|^2 \leq (\|x_n - p\| + \|x - p\|)^2 \leq (N + 2N)^2 = d.$$

With $n := \max\{n_0, n_1\}$ by (3) and (Q₂), for all $m \geq n$,

$$r_m = 2\beta_m\langle x, x - x_m \rangle + (1 - \beta_m)\|x\|^2 \leq \frac{2}{12(\tilde{k} + 1)} + \frac{(\|x - p\| + \|p\|)^2}{54N^2(\tilde{k} + 1)} \leq \frac{1}{3(\tilde{k} + 1)}.$$

Observe that $2 \|\beta_n x_n - x\| + \|T(x) - x\| \leq 2(\|x_n\| + \|x - p\| + \|p\|) + 1 \leq 10N + 1$. For $m \leq \bar{f}(n_0) = f(\sigma(\tilde{k}, n))$, using the fact that the function h is monotone we have

$$\begin{aligned} v_m &= h(m) \|T(x) - x\| (2 \|\beta_m x_m - x\| + \|T(x) - x\|) \leq \frac{h(\bar{f}(n_0))(10N + 1)}{\bar{f}(n_0) + 1} \\ &= \frac{h(\bar{f}(n_0))(10N + 1)}{3(10N + 1)(\tilde{k} + 1)(\bar{f}(n_0) + 1)h(\bar{f}(n_0))} = \frac{1}{3(\tilde{k} + 1)(f(\sigma(\tilde{k}, n)) + 1)} \end{aligned}$$

This shows that $v_m \leq \frac{1}{3(\tilde{k} + 1)(f(\sigma(\tilde{k}, n)) + 1)}$ and $r_m \leq \frac{1}{3(\tilde{k} + 1)}$, for all $m \in [n, f(\sigma(\tilde{k}, n))]$. Hence, by Lemma 2.6

$$\forall m \in [\sigma(\tilde{k}, n), f(\sigma(\tilde{k}, n))] \left(\|x_m - x\|^2 \leq \frac{1}{\tilde{k} + 1} \right). \quad (4)$$

We conclude that for $i, j \in [\sigma(\tilde{k}, n), f(\sigma(\tilde{k}, n))]$

$$\|x_i - x_j\| \leq \|x_i - x\| + \|x_j - x\| \leq \frac{1}{2(\tilde{k} + 1)} + \frac{1}{2(\tilde{k} + 1)} = \frac{1}{\tilde{k} + 1},$$

which entails the result since $\sigma(\tilde{k}, n) \leq \mu(\tilde{k}, f)$. \square

Remark 3.6. The proof of metastability result above does not require the projection argument neither sequential weak compactness. Nevertheless, in the presence of the projection point $\text{proj}_{\text{Fix } T}(0)$, following the arguments culminating in (4) one can say that Theorem 3.5 corresponds to a quantitative version of Theorem 2.1.

Corollary 3.7. Let $\alpha \in (0, 1]$ and $T : H \rightarrow H$ be α -averaged. Given $(\beta_n) \subset (0, 1]$, $(\lambda_n) \subset (0, \frac{1}{\alpha}]$ and $x_0 \in H$, consider (x_n) generated by (T-KM). Let $a \in \mathbb{N} \setminus \{0\}$ be such that $\alpha \geq \frac{1}{a}$. Assume that there exist $\ell \in \mathbb{N} \setminus \{0\}$ and monotone functions $b, D, B, L : \mathbb{N} \rightarrow \mathbb{N}$ and $h : \mathbb{N} \rightarrow \mathbb{N} \setminus \{0\}$ such that the conditions (Q₁) – (Q₆) hold.

Assume that there exists $N \in \mathbb{N} \setminus \{0\}$ such that $N \geq \max\{\|x_0 - p\|, \|p\|\}$, for some $p \in \text{Fix } T$. Then for all $k \in \mathbb{N}$ and monotone function $f : \mathbb{N} \rightarrow \mathbb{N}$

$$\exists n \leq \mu_1(k, f) \forall i, j \in [n, f(n)] (\|x_i - x_j\| \leq \frac{1}{k+1}),$$

where $\mu_1(k, f) := \mu_1[a, N, \ell, b, h, D, B, L](k, f) := \mu[N, a\ell, b, h, D, B, L](k, f)$, with μ as in Theorem 3.5.

Proof. Since T is α -averaged, there exists T' nonexpansive and such that $T = (1 - \alpha)\text{Id} + \alpha T'$. It is easy to see that (x_n) is generated by (T-KM), with starting point x_0 , using the sequences $(\beta_n), (\alpha\lambda_n) \subset (0, 1]$ and T' . Since $\text{Fix } T = \text{Fix } T'$, by Theorem 3.5, we just need to check that the conditions (Q5) and (Q6) still hold using the sequence $(\alpha\lambda_n)$ instead of the sequence (λ_n) . The condition (Q5) holds with $\ell := a\ell$, since for all $n \in \mathbb{N}$ we have $\alpha\lambda_n \geq \frac{\alpha}{\ell} \geq \frac{1}{a\ell}$. We have, for all $k, n \in \mathbb{N}$ that since $\alpha \in (0, 1]$

$$\sum_{i=L(k)+1}^{L(k)+n} |\alpha\lambda_i - \alpha\lambda_{i-1}| \leq \sum_{i=L(k)+1}^{L(k)+n} |\lambda_i - \lambda_{i-1}| \leq \frac{1}{k+1}.$$

Hence (Q6) is also satisfied for the sequence $(\alpha\lambda_n)$ with the function L . \square

Corollary 3.8. Let $T_1 : H \rightrightarrows H$ be maximal monotone and $T_2 : H \rightarrow H$ be δ -cocoercive, for some $\delta > 0$. Let $\gamma \in (0, 2\delta]$. Given $(\beta_n) \subset (0, 1]$, $(\lambda_n) \subset (0, \frac{4\delta - \gamma}{2\delta}]$ and $x_0 \in H$, consider (x_n) generated, for all $n \in \mathbb{N}$, by

$$x_{n+1} = (1 - \lambda_n)\beta_n x_n + \lambda_n J_{\gamma T_1}(\beta_n x_n - \gamma T_2(\beta_n x_n)) \quad (\text{T-FB})$$

Assume that there exist $\ell \in \mathbb{N} \setminus \{0\}$ and monotone functions $b, D, B, L : \mathbb{N} \rightarrow \mathbb{N}$ and $h : \mathbb{N} \rightarrow \mathbb{N} \setminus \{0\}$ such that the conditions (Q1) – (Q6) hold. Assume that there exists $N \in \mathbb{N} \setminus \{0\}$ such that $N \geq \max\{\|x_0 - p\|, \|p\|\}$, for some $p \in \text{zer}(T_1 + T_2)$. Then for all $k \in \mathbb{N}$ and monotone function $f : \mathbb{N} \rightarrow \mathbb{N}$

$$\exists n \leq \mu_2(k, f) \forall i, j \in [n, f(n)] (\|x_i - x_j\| \leq \frac{1}{k+1}),$$

where $\mu_2(k, f) := \mu_2[N, \ell, b, h, D, B, L](k, f) := \mu[N, 2\ell, b, h, D, B, L](k, f)$, with μ as in Theorem 3.5.

Proof. It is straightforward to see that the iteration (x_n) is generated by (T-KM) with $T := J_{\gamma T_1} \circ (\text{Id} - \gamma T_2)$. The resolvent function $J_{\gamma T_1}$ is firmly nonexpansive, i.e. $\frac{1}{2}$ -averaged. Furthermore, we have $\frac{\gamma}{2\delta} \in (0, 1]$ and, using [1, Proposition 4.33] it follows that $(\text{Id} - \gamma T_2)$ is $\frac{\gamma}{2\delta}$ -averaged. If $\gamma < 2\delta$, then using [13, Theorem 3(b)] it follows that T is $\frac{2\delta}{4\delta - \gamma}$ -averaged. If $\gamma = 2\delta$, then T is nonexpansive and therefore also $\frac{2\delta}{4\delta - \gamma}$ -averaged. Since $\text{Fix } T = \text{zer}(T_1 + T_2)$ [1, Proposition 25.1(iv)] we have $N \geq \max\{\|x_0 - p\|, \|p\|\}$, for some $p \in \text{Fix } T$. Noting that $\frac{2\delta}{4\delta - \gamma} \in [\frac{1}{2}, 1]$, we may apply Corollary 3.7 with $a = 2$ and the result follows. \square

Remark 3.9. If in Corollary 3.8 the resolvent function $J_{\gamma T_1}$ is replaced by an arbitrary firmly nonexpansive mapping T_1 , and p is some point in $\text{Fix}(T_1 \circ (\text{Id} - \gamma T_2))$, then the result holds also for $\gamma = 0$, by Corollary 3.7. In such case, the sequence (λ_n) is allowed to vary in the interval $(0, 2]$.

Corollary 3.10. Let $T_1, T_2 : H \rightrightarrows H$ be two maximal monotone operators and $\gamma > 0$. Given $(\beta_n) \subset (0, 1]$, $(\lambda_n) \subset (0, 2]$ and $x_0 \in H$, consider (x_n) generated, for all $n \in \mathbb{N}$, by

$$\begin{cases} y_n = J_{\gamma T_2}(\beta_n x_n) \\ z_n = J_{\gamma T_1}(2y_n - \beta_n x_n) \\ x_{n+1} = \beta_n x_n + \lambda_n(z_n - y_n) \end{cases} \quad (\text{T-DR})$$

Assume that there exist $\ell \in \mathbb{N} \setminus \{0\}$ and monotone functions $b, D, B, L : \mathbb{N} \rightarrow \mathbb{N}$ and $h : \mathbb{N} \rightarrow \mathbb{N} \setminus \{0\}$ such that the conditions (Q1) – (Q6) hold. Assume that there exists $N \in \mathbb{N} \setminus \{0\}$ such that $N \geq \max\{\|x_0 - p\|, \|p\|\}$, for some $p \in \text{Fix}(R_{\gamma T_1} \circ R_{\gamma T_2})$. Then, for all $k \in \mathbb{N}$ and monotone function $f : \mathbb{N} \rightarrow \mathbb{N}$

$$(i) \exists n \leq \mu_3(k, f) \forall i, j \in [n, f(n)] (\|x_i - x_j\| \leq \frac{1}{k+1}),$$

$$(ii) \exists n \leq \mu_4(k, f) \forall i, j \in [n, f(n)] (\|y_i - y_j\| \leq \frac{1}{k+1}),$$

$$(iii) \exists n \leq \mu_5(k, f) \forall i, j \in [n, f(n)] (\|z_i - z_j\| \leq \frac{1}{k+1}),$$

where $\mu_3(k, f) := \mu_3[N, \ell, b, h, D, B, L](k, f) := \mu[N, 2\ell, b, h, D, B, L](k, f)$, $\mu_4(k, f) := \max\{\mu_3(2k+1, g_1), b(8N(k+1) - 1)\}$ and

$$\mu_5(k, f) := \max\{\mu_4(3k+2, g_2), \nu_1(6\ell(k+1) - 1), b(12\ell N(k+1) - 1)\},$$

with μ as in Theorem 3.5, and $g_1(m) := f(\max\{m, b(8N(k+1) - 1)\})$ and $g_2(m) := f(\max\{m, \nu_1(6\ell(k+1) - 1), b(12\ell N(k+1) - 1)\})$.

Proof. It is easy to see that the iteration (x_n) is generated by (T-KM) using the sequences (β_n) and $(\frac{\lambda_n}{2})$, and with $T := R_{\gamma T_1} \circ R_{\gamma T_2}$, which is nonexpansive. Hence, applying Theorem 3.5 we obtain Part (i).

By Part (i), there exists $n_0 \leq \mu_3(2k+1, g_1)$ such that $\|x_i - x_j\| \leq \frac{1}{2(k+1)}$, for all $i, j \in [n_0, g_1(n_0)]$. Define $n := \max\{n_0, b(8N(k+1) - 1)\} (\leq \mu_4(k, f))$. Since $[n, f(n)] \subseteq [n_0, g_1(n_0)]$, we have for $i, j \in [n, f(n)]$

$$\begin{aligned} \|y_i - y_j\| &\leq \|\beta_i x_i - \beta_j x_j\| \leq \|x_i - x_j\| + |\beta_i - \beta_j| \|x_j\| \\ &\leq \frac{1}{2(k+1)} + \left(\frac{1}{8N(k+1)} + \frac{1}{8N(k+1)} \right) 2N = \frac{1}{k+1}, \end{aligned}$$

which shows Part (ii).

Define $n_1 := \max\{\nu_1(6\ell(k+1) - 1), b(12\ell N(k+1) - 1)\}$, where ν_1 is as in Lemma 3.3. By the definition of x_{n+1} , for $i \geq n_1$

$$\begin{aligned} \|z_i - y_i\| &\leq \ell \|x_{i+1} - \beta_i x_i\| \leq \ell(\|x_{i+1} - x_i\| + 2N(1 - \beta_i)) \\ &\leq \frac{\ell}{6\ell(k+1)} + \frac{2\ell N}{12\ell N(k+1)} = \frac{1}{3(k+1)} \end{aligned}$$

By Part (ii), there exists $n_2 \leq \mu_4(3k+2, g_2)$ such that $\|y_i - y_j\| \leq \frac{1}{3(k+1)}$, for all $i, j \in [n_2, g_2(n_2)]$. Define $n := \max\{n_1, n_2\} (\leq \mu_5(k, f))$. Since $[n, f(n)] \subseteq [n_2, g_2(n_2)]$, we have for $i, j \in [n, f(n)]$

$$\|z_i - z_j\| \leq \|z_i - y_i\| + \|y_i - y_j\| + \|y_j - z_j\| \leq \frac{1}{k+1},$$

which concludes the result. □

References

- [1] Heinz Bauschke and Patrick Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, New York, 2011. With a foreword by Hedy Attouch.
- [2] Marc Bezem. Strongly majorizable functionals of finite type: A model for barrecursion containing discontinuous functionals. *J. Symb. Log.*, 50(3):652–660, 1985.
- [3] Radu Ioan Boț, Ernő Robert Csetnek, and Dennis Meier. Inducing strong convergence into the asymptotic behaviour of proximal splitting algorithms in Hilbert spaces. *Optim. Methods Softw.*, 34(3):489–514, 2019.
- [4] Bruno Dinis and Pedro Pinto. Metastability of the proximal point algorithm with multi-parameters. (submitted) arXiv:1906.09129v2.
- [5] Bruno Dinis and Pedro Pinto. Quantitative results on the multi-parameters proximal point algorithm. (To appear in *Journal of Convex Analysis*).
- [6] Fernando Ferreira, Laurențiu Leuştean, and Pedro Pinto. On the removal of weak compactness arguments in proof mining. *Advances in Mathematics*, 354:106728, 2019.
- [7] Ulrich Kohlenbach. *Applied proof theory: proof interpretations and their use in mathematics*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2008.
- [8] Ulrich Kohlenbach. On quantitative versions of theorems due to F. E. Browder and R. Wittmann. *Adv. Math.*, 226(3):2764–2795, 2011.
- [9] Ulrich Kohlenbach. Recent progress in proof mining in nonlinear analysis. *IFCoLog Journal of Logics and their Applications*, 10:3357–3406, 2017.
- [10] Noureddine Lehdili and Abdellatif Moudafi. Combining the proximal algorithm and tikhonov regularization. *Optimization*, 37(3):239–252, 1996.
- [11] Laurențiu Leuştean and Pedro Pinto. Quantitative results on Halpern type proximal point algorithms. (submitted) arXiv:2001.10040.
- [12] Bernard Martinet. Régularisation d'inéquations variationnelles par approximations successives. *Rev. Française Informat. Recherche Opérationnelle*, 4(Sér. R-3):154–158, 1970.
- [13] Nobuhiko Ogura and Isao Yamada. Non-strictly convex minimization over the fixed point set of an asymptotically shrinking nonexpansive mapping. *Numer. Funct. Anal. Optim.*, 23(1-2):113–137, 2002.

- [14] Pedro Pinto. A rate of metastability for the Halpern type Proximal Point Algorithm. (submitted) arXiv:1912.12468.
- [15] Ralph Rockafellar. Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization*, 14(5):877–898, 1976.