Quasi-Nonexpansive Iterations on the Affine Hull of Orbits: From Mann's Mean Value Algorithm to Inertial Methods

Patrick L Combettes, Lilian E. Glaudin

SIAM J. OPTIM.
Vol. 27, No. 4, pp. 2356–2380

QUASI-NONEXPANSIVE ITERATIONS ON THE AFFINE HULL OF ORBITS: FROM MANN'S MEAN VALUE ALGORITHM TO INERTIAL METHODS*

PATRICK L. COMBETTES† AND LILIAN E. GLAUDIN‡

Dedicated to the memory of Felipe Álvarez 1972–2017

Abstract. Fixed point iterations play a central role in the design and the analysis of a large number of optimization algorithms. We study a new iterative scheme in which the update is obtained by applying a composition of quasi-nonexpansive operators to a point in the affine hull of the orbit generated up to the current iterate. This investigation unifies several algorithmic constructs, including Mann's mean value method, inertial methods, and multilayer memoryless methods. It also provides a framework for the development of new algorithms, such as those we propose for solving monotone inclusion and minimization problems.

Key words. averaged operator, fixed point iteration, forward-backward algorithm, inertial algorithm, mean value iterations, monotone operator splitting, nonsmooth minimization, Peaceman–Rachford algorithm, proximal algorithm

AMS subject classifications. Primary, 65J15, 47H09; Secondary, 47H05, 65K05, 90C25

DOI. 10.1137/17M112806X

1. Introduction.

Algorithms arising in various branches of optimization can be efficiently modeled and analyzed as fixed point iterations in a real Hilbert space $\mathcal{H}$; see, e.g., [9, 10, 13, 16, 18, 19, 22, 26, 43]. Our paper unifies three important algorithmic fixed point frameworks that coexist in the literature: mean value methods, inertial methods, and multilayer memoryless methods.

Let $T\colon \mathcal{H} \to \mathcal{H}$ be an operator with fixed point set $\operatorname{Fix} T$. In 1953, inspired by classical results on the summation of divergent series [11, 29, 44], Mann [34] proposed to extend the standard successive approximation scheme

$$x_0 \in \mathcal{H} \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad x_{n+1} = Tx_n \tag{1.1}$$

to the mean value algorithm

$$x_0 \in \mathcal{H} \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad x_{n+1} = T\overline{x}_n, \quad \text{where} \quad \overline{x}_n \in \operatorname{conv}\big(x_j\big)_{0\leqslant j\leqslant n}. \tag{1.2}$$

In other words, the operator T is not applied to the most current iterate as in the memoryless (single step) process (1.1), but to a point in the convex hull of the orbit $(x_j)_{0\leqslant j\leqslant n}$ generated so far. His motivation was that, although the sequence generated by (1.1) may fail to converge to a fixed point of T, that generated by (1.2) can, under suitable conditions. This work was followed by interesting developments and analyses of such mean value iterations (e.g., [8, 12, 15, 28, 30, 32, 35, 37, 42]), especially in the case when T is nonexpansive (1-Lipschitzian) or merely quasi-nonexpansive, that is (this notion was essentially introduced in [27]),

$$(\forall x \in \mathcal{H})(\forall y \in \operatorname{Fix} T) \quad \|Tx - y\| \leqslant \|x - y\|. \tag{1.3}$$

*Received by the editors May 1, 2017; accepted for publication (in revised form) August 7, 2017; published electronically November 16, 2017.
http://www.siam.org/journals/siopt/27-4/M112806.html
†Corresponding author. North Carolina State University, Department of Mathematics, Raleigh, NC 27695-8205 (plc@math.ncsu.edu).
‡Sorbonne Universités – UPMC Univ. Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France (glaudin@ljll.math.upmc.fr).

In [21], the asymptotic behavior of the mean value process

$$x_0 \in \mathcal{H} \text{ and } (\forall n \in \mathbb{N}) \quad x_{n+1} = \overline{x}_n + \lambda_n\big(T_n\overline{x}_n + e_n - \overline{x}_n\big), \text{ where } \overline{x}_n \in \operatorname{conv}\big(x_j\big)_{0\leqslant j\leqslant n}, \tag{1.4}$$

was investigated under general conditions on the construction of the averaging process $(\overline{x}_n)_{n\in\mathbb{N}}$ and the assumptions that, for every $n \in \mathbb{N}$, $e_n \in \mathcal{H}$ models a possible error made in the computation of $T_n\overline{x}_n$, $\lambda_n \in \left]0,2\right[$, and $T_n\colon \mathcal{H} \to \mathcal{H}$ is firmly quasi-nonexpansive, i.e., $2T_n - \mathrm{Id}$ is quasi-nonexpansive or, equivalently [10],

$$(\forall x \in \mathcal{H})(\forall y \in \operatorname{Fix} T_n) \quad \langle y - T_n x \mid x - T_n x\rangle \leqslant 0. \tag{1.5}$$

The idea of using the past of the orbit generated by an algorithm can also be found in the work of Polyak [39, 41], who drew inspiration from classical multistep methods in numerical analysis. His motivation was to improve the speed of convergence over memoryless methods. For instance, the classical gradient method [38] for minimizing a smooth convex function $f\colon \mathcal{H} \to \mathbb{R}$ is an explicit discretization of the continuous-time process $-\dot{x}(t) = \nabla f(x(t))$. Polyak [39] proposed to consider instead the process $-\ddot{x}(t) - \beta\dot{x}(t) = \nabla f(x(t))$, where $\beta \in \left]0,+\infty\right[$, and studied the algorithm resulting from its explicit discretization. He observed that, from a mechanical viewpoint, the term $\ddot{x}(t)$ can be interpreted as an inertial component. More generally, for a proper lower semicontinuous convex function $f\colon \mathcal{H} \to \left]-\infty,+\infty\right]$, Álvarez investigated in [1] an implicit discretization of the inertial differential inclusion $-\ddot{x}(t) - \beta\dot{x}(t) \in \partial f(x(t))$, namely,

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \operatorname{prox}_{\gamma_n f}\overline{x}_n + e_n, \quad \text{where} \quad \begin{cases} \overline{x}_n = (1+\eta_n)x_n - \eta_n x_{n-1}, \\ \eta_n \in \left[0,1\right[, \\ \gamma_n \in \left]0,+\infty\right[, \end{cases} \tag{1.6}$$

and where prox_f is the proximity operator of f [10, 36]. The inertial proximal point algorithm (1.6) has been extended in various directions; e.g., [3, 14, 17]; see also [5] for further motivation in the context of nonconvex minimization problems.

Working from a different perspective, a structured extension of (1.1) involving the composition of m averaged nonexpansive operators was proposed in [19]. This m-layer algorithm is governed by the memoryless recursion

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n\big(T_{1,n}\cdots T_{m,n}x_n + e_n - x_n\big), \quad \text{where} \quad \lambda_n \in \left]0,1\right]. \tag{1.7}$$

Recall that a nonexpansive operator $T\colon \mathcal{H} \to \mathcal{H}$ is averaged with constant $\alpha \in \left]0,1\right[$ if there exists a nonexpansive operator $R\colon \mathcal{H} \to \mathcal{H}$ such that $T = (1-\alpha)\,\mathrm{Id} + \alpha R$ [7, 10]. The multilayer iteration process (1.7) was shown in [19] to provide a synthetic analysis of various algorithms, in particular in the area of monotone operator splitting methods. It was extended in [25] to an overrelaxed method, i.e., one with parameters $(\lambda_n)_{n\in\mathbb{N}}$ possibly larger than 1.

In the literature, the asymptotic analysis of the above methods has been carried out independently because of their apparent lack of common structure. In the present paper, we exhibit a structure that unifies (1.1), (1.2), (1.4), (1.6), and (1.7) in a single algorithm of the form

$$\begin{aligned} &x_0 \in \mathcal{H} \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad x_{n+1} = \overline{x}_n + \lambda_n\big(T_{1,n}\cdots T_{m,n}\overline{x}_n + e_n - \overline{x}_n\big), \\ &\qquad\qquad \text{where} \quad \overline{x}_n \in \operatorname{aff}\big(x_j\big)_{0\leqslant j\leqslant n} \quad \text{and} \quad \lambda_n \in \left]0,+\infty\right[\end{aligned} \tag{1.8}$$

under the assumption that each operator $T_{i,n}$ is $\alpha_{i,n}$-averaged quasi-nonexpansive, i.e.,

$$\begin{aligned} &(\forall x \in \mathcal{H})(\forall y \in \operatorname{Fix} T_{i,n}) \\ &\qquad 2(1-\alpha_{i,n})\langle y - T_{i,n}x \mid x - T_{i,n}x\rangle \leqslant (2\alpha_{i,n}-1)\big(\|x-y\|^2 - \|T_{i,n}x - y\|^2\big) \end{aligned} \tag{1.9}$$

for some $\alpha_{i,n} \in]0,1]$, which means that the operator $(1-1/\alpha_{i,n})\,\mathrm{Id} + (1/\alpha_{i,n})T_{i,n}$ is quasi-nonexpansive. In words, at iteration n, a point $\overline{x}_n$ is picked in the affine hull of the orbit $(x_j)_{0\leqslant j\leqslant n}$ generated so far, a composition of quasi-nonexpansive operators is applied to it, up to some error e_n, and the update x_{n+1} is obtained via a relaxation with parameter λ_n. Note that (1.8)–(1.9) not only brings together mean value iterations, inertial methods, and the memoryless multilayer setting of [19, 25], but also provides a flexible framework to design new iterative methods.

The fixed point problem under consideration will be the following (note that we allow 1 as an averaging constant for added flexibility).

Problem 1.1. Let m be a strictly positive integer. For every $n \in \mathbb{N}$ and every $i \in \{1,\dots,m\}$, $\alpha_{i,n} \in]0,1]$ and $T_{i,n} \colon \mathcal{H} \to \mathcal{H}$ is $\alpha_{i,n}$-averaged nonexpansive if $i < m$, and $\alpha_{m,n}$-averaged quasi-nonexpansive if $i = m$. In addition,

$$S = \bigcap_{n\in\mathbb{N}} \operatorname{Fix} T_n \neq \varnothing, \quad \text{where} \quad (\forall n \in \mathbb{N}) \quad T_n = T_{1,n} \cdots T_{m,n}, \tag{1.10}$$

and one of the following holds:

(a) For every $n \in \mathbb{N}$, $T_{m,n}$ is $\alpha_{m,n}$-averaged nonexpansive.

(b) $m > 1$, and, for every $n \in \mathbb{N}$, $\alpha_{m,n} < 1$ and $\bigcap_{i=1}^m \operatorname{Fix} T_{i,n} \neq \varnothing$.

(c) $m = 1$.

The problem is to find a point in S.

To solve Problem 1.1, we are going to employ (1.8), which we now formulate more formally.

Algorithm 1.2.

Consider the setting of Problem 1.1. For every $n \in \mathbb{N}$, let ϕ_n be an averaging constant of T_n, let $\lambda_n \in]0, 1/\phi_n]$ and for every $i \in \{1,\dots,m\}$, let $e_{i,n} \in \mathcal{H}$. Let $(\mu_{n,j})_{n\in\mathbb{N}, 0\leqslant j\leqslant n}$ be a real array which satisfies the following:

(a) $\sup_{n\in\mathbb{N}} \sum_{j=0}^n |\mu_{n,j}| < +\infty$.

(b) $(\forall n \in \mathbb{N})\ \sum_{j=0}^n \mu_{n,j} = 1$.

(c) $(\forall j \in \mathbb{N})\ \lim_{n\to+\infty} \mu_{n,j} = 0$.

(d) There exists a sequence $(\chi_n)_{n\in\mathbb{N}}$ in $]0,+\infty[$ such that $\inf_{n\in\mathbb{N}} \chi_n > 0$ and every sequence $(\xi_n)_{n\in\mathbb{N}}$ in $[0,+\infty[$ that satisfies

$$\Big(\exists\, (\varepsilon_n)_{n\in\mathbb{N}} \in [0,+\infty[^{\mathbb{N}}\Big) \quad \begin{cases} \sum_{n\in\mathbb{N}} \chi_n \varepsilon_n < +\infty, \\ (\forall n \in \mathbb{N}) \quad \xi_{n+1} \leqslant \sum_{j=0}^n \mu_{n,j}\xi_j + \varepsilon_n \end{cases} \tag{1.11}$$

converges.

Let $x_0 \in \mathcal{H}$ and set

(1.12)

$$\begin{array}{l} \text{for } n = 0,1,\ldots \\ \left\lfloor \begin{array}{l} \overline{x}_n = \displaystyle\sum_{j=0}^n \mu_{n,j} x_j, \\ x_{n+1} = \overline{x}_n + \lambda_n\Big(T_{1,n}\Big(T_{2,n}\big(\cdots T_{m-1,n}(T_{m,n}\overline{x}_n + e_{m,n}) + e_{m-1,n}\cdots\big) + e_{2,n}\Big) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + e_{1,n} - \overline{x}_n\Big). \end{array} \right. \end{array}$$

Remark 1.3. Here are some comments about the parameters appearing in Problem 1.1 and Algorithm 1.2.

(i) The composite operator T_n of (1.10) is averaged quasi-nonexpansive with constant

$$\phi_n = \begin{cases} \left(1 + \left(\displaystyle\sum_{i=1}^{m} \frac{\alpha_{i,n}}{1-\alpha_{i,n}}\right)^{-1}\right)^{-1} & \text{if } \displaystyle\max_{1\leqslant i\leqslant m} \alpha_{i,n} < 1, \\ 1 & \text{otherwise.} \end{cases} \tag{1.13}$$

The proof is given in [25, Proposition 2.5] for case (a) of Problem 1.1. It easily extends to case (b), while case (c) is trivial.

(ii) Examples of arrays $(\mu_{n,j})_{n\in\mathbb{N},0\leqslant j\leqslant n}$ that satisfy conditions (a)–(d) in Algorithm 1.2 are provided in [21, section 2] in the case of mean value iterations, i.e., $\inf_{n\in\mathbb{N}} \min_{0\leqslant j\leqslant n} \mu_{n,j} \geqslant 0$ with $\chi_n \equiv 1$. An important instance with negative coefficients will be presented in Example 2.5.

(iii) The term $e_{i,n}$ in (1.12) models a possible numerical error in the implementation of the operator $T_{i,n}$.

The material is organized as follows. In section 2 we provide preliminary results. The main results on the convergence of the orbits of Algorithm 1.2 are presented in section 3. Section 4 is dedicated to new algorithms for fixed point computation, monotone operator splitting, and nonsmooth minimization based on the proposed framework.

Notation. $\mathcal{H}$ is a real Hilbert space with scalar product $\langle\cdot \mid \cdot\rangle$ and associated norm $\|\cdot\|$. We denote by Id the identity operator on $\mathcal{H}$; $\rightharpoonup$ and $\to$ denote, respectively, weak and strong convergence in $\mathcal{H}$. The positive and negative parts of $\xi \in \mathbb{R}$ are, respectively, $\xi^+ = \max\{0,\xi\}$ and $\xi^- = -\min\{0,\xi\}$. Finally, $\delta_{n,j}$ is the Kronecker delta: it takes on the value 1 if $n = j$, and 0 otherwise.

2. Preliminary results.

In this section we establish some technical facts that will be used subsequently. We start with a Grönwall-type result.

Lemma 2.1. *Let $(\theta_n)_{n\in\mathbb{N}}$ and $(\varepsilon_n)_{n\in\mathbb{N}}$ be sequences in $[0,+\infty[$, and let $(\nu_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$ such that $(\forall n \in \mathbb{N})\ \theta_{n+1} \leqslant (1+\nu_n)\theta_n + \varepsilon_n$. Then*

$$(\forall n \in \mathbb{N}) \quad \theta_{n+1} \leqslant \theta_0 \exp\left(\sum_{k=0}^{n} \nu_k\right) + \sum_{j=0}^{n-1} \varepsilon_j \exp\left(\sum_{k=j+1}^{n} \nu_k\right) + \varepsilon_n. \tag{2.1}$$

Proof. We have $(\forall n \in \mathbb{N})\ 1+\nu_n \leqslant \exp(\nu_n)$. Therefore $\theta_1 \leqslant \theta_0\exp(\nu_0) + \varepsilon_0$ and

$$\begin{aligned}(\forall n \in \mathbb{N}\smallsetminus\{0\}) \quad \theta_{n+1} &\leqslant \theta_n \exp(\nu_n) + \varepsilon_n \\ &\leqslant \theta_{n-1}\exp(\nu_n)\exp(\nu_{n-1}) + \varepsilon_{n-1}\exp(\nu_n) + \varepsilon_n \\ &\leqslant \theta_0 \prod_{k=0}^{n} \exp(\nu_k) + \sum_{j=0}^{n-1} \varepsilon_j \prod_{k=j+1}^{n} \exp(\nu_k) + \varepsilon_n \\ &= \theta_0 \exp\left(\sum_{k=0}^{n} \nu_k\right) + \sum_{j=0}^{n-1} \varepsilon_j \exp\left(\sum_{k=j+1}^{n} \nu_k\right) + \varepsilon_n, \qquad (2.2)\end{aligned}$$

as claimed. ∎

LEMMA 2.2 (see [31, Theorem 43.5]). *Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$, let $\xi \in \mathbb{R}$, suppose that $(\mu_{n,j})_{n\in\mathbb{N},0\leqslant j\leqslant n}$ is a real array that satisfies conditions* (a)–(c) *in Algorithm* 1.2. *Then $\xi_n \to \xi \Rightarrow \sum_{j=0}^n \mu_{n,j}\xi_j \to \xi$.*

LEMMA 2.3. *Let $(\beta_n)_{n\in\mathbb{N}}$, $(\gamma_n)_{n\in\mathbb{N}}$, $(\delta_n)_{n\in\mathbb{N}}$, $(\eta_n)_{n\in\mathbb{N}}$, and $(\lambda_n)_{n\in\mathbb{N}}$ be sequences in $[0,+\infty[$, let $(\phi_n)_{n\in\mathbb{N}}$ be a sequence in $]0,1]$, let $(\vartheta,\sigma) \in]0,+\infty[^2$, and let $\eta \in]0,1[$. Set $\beta_{-1} = \beta_0$ and*

$$(\forall n \in \mathbb{N}) \quad \omega_n = \frac{1}{\phi_n} - \lambda_n, \tag{2.3}$$

and suppose that the following hold:

(a) $(\forall n \in \mathbb{N})$ $\eta_n \leqslant \eta_{n+1} \leqslant \eta$.
(b) $(\forall n \in \mathbb{N})$ $\gamma_n \leqslant \eta(1+\eta) + \eta\vartheta\omega_n$.
(c) $(\forall n \in \mathbb{N})$ $\dfrac{\eta^2(1+\eta)+\eta\sigma}{\vartheta} < \dfrac{1}{\phi_n} - \eta^2\omega_{n+1}$.
(d) $(\forall n \in \mathbb{N})$ $0 < \lambda_n \leqslant \dfrac{\vartheta/\phi_n - \eta\big(\eta(1+\eta)+\eta\vartheta\omega_{n+1}+\sigma\big)}{\vartheta\big(1+\eta(1+\eta)+\eta\vartheta\omega_{n+1}+\sigma\big)}$.
(e) $(\forall n \in \mathbb{N})$ $\beta_{n+1} - \beta_n - \eta_n(\beta_n - \beta_{n-1}) \leqslant \dfrac{(1/\phi_n - \lambda_n)\big(\eta_n/(\eta_n+\vartheta\lambda_n)-1\big)}{\lambda_n}\delta_{n+1} + \gamma_n\delta_n$.

Then $\sum_{n\in\mathbb{N}} \delta_n < +\infty$.

Proof. We use arguments similar to those used in [3, 14]. It follows from (c) that $(\forall n \in \mathbb{N})$ $0 < \vartheta/\phi_n - \eta^2\omega_{n+1}\vartheta - \eta^2(1+\eta) - \eta\sigma$. This shows that $(\lambda_n)_{n\in\mathbb{N}}$ is well defined. Now set $(\forall n \in \mathbb{N})$ $\rho_n = 1/(\eta_n + \vartheta\lambda_n)$ and $\kappa_n = \beta_n - \eta_n\beta_{n-1} + \gamma_n\delta_n$. We derive from (a) and (e) that

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \kappa_{n+1} - \kappa_n &\leqslant \beta_{n+1} - \eta_n\beta_n - \beta_n + \eta_n\beta_{n-1} + \gamma_{n+1}\delta_{n+1} - \gamma_n\delta_n\\ &\leqslant \left(\frac{(1/\phi_n - \lambda_n)(\eta_n\rho_n - 1)}{\lambda_n} + \gamma_{n+1}\right)\delta_{n+1}.\end{aligned} \tag{2.4}$$

On the other hand, $(\forall n \in \mathbb{N})$ $\vartheta(1 + (\eta(1+\eta) + \eta\vartheta\omega_{n+1} + \sigma)) > 0$. Consequently, (d) can be written as

$$(\forall n \in \mathbb{N}) \quad \vartheta\lambda_n + \vartheta\lambda_n\big(\eta(1+\eta)+\eta\vartheta\omega_{n+1}+\sigma\big) \leqslant \frac{\vartheta}{\phi_n} - \eta\big(\eta(1+\eta)+\eta\vartheta\omega_{n+1}+\sigma\big). \tag{2.5}$$

Using (a) and (b), and then (2.5), we get

$$(\forall n \in \mathbb{N}) \quad (\eta_n+\vartheta\lambda_n)(\gamma_{n+1}+\sigma)+\vartheta\lambda_n \leqslant (\eta+\vartheta\lambda_n)\big(\eta(1+\eta)+\eta\vartheta\omega_{n+1}+\sigma\big)+\vartheta\lambda_n \leqslant \frac{\vartheta}{\phi_n}. \tag{2.6}$$

However,

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad (\eta_n + \vartheta\lambda_n)(\gamma_{n+1} + \sigma) + \vartheta\lambda_n &\leqslant \frac{\vartheta}{\phi_n}\\ &\Leftrightarrow (\eta_n + \vartheta\lambda_n)(\gamma_{n+1} + \sigma) - (1/\phi_n - \lambda_n)\vartheta \leqslant 0\\ &\Leftrightarrow (1/\phi_n - \lambda_n)\left(\frac{-\vartheta}{\eta_n + \vartheta\lambda_n}\right) \leqslant -(\gamma_{n+1} + \sigma)\\ &\Leftrightarrow \frac{(1/\phi_n - \lambda_n)(\eta_n\rho_n - 1)}{\lambda_n} + \gamma_{n+1} \leqslant -\sigma.\end{aligned} \tag{2.7}$$

It therefore follows from (2.4) and (2.6) that

$$(\forall n \in \mathbb{N}) \quad \kappa_{n+1} - \kappa_n \leqslant -\sigma\delta_{n+1}. \tag{2.8}$$

Thus, $(\kappa_n)_{n\in\mathbb{N}}$ is decreasing and

$$(\forall n \in \mathbb{N}) \quad \beta_n - \eta\beta_{n-1} = \kappa_n - \gamma_n\delta_n \leqslant \kappa_n \leqslant \kappa_0 \tag{2.9}$$

from which we infer that $(\forall n \in \mathbb{N})\ \beta_n \leqslant \kappa_0 + \eta\beta_{n-1}$. In turn,

$$(\forall n \in \mathbb{N} \smallsetminus \{0\}) \quad \beta_n \leqslant \eta^n\beta_0 + \kappa_0\sum_{j=0}^{n-1}\eta^j \leqslant \eta^n\beta_0 + \frac{\kappa_0}{1-\eta}. \tag{2.10}$$

Altogether, we derive from (2.8), (2.9), and (2.10) that

$$(\forall n \in \mathbb{N}) \quad \sigma\sum_{j=0}^{n}\delta_{j+1} \leqslant \kappa_0 - \kappa_{n+1} \leqslant \kappa_0 + \eta\beta_n \leqslant \frac{\kappa_0}{1-\eta} + \eta^{n+1}\beta_0. \tag{2.11}$$

Hence, $\sum_{j\geqslant 1}\delta_j \leqslant \kappa_0/((1-\eta)\sigma) < +\infty$, and the proof is complete. ∎

Lemma 2.4. *Let $(\eta_n)_{n\in\mathbb{N}}$ be a sequence in $[0,1[$. For every $n \in \mathbb{N}$, set*

$$(\forall k \in \mathbb{N}) \quad \zeta_{k,n} = \begin{cases} 0 & \text{if } k \leqslant n, \\ \displaystyle\sum_{j=n+1}^{k}(\eta_j - 1) & \text{if } k > n, \end{cases} \tag{2.12}$$

and $\chi_n = \sum_{k\geqslant n}\exp(\zeta_{k,n})$. Then the following hold:

(i) *Let $\tau \in [2,+\infty[$ and suppose that $(\forall n \in \mathbb{N})\ \eta_{n+1} = n/(n+1+\tau)$. Then $(\forall n \in \mathbb{N})\ \chi_n \leqslant (n+7)/2$.*

(ii) *Suppose that $(\exists\, \eta \in [0,1[)(\forall n \in \mathbb{N})\ \eta_n \leqslant \eta$. Then $(\forall n \in \mathbb{N})\ \chi_n \leqslant e/(1-\eta)$.*

Proof. (i): We have $(\forall n \in \mathbb{N})(\forall k \in \{n+1, n+2, \ldots\})\ \zeta_{k,n} = -(1+\tau)\sum_{j=n+1}^{k} 1/(j+\tau) \leqslant -3\sum_{j=n+1}^{k} 1/(j+2)$. Since $\xi \mapsto 1/(\xi+2)$ is decreasing on $[1,+\infty[$, it follows that

$$(\forall n \in \mathbb{N})(\forall k \in \{n+1, n+2, \ldots\}) \quad \zeta_{k,n} \leqslant -3\int_{n+1}^{k+1}\frac{d\xi}{\xi+2} = \ln\frac{(n+3)^3}{(k+3)^3}. \tag{2.13}$$

Furthermore, since $\xi \mapsto 1/(\xi+3)^3$ is decreasing on $]-1,+\infty[$, (2.12) yields

$$(\forall n \in \mathbb{N}) \quad \chi_n \leqslant \sum_{k\geqslant n}\frac{(n+3)^3}{(k+3)^3} \leqslant (n+3)^3\int_{n-1}^{+\infty}\frac{d\xi}{(\xi+3)^3} = \frac{(n+3)^3}{2(n+2)^2} \leqslant \frac{n+7}{2}. \tag{2.14}$$

(ii): Note that

$$(\forall n \in \mathbb{N})(\forall k \in \{n+1, n+2, \ldots\}) \quad \zeta_{k,n} = \sum_{j=n+1}^{k}(\eta_j - 1) \leqslant \sum_{j=n+1}^{k}(\eta - 1) = (\eta-1)(k-n). \tag{2.15}$$

Since $\xi \mapsto \exp((\eta-1)\xi)$ is decreasing on $]-1,+\infty[$, it follows that

$$(\forall n \in \mathbb{N}) \quad \chi_n \leqslant \sum_{k\geqslant n}\exp\big((\eta-1)(k-n)\big) \leqslant \int_{n-1}^{+\infty}\exp\big((\eta-1)(\xi-n)\big)d\xi = \frac{\exp(1-\eta)}{1-\eta}, \tag{2.16}$$

which proves the assertion. ∎

The next example provides an instance of an array $(\mu_{n,j})_{n\in\mathbb{N},0\leqslant j\leqslant n}$ satisfying the conditions of Algorithm 1.2 with negative entries. This example will be central to the study of the convergence of some inertial methods.

Example 2.5. Let $(\mu_{n,j})_{n\in\mathbb{N},0\leqslant j\leqslant n}$ be a real array such that $\mu_{0,0}=1$ and

$$(\forall n\in\mathbb{N})\quad 1\leqslant\mu_{n,n}<2\quad\text{and}\quad(\forall j\in\{0,\ldots,n\})\quad \mu_{n,j}=\begin{cases}1-\mu_{n,n} & \text{if } j=n-1,\\ 0 & \text{if } j<n-1.\end{cases}\tag{2.17}$$

For every $n\in\mathbb{N}$, set

$$(\forall k\in\mathbb{N})\quad \zeta_{k,n}=\begin{cases}0 & \text{if } k\leqslant n,\\ \displaystyle\sum_{j=n+1}^{k}(\mu_{j,j}-2) & \text{if } k>n,\end{cases}\tag{2.18}$$

and suppose that $\chi_n=\sum_{k\geqslant n}\exp(\zeta_{k,n})<+\infty$. Then $(\mu_{n,j})_{n\in\mathbb{N},0\leqslant j\leqslant n}$ satisfies conditions (a)–(d) in Algorithm 1.2.

Proof.

(a): $(\forall n\in\mathbb{N})$ $\sum_{j=0}^{n}|\mu_{n,j}|=\mu_{n,n}+|1-\mu_{n,n}|\leqslant 3$.

(b): $(\forall n\in\mathbb{N})$ $\sum_{j=0}^{n}\mu_{n,j}=(1-\mu_{n,n})+\mu_{n,n}=1$.

(c): Let $j\in\mathbb{N}$. Then $(\forall n\in\mathbb{N})$ $n>j+1\Rightarrow\mu_{n,j}=0$. Hence, $\lim_{n\to+\infty}\mu_{n,j}=0$.

(d): We have $(\forall n\in\mathbb{N})$ $\chi_n=\sum_{k\geqslant n}\exp(\zeta_{k,n})\geqslant\exp(\zeta_{n,n})=1$. Now suppose that $(\xi_n)_{n\in\mathbb{N}}$ is a sequence in $[0,+\infty[$ such that there exists a sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ in $[0,+\infty[$ that satisfies

$$\sum_{n\in\mathbb{N}}\chi_n\varepsilon_n<+\infty\quad\text{and}\quad(\forall n\in\mathbb{N})\quad \xi_{n+1}\leqslant\sum_{j=0}^{n}\mu_{n,j}\xi_j+\varepsilon_n.\tag{2.19}$$

Set $\theta_0=0$ and $(\forall n\in\mathbb{N})$ $\theta_{n+1}=[\xi_{n+1}-\xi_n]^+$ and $\nu_n=\mu_{n,n}-2$. It results from (2.19), (2.17), and the inequalities $\xi_1-\xi_0\leqslant(\mu_{0,0}-1)\xi_0+\varepsilon_0$ and

$$\begin{aligned}(\forall n\in\mathbb{N}\smallsetminus\{0\})\quad \xi_{n+1}-\xi_n&\leqslant(\mu_{n,n}-1)\xi_n+(1-\mu_{n,n})\xi_{n-1}+\varepsilon_n\\ &=(\mu_{n,n}-1)(\xi_n-\xi_{n-1})+\varepsilon_n,\end{aligned}\tag{2.20}$$

that $(\forall n\in\mathbb{N})$ $\theta_{n+1}\leqslant(\mu_{n,n}-1)\theta_n+\varepsilon_n=(1+\nu_n)\theta_n+\varepsilon_n$. Consequently, we derive from Lemma 2.1 and (2.18) that $(\forall n\in\mathbb{N})$ $\theta_{n+1}\leqslant\sum_{k=0}^{n}\varepsilon_k\exp(\zeta_{n,k})$. Using [31, Theorem 141], this yields

$$\sum_{n\in\mathbb{N}}\theta_{n+1}\leqslant\sum_{n\in\mathbb{N}}\sum_{k=0}^{n}\varepsilon_k\exp(\zeta_{n,k})=\sum_{k\in\mathbb{N}}\varepsilon_k\sum_{n\geqslant k}\exp(\zeta_{n,k})=\sum_{k\in\mathbb{N}}\varepsilon_k\chi_k.\tag{2.21}$$

Now set $(\forall n\in\mathbb{N})$ $\omega_n=\xi_n-\sum_{k=0}^{n}\theta_k$. Since $\sum_{k\in\mathbb{N}}\chi_k\varepsilon_k<+\infty$, we infer from (2.21) that $\sum_{n\in\mathbb{N}}\theta_n<+\infty$. Thus, since $\inf_{n\in\mathbb{N}}\xi_n\geqslant 0$, $(\omega_n)_{n\in\mathbb{N}}$ is bounded below and

$$(\forall n\in\mathbb{N})\quad \omega_{n+1}=\xi_{n+1}-\theta_{n+1}-\sum_{k=0}^{n}\theta_k\leqslant\xi_{n+1}-\xi_{n+1}+\xi_n-\sum_{k=0}^{n}\theta_k=\omega_n.\tag{2.22}$$

Altogether, $(\omega_n)_{n\in\mathbb{N}}$ converges, and so does therefore $(\xi_n)_{n\in\mathbb{N}}$. ∎

3. Asymptotic behavior of Algorithm 1.2. The main result of the paper is the following theorem, which analyzes the asymptotic behavior of Algorithm 1.2.

THEOREM 3.1. *Consider the setting of Algorithm* 1.2. *For every* $n \in \mathbb{N}$, *define*

$$\vartheta_n = \lambda_n \sum_{i=1}^{m} \|e_{i,n}\| \quad \text{and} \quad (\forall i \in \{1,\ldots,m\}) \quad T_{i+,n} = \begin{cases} T_{i+1,n} \cdots T_{m,n} & \text{if } i \neq m, \\ \mathrm{Id} & \text{if } i = m, \end{cases} \tag{3.1}$$

and set

$$\nu_n \colon S \to [0,+\infty[\, \colon x \mapsto \vartheta_n\big(2\|\overline{x}_n - x\| + \vartheta_n\big). \tag{3.2}$$

Then the following hold:

(i) *Let* $n \in \mathbb{N}$ *and* $x \in S$. *Then* $\|x_{n+1} - x\| \leqslant \sum_{j=0}^{n} |\mu_{n,j}|\, \|x_j - x\| + \vartheta_n$.

(ii) *Let* $n \in \mathbb{N}$ *and* $x \in S$. *Then*

$$\begin{aligned}\|x_{n+1} - x\|^2 \leqslant \sum_{j=0}^{n} \mu_{n,j} \|x_j - x\|^2 - \frac{1}{2} \sum_{j=0}^{n} \sum_{k=0}^{n} \mu_{n,j}\mu_{n,k} \|x_j - x_k\|^2 \\ - \lambda_n(1/\phi_n - \lambda_n)\|T_n\overline{x}_n - \overline{x}_n\|^2 + \nu_n(x).\end{aligned}$$

(iii) *Let* $n \in \mathbb{N}$ *and* $x \in S$. *Then*

$$\begin{aligned}&\|x_{n+1} - x\|^2 \\ &\quad \leqslant \sum_{j=0}^{n} \mu_{n,j} \|x_j - x\|^2 - \frac{1}{2} \sum_{j=0}^{n} \sum_{k=0}^{n} \mu_{n,j}\mu_{n,k} \|x_j - x_k\|^2 \\ &\quad + \lambda_n(\lambda_n - 1)\|T_n\overline{x}_n - \overline{x}_n\|^2 \\ &\quad - \lambda_n \max_{1 \leqslant i \leqslant m} \left(\frac{1 - \alpha_{i,n}}{\alpha_{i,n}} \|(\mathrm{Id} - T_{i,n})T_{i+,n}\overline{x}_n - (\mathrm{Id} - T_{i,n})T_{i+,n}x\|^2 \right) + \nu_n(x).\end{aligned}$$

Now assume that, in addition,

$$\sum_{n \in \mathbb{N}} \chi_n \sum_{j=0}^{n} \sum_{k=0}^{n} [\mu_{n,j}\mu_{n,k}]^- \|x_j - x_k\|^2 < +\infty \quad \text{and} \quad (\forall x \in S) \quad \sum_{n \in \mathbb{N}} \chi_n \nu_n(x) < +\infty. \tag{3.3}$$

Then the following hold:

(iv) *Let* $x \in S$. *Then* $(\|x_n - x\|)_{n \in \mathbb{N}}$ *converges.*

(v) $\lambda_n(1/\phi_n - \lambda_n)\|T_n\overline{x}_n - \overline{x}_n\|^2 \to 0$.

(vi) $\sum_{j=0}^{n} \sum_{k=0}^{n} [\mu_{n,j}\mu_{n,k}]^+ \|x_j - x_k\|^2 \to 0$.

(vii) *Suppose that*

$$(\exists\, \varepsilon \in \,]0,1[)(\forall n \in \mathbb{N}) \quad \lambda_n \leqslant (1 - \varepsilon)/\phi_n. \tag{3.4}$$

Then $x_{n+1} - \overline{x}_n \to 0$. *In addition, if every weak sequential cluster point of* $(\overline{x}_n)_{n \in \mathbb{N}}$ *is in* S, *then there exists* $x \in S$ *such that* $x_n \rightharpoonup x$.

(viii) *Suppose that* $(\overline{x}_n)_{n \in \mathbb{N}}$ *has a strong cluster point* x *in* S *and that* (3.4) *holds. Then* $x_n \to x$.

(ix) *Let* $x \in S$ *and suppose that* $(\exists\, \varepsilon \in \,]0,1[)(\forall n \in \mathbb{N})$ $\lambda_n \leqslant \varepsilon + (1 - \varepsilon)/\phi_n$. *Then*

$$\lambda_n \max_{1 \leqslant i \leqslant m} \frac{1 - \alpha_{i,n}}{\alpha_{i,n}} \|(\mathrm{Id} - T_{i,n})T_{i+,n}\overline{x}_n - (\mathrm{Id} - T_{i,n})T_{i+,n}x\|^2 \to 0.$$

Proof. Let $n \in \mathbb{N}$ and set

$$e_n = T_{1,n}\Big(T_{2,n}\big(\cdots T_{m-1,n}(T_{m,n}\overline{x}_n + e_{m,n}) + e_{m-1,n}\cdots\big) + e_{2,n}\Big) + e_{1,n} - T_n\overline{x}_n. \tag{3.5}$$

If $m > 1$, using the nonexpansiveness of the operators $(T_{i,n})_{1\leqslant i\leqslant m-1}$, we obtain

$$\begin{aligned}
\|e_n\| &\leqslant \|e_{1,n}\| + \bigg\|T_{1,n}\bigg(T_{2,n}\big(\cdots T_{m-1,n}(T_{m,n}\overline{x}_n + e_{m,n}) + e_{m-1,n}\cdots\big)\\
&\qquad\qquad + e_{2,n}\bigg) - T_{1,n}\cdots T_{m,n}\overline{x}_n\bigg\|\\
&\leqslant \|e_{1,n}\| + \bigg\|T_{2,n}\bigg(T_{3,n}\big(\cdots T_{m-1,n}(T_{m,n}\overline{x}_n + e_{m,n}) + e_{m-1,n}\cdots\big) + e_{3,n}\bigg)\\
&\qquad\qquad + e_{2,n} - T_{2,n}\cdots T_{m,n}\overline{x}_n\bigg\|\\
&\leqslant \|e_{1,n}\| + \|e_{2,n}\| + \bigg\|T_{3,n}\bigg(T_{4,n}\big(\cdots T_{m-1,n}(T_{m,n}\overline{x}_n + e_{m,n}) + e_{m-1,n}\cdots\big)\\
&\qquad\qquad + e_{4,n}\bigg) + e_{3,n} - T_{3,n}\cdots T_{m,n}\overline{x}_n\bigg\|\\
&\;\;\vdots\\
&\leqslant \sum_{i=1}^{m}\|e_{i,n}\|.
\end{aligned}\tag{3.6}$$

Thus, we infer from (3.1) that

$$\lambda_n\|e_n\| \leqslant \vartheta_n. \tag{3.7}$$

On the other hand, we derive from (1.12) and (3.5) that

$$x_{n+1} = \overline{x}_n + \lambda_n\big(T_n\overline{x}_n + e_n - \overline{x}_n\big). \tag{3.8}$$

Now set

$$R_n = \frac{1}{\phi_n}T_n - \frac{1-\phi_n}{\phi_n}\,\mathrm{Id} \quad\text{and}\quad \eta_n = \lambda_n\phi_n. \tag{3.9}$$

Then $\eta_n \in \,]0,1]$, $\operatorname{Fix} R_n = \operatorname{Fix} T_n$, and R_n is quasi-nonexpansive since T_n is averaged quasi-nonexpansive with constant ϕ_n by Remark 1.3(i). Furthermore, (3.8) can be written as

$$x_{n+1} = \overline{x}_n + \eta_n\big(R_n\overline{x}_n - \overline{x}_n\big) + \lambda_n e_n. \tag{3.10}$$

Next, we define

$$z_n = \overline{x}_n + \lambda_n(T_n\overline{x}_n - \overline{x}_n) = \overline{x}_n + \eta_n(R_n\overline{x}_n - \overline{x}_n). \tag{3.11}$$

Let $x \in S$. Since $x \in \operatorname{Fix} R_n$ and R_n is quasi-nonexpansive, we have

$$\begin{aligned}
\|z_n - x\| &= \|(1-\eta_n)(\overline{x}_n - x) + \eta_n(R_n\overline{x}_n - x)\|\\
&\leqslant (1-\eta_n)\|\overline{x}_n - x\| + \eta_n\|R_n\overline{x}_n - x\|\\
&\leqslant \|\overline{x}_n - x\|.
\end{aligned}\tag{3.12}$$

Hence, (3.10) and (3.7) yield

$$\|x_{n+1} - x\| \leqslant \|z_n - x\| + \lambda_n\|e_n\| \leqslant \|z_n - x\| + \vartheta_n. \tag{3.13}$$

In turn, it follows from (3.12) and (3.2) that

$$\|x_{n+1}-x\|^2 \leqslant \|z_n-x\|^2+2\vartheta_n\|z_n-x\|+\vartheta_n^2 \leqslant \|z_n-x\|^2+\nu_n(x). \tag{3.14}$$

In addition, [10, Lemma 2.14(ii)] yields

$$\|\overline{x}_n-x\|^2=\left\|\sum_{j=0}^n \mu_{n,j}(x_j-x)\right\|^2=\sum_{j=0}^n \mu_{n,j}\|x_j-x\|^2-\frac{1}{2}\sum_{j=0}^n\sum_{k=0}^n \mu_{n,j}\mu_{n,k}\|x_j-x_k\|^2. \tag{3.15}$$

(i): By (3.13) and (3.12),

$$(\forall n\in\mathbb{N})(\forall x\in S)\quad \|x_{n+1}-x\| \leqslant \|\overline{x}_n-x\|+\vartheta_n \leqslant \sum_{j=0}^n |\mu_{n,j}|\,\|x_j-x\|+\vartheta_n. \tag{3.16}$$

(ii): Let $n\in\mathbb{N}$ and $x\in S$. Since

$$\begin{aligned}\|z_n-x\|^2&=\|(1-\eta_n)(\overline{x}_n-x)+\eta_n(R_n\overline{x}_n-x)\|^2\\&=(1-\eta_n)\|\overline{x}_n-x\|^2+\eta_n\|R_n\overline{x}_n-x\|^2-\eta_n(1-\eta_n)\|R_n\overline{x}_n-\overline{x}_n\|^2\\&\leqslant \|\overline{x}_n-x\|^2-\eta_n(1-\eta_n)\|R_n\overline{x}_n-\overline{x}_n\|^2, \qquad (3.17)\end{aligned}$$

we deduce from (3.14) and (3.9) that

$$\begin{aligned}\|x_{n+1}-x\|^2&\leqslant \|z_n-x\|^2+\nu_n(x)\\&\leqslant \|\overline{x}_n-x\|^2-\eta_n(1-\eta_n)\|R_n\overline{x}_n-\overline{x}_n\|^2+\nu_n(x)\\&=\|\overline{x}_n-x\|^2-\lambda_n(1/\phi_n-\lambda_n)\|T_n\overline{x}_n-\overline{x}_n\|^2+\nu_n(x). \qquad (3.18)\end{aligned}$$

In view of (3.15), we obtain the announced inequality.

(iii): Let $n\in\mathbb{N}$ and $x\in S$. We derive from [10, Proposition 4.35] that

$$\begin{aligned}&(\forall i\in\{1,\ldots,m-1\})\ \big(\forall(u,v)\in\mathcal{H}^2\big)\\&\|T_{i,n}u-T_{i,n}v\|^2\leqslant\|u-v\|^2-\frac{1-\alpha_{i,n}}{\alpha_{i,n}}\|(\mathrm{Id}-T_{i,n})u-(\mathrm{Id}-T_{i,n})v\|^2. \qquad (3.19)\end{aligned}$$

If $m>1$, using this inequality successively for $i=1,\ldots,m-1$ leads to

$$\begin{aligned}\|T_n\overline{x}_n-x\|^2&=\|T_{1,n}\cdots T_{m,n}\overline{x}_n-T_{1,n}\cdots T_{m,n}x\|^2\\&\leqslant\|T_{m,n}\overline{x}_n-T_{m,n}x\|^2\\&\quad-\sum_{i=1}^{m-1}\frac{1-\alpha_{i,n}}{\alpha_{i,n}}\,\|(\mathrm{Id}-T_{i,n})T_{i+,n}\overline{x}_n-(\mathrm{Id}-T_{i,n})T_{i+,n}x\|^2\\&\leqslant\|T_{m,n}\overline{x}_n-T_{m,n}x\|^2\\&\quad-\max_{1\leqslant i\leqslant m-1}\frac{1-\alpha_{i,n}}{\alpha_{i,n}}\,\|(\mathrm{Id}-T_{i,n})T_{i+,n}\overline{x}_n-(\mathrm{Id}-T_{i,n})T_{i+,n}x\|^2. \qquad (3.20)\end{aligned}$$

Note that, in cases (a) and (c) of Problem 1.1,

$$\|T_{m,n}\overline{x}_n-T_{m,n}x\|^2\leqslant\|\overline{x}_n-x\|^2-\frac{1-\alpha_{m,n}}{\alpha_{m,n}}\|(\mathrm{Id}-T_{m,n})\overline{x}_n-(\mathrm{Id}-T_{m,n})x\|^2. \tag{3.21}$$

This inequality remains valid in case (b) of Problem 1.1 since [10, Proposition 4.49(i)] implies that

$$\operatorname{Fix}(T_{1,n}\cdots T_{m,n}) = \bigcap_{i=1}^{m} \operatorname{Fix} T_{i,n} \tag{3.22}$$

and, therefore, that $x \in \operatorname{Fix} T_{m,n}$. Altogether, we deduce from (3.20) and (3.21) that

$$\|T_n\overline{x}_n - x\|^2 \leqslant \|\overline{x}_n - x\|^2 - \max_{1\leqslant i\leqslant m} \frac{1-\alpha_{i,n}}{\alpha_{i,n}} \|(\mathrm{Id} - T_{i,n})T_{i+,n}\overline{x}_n - (\mathrm{Id} - T_{i,n})T_{i+,n}x\|^2 . \tag{3.23}$$

Hence, it follows from (3.11) that

$$\begin{aligned}\|z_n - x\|^2 &= \|(1-\lambda_n)(\overline{x}_n - x) + \lambda_n(T_n\overline{x}_n - x)\|^2\\ &= (1-\lambda_n)\|\overline{x}_n - x\|^2 + \lambda_n\|T_n\overline{x}_n - x\|^2 + \lambda_n(\lambda_n - 1)\|T_n\overline{x}_n - \overline{x}_n\|^2\\ &\leqslant \|\overline{x}_n - x\|^2 - \lambda_n \max_{1\leqslant i\leqslant m} \frac{1-\alpha_{i,n}}{\alpha_{i,n}} \|(\mathrm{Id} - T_{i,n})T_{i+,n}\overline{x}_n - (\mathrm{Id} - T_{i,n})T_{i+,n}x\|^2\\ &\quad + \lambda_n(\lambda_n - 1)\|T_n\overline{x}_n - \overline{x}_n\|^2.\end{aligned} \tag{3.24}$$

In view of (3.14) and (3.15), the inequality is established.

(iv): Let $x \in S$ and set

$$(\forall n \in \mathbb{N}) \quad \begin{cases}\xi_n = \|x_n - x\|^2,\\ \varepsilon_n = \nu_n(x) + \dfrac{1}{2}\displaystyle\sum_{j=0}^{n}\sum_{k=0}^{n}[\mu_{n,j}\mu_{n,k}]^-\|x_j - x_k\|^2.\end{cases} \tag{3.25}$$

Since $\inf_{n\in\mathbb{N}} \lambda_n(1/\phi_n - \lambda_n) \geqslant 0$, (3.3) and (ii) imply that

$$\sum_{n\in\mathbb{N}} \chi_n\varepsilon_n < +\infty \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad \xi_{n+1} \leqslant \sum_{j=0}^{n} \mu_{n,j}\xi_j + \varepsilon_n. \tag{3.26}$$

In turn, it follows from (3.3) and condition (d) in Algorithm 1.2 that $(\|x_n - x\|)_{n\in\mathbb{N}}$ converges.

(v)–(vi): Let $x \in S$. Then it follows from (iv) that $\rho = \lim_{n\to+\infty} \|x_n - x\|$ is well defined. Hence, Lemma 2.2 implies that $\sum_{j=0}^{n} \mu_{n,j}\|x_j - x\|^2 \to \rho^2$ and therefore that

$$\sum_{j=0}^{n} \mu_{n,j}\|x_j - x\|^2 - \|x_{n+1} - x\|^2 \to 0. \tag{3.27}$$

Since $\inf_{n\in\mathbb{N}} \chi_n > 0$, (3.3) yields

$$\nu_n(x) \to 0 \quad \text{and} \quad \sum_{j=0}^{n}\sum_{k=0}^{n}[\mu_{n,j}\mu_{n,k}]^-\|x_j - x_k\|^2 \to 0. \tag{3.28}$$

It follows from (ii), (3.27), and (3.28) that

$$\begin{aligned}0 &\leqslant \lambda_n(1/\phi_n - \lambda_n)\|T_n\overline{x}_n - \overline{x}_n\|^2 + \frac{1}{2}\sum_{j=0}^{n}\sum_{k=0}^{n}[\mu_{n,j}\mu_{n,k}]^+\|x_j - x_k\|^2\\ &\leqslant \sum_{j=0}^{n} \mu_{n,j}\|x_j - x\|^2 - \|x_{n+1} - x\|^2 + \frac{1}{2}\sum_{j=0}^{n}\sum_{k=0}^{n}[\mu_{n,j}\mu_{n,k}]^-\|x_j - x_k\|^2 + \nu_n(x)\\ &\to 0,\end{aligned} \tag{3.29}$$

which gives the desired conclusions.

(vii): Set $\zeta = 1/\varepsilon - 1$. We deduce from (3.3) and (3.2) that $\sum_{n\in\mathbb{N}} \vartheta_n^2 < +\infty$. Hence, it follows from (3.8), (3.7), (3.4), and (v) that

$$\begin{aligned}\|x_{n+1} - \overline{x}_n\|^2 &\leqslant 2\Big(\lambda_n^2\big\|T_n\overline{x}_n - \overline{x}_n\big\|^2 + \lambda_n^2\|e_n\|^2\Big)\\ &\leqslant 2\bigg(\frac{\lambda_n}{1/\phi_n - \lambda_n}\lambda_n(1/\phi_n - \lambda_n)\big\|T_n\overline{x}_n - \overline{x}_n\big\|^2 + \vartheta_n^2\bigg)\\ &\leqslant 2\Big(\zeta\lambda_n(1/\phi_n - \lambda_n)\big\|T_n\overline{x}_n - \overline{x}_n\big\|^2 + \vartheta_n^2\Big)\\ &\to 0. \end{aligned} \tag{3.30}$$

Therefore $x_{n+1} - \overline{x}_n \rightharpoonup 0$ and hence the weak sequential cluster points of $(x_n)_{n\in\mathbb{N}}$ lie in S. In view of (iv) and [10, Lemma 2.47], the claim is proved.

(viii): Since $x_{n+1} - \overline{x}_n \to 0$ by (3.30), $(x_n)_{n\in\mathbb{N}}$ has a strong cluster point $x \in S$. In view of (iv), $x_n \to x$.

(ix): Set $\zeta = 1/\varepsilon - 1$. Then, for every $n \in \mathbb{N}$, $\lambda_n \leqslant 1/(1+\zeta) + \zeta/(\phi_n(1+\zeta))$ and hence $(1+\zeta)\lambda_n - 1 \leqslant \zeta/\phi_n$, i.e., $\lambda_n - 1 \leqslant \zeta(1/\phi_n - \lambda_n)$. We therefore derive from (iii), (3.27), (3.3), (v), and (3.28) that

$$\begin{aligned}0 &\leqslant \lambda_n \max_{1\leqslant i\leqslant m} \frac{1-\alpha_{i,n}}{\alpha_{i,n}} \|(\mathrm{Id} - T_{i,n})T_{i+,n}\overline{x}_n - (\mathrm{Id} - T_{i,n})T_{i+,n}x\|^2\\ &\leqslant \sum_{j=0}^n \mu_{n,j}\|x_j - x\|^2 - \|x_{n+1} - x\|^2 + \frac{1}{2}\sum_{j=0}^n\sum_{k=0}^n [\mu_{n,j}\mu_{n,k}]^-\|x_j - x_k\|^2\\ &\quad + \lambda_n(\lambda_n - 1)\|T_n\overline{x}_n - \overline{x}_n\|^2 + \nu_n(x)\\ &\leqslant \sum_{j=0}^n \mu_{n,j}\|x_j - x\|^2 - \|x_{n+1} - x\|^2 + \frac{1}{2}\sum_{j=0}^n\sum_{k=0}^n [\mu_{n,j}\mu_{n,k}]^-\|x_j - x_k\|^2\\ &\quad + \zeta\lambda_n(1/\phi_n - \lambda_n)\|T_n\overline{x}_n - \overline{x}_n\|^2 + \nu_n(x)\\ &\to 0, \end{aligned} \tag{3.31}$$

which shows the assertion. ∎

Next, we present two corollaries that are instrumental in the analysis of two important special cases of our framework: mean value and inertial multilayer algorithms.

COROLLARY 3.2. *Consider the setting of Algorithm* 1.2 *and define* $(\vartheta_n)_{n\in\mathbb{N}}$ *as in* (3.1). *Assume that*

$$\inf_{n\in\mathbb{N}} \min_{0\leqslant j\leqslant n} \mu_{n,j} \geqslant 0 \quad \text{and} \quad \sum_{n\in\mathbb{N}} \chi_n\vartheta_n < +\infty. \tag{3.32}$$

Then the following hold:

(i) $\sum_{j=0}^n \sum_{k=0}^n \mu_{n,j}\mu_{n,k}\|x_j - x_k\|^2 \to 0$.

(ii) *Let* $x \in S$ *and suppose that* $(\exists\, \varepsilon \in\,]0,1[)(\forall n \in \mathbb{N})\ \lambda_n \leqslant \varepsilon + (1-\varepsilon)/\phi_n$. *Then*

$$\lambda_n \max_{1\leqslant i\leqslant m} \frac{1-\alpha_{i,n}}{\alpha_{i,n}} \|(\mathrm{Id} - T_{i,n})T_{i+,n}\overline{x}_n - (\mathrm{Id} - T_{i,n})T_{i+,n}x\|^2 \to 0.$$

(iii) *Suppose that every weak sequential cluster point of* $(\overline{x}_n)_{n\in\mathbb{N}}$ *is in* S *and that* $(\exists\, \varepsilon \in\,]0,1[)(\forall n \in \mathbb{N})\ \lambda_n \leqslant (1-\varepsilon)/\phi_n$. *Then* $x_{n+1} - \overline{x}_n \to 0$ *and there exists* $x \in S$ *such that* $x_n \rightharpoonup x$.

(iv) *Suppose that every weak sequential cluster point of* $(\overline{x}_n)_{n\in\mathbb{N}}$ *is in* S, *that* $\sup_{n\in\mathbb{N}} \phi_n < 1$, *and that* $(\exists\, \varepsilon \in\,]0,1[)(\forall n \in \mathbb{N})\ \lambda_n \leqslant \varepsilon + (1-\varepsilon)/\phi_n$. *Then* $x_{n+1} - \overline{x}_n \to 0$ *and there exists* $x \in S$ *such that* $x_n \rightharpoonup x$.

(v) *Suppose that every weak sequential cluster point of* $(\overline{x}_n)_{n\in\mathbb{N}}$ *is in* S *and that* $\inf_{n\in\mathbb{N}}\mu_{n,n}>0$. *Then* $x_n-\overline{x}_n\to 0$ *and there exists* $x\in S$ *such that* $x_n \rightharpoonup x$.

Proof. We derive from Theorem 3.1(i) that $(\forall n\in\mathbb{N})\ \|x_{n+1}-x\|\leqslant\sum_{j=0}^n\mu_{n,j}\|x_j-x\|+\vartheta_n$. In turn, it follows from condition (d) in Algorithm 1.2 that $(\|x_n-x\|)_{n\in\mathbb{N}}$ converges. As a result, $(x_n)_{n\in\mathbb{N}}$ is bounded and (3.32) therefore implies (3.3).

(i)–(iii): These follow, respectively, from items (vi), (ix), and (vii) in Theorem 3.1.

(iv): Set $\delta=\varepsilon(1-\sup_{n\in\mathbb{N}}\phi_n)$. Then $\delta\in\,]0,\varepsilon[$ and $(\forall n\in\mathbb{N})\ (\varepsilon-\delta)/\phi_n\geqslant\varepsilon$. Hence,

$$(\forall n\in\mathbb{N})\quad \lambda_n\leqslant\varepsilon+\frac{1-\varepsilon}{\phi_n}=\varepsilon+\frac{1-\delta}{\phi_n}-\frac{\varepsilon-\delta}{\phi_n}\leqslant\frac{1-\delta}{\phi_n}. \tag{3.33}$$

The claim therefore follows from (iii).

(v): Set

$$\theta=\frac{1}{\inf\limits_{n\in\mathbb{N}}\mu_{n,n}}\quad\text{and}\quad(\forall n\in\mathbb{N})(\forall j\in\{0,\dots,n\})\quad\gamma_{n,j}=\begin{cases}\dfrac{\mu_{n,n}+1}{2} & \text{if } j=n,\\[2mm] \dfrac{\mu_{n,j}}{2} & \text{if } j<n.\end{cases} \tag{3.34}$$

Then, using Apollonius' identity, [10, Lemma 2.12(iv)], and (3.34), we obtain

$$\begin{aligned}
&\frac14\|\overline{x}_n-x_n\|^2\\
&\quad=\frac12\Big(\|\overline{x}_n-x\|^2+\|x_n-x\|^2\Big)-\left\|\frac{\overline{x}_n+x_n}{2}-x\right\|^2\\
&\quad=\frac12\Big(\|\overline{x}_n-x\|^2+\|x_n-x\|^2\Big)-\left\|\sum_{j=0}^n\gamma_{n,j}(x_j-x)\right\|^2\\
&\quad=\frac12\Big(\|\overline{x}_n-x\|^2+\|x_n-x\|^2\Big)-\sum_{j=0}^n\gamma_{n,j}\|x_j-x\|^2+\sum_{0\leqslant j<k\leqslant n}\gamma_{n,j}\gamma_{n,k}\|x_j-x_k\|^2\\
&\quad\leqslant\frac12\left(\sum_{j=0}^n\mu_{n,j}\|x_j-x\|^2+\|x_n-x\|^2\right)-\sum_{j=0}^n\gamma_{n,j}\|x_j-x\|^2\\
&\qquad+\frac14\left(\sum_{0\leqslant j<k<n}\mu_{n,j}\mu_{n,k}\|x_j-x_k\|^2+\sum_{j=0}^{n-1}\mu_{n,j}(\mu_{n,n}+1)\|x_j-x_n\|^2\right)\\
&\quad\leqslant\frac12\left(\sum_{j=0}^n\mu_{n,j}\|x_j-x\|^2+\|x_n-x\|^2\right)-\sum_{j=0}^n\gamma_{n,j}\|x_j-x\|^2\\
&\qquad+\frac14\left(\sum_{0\leqslant j<k\leqslant n}\mu_{n,j}\mu_{n,k}\|x_j-x_k\|^2+\theta\sum_{j=0}^{n-1}\mu_{n,j}\mu_{n,n}\|x_j-x_n\|^2\right)\\
&\quad\leqslant\frac12\left(\sum_{j=0}^n\mu_{n,j}\|x_j-x\|^2+\|x_n-x\|^2\right)-\sum_{j=0}^n\gamma_{n,j}\|x_j-x\|^2\\
&\qquad+\frac{1+\theta}{4}\sum_{0\leqslant j<k\leqslant n}\mu_{n,j}\mu_{n,k}\|x_j-x_k\|^2.
\end{aligned} \tag{3.35}$$

Next, let us set $\rho = \lim \|x_n - x\|^2$. Then it follows from Lemma 2.2 that $\sum_{j=0}^n \mu_{n,j}\|x_j - x\|^2 \to \rho$ and $\sum_{j=0}^n \gamma_{n,j}\|x_j - x\|^2 \to \rho$. On the other hand, (i) asserts that $\sum_{0\leqslant j<k\leqslant n} \mu_{n,j}\mu_{n,k}\|x_j - x_k\|^2 \to 0$. Altogether, (3.35) yields $\|\overline{x}_n - x_n\| \to 0$. Thus, the weak sequential cluster points of $(x_n)_{n\in\mathbb{N}}$ belong to S, and the conclusion follows from the fact that $(\|x_n - x\|)_{n\in\mathbb{N}}$ converges and from [10, Lemma 2.47]. ∎

COROLLARY 3.3. *Consider the setting of Algorithm* 1.2 *with*

$$(\forall i \in \{1,\dots,m\})(\forall n \in \mathbb{N}) \quad e_{i,n} = 0. \tag{3.36}$$

Set $x_{-1} = x_0$ *and suppose that there exists a sequence* $(\eta_n)_{n\in\mathbb{N}}$ *in* $[0,1[$ *such that* $\eta_0 = 0$ *and*

$$(\forall n \in \mathbb{N})(\forall j \in \{0,\dots,n\}) \quad \mu_{n,j} = \begin{cases} 1+\eta_n & \text{if } j = n,\\ -\eta_n & \text{if } j = n-1,\\ 0 & \text{if } j < n-1. \end{cases} \tag{3.37}$$

For every $n \in \mathbb{N}$, *set*

$$(\forall k \in \mathbb{N}) \quad \zeta_{k,n} = \begin{cases} 0 & \text{if } k \leqslant n,\\ \displaystyle\sum_{j=n+1}^{k} (\eta_j - 1) & \text{if } k > n, \end{cases} \tag{3.38}$$

and assume that $\chi_n = \sum_{k\geqslant n} \exp(\zeta_{k,n})$. *Suppose that one of the following is satisfied:*

(a) $\sum_{n\in\mathbb{N}} \chi_n\eta_n\|x_n - x_{n-1}\|^2 < +\infty$.

(b) $\sum_{n\in\mathbb{N}} n\|x_n - x_{n-1}\|^2 < +\infty$ *and there exists* $\tau \in [2,+\infty[$ *such that* $(\forall n \in \mathbb{N}\smallsetminus\{0\})$ $\eta_n = (n-1)/(n+\tau)$.

(c) $\sum_{n\in\mathbb{N}} \eta_n\|x_n - x_{n-1}\|^2 < +\infty$ *and there exists* $\eta \in [0,1[$ *such that* $(\forall n \in \mathbb{N})$ $\eta_n \leqslant \eta$.

(d) *Set* $(\forall n \in \mathbb{N})$ $\omega_n = 1/\phi_n - \lambda_n$. *There exist* $(\sigma,\vartheta) \in\,]0,+\infty[^2$ *and* $\eta \in\,]0,1[$ *such that*

$$(\forall n \in \mathbb{N}) \quad \begin{cases} \eta_n \leqslant \eta_{n+1} \leqslant \eta,\\ \lambda_n \leqslant \dfrac{\vartheta/\phi_n - \eta\big(\eta(1+\eta) + \eta\vartheta\omega_{n+1} + \sigma\big)}{\vartheta\big(1+\eta(1+\eta) + \eta\vartheta\omega_{n+1} + \sigma\big)},\\ \dfrac{\eta^2(1+\eta) + \eta\sigma}{\vartheta} < \dfrac{1}{\phi_n} - \eta^2\omega_{n+1}. \end{cases} \tag{3.39}$$

Then the following hold:

(i) $\lambda_n(1/\phi_n - \lambda_n)\|T_n\overline{x}_n - \overline{x}_n\|^2 \to 0$.

(ii) *Let* $x \in S$ *and suppose that* $(\exists\,\varepsilon \in\,]0,1[)(\forall n \in \mathbb{N})$ $\lambda_n \leqslant \varepsilon + (1-\varepsilon)/\phi_n$. *Then*

$$\lambda_n \max_{1\leqslant i\leqslant m} \frac{1-\alpha_{i,n}}{\alpha_{i,n}} \|(\mathrm{Id} - T_{i,n})T_{i+,n}\overline{x}_n - (\mathrm{Id} - T_{i,n})T_{i+,n}x\|^2 \to 0.$$

(iii) $\overline{x}_n - x_n \to 0$.

(iv) *Suppose that every weak sequential cluster point of* $(\overline{x}_n)_{n\in\mathbb{N}}$ *is in* S. *Then there exists* $x \in S$ *such that* $x_n \rightharpoonup x$.

Proof. In view of Example 2.5, $(\mu_{n,j})_{n\in\mathbb{N},0\leqslant j\leqslant n}$ satisfies conditions (a)–(d) in Algorithm 1.2.

(a): Set $\chi = \inf_{n\in\mathbb{N}} \chi_n$ and define $(\nu_n)_{n\in\mathbb{N}}$ as in (3.2). We have $\sup_{n\in\mathbb{N}}(1+\eta_n) \leqslant 2$ and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} \chi_n \sum_{j=0}^n \sum_{k=0}^n [\mu_{n,j}\mu_{n,k}]^- \|x_j - x_k\|^2 = (1+\eta_n)\chi_n\eta_n\|x_n - x_{n-1}\|^2, \\ (\forall x \in S) \quad \chi_n \nu_n(x) = 0. \end{cases} \tag{3.40}$$

Hence (3.3) holds, and (i) and (ii) follow from items (v) and (ix) in Theorem 3.1, respectively. Furthermore, (3.37) implies that

$$\|\overline{x}_n - x_n\|^2 \leqslant \eta_n^2\|x_n - x_{n-1}\|^2 \leqslant \eta_n\|x_n - x_{n-1}\|^2 \leqslant \frac{\chi_n\eta_n}{\chi}\|x_n - x_{n-1}\|^2 \to 0. \tag{3.41}$$

Thus, (iii) holds. In turn, the weak sequential cluster points of $(x_n)_{n\in\mathbb{N}}$ belong to S and (iv) therefore follows from Theorem 3.1(iv) and [10, Lemma 2.47].

(b)⇒(a): It follows from Lemma 2.4(i) that

$$\sum_{n\in\mathbb{N}} \chi_n\eta_n\|x_n - x_{n-1}\|^2 \leqslant \sum_{n\in\mathbb{N}} \frac{n+7}{2}\|x_n - x_{n-1}\|^2 < +\infty. \tag{3.42}$$

(c)⇒(a): Lemma 2.4(ii) asserts that $\sup_{n\in\mathbb{N}} \chi_n \leqslant e/(1-\eta)$.

(d)⇒(c): Let $x \in S$. It follows from Theorem 3.1(ii) that

$$\begin{aligned}(\forall n \in \mathbb{N}) \ \|x_{n+1} - x\|^2 \leqslant (1+\eta_n)\|x_n - x\|^2 - \eta_n\|x_{n-1} - x\|^2 + \eta_n(1+\eta_n)\|x_n - x_{n-1}\|^2 \\ - \lambda_n\left(\frac{1}{\phi_n} - \lambda_n\right)\|T_n\overline{x}_n - \overline{x}_n\|^2.\end{aligned} \tag{3.43}$$

Now set $\beta_{-1} = \|x_0 - x\|^2$ and

$$(\forall n \in \mathbb{N}) \quad \beta_n = \|x_n - x\|^2, \quad \delta_n = \|x_n - x_{n-1}\|^2, \quad \text{and} \quad \rho_n = \frac{1}{\eta_n + \lambda_n\vartheta}. \tag{3.44}$$

Then

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \|T_n\overline{x}_n - \overline{x}_n\|^2 &= \frac{1}{\lambda_n^2}\big\|(x_{n+1} - x_n) + \eta_n(x_{n-1} - x_n)\big\|^2 \\ &= \frac{1}{\lambda_n^2}\left(\delta_{n+1} + \eta_n^2\delta_n + \eta_n\left(2\left\langle \sqrt{\rho_n}(x_{n+1} - x_n) \;\middle|\; \frac{x_{n-1} - x_n}{\sqrt{\rho_n}}\right\rangle\right)\right) \\ &\geqslant \frac{1}{\lambda_n^2}\left(\delta_{n+1} + \eta_n^2\delta_n - \eta_n\left(\rho_n\delta_{n+1} + \frac{\delta_n}{\rho_n}\right)\right).\end{aligned} \tag{3.45}$$

Thus, we derive from (3.43) that

$$(\forall n \in \mathbb{N}) \quad \beta_{n+1} - \beta_n - \eta_n(\beta_n - \beta_{n-1}) \leqslant \frac{(1/\phi_n - \lambda_n)(\eta_n\rho_n - 1)}{\lambda_n}\delta_{n+1} + \gamma_n\delta_n, \tag{3.46}$$

where

$$(\forall n \in \mathbb{N}) \quad \gamma_n = \eta_n(1+\eta_n) + \eta_n\left(\frac{1}{\phi_n} - \lambda_n\right)\frac{1 - \rho_n\eta_n}{\rho_n\lambda_n} \geqslant 0. \tag{3.47}$$

However, it follows from (3.44) that $(\forall n \in \mathbb{N})$ $\vartheta = (1 - \rho_n\eta_n)/(\rho_n\lambda_n)$. Hence, (3.47) yields

$$(\forall n \in \mathbb{N}) \quad \gamma_n = \eta_n(1+\eta_n) + \eta_n\left(\frac{1}{\phi_n} - \lambda_n\right)\vartheta \leqslant \eta(1+\eta) + \eta\vartheta\omega_n. \tag{3.48}$$

Thus, by Lemma 2.3, $\sum_{n\in\mathbb{N}} \eta_n\delta_n \leqslant \sum_{n\in\mathbb{N}} \delta_n < +\infty$ and we conclude that (c) is satisfied. ∎

Remark 3.4. In Corollary 3.3, no errors were allowed in the implementation of the operators. It is however possible to allow errors in multilayer inertial methods in certain scenarios. For instance, suppose that in Corollary 3.3 we make the additional assumptions that $\lambda_n \equiv 1$ and that $\bigcup_{n\in\mathbb{N}} \operatorname{ran} T_{1,n}$ is bounded. At the same time, let us introduce errors of such that $(\forall i \in \{1,\ldots,m\})\ \sum_{n\in\mathbb{N}} \chi_n \|e_{i,n}\| < +\infty$. Note that (1.12) becomes

(3.49)

$$\begin{array}{l}\text{for } n = 0,1,\ldots\\ \left\lfloor \begin{array}{l} \overline{x}_n = (1+\eta_n)x_n - \eta_n x_{n-1},\\ x_{n+1} = T_{1,n}\Big(T_{2,n}\big(\cdots T_{m-1,n}(T_{m,n}\overline{x}_n + e_{m,n}) + e_{m-1,n}\cdots\big) + e_{2,n}\Big) + e_{1,n}.\end{array}\right.\end{array}$$

Hence, the assumptions imply that $(x_n)_{n\in\mathbb{N}}$ is bounded. In turn, $(\overline{x}_n)_{n\in\mathbb{N}}$ is bounded and it follows from (3.2) that $(\forall x \in S)\ \sum_{n\in\mathbb{N}} \chi_n \nu_n(x) < +\infty$. An inspection of the proof of Corollary 3.3 then reveals immediately that its conclusions under any of assumptions (a)–(c) remain true.

4. Examples and applications. In this section we exhibit various existing results as special cases of our framework. Our purpose is not to exploit it to its full capacity but rather to illustrate its potential on simple instances. We first recover the main result of [21] on algorithm (1.4).

Example 4.1. We consider the setting studied in [21]. Let $(T_n)_{n\in\mathbb{N}}$ be a sequence of firmly quasi-nonexpansive operators from $\mathcal{H}$ to $\mathcal{H}$ such that $S = \bigcap_{n\in\mathbb{N}} \operatorname{Fix} T_n \neq \varnothing$. Then the problem of finding a point in S is a special case of Problem 1.1(c), where we assume that $\alpha_{1,n} \equiv 1/2$. In addition, let $(e_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that $\sum_{n\in\mathbb{N}} \|e_n\| < +\infty$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $]0,2[$ such that $0 < \inf_{n\in\mathbb{N}} \lambda_n \leqslant \sup_{n\in\mathbb{N}} \lambda_n < 2$, and let $(\mu_{n,j})_{n\in\mathbb{N},0\leqslant j\leqslant n}$ be an array with entries in $[0,+\infty[$ which satisfies the following:

(a) $(\forall n \in \mathbb{N})\ \sum_{j=0}^{n} \mu_{n,j} = 1$.
(b) $(\forall j \in \mathbb{N})\ \lim_{n\to+\infty} \mu_{n,j} = 0$.
(c) Every sequence $(\xi_n)_{n\in\mathbb{N}}$ in $[0,+\infty[$ such that

$$(4.1)\qquad \Big(\exists\, (\varepsilon_n)_{n\in\mathbb{N}} \in [0,+\infty[^{\mathbb{N}}\Big)\quad \begin{cases} \sum_{n\in\mathbb{N}} \varepsilon_n < +\infty,\\ (\forall n\in\mathbb{N})\quad \xi_{n+1} \leqslant \sum_{j=0}^{n} \mu_{n,j}\xi_j + \varepsilon_n,\end{cases}$$

converges.

Clearly, conditions (a)–(c) above imply that, in Algorithm 1.2, conditions (a)–(d) are satisfied. Now let $x_0 \in \mathcal{H}$, and define a sequence $(x_n)_{n\in\mathbb{N}}$ by

$$(4.2)\qquad (\forall n\in\mathbb{N})\quad x_{n+1} = \overline{x}_n + \lambda_n\big(T_n\overline{x}_n + e_n - \overline{x}_n\big),\ \text{ where }\ \overline{x}_n = \sum_{j=0}^{n} \mu_{n,j}x_j,$$

which corresponds to a 1-layer instance of (1.12). This mean iteration process was seen in [21] to cover several classical mean iteration methods, as well as memoryless convex feasibility algorithms [18] (see also [13]). The result obtained in [21, Theorem 3.5(i)] on the weak convergence of $(x_n)_{n\in\mathbb{N}}$ to a point in S corresponds to the special case of Corollary 3.2(iii) in which we further set $\chi_n \equiv 1$.

Next, we retrieve the main result of [25] on the convergence of an overrelaxed version of (1.7) and the special cases discussed there, in particular, those of [19].

Example 4.2. We consider the setting studied in [25], which corresponds to Problem 1.1(a). Given $x_0 \in \mathcal{H}$ and sequences $(e_{1,n})_{n\in\mathbb{N}}, \ldots, (e_{m,n})_{n\in\mathbb{N}}$ in $\mathcal{H}$ such that $\sum_{n\in\mathbb{N}} \lambda_n \sum_{i=1}^m \|e_{i,n}\| < +\infty$, construct a sequence $(x_n)_{n\in\mathbb{N}}$ via the m-layer recursion

$$
\begin{aligned}
(4.3)\quad (\forall n \in \mathbb{N})\quad x_{n+1} &= x_n \\
&+ \lambda_n\Big(T_{1,n}\Big(T_{2,n}\big(\cdots T_{m-1,n}(T_{m,n}x_n + e_{m,n}) + e_{m-1,n}\cdots\big) + e_{2,n}\Big) + e_{1,n} - x_n\Big), \\
&\text{where}\quad 0 < \lambda_n \leqslant \varepsilon + (1-\varepsilon)/\phi_n.
\end{aligned}
$$

Note that (4.3) corresponds to the memoryless version of (1.12). The result on the weak convergence of $(x_n)_{n\in\mathbb{N}}$ obtained in [25, Theorem 3.5(iii)] corresponds to the special case of Corollary 3.2(iv) in which the following additional assumptions are made:

(a) $(\forall n \in \mathbb{N})$ $\chi_n = 1$ and $(\forall j \in \{0, \ldots, n\})$ $\mu_{n,j} = \delta_{n,j}$.

(b) $(\exists\, \varepsilon \in \left]0,1\right[)(\forall n \in \mathbb{N})$ $\lambda_n \leqslant (1-\varepsilon)(1/\phi_n + \varepsilon)$.

Note that condition (a) above implies that, in Algorithm 1.2, conditions (a)–(c) trivially hold, while condition (d) follows from [10, Lemma 5.31]. We also observe that [25, Theorem 3.5(iii)] itself extends the results of [19, section 3], where the relaxation parameters $(\lambda_n)_{n\in\mathbb{N}}$ are confined to $\left]0,1\right]$.

The next two examples feature mean value and inertial iterations in the case of a single quasi-nonexpansive operator. As is easily seen, the memoryless algorithm (1.1) can fail to produce a convergent sequence in this scenario.

Example 4.3. Let $T\colon \mathcal{H} \to \mathcal{H}$ be a quasi-nonexpansive operator such that $\mathrm{Id} - T$ is demiclosed at 0 and $\operatorname{Fix} T \neq \varnothing$, let $(\mu_{n,j})_{n\in\mathbb{N}, 0\leqslant j\leqslant n}$ be an array in $\left[0,+\infty\right[$ that satisfies conditions (a)–(d) in Algorithm 1.2 with $\chi_n \equiv 1$ and such that $\inf_{n\in\mathbb{N}}\ \mu_{n+1,n}\mu_{n+1,n+1} > 0$, and let $(e_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that $\sum_{n\in\mathbb{N}} \|e_n\| < +\infty$. Let $x_0 \in \mathcal{H}$ and iterate

$$
(4.4)\qquad (\forall n \in \mathbb{N})\quad x_{n+1} = T\overline{x}_n + e_n, \quad \text{where}\quad \overline{x}_n = \sum_{j=0}^n \mu_{n,j}x_j.
$$

Then $T\overline{x}_n - \overline{x}_n \to 0$ and there exists $x \in \operatorname{Fix} T$ such that $x_n \rightharpoonup x$ and $\overline{x}_n \rightharpoonup x$.

Proof. We apply Corollary 3.2 in the setting of Problem 1.1(c) with $T_{1,n} \equiv T$, $\alpha_{1,n} \equiv 1$, $\phi_n \equiv 1$, and $\lambda_n \equiv 1$. First, note that (3.32) is satisfied. Furthermore, Corollary 3.2(v) entails that $\overline{x}_n - x_n \to 0$, while Corollary 3.2(i) yields $\mu_{n+1,n}\mu_{n+1,n+1}\|x_{n+1} - x_n\|^2 \to 0$ and hence $x_{n+1} - x_n \to 0$. Therefore $\overline{x}_n - T\overline{x}_n = (\overline{x}_n - x_n) + (x_n - x_{n+1}) + e_n \to 0$. Since $\mathrm{Id} - T$ is demiclosed at 0, it follows that every weak sequential cluster point of $(\overline{x}_n)_{n\in\mathbb{N}}$ is in $\operatorname{Fix} T$. In view of Corollary 3.2(v), the proof is complete. ∎

Example 4.4. Let $T\colon \mathcal{H} \to \mathcal{H}$ be a quasi-nonexpansive operator such that $\mathrm{Id} - T$ is demiclosed at 0 and $\operatorname{Fix} T \neq \varnothing$, and let $(\eta_n)_{n\in\mathbb{N}}$ be a sequence in $\left[0,1\right[$ such that $\eta_0 = 0$, $\eta = \sup_{n\in\mathbb{N}} \eta_n < 1$, and $(\forall n \in \mathbb{N})$ $\eta_n \leqslant \eta_{n+1}$. Let $(\sigma, \vartheta) \in \left]0,+\infty\right[^2$ be such that $(\eta^2(1+\eta) + \eta\sigma)/\vartheta < 1 - \eta^2$, and let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $\left]0,1\right[$ such that $0 < \inf_{n\in\mathbb{N}} \lambda_n \leqslant \sup_{n\in\mathbb{N}} \lambda_n \leqslant (\vartheta - \eta(\eta(1+\eta) + \eta\vartheta + \sigma))/(\vartheta(1 + \eta(1+\eta) + \eta\vartheta + \sigma))$. Let $x_0 \in \mathcal{H}$, set $x_{-1} = x_0$, and iterate

$$
(4.5)\quad (\forall n \in \mathbb{N})\quad x_{n+1} = \overline{x}_n + \lambda_n\big(T\overline{x}_n - \overline{x}_n\big), \quad \text{where}\quad \overline{x}_n = (1+\eta_n)x_n - \eta_n x_{n-1}.
$$

Then $T\overline{x}_n - \overline{x}_n \to 0$ and there exists $x \in \operatorname{Fix} T$ such that $x_n \rightharpoonup x$. In the case when T is nonexpansive, this result appears in [14, Theorem 5].

Proof. This is an instance of items (i) and (iv) in Corollary 3.3(d) and Problem 1.1(c) in which $T_{1,n} \equiv T$, $\alpha_{1,n} \equiv 1$, and $\phi_n \equiv 1$. Note that condition (d) in Corollary 3.3 is satisfied since $(\forall n \in \mathbb{N})\ \omega_n = 1 - \lambda_n < 1$. ∎

Next, we consider applications to monotone operator splitting. Let us recall basic notions about a set-valued operator $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ [10]. We denote by $\operatorname{ran} A = \big\{u \in \mathcal{H} \;\big|\; (\exists\, x \in \mathcal{H})\ u \in Ax\big\}$ the range of A, by $\operatorname{dom} A = \big\{x \in \mathcal{H} \;\big|\; Ax \neq \varnothing\big\}$ the domain of A, by $\operatorname{zer} A = \big\{x \in \mathcal{H} \;\big|\; 0 \in Ax\big\}$ the set of zeros of A, by $\operatorname{gra} A = \big\{(x,u) \in \mathcal{H} \times \mathcal{H} \;\big|\; u \in Ax\big\}$ the graph of A, and by A^{-1} the inverse of A, i.e., the operator with graph $\big\{(u,x) \in \mathcal{H} \times \mathcal{H} \;\big|\; u \in Ax\big\}$. The resolvent of A is $J_A = (\mathrm{Id} + A)^{-1}$ and $s\colon \operatorname{dom} A \to \mathcal{H}$ is a selection of A if $(\forall x \in \operatorname{dom} A)\ s(x) \in Ax$. Moreover, A is monotone if

$$(\forall (x,u) \in \operatorname{gra} A)(\forall (y,v) \in \operatorname{gra} A) \quad \langle x - y \mid u - v \rangle \geqslant 0, \tag{4.6}$$

and maximally monotone if there exists no monotone operator $B\colon \mathcal{H} \to 2^{\mathcal{H}}$ such that $\operatorname{gra} A \subset \operatorname{gra} B \neq \operatorname{gra} A$. In this case, J_A is a firmly nonexpansive operator defined everywhere on $\mathcal{H}$ and the reflector $R_A = 2J_A - \mathrm{Id}$ is nonexpansive. We denote by $\Gamma_0(\mathcal{H})$ the class of proper lower semicontinuous convex functions from $\mathcal{H}$ to $\left]-\infty, +\infty\right]$. Let $f \in \Gamma_0(\mathcal{H})$. For every $x \in \mathcal{H}$, $f + \|x - \cdot\|^2/2$ possesses a unique minimizer, which is denoted by $\operatorname{prox}_f x$. We have $\operatorname{prox}_f = J_{\partial f}$, where

$$\partial f\colon \mathcal{H} \to 2^{\mathcal{H}}\colon x \mapsto \big\{u \in \mathcal{H} \;\big|\; (\forall y \in \mathcal{H})\ \langle y - x \mid u \rangle + f(x) \leqslant f(y)\big\} \tag{4.7}$$

is the Moreau subdifferential of f. Our convergence results will rest on the following asymptotic principle.

LEMMA 4.5. *Let A and B be maximally monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$, let $(x_n, u_n)_{n \in \mathbb{N}}$ be a sequence in $\operatorname{gra} A$, let $(y_n, v_n)_{n \in \mathbb{N}}$ be a sequence in $\operatorname{gra} B$, let $x \in \mathcal{H}$, and let $v \in \mathcal{H}$. Suppose that $x_n \rightharpoonup x$, $v_n \rightharpoonup v$, $x_n - y_n \to 0$, and $u_n + v_n \to 0$. Then the following hold:*
(i) $(x, -v) \in \operatorname{gra} A$ *and* $(x, v) \in \operatorname{gra} B$.
(ii) $0 \in Ax + Bx$ *and* $0 \in -A^{-1}(-v) + B^{-1}v$.

Proof. Apply [10, Proposition 26.5] with $\mathcal{K} = \mathcal{H}$ and $L = \mathrm{Id}$. ∎

As discussed in [19], many splitting methods can be analyzed within the powerful framework of fixed point methods for averaged operators. The analysis provided in the present paper therefore makes it possible to develop new methods in this framework, for instance, mean value or inertial versions of standard splitting methods. We provide two such examples below. First, we consider the Peaceman–Rachford splitting method, which typically does not converge unless strong requirements are imposed on the underlying operators [20]. In the spirit of Mann's work [34], we show that mean iterations induce the convergence of this algorithm.

PROPOSITION 4.6. *Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ and $B\colon \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone operators such that $\operatorname{zer}(A + B) \neq \varnothing$ and let $\gamma \in \left]0, +\infty\right[$. Let $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ be sequences in $\mathcal{H}$ such that $\sum_{n \in \mathbb{N}} \|a_n\| < +\infty$ and $\sum_{n \in \mathbb{N}} \|b_n\| < +\infty$, let $x_0 \in \mathcal{H}$, and let $(\mu_{n,j})_{n \in \mathbb{N}, 0 \leqslant j \leqslant n}$ be an array in $\left[0, +\infty\right[$ that satisfies conditions* (a)–(d) *in*

Algorithm 1.2 *with* $\chi_n \equiv 1$ *and such that* $\inf_{n\in\mathbb{N}} \mu_{n+1,n}\mu_{n+1,n+1} > 0$. *Iterate*

$$\text{for } n = 0,1,\ldots \left\lfloor \begin{array}{l} \overline{x}_n = \sum_{j=0}^{n} \mu_{n,j} x_j, \\ y_n = J_{\gamma B}\overline{x}_n + b_n, \\ z_n = J_{\gamma A}(2y_n - \overline{x}_n) + a_n, \\ x_{n+1} = \overline{x}_n + 2(z_n - y_n). \end{array} \right. \tag{4.8}$$

Then there exists $x \in \operatorname{Fix} R_{\gamma A}R_{\gamma B}$ *such that* $x_n \rightharpoonup x$ *and* $\overline{x}_n \rightharpoonup x$. *Now set* $y = J_{\gamma B}x$. *Then* $y \in \operatorname{zer}(A+B)$, $z_n - y_n \to 0$, $y_n \rightharpoonup y$, *and* $z_n \rightharpoonup y$.

Proof. Set $T = R_{\gamma A}R_{\gamma B}$ and $(\forall n \in \mathbb{N})$ $e_n = 2a_n + R_{\gamma A}(R_{\gamma B}\overline{x}_n + 2b_n) - R_{\gamma A}(R_{\gamma B}\overline{x}_n)$. Then T is nonexpansive, $\operatorname{Id} - T$ is therefore demiclosed, and, since $\operatorname{zer}(A+B) \neq \varnothing$, [10, Proposition 26.1(iii)(b)] yields $\operatorname{Fix} T = \operatorname{Fix} R_{\gamma A}R_{\gamma B} \neq \varnothing$. In addition, we derive from (4.8) that

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = T\overline{x}_n + e_n, \tag{4.9}$$

where

$$\begin{aligned} \sum_{n\in\mathbb{N}} \|e_n\| &\leqslant \sum_{n\in\mathbb{N}} \big(2\|a_n\| + \|R_{\gamma A}(R_{\gamma B}\overline{x}_n + 2b_n) - R_{\gamma A}(R_{\gamma B}\overline{x}_n)\|\big) \\ &\leqslant \sum_{n\in\mathbb{N}} 2\big(\|a_n\| + \|b_n\|\big) \\ &< +\infty. \end{aligned} \tag{4.10}$$

Consequently, we deduce from Example 4.3 that $(x_n)_{n\in\mathbb{N}}$ and $(\overline{x}_n)_{n\in\mathbb{N}}$ converge weakly to a point $x \in \operatorname{Fix} T = \operatorname{Fix} R_{\gamma A}R_{\gamma B}$, and that $T\overline{x}_n - \overline{x}_n \to 0$. In addition, [10, Proposition 26.1(iii)(b)] asserts that $y \in \operatorname{zer}(A+B)$. Next, we derive from (4.8) and (4.9) that $2(z_n - y_n) = x_{n+1} - \overline{x}_n = (T\overline{x}_n - \overline{x}_n) + e_n \to 0$. It remains to show that $y_n \rightharpoonup y$. Since $(\overline{x}_n)_{n\in\mathbb{N}}$ converges weakly, it is bounded. However, $(\forall n \in \mathbb{N})$ $\|y_n - y_0\| = \|J_{\gamma B}\overline{x}_n - J_{\gamma B}x_0 + b_n\| \leqslant \|\overline{x}_n - x_0\| + \|b_n\|$. Therefore $(y_n)_{n\in\mathbb{N}}$ is bounded. Now let z be a weak sequential cluster point of $(y_n)_{n\in\mathbb{N}}$, say $y_{k_n} \rightharpoonup z$. In view of [10, Lemma 2.46], it is enough to show that $z = y$. To this end, set $(\forall n \in \mathbb{N})$ $v_n = \gamma^{-1}(\overline{x}_n - y_n + b_n)$ and $w_n = \gamma^{-1}(2y_n - \overline{x}_n - z_n + a_n)$. Then $(\forall n \in \mathbb{N})$ $(z_n - a_n, w_n) \in \operatorname{gra} A$ and $(y_n - b_n, v_n) \in \operatorname{gra} B$. In addition, we have $\overline{x}_{k_n} \rightharpoonup x$, $z_{k_n} - a_{k_n} \rightharpoonup z$, $v_{k_n} \rightharpoonup \gamma^{-1}(x - z)$, $(z_{k_n} - a_{k_n}) - (y_{k_n} - b_{k_n}) \to 0$, and $v_{k_n} + w_{k_n} = \gamma^{-1}(y_{k_n} - z_{k_n} + a_{k_n} + b_{k_n}) \to 0$. Hence, we derive from Lemma 4.5(i) that $(z, \gamma^{-1}(x - z)) \in \operatorname{gra} B$, i.e., $z = J_{\gamma B}x = y$. ∎

Remark 4.7. Let f and g be functions in $\Gamma_0(\mathcal{H})$, and specialize Proposition 4.6 to $A = \partial f$ and $B = \partial g$. Then $\operatorname{zer}(A+B) = \operatorname{Argmin}(f+g)$. Moreover, (4.8) becomes

$$\text{for } n = 0,1,\ldots \left\lfloor \begin{array}{l} \overline{x}_n = \sum_{j=0}^{n} \mu_{n,j} x_j, \\ y_n = \operatorname{prox}_{\gamma g}\overline{x}_n + b_n, \\ z_n = \operatorname{prox}_{\gamma f}(2y_n - \overline{x}_n) + a_n, \\ x_{n+1} = \overline{x}_n + 2(z_n - y_n), \end{array} \right. \tag{4.11}$$

and we conclude that there exists a point $y \in \operatorname{Argmin}(f+g)$ such that $y_n \rightharpoonup y$ and $z_n \rightharpoonup y$.

We now propose a new forward-backward splitting framework which includes existing instances as special cases. The following notion will be needed to establish strong convergence properties (see [4, Proposition 2.4] for special cases).

DEFINITION 4.8 (see [4, Definition 2.3]). *An operator* $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ *is* demiregular *at* $x \in \operatorname{dom} A$ *if, for every sequence* $(x_n, u_n)_{n\in\mathbb{N}}$ *in* $\operatorname{gra} A$ *and every* $u \in Ax$ *such that* $x_n \rightharpoonup x$ *and* $u_n \to u$*, we have* $x_n \to x$.

PROPOSITION 4.9. *Let* $\beta \in \left]0,+\infty\right[$, *let* $\varepsilon \in \left]0,\min\{1/2,\beta\}\right[$, *let* $x_0 \in \mathcal{H}$, *let* $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ *be maximally monotone, and let* $B\colon \mathcal{H} \to \mathcal{H}$ *be* β*-cocoercive, i.e.,*

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \langle x - y \mid Bx - By\rangle/\beta \geqslant \|Bx - By\|^2. \tag{4.12}$$

Furthermore, let $(\gamma_n)_{n\in\mathbb{N}}$ *be a sequence in* $[\varepsilon, 2\beta/(1+\varepsilon)]$, *let* $(\mu_{n,j})_{n\in\mathbb{N},0\leqslant j\leqslant n}$ *be a real array that satisfies conditions* (a)–(d) *in Algorithm* 1.2, *and let* $(a_n)_{n\in\mathbb{N}}$ *and* $(b_n)_{n\in\mathbb{N}}$ *be sequences in* $\mathcal{H}$ *such that* $\sum_{n\in\mathbb{N}} \chi_n\|a_n\| < +\infty$ *and* $\sum_{n\in\mathbb{N}} \chi_n\|b_n\| < +\infty$. *Suppose that* $\operatorname{zer}(A+B) \neq \varnothing$ *and that*

$$(\forall n \in \mathbb{N}) \quad \lambda_n \in \left[\varepsilon, 1 + (1-\varepsilon)\left(1 - \frac{\gamma_n}{2\beta}\right)\right], \tag{4.13}$$

and set $(\forall n \in \mathbb{N})$ $\phi_n = 2/(4 - \gamma_n/\beta)$. *For every* $n \in \mathbb{N}$, *iterate*

$$x_{n+1} = \overline{x}_n + \lambda_n\Big(J_{\gamma_n A}\big(\overline{x}_n - \gamma_n(B\overline{x}_n + b_n)\big) + a_n - \overline{x}_n\Big), \quad \textit{where} \quad \overline{x}_n = \sum_{j=0}^{n} \mu_{n,j} x_j. \tag{4.14}$$

Suppose that one of the following is satisfied:

(a) $\inf_{n\in\mathbb{N}} \min_{0\leqslant j\leqslant n} \mu_{n,j} \geqslant 0$.

(b) $a_n \equiv b_n \equiv 0$, $(\mu_{n,j})_{n\in\mathbb{N},0\leqslant j\leqslant n}$ *satisfies* (3.37), *and one of conditions* (a)–(d) *in Corollary* 3.3 *is satisfied.*

Then the following hold:

(i) $J_{\gamma_n A}(\overline{x}_n - \gamma_n B\overline{x}_n) - \overline{x}_n \to 0$.

(ii) *Let* $z \in \operatorname{zer}(A+B)$. *Then* $B\overline{x}_n \to Bz$.

(iii) *There exists* $x \in \operatorname{zer}(A+B)$ *such that* $x_n \rightharpoonup x$.

(iv) *Suppose that* A *or* B *is demiregular at every point in* $\operatorname{zer}(A+B)$. *Then there exists* $x \in \operatorname{zer}(A+B)$ *such that* $x_n \to x$.

Proof. We apply Corollary 3.2 in case (a) and from Corollary 3.3 in case (b). We first note that (4.14) is an instance of Algorithm 1.2 with $m = 2$ and $(\forall n \in \mathbb{N})$ $T_{1,n} = J_{\gamma_n A}$, $T_{2,n} = \operatorname{Id} - \gamma_n B$, $e_{1,n} = a_n$, and $e_{2,n} = -\gamma_n b_n$. Indeed, for every $n \in \mathbb{N}$, $T_{1,n}$ is $\alpha_{1,n}$-averaged with $\alpha_{1,n} = 1/2$ [10, Remark 4.34(iii) and Corollary 23.9], $T_{2,n}$ is $\alpha_{2,n}$-averaged with $\alpha_{2,n} = \gamma_n/(2\beta)$ [10, Proposition 4.39], and the averaging constant of $T_{1,n}T_{2,n}$ is therefore given by (1.13) as

$$\frac{\alpha_{1,n} + \alpha_{2,n} - 2\alpha_{1,n}\alpha_{2,n}}{1 - \alpha_{1,n}\alpha_{2,n}} = \frac{2}{4 - \gamma_n/\beta} = \phi_n. \tag{4.15}$$

On the other hand, we are in the setting of Problem 1.1(a) since [10, Proposition 26.1(iv)(a)] yields $(\forall n \in \mathbb{N})$ $\operatorname{Fix}(T_{1,n}T_{2,n}) = \operatorname{zer}(A+B) \neq \varnothing$. We also observe that, in view of (4.13),

$$(\forall n \in \mathbb{N}) \quad \varepsilon \leqslant \lambda_n \leqslant \varepsilon + \frac{1-\varepsilon}{\phi_n}, \quad \frac{1-\alpha_{1,n}}{\alpha_{1,n}} = 1, \quad \text{and} \quad \frac{1-\alpha_{2,n}}{\alpha_{2,n}} \geqslant \varepsilon \tag{4.16}$$

which, by (1.13), yields

$$\sup_{n\in\mathbb{N}} \phi_n \leqslant \frac{1+\varepsilon}{1+2\varepsilon} < 1. \tag{4.17}$$

In addition, it results from (4.15) that $(\forall n \in \mathbb{N})\ \lambda_n \leqslant 1/\phi_n+\varepsilon \leqslant 2-\gamma_n/(2\beta)+\varepsilon \leqslant 2+\varepsilon$. Therefore,

$$\begin{cases}\sum_{n\in\mathbb{N}} \chi_n\lambda_n\|e_{1,n}\| = (2+\varepsilon)\sum_{n\in\mathbb{N}} \chi_n\|a_n\| < +\infty,\\ \sum_{n\in\mathbb{N}} \chi_n\lambda_n\|e_{2,n}\| \leqslant 2\beta(2+\varepsilon)\sum_{n\in\mathbb{N}} \chi_n\|b_n\| < +\infty,\end{cases} \tag{4.18}$$

which establishes (3.32) for case (a). Altogether, (4.16), Corollary 3.2(ii), and Corollary 3.3(ii) imply that, for every $z \in \operatorname{zer}(A+B)$,

$$\begin{cases}(T_{1,n}-\mathrm{Id}\,)T_{2,n}\overline{x}_n + (T_{2,n}-\mathrm{Id}\,)z = (T_{1,n}-\mathrm{Id}\,)T_{2,n}\overline{x}_n - (T_{1,n}-\mathrm{Id}\,)T_{2,n}z \to 0,\\ (T_{2,n}-\mathrm{Id}\,)\overline{x}_n - (T_{2,n}-\mathrm{Id}\,)z \to 0.\end{cases} \tag{4.19}$$

Now set

$$(\forall n \in \mathbb{N}) \quad y_n = J_{\gamma_n A}(\overline{x}_n - \gamma_n B\overline{x}_n), \quad u_n = \frac{\overline{x}_n - y_n}{\gamma_n} - B\overline{x}_n, \quad \text{and} \quad v_n = B\overline{x}_n, \tag{4.20}$$

and note that

$$(\forall n \in \mathbb{N}) \quad u_n \in Ay_n. \tag{4.21}$$

(i): Let $z \in \operatorname{zer}(A+B)$. Then (4.19) yields $J_{\gamma_n A}(\overline{x}_n - \gamma_n B\overline{x}_n) - \overline{x}_n = (T_{1,n}-\mathrm{Id}\,)T_{2,n}\overline{x}_n + (T_{2,n}-\mathrm{Id}\,)z + (T_{2,n}-\mathrm{Id}\,)\overline{x}_n - (T_{2,n}-\mathrm{Id}\,)z \to 0$.

(ii): We derive from (4.19) that

$$\|B\overline{x}_n - Bz\| = \gamma_n^{-1}\|T_{2,n}\overline{x}_n - \overline{x}_n - T_{2,n}z + z\| \leqslant \varepsilon^{-1}\|T_{2,n}\overline{x}_n - \overline{x}_n - T_{2,n}z + z\| \to 0. \tag{4.22}$$

(iii): Let $(k_n)_{n\in\mathbb{N}}$ be a strictly increasing sequence in $\mathbb{N}$ and let $y \in \mathcal{H}$ be such that $\overline{x}_{k_n} \rightharpoonup y$. In view of Corollary 3.2(iv) in case (a), and Corollary 3.3(iv) in case (b), it remains to show that $y \in \operatorname{zer}(A+B)$. We derive from (i) that $y_n - \overline{x}_n \to 0$. Hence $y_{k_n} \rightharpoonup y$. Now let $z \in \operatorname{zer}(A+B)$. Then (ii) implies that $B\overline{x}_n \to Bz$. Altogether, $y_{k_n} \rightharpoonup y$, $v_{k_n} \rightharpoonup Bz$, $y_{k_n} - \overline{x}_{k_n} \to 0$, $u_{k_n} + v_{k_n} \to 0$, and, for every $n \in \mathbb{N}$, $u_{k_n} \in Ay_{k_n}$ and $v_{k_n} \in B\overline{x}_{k_n}$. It therefore follows from Lemma 4.5(ii) that $y \in \operatorname{zer}(A+B)$.

(iv): By (iii), there exists $x \in \operatorname{zer}(A+B)$ such that $x_n \rightharpoonup x$. In addition, we derive from Corollary 3.2(iv) or Corollary 3.3(iii) that

$$\overline{x}_n - x_{n+1} \to 0 \quad \text{or} \quad \overline{x}_n - x_n \to 0. \tag{4.23}$$

Hence it follows from (4.20) and (i) that $y_n \rightharpoonup x$, and, from (4.20) and (ii), that $u_n \to -Bx$. In turn, if A is demiregular on $\operatorname{zer}(A+B)$, we derive from (4.21) that $y_n \to x$. Since $y_n - \overline{x}_n \to 0$, (4.23) yields $x_n \to x$. Now suppose that B is demiregular on $\operatorname{zer}(A+B)$. Since $\overline{x}_n \rightharpoonup x$, (ii) implies that $\overline{x}_n \to x$ and it follows from (4.23) that $x_n \to x$. ∎

Remark 4.10. As noted in Remark 3.4, we can allow errors in inertial multilayer methods and, in particular, in the inertial forward-backward algorithm. Thus, suppose that, in Proposition 4.9, $\lambda_n \equiv 1$ and A has bounded domain. Then $\bigcup_{n\in\mathbb{N}} \operatorname{ran} T_{1,n} = \bigcup_{n\in\mathbb{N}} \operatorname{ran} (\mathrm{Id} + \gamma_n A)^{-1} = \operatorname{dom} A$ is bounded. Hence, it follows from Remark 3.4 that, if $\sum_{n\in\mathbb{N}} \chi_n \|a_n\| < +\infty$ and $\sum_{n\in\mathbb{N}} \chi_n \|b_n\| < +\infty$, the conclusions of Proposition 4.9(b) under any of assumptions (a)–(c) of Corollary 3.3 remain true for the inertial forward-backward algorithm

$$\begin{array}{l} \text{for } n = 0, 1, \ldots \\ \left\lfloor \begin{array}{l} \overline{x}_n = (1+\eta_n)x_n - \eta_n x_{n-1}, \\ x_{n+1} = J_{\gamma_n A}\big(\overline{x}_n - \gamma_n(B\overline{x}_n + b_n)\big) + a_n. \end{array} \right. \end{array} \tag{4.24}$$

Remark 4.11. Let $f \in \Gamma_0(\mathcal{H})$, let $g\colon \mathcal{H} \to \mathbb{R}$ be convex and differentiable with a $1/\beta$-Lipschitzian gradient, and suppose that $\operatorname{Argmin}(f+g) \neq \varnothing$. Then ∇g is β-cocoercive [10, Corollary 18.17]. Upon setting $A = \partial f$ and $B = \nabla g$ in Proposition 4.9, we see that, for every $n \in \mathbb{N}$, (4.14) becomes

$$x_{n+1} = \overline{x}_n + \lambda_n \Big(\operatorname{prox}_{\gamma_n f}\big(\overline{x}_n - \gamma_n(\nabla g(\overline{x}_n) + b_n)\big) + a_n - \overline{x}_n\Big), \quad \text{where} \quad \overline{x}_n = \sum_{j=0}^{n} \mu_{n,j} x_j, \tag{4.25}$$

and we conclude that there exists $x \in \operatorname{Argmin}(f+g)$ such that $x_n \rightharpoonup x$ and $\nabla g(\overline{x}_n) \to \nabla g(x)$.

Remark 4.12. Various results on the convergence of the forward-backward splitting algorithm can be recovered from Proposition 4.9.

(i) Suppose that $(\forall n \in \mathbb{N})$ $\lambda_n \leqslant 1$ and $(\forall j \in \{0, \ldots, n\})$ $\mu_{n,j} = \delta_{n,j}$. Then conditions (a)–(d) in Algorithm 1.2 hold with $\chi_n \equiv 1$ by [10, Lemma 5.31]. In turn, Proposition 4.9(a)(iii) reduces to [19, Corollary 6.5]. In the context of Remark 4.11, Proposition 4.9(a)(iii) captures [24, Theorem 3.4(i)]. In this setting, under suitable conditions on the errors, it is shown in [23, Theorem 3(vi)] that $(f+g)(x_n) - \min(f+g)(\mathcal{H}) = o(1/n)$.

(ii) In the context of Remark 4.11, let $(\eta_n)_{n\in\mathbb{N}}$ be a sequence in $[0,1[$ that satisfies condition (b) in Corollary 3.3, and set $x_{-1} = x_0$ and $\eta_0 = 0$. If $\gamma_n \equiv \gamma_0 \leqslant \beta$ and $\lambda_n \equiv 1$, Proposition 4.9(b)(iii) covers the scheme

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \operatorname{prox}_{\gamma_0 f}\Big(x_n + \eta_n(x_n - x_{n-1}) - \gamma_0 \nabla g\big(x_n + \eta_n(x_n - x_{n-1})\big)\Big) \tag{4.26}$$

studied in [17, Theorem 4.1], where it was established that

$$\sum_{n\in\mathbb{N}} n\|x_n - x_{n-1}\|^2 < +\infty. \tag{4.27}$$

In this case, it is shown in [6, Theorem 1] that $(f+g)(x_n) - \min(f+g)(\mathcal{H}) = o(1/n^2)$.

(iii) If $\lambda_n \equiv 1$, then Proposition 4.9(b)(iii) under hypothesis (c) of Corollary 3.3 establishes a statement made in [33, Theorem 1]. Let us note, however, that the proof of [33] is not convincing as the authors appear to use the weak continuity of some operators which are merely strongly continuous.

(iv) Suppose that $(\forall n \in \mathbb{N})$ $\lambda_n \leqslant (1-\varepsilon)(2+\varepsilon-\gamma_n/(2\beta))$ and $(\forall j \in \{0, \ldots, n\})$ $\mu_{n,j} = \delta_{n,j}$. Then items (a)(iii) and (a)(iv) of Proposition 4.9 capture, respectively, items (iii) and (iv)(a)–(b) of [25, Proposition 4.4]. In addition, in the context of Remark 4.11, Proposition 4.9(a)(iii) captures [25, Proposition 4.7(iii)].

(v) Proposition 4.9 also applies to the proximal point algorithm. Indeed, when $B = 0$, it suffices to allow $\beta = +\infty$ and $\alpha_{2,n} \equiv 0$, to set $1/+\infty = 0$ and $1/0 = +\infty$, and to take $(\gamma_n)_{n\in\mathbb{N}}$ in $[\varepsilon, +\infty[$. In this setting, the proof remains valid and

(a) Items (i) and (iii) of Proposition 4.9(b) under hypothesis (c) of Corollary 3.3 capture the error-free case of [2, Theorem 3.1], while Theorem 3.1 covers its general case.

(b) Let $\eta \in]0, 1/3[$, set $\sigma = (1-3\eta)/2$ and $\vartheta = 2/3$, and suppose that $\lambda_n \equiv 1$. Then Proposition 4.9(b) under hypothesis (d) of Corollary 3.3 yields [3, Proposition 2.1].

Next, we derive from Corollary 3.2 a mean value extension of Polyak's subgradient projection method [40] (likewise, an inertial version can be derived from Corollary 3.3).

Example 4.13. Let C be a nonempty closed convex subset of $\mathcal{H}$ with projector P_C, let $f\colon \mathcal{H} \to \mathbb{R}$ be a continuous convex function such that $\operatorname{Argmin}_C f \neq \varnothing$ and $\theta = \min f(C)$ is known. Suppose that one of the following holds:

(i) f is bounded on every bounded subset of $\mathcal{H}$.

(ii) The conjugate f^* of f is supercoercive, i.e., $\lim_{\|u\|\to+\infty} f^*(u)/\|u\| = +\infty$.

(iii) $\mathcal{H}$ is finite dimensional.

Let $\eta \in]0, 1[$, let $\varepsilon \in]0, \eta/(2+\eta)[$, let $(\mu_{n,j})_{n\in\mathbb{N},0\leqslant j\leqslant n}$ be an array in $[0, +\infty[$ that satisfies conditions (a)–(d) in Algorithm 1.2, let $(\xi_n)_{n\in\mathbb{N}}$ be in $[\eta, 2-\eta]$, let $(\lambda_n)_{n\in\mathbb{N}}$ be in $[\varepsilon, (1-\varepsilon)(2-\xi_n/2)]$, let s be a selection of ∂f, and let $x_0 \in C$. Iterate

(4.28)
$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots\\
\left\lfloor
\begin{array}{l}
\overline{x}_n = \displaystyle\sum_{j=0}^{n} \mu_{n,j} x_j,\\
x_{n+1} = \begin{cases} \overline{x}_n + \lambda_n\left(P_C\left(\overline{x}_n + \xi_n \dfrac{\theta - f(\overline{x}_n)}{\|s(\overline{x}_n)\|^2} s(\overline{x}_n)\right) - \overline{x}_n\right) & \text{if } s(\overline{x}_n) \neq 0,\\ \overline{x}_n & \text{if } s(\overline{x}_n) = 0. \end{cases}
\end{array}
\right.
\end{array}
$$

Then there exists $x \in \operatorname{Argmin}_C f$ such that $x_n \rightharpoonup x$.

Proof. Let G be the subgradient projector onto $D = \{x \in \mathcal{H} \mid f(x) \leqslant \theta\}$ associated with s, that is,

$$
G\colon \mathcal{H} \to \mathcal{H}\colon x \mapsto \begin{cases} x + \dfrac{\theta - f(x)}{\|s(x)\|^2} s(x) & \text{if } f(x) > \theta,\\ x & \text{if } f(x) \leqslant \theta. \end{cases} \tag{4.29}
$$

Then G is firmly quasi-nonexpansive [10, Proposition 29.41(iii)]. Now set $(\forall n \in \mathbb{N})$ $T_{1,n} = P_C$, $T_{2,n} = \mathrm{Id} + \xi_n(G - \mathrm{Id})$, $\alpha_{1,n} = 1/2$, and $\alpha_{2,n} = \xi_n/2$. Then, for every $n \in \mathcal{H}$, $T_{1,n}$ is an $\alpha_{1,n}$-averaged nonexpansive operator [10, Proposition 4.16], $T_{2,n}$ is an $\alpha_{2,n}$-averaged quasi-nonexpansive operator, [10, Proposition 4.49(i)] yields

$$
\operatorname{Fix} T_{1,n}T_{2,n} = \operatorname{Fix} T_{1,n} \cap \operatorname{Fix} T_{2,n} = C \cap D = \operatorname{Argmin}_C f, \tag{4.30}
$$

and Remark 1.3(i) asserts that $T_{1,n}T_{2,n}$ is an averaged quasi-nonexpansive operator with constant $\phi_n = 2/(4-\xi_n)$ and $\lambda_n \leqslant (1-\varepsilon)/\phi_n$. Thus, the problem of minimizing f over C is a special case of Problem 1.1(b) with $m = 2$, and (4.28) is a special case of Algorithm 1.2 with $e_{1,n} \equiv 0$ and $e_{2,n} \equiv 0$. Now let $(k_n)_{n\in\mathbb{N}}$ be a strictly increasing sequence in $\mathbb{N}$ and let $x \in \mathcal{H}$ be such that $\overline{x}_{k_n} \rightharpoonup x$. Then, by Corollary 3.2(iii), it remains to show that $x \in C \cap D$. We derive from Corollary 3.2(ii) and (4.30) that $T_{2,k_n}\overline{x}_{k_n} - P_C T_{2,k_n}\overline{x}_{k_n} \to 0$ and $\overline{x}_{k_n} - T_{2,k_n}\overline{x}_{k_n} \to 0$. Therefore $C \ni P_C T_{2,k_n}\overline{x}_{k_n} = (P_C T_{2,k_n}\overline{x}_{k_n} - T_{2,n}\overline{x}_{k_n}) + (T_{2,k_n}\overline{x}_{k_n} - \overline{x}_{k_n}) + \overline{x}_{k_n} \rightharpoonup x$ and, since C is weakly closed, $x \in C$. On the other hand, $\|G\overline{x}_n - \overline{x}_n\| = \|T_{2,n}\overline{x}_n - \overline{x}_n\|/\xi_n \leqslant \|T_{2,n}\overline{x}_n - \overline{x}_n\|/\eta \to 0$. Since (iii)⇒(ii)⇔(i) [10, Proposition 16.20] and (i) imply that $\mathrm{Id} - G$ is demiclosed at 0 [10, Proposition 29.41(vii)], we conclude that $x \in \operatorname{Fix} G = D$. ∎

Remark 4.14. Example 4.13 reverts to Polyak's classical result [40, Theorem 1] in the case when $(\forall n \in \mathbb{N})$ $\lambda_n = 1$ and $(\forall j \in \{0, \ldots, n\})$ $\mu_{n,j} = \delta_{n,j}$. The unrelaxed pattern $\lambda_n \equiv 1$ is indeed achievable because $(\forall n \in \mathbb{N})$ $\lambda_n \in [\varepsilon, (1-\varepsilon)(2-\xi_n/2)]$ and $(1-\varepsilon)(2-\xi_n/2) \geqslant (1-\varepsilon)(2-(2-\eta)/2) > (1-\eta/(2+\eta))(1+\eta/2) = 1$.

REFERENCES

[1] F. Alvarez, *On the minimizing property of a second order dissipative system in Hilbert spaces,* SIAM J. Control Optim., 38 (2000), pp. 1102–1119.

[2] F. Alvarez, *Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space,* SIAM J. Optim., 14 (2004), pp. 773–782.

[3] F. Alvarez and H. Attouch, *An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping,* Set-Valued Anal., 9 (2001), pp. 3–11.

[4] H. Attouch, L. M. Briceño-Arias, and P. L. Combettes, *A parallel splitting method for coupled monotone inclusions,* SIAM J. Control Optim., 48 (2010), pp. 3246–3270.

[5] H. Attouch, X. Goudou, and P. Redont, *The heavy ball with friction method,* I. *The continuous dynamical system: Global exploration of the local minima of real-valued function by asymptotic analysis of a dissipative dynamical system,* Commun. Contemp. Math., 2 (2000), pp. 1–34.

[6] H. Attouch and J. Peypouquet, *The rate of convergence of Nesterov's accelerated forward-backward method is actually faster than $1/k^2$,* SIAM J. Optim., 26 (2016), pp. 1824–1834.

[7] J. B. Baillon, R. E. Bruck, and S. Reich, *On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces,* Houston J. Math., 4 (1978), pp. 1–9.

[8] D. Barro and E. Canestrelli, *Combining stochastic programming and optimal control to decompose multistage stochastic optimization problems,* OR Spectrum, 38 (2016), pp. 711–742.

[9] H. H. Bauschke and J. M. Borwein, *On projection algorithms for solving convex feasibility problems,* SIAM Rev., 38 (1996), pp. 367–426.

[10] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces,* 2nd ed., Springer, New York, 2017.

[11] E. Borel, *Leçons sur les Séries Divergentes,* Gauthier-Villars, Paris, 1901.

[12] D. Borwein and J. Borwein, *Fixed point iterations for real functions,* J. Math. Anal. Appl., 157 (1991), pp. 112–126.

[13] J. M. Borwein, G. Li, and M. K. Tam, *Convergence rate analysis for averaged fixed point iterations in common fixed point problems,* SIAM J. Optim., 27 (2017), pp. 1–33.

[14] R. I. Boţ, E. R. Csetnek, and C. Hendrich, *Inertial Douglas-Rachford splitting for monotone inclusion problems,* Appl. Math. Comput., 256 (2015), pp. 472–487.

[15] C. Brezinski and J.-P. Chehab, *Nonlinear hybrid procedures and fixed point iterations,* Numer. Funct. Anal. Optim., 19 (1998), pp. 465–487.

[16] C. L. Byrne, *Iterative Optimization in Inverse Problems,* CRC Press, Boca Raton, FL, 2014.

[17] A. Chambolle and C. Dossal, *On the convergence of the iterates of the "Fast iterative shrinkage/thresholding algorithm,"* J. Optim. Theory Appl., 166 (2015), pp. 968–982.

[18] P. L. Combettes, *Quasi-Fejérian analysis of some optimization algorithms,* in Inherently Parallel Algorithms for Feasibility and Optimization, D. Butnariu, Y. Censor, and S. Reich, eds., Elsevier, New York, 2001, pp. 115–152.

[19] P. L. Combettes, *Solving monotone inclusions via compositions of nonexpansive averaged operators,* Optimization, 53 (2004), pp. 475–504.

[20] P. L. Combettes, *Iterative construction of the resolvent of a sum of maximal monotone operators,* J. Convex Anal., 16 (2009), pp. 727–748.

[21] P. L. Combettes and T. Pennanen, *Generalized Mann iterates for constructing fixed points in Hilbert spaces,* J. Math. Anal. Appl., 275 (2002), pp. 521–536.

[22] P. L. Combettes and J.-C. Pesquet, *Stochastic quasi-Fejér block-coordinate fixed point iterations with random sweeping,* SIAM J. Optim., 25 (2015), pp. 1221–1248.
[23] P. L. Combettes, S. Salzo, and S. Villa, *Consistent learning by composite proximal thresholding,* Math. Program., to appear.
[24] P. L. Combettes and V. R. Wajs, *Signal recovery by proximal forward-backward splitting,* Multiscale Model. Simul., 4 (2005), pp. 1168–1200.
[25] P. L. Combettes and I. Yamada, *Compositions and convex combinations of averaged nonexpansive operators,* J. Math. Anal. Appl., 425 (2015), pp. 55–70.
[26] D. Davis, *Convergence rate analysis of primal-dual splitting schemes,* SIAM J. Optim., 25 (2015), pp. 1912–1943.
[27] J. B. Diaz and F. T. Metcalf, *On the structure of the set of subsequential limit points of successive approximations,* Bull. Amer. Math. Soc. (N.S), 73 (1967), pp. 516–519.
[28] W. G. Dotson, *On the Mann iterative process,* Trans. Amer. Math. Soc., 149 (1970), pp. 65–73.
[29] L. Fejér, *Untersuchungen über Fouriersche Reihen,* Math. Ann., 58 (1903), pp. 51–69.
[30] C. W. Groetsch, *A note on segmenting Mann iterates,* J. Math. Anal. Appl., 40 (1972), pp. 369–372.
[31] K. Knopp, *Theory and Application of Infinite Series,* Blackie & Son, London, 1954.
[32] P. Kügler and A. Leitão, *Mean value iterations for nonlinear elliptic Cauchy problems,* Numer. Math., 96 (2003), pp. 269–293.
[33] D. A. Lorenz and T. Pock, *An inertial forward-backward algorithm for monotone inclusions,* J. Math. Imaging Vision, 51 (2015), pp. 311–325.
[34] W. R. Mann, *Mean value methods in iteration,* Proc. Amer. Math. Soc., 4 (1953), pp. 506–510.
[35] W. R. Mann, *Averaging to improve convergence of iterative processes,* in Fundamental Analysus Methods in Numerical Analysis, Lecture Notes in Math. 701, Springer, Berlin, 1979, pp. 169–179.
[36] J. J. Moreau, *Fonctions convexes duales et points proximaux dans un espace Hilbertien,* C. R. Acad. Sci. Paris Sér. A, 255 (1962), pp. 2897–2899.
[37] C. Outlaw and C. W. Groetsch, *Averaging iteration in a Banach space,* Bull. Amer. Math. Soc. (N.S.), 75 (1969), pp. 430–432.
[38] B. T. Polyak, *Gradient methods for minimizing functionals,* USSR Comput. Math. Math. Phys., 3 (1963), pp. 864–878.
[39] B. T. Polyak, *Some methods of speeding up the convergence of iteration methods,* USSR Comput. Math. Math. Phys., 4 (1964), pp. 1–17.
[40] B. T. Polyak, *Minimization of unsmooth functionals,* USSR Comput. Math. Math. Phys., 9 (1969), pp. 14–29.
[41] B. T. Polyak, *Comparison of the speed of convergence of single-step and multistep algorithms of optimization when there are noises,* Engrg. Cybernet., 15 (1977), pp. 6–10.
[42] B. E. Rhoades, *Fixed point iterations using infinite matrices,* Trans. Amer. Math. Soc., 196 (1974), pp. 161–176.
[43] K. Slavakis and I. Yamada, *The adaptive projected subgradient method constrained by families of quasi-nonexpansive mappings and its application to online learning,* SIAM J. Optim., 23 (2013), pp. 126–152.
[44] O. Toeplitz, *Über allgemeine lineare Mittelbildungen,* Prace Mat.-Fiz., 22 (1911), pp. 113–119.