(Article begins on next page)

Scaled, inexact and adaptive generalized FISTA for strongly convex optimization

S. Rebegoldi (Dipartimento di Ingegneria Industriale, UniFi, Italy),
L. Calatroni (I3S, CNRS, UCA, Inria, France).

Abstract

We consider a variable metric and inexact version of the FISTA-type algorithm considered in [11, 13] for the minimization of the sum of two (possibly strongly) convex functions. The proposed algorithm is combined with an adaptive (non-monotone) backtracking strategy, which allows for the adjustment of the algorithmic step-size along the iterations in order to improve the convergence speed. We prove a linear convergence result for the function values, which depends on both the strong convexity moduli of the two functions and the upper and lower bounds on the spectrum of the variable metric operators. We validate the proposed algorithm, named Scaled Adaptive GEneralized FISTA (SAGE-FISTA), on exemplar image denoising and deblurring problems where edge-preserving Total Variation (TV) regularization is combined with Kullback-Leibler-type fidelity terms, as it is common in applications where signal-dependent Poisson noise is assumed in the data.

Keywords: Convex optimization, inertial forward-backward splitting, variable metric, adaptive backtracking, image restoration.

1 Introduction

In recent years, inertial Forward–Backward (FB) methods have become standard tools for solving composite convex optimization problems of the form

$$\min_{x \in \mathcal{H}} F(x) \equiv f(x) + g(x), \tag{1}$$

where $\mathcal{H}$ is an Hilbert space, $f, g : \mathcal{H} \to \mathbb{R} \cup \{\infty\}$ are convex functions and f is continuously differentiable on a subset $Y \subseteq \mathcal{H}$. In its standard form, the FB algorithm [14, 16] is defined as the sequential application of a (forward) gradient step on the smooth part f followed by a (backward) proximal step on the non-smooth term g, thus resulting in the following iterative scheme defined for all $k \geq 0$

$$x^{(k+1)} = \text{prox}_{\tau_k g}(x^{(k)} - \tau_k \nabla f(x^{(k)})) = \underset{z \in \mathcal{H}}{\text{argmin}} \; g(z) + \frac{1}{2\tau_k} \| z - x^{(k)} + \tau_k \nabla f(x^{(k)}) \|^2, \tag{2}$$

where $\tau_k > 0$ is the algorithmic step-size and $\text{prox}_{\tau_k g}$ denotes the proximal operator associated to the convex function $\tau_k g$ [25]. In order to improve the convergence rate of (2) to a minimum point $x^* \in \text{argmin } F$, one can consider the following inertial variant of the FB scheme [3, 26]:

$$x^{(k+1)} = \text{prox}_{\tau_{k+1} g}(y^{(k+1)} - \tau_{k+1} \nabla f(y^{(k+1)})), \quad \text{where } y^{(k+1)} = x^{(k)} + \beta_{k+1}(x^{(k)} - x^{(k-1)}). \tag{3}$$

Here, $y^{(k+1)}$ denotes the inertial point obtained by extrapolation of the sequence $\{x^{(k)}\}_{k\in\mathbb{N}}$ and β_{k+1} is the inertial parameter. The idea of introducing inertia in (2) traces back to the pioneering work of Nesterov [26], where the author proposes an inertial gradient method suited for the case $g \equiv 0$ with optimal convergence rate $\mathcal{O}(1/k^2)$, that is one order higher than the standard $\mathcal{O}(1/k)$ rate available for the method (2). This approach has been extended by Nesterov in [27] to projected gradient descent algorithms and, later, to a general non-smooth convex function g in [3], under the popular name of "Fast Iterative Soft-Thresholding Algorithm" (FISTA). In spite of its optimal convergence rate, the practical acceleration of FISTA is often limited by the selection of the algorithmic step-sizes τ_k, which, typically, are computed either as the inverse of a (pessimistic) estimate of the Lipschitz constant L_f of the gradient of f, or through a monotone Armijo-type backtracking procedure when such an estimate is not available [3, Section 4]. Both choices usually generate a sequence of admissible step-sizes which, in case they are too small, may lead to a severe slowdown in the numerical performance of the algorithm. Possible remedies to this issue consist in either adopting an adaptive (non-monotone) backtracking procedure allowing for the local increasing of the step-size [28, 32], or computing the proximal–gradient operator in (3) with respect to the norm induced by a suitably-chosen variable linear operator, which depends on some second-order information of the function f at the current inertial point. Such variable metric strategy has been popularized in recent years and thoroughly studied in a series of works, such as, e.g., [5, 7, 8, 15].

Furthermore, if one or both functions f, g in (1) are strongly convex, then it is possible to incorporate the strong convexity moduli inside the definition of the inertial parameters β_{k+1} in order to obtain linear convergence rates. In the case of smooth optimization problems, this feature is well-known. In [27, Section 2.2.1], under the assumption that f is μ_f-strongly convex, ∇f is L_f-Lipschitz continuous and $g \equiv 0$, Nesterov considered the scheme (3) with $\tau_{k+1} \equiv 1/L_f$ and $\beta_{k+1} \equiv (\sqrt{L_f} - \sqrt{\mu_f})/(\sqrt{L_f} + \sqrt{\mu_f})$, proving that [27, Theorem 2.2.3]

$$f(x^{(k)}) - f(x^*) = \mathcal{O}\left(\left(1 - \sqrt{\frac{\mu_f}{L_f}}\right)^k\right).$$

A first extension of Nesterov's algorithm to a general non-smooth g is studied in [33], where strong convexity is assumed only on the smooth component f. In [13, Algorithm 5], the authors consider a variant of this algorithm for problems where either one or both terms f, g are strongly convex. The same algorithm is denominated Generalised FISTA (GFISTA) in [11] and equipped with an adaptive backtracking strategy allowing for local increasing of the step-size, similarly to the one proposed in [32]. In [19, 20] similar algorithms are considered and thoroughly analyzed by means of generalized estimate sequences arguments. However, the convergence analysis carried out in [11, 13, 19, 20] does not treat explicitly the case where the proximal operator of g is computed inexactly due to the lack of an exact analytical formula. This happens frequently in signal and image processing problems, especially in the presence of sparsity-inducing priors [1]. The practical performance of GFISTA may be in fact severely affected by inexact proximal evaluations, since inertial FB variants are usually more sensitive to computational errors than the standard FB scheme, see, e.g. [33, 34]. In this context, the adoption of a variable metric strategy in (3) is expected to compensate for the proximal errors, thus favoring the recovery of the expected linear rate of GFISTA, as observed also for standard FISTA [8].

Contribution and structure of the paper. In this paper, we consider an extension of the GFISTA algorithm [11, 13] suited for the minimization of composite (possibly strongly) convex functionals, which we name SAGE-FISTA (Scaled Adaptive GEneralized FISTA). The proposed algorithm computes the FB iterates inexactly with increasing accuracy and with respect to a sequence of suitably-chosen variable metric operators. In order to limit the computational slowdowns due to the use of possibly small step-sizes generated by Armijo-type monotone backtracking strategies, we further consider an adaptive backtracking procedure allowing for local adjustment of such parameter [11, 32]. Under the assumption that one or both functions f, g in (1) are strongly convex, we prove the linear convergence of the sequence of function values generated by SAGE-FISTA. Furthermore, we assess the numerical performance of SAGE-FISTA on some image restoration problems, showing that the combination of the scaling technique with the adaptive backtracking strategy significantly reduces the computational times and avoids misspecifications of the algorithmic step-size.

The paper is organized as follows. In Section 2 we set the notation and report some auxiliary technical results. In Section 3 we describe SAGE-FISTA in detail and provide a rigorous study of its convergence properties, showing, in particular, that linear convergence is achieved in strongly convex scenarios. Furthermore we report some useful results for the practical definition of both the inexactness parameter sequence and the variable metric operators. Finally, in Section 4, we apply SAGE-FISTA to exemplar image restoration problems such as Poisson image denoising and deblurring.

2 Problem setting and preliminaries

In this section, we fix the notation, formulate the general optimization problem we aim to solve and prove some important preliminary results for the convergence analysis that follows.

2.1 Notations

For a given Hilbert space $\mathcal{H}$, we denote by $\|\cdot\|$ the norm induced by the inner product $\langle \cdot, \cdot \rangle$ defined on $\mathcal{H}$. For any function $f : \mathcal{H} \to \mathbb{R} \cup \{\infty\}$, we denote the domain of f by the set $\mathrm{dom}(f) := \{x \in \mathcal{H} : f(x) < \infty\}$. If $f : \mathcal{H} \to \mathbb{R}$ is a continuously differentiable function, then by $\mathbb{D}_f(x, y) = f(x) - f(y) - \langle \nabla f(y), x - y \rangle$ we denote the Bregman distance of f between the two points $x, y \in \mathcal{H}$.

Let $\mathcal{S}(\mathcal{H})$ be the set of linear, bounded and self-adjoint operators from $\mathcal{H}$ to $\mathcal{H}$, and $\mathcal{I} \in \mathcal{S}(\mathcal{H})$ the identity operator on $\mathcal{H}$. The Loewner partial ordering relation on $\mathcal{S}(\mathcal{H})$ is defined as follows:

$$\forall D_1, D_2 \in \mathcal{S}(\mathcal{H}) \quad D_1 \preceq D_2 \Leftrightarrow \langle D_1 x, x \rangle \leq \langle D_2 x, x \rangle \; \forall x \in \mathcal{H}.$$

For any η_{inf}, η_{sup} with $0 < \eta_{inf} \leq \eta_{sup}$, we introduce the following sets

$$\mathcal{D}_{\eta_{inf}} := \{D \in \mathcal{S}(\mathcal{H}) : \eta_{inf}\mathcal{I} \preceq D\}, \quad \mathcal{D}_{\eta_{inf}}^{\eta_{sup}} := \{D \in \mathcal{S}(\mathcal{H}) : \eta_{inf}\mathcal{I} \preceq D \preceq \eta_{sup}\mathcal{I}\}.$$

By definition, it follows that $\mathcal{D}_{\eta_{inf}}^{\eta_{sup}} \subseteq \mathcal{D}_{\eta_{inf}}$. Furthermore, following [17, Theorem 4.6.11], we have that if $D \in \mathcal{D}_{\eta_{inf}}$, then D is invertible and, if also $D \in \mathcal{D}_{\eta_{inf}}^{\eta_{sup}}$, then we have $D^{-1} \in \mathcal{D}_{1/\eta_{sup}}^{1/\eta_{inf}}$. If $D \in \mathcal{D}_{\eta_{inf}}$, then

$$(x, y) := \langle Dx, y \rangle, \tag{4}$$

defines an inner product on $\mathcal{H}$, and $\|x\|_D = \sqrt{\langle Dx, x \rangle}$ denotes the $D-$norm induced by (4). Hence, if $D \in \mathcal{D}_{\eta_{inf}}^{\eta_{sup}}$, the following inequality holds

$$\eta_{inf}\|x\|^2 \leq \|x\|_D^2 \leq \eta_{sup}\|x\|^2, \quad \forall\, x \in \mathcal{H}. \tag{5}$$

Given $Y \subseteq \mathcal{H}$ nonempty, closed convex set and $D \in \mathcal{D}_{\eta_{inf}}$, we define the projection operator onto Y in the metric induced by D as

$$P_{Y,D}(x) = \underset{z \in Y}{\operatorname{argmin}}\, \|z - x\|_D^2, \quad \forall\, x \in \mathcal{H}.$$

2.2 Problem formulation and preliminary results

We are interested in solving the optimization problem

$$\min_{x \in \mathcal{H}} F(x) \equiv f(x) + g(x), \tag{6}$$

where

- $f : \mathcal{H} \to \mathbb{R}$ is μ_f-strongly convex ($\mu_f \geq 0$) and continuously differentiable with L_f-Lipschitz continuous gradient on a nonempty, closed convex set Y such that $\operatorname{dom}(g) \subseteq Y \subseteq \operatorname{dom}(f)$;

- $g : \mathcal{H} \to \mathbb{R} \cup \{\infty\}$ is proper, μ_g-strongly convex ($\mu_g \geq 0$), and lower semicontinuous.

We now prove a technical descent inequality that is crucial for the convergence analysis carried out in Section 3. We start noting that, under the assumption that f is μ_f-strongly convex, then f is also strongly convex with respect to the norm induced by any operator $D \in \mathcal{D}_{\eta_{inf}}^{\eta_{sup}}$.

Lemma 2.1. *Let $f : \mathcal{H} \to \mathbb{R} \cup \{\infty\}$ be a μ_f-strongly convex function on a nonempty convex set $Y \subseteq \operatorname{dom}(f)$. For all $D \in \mathcal{D}_{\eta_{inf}}^{\eta_{sup}}$ and $\mu_{f,D} \leq \frac{\mu_f}{\eta_{sup}}$, we have that f is $\mu_{f,D}-$strongly convex with respect to the $D-$norm, that is*

$$f(y) \geq f(x) + \langle w, y - x \rangle + \frac{\mu_{f,D}}{2}\|y - x\|_D^2, \quad \forall\, y, x \in Y,\ w \in \partial f(x). \tag{7}$$

Proof. In view of the assumption on $\mu_{f,D}$, the μ_f-strong convexity of f and (5), we have

$$\begin{aligned} f(y) &\geq f(x) + \langle w, y - x \rangle + \frac{\mu_f}{2\eta_{sup}}\eta_{sup}\|y - x\|^2 \\ &\geq f(x) + \langle w, y - x \rangle + \frac{\mu_{f,D}}{2}\|y - x\|_D^2, \quad \forall\, y, x \in Y,\ w \in \partial f(x). \end{aligned}$$

$\square$

Furthermore we report the statement of a scaled version of the classical descent lemma whose proof can be found in [7].

Lemma 2.2. *[7, Lemma 6] Let $f : \mathcal{H} \to \mathbb{R} \cup \{\infty\}$ be a continuously differentiable function with L_f-Lipschitz continuous gradient on $Y \subseteq \operatorname{dom}(f)$ and $D \in \mathcal{D}_{\eta_{inf}}$. If $0 < \tau \leq \frac{\eta_{inf}}{L_f}$, we have*

$$f(y) \leq f(x) + <\nabla f(x), y - x> + \frac{1}{2\tau}\|x - y\|_D^2 \quad \forall\, x, y \in Y. \tag{8}$$

4

We now introduce the notion of proximal–gradient point, which can be computed by alternating a (forward) gradient step performed w.r.t. the differentiable function f with a (backward) proximal step on the non-smooth term g. To this aim, for any given $\bar{x} \in Y$, $\tau > 0$ and $D \in \mathcal{D}_{\eta_{inf}}^{\eta_{sup}}$, we define the function $h_{\tau,D}(\cdot; \bar{x}) : \mathcal{H} \to \mathbb{R} \cup \{\infty\}$ as

$$h_{\tau,D}(z; \bar{x}) := f(\bar{x}) + \langle \nabla f(\bar{x}), z - \bar{x} \rangle + \frac{1}{2\tau}\|z - \bar{x}\|_D^2 + g(z), \quad \forall\, z \in \mathcal{H}. \tag{9}$$

Since g is strongly convex w.r.t. the D-norm with modulus $\mu_{g,D} \leq \mu_g/\eta_{sup}$ (see Lemma 2.1), it is easy to show that $h_{\tau,D}(\cdot; \bar{x})$ is also strongly convex w.r.t. the D-norm with modulus $\frac{1}{\tau} + \mu_{g,D}$. Consequently, $h_{\tau,D}(\cdot; \bar{x})$ has a unique minimizer.

Definition 2.1. *Given $\tau > 0$ and $D \in \mathcal{D}_{\eta_{inf}}^{\eta_{sup}}$, we define the proximal operator associated to g in the metric induced by D as*

$$\mathrm{prox}_g^D(x) = \operatorname*{argmin}_{z \in \mathcal{H}}\, g(z) + \frac{1}{2}\|z - x\|_D^2, \quad \forall\, x \in \mathcal{H}.$$

For any $\bar{x} \in Y$, the proximal–gradient point $\hat{x}$ with parameters τ and D is thus given by

$$\hat{x} := \mathrm{prox}_{\tau g}^D(\bar{x} - \tau D^{-1}\nabla f(\bar{x})) = \operatorname*{argmin}_{z \in \mathcal{H}}\, h_{\tau,D}(z; \bar{x}). \tag{10}$$

For our purposes, it will be convenient to consider a suitable approximation of the point $\hat{x}$ in (10) computed by means of a fixed positive tolerance parameter.

Definition 2.2. *Given $\bar{x} \in Y$, $\tau > 0$, $D \in \mathcal{D}_{\eta_{inf}}^{\eta_{sup}}$ and $\epsilon \geq 0$, a point $\tilde{x} \in dom(g)$ is called an $\epsilon-$approximation of the proximal–gradient point $\hat{x}$ if the following condition holds*

$$h_{\tau,D}(\tilde{x}; \bar{x}) - h_{\tau,D}(\hat{x}; \bar{x}) \leq \epsilon. \tag{11}$$

In this case, we use the notation $\tilde{x} \approx_\epsilon \hat{x}$.

We are now ready to state and prove the promised generalized descent inequality, adapted to the scaled and inexact framework introduced above.

Lemma 2.3. *Given a point $\bar{x} \in Y$, $\tau > 0$, $D \in \mathcal{D}_{\eta_{inf}}^{\eta_{sup}}$, $\epsilon \geq 0$ and a point $\tilde{x} \approx_\epsilon \hat{x}$, the following inequality holds*

$$F(\tilde{x}) + (1 + \tau\mu_{g,D})\frac{\|x - \tilde{x}\|_D^2}{2\tau} + \left(\frac{\|\tilde{x} - \bar{x}\|_D^2}{2\tau} - \mathbb{D}_f(\tilde{x}, \bar{x})\right)$$

$$\leq F(x) + (1 - \tau\mu_{f,D})\frac{\|x - \bar{x}\|_D^2}{2\tau} + \epsilon + \frac{\sqrt{2\epsilon\tau(1 + \tau\mu_{g,D})}}{\tau}\|x - \tilde{x}\|_D, \quad \forall\, x \in \mathcal{H}, \tag{12}$$

where $\mu_{g,D} \leq \frac{\mu_g}{\eta_{sup}}$ and $\mu_{f,D} \leq \frac{\mu_f}{\eta_{sup}}$.

Proof. Let $\mu_{h,D} = \frac{1}{\tau} + \mu_{g,D}$ be the strong convexity modulus of $h_{\tau,D}(\cdot; \bar{x})$ with respect to the $D-$norm. Combining the strong convexity of $h_{\tau,D}(\cdot; \bar{x})$ with the optimality of the proximal–gradient point $\hat{x}$, we obtain

$$h_{\tau,D}(x; \bar{x}) \geq h_{\tau,D}(\hat{x}; \bar{x}) + \frac{\mu_{h,D}}{2}\|x - \hat{x}\|_D^2, \quad \forall x \in \mathcal{H} \tag{13}$$

or, equivalently,

$$\|x - \hat{x}\|_D^2 \leq \frac{2}{\mu_{h,D}}(h_{\tau,D}(x; \bar{x}) - h_{\tau,D}(\hat{x}; \bar{x})), \quad \forall x \in \mathcal{H}.$$

Choosing $x = \tilde{x}$ above and combining it with (11) leads to

$$\|\tilde{x} - \hat{x}\|_D^2 \leq \frac{2}{\mu_{h,D}}\epsilon. \tag{14}$$

Starting again from (13), we can thus write the following chain of inequalities

$$
\begin{aligned}
h_{\tau,D}(x; \bar{x}) &\geq h_{\tau,D}(\hat{x}; \bar{x}) + \frac{\mu_{h,D}}{2}\|x - \hat{x}\|_D^2 \\
&= h_{\tau,D}(\tilde{x}; \bar{x}) + h_{\tau,D}(\hat{x}; \bar{x}) - h_{\tau,D}(\tilde{x}; \bar{x}) + \frac{\mu_{h,D}}{2}\|x - \hat{x}\|_D^2 \\
&\geq h_{\tau,D}(\tilde{x}; \bar{x}) - \epsilon + \frac{\mu_{h,D}}{2}\|x - \tilde{x}\|_D^2 + \frac{\mu_{h,D}}{2}\|\tilde{x} - \hat{x}\|_D^2 + \mu_{h,D}\langle D(x - \tilde{x}), \tilde{x} - \hat{x}\rangle \\
&\geq h_{\tau,D}(\tilde{x}; \bar{x}) + \frac{\mu_{h,D}}{2}\|x - \tilde{x}\|_D^2 - \epsilon - \mu_{h,D}\sqrt{\frac{2\epsilon}{\mu_{h,D}}}\|x - \tilde{x}\|_D, \quad \forall x \in \mathcal{H}
\end{aligned}
$$

where the third inequality follows from the definition of $\epsilon-$approximation (11) and the three-point-equality $\|a - c\|_D^2 = \|a - b\|_D^2 + \|b - c\|_D^2 + 2\langle D(a - b), b - c\rangle$ and the fourth one from the Cauchy-Schwarz inequality, combined with relation (14) and the non-negativity of the quantity $\|\tilde{x} - \hat{x}\|_D^2$. Hence, we deduce

$$h_{\tau,D}(x; \bar{x}) \geq h_{\tau,D}(\tilde{x}; \bar{x}) + \frac{\mu_{h,D}}{2}\|\tilde{x} - x\|^2 - \epsilon - \mu_{h,D}\sqrt{\frac{2\epsilon}{\mu_{h,D}}}\|x - \tilde{x}\|_D, \quad \forall x \in \mathcal{H}. \tag{15}$$

Now, for all $x \in \mathcal{H}$, we have

$$
\begin{aligned}
F(x) + (1 - \tau\mu_{f,D})\frac{\|x - \bar{x}\|_D^2}{2\tau} &\geq g(x) + f(\bar{x}) + \langle \nabla f(\bar{x}), x - \bar{x}\rangle + \frac{\|x - \bar{x}\|_D^2}{2\tau} = h_{\tau,D}(x; \bar{x}) \\
&\overset{(15)}{\geq} h_{\tau,D}(\tilde{x}; \bar{x}) + \frac{\mu_{h,D}}{2}\|\tilde{x} - x\|^2 - \epsilon - \mu_{h,D}\sqrt{\frac{2\epsilon}{\mu_{h,D}}}\|x - \tilde{x}\|_D \\
&= g(\tilde{x}) + f(\bar{x}) + \langle \nabla f(\bar{x}), \tilde{x} - \bar{x}\rangle + \frac{\|\tilde{x} - \bar{x}\|_D^2}{2\tau} + \frac{\tau\mu_{g,D} + 1}{2\tau}\|\tilde{x} - x\|_D^2 - \epsilon - \mu_{h,D}\sqrt{\frac{2\epsilon}{\mu_{h,D}}}\|x - \tilde{x}\|_D \\
&= F(\tilde{x}) + \frac{\|\tilde{x} - \bar{x}\|_D^2}{2\tau} - \mathbb{D}_f(\tilde{x}, \bar{x}) + \frac{\tau\mu_{g,D} + 1}{2\tau}\|\tilde{x} - x\|_D^2 - \epsilon - \sqrt{2\epsilon\left(\frac{1}{\tau} + \mu_{g,D}\right)}\|x - \tilde{x}\|_D,
\end{aligned}
$$

which concludes the proof. $\qquad\square$

3 SAGE-FISTA: description and convergence analysis

In the following, we describe a generalized version of the popular "Fast Iterative Soft-Thresholding Algorithm" (FISTA) proposed in [3], which is suited for solving the (possibly strongly) convex problem (6) by means of an appropriate scaled and inexact inertial forward-backward splitting.

6

We name our algorithm SAGE-FISTA (Scaled Adaptive GEneralized FISTA); it is obtained by incorporating in the iterative scheme (3) the following features:

- the use of a variable non-Euclidean metric in the computation of the proximal–gradient point, which is induced along the iterations by the elements of a sequence of linear, bounded and self-adjoint positive operators $\{D_k\}_{k\in\mathbb{N}}$; such operators are typically chosen in order to capture some second order information of the differentiable part f at the current iterate $x^{(k)}$ (see e.g. [6, 8, 9, 23]);

- the approximate computation of the proximal-gradient point according to the notion of ϵ -approximation given in Definition 2.2;

- a non-monotone backtracking strategy similar to the one considered in [11], which allows for possible increasing of the step-size τ_{k+1} at each iteration; differently from monotone Armijo-type approaches, this strategy is particularly helpful when the initial τ_0 is chosen to be extremely small, which corresponds to a (pessimistic) large estimate L_0 of the Lipschitz constant value L_f;

- a novel selection rule for the inertial parameter β_{k+1}, which depends on both the strong convexity moduli μ_f, μ_g and the bounds on the spectrum of the operators D_{k+1}.

The proposed SAGE-FISTA is fully reported in Algorithm 1.

Algorithm 1 Scaled Adaptive GEneralized FISTA (SAGE-FISTA)

Parameters: $\rho \in (0,1)$, $\delta \in (0,1]$, $\{\eta_{sup}^k\}_{k\in\mathbb{N}}$ s.t. $0 < \eta_{inf} \leq \eta_{sup}^k \leq \eta_{sup}$, $\{\mu_{f,k}\}_{k\in\mathbb{N}}$, $\{\mu_{g,k}\}_{k\in\mathbb{N}}$, $\{\mu_k\}_{k\in\mathbb{N}}$ s.t. $\mu_{f,k} = \frac{\mu_f}{\eta_{sup}^k}$, $\mu_{g,k} = \frac{\mu_g}{\eta_{sup}^k}$, $\mu_k = \mu_{f,k} + \mu_{g,k}$.

Initialization: $\tau_0 \in \mathbb{R}$ s.t. $0 < \tau_0\mu_{f,0} < 1$, $q_0 = \frac{\tau_0\mu_0}{1+\tau_0\mu_{g,0}}$, $x^{(-1)} \in \mathcal{H}$, $x^{(0)} = x^{(-1)}$, and $t_0 \in \mathbb{R}$ s.t. $1 \leq t_0 \leq \frac{1}{\sqrt{q_0}}$.

FOR $k = 0, 1, 2, \ldots$

    Choose $D_{k+1} \in \mathcal{D}_{\eta_{inf}}^{\eta_{sup}^k}$ and set

$$\tau_{k+1}^0 = \frac{\tau_k}{\delta}.$$

    FOR $i = 0, 1, \ldots$ REPEAT

    STEP 1. Set $\tau_{k+1} = \rho^i \tau_{k+1}^0$;

    STEP 2. Set $q_{k+1} = \frac{\tau_{k+1}\mu_{k+1}}{1+\tau_{k+1}\mu_{g,k+1}}$;

    STEP 3. Set $t_{k+1} = \frac{1 - q_k t_k^2 + \sqrt{(1-q_k t_k^2)^2 + 4\frac{q_k}{q_{k+1}}t_k^2}}{2}$.

    STEP 4. Set $\beta_{k+1} = \left(\frac{t_k - 1}{t_{k+1}}\right)\frac{1+\tau_{k+1}\mu_{g,k+1} - t_{k+1}\tau_{k+1}\mu_{k+1}}{1-\tau_{k+1}\mu_{f,k+1}}$.

    STEP 5. Set $y^{(k+1)} = P_{Y,D_{k+1}}(x^{(k)} + \beta_{k+1}(x^{(k)} - x^{(k-1)}))$.

    STEP 6. Set $\hat{x}^{(k+1)} = \text{prox}_{\tau_{k+1}g}^{D_{k+1}}(y^{(k+1)} - \tau_{k+1}D_{k+1}^{-1}\nabla f(y^{(k+1)}))$, choose $\epsilon_{k+1} \geq 0$ and compute $x^{(k+1)} \in \text{dom}(g)$ such that

$$x^{(k+1)} \approx_{\epsilon_{k+1}} \hat{x}^{(k+1)}.$$

    UNTIL $\mathbb{D}_f(x^{(k+1)}, y^{(k+1)}) < \frac{\|x^{(k+1)} - y^{(k+1)}\|_{D_{k+1}}^2}{2\tau_{k+1}}$.

END

At each iteration $k \geq 0$, the preliminary step of SAGE-FISTA is the choice of both the linear, bounded and self-adjoint operator $D_{k+1} \in \mathcal{D}_{\eta_{inf}}^{\eta_{sup}^k}$ and the tentative step-size $\tau_{k+1}^0 = \tau_k/\delta$. If $\delta < 1$, then an adaptive increasing of τ_k is allowed (see [11, 20] for analogous strategies applied to GFISTA and [32] for FISTA), while if $\delta = 1$, a classical Armijo-type backtracking is performed (as in [3] for FISTA). Then, an inner backtracking procedure computing the next iterate $x^{(k+1)}$ follows. At each inner backtracking iteration $i \geq 0$, the iterates q_{k+1}, t_{k+1} are first updated (STEPS 2–3); such quantities depend not only on the scaled quantities $\mu_{f,k+1} = \mu_f/\eta_{sup}^k$, $\mu_{g,k+1} = \mu_g/\eta_{sup}^k$ and $\mu_{k+1} = \mu_{f,k+1} + \mu_{g,k+1}$, but also on the tentative step-size τ_{k+1}. Then, the inertial parameter β_{k+1} and the projected inertial point $y^{(k+1)}$ are computed (STEPS 4–5). Finally, an approximated point $x^{(k+1)}$ such that $x^{(k+1)} \approx_{\epsilon_{k+1}} \text{prox}_{\tau_{k+1}g}^{D_{k+1}}(y^{(k+1)} - \tau_{k+1}D_{k+1}^{-1}\nabla f(y^{(k+1)}))$ is computed. For each backtracking iteration, a check on the condition

$$\mathbb{D}_f(x^{(k+1)}, y^{(k+1)}) < \frac{\|x^{(k+1)} - y^{(k+1)}\|_{D_{k+1}}^2}{2\tau_{k+1}} \tag{16}$$

is performed (STEP 6). If such condition is not satisfied, then the step-size is reduced by a factor ρ (STEP 1) and STEPS 2–6 are performed until the condition is satisfied. Note that this procedure ends in a finite number of steps, thanks to Lemma 2.2.

When $\eta_{inf}^k = \eta_{sup}^k = 1$, $D_k \equiv \mathcal{I}$ and $Y = \mathcal{H}$, Algorithm 1 reduces to the Generalised FISTA (GFISTA) algorithm studied in [11]. The presence of the strong convexity moduli μ_f, μ_g in the update rule of the inertial parameter β_{k+1} allows to prove a linear convergence rate for the function values also when an adaptive backtracking strategy is adopted, see [11, Theorem 4.6].

When $D_k \neq \mathcal{I}$, Algorithm 1 can be considered as a scaled version of GFISTA where the inexact computation of the proximal-gradient point is explicitly taken into account in STEP 6. Note that the introduction of the positive operator D_k further imposes a modification of the parameters q_{k+1} and t_{k+1} in STEP 2 and STEP 3, respectively. However, linear convergence rates still hold, as we will rigorously show in Section 3.

Finally, note that when $\mu = \mu_f + \mu_g = 0$ and $\delta = 1$, STEP 3 reduces to

$$t_{k+1} = \frac{1 + \sqrt{1 + 4\frac{\tau_k}{\tau_{k+1}} t_k^2}}{2},$$

and Algorithm 1 becomes the inexact scaled forward-backward extrapolation method equipped with Armijo-type backtracking proposed in [8]. In this case, standard $\mathcal{O}(1/k^2)$ convergence rates can be proved, coherently with the result obtained in [8, Theorem 3.1].

Remark 3.1. *We observe that*

$$0 \leq q_k < 1, \quad t_k \geq 1, \quad \forall\, k \geq 1.$$

Indeed q_k is nonnegative due to the nonnegativity of μ_k, μ_g, τ_k, and combining the $\mu_{f,k}-$strong convexity of f with respect to the D_k-norm with (16), we have

$$\frac{\mu_{f,k}}{2}\|x^{(k)} - y^{(k)}\|_{D_k}^2 \leq \mathbb{D}_f(x^{(k)}, y^{(k)}) < \frac{\|x^{(k)} - y^{(k)}\|_{D_k}^2}{2\tau_k},$$

which implies $\mu_{f,k}\tau_k < 1$ and thus $q_k < 1$. Furthermore, proceeding as in [11, Lemma 4.4], we have

$$t_k \geq \frac{1 - q_{k-1}t_{k-1}^2 + \sqrt{\left(1 - q_{k-1}t_{k-1}^2\right)^2 + 4q_{k-1}t_{k-1}^2}}{2} = \frac{1 - q_{k-1}t_{k-1}^2 + \sqrt{\left(1 + q_{k-1}t_{k-1}^2\right)^2}}{2} = 1.$$

Remark 3.2. *When $\mu = \mu_f + \mu_g > 0$, the update rule for the sequence $\{t_k\}_k$ is obtained by imposing the following condition at each iteration:*

$$\tau_{k+1}' t_{k+1}(t_{k+1} - 1) = \frac{\mu_k}{\mu_{k+1}}\omega_{k+1}\tau_k' t_k^2, \quad \forall\, k \geq 0 \tag{17}$$

where we have set:

$$\tau_k' := \frac{\tau_k}{1 + \tau_k\mu_{g,k}}, \quad \omega_k := 1 - t_k q_k = \frac{1 + \tau_k\mu_{g,k} - \tau_k t_k \mu_k}{1 + \tau_k\mu_{g,k}}. \tag{18}$$

Indeed, note that t_{k+1} is defined as the solution of the following second-degree equation

$$t^2 - (1 - q_k t_k^2)t - \frac{q_k}{q_{k+1}}t_k^2 = 0.$$

9

By now observing that $q_k/q_{k+1} = (\mu_k \tau_k')/(\mu_{k+1} \tau_{k+1}')$, we can rewrite the equation above as

$$t^2 - \left(1 - \frac{\mu_k \tau_k' t_k^2}{\mu_{k+1} \tau_{k+1}'} q_{k+1}\right) t - \frac{\mu_k \tau_k' t_k^2}{\mu_{k+1} \tau_{k+1}'} = 0.$$

Then, choosing $t = t_{k+1}$, multiplying by τ_{k+1}' and rearranging terms leads to (17). When $\mu = 0$, equality (17) holds by neglecting the quantity μ_k/μ_{k+1}.

We conclude this section with the following technical result holding for the sequences $\{q_k t_k^2\}_{k \in \mathbb{N}}$ and $\{q_k t_k\}_{k \in \mathbb{N}}$.

Lemma 3.1. *If $0 \leq t_0 \leq \frac{1}{\sqrt{q_0}}$ (as assumed in Algorithm 1), then it follows that*

$$q_k t_k^2 \leq 1 \quad and \quad q_k t_k < 1, \quad \forall\, k \geq 0. \tag{19}$$

Proof. If $\mu = \mu_f + \mu_g = 0$, the thesis follows trivially. Then, we assume $\mu > 0$ and proceed by induction as in [11, Lemma 4.5]. For $k = 0$, the assumption on t_0 reads $q_0 t_0^2 \leq 1$, which entails $q_0 t_0 \leq \sqrt{q_0} < 1$, so the thesis holds. Then, let us assume that (19) holds for some $k \geq 0$. Recalling that t_{k+1} is the solution of (17), we have:

$$q_{k+1} t_{k+1}^2 = q_{k+1} t_{k+1} + q_{k+1} \omega_{k+1} \frac{\mu_k \tau_k'}{\mu_{k+1} \tau_{k+1}'} t_k^2 = q_{k+1} t_{k+1} + \omega_{k+1} q_k t_k^2 = 1 + \omega_{k+1}(q_k t_k^2 - 1) \leq 1,$$

which follows by simply applying the induction assumption. Thus, we get $q_{k+1} t_{k+1}^2 \leq 1$. Notice that the same chain of equalities also implies:

$$q_{k+1} t_{k+1}^2 = q_{k+1} t_{k+1} + (1 - q_{k+1} t_{k+1}) t_k^2 q_k.$$

By contradiction, if $q_{k+1} t_{k+1} \geq 1$, then the previous inequality would imply

$$q_{k+1} t_{k+1}^2 \leq q_{k+1} t_{k+1} \quad \Leftrightarrow \quad q_{k+1} t_{k+1} \leq q_{k+1} < 1,$$

which is absurd. Thus, we deduce $q_{k+1} t_{k+1} < 1$, as required. $\qquad\square$

3.1 Convergence rates

We now show that SAGE-FISTA enjoys a convergence rate in function values that is either $R-$linear or proportional to $1/k^2$, depending on whether the strong convexity modulus $\mu = \mu_f + \mu_g$ is positive or null. A couple of technical results holding for nonnegative sequences are needed.

Lemma 3.2. *[33, Lemma 1] Let $\{p_k\}_{k \in \mathbb{N}}$, $\{q_k\}_{k \in \mathbb{N}}$, $\{\lambda_k\}_{k \in \mathbb{N}}$ be sequences of real nonnegative numbers, with $\{q_k\}_{k \in \mathbb{N}}$ being a monotone nondecreasing sequence, such that the following recursive property is satisfied:*

$$p_k^2 \leq q_k + \sum_{i=1}^{k} \lambda_i p_i, \quad \forall\, k \geq 1. \tag{20}$$

Then we have

$$p_k \leq \frac{1}{2} \sum_{i=1}^{k} \lambda_i + \left(q_k + \left(\frac{1}{2}\sum_{i=1}^{k}\lambda_i\right)^2\right)^{\frac{1}{2}}, \quad \forall\, k \geq 1. \tag{21}$$

10

Lemma 3.3. *[30] Let $\{p_k\}_{k\in\mathbb{N}}$, $\{\zeta_k\}_{k\in\mathbb{N}}$, $\{\xi_k\}_{k\in\mathbb{N}}$ be sequences of real nonnegative numbers such that $\sum_{k=0}^{\infty}\zeta_k < \infty$, $\sum_{k=0}^{\infty}\xi_k < \infty$ and the following inequality is satisfied:*

$$p_{k+1} \leq (1+\zeta_k)p_k + \xi_k, \quad \forall\, k \geq 0.$$

Then $\{p_k\}_{k\in\mathbb{N}}$ is a convergent sequence.

Furthermore, we make the following assumption on the sequence of operators $\{D_k\}_{k\in\mathbb{N}}$ and the sequence of upper bounding values $\{\eta_{sup}^k\}_{k\in\mathbb{N}}$.

Assumption 1. *There exists a sequence of real nonnegative numbers $\{\gamma_k\}_{k\in\mathbb{N}}$ s.t. $\sum_{k=0}^{\infty}\gamma_k < \infty$ and, for all $k \geq 0$, the following conditions hold*

$$\begin{align}
D_{k+1} &\preceq (1+\gamma_{k+1})D_k, \tag{22}\\
\frac{\eta_{sup}^{k+1}}{\eta_{sup}^k} &\leq 1 + \gamma_{k+1}. \tag{23}
\end{align}$$

Remark 3.3. *Condition (22) has been often used to prove the convergence of variable metric FB methods in the convex setting, see for instance [5, 7, 8, 15]. It is easy to see that (22) implies (23) for the particular choice $\eta_{sup}^k = \|D_k^{\frac{1}{2}}\|^2$, with $D_k^{\frac{1}{2}}$ being the square root operator applied to D_k; when $\|\cdot\| = \|\cdot\|_2$, this amounts to choosing $\eta_{sup}^k = \|D_k\|_2$, i.e. the spectral radius of D_k. However, in general, we need to impose condition (23) explicitly. Let us mention two particular examples of sequences $\{D_k\}_{k\in\mathbb{N}}$ and $\{\eta_{sup}^k\}_{k\in\mathbb{N}}$ satisfying Assumption 1, which we will use in the following numerical experiments.*

- *Suppose $D_k \equiv D$ and $\eta_{sup}^k \equiv \eta > 0$, with $D \in \mathcal{D}_{\eta_{inf}}^{\eta}$. Then Assumption 1 holds with $\gamma_k \equiv 0$.*

- *Suppose $D_k \in \mathcal{D}_{\eta_{inf}^k}^{\eta_{sup}^k}$ for all $k \geq 0$, and the upper and lower bounds converge to the same positive value at a sufficiently fast rate, that is*

$$\eta_{inf}^k = \eta - \nu_{inf}^k, \quad \eta_{sup}^k = \eta + \nu_{sup}^k, \tag{24}$$

where $0 \leq \nu_{inf}^k < \eta$, $\nu_{sup}^k \geq 0$, and $\sum_{k=0}^{\infty}\nu_{inf}^k < \infty$, $\sum_{k=0}^{\infty}\nu_{sup}^k < \infty$. In this case, recalling (5), we can write

$$\|x\|_{D_{k+1}}^2 \leq \eta_{sup}^{k+1}\|x\|^2 = \frac{\eta_{sup}^{k+1}}{\eta_{inf}^k}\eta_{inf}^k\|x\|^2 \leq \frac{\eta_{sup}^{k+1}}{\eta_{inf}^k}\|x\|_{D_k}^2, \quad \forall\, x \in \mathcal{H}. \tag{25}$$

Thanks to (24), the factor multiplying $\|x\|_{D_k}^2$ can be rewritten as

$$\frac{\eta_{sup}^{k+1}}{\eta_{inf}^k} = 1 + \gamma_{k+1}, \quad \text{where} \quad \gamma_{k+1} = \frac{\nu_{sup}^{k+1} + \nu_{inf}^k}{\eta - \nu_{inf}^k}.$$

Since $\gamma_{k+1} \leq (\eta - \max_k \nu_{inf}^k)^{-1}(\nu_{sup}^{k+1} + \nu_{inf}^k)$ and the sequences $\{\nu_{inf}^k\}_{k\in\mathbb{N}}$, $\{\nu_{sup}^k\}_{k\in\mathbb{N}}$ are summable, we can conclude that condition (22) is satisfied. Furthermore, we also have

$$\frac{\eta_{sup}^{k+1}}{\eta_{sup}^k} = \frac{\eta_{sup}^{k+1}}{\eta_{inf}^k} \cdot \frac{\eta_{inf}^k}{\eta_{sup}^k} \leq \frac{\eta_{sup}^{k+1}}{\eta_{inf}^k} = 1 + \gamma_{k+1},$$

that is condition (23) also holds.

Note that, if $\mathcal{H} = \mathbb{R}^n$, $\|\cdot\| = \|\cdot\|_2$, and $\{D_k\}_{k \in \mathbb{N}}$ are diagonal matrices, then we can always impose condition (24) by constraining the diagonal elements of D_k in the interval $[\eta - \nu_{inf}^k, \eta + \nu_{sup}^k]$ (see Section 4 for a practical implementation). By doing so, we progressively "squeeze" the scaling matrices to a multiple of the identity matrix as the iterations increase.

We now provide a key descent inequality that holds for the iterates of SAGE-FISTA Algorithm 1, in the same spirit of the one employed in [11, p. 8]. The result follows by combining Lemma 2.3 with the backtracking condition of Algorithm 1 and Assumption 1.

Lemma 3.4. *Suppose that $F = f + g$ is $\mu-$strongly convex with $\mu \geq 0$ and Assumption 1 holds. Then, for all $k \geq 0$, we have*

$$
\tau_{k+1}' t_{k+1}^2 (F(x^{(k+1)}) - F(x^*)) + \frac{1}{2}\|x^* - x^{(k+1)} - (t_{k+1} - 1)(x^{(k+1)} - x^{(k)})\|_{D_{k+1}}^2
$$

$$
\leq \omega_{k+1}(1 + \gamma_{k+1})\left(\tau_k' t_k^2(F(x^{(k)}) - F(x^*)) + \frac{1}{2}\|x^* - x^{(k)} - (t_k - 1)(x^{(k)} - x^{(k-1)})\|_{D_k}^2\right)
$$

$$
+ \tau_{k+1}' t_{k+1}^2 \epsilon_{k+1} + t_{k+1}\sqrt{2\epsilon_{k+1}\tau_{k+1}'}\|x^* - x^{(k+1)} - (t_{k+1} - 1)(x^{(k+1)} - x^{(k)})\|_{D_{k+1}}.
$$

Proof. We apply Lemma 2.3 with the choices $\tau = \tau_{k+1}$, $\bar{x} = y^{(k+1)}$, $\tilde{x} = x^{(k+1)}$, $\epsilon = \epsilon_{k+1}$, $D = D_{k+1}$, $\mu_{f,D} = \mu_{f,k+1}$, $\mu_{g,D} = \mu_{g,k+1}$ and combine it with the backtracking condition (16) so as to obtain:

$$
F(x^{(k+1)}) + (1 + \tau_{k+1}\mu_{g,k+1})\frac{\|x - x^{(k+1)}\|_{D_{k+1}}^2}{2\tau_{k+1}} \leq F(x) + (1 - \tau_{k+1}\mu_{f,k+1})\frac{\|x - y^{(k+1)}\|_{D_{k+1}}^2}{2\tau_{k+1}}
$$

$$
+ \epsilon_{k+1} + \frac{\sqrt{2\epsilon_{k+1}\tau_{k+1}(1 + \tau_{k+1}\mu_{g,k+1})}}{\tau_{k+1}}\|x - x^{(k+1)}\|_{D_{k+1}}.
$$

Let us now define $x = \frac{(t_{k+1} - 1)x^{(k)} + x^*}{t_{k+1}}$ with $x^{(k)}$ being the $k-$th iterate of the algorithm and $x^* \in \mathcal{H}$ a solution of (6), and set $\bar{y}^{(k+1)} = x^{(k)} + \beta_{k+1}(x^{(k)} - x^{(k-1)})$. Since the operator $P_{Y, D_{k+1}}$ is firmly non-expansive [8, Remark 2.2] and $x \in Y$, we have $\|x - y^{(k+1)}\|_{D_{k+1}}^2 \leq \|x - \bar{y}^{(k+1)}\|_{D_{k+1}}^2$. Then, the previous inequality becomes

$$
F(x^{(k+1)}) + (1 + \tau_{k+1}\mu_{g,k+1})\frac{\|(t_{k+1} - 1)x^{(k)} + x^* - t_{k+1}x^{(k+1)}\|_{D_{k+1}}^2}{2\tau_{k+1}t_{k+1}^2} \leq F\left(\frac{(t_{k+1} - 1)x^{(k)} + x^*}{t_{k+1}}\right)
$$

$$
+ (1 - \tau_{k+1}\mu_{f,k+1})\frac{\|(t_{k+1} - 1)x^{(k)} + x^* - t_{k+1}\bar{y}^{(k+1)}\|_{D_{k+1}}^2}{2\tau_{k+1}t_{k+1}^2}
$$

$$
+ \epsilon_{k+1} + \frac{\sqrt{2\epsilon_{k+1}\tau_{k+1}(1 + \tau_{k+1}\mu_{g,k+1})}}{\tau_{k+1}t_{k+1}}\|(t_{k+1} - 1)x^{(k)} + x^* - t_{k+1}x^{(k+1)}\|_{D_{k+1}}. \quad (26)
$$

We now multiply both sides of (26) by t_{k+1}^2, apply the $\mu_{k+1}-$strong convexity of F with respect to the $D_{k+1}-$norm on the right-hand side (the point $\frac{(t_{k+1} - 1)x^{(k)} + x^*}{t_{k+1}}$ is a convex combination of

12

$x^{(k)}$ and x^*) and subtract the term $t_{k+1}^2 F(x^*)$ to both sides, thus obtaining

$$t_{k+1}^2(F(x^{(k+1)}) - F(x^*)) + \frac{1 + \tau_{k+1}\mu_{g,k+1}}{2\tau_{k+1}}\|(t_{k+1} - 1)x^{(k)} + x^* - t_{k+1}x^{(k+1)}\|_{D_{k+1}}^2$$

$$\leq t_{k+1}(t_{k+1} - 1)(F(x^{(k)}) - F(x^*)) - \frac{\mu_{k+1}(t_{k+1} - 1)}{2}\|x^{(k)} - x^*\|_{D_{k+1}}^2$$

$$+ (1 - \tau_{k+1}\mu_{f,k+1})\frac{\|(t_{k+1} - 1)x^{(k)} + x^* - t_{k+1}\bar{y}^{(k+1)}\|_{D_{k+1}}^2}{2\tau_{k+1}}$$

$$+ t_{k+1}^2\epsilon_{k+1} + t_{k+1}\frac{\sqrt{2\epsilon_{k+1}\tau_{k+1}(1 + \tau_{k+1}\mu_{g,k+1})}}{\tau_{k+1}}\|(t_{k+1} - 1)x^{(k)} + x^* - t_{k+1}x^{(k+1)}\|_{D_{k+1}}.$$
$$\tag{27}$$

By proceeding similarly as in [13, p. 292], we get:

$$t_{k+1}^2(F(x^{(k+1)}) - F(x^*)) + \frac{1 + \tau_{k+1}\mu_{g,k+1}}{2\tau_{k+1}}\|x^* - x^{(k+1)} - (t_{k+1} - 1)(x^{(k+1)} - x^{(k)})\|_{D_{k+1}}^2$$

$$+ \frac{t_{k+1}^2(t_{k+1} - 1)\mu_{k+1}(1 - \tau_{k+1}\mu_{f,k+1})}{2(1 + \tau_{k+1}\mu_{g,k+1} - t_{k+1}\tau_{k+1}\mu_{k+1})}\|x^{(k)} - \bar{y}^{(k+1)}\|_{D_{k+1}}^2 \leq t_{k+1}(t_{k+1} - 1)(F(x^{(k)}) - F(x^*))$$

$$+ \frac{1 + \tau_{k+1}\mu_{g,k+1} - t_{k+1}\tau_{k+1}\mu_{k+1}}{2\tau_{k+1}}\left\|x^* - x^{(k)} - \frac{t_{k+1}(1 - \tau_{k+1}\mu_{f,k+1})}{1 + \tau_{k+1}\mu_{g,k+1} - t_{k+1}\tau_{k+1}\mu_{k+1}}(\bar{y}^{(k+1)} - x^{(k)})\right\|_{D_{k+1}}^2$$

$$+ t_{k+1}^2\epsilon_{k+1} + t_{k+1}\frac{\sqrt{2\epsilon_{k+1}\tau_{k+1}(1 + \tau_{k+1}\mu_{g,k+1})}}{\tau_{k+1}}\|x^* - x^{(k+1)} - (t_{k+1} - 1)(x^{(k+1)} - x^{(k)})\|_{D_{k+1}}.$$

Recalling the definitions of ω_{k+1} and τ'_{k+1} given in (18), we observe that $0 < \omega_{k+1} \leq 1$ since $0 \leq q_{k+1}t_{k+1} < 1$ (see Lemma 3.1). Furthermore, the parameters β_{k+1} appearing in Algorithm 1 can be rewritten in terms of ω_{k+1} as

$$\beta_{k+1} = \omega_{k+1}\frac{t_k - 1}{t_{k+1}}\frac{1 + \tau_{k+1}\mu_{g,k+1}}{1 - \tau_{k+1}\mu_{f,k+1}}.\tag{28}$$

Recalling also the definition of $\bar{y}^{(k+1)}$, plugging (18) and (28) inside the previous inequality and multiplying it by τ'_{k+1}, we deduce:

$$\tau'_{k+1}t_{k+1}^2(F(x^{(k+1)}) - F(x^*)) + \frac{1}{2}\|x^* - x^{(k+1)} - (t_{k+1} - 1)(x^{(k+1)} - x^{(k)})\|_{D_{k+1}}^2$$

$$\leq \tau'_{k+1}t_{k+1}(t_{k+1} - 1)(F(x^{(k)}) - F(x^*)) + \frac{\omega_{k+1}}{2}\|x^* - x^{(k)} - (t_k - 1)(x^{(k)} - x^{(k-1)})\|_{D_{k+1}}^2$$

$$+ \tau'_{k+1}t_{k+1}^2\epsilon_{k+1} + t_{k+1}\sqrt{2\epsilon_{k+1}\tau'_{k+1}}\|x^* - x^{(k+1)} - (t_{k+1} - 1)(x^{(k+1)} - x^{(k)})\|_{D_{k+1}}.\tag{29}$$

Combining equation (17) with relation $\mu_k/\mu_{k+1} = \eta_{sup}^{k+1}/\eta_{sup}^k$, we deduce from (29) the following:

$$\tau'_{k+1}t_{k+1}^2(F(x^{(k+1)}) - F(x^*)) + \frac{1}{2}\|x^* - x^{(k+1)} - (t_{k+1} - 1)(x^{(k+1)} - x^{(k)})\|_{D_{k+1}}^2$$

$$\leq \omega_{k+1}\left(\frac{\eta_{sup}^{k+1}}{\eta_{sup}^k}\tau'_k t_k^2(F(x^{(k)}) - F(x^*)) + \frac{1}{2}\|x^* - x^{(k)} - (t_k - 1)(x^{(k)} - x^{(k-1)})\|_{D_{k+1}}^2\right)$$

$$+ \tau'_{k+1}t_{k+1}^2\epsilon_{k+1} + t_{k+1}\sqrt{2\epsilon_{k+1}\tau'_{k+1}}\|x^* - x^{(k+1)} - (t_{k+1} - 1)(x^{(k+1)} - x^{(k)})\|_{D_{k+1}}$$

whence the thesis follows by applying Assumption 1. □

Let us now define the sequence $\{\theta_k\}_{k \in \mathbb{N}}$ by

$$\theta_k := \frac{\prod_{i=0}^{k} \omega_i}{\tau'_k t_k^2}, \quad \forall \ k \geq 0. \tag{30}$$

The convergence rate for the function values of the iterates of Algorithm 1 can be computed by a careful study of the factor θ_k, which decays linearly if $\mu > 0$ and quadratically if $\mu = 0$. We prove this statement with the following.

Theorem 3.1. *Suppose $F = f + g$ is $\mu-$strongly convex with $\mu = \mu_f + \mu_g \geq 0$, Assumption 1 holds and $1 \leq t_0 \leq 1/\sqrt{q_0}$. Let x^* be a solution of (6), recall definitions (18) and (30), set $\gamma = \lim_{k \to \infty} \prod_{i=1}^{k} (1 + \gamma_i) < \infty$ and*

$$E_1^k = \sum_{i=0}^{k} \sqrt{\theta_{i+1}^{-1} \epsilon_{i+1}}, \qquad E_2^k = \sum_{i=0}^{k} \theta_{i+1}^{-1} \epsilon_{i+1}. \tag{31}$$

Then, for all $k \geq 0$, the following upper bound for the function values holds:

$$F(x^{(k+1)}) - F(x^*) \leq \gamma \theta_{k+1} \left(\sqrt{\frac{\omega_0}{2}} \|x^{(0)} - x^*\|_{D_0} + \sqrt{\tau'_0 t_0^2 \omega_0 \left(F(x^{(0)}) - F(x^*) \right)} + 2\sqrt{\gamma} E_1^k + \sqrt{E_2^k} \right)^2. \tag{32}$$

Furthermore, the factor θ_{k+1} can be bounded as follows:

$$\theta_{k+1} \leq \frac{\eta_{sup}}{\eta_{inf}} \cdot \min \left\{ \left(1 - \sqrt{\frac{\mu \rho}{L_f \eta_{sup} + \mu_g \rho}} \right)^{k+1} \left(\frac{1}{\tau_0} - \mu_{f,0} \right), \frac{4}{(k+2)^2} \sqrt{\frac{L_f \eta_{sup} - \rho \mu_f}{\rho \eta_{sup}}} \right\}. \tag{33}$$

Proof. The result is obtained by combining the proof in [11, Theorem 4.6] with some arguments employed in [33, Propositions 2-4]. By using the following short-hand notations

$$u^{(k)} := x^* - x^{(k)} - (t_k - 1)(x^{(k)} - x^{(k-1)}), \quad v_k := F(x^{(k)}) - F(x^*),$$

we have that the inequality in Lemma 3.4 can be reformulated in a compact form as

$$\tau'_{k+1} t_{k+1}^2 v_{k+1} + \frac{1}{2} \|u^{(k+1)}\|_{D_{k+1}}^2 \leq \omega_{k+1} (1 + \gamma_{k+1}) \left(\tau'_k t_k^2 v_k + \frac{1}{2} \|u^{(k)}\|_{D_k}^2 \right)$$
$$+ \tau'_{k+1} t_{k+1}^2 \epsilon_{k+1} + t_{k+1} \sqrt{2\epsilon_{k+1} \tau'_{k+1}} \|u^{(k+1)}\|_{D_{k+1}}. \tag{34}$$

By recursively applying (34), we obtain

$$\tau'_{k+1} t_{k+1}^2 v_{k+1} + \frac{1}{2} \|u^{(k+1)}\|_{D_{k+1}}^2 \leq \left(\prod_{i=1}^{k+1} \omega_i (1 + \gamma_i) \right) \left(\tau'_0 t_0^2 v_0 + \frac{1}{2} \|u^{(0)}\|_{D_0}^2 \right)$$
$$+ \sum_{i=0}^{k} \left(\prod_{j=i+2}^{k+1} \omega_j (1 + \gamma_j) \right) \left(\tau'_{i+1} t_{i+1}^2 \epsilon_{i+1} + t_{i+1} \sqrt{2\epsilon_{i+1} \tau'_{i+1}} \|u^{(i+1)}\|_{D_{i+1}} \right), \tag{35}$$

14

where we used the notation $\prod_{j=i+2}^{k+1} \omega_j(1+\gamma_j) = 1$ whenever $i > k-1$.

Let us define the sequence $\{p_k\}_{k \in \mathbb{N}}$ with elements $p_k = \prod_{i=1}^{k}(1+\gamma_i)$ for all $k \geq 1$. We notice that $p_{k+1} = (1+\gamma_{k+1})p_k$. Recalling the summability of the sequence $\{\gamma_k\}_{k \in \mathbb{N}}$ and applying Lemma 3.3, it follows that $\{p_k\}_{k \in \mathbb{N}}$ converges. Setting $\gamma = \lim_{k \to \infty} \prod_{i=1}^{k}(1+\gamma_i) < \infty$, we deduce from equation (35)

$$\tau'_{k+1}t_{k+1}^2 v_{k+1} + \frac{1}{2}\|u^{(k+1)}\|_{D_{k+1}}^2 \leq \left(\prod_{i=1}^{k+1}\omega_i\right)\left(\gamma\tau'_0 t_0^2 v_0 + \frac{\gamma}{2}\|u^{(0)}\|_{D_0}^2\right)$$
$$+ \sum_{i=0}^{k}\left(\prod_{j=i+2}^{k+1}\omega_j\right)\left(\gamma\tau'_{i+1}t_{i+1}^2 \epsilon_{i+1} + \gamma t_{i+1}\sqrt{2\epsilon_{i+1}\tau'_{i+1}}\|u^{(i+1)}\|_{D_{i+1}}\right).$$

Let us now define $\Omega_{k+1} := \prod_{i=0}^{k+1}\omega_i \in (0,1]$. Recalling the definition of θ_k in (30), we can rewrite the previous inequality as:

$$\tau'_{k+1}t_{k+1}^2 v_{k+1} + \frac{1}{2}\|u^{(k+1)}\|_{D_{k+1}}^2$$
$$\leq \gamma\Omega_{k+1}\left(\tau'_0 t_0^2 v_0\omega_0 + \frac{\omega_0}{2}\|u^{(0)}\|_{D_0}^2 + \sum_{i=0}^{k}\theta_{i+1}^{-1}\epsilon_{i+1} + \sqrt{2\theta_{i+1}^{-1}\epsilon_{i+1}}\,\Omega_{i+1}^{-\frac{1}{2}}\|u^{(i+1)}\|_{D_{i+1}}\right). \quad (36)$$

We can now proceed as in [33, Proposition 2]. First, we discard the term $\tau'_{k+1}t_{k+1}^2 v_{k+1}$ on the left-hand side, so as to deduce the following inequality

$$\Omega_{k+1}^{-1}\|u^{(k+1)}\|_{D_{k+1}}^2 \leq 2\gamma\left(\frac{\omega_0\|u^{(0)}\|_{D_0}^2}{2} + \tau'_0 t_0^2 v_0\omega_0 + \sum_{i=0}^{k}\theta_{i+1}^{-1}\epsilon_{i+1}\right) + \sum_{i=0}^{k}2\gamma\sqrt{2\theta_{i+1}^{-1}\epsilon_{i+1}}\,\Omega_{i+1}^{-\frac{1}{2}}\|u^{(i+1)}\|_{D_{i+1}}.$$

An application of Lemma 3.2 to this relation leads to

$$\Omega_{k+1}^{-\frac{1}{2}}\|u^{(k+1)}\|_{D_{k+1}} \leq \sum_{i=0}^{k}\gamma\sqrt{2\theta_{i+1}^{-1}\epsilon_{i+1}}$$
$$+ \left(2\gamma\left(\frac{\omega_0\|u^{(0)}\|_{D_0}^2}{2} + \tau'_0 t_0^2 v_0\omega_0 + \sum_{i=0}^{k}\theta_{i+1}^{-1}\epsilon_{i+1}\right) + \left(\sum_{i=0}^{k}\gamma\sqrt{2\theta_{i+1}^{-1}\epsilon_{i+1}}\right)^2\right)^{\frac{1}{2}}.$$

Therefore, by standard inequalities on the square root of the sum of positive terms, we can deduce that for all $j = 0, \ldots, k$

$$\Omega_{j+1}^{-\frac{1}{2}}\|u^{(j+1)}\|_{D_{j+1}} \leq \sqrt{\gamma\omega_0}\|u^{(0)}\|_{D_0} + \sqrt{2\gamma\tau'_0 t_0^2 v_0\omega_0} + \sqrt{2\sum_{i=0}^{k}\gamma\theta_{i+1}^{-1}\epsilon_{i+1}} + 2\sum_{i=0}^{k}\gamma\sqrt{2\theta_{i+1}^{-1}\epsilon_{i+1}}. \quad (37)$$

We now go back to (36), discard the term $\frac{1}{2}\|u^{(k+1)}\|_{D_{k+1}}^2$ on the left-hand side, plug the relation

(37) and write the following chain of inequalities

$$
\theta_{k+1}^{-1} v_{k+1} \le \gamma \tau_0' t_0^2 v_0 \omega_0 + \frac{\gamma \omega_0}{2} \|u^{(0)}\|_{D_0}^2 + \sum_{i=0}^{k} \gamma \theta_{i+1}^{-1} \epsilon_{i+1}
$$

$$
+ \left(2 \sum_{i=0}^{k} \gamma \sqrt{\theta_{i+1}^{-1} \epsilon_{i+1}} \right) \left(\sqrt{\frac{\gamma \omega_0}{2}} \|u^{(0)}\|_{D_0} + \sqrt{\gamma \tau_0' t_0^2 v_0 \omega_0} + \sqrt{\sum_{i=0}^{k} \gamma \theta_{i+1}^{-1} \epsilon_{i+1}} + 2 \sum_{i=0}^{k} \gamma \sqrt{\theta_{i+1}^{-1} \epsilon_{i+1}} \right)
$$

$$
\le \left(\sqrt{\frac{\gamma \omega_0}{2}} \|u^{(0)}\|_{D_0} + \sqrt{\gamma \tau_0' t_0^2 v_0 \omega_0} + 2 \sum_{i=0}^{k} \gamma \sqrt{\theta_{i+1}^{-1} \epsilon_{i+1}} + \sqrt{\sum_{i=0}^{k} \gamma \theta_{i+1}^{-1} \epsilon_{i+1}} \right)^2 .
$$

We thus deduce

$$
v_{k+1} \le \gamma \theta_{k+1} \left(\sqrt{\frac{\omega_0}{2}} \|u^{(0)}\|_{D_0} + \sqrt{\tau_0' t_0^2 v_0 \omega_0} + 2 \sum_{i=0}^{k} \sqrt{\gamma \theta_{i+1}^{-1} \epsilon_{i+1}} + \sqrt{\sum_{i=0}^{k} \theta_{i+1}^{-1} \epsilon_{i+1}} \right)^2 . \tag{38}
$$

At this point, proceeding as in [11, Theorem 4.6], we need to provide an upper bound for the factor θ_k. To this aim, we introduce the following average quantities:

$$
\sqrt{\bar{L}_{k+1}} := \frac{1}{\frac{1}{k+2} \sum\limits_{i=0}^{k+1} \frac{1}{\sqrt{L_i - \mu_{f,i}}}}, \qquad \sqrt{\bar{q}_{k+1}} := \frac{1}{k+1} \sum_{i=1}^{k+1} \sqrt{\frac{\mu_i}{L_i + \mu_{g,i}}}, \tag{39}
$$

where $L_i := 1/\tau_i$.

In order to show the quadratic decay of θ_k, we first prove the following inequality:

$$
\frac{1}{\sqrt{\eta_{sup}^k \theta_k}} \ge \frac{1}{2} \sum_{i=0}^{k} \frac{1}{\sqrt{\eta_{sup}^i \left(\frac{1}{\tau_i} - \mu_{f,i} \right)}}, \quad \forall \, k \ge 0. \tag{40}
$$

The proof follows by induction: for $k = 0$ we have

$$
\theta_0 = \frac{1 - \mu_0 \tau_0' t_0}{\tau_0' t_0^2} = \frac{(1 - \mu_0 \tau_0' t_0)(1 + \tau_0 \mu_{g,0})}{\tau_0 t_0^2} = \frac{1 - (t_0 - 1)\tau_0 \mu_{g,0} - \mu_{f,0} \tau_0 t_0}{\tau_0 t_0^2} \le \frac{1}{\tau_0} - \mu_{f,0}, \tag{41}
$$

which holds since, by assumption, $t_0 \ge 1$. We observe that this is nothing but (40) for $k = 0$. By induction, suppose now that (40) holds for some $k \ge 0$. Then, from (17), we deduce:

$$
1 - \frac{1}{t_{k+1}} = \frac{\mu_k}{\mu_{k+1}} \omega_{k+1} \frac{\tau_k' t_k^2}{\tau_{k+1}' t_{k+1}^2} = \frac{\eta_{sup}^{k+1}}{\eta_{sup}^k} \omega_{k+1} \frac{\tau_k' t_k^2}{\tau_{k+1}' t_{k+1}^2} = \frac{\eta_{sup}^{k+1}}{\eta_{sup}^k} \frac{\theta_{k+1}}{\theta_k} \le 1, \tag{42}
$$

which implies that the sequence $\{\eta_{sup}^k \theta_k\}_{k \in \mathbb{N}}$ is non-increasing. Therefore, we have

$$
\frac{1}{\sqrt{\eta_{sup}^{k+1} \theta_{k+1}}} - \frac{1}{\sqrt{\eta_{sup}^k \theta_k}} = \frac{\eta_{sup}^k \theta_k - \eta_{sup}^{k+1} \theta_{k+1}}{\sqrt{\eta_{sup}^k \eta_{sup}^{k+1} \theta_k \theta_{k+1}} \left(\sqrt{\eta_{sup}^k \theta_k} + \sqrt{\eta_{sup}^{k+1} \theta_{k+1}} \right)} \ge \frac{\eta_{sup}^k \theta_k - \eta_{sup}^{k+1} \theta_{k+1}}{2 \eta_{sup}^k \theta_k \sqrt{\eta_{sup}^{k+1} \theta_{k+1}}}.
$$

16

By using again (42), we further obtain

$$\frac{1}{\sqrt{\eta_{sup}^{k+1}\theta_{k+1}}} - \frac{1}{\sqrt{\eta_{sup}^k \theta_k}} \geq \frac{1}{2t_{k+1}\sqrt{\eta_{sup}^{k+1}\theta_{k+1}}}.$$

Now, recalling the definitions of the parameters θ_{k+1} and ω_{k+1}, we can write the following chain of inequalities:

$$t_{k+1}\sqrt{\eta_{sup}^{k+1}\theta_{k+1}} = \sqrt{\frac{\eta_{sup}^{k+1}}{\tau_{k+1}'}\prod_{i=0}^{k+1}\sqrt{\omega_i}} \leq \sqrt{\frac{\eta_{sup}^{k+1}\omega_{k+1}}{\tau_{k+1}'}} = \sqrt{\eta_{sup}^{k+1}\left(\frac{1}{\tau_{k+1}'} - \mu_{k+1}t_{k+1}\right)}$$

$$\leq \sqrt{\eta_{sup}^{k+1}\left(\frac{1}{\tau_{k+1}'} - \mu_{k+1}\right)} = \sqrt{\eta_{sup}^{k+1}\left(\frac{1}{\tau_{k+1}} - \mu_{f,k+1}\right)}.$$

Thus, we get

$$\frac{1}{\sqrt{\eta_{sup}^{k+1}\theta_{k+1}}} - \frac{1}{\sqrt{\eta_{sup}^k \theta_k}} \geq \frac{1}{2\sqrt{\eta_{sup}^{k+1}\left(\frac{1}{\tau_{k+1}} - \mu_{f,k+1}\right)}}.$$

By induction, the previous inequality yields (40) with $k = k + 1$. Recalling the definition of $\bar{L}_{k+1}$ and observing that $\eta_{inf} \leq \eta_{sup}^i \leq \eta_{sup}$ for $i = 0, \ldots, k$, we get

$$\sqrt{\theta_{k+1}} \leq \frac{2}{k+2}\sqrt{\frac{\eta_{sup}}{\eta_{inf}}\bar{L}_{k+1}}, \quad \forall\, k \geq 0. \tag{43}$$

To get the linear rate, we use (42) and (19) and deduce the following relations:

$$\theta_{k+1} = \theta_0 \prod_{i=1}^{k+1}\frac{\eta_{sup}^{i-1}}{\eta_{sup}^i}\left(1 - \frac{1}{t_i}\right) \leq \frac{\eta_{sup}^0 \theta_0}{\eta_{sup}^{k+1}}\prod_{i=1}^{k+1}(1 - \sqrt{q_i})$$

$$= \frac{\eta_{sup}^0 \theta_0}{\eta_{sup}^{k+1}}\prod_{i=1}^{k+1}\left(1 - \sqrt{\frac{\mu_i}{L_i + \mu_{g,i}}}\right) \leq \frac{\eta_{sup}}{\eta_{inf}}(L_0 - \mu_{f,0})\left(1 - \sqrt{\bar{q}_{k+1}}\right)^{k+1}, \tag{44}$$

where the last inequality follows from the concavity of the logarithmic function, the definition of $\bar{q}_{k+1}$, (41) and the fact that $\eta_{inf} \leq \eta_{sup}^i \leq \eta_{sup}$ for $i = 0, \ldots, k + 1$.

Finally, we note that the averaging term $\bar{L}_{k+1}$ appearing in (43) is always smaller than the actual average of the terms $(L_i - \mu_{f,i})$, since:

$$\sqrt{\bar{L}_k} \leq \frac{1}{k+1}\sum_{i=0}^k \sqrt{L_i - \mu_{f,i}} \leq \sqrt{\frac{1}{k+1}\sum_{i=0}^k (L_i - \mu_{f,i})}. \tag{45}$$

Then, since the use of the backtracking strategy at any iteration implies that $\tau_k \geq \rho/L_f$, we can deduce the following uniform bounds for the averaging terms:

$$\sqrt{\bar{L}_{k+1}} \leq \sqrt{\frac{L_f}{\rho} - \frac{\mu_f}{\eta_{sup}}}, \qquad \sqrt{\bar{q}_{k+1}} \geq \sqrt{\frac{\mu\rho}{L_f\eta_{sup} + \mu_g\rho}},$$

by which (33) follows. □

Remark 3.4. *By (43)-(44), we deduce that the factor θ_{k+1} can be estimated just in terms of the average quantities $\overline{L}_{k+1}$ and $\overline{q}_{k+1}$ defined in (39), thus avoiding the dependence on the (possibly unknown) value L_f. In this case, note that the storage of all the values L_i computed along the iterations by backtracking strategy is required.*

The boundedness of the sequences E_1^k and E_2^k can be guaranteed by carefully choosing the error parameters ϵ_{k+1} controlling the accuracy of the inexact proximal evaluations. If $\mu > 0$ the following corollary provides a sufficient condition for this choice.

Corollary 3.1. *Suppose that $F = f + g$ is $\mu-$strongly convex with $\mu > 0$ and let x^* be the solution of (6). Suppose that Assumption 1 holds and that the sequence of errors $\{\epsilon_k\}_{k\in\mathbb{N}}$ is chosen as*

$$\epsilon_{k+1} = \theta_{k+1}a_{k+1}, \tag{46}$$

where $\{\theta_k\}_{k\in\mathbb{N}}$ is defined as in (30) and $\{a_k\}_{k\in\mathbb{N}}$ is a sequence of nonnegative numbers such that $\sum_{k=0}^{\infty} \sqrt{a_k} < \infty$. Then, for all sufficiently large k, we have

$$F(x^{(k+1)}) - F(x^*) = \mathcal{O}\left(\left(1 - \sqrt{\frac{\mu\rho}{L_f\eta_{sup} + \mu_g\rho}}\right)^{k+1}\right).$$

Note that the factor θ_{k+1} depends on q_{k+1} and t_{k+1}, which are computed at STEP 2 and STEP 3 of Algorithm 1. Therefore, in order to ensure the linear rate, the parameter ϵ_{k+1} needs to be recomputed according to (46) at each trial step of the backtracking procedure.

In the particular case $\mu = \mu_g > 0$, we are able to pre-identify a sequence of error parameters that guarantees a linear rate without requiring any inner update, as shown below.

Corollary 3.2. *Suppose that $F = f + g$ is $\mu-$strongly convex with $\mu = \mu_g > 0$ and let x^* be the solution of (6). Suppose that Assumption 1 holds and that the sequence of errors $\{\epsilon_k\}_{k\in\mathbb{N}}$ is chosen as*

$$\epsilon_{k+1} = \mathcal{O}((ab^k)^{k+1}), \tag{47}$$

where $a < \frac{\delta}{2}\min\{1, \eta_{inf}/(\tau_0\mu_g)\}$ and $b < \sqrt{\delta}$. Then, for all sufficiently large k, we have

$$F(x^{(k+1)}) - F(x^*) = \mathcal{O}\left(\left(1 - \sqrt{\frac{\mu_g\rho}{L_f\eta_{sup} + \mu_g\rho}}\right)^{k+1}\right).$$

Proof. Let us provide an upper bound for the quantity θ_{i+1}^{-1} for $i = 0, \ldots, k$. Recalling the relation $\tau'_{i+1} = q_{i+1}/\mu_{i+1}$ and the fact that $q_{i+1}t_{i+1}^2 \leq 1$ for all i (see Lemma 3.1), we have

$$\theta_{i+1}^{-1} = \frac{q_{i+1}t_{i+1}^2}{\mu_{i+1}\prod_{j=1}^{i+1}(1 - q_jt_j)} \leq \frac{1}{\mu_{i+1}}\prod_{j=1}^{i+1}\frac{1}{1 - \sqrt{q_j}} = \frac{1}{\mu_{g,i+1}}\prod_{j=1}^{i+1}\sqrt{1 + \tau_j\mu_{g,j}}(\sqrt{1 + \tau_j\mu_{g,j}} + \sqrt{\tau_j\mu_{g,j}}),$$

where the second equality follows from STEP 2 of Algorithm 1 and the assumption $\mu = \mu_g$. Note that we can recursively apply the inequality $\tau_j \leq \tau_{j-1}/\delta$ so as to obtain

$$\tau_j \leq \tau_0\left(\frac{1}{\delta}\right)^j, \quad j \geq 0. \tag{48}$$

Then, we can continue as follows:

$$\theta_{i+1}^{-1} \leq \frac{2^{i+1}}{\mu_{g,i+1}} \prod_{j=1}^{i+1} 1 + \tau_j \mu_{g,j} \leq \frac{\eta_{sup}(2 \max\{1, (\tau_0 \mu_g)/\eta_{inf}\})^{i+1}}{\mu_g} \prod_{j=1}^{i+1} \left(\frac{1}{\delta}\right)^j$$

$$= \frac{\eta_{sup}(2 \max\{1, (\tau_0 \mu_g)/\eta_{inf}\})^{i+1}}{\mu_g} \left(\frac{1}{\sqrt{\delta}}\right)^{(i+2)(i+1)},$$

where the second inequality follows from (48), the definition of $\mu_{g,i+1}$ and $\eta_{inf} \leq \eta_{sup}^j \leq \eta_{sup}$. Then, the thesis follows by applying the above inequality to (32) in Theorem 3.1. $\qquad \square$

We remark that condition (47) is more severe when $\delta < 1$, i.e. whenever an adaptive backtracking procedure is considered.

If $\mu = 0$, the SAGE-FISTA Algorithm 1 reduces to an inexact variable metric version of FISTA. Consequently, Theorem 3.1 recovers the well-known $\mathcal{O}(1/k^2)$ convergence rate result available for FISTA and for its several inexact variants in the convex case [3,8,33]. Similarly to the case $\mu = \mu_g$, we can devise some prefixed rule for computing the errors ϵ_{k+1} in order to get the expected quadratic rate.

Corollary 3.3. *Suppose that $F = f + g$ is $\mu-$strongly convex with $\mu = 0$ and let x^* be a solution of (6). Suppose Assumption 1 holds and the sequence of errors $\{\epsilon_k\}_{k \in \mathbb{N}}$ is chosen as*

$$\epsilon_{k+1} = \begin{cases} \mathcal{O}(a^{k+1}), & \text{if } \delta < 1 \\ \dfrac{b_{k+1}}{(k+1+t_0)^2}, & \text{if } \delta = 1 \end{cases}, \tag{49}$$

where $a < \delta$ and $\{b_k\}_{k \in \mathbb{N}}$ is a sequence of nonnegative numbers such that $\sum_{k=0}^{\infty} \sqrt{b_k} < \infty$. Then, for all $k \geq 0$, we have

$$F(x^{(k+1)}) - F(x^*) = \mathcal{O}\left(\frac{1}{(k+1)^2}\right).$$

Proof. If $\mu = 0$, we can majorize the quantity $\sqrt{\theta_{i+1}^{-1}}$ as follows

$$\sqrt{\theta_{i+1}^{-1}} = \sqrt{\tau_{i+1}} t_{i+1} \leq \sqrt{\tau_{i+1}}\left(1 + t_i \sqrt{\frac{\tau_i}{\tau_{i+1}}}\right) = \sqrt{\tau_{i+1}} + t_i \sqrt{\tau_i} \leq \ldots \leq \sum_{j=1}^{i+1} \sqrt{\tau_j} + t_0 \sqrt{\tau_0}, \quad (50)$$

which follows by applying recursively the property:

$$t_{j+1} \leq 1 + t_j \sqrt{\frac{\tau_j}{\tau_{j+1}}}, \quad j = 0, \ldots, i.$$

Combining (50) with (48) leads to

$$\theta_{i+1}^{-1} \leq \left(\sum_{j=1}^{i+1} \sqrt{\tau_0} \left(\frac{1}{\sqrt{\delta}}\right)^j + t_0 \sqrt{\tau_0}\right)^2 = \begin{cases} \tau_0 \left(\frac{(1/\sqrt{\delta})^{i+2} - (1/\sqrt{\delta})}{(1/\sqrt{\delta}) - 1} + t_0\right)^2, & \text{if } \delta < 1, \\ \tau_0 \left(i + 1 + t_0\right)^2, & \text{if } \delta = 1. \end{cases}$$

Using this upper bound in (32), we get the thesis. $\qquad \square$

Note that when $\delta = 1$ condition (49) is analogous to the one imposed in [8, Corollary 3.1] to get the quadratic rate for an inexact, variable metric version of the FISTA algorithm. When $\delta < 1$, (49) is more restrictive, as it requires the errors to decay at a linear rate.

Finally, we consider the case where no backtracking is performed, i.e. the step-size τ_0 is chosen in order to satisfy the descent Lemma 2.2 and $\delta = 1$.

Corollary 3.4. *Suppose that $F = f + g$ is $\mu-$strongly convex with $\mu > 0$ and let x^* be a solution of (6). Suppose Assumption 1 holds, $\delta = 1$ and $0 < \tau_0 < \frac{\eta_{inf}}{L_f}$. Define the parameter $0 < q < 1$ as $q = (\tau_0\mu)/(\eta_{inf} + \tau_0\mu_g)$, choose $a < 1 - \sqrt{q}$ and the sequence of errors $\{\epsilon_k\}_{k\in\mathbb{N}}$ as*

$$\epsilon_{k+1} = \mathcal{O}(a^{k+1}). \tag{51}$$

Then, for all $k \geq 0$, we have $\tau_k = \tau_0$ and

$$F(x^{(k+1)}) - F(x^*) = \mathcal{O}\left(\left(1 - \sqrt{\frac{\tau_0\mu}{\eta_{sup} + \tau_0\mu_g}}\right)^{k+1}\right).$$

Proof. Since $\tau_0 < \eta_{inf}/L_f$ and $\mu_f \leq L_f$, there holds $q < 1$. Thanks to $\tau_0 < \eta_{inf}/L_f$, $\delta = 1$ and Lemma 2.2, we also have that $\tau_k = \tau_0$ for all $k \geq 0$. Furthermore, $\eta_{sup}^k \geq \eta_{inf}$ implies $q_k \leq q$ for all $k \geq 0$. Then, the following upper bound holds for θ_{i+1}^{-1}:

$$\theta_{i+1}^{-1} \leq \frac{1}{\mu_{i+1}} \prod_{j=1}^{i+1} \frac{1}{1 - \sqrt{q_j}} \leq \frac{\eta_{sup}}{\mu} \left(\frac{1}{1 - \sqrt{q}}\right)^{i+1}. \tag{52}$$

Finally, since $\tau_k \equiv \tau_0$, we can provide the following lower bound for the average quantity $\bar{q}_{k+1}$:

$$\sqrt{\bar{q}_{k+1}} \geq \sqrt{\frac{\mu\tau_0}{\eta_{sup} + \tau_0\mu_g}}. \tag{53}$$

Then, choosing ϵ_{k+1} as in (51), and applying (44) combined with (53) in (32), the thesis follows. □

Corollary 3.4 guarantees that the function values of the iterates of the SAGE-FISTA Algorithm 1 converge to the optimal value at a $R-$linear rate, provided that the errors sequence $\{\epsilon_k\}_{k\in\mathbb{N}}$ converges linearly to 0 at a sufficiently small rate. We remark that the result obtained in Corollary 3.4 is similar to the one provided in [33, Proposition 4] for a basic variant of the inexact FISTA algorithm, which is specifically designed for strongly convex problems. In contrast with our work, the authors in [33] assume strong convexity only on the differentiable part and, unlike SAGE-FISTA, their algorithm makes use of a constant inertial parameter defined by $\beta_k = (1 - \sqrt{\mu_f/L_f})/(1 + \sqrt{\mu_f/L_f})$.

4   Numerical experiments

In this section, we report some numerical experiments where SAGE-FISTA (Algorithm 1) is applied to some image restoration problems such as edge-preserving image denoising and image

deblurring. As observed in several related works (see, e.g., [6] for a review), the scaling framework discussed in Section 3 turns out to be particularly effective whenever a modeling encoding signal-dependent Poisson noise is considered, as it is common in microscopy and astronomy imaging applications [4].

Given an observed image $z \in \mathbb{R}^n_{\geq 0}$ with nonnegative pixels, we thus consider the following ill-posed image restoration problem:

$$z = \mathcal{P}(Hx + b), \tag{54}$$

where $x \in \mathbb{R}^n_{\geq 0}$ is the desired solution, $H \in \mathbb{R}^{n \times n}$ represents a blur (convolution) operator computed for a given Point Spread Function (PSF), the term $b \in \mathbb{R}^n_{>0}$ stands for a positive background and $\mathcal{P}(w)$ denotes a realization of a Poisson-distributed n-dimensional random vector with parameter $w \in \mathbb{R}^n_{>0}$. Following the standard Bayesian Maximum A Posteriori (MAP) paradigm, one finds that the data fidelity term corresponding to the negative log-likelihood of the Poisson p.d.f. is the generalized Kullback-Leibler (KL) divergence functional defined by:

$$KL(Hx + b; z) := \sum_{i=1}^{n} \left(z_i \log \frac{z_i}{(Hx)_i + b_i} + (Hx)_i + b_i - z_i \right), \tag{55}$$

where the convention $0 \log 0 = 0$ is adopted. If the operator H has nonnegative entries and if it has at least one strictly positive entry for each row and column (i.e. $He > 0$ and $H^T e > 0$ for $e \in \mathbb{R}^n$ being the vector of all ones), the KL divergence is a nonnegative, convex and coercive function on the non-negative orthant $Y := \{x \geq 0\}$, see, e.g., [21]. The generalized KL divergence can be then coupled with convex and non-smooth image regularization terms, such as the isotropic Total Variation (TV) regularization:

$$TV(x) := \|\nabla x\|_{2,1} = \sum_{i=1}^{n} \|\nabla_i x\|_2, \tag{56}$$

where, for all $i = 1, \ldots, n$, the term $\nabla_i x \in \mathbb{R}^2$ stands for the standard forward-difference discretization of the image gradient along the two horizontal and vertical directions at pixel i. For simplicity, we will consider reflexive boundary conditions for the computation of such discretization. Due to the non-smoothness of the TV term (56), first-order optimization schemes based on forward-backward splitting are often used for solving the composite optimization problem:

$$\min_{x \in \mathbb{R}^n} KL(Hx + b; z) + \lambda TV(x) + \iota_Y(x) \tag{57}$$

where the parameter $\lambda > 0$ balances the effect of the regularization against data fidelity and where $\iota_Y(\cdot)$ denotes the indicator function of Y, thus enforcing a non-negative constraint on $Y \subset \mathbb{R}^n$ for the desired solution x.

In the following sections, we consider two models analogous to (57), which we adapt to fit with the strongly-convex scenarios suitable for SAGE-FISTA. For both models, we simulate Poisson noisy acquisitions by means of the MATLAB `imnoise` routine. We report in Figure 1 the blurred and noisy images used in our numerical experiments, while in Table 1 we specify the parameters involved in both the simulation and the reconstruction for each image considered.

For all tests, an approximation x^* of the desired solution is pre-computed by running standard FISTA for 5000 iterations. Our results are validated by computing the relative objective error:

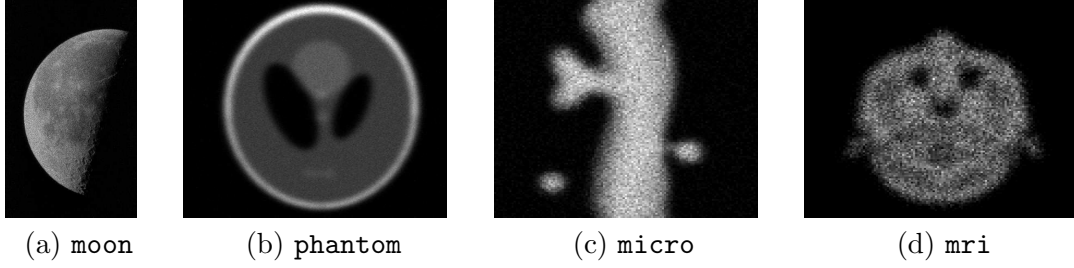$$eF_k = \frac{F(x^{(k)}) - F(x^*)}{F(x^*)}, \qquad k \geq 1 \tag{58}$$

21

(a) `moon` (b) `phantom` (c) `micro` (d) `mri`

Figure 1: Blurred and noisy test images used for the numerical tests.

Image	Size	Range	σ_{PSF}	b	λ	μ	L_f
`moon`	537×358	$[0, 400]$	0	0.01	0.15	$\mu_f \in \{0, 0.0025\}$	100
`phantom`	256×256	$[0, 1]$	1.4	0.01	0.004	$\mu_g = 0.0001$	7701
`micro`	128×128	$[0, 92]$	3.2	0.5	0.092	$\mu_g = 0.0001$	368
`mri`	128×128	$[0, 170]$	3.2	0.5	0.001	$\mu_g = 0.0005$	680

Table 1: Simulation and model parameters for the test images in Figure 1.

both along the iterations and with respect to computational times. In the latter case, we show (at most) the first 30 seconds of run, which for our examples is a time frame long enough to reach a significant level of accuracy. The other algorithmic parameters (backtracking factors and initializations) are specified in each test.

Inexact computation of proximal points with ϵ_k-accuracy. Following [8, Section 4.2], we now sketch a general strategy ensuring that the inexactness condition required by Definition 2.2 is fulfilled in practice. For that, we require that the non-smooth function g in (6) can be expressed in the form:

$$g(x) = \sum_{i=1}^{p} \phi_i(M_i x) + \psi(x), \tag{59}$$

where $M_i : \mathcal{H} \to \mathcal{Z}_i$ are linear bounded operators between Hilbert spaces and $\phi_i : \mathcal{Z}_i \to \mathbb{R} \cup \{\infty\}$, $\psi : \mathcal{H} \to \mathbb{R} \cup \{\infty\}$ are proper, l.s.c. and convex functions. For image restoration problems, the couple (ϕ_i, M_i) is typically associated to the use of edge-preserving regularization terms (e.g. (56) by setting $\phi_i : \mathbb{R}^2 \to \mathbb{R}, \phi_i(v) = \lambda \|v\|$ and $M_i = \nabla_i$), while the function ψ may encode further requirements on the desired signal such as a positivity constraint or a convex perturbation term.

At each iteration $k \geq 1$ of SAGE-FISTA, the inner (primal) subproblem that needs to be solved to compute the proximal point $x^{(k)}$ has the form:

$$\min_{x \in \mathcal{H}} \left\{ \mathcal{P}_{\tau_k, D_k}(x) := \sum_{i=1}^{p} \phi_i(M_i x) + \psi(x) + \frac{1}{2\tau_k} \|x - \bar{y}^{(k)}\|_{D_k}^2 \right\}, \tag{60}$$

where $\bar{y}^{(k)} := y^{(k)} - \tau_k D_k^{-1} \nabla f(y^{(k)})$. Denoting by $\mathcal{Z} = \mathcal{Z}_1 \times \ldots \times \mathcal{Z}_p$ the Hilbert product space equipped with the inner product $\langle w, v \rangle = \sum_{i=1}^{p} \langle w_i, z_i \rangle$ and by $M : \mathcal{H} \to \mathcal{Z}$ the linear bounded

operator $Mx = (M_1 x, \ldots, M_p x) \in \mathcal{Z}$, the dual formulation of (60) reads as

$$\max_{w \in \mathcal{Z}} \left\{ \mathcal{Q}_{\tau_k, D_k}(w) := -\sum_{i=1}^{p} \phi_i^*(w_i) + \Phi_{\tau_k, D_k, \bar{y}^{(k)}}(w) \right\} \tag{61}$$

where ϕ_i^* is the Fenchel conjugate of ϕ_i and $\Phi_{\tau_k, D_k, \bar{y}^{(k)}}(w)$ is defined as:

$$\Phi_{\tau_k, D_k, \bar{y}^{(k)}}(w) := \psi \left(\text{prox}_{\tau_k \psi}^{D_k}(\bar{y}^{(k)} - \tau_k D_k^{-1} M^* w) \right) - \frac{1}{2\tau_k} \|\bar{y}^{(k)} - \tau_k D_k^{-1} M^* w\|_{D_k}^2 + \frac{1}{2\tau_k} \|\bar{y}^{(k)}\|_{D_k}^2$$

$$+ \frac{1}{2\tau_k} \|\text{prox}_{\tau_k \psi}^{D_k}(\bar{y}^{(k)} - \tau_k D_k^{-1} M^* w) - (\bar{y}^{(k)} - \tau_k D_k^{-1} M^* w)\|_{D_k}^2.$$

Upon suitable assumptions (see, e.g. [8, Section 4.1]) problems (60) and (61) are equivalent. Recalling definition (9), one can further deduce the following inequality:

$$h_{\tau_k, D_k}(x; \bar{y}^{(k)}) - h_{\tau_k, D_k}(\text{prox}_{\tau_k \psi}^{D_k}(\bar{y}^{(k)}); \bar{y}^{(k)}) = \mathcal{P}_{\tau_k, D_k}(x) - \mathcal{P}_{\tau_k, D_k}\left(\text{prox}_{\tau_k \psi}^{D_k}(\bar{y}^{(k)}) \right)$$

$$\leq \mathcal{P}_{\tau_k, D_k}(x) - \mathcal{Q}_{\tau_k, D_k}(w), \quad \forall w \in \mathcal{Z}.$$

Hence, a sufficient condition for a point x to be an ϵ_k-approximation as in Definition 2.2 is the existence of a dual point w such that $\mathcal{P}_{\tau_k, D_k}(x) - \mathcal{Q}_{\tau_k, D_k}(w) \leq \epsilon_k$. Assuming that g is continuous on its domain and $\text{dom}(\psi) = \text{dom}(g)$, such pair (x, w) can be computed by generating a dual sequence $\{w^{(k,l)}\}_{l \in \mathbb{N}} \subset \mathcal{Z}$ converging to the solution of (61) with

$$\lim_{l \to \infty} \mathcal{Q}_{\tau_k, D_k}\left(w^{(k,l)} \right) = \max_{z \in \mathcal{Z}} \mathcal{Q}_{\tau_k, D_k}(z),$$

a corresponding primal sequence $\{x^{(k,l)}\}_{l \in \mathbb{N}} \subset \mathcal{H}$ computed as

$$x^{(k,l)} := \text{prox}_{\tau_k \psi}^{D_k}\left(\bar{y}^{(k)} - \tau_k D_k^{-1} M^* w^{(k,l)} \right), \qquad \forall l \in \mathbb{N}, \tag{62}$$

and then stopping the iterates at the first nonnegative integer l such that

$$\mathcal{P}_{\tau_k, D_k}(x^{(k,l)}) - \mathcal{Q}_{\tau_k, D_k}(w^{(k,l)}) \leq \epsilon_k. \tag{63}$$

Such a procedure is well-defined, as condition (63) holds for all sufficiently large l [8, Proposition 4.2]. Note that the dual sequence $\{w^{(k,l)}\}_{l \in \mathbb{N}}$ can be generated by means of a forward–backward scheme solving (61), e.g. using an efficient inner FISTA routine, provided that the extrapolation parameters are chosen in a way that weak convergence of the iterates is guaranteed (see, e.g. [12]).

Variable metric selection As far as the choice of the scaling matrices $\{D_k\}_{k \in \mathbb{N}}$ is concerned, we exploit the split-gradient technique studied in [23] and later used in several works (see, e.g., [5, 8]) where a decomposition of ∇f into a non-negative part and a positive one is used. The idea beyond this strategy consists in decomposing the gradient of the smooth component f of problem (6) as $-\nabla f(x) = U(x) - V(x)$, where $U(x) \geq 0$ and $V(x) > 0$, and then defining for each $k \geq 1$ the scaling matrix $D_k = \text{diag}\left(y^{(k)}/V(y^{(k)}) \right)^{-1}$, so that the forward iteration can be written as:

$$\bar{y}^{(k)} = y^{(k)} - \tau_k D_k^{-1} \nabla f(y^{(k)}) = y^{(k)} + \tau_k y^{(k)} \cdot \left(\frac{U(x^{(k)})}{V(x^{(k)})} - 1 \right), \tag{64}$$

23

where the fraction and product symbols denote component-wise division and product, respectively. Note that several classical reconstruction algorithms in imaging can be cast in the form (64), such as the Richardson-Lucy (RL) algorithm for Poisson image deconvolution (see, e.g., [6, 23]). In order to ensure the conditions required by Assumption 1, we introduce an appropriate thresholding parameter in the definition of D_k, thus obtaining

$$D_k = \text{diag}\left(\max\left(\frac{1}{\gamma_k}, \min\left(\gamma_k, \frac{y^{(k)}}{V(y^{(k)})}\right)\right)\right)^{-1},\qquad(65)$$

for parameters γ_k defined as in [5, 8] by:

$$\gamma_k = \sqrt{1 + \frac{s_1}{(k+1)^{s_2}}}, \quad \text{where } s_1 > 0, \ s_2 > 1.\qquad(66)$$

Clearly, the choice (65) is problem-dependent, due to presence of the the function $V(\cdot)$. Note that in (66), when $s_1 = 0$, we have $D_k \equiv \mathcal{I}$, hence the standard Euclidean metric is recovered. In general, it is good practice to choose a large value of s_1 in order to benefit from the use of the variable metric in the early iterations of the algorithm, while leaving the asymptotic convergence behavior to be driven by s_2.

4.1 Weighted-ℓ^2-TV image denoising

As a first test, we set $H = \mathcal{I} \in \mathbb{R}^{n \times n}$ in (54) and consider a Poisson image denoising problem where TV regularization (56) is combined with a data fidelity term corresponding to the second order Taylor approximation of the KL divergence (55) around the noisy data z, see, e.g. [10, 24, 31]. For a given Poisson noisy image $z \in \mathbb{R}^n_{\geq 0}$ and for a fixed regularization parameter $\lambda > 0$, we consider the problem:

$$x^* = \underset{x \in \mathbb{R}^n}{\text{argmin}} \left\{ F(x) := \frac{1}{2} \sum_{i=1}^{n} \frac{(x_i - z_i + b)^2}{z_i + b} + \lambda TV(x) + \iota_Y(x) \right\},\qquad(67)$$

where $b > 0$ is a constant background term. By weighting the least-square residuals by the inverse values of the noisy image, the use of the data fidelity term in (67) promotes high fidelity in low-intensity pixels (low noise) and large regularization in high-intensity pixels (high noise). Applying the splitting (6) to problem (67), we can thus set:

$$f(x) := \frac{1}{2} \sum_{i=1}^{n} \frac{(x_i - z_i + b)^2}{z_i + b}, \qquad g(x) := \lambda TV(x) + \iota_Y(x)\qquad(68)$$

and observe that f is σ_f-strongly convex with L_f-Lipschitz continuous gradient with:

$$\sigma_f = \frac{1}{\max\limits_{i=1,\dots,n} z_i + b}, \qquad \nabla f(x) = \frac{x - z + b}{z + b}, \qquad L_f = \frac{1}{\min\limits_{i=1,\dots,n} z_i + b}.\qquad(69)$$

We now apply SAGE-FISTA Algorithm 1 with $\mu_g = 0$ and either $\mu_f = \sigma_f$ or $\mu_f = 0$. In particular, we report the results obtained when strong convexity is taken into account (SAGE-FISTA with $\mu_f = \sigma_f$) and when it is not (S-FISTA with $\mu_f = 0$). For this first test, we consider the `moon` noisy image in Figure 1. From (69) and Table 1, we have that $L_f = 100$ in this case.

Scaled VS. non-scaled SAGE-FISTA with backtracking. First, we compare the performance of the non-scaled version of SAGE-FISTA with its scaled counterparts for different choices of the scaling matrices $\{D_k\}_{k\in\mathbb{N}}$. We apply Algorithm 1 with $\lambda = 0.15$, maximum number of outer iterations `maxiter`=500, maximum number of inner backtracking iterations `max_bt`=10, $\rho = 0.8$ (decreasing backtracking factor), $L_0 = 30$ and $t_0 = 1.01$. In order to compare the adaptive backtracking strategy with the non-adaptive one, we choose $\delta \in \{0.99, 1\}$. We recall that for $\delta = 1$, a standard Armijo backtracking is performed, while $\delta < 1$ allows for local increasing of the algorithmic step-size. Finally, SAGE-FISTA is initialized by setting $x^{(0)} = z$.

In order to compute a suitable ϵ_k-approximation $x^{(k)}$ of the proximal-gradient point, we follow the strategy described above with the choice $\phi_i : \mathbb{R}^2 \to \mathbb{R}$, $\phi_i(v) = \lambda\|v\|_2$, $M_i = \nabla_i, i = 1, \ldots, n$ and $\psi : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}, \psi(x) = \iota_Y(x)$. Then, for each inner iteration $l \in \mathbb{N}$ of the primal-dual subroutine, the primal iterate (62) is given by

$$x^{(k,l)} = P_{Y,D_k}\left(y^{(k)} - \tau_k D_k^{-1}\left(\nabla f(y^{(k)}) + M^* w^{(k,l)}\right)\right). \tag{70}$$

Concerning the sequence $\{\epsilon_k\}_{k\in\mathbb{N}}$, we apply the result provided by Corollary 3.1, choosing in particular $a_k = 1/k^{2.1}$ for all $k \geq 1$ and updating θ_k in the inner backtracking routine by using the current estimates of t_k and q_k appearing in the definition (18) of the elements ω_k.

The sequence of scaling matrices $\{D_k\}_{k\in\mathbb{N}}$ is chosen by applying the diagonal split-gradient strategy described above, which, for $k \geq 1$, corresponds to the choice:

$$D_k = \text{diag}\left(\max\left(\frac{1}{\gamma_k}, \min\left(\gamma_k, z + b\right)\right)\right)^{-1}, \tag{71}$$

where γ_k are the thresholding parameters in (66). We notice that since $H = \mathcal{I}$ in this example, the expression of the terms D_k in (71) is very handy as it depends only on the noisy image z and on the background term b.

We now apply SAGE-FISTA with and without scaling for solving problem (67) whenever different choices of parameters s_1 and s_2 in (66) are considered. In Figure 2, we compare the results obtained with standard Armijo backtracking ($\delta = 1$) and adaptive backtracking ($\delta = 0.99$), respectively. This choice allows for local non-monotone adjustments of the estimates L_k of L_f, see the plots for the Lipschitz estimates in Figure 2c. We observe that the scaling procedure significantly improves convergence rates for any selected combination of the thresholding parameters (s_1, s_2), while remaining affordable in terms of computational times. This is a well-known behavior for scaled non strongly-convex variants of FISTA [8]; our experiments confirm that the same holds in strongly-convex regimes.

SAGE-FISTA VS. S-FISTA with backtracking. In order to show the effectiveness of SAGE-FISTA also when strong convexity is not directly taken into account, we now run the same experiments as above by setting $\mu_f = 0$ while keeping all the other parameters fixed. Note that this corresponds to apply the inexact and scaled FISTA (denoted here by S-FISTA) algorithm considered in [8] in analogous contexts, the novelty being here the use of the adaptive backtracking strategy in comparison to the standard Armijo one.

For S-FISTA, we compute the scaling matrices $\{D_k\}_{k\in\mathbb{N}}$ either according to the variable strategy (71) or setting $D_k \equiv D$ with $D = \text{diag}\,(z + b)^{-1}$. This constant choice is motivated by the fact that, for problem (67), the update of the matrices D_k in (71) depends only on the presence of the parameters γ_k, which can be possibly kept fixed (for instance by setting $s_1 \gg 1$

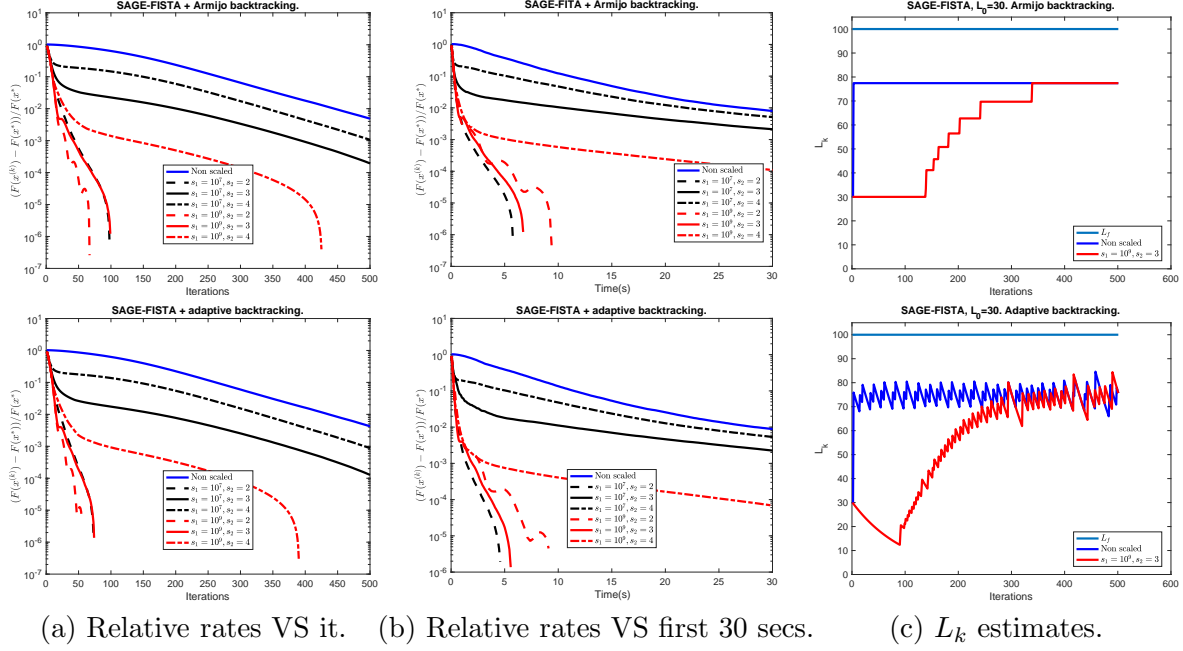(a) Relative rates VS it.   (b) Relative rates VS first 30 secs.   (c) L_k estimates.

Figure 2: Performance of non-scaled VS. scaled SAGE-FISTA with Armijo backtracking (top row) and adaptive backtracking (bottom row) with $L_0 = 30$ for different choices of s_1 and s_2 for problem (67) on `moon` image.

and $s_2 = 0$), thus imposing a constant Newton-type scaling matrix along the iterations. By Remark 3.3, this choice still satisfies Assumption 1, thus the convergence result of Theorem 3.1 still applies. Our results are reported in Figure 3. As expected, this choice improves even more the convergence rate with respect to the number of iterations, at the price of significantly higher computational costs. Indeed, since the matrices D_k do not converge any longer to the identity matrix, the Lipschitz constant of the inner dual problem (61) tends to be bigger, being proportional to $\|D_k^{-1}\|$ (see [8, Section 4.2]), thus affecting negatively the inner FISTA subroutine for the computation of the ϵ_k−approximation when the parameter ϵ_k is small. In support of our remark, we report the total CPU time to execute 500 iterations: S-FISTA with Armijo backtracking takes 106.5 s (variable metric) and 3591 s (constant metric), whereas S-FISTA with adaptive backtracking takes 238.4 s (variable metric) and 2672.6 s (constant metric). We conclude that, for suitable choices of s_1 and s_2, the variable metric strategy (71) represents a good balance between the good convergence properties of Newton methods and the higher computational costs.

As a final comparison, we report in Figure 3b the relative convergence rates for the various scaled and non-scaled versions of S-FISTA and SAGE-FISTA, combined with either an Armijo or an adaptive backtracking strategy. As it is expected, encoding explicitly the prior knowledge of strong convexity (SAGE-FISTA) improves convergence rates in comparison with its convex variant (S-FISTA). As far as the two backtracking strategies are concerned, the use of an adaptive strategy favoring non-monotone adjustments of the values L_k further improves convergence speed with respect to standard monotone strategies, as already observed in [11] for GFISTA.

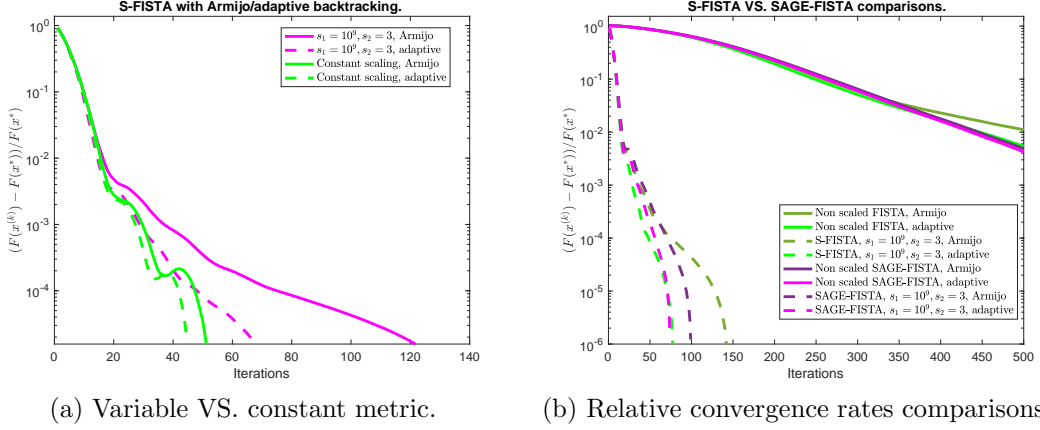(a) Variable VS. constant metric.   (b) Relative convergence rates comparisons.

Figure 3: Figure 3a: S-FISTA with Armijo/adaptive backtracking for variable and constant metric. Figure 3b: comparison between non-scaled and scaled convex (S-FISTA) and strongly-convex (SAGE-FISTA) algorithms with Armijo/adaptive backtracking. Problem (67) on `moon` image.

4.2 Poisson image deblurring with strongly convex TV regularization

As a second example, we consider a TV image deblurring problem for images corrupted by Poisson noise. For this test, we consider the KL divergence functional (55) and combine it with a TV-type regularization term, where a quadratic perturbation weighted by a parameter $\varepsilon \ll 1$ is added to make the composite problem strongly-convex. By further encoding also a non-negativity constraint, the final problem reads

$$x^* = \underset{x \in \mathbb{R}^n}{\operatorname{argmin}} \left\{ F(x) := KL(Hx + b; z) + \lambda TV(x) + \frac{\varepsilon}{2} \|x\|_2^2 + \iota_Y(x) \right\}. \tag{72}$$

Recalling splitting (6), we set here:

$$f(x) := KL(Hx + b; z), \qquad g(x) := \lambda TV(x) + \frac{\varepsilon}{2} \|x\|_2^2 + \iota_Y(x). \tag{73}$$

It is well-known that ∇f is Lipschitz continuous on Y and that an over-estimation of its Lipschitz constant L_f can be provided [21]. Furthermore, the function g is ε-strongly convex. We thus have:

$$\nabla f(x) = H^T e - H^T \left(\frac{z}{Hx + b} \right), \qquad L_f = \frac{\max z_i}{b^2} \max(H^T e) \max(He), \qquad \mu_g = \varepsilon, \tag{74}$$

where $e \in \mathbb{R}^n$ is a vector whose components are all equal to one. Due to (74), we observe that the variability in the estimation of L_f (see Table 1) depends on the range on the data z, which in our examples is intentionally allowed to vary. By assuming reflexive boundary conditions, the matrix-vector products defined in terms of H and H^T can be performed efficiently via discrete cosine transform. Under these choices, we apply the SAGE-FISTA Algorithm 1 with $\mu_f = 0$ and $\mu = \mu_g$ on the `phantom`, `micro` and `mri` images reported in Figure 1. For each image, the choice of model parameters λ and ε is reported in Table 1. As far as the algorithmic parameters are concerned, we set for all the following tests `maxiter`= 300, `max_bt`= 10, $\rho = 0.85$ and

$\delta \in \{1, 0.98\}$ depending on whether Armijo/adaptive backtracking is considered, respectively. As for the previous test, we set $t_0 = 1.01$ and $x^{(0)} = z$. SAGE-FISTA is run with an initial estimate of L_f chosen as $L_0 = 0.1$ for `phantom`, $L_0 = 1$ for `micro` and $L_0 = 100$ for `mri`.

Regarding the sequence of scaling matrices $\{D_k\}_{k \in \mathbb{N}}$, we consider the diagonal split-gradient strategy (65) setting $V(y^{(k)}) = H^T e$. In this case, unlike in (71), the scaling matrix D_k explicitly depends on the extrapolated point $y^{(k)}$.

Concerning the sequence $\{\epsilon_k\}_{k \in \mathbb{N}}$, we apply again the result provided by Corollary 3.1, with the same choice of $\{\theta_k\}_{k \in \mathbb{N}}$ and $\{a_k\}_{k \in \mathbb{N}}$ adopted in the previous test. As far as the inexact computation of the proximal operator of g is concerned, we start noticing that the function g in (73) can be cast in the form (59) with the choices $\phi_i(v) = \lambda \|v\|_2, M_i = \nabla_i, i = 1, \ldots, n$ and $\psi(x) = \frac{\varepsilon}{2}\|x\|_2^2 + \iota_Y(x)$. Due to the presence of the quadratic perturbation, the computation of the sequence $x^{(k,l)}$ in (62) can be performed similarly as in (70), but after a suitable rescaling defined in terms of the operators D_k. In particular, denoting by $\{w^{(k,l)}\}_{l \in \mathbb{N}}$ the dual sequence, we have that the inner primal iterates can be computed as follows:

$$x^{(k,l)} = \text{prox}_{\tau_k \psi}^{D_k} \left(y^{(k)} - \tau_k D_k^{-1} \left(\nabla f(y^{(k)}) + M^* w^{(k,l)} \right) \right) \tag{75}$$

$$= \text{prox}_{\tau_k \iota_Y}^{D_k + \tau_k \varepsilon \mathcal{I}} \left(\left(\frac{D_k}{D_k + \tau_k \varepsilon \mathcal{I}} \right) \left(y^{(k)} - \tau_k D_k^{-1} \left(\nabla f(y^{(k)}) + M^* w^{(k,l)} \right) \right) \right) \tag{76}$$

$$= P_{Y, D_k} \left(\left(\frac{D_k}{D_k + \tau_k \varepsilon \mathcal{I}} \right) \left(y^{(k)} - \tau_k D_k^{-1} \left(\nabla f(y^{(k)}) + M^* w^{(k,l)} \right) \right) \right), \qquad \forall l \in \mathbb{N},$$

where the fraction $D_k/(D_k + \tau_k \varepsilon \mathcal{I})$ has to be intended element-wise. We report the proof of the result that guarantees the equality between (75) and (76) in Lemma A.1 in Appendix A.

We now report the results obtained by the use of SAGE-FISTA with Armijo and adaptive backtracking in correspondence of the images `phantom` (Figure 4) `micro` (Figure 5) and `mri` (Figure 6). For all the experiments, we observe a clear advantage in the use of the variable scaling algorithm in comparison to its non-scaled variant. This is particularly clear from plots (a) in the Figures, where relative error is shown to decay much faster along the iterations. Regarding the computational times (Plots (b)), we notice that, depending on the particular data at hand, the performance may vary, especially depending on the the choice of the parameters s_1 and s_2. Overall, however, we observe that the use of a scaling procedure allows to compute highly accurate solutions in a limited amount of CPU time. This is particularly useful in applications such as medical imaging where the need of fast image analysis strongly benefits from the ability of computing an accurate result in a pre-determined amount of time (usually 5/10 secs). As far as the choice of the backtracking strategy is concerned, we observe from the tests performed on the `micro` and `mri` images that whenever the initial value $L_0 = 1/\tau_0$ is too large, the use of a monotone Armijo strategy does not allow for its adjustment. In this case, even in the presence of an appropriate scaling strategy, convergence speed may suffer. The use of an adaptive backtracking corrects this inconvenience by adjusting the values L_k, which corresponds to considering larger step-sizes τ_k favoring faster convergence. This is particularly evident for problems like (72), where the estimation L_f in (74) significantly overestimates the admissible upper value for which convergence is guaranteed.
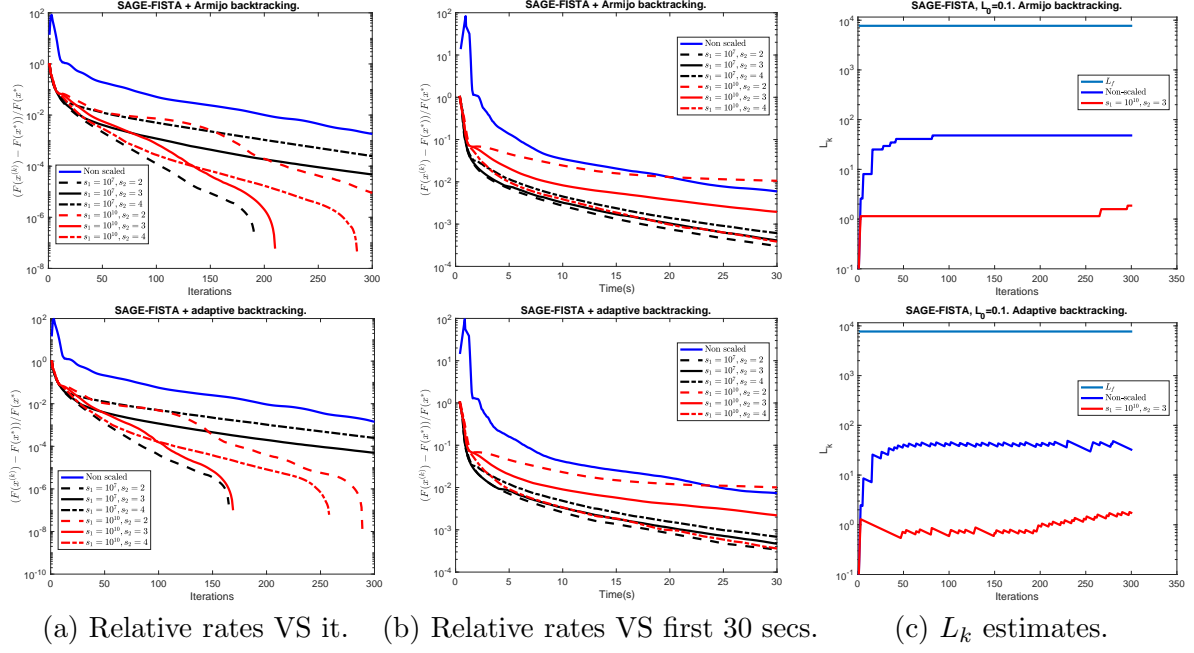
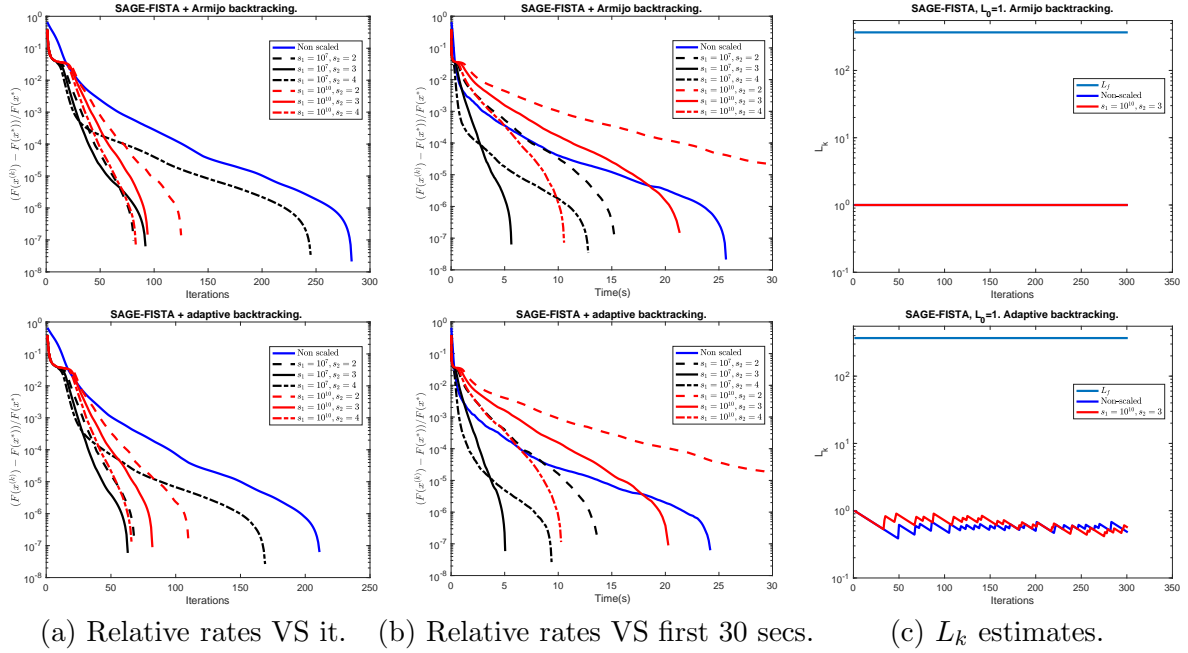(a) Relative rates VS it. (b) Relative rates VS first 30 secs. (c) L_k estimates.

Figure 4: Performance of SAGE-FISTA with Armijo backtracking (top row) and adaptive-backtracking (bottom row) for different choices of s_1 and s_2 for problem (72) on `phantom` image.

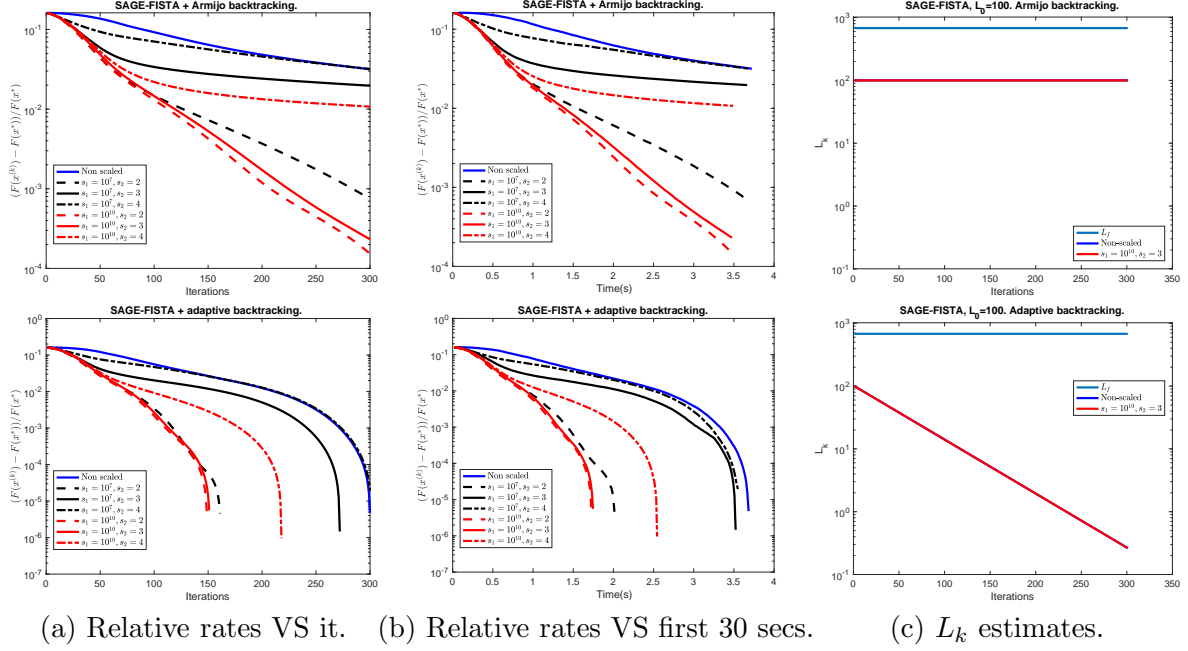(a) Relative rates VS it. (b) Relative rates VS first 30 secs. (c) L_k estimates.

Figure 5: Performance of SAGE-FISTA with Armijo backtracking (top row) and adaptive-backtracking (bottom row) for different choices of s_1 and s_2 for problem (72) on `micro` image.

(a) Relative rates VS it.   (b) Relative rates VS first 30 secs.   (c) L_k estimates.

Figure 6: Performance of SAGE-FISTA with Armijo backtracking (top row) and adaptive-backtracking (bottom row) for different choices of s_1 and s_2 for problem (72) on `mri` image.

5 Conclusions and outlook

We studied a Scaled, inexact and Adaptive GEneralized FISTA (named SAGE-FISTA) algorithm for solving (strongly) convex composite optimization problems via inertial forward-backward splitting. In order to take into account possible errors in the computation of the proximal-gradient points, an appropriate notion of inexactness is introduced and combined with the use of suitable variable-metric operators. To avoid the limitations due to the use of monotone Armijo-type backtracking strategies, we considered an adaptive backtracking strategy that allows for local increasing/decreasing of the algorithmic step-size. A linear convergence result in function values is given, and numerical experiments on some exemplar Poisson image restoration problems are reported. The combination of the scaling procedure with the adaptive backtracking strategy allows to compute much more accurate solutions in shorter computational times, which is of particular interest in real-world scenarios where limited computational resources are available.

Possible generalizations of this work shall consider the use of Bregman distances in the computation of the proximity operator, in the same spirit of [2]. For the estimation of the strong convexity moduli, the use of restarting strategies similar to the ones proposed in [18, 29] should be also explored. Finally, in order to relax the hypothesis on strong convexity, the study of SAGE-FISTA under the proximal Polyak-Łojasiewicz condition [22] is envisaged.

A Computation of perturbed scaled projections

We report a useful Lemma employed in Section 4.2 to define appropriately the primal sequence $\left\{x^{(k,l)}\right\}_{l \in \mathbb{N}}$ in (75) for problem (72). We prove the result under the assumption that, at each iteration $k \geq 1$, the scaling matrix D_k is diagonal and positive definite. Whenever $D_k = \mathcal{I}$, the

identity matrix, the result is a standard property of proximal operators, see, e.g., [14].

Lemma A.1. *Let $h : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be a proper, convex and lower semicontinuous function, $D \in \mathbb{R}^{n \times n}$ a diagonal positive definite matrix and let $\varepsilon > 0$. Then, defining the function $g : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ as $g(x) := h(x) + \frac{\varepsilon}{2}\|x\|_2^2$, there holds:*

$$\mathrm{prox}_{\tau g}^D(z) = \mathrm{prox}_{\tau h}^{D+\tau\varepsilon\mathcal{I}}\left(\left(\frac{D}{D+\tau\varepsilon\mathcal{I}}\right)z\right), \qquad \forall z \in \mathbb{R}^n, \quad \tau > 0, \tag{77}$$

where $\mathcal{I} \in \mathbb{R}^{n \times n}$ is the identity matrix and the fraction symbol is intended element-wise.

Proof. We have the following chain of equalities holding for all $z \in \mathbb{R}^n$:

$$\mathrm{prox}_{\tau g}^D(z) = \underset{y \in \mathbb{R}^n}{\mathrm{argmin}} \; h(y) + \frac{\varepsilon}{2}\|y\|_2^2 + \frac{1}{2\tau}\|y-z\|_D^2$$

$$= \underset{y \in \mathbb{R}^n}{\mathrm{argmin}} \; h(y) + \sum_{i=1}^n \left(\frac{D_{i,i}+\tau\varepsilon}{2\tau}y_i^2 + \frac{D_{i,i}}{2\tau}z_i^2 - \frac{D_{i,i}}{\tau}y_i z_i\right)$$

$$= \underset{y \in \mathbb{R}^n}{\mathrm{argmin}} \; h(y) + \sum_{i=1}^n \left(\frac{D_{i,i}+\tau\varepsilon}{2\tau}y_i^2 + \left(\frac{D_{i,i}^2}{2\tau(D_{i,i}+\tau\varepsilon)} + \frac{\varepsilon D_{i,i}}{2(D_{i,i}+\tau\varepsilon)}\right)z_i^2 - \frac{D_{i,i}}{\tau}y_i z_i\right).$$

By neglecting the constant term $(\varepsilon D_{i,i} z_i^2)/(2(D_{i,i}+\tau\varepsilon))$, we obtain:

$$\mathrm{prox}_{\tau g}^D(z) = \underset{y \in \mathbb{R}^n}{\mathrm{argmin}} \; h(y) + \sum_{i=1}^n \frac{D_{i,i}+\tau\varepsilon}{2\tau}\left(y_i^2 + \left(\frac{D_{i,i}z_i}{D_{i,i}+\tau\varepsilon}\right)^2 - 2y_i\left(\frac{D_{i,i}z_i}{D_{i,i}+\tau\varepsilon}\right)\right)$$

$$= \underset{y \in \mathbb{R}^n}{\mathrm{argmin}} \; h(y) + \frac{1}{2\tau}\left\|y - \left(\frac{D}{D+\tau\varepsilon\mathcal{I}}\right)z\right\|_{D+\tau\varepsilon\mathcal{I}}^2 = \mathrm{prox}_{\tau h}^{D+\tau\varepsilon\mathcal{I}}\left(\left(\frac{D}{D+\tau\varepsilon\mathcal{I}}\right)z\right).$$

$\square$

References

[1] F. Bach, R. Jenatton, J. Mairal, and G. Obozinski. Structured sparsity through convex optimization. *Stat. Sci.*, 27(4):450–468, 2012.

[2] H. H. Bauschke, J. Bolte, and M. Teboulle. A descent lemma beyond Lipschitz gradient continuity: First-order methods revisited and applications. *Math. Oper. Res.*, 4(1):330–348, November 2016.

[3] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.*, 2:183–202, 2009.

[4] M. Bertero, P. Boccacci, and V. Ruggiero. *Inverse Imaging with Poisson Data.* 2053-2563. IOP Publishing, 2018.

[5] S. Bonettini, I. Loris, F. Porta, and M. Prato. Variable metric inexact line–search based methods for nonsmooth optimization. *SIAM J. Optim.*, 26(2):891–921, 2016.

[6] S. Bonettini, F. Porta, M. Prato, S. Rebegoldi, V. Ruggiero, and L. Zanni. *Recent Advances in Variable Metric First-Order Methods*, pages 1–31. Springer International Publishing, Cham, 2019.

[7] S. Bonettini, F. Porta, and V. Ruggiero. A variable metric forward-backward method with extrapolation. *SIAM J. Sci. Comput.*, 38:A2558–A2584, 2016.

[8] S. Bonettini, S. Rebegoldi, and V. Ruggiero. Inertial variable metric techniques for the inexact forward–backward algorithm. *SIAM J. Sci. Comput.*, 40(5):A3180–A3210, 2018.

[9] S. Bonettini, R. Zanella, and L. Zanni. A scaled gradient projection method for constrained image deblurring. *Inverse Probl.*, 25(1), January 2009.

[10] M. Burger, J. Müller, E. Papoutsellis, and C.-B. Schönlieb. Total variation regularisation in measurement and image space for PET reconstruction. *Inverse Probl.*, 30(10), sep 2014.

[11] L. Calatroni and A. Chambolle. Backtracking strategies for accelerated descent methods with smooth composite objectives. *SIAM J. Optim.*, 29(3):1772–1798, 2019.

[12] A. Chambolle and C. Dossal. On the convergence of the iterates of the "Fast Iterative Shrinkage/Thresholding Algorithm". *J. Optim. Theory Appl.*, 166(3):968–982, September 2015.

[13] A. Chambolle and T. Pock. An introduction to continuous optimization for imaging. *Acta Numer.*, 25:161–319, 2016.

[14] P.L. Combettes and J.-C. Pesquet. Proximal splitting methods in signal processing. In H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, editors, *Fixed-point algorithms for inverse problems in science and engineering*, Springer Optimization and Its Applications, pages 185–212. Springer, New York NY, 2011.

[15] P.L. Combettes and B.C. Vũ. Variable metric forward-backward splitting with applications to monotone inclusions in duality. *Optimization*, 63(9):1289–1318, 2014.

[16] P.L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Model. Sim.*, 4(4):1168–1200, 2005.

[17] L. Debnath and P. Mikusiński. *Introduction to Hilbert spaces with applications*. Academic Press, Boston, 1990.

[18] O. Fercoq and Z. Qu. Restarting the accelerated coordinate descent method with a rough strong convexity estimate. *Comput. Optim. Appl.*, 75(1):63–91, 2020.

[19] M. I. Florea and S. A. Vorobyov. An accelerated composite gradient method for large-scale composite objective problems. *IEEE Trans. Signal Process.*, 67(2):444–459, 2019.

[20] M. I. Florea and S. A. Vorobyov. A generalized accelerated composite gradient method: Uniting nesterov's fast gradient method and fista. *IEEE Trans. Signal Process.*, 68:3033–3048, 2020.

[21] Z.T. Harmany, R.F. Marcia, and R.M. Willett. This is SPIRAL-TAP: Sparse Poisson Intensity Reconstruction ALgorithms - Theory and practice. *IEEE Trans. Image Process.*, 21(3):1084–1096, 2012.

[22] H. Karimi, J. Nutini, and M. W. Schmidt. Linear convergence of gradient and proximal–gradient methods under the Polyak-Łojasiewicz condition. In *European Conference on Machine Learning and Knowledge Discovery in Databases*, volume 9851, pages 795–811, 2016.

[23] H. Lantéri, M. Roche, O. Cuevas, and C. Aime. A general method to devise maximum likelihood signal restoration multiplicative algorithms with non-negativity constraints. *Signal Process.*, 81(5):945–974, May 2001.

[24] M. Lazzaretti, L. Calatroni, and C. Estatico. Weighted-CEL0 sparse regularisation for molecule localisation in super-resolution microscopy with Poisson data. *https://arxiv.org/abs/2010.13173*, 2020.

[25] J. J. Moreau. Proximité et dualité dans un espace hilbertien. *Bull. Soc. Math. France*, 93:273–299, 1965.

[26] Y. Nesterov. A method for solving the convex programming problem with convergence rate $O(1/k^2)$. *Soviet Math. Dokl.*, 269:543–547, 1983.

[27] Y. Nesterov. *Introductory lectures on convex optimization: a basic course.* Applied optimization. Kluwer Academic Publ., Boston, Dordrecht, London, 2004.

[28] Y. Nesterov. Gradient methods for minimizing composite functions. *Math. Program.*, 140:125–161, 2013.

[29] B. O'Donoghue and E. Candès. Adaptive restart for accelerated gradient schemes. *FoCM*, 15(3):715–732, 2015.

[30] B. Polyak. *Introduction to Optimization.* Optimization Software - Inc., Publication Division, N.Y., 1987.

[31] A. Sawatzky. *(Nonlocal) Total Variation in Medical Imaging.* PhD thesis, 2011. University of Münster.

[32] K. Scheinberg, D. Goldfarb, and X. Bai. Fast first–order methods for composite convex optimization with backtracking. *Found. Comput. Math.*, 14:389–417, 2014.

[33] M. Schmidt, N. Le Roux, and F. Bach. Convergence rates of inexact proximal-gradient methods for convex optimization. *arXiv:1109.2415v2*, 2011.

[34] S. Villa, S. Salzo, L. Baldassarre, and A. Verri. Accelerated and inexact forward-backward algorithms. *SIAM J. Optim.*, 23(3):1607–1633, 2013.