

This is the peer reviewed version of the following article:

Scaling techniques for  $\epsilon$ -subgradient methods / Bonettini, Silvia; Benfenati, Alessandro; Ruggiero, Valeria. - In: SIAM JOURNAL ON OPTIMIZATION. - ISSN 1052-6234. - 26:3(2016), pp. 1741-1772. [10.1137/14097642X]

*Terms of use:*

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

19/12/2025 04:19

# Scaling techniques for $\epsilon$ -subgradient projection methods\*

S. Bonettini<sup>†</sup>, A. Benfenati<sup>‡</sup>, V. Ruggiero<sup>§</sup>

August 26, 2018

## Abstract

The recent literature on first order methods for smooth optimization shows that significant improvements on the practical convergence behaviour can be achieved with variable stepsize and scaling for the gradient, making this class of algorithms attractive for a variety of relevant applications. In this paper we introduce a variable metric in the context of the  $\epsilon$ -subgradient projection methods for nonsmooth, constrained, convex problems, in combination with two different stepsize selection strategies. We develop the theoretical convergence analysis of the proposed approach and we also discuss practical implementation issues, as the choice of the scaling matrix. In order to illustrate the effectiveness of the method, we consider a specific problem in the image restoration framework and we numerically evaluate the effects of a variable scaling and of the steplength selection strategy on the convergence behaviour.

## 1 Introduction

In this paper we consider the problem

$$\min_{x \in X} f(x) \quad (1)$$

where  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$  is a convex, proper, lower semicontinuous function and  $X$  is a nonempty, closed, convex subset of  $\mathbb{R}^n$  contained in the domain of  $f$ . We denote by  $X^*$  the set of the solutions of (1). We are interested in the case where  $f$  is nondifferentiable and a subgradient or an approximate subgradient of  $f$  can be easily computed. This arises for example in duality and minimax contexts. A well-known method to solve the problem (1) is the  $\epsilon$ -subgradient method:

$$x^{(k+1)} = P_X \left( x^{(k)} - \alpha_k u^{(k)} \right) \quad (2)$$

where  $u^{(k)} \in \partial_{\epsilon_k} f(x^{(k)})$  for some  $\epsilon_k \geq 0$ ,  $\alpha_k$  is a positive stepsize and  $P_X(\cdot)$  is the Euclidean projection operator on the set  $X$ . The choice  $\epsilon_k = 0$  for all  $k$  corresponds to the subgradient

---

\*This work has been partially supported by MIUR (Italian Ministry for University and Research), under the projects FIRB - Futuro in Ricerca 2012, contract RBFR12M3AC, and by the Italian Spinner2013 PhD project *High-complexity inverse problems in biomedical applications and social systems*. The Italian GNCS - INdAM (Gruppo Nazionale per il Calcolo Scientifico - Istituto Nazionale di Alta Matematica) is also acknowledged.

<sup>†</sup>Dipartimento di Matematica e di Informatica, Università di Ferrara, Polo Scientifico Tecnologico, Blocco B, Via Saragat 1, I-44122 Ferrara, Italy ([silvia.bonettini@unife.it](mailto:silvia.bonettini@unife.it)).

<sup>‡</sup>Dipartimento di Matematica e di Informatica, Università di Ferrara, Polo Scientifico Tecnologico, Blocco B, Via Saragat 1, I-44122 Ferrara, Italy ([alessandro.benfenati@unife.it](mailto:alessandro.benfenati@unife.it)).

<sup>§</sup>Dipartimento di Matematica e di Informatica, Università di Ferrara, Polo Scientifico Tecnologico, Blocco B, Via Saragat 1, I-44122 Ferrara, Italy ([valeria.ruggiero@unife.it](mailto:valeria.ruggiero@unife.it)).

method, which has been extensively investigated (see for example the contributions collected in [26, 32, 43, 48]).

The more general case allowing  $\epsilon_k > 0$  was introduced and developed in [26, 43], while more recent convergence results under different assumptions are given in [1, 22, 35, 37, 45, 49]. A typical assumption on the sequence  $\{\epsilon_k\}$  is that

$$\lim_{k \rightarrow \infty} \epsilon_k = 0 \quad (3)$$

and, in this case, the subgradient and the  $\epsilon$ -subgradient methods have very similar convergence properties. In the following discussion we assume that (3) holds.

The  $\epsilon$ -subgradient method is interesting in itself, since, when the projection on  $X$  is easy to compute and an approximate subgradient is available, it can be easily implemented, but it is also a useful tool to analyze the theoretical convergence properties of a variety of algorithms [15, 27, 42, 45].

We can distinguish different variants of the method (2) according to the rule adopted to select the stepsize. We list below the most studied stepsize choices for subgradient methods, which, with minor modifications, could be applied also to the case  $\epsilon_k > 0$ :

( $\mathcal{R}_1$ ) the *constant stepsize* rule  $\alpha_k = \alpha > 0$ ;

( $\mathcal{R}_2$ ) the *Polyak rule*

$$\alpha_k = c_k \frac{f(x^{(k)}) - f^*}{\|u^{(k)}\|^2} \quad \text{or} \quad \alpha_k = c_k \frac{f(x^{(k)}) - f^*}{\max\{1, \|u^{(k)}\|^2\}} \quad c_k \in (0, 2)$$

where  $f^*$  is the minimum of  $f$

( $\mathcal{R}_3$ ) the *Ermoliev or diminishing, divergent series* stepsize rule, which includes any sequences  $\{\alpha_k\}$  such that

$$\alpha_k > 0 \quad \lim_{k \rightarrow \infty} \alpha_k = 0 \quad \sum_{k=0}^{\infty} \alpha_k = \infty \quad (4)$$

( $\mathcal{R}_4$ ) the *diminishing, divergent series, square summable* stepsize rule, which, in addition to (4), also requires  $\sum_{k=0}^{\infty} \alpha_k^2 < \infty$ .

( $\mathcal{R}_5$ ) the *dynamic or adaptive* stepsize rule

$$\alpha_k = \frac{f(x^{(k)}) - f_k}{\|u^{(k)}\|^2} \quad \text{or} \quad \alpha_k = \frac{f(x^{(k)}) - f_k}{\max\{1, \|u^{(k)}\|^2\}} \quad (5)$$

where  $f_k$  is an adaptively computed estimate of  $f^*$ ; several further variants of this rule, which can be considered as an approximation of ( $\mathcal{R}_2$ ) when  $f^*$  is not known, depend on how  $f_k$  is defined.

Keeping a constant stepsize as in ( $\mathcal{R}_1$ ), only the convergence of a subsequence of  $\{f(x^{(k)})\}$  to a possibly suboptimal value is established, i.e.  $\liminf_k f(x^{(k)}) \leq f^* + C\alpha$ , for some positive constant  $C$  [10, 40]; for rules ( $\mathcal{R}_3$ ) and ( $\mathcal{R}_5$ ) stronger results have been proved [32, 37, 40], showing that  $\lim_k f(x^{(k)}) = f^*$  and  $\min_{x^* \in X^*} \|x^* - x^{(k)}\| \rightarrow 0$  (with the assumption  $\epsilon_k = 0$

for the latter case). Finally, the convergence of the sequence  $\{x^{(k)}\}$  to a solution of (1) can be proved in the cases  $(\mathcal{R}_2)$ , with  $\epsilon_k = 0$ , and  $(\mathcal{R}_4)$  [1, 37].

Results about an optimal stepsize choice to obtain a suboptimal rate of convergence for the subgradient method are reported in [41]; in [10, Sec. 6.3], a convergence analysis is performed for different stepsize choices while in [3] analogous results are obtained with respect to non Euclidean metrics.

The key property that the stepsize parameter has to induce on the iterates (2) which is exploited in the standard convergence analysis for subgradient methods is the *quasi-Féjer monotonicity* with respect to the set  $X^*$

$$\|x^{(k+1)} - x^*\|^2 \leq \|x^{(k)} - x^*\|^2 + \eta_k \quad \forall x^* \in X^*$$

for some nonnegative sequence  $\{\eta_k\}$  such that  $\sum \eta_k < \infty$  (see [7, 18]).

It is worth noticing that the stepsize in subgradient methods plays quite a different role than in the smooth case, where an analogous parameter is employed to ensure the sufficient decrease of the objective function, and, in some kind of schemes, also to accelerate the convergence, for example by means of the well known Barzilai–Borwein rules [6, 23, 30] (see also [2, 29] for recent developments in this field). Thus, these valid approaches to the stepsize selection for the smooth case are difficult to extend to the method (2).

On the other side, recent advances in the context of gradient based methods show that introducing a variable scaling matrix for the gradient can lead to significant improvements on the practical performances [5, 16, 12, 34, 50], especially on large scale and ill conditioned problems. Variable metric was introduced also in [20] in the context of monotone operators and in [39, Chapter 5] in the subgradient methods for unconstrained optimization.

Motivated by this, we propose to introduce a variable scaling matrix for the  $\epsilon$ -subgradient vector in the iteration (2). More precisely, the contribution of this paper is to provide the convergence analysis, under standard assumptions, of the following scaled  $\epsilon$ -subgradient scheme:

$$x^{(k+1)} = P_{X, D_k^{-1}} \left( x^{(k)} - \alpha_k D_k u^{(k)} \right) \quad (6)$$

where  $D_k$  is a symmetric positive definite matrix with bounded eigenvalues, the projection operator is defined as

$$P_{X, D_k^{-1}}(x) = \operatorname{argmin}_{z \in X} (x - z) D_k^{-1} (x - z) \quad (7)$$

and  $\alpha_k$  is chosen either as an *a priori* selected sequence obeying to the diminishing, divergent series, summable squares stepsize rule  $(\mathcal{R}_4)$ , or with an adaptive rule  $(\mathcal{R}_5)$  of Brännlund's type [17, 33, 40].

In the first case, assuming that the set  $X^*$  is nonempty, we prove the convergence of the sequence  $\{x^{(k)}\}_{k \in \mathbb{N}}$  to a point  $x^* \in X^*$ , while in latter one we prove the convergence of the sequence  $\{f(x^{(k)})\}_{k \in \mathbb{N}}$  to the minimum value  $f^*$ .

A further contribution of the paper is to introduce, as special case of (6) and also as a generalization of the method in [15], a Scaled Primal–Dual Hybrid Gradient (SPDHG) method, which applies to the case

$$\min_{x \in X} f_0(x) + f_1(Ax) \quad (8)$$

where  $f_0$  and  $f_1$  are convex, proper, lower semicontinuous functions and  $A$  is a linear operator. Many relevant problems can be modeled, as, for example, the restoration of images approached from the Bayesian paradigm [31].

Several methods for the solution of (8) have been developed in the recent literature: in particular, when  $f_0$  is continuously differentiable with Lipschitz continuous gradient, suitable splitting and forward-backward methods can be applied [20, 21, 38]. As we will show in Section 4, problem (8) can be handled also by SPDHG, even when  $f_0$  is non differentiable or its gradient is not Lipschitz continuous on  $\text{dom}(f_0)$ . In particular, we provide an especially tailored implementation of SPDHG for the Total Variation (TV) restoration of images corrupted by Poisson noise. This problem is related to several applications such as astronomical imaging, electronic microscopy, single particle emission computed tomography (SPECT) and positron emission tomography (PET), and a variety of specialized methods have been proposed for its solution (see [4, 14, 24, 25, 28, 47] and references therein). For this special case of SPDHG, we devise an effective strategy to choose the scaling matrix, showing that significant improvements on the practical convergence speed can be obtained.

The paper is organized as follows. In Section 2 we present a convergence analysis for the scaled  $\epsilon$ -subgradient method (6) when the stepsizes are chosen according to a diminishing, divergent series, square summable stepsize rule. At the end of the section our results are also compared to the very recent works [19, 20], where variable metrics are studied from the point of view of more general operators.

Building on this material, in Section 3 we propose a generalization to variable scaling and approximate subgradients of the *level algorithm* in [33, 40], based on a dynamic stepsize selection rule, showing that in our more general settings the main properties still hold. Further, in Section 4 we consider problem (8) and we present the SPDHG method, proving its convergence as special case of an  $\epsilon$ -subgradient scheme. In order to illustrate a practical implementation of SPDHG, in Section 5 we describe the problem of deblurring an image corrupted by Poisson noise via the TV regularization. A suitable scaling for the SPDHG method is discussed and an algorithm for its computation is detailed. In Section 6, we describe some numerical simulations concerning the considered application, with the aim to evaluate the effectiveness of the scaling technique in the  $\epsilon$ -subgradient methods in combination with the two stepsize selection strategies analyzed in the previous sections. The numerical experiments show that a suitable selection of the scaling matrix can be a very effective tool to improve the convergence behaviour also in nonsmooth methods. Finally, some concluding remarks are given in Section 7.

**Notations and definitions.** In the following,  $\|\cdot\|$  denotes the Euclidean vector or matrix norm. Given  $x \in \mathbb{R}^n$  and a symmetric and positive definite matrix  $D$  of order  $n$ ,  $\|x\|_D$  denotes the energy norm, i.e.  $\|x\|_D = \sqrt{x^T D x}$ , and  $P_{X,D}(x) = \arg\min_{y \in X} \|y - x\|_D^2$ . By  $\text{dom}(f)$  we indicate the domain of any function  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ , i.e.  $\text{dom}(f) = \{x \in \mathbb{R}^n : f(x) < \infty\}$ , while  $\text{diam}(X)$  denotes the diameter of the closed, convex set  $X \subset \mathbb{R}^n$ ,  $\text{diam}(X) = \max_{x,z \in X} \|x - z\|$ . The Fenchel dual or conjugate of  $f$  is defined as  $f^*(y) = \sup_{x \in \mathbb{R}^n} x^T y - f(x)$ . Furthermore, we recall that the  $\epsilon$ -subdifferential of  $f$  at  $x \in \text{dom}(f)$  for some  $\epsilon \in \mathbb{R}$ ,  $\epsilon \geq 0$ , is the set

$$\partial_\epsilon f(x) = \{p \in \mathbb{R}^n : f(z) \geq f(x) + p^T(z - x) - \epsilon, \quad \forall z \in \mathbb{R}^n\}$$

Any vector  $p \in \partial_\epsilon f(x)$  is an  $\epsilon$ -subgradient of  $f$  at  $x$  [46, §23].

If  $f(x) = \sum_{i=1}^n \beta_i f_i(x)$ , where  $\beta_i \geq 0$ ,  $u_i \in \partial_{\epsilon_i} f_i(x)$  and  $x \in \bigcap_{i=1}^n \text{dom}(f_i)$ , then  $\sum_{i=1}^n \beta_i u_i \in \partial_\epsilon f(x)$ , where  $\epsilon = \sum_{i=1}^n \epsilon_i$ . The proof of this property can be found in [15, 46].

## 2 Convergence analysis with square summable stepsize sequences

In this section we show that with the rule  $(\mathcal{R}_4)$  the method (6) generates a sequence of points converging to a solution of (1), under standard assumptions on the error sequence  $\epsilon_k$  and the scaling matrices  $D_k$ .

Before giving the main convergence result, we prove the following two useful lemmata concerning the energy norm, the corresponding projection operator and the summable sequences.

LEMMA 2.1 *Let  $D$  be a symmetric positive definite matrix. The following relations hold.*

(i) *For all  $a, b, c \in \mathbb{R}^n$ , we have*

$$\|a - b\|_D^2 + \|b - c\|_D^2 - \|a - c\|_D^2 = 2(b - a)^T D(b - c)$$

(ii) *For any  $x \in \mathbb{R}^n$ ,  $z \in X$  we have*

$$(P_{X,D^{-1}}(x) - x)^T D^{-1}(z - P_{X,D^{-1}}(x)) \geq 0 \quad (9)$$

$$(P_{X,D^{-1}}(x) - x)^T D^{-1}(z - x) \geq 0 \quad (10)$$

(iii) *For any  $x, z \in \mathbb{R}^n$  we have*

$$\|P_{X,D^{-1}}(x) - P_{X,D^{-1}}(z)\|_{D^{-1}} \leq \|x - z\|_{D^{-1}}$$

(iv) *Let  $L$  a positive number such that  $\|D\| = \lambda_{\max}(D) \leq L$  and  $\|D^{-1}\| = \frac{1}{\lambda_{\min}(D)} \leq L$ , then*

$$\|P_{X,D^{-1}}(x) - P_{X,D^{-1}}(z)\| \leq L\|x - z\| \quad (11)$$

*Proof.* Part (i) directly follows from the definition of  $\|\cdot\|_D$ , while the optimality conditions of the minimum problem (7) yield part (ii) (see also [11, Proposition 3.7]). Part (iii) is a consequence of (9) and of the Cauchy–Schwartz inequality. For (iv), see [16, Lemma 2.1].  $\square$

LEMMA 2.2 *Let  $\{L_k\}$  be a sequence of positive numbers such that  $L_k^2 = 1 + \gamma_k$ ,  $\gamma_k \geq 0$ , where  $\sum_{k=0}^{\infty} \gamma_k < \infty$ . Let  $\theta_k = \prod_{j=0}^k L_j^2$  for any  $k \geq 0$ . Then the sequence  $\{\theta_k\}$  is bounded.*

*Proof.* We want to show that there exists a constant  $M > 0$  such that  $\theta_k \leq M$  for all  $k \geq 0$ . By the monotonicity of the logarithm, this is true if and only if  $\log(\theta_k) \leq \log(M)$ . By definition of  $\theta_k$  we have

$$\log(\theta_k) = \sum_{j=0}^k \log(L_j^2) \leq \sum_{j=0}^{\infty} \log(L_j^2) \quad (12)$$

Thus, if the series on the right hand side of (12) converges, the quantities  $\theta_k$  are bounded for all  $k$ . We observe that, since  $L_k^2 = 1 + \gamma_k$ , where  $\gamma_k \rightarrow 0$  as  $k$  diverges, by the known limit  $\lim_{k \rightarrow \infty} \frac{\log(1+\gamma_k)}{\gamma_k} = 1$ , the series  $\sum_{j=0}^{\infty} \log(L_j^2)$  and  $\sum_{j=0}^{\infty} \gamma_j$  have the same behavior. Thus, since by hypothesis the latter one is convergent, the theorem follows.  $\square$

We are now ready to present the main convergence result about the method (6) with diminishing, divergent series, summable square stepsize sequences, whose proof is developed using similar techniques as in [1, Lemma 1] and [37, Theorem 8].

**THEOREM 2.1** *Let  $\{x^{(k)}\} \subset \Omega$  be the sequence generated by iteration (6), where  $u^{(k)} \in \partial_{\epsilon_k} f(x^{(k)})$ , for a given sequence  $\{\epsilon_k\} \subset \mathbb{R}$ ,  $\epsilon_k \geq 0$ . Assume that the set of the solutions of (1)  $X^*$  is nonempty and that there exists a positive constant  $\rho$  such that  $\|u^{(k)}\| \leq \rho$  and a sequence of positive numbers  $\{L_k\}$  such that  $\|D_k\| \leq L_k$ ,  $\|D_k^{-1}\| \leq L_k$ , with  $1 \leq L_k \leq L$  for some positive constant  $L$ , for all  $k \geq 0$ . If the following conditions holds*

$$\lim_{k \rightarrow \infty} \epsilon_k = 0 \quad (13)$$

$$\sum_{k=0}^{\infty} \alpha_k = \infty \quad (14)$$

$$\sum_{k=0}^{\infty} \alpha_k^2 < \infty \quad \sum_{k=0}^{\infty} \epsilon_k \alpha_k < \infty \quad (15)$$

$$L_k^2 = 1 + \gamma_k \quad \sum_{k=0}^{\infty} \gamma_k < \infty \quad (16)$$

then, the sequence  $\{x^{(k)}\}$  converges to a solution of (1).

*Proof.* For all  $k$  let us define  $z^{(k)} = x^{(k)} - \alpha_k D_k u^{(k)}$ . By Lemma 2.1 part (iii) we have that

$$\|x^{(k+1)} - x^{(k)}\|_{D_k^{-1}} \leq \|z^{(k)} - x^{(k)}\|_{D_k^{-1}} \leq L_k^{\frac{1}{2}} \alpha_k \|u^{(k)}\| \leq L_k^{\frac{1}{2}} \alpha_k \rho \quad (17)$$

Then, thanks to Lemma 2.1, part (i), for any  $\tilde{x} \in X^*$  we can write

$$\begin{aligned} L_k \alpha_k^2 \rho^2 &+ \|x^{(k)} - \tilde{x}\|_{D_k^{-1}}^2 - \|x^{(k+1)} - \tilde{x}\|_{D_k^{-1}}^2 \geq \\ &\geq \|x^{(k+1)} - x^{(k)}\|_{D_k^{-1}}^2 + \|x^{(k)} - \tilde{x}\|_{D_k^{-1}}^2 - \|x^{(k+1)} - \tilde{x}\|_{D_k^{-1}}^2 \\ &= 2(x^{(k)} - \tilde{x})^T D_k^{-1} (x^{(k)} - x^{(k+1)}) \\ &= 2(x^{(k)} - \tilde{x})^T D_k^{-1} (x^{(k)} - z^{(k)}) + 2(x^{(k)} - \tilde{x})^T D_k^{-1} (z^{(k)} - x^{(k+1)}) \\ &= 2\alpha_k (x^{(k)} - \tilde{x})^T u^{(k)} + 2(x^{(k)} - z^{(k)})^T D_k^{-1} (z^{(k)} - x^{(k+1)}) + 2(z^{(k)} - \tilde{x})^T D_k^{-1} (z^{(k)} - x^{(k+1)}) \\ &= 2\alpha_k (x^{(k)} - \tilde{x})^T u^{(k)} + 2(x^{(k)} - z^{(k)})^T D_k^{-1} (z^{(k)} - x^{(k+1)}) + \\ &\quad + 2(z^{(k)} - \tilde{x})^T D_k^{-1} (z^{(k)} - P_{X, D_k^{-1}}(z^{(k)})) \\ &\geq 2\alpha_k (x^{(k)} - \tilde{x})^T u^{(k)} + 2(x^{(k)} - z^{(k)})^T D_k^{-1} (z^{(k)} - x^{(k+1)}) \\ &= 2\alpha_k (x^{(k)} - \tilde{x})^T u^{(k)} + 2(x^{(k)} - z^{(k)})^T D_k^{-1} (z^{(k)} - x^{(k)}) + 2(x^{(k)} - z^{(k)})^T D_k^{-1} (x^{(k)} - x^{(k+1)}) \\ &= 2\alpha_k (x^{(k)} - \tilde{x})^T u^{(k)} - 2(x^{(k)} - z^{(k)})^T D_k^{-1} (x^{(k)} - z^{(k)}) + 2\alpha_k (x^{(k)} - x^{(k+1)})^T u^{(k)} \\ &\geq 2\alpha_k (x^{(k)} - \tilde{x})^T u^{(k)} - 2\alpha_k^2 \|u^{(k)}\|^2 \cdot \|D_k\| - 2\alpha_k \|x^{(k)} - x^{(k+1)}\| \cdot \|u^{(k)}\| \\ &\geq 2\alpha_k (x^{(k)} - \tilde{x})^T u^{(k)} - 2L_k \alpha_k^2 \|u^{(k)}\|^2 - 2\alpha_k^2 L_k \|u^{(k)}\|^2 \\ &\geq 2\alpha_k (x^{(k)} - \tilde{x})^T u^{(k)} - 4\alpha_k^2 L_k \rho^2 \\ &\geq 2\alpha_k (f(x^{(k)}) - f(\tilde{x}) - \epsilon_k) - 4\alpha_k^2 L_k \rho^2 \\ &\geq -2\alpha_k \epsilon_k - 4\alpha_k^2 L_k \rho^2 \\ &\geq -2L_k \alpha_k \epsilon_k - 4\alpha_k^2 L_k \rho^2 \end{aligned} \quad (18)$$

where the first inequality follows from (17), while we use Lemma 2.1 (ii) in the second one, the definition of  $z^{(k)}$  and the Cauchy-Schwartz inequality in the third one, the assumption

$\|D_k\| \leq L_k$  and (11) in the fourth one, the bound  $\|u^{(k)}\| \leq \rho$  in the fifth one, the definition of  $\epsilon$ -subgradient in the sixth one, the fact that  $\tilde{x} \in X^*$  in the seventh one and the inequality  $L_k \geq 1$  in the last one.

Upon rearranging terms, this yields

$$\|x^{(k+1)} - \tilde{x}\|_{D_k^{-1}}^2 \leq \|x^{(k)} - \tilde{x}\|_{D_k^{-1}}^2 + 5L_k\alpha_k^2\rho^2 + 2L_k\alpha_k\epsilon_k \quad (19)$$

Since we have

$$\begin{aligned} \|x^{(k+1)} - \tilde{x}\|_{D_k^{-1}}^2 &\geq \lambda_{\min}(D_k^{-1})\|x^{(k+1)} - \tilde{x}\|^2 \geq \frac{1}{L_k}\|x^{(k+1)} - \tilde{x}\|^2 \\ \|x^{(k)} - \tilde{x}\|_{D_k^{-1}}^2 &\leq \lambda_{\max}(D_k^{-1})\|x^{(k)} - \tilde{x}\|^2 \leq L_k\|x^{(k)} - \tilde{x}\|^2 \end{aligned} \quad (20)$$

we can also write

$$\begin{aligned} \|x^{(k+1)} - \tilde{x}\|^2 &\leq L_k^2\|x^{(k)} - \tilde{x}\|^2 + 5L_k^2\alpha_k^2\rho^2 + 2L_k^2\alpha_k\epsilon_k \\ &\leq L_k^2\|x^{(k)} - \tilde{x}\|^2 + \sigma L_k^2\alpha_k^2 + 2L_k^2\alpha_k\epsilon_k \end{aligned}$$

where  $\sigma = 5\rho^2$ . By repeatedly applying the previous inequality we obtain

$$\begin{aligned} \|x^{(k+1)} - \tilde{x}\|^2 &\leq L_k^2\|x^{(k)} - \tilde{x}\|^2 + \sigma L_k^2\alpha_k^2 + 2L_k^2\alpha_k\epsilon_k \\ &\leq L_k^2(L_{k-1}^2\|x^{(k-1)} - \tilde{x}\|^2 + \sigma L_{k-1}^2\alpha_{k-1}^2 + 2L_{k-1}^2\alpha_{k-1}\epsilon_{k-1}) + \sigma L_k^2\alpha_k^2 + 2L_k^2\alpha_k\epsilon_k \\ &\leq \theta_0^k\|x^{(0)} - \tilde{x}\|^2 + \sigma \sum_{j=0}^k \theta_j^k \alpha_j^2 + 2 \sum_{j=0}^k \theta_j^k \alpha_j \epsilon_j \end{aligned} \quad (21)$$

where  $\theta_j^k = \prod_{i=j}^k L_i^2$ ,  $j \leq k$ . Since  $L_i^2 \geq 1$  we have  $1 \leq \theta_j^k \leq \theta_{j-1}^k \leq \theta_0^k$ , which implies

$$\begin{aligned} \|x^{(k+1)} - \tilde{x}\|^2 &\leq \theta_0^k\|x^{(0)} - \tilde{x}\|^2 + \sigma \theta_0^k \sum_{j=0}^k \alpha_j^2 + 2\theta_0^k \sum_{j=0}^k \alpha_j \epsilon_j \\ &\leq M \left( \|x^{(0)} - \tilde{x}\|^2 + \sigma \sum_{j=0}^k \alpha_j^2 + 2 \sum_{j=0}^k \alpha_j \epsilon_j \right) \end{aligned}$$

where the last inequality follows from Lemma 2.2. Thus, by conditions (15), the sequence  $\{x^{(k)}\}$  is bounded. In order to show that  $\{x^{(k)}\}$  converges to a solution of (1), we now consider inequality (18), which, in view of (20), results in

$$\begin{aligned} \|x^{(k+1)} - \tilde{x}\|^2 &\leq L_k^2\|x^{(k)} - \tilde{x}\|^2 + L_k^8\alpha_k^2\rho^2 + 4\alpha_k^2L_k^4\rho^2 - 2\alpha_kL_k(x^{(k)} - \tilde{x})^T u^{(k)} \\ &\leq L_k^2\|x^{(k)} - \tilde{x}\|^2 + \sigma L_k^2\alpha_k^2 + 2\alpha_kL_k(\tilde{x} - x^{(k)})^T u^{(k)} \end{aligned}$$

By repeatedly applying the previous inequality we obtain

$$\|x^{(k+1)} - \tilde{x}\|^2 \leq \theta_0^k\|x^{(0)} - \tilde{x}\|^2 + \sigma \sum_{j=0}^k \theta_j^k \alpha_j^2 + 2 \sum_{j=0}^k \tilde{\theta}_j^k \alpha_j (\tilde{x} - x^{(j)})^T u^{(j)} \quad (22)$$

where  $\tilde{\theta}_j^k = \theta_j^k/L_j$ . Since  $\tilde{x} \in X^*$  and  $u^{(j)} \in \partial_{\epsilon_k} f(x^{(j)})$ , for all  $j \geq 0$  we have

$$f(x^{(j)}) \geq f(\tilde{x}) \geq f(x^{(j)}) + (\tilde{x} - x^{(j)})^T u^{(j)} - \epsilon_j \quad (23)$$



Hence,  $(\tilde{x} - x^{(j)})^T u^{(j)} \leq \epsilon_j$  for all  $j \geq 0$ .

Now, we show that  $(\tilde{x} - x^{(j)})^T u^{(j)} \rightarrow 0$  for  $j \rightarrow \infty$ . To this end, assume by contradiction that  $(\tilde{x} - x^{(j)})^T u^{(j)} < -\epsilon$  for some  $\epsilon > 0$ . Then, by inequality (22) we obtain

$$\|x^{(k+1)} - \tilde{x}\|^2 \leq M\|x^{(0)} - \tilde{x}\|^2 + \sigma M \sum_{j=0}^k \alpha_k^2 - 2\frac{\epsilon}{L} \sum_{j=0}^k \alpha_j$$

where we applied the inequalities  $1 \leq \theta_j^k \leq \theta_{j-1}^k \leq \theta_0^k \leq M$ ,  $1 \leq L_k \leq L$ . Then, taking limits for  $k \rightarrow \infty$ , by assumption (14) we have an absurdum. Thus, there exists a subsequence  $\{x^{(k_i)}\}$  such that  $\lim_i (\tilde{x} - x^{(k_i)})^T u^{(k_i)} = 0$ . Then, from (23) and the assumption (15) it follows that  $\lim_i f(x^{(k_i)}) = f^*$ . Since  $\{x^{(k)}\}$  is bounded,  $\{x^{(k_i)}\}$  is also bounded and, thus it has an accumulation point  $x^\infty$ . By the continuity of  $f(x)$  and by inequality (23) we can conclude that  $x^\infty \in X^*$ .

Now we show that the whole sequence  $\{x^{(k)}\}$  converges to  $x^\infty$ . Let  $\delta > 0$ ; since  $x^\infty$  is an accumulation point of  $\{x^{(k)}\}$  and from (15), there exists a positive integer  $m_\delta$  such that  $\|x^\infty - x^{(m_\delta)}\|^2 \leq \delta/(3M)$ ,  $\sum_{j=m_\delta}^\infty \alpha_k^2 < \delta/(3\sigma M)$  and  $\sum_{j=m_\delta}^\infty \alpha_k \epsilon_k < \delta/(6M)$ . Then, for any  $k > m_\delta$ , using the same arguments as in (21), we obtain

$$\begin{aligned} \|x^{(k)} - x^\infty\|^2 &\leq \theta_{m_\delta}^{k-1} \|x^{(m_\delta)} - x^\infty\|^2 + \sigma \sum_{j=m_\delta}^{k-1} \theta_j^{k-1} \alpha_j^2 + 2 \sum_{j=m_\delta}^{k-1} \theta_j^{k-1} \alpha_j \epsilon_j \\ &\leq M \|x^{(m_\delta)} - x^\infty\|^2 + \sigma M \sum_{j=m_\delta}^\infty \alpha_j^2 + 2M \sum_{j=m_\delta}^\infty \alpha_j \epsilon_j \\ &\leq \delta \end{aligned}$$

Since  $\delta$  can be chosen arbitrarily small, then  $\{x^{(k)}\}$  converges to  $x^\infty$ .  $\square$

We now discuss some further issues about Theorem 2.1, relating our results with the recent literature, in particular with the papers [19, 20] and [39].

**Remark 1.** From equation (19), observing that

$$\begin{aligned} \|x^{(k+1)} - \tilde{x}\|_{D_k^{-1}}^2 &\geq \lambda_{\min}(D_k^{-1}) \|x^{(k+1)} - \tilde{x}\|^2 \\ &= \frac{\lambda_{\min}(D_k^{-1})}{\lambda_{\max}(D_{k+1}^{-1})} \lambda_{\max}(D_{k+1}^{-1}) \|x^{(k+1)} - \tilde{x}\|^2 \\ &\geq \lambda_{\min}(D_k^{-1}) \lambda_{\min}(D_{k+1}) \|x^{(k+1)} - \tilde{x}\|_{D_{k+1}^{-1}}^2 \\ &\geq \frac{1}{L_k L_{k+1}} \|x^{(k+1)} - \tilde{x}\|_{D_{k+1}^{-1}}^2 \end{aligned}$$

we obtain

$$\|x^{(k+1)} - \tilde{x}\|_{D_{k+1}^{-1}}^2 \leq \zeta_k \|x^{(k)} - \tilde{x}\|_{D_k^{-1}}^2 + \xi \zeta_k \alpha_k^2 + 2L \zeta_k \alpha_k \epsilon_k$$

where  $\zeta_k = \sqrt{(1 + \gamma_k)(1 + \gamma_{k+1})}$  and  $\xi = 5L\rho^2$ . By the assumptions made on  $\{\gamma_k\}$ , the sequence  $\{\zeta_k\}$  is bounded. We can also set  $\zeta_k = 1 + \eta_k$ , with  $\eta_k = \sqrt{(1 + \gamma_k)(1 + \gamma_{k+1})} - 1$ , and observe that the series  $\sum \eta_k$  and  $\sum \gamma_k$  have the same behaviour, thanks to the limit  $\lim_{z \rightarrow 0} (\sqrt{1+z} - 1)/z = 1/2$ . Then, since the assumption (16), we can conclude that  $\sum \eta_k$  is a convergent series.

Thus, the sequence  $\{x^{(k)}\}_{k \in \mathbb{N}}$  is *quasi-Fejér monotone* with respect to  $X^*$  relative to  $\{D_k^{-1}\}$ , in the sense of [19, Definition 3.1] and we could apply Proposition 3.2 in [19] (see also [20]) to obtain that  $\{\|x^{(k)} - \tilde{x}\|_{D_k^{-1}}\}_{k \in \mathbb{N}}$  converges and, thus,  $\{x^{(k)}\}_{k \in \mathbb{N}}$  is bounded.

However, the techniques we used in the proof of Theorem 2.1 to show the boundedness of  $\{x^{(k)}\}$  allow to prove the convergence of the whole sequence  $\{x^{(k)}\}$  to a solution of (1) and are employed also in the following section.

Variable metric was introduced also in [39, Chapter 5] in the context of subgradient methods for unconstrained problems (i.e.  $X = \mathbb{R}^n$ ). In this case, setting  $D_k = B_k B_k^T$ , the scaling matrices are assumed to satisfy  $\|B_{k+1}^{-1} B_k\| \geq 1$  and  $\prod_{k=0}^{\infty} \|B_{k+1}^{-1} B_k\|^2 < \infty$ . Even if the second condition is verified under the assumptions of Theorem 2.1, we observe that the requirement  $\|B_{k+1}^{-1} B_k\| \geq 1$  restricts the choice of the scaling matrix, strictly connecting the metrics adopted in two successive iterates.

**Remark 2.** One of the crucial assumptions of Theorem 2.1 concerns the boundedness of the sequence  $\{u^{(k)}\}$ ,  $u^{(k)} \in \partial_{\epsilon_k} f(x^{(k)})$ . This assumption is satisfied for example when  $\text{diam}(\text{dom}(f^*))$  is finite. Indeed, it holds  $\text{dom}(f^*) = \bigcup_{x \in \mathbb{R}^n} \partial_{\epsilon} f(x)$ , for every  $\epsilon > 0$  (see Remark 2 in [51]). Moreover, if  $\text{diam}(\text{dom}(f^*)) = M$  for some  $M > 0$ , by definition of  $\epsilon$ -subdifferential we obtain  $|f(x) - f(z)| \leq M\|x - z\| + \epsilon$ . Since  $\epsilon$  is arbitrary, it follows that  $f$  is Lipschitz continuous with constant  $M$ .

**Remark 3.** The convergence analysis of Theorem 2.1 applies also to further variants of the iteration (6), as for example the following one:

$$x^{(k+1)} = P_{X, D_k^{-1}} \left( x^{(k)} - \frac{\alpha_k}{\max(1, \|u^{(k)}\|_{D_k})} D_k u^{(k)} \right) \quad (24)$$

The convergence of the method (24) can be analyzed as follows.

**COROLLARY 2.1** *Let  $\{x^{(k)}\}_{k \in \mathbb{N}}$  a sequence satisfying (24) such that the assumptions of Theorem 2.1 are satisfied. Then,  $\{x^{(k)}\}_{k \in \mathbb{N}}$  converges to a solution of problem (1).*

*Proof.* Let us define  $\bar{\alpha}_k = \frac{\alpha_k}{\max(1, \|u^{(k)}\|_{D_k})}$ . From the inequalities

$$1 \leq \max(1, \|u^{(k)}\|_{D_k}) \leq \max(L^{\frac{1}{2}} \rho, 1)$$

it follows that

$$\begin{aligned} \sum \bar{\alpha}_k &= \sum \frac{\alpha_k}{\max(1, \|u^{(k)}\|_{D_k})} \geq \frac{1}{\max(L^{\frac{1}{2}} \rho, 1)} \sum \alpha_k = \infty \\ \sum \bar{\alpha}_k^2 &= \sum \frac{\alpha_k^2}{\max(1, \|u^{(k)}\|_{D_k})^2} \leq \sum \alpha_k^2 < \infty \\ \sum \bar{\alpha}_k \epsilon_k &= \sum \frac{\alpha_k}{\max(1, \|u^{(k)}\|_{D_k})} \epsilon_k \leq \sum \alpha_k \epsilon_k < \infty \end{aligned}$$

Then, we can invoke Theorem 2.1 to obtain the result.  $\square$

The results in this Section can be exploited for the practical implementation of the methods (6) and (24), since they indicate how to choose the sequence  $\{\alpha_k\}$ , and they are employed also in the convergence analysis of (24) equipped with an adaptive stepsize rule, as the one described in the following.

### 3 Convergence analysis with dynamic stepsize rule

A critical point for the implementation of the methods (6) and (24) is how to select the sequence  $\{\alpha_k\}$ ; a practical strategy to obtain good performances is still an open problem, since they are in general quite sensitive to this choice [15]. Borrowing the ideas of [17] and [33], in this section we describe a *level algorithm* that allows to adaptively compute a stepsize  $\alpha_k$  of the form (5) in the iteration (24). In our scheme we introduce the use of the  $\epsilon$ -subgradient of  $f$  at the current iterate (instead of the subgradient) and a variable metric. The resulting algorithm is detailed in Algorithm 1, whose underlying assumption is that, for any given  $\epsilon_k \geq 0$ , we are able to provide an element  $u^{(k)}$  of the set  $\partial_{\epsilon_k} f(x^{(k)})$ . In Algorithm 1, we have  $f_k^{rec} = \min_{i=0,\dots,k} f(x^{(i)})$ , while

---

**Algorithm 1** Scaled  $\epsilon$ -Subgradient Level Algorithm (SSL)

---

Choose  $B > 0$ ,  $\nu_1, \nu_2 \in (0, 1)$ ,  $f_{-1}^{rec} = \infty$ ;  $k = 0$ ,  $l = 0$ ,  $k(l) = 0$ ,  $\delta_0 > 0$ ; choose  $x^{(0)} \in X$ .

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

STEP 1. Computation of  $f(x^{(k)})$

STEP 2. If  $f(x^{(k)}) < f_{k-1}^{rec}$ , then  $f_k^{rec} = f(x^{(k)})$  else  $f_k^{rec} = f_{k-1}^{rec}$

STEP 3. If  $f(x^{(k)}) < f_{k(l)}^{rec} - \nu_1 \delta_l$ , then  $k(l+1) = k$ ,  $\sigma_k = 0$ ,  $\delta_{l+1} = \delta_l$ ,  $l = l + 1$  and go to Step 5.

STEP 4. If  $\sigma_k > B$ , then  $k(l+1) = k$ ,  $\sigma_k = 0$ ,  $\delta_{l+1} = \nu_2 \delta_l$ ,  $l = l + 1$ .

STEP 5. Set  $f_k^{lev} = f_{k(l)}^{rec} - \delta_l$

STEP 6. Update the stepsize and compute the new iterate

$$\alpha_k = \frac{f(x^{(k)}) - f_k^{lev}}{\max(1, \|u^{(k)}\|_{D_k})}$$

$$x^{(k+1)} = P_{X, D_k^{-1}} \left( x^{(k)} - \alpha_k D_k \frac{u^{(k)}}{\max(1, \|u^{(k)}\|_{D_k})} \right) \quad (25)$$

STEP 7.  $\sigma_{k+1} = \sigma_k + \alpha_k$  and go to Step 1.

END

---

$l$  is the number of times that the value  $f^{lev}$  has been updated and  $k(l)$  is the iteration where the  $l$ -th updating occurred. Finally,  $\sigma_k$  is the cumulative path length between two successive updates of  $f^{lev}$ .

Steps 2-5 aim to provide in  $f_k^{lev}$  an estimate of the optimal function value at the iterate  $k$ , which is used as target level for the successive iterates until the objective function value is sufficiently close to it or the iterates move through a long path without approaching it. In the first case, i.e. when the inequality at Step 3 is satisfied,  $f_k^{lev}$  is reduced at Step 5 by subtracting the positive quantity  $\delta_l$  to the best value obtained so far,  $f_k^{rec}$ . In the other case, when the inequality at Step 4 is satisfied, the estimated difference from the optimal value  $\delta_l$  is reduced and, as a consequence of Step 5, the target level  $f_k^{lev}$  is increased.

One of the main differences between the stepsize computed by Algorithm 1 and the square summable sequence considered in the previous section is that the former one does not necessarily

converge to zero.

In the rest of this section we prove that the sequence  $\{f(x^{(k)})\}$ , where  $x^{(k)}$  is computed by Algorithm 1, converges to the minimum of  $f$ , using similar techniques as in [40]. Before giving the main result, whose proof also exploits the results in Section 2, we recall the following technical lemma. We omit the proof, since it runs as that of Lemma 2.2 in [40].

**LEMMA 3.1** *Assume that the set  $X^*$  of the solutions of (1) is nonempty. Assume that there exists a positive constant  $\rho$  such that  $\|u^{(k)}\| \leq \rho$ ,  $u^{(k)} \in \partial_{\epsilon_k} f(x^{(k)})$ , and a sequence of positive numbers  $\{L_k\}$  such that  $\|D_k\| \leq L_k$ ,  $\|D_k^{-1}\| \leq L_k$ , with  $1 \leq L_k \leq L$  for some positive constant  $L$ , for all  $k \geq 0$ . Given  $B > 0$  and  $\{\epsilon_k\}$  such that  $\epsilon_k \rightarrow 0$  as  $k \rightarrow \infty$  in Algorithm 1, we have  $l \rightarrow \infty$  and  $\delta_l \rightarrow 0$  as  $l \rightarrow \infty$ .*

The following Theorem can be considered as a generalization of Proposition 2.7 in [40], which only deals with the case  $D_k = I$ ,  $\epsilon_k = 0$  for all  $k$ .

**THEOREM 3.1** *Under the same assumptions of the previous Lemma, if (16) holds, for SSL we have  $\tilde{f} = \inf_{k \geq 0} f(x^{(k)}) = f(x^*)$ , with  $x^* \in X^*$ .*

*Proof.* The first part of the proof aims to show that  $\sum_j \alpha_j = \infty$  and runs as Proposition 2.7 in [40]. For sake of completeness, we report below the detailed derivation of the result.

From Lemma 3.1,  $\delta_l \rightarrow 0$  as  $l \rightarrow \infty$ . Let  $S$  be given by  $S = \{l \in \{1, 2, \dots\}, \delta_l = \nu_2 \delta_{l-1}\}$ . Then, from Step 4 and 6 of Algorithm 1, we obtain

$$\sigma_k = \sigma_{k-1} + \alpha_{k-1} = \sum_{j=k(l)}^{k-1} \alpha_j$$

so that  $k(l+1) = k$  and  $l+1 \in S$  whenever  $\sum_{j=k(l)}^{k-1} \alpha_j > B$  at Step 4. Hence

$$\sum_{j=k(l-1)}^{k(l)-1} \alpha_j > B \quad \forall l \in S$$

and since the cardinality of  $S$  is infinite, we have

$$\sum_{k=k(l)}^{\infty} \alpha_k \geq \sum_{l \geq \bar{l}, l \in L} \sum_{j=k(l-1)}^{k(l)-1} \alpha_j > \sum_{l \geq \bar{l}, l \in S} B = \infty \quad (26)$$

Now in order to obtain a contradiction, assume that  $\tilde{f} > f(x^*)$ , so that for some  $\tilde{y} \in X$  and some  $\eta > 0$

$$\tilde{f} - \eta \geq f(\tilde{y}) \quad (27)$$

Since  $\delta_l \rightarrow 0$  and  $\epsilon_k \rightarrow 0$ , there are large enough  $\bar{l}$  and  $\bar{k}$  such that, for all  $l \geq \bar{l}$  and  $k \geq \bar{k}$ , we have  $\delta_l < \eta/2$  and  $\epsilon_k < \eta/2$ ; then for all  $k \geq \tilde{k} = \max(k(\bar{l}), \bar{k})$

$$f_k^{lev} - \epsilon_k = f_{k(l)}^{rec} - \epsilon_k - \delta_l > \tilde{f} - \eta \geq f(\tilde{y})$$

From this inequality, by Lemma 2.1 (iii), the definition of  $\epsilon$ -subgradient, the definition of  $\alpha_k$  we obtain

$$\begin{aligned} \|x^{(k+1)} - \tilde{y}\|_{D_k^{-1}}^2 &\leq \|x^{(k)} - \tilde{y}\|_{D_k^{-1}}^2 - 2\frac{\alpha_k}{\max(1, \|u^{(k)}\|_{D_k})}(f(x^{(k)}) - f(\tilde{y}) - \epsilon_k) + \alpha_k^2 \\ &\leq \|x^{(k)} - \tilde{y}\|_{D_k^{-1}}^2 - 2\frac{\alpha_k}{\max(1, \|u^{(k)}\|_{D_k})}(f(x^{(k)}) - f_k^{lev}) + \alpha_k^2 \\ &\leq \|x^{(k)} - \tilde{y}\|_{D_k^{-1}}^2 - \alpha_k^2 \end{aligned}$$

In view of (20) with  $\tilde{x} = \tilde{y}$  and  $L_k \geq 1$ , we can write

$$\|x^{(k+1)} - \tilde{y}\|^2 \leq L_k^2 \|x^{(k)} - \tilde{y}\|^2 - \alpha_k^2 \quad (28)$$

By repeatedly applying the previous inequality we obtain

$$\|x^{(k+1)} - \tilde{y}\|^2 \leq \theta_{\tilde{k}}^k \|x^{(\tilde{k})} - \tilde{y}\|^2 - \sum_{j=\tilde{k}}^k \theta_{j+1}^k \alpha_j^2$$

where  $\theta_j^k = L_j^2 \cdot \dots \cdot L_k^2$ ; since  $\theta_j^k \leq \theta_0^k \leq M$ , where  $M$  is a positive constant (see Lemma 2.2) and  $\theta_j^k \geq 1$  we have

$$\sum_{\tilde{k}}^{\infty} \alpha_j^2 \leq M \|x^{(\tilde{k})} - \tilde{y}\|^2$$

and consequently  $\sum_{k=\tilde{k}}^{\infty} \alpha_k^2 < \infty$ . Then  $\alpha_k \rightarrow 0$  as  $k \rightarrow \infty$  and, from (26),  $\sum_{k=\tilde{k}}^{\infty} \alpha_k = \infty$ . Now we show that  $\sum \alpha_k \epsilon_k < \infty$ . Indeed, since  $\epsilon_k \rightarrow 0$  as  $k \rightarrow \infty$ , there exists  $\bar{k}$  such that  $2\epsilon_k < \eta$  for  $k \geq \bar{k}$ , where  $\eta$  is such that (27) holds. We consider the inequality

$$\|x^{(k+1)} - \tilde{y}\|_{D_k^{-1}}^2 \leq \|x^{(k)} - \tilde{y}\|_{D_k^{-1}}^2 + \alpha_k^2 - 2\frac{\alpha_k}{\max(1, \|u^{(k)}\|_{D_k})} u^{(k)T} (x^{(k)} - \tilde{y}) \quad (29)$$

For the convexity of  $f$ , the inequality (27) and inequality  $2\epsilon_k < \eta$ , we have

$$f(x^{(k)}) + u^{(k)T} (\tilde{y} - x^{(k)}) - \epsilon_k \leq f(\tilde{y}) \leq \tilde{f} - \eta \leq f(x^{(k)}) - 2\epsilon_k$$

Then we have

$$u^{(k)T} (\tilde{y} - x^{(k)}) \leq -\epsilon_k$$

Using this inequality in (29), we obtain

$$\|x^{(k+1)} - \tilde{y}\|_{D_k^{-1}}^2 \leq \|x^{(k)} - \tilde{y}\|_{D_k^{-1}}^2 + \alpha_k^2 - 2\frac{\alpha_k \epsilon_k}{\max(1, \|u^{(k)}\|_{D_k})}$$

Using the same arguments as above, we obtain

$$\|x^{(k+1)} - \tilde{y}\|^2 \leq L_k^2 \|x^{(k)} - \tilde{y}\|^2 + L_k^2 \alpha_k^2 - 2\frac{\alpha_k \epsilon_k}{\max(1, L^{\frac{1}{2}} \rho)}$$

By repeatedly applying the previous inequality we have

$$\begin{aligned} \|x^{(k+1)} - \tilde{y}\|^2 &\leq \theta_k^k \|x^{(\bar{k})} - \tilde{y}\|^2 + \theta_k^k \sum_{j=\bar{k}}^k \alpha_j^2 - \frac{2}{\max(1, L^{\frac{1}{2}}\rho)} \sum_{j=\bar{k}}^k \alpha_j \epsilon_j \\ &\leq M \left( \|x^{(\bar{k})} - \tilde{y}\|^2 + \sum_{j=\bar{k}}^k \alpha_j^2 \right) - \frac{2}{\max(1, L^{\frac{1}{2}}\rho)} \sum_{j=\bar{k}}^k \alpha_j \epsilon_j \end{aligned}$$

Then we have

$$\sum_{j=\bar{k}}^{\infty} \alpha_j \epsilon_j \leq \frac{M}{2} \max(1, L^{\frac{1}{2}}C) \left( \|x^{(\bar{k})} - \tilde{y}\|^2 + \sum_{j=\bar{k}}^{\infty} \alpha_j^2 \right) < \infty$$

According to Corollary 2.1 we have  $\tilde{f} = f^*$  that contradicts (27).  $\square$

Further generalizations of Algorithm 1 could be included in the analysis of the previous theorem following [40, p.122], where the authors suggest some modifications of Steps 2, 3 and 4 allowing a variable path bound  $B$  and different strategies to update the parameter  $\delta_l$ . For sake of simplicity we omit here these details.

## 4 A Scaled Primal–Dual Hybrid Gradient Method

The aim of this Section is to present a concrete example of the method (6) for the problem (8), which for convenience is reproduced below,

$$\min_{x \in X} f(x) \equiv f_0(x) + f_1(Ax) \quad (30)$$

where  $A \in \mathbb{R}^{m \times n}$ ,  $f_0(x), f_1(x)$  are convex, proper, lower semicontinuous functions such that  $\text{diam}(\text{dom}(f_1^*))$  is finite and  $f_1^*(y)$  is the Fenchel dual of  $f_1$ .

We propose the following Scaled Primal–Dual Hybrid Gradient (SPDHG) method for the solution of (30)

$$y^{(k+1)} = (I + \tau_k \partial f_1^*)^{-1}(y^{(k)} + \tau_k A x^{(k)}) \quad (31)$$

$$x^{(k+1)} = P_{X, D_k^{-1}}(x^{(k)} - \alpha_k D_k(w^{(k)} + A^T y^{(k+1)})) \quad (32)$$

where  $w^{(k)} \in \partial_{\delta_k} f_0(x^{(k)})$ , for some  $\delta_k \geq 0$ ,  $\{\tau_k\}$ ,  $\{\alpha_k\}$  are the dual and primal steplength sequences respectively and  $(I + \tau_k \partial f_1^*)^{-1}$  is the resolvent operator [46] of  $f_1^*$  defined as

$$(I + \tau_k \partial f_1^*)^{-1}(x) = \arg \min_z f_1^*(z) + \frac{1}{2\tau_k} \|z - x\|^2.$$

The basic result allowing to consider the method (31)–(32) as a special case of a scaled  $\epsilon$ -subgradient method (6) is the following Lemma (see [15, Lemma 1]).

**LEMMA 4.1** *Let  $y^{(k+1)}$  defined as in (31). Then,  $y^{(k+1)} \in \text{dom}(f_1^*)$  and, thus,  $A^T y^{(k+1)} \in \partial_{\sigma_k}(f_1 \circ A)(x^{(k)})$ , where  $\sigma_k = f_1(Ax^{(k)}) + f_1^*(y^{(k+1)}) - y^{(k+1)T} Ax^{(k)}$ . Moreover, if there exists a positive number  $D$  such that  $\text{diam}(\text{dom}(f_1^*)) \leq D$ , then  $\sigma_k \leq (2\tau_k)^{-1} D^2$ .*

Thus, recalling the additivity of the  $\epsilon$ -subgradient, we can conclude that  $u^{(k)} = w^{(k)} + A^T y^{(k+1)} \in \partial_{\epsilon_k} f(x^{(k)})$ , where  $\epsilon_k = \delta_k + \sigma_k$ .

Motivated by the previous observation, building on the material developed in Sections 2 and 3, we discuss two stepsize selection strategies for the method (31)–(32), providing two different SPDHG implementations. In the first case, we assume that  $\{\tau_k\}$ ,  $\{\alpha_k\}$ ,  $\{L_k\}$  are prefixed sequences. The following Corollary, which is a consequence of Lemma 4.1 and of Theorem 2.1, gives conditions on these sequences ensuring the convergence of the sequence  $\{x^{(k)}\}$  generated by SPDHG to a point  $x^* \in X^*$ .

**COROLLARY 4.1** *Let  $\{x^{(k)}\}$  be the sequence generated by iteration (31)–(32). Assume that  $w^{(k)} \in \partial_{\delta_k} f_0(x^{(k)})$  and that there exists  $\rho > 0$  such that  $\|w^{(k)}\| \leq \rho$  for all  $k$ . Let the steplength sequences  $\{\tau_k\}$ ,  $\{\alpha_k\}$  and the scaling matrix bounds  $\{L_k\}$  satisfy*

$$\alpha_k = \mathcal{O}\left(\frac{1}{k^p}\right), \quad \tau_k = \mathcal{O}(k^p), \quad L_k = \sqrt{1 + \mathcal{O}\left(\frac{1}{k^q}\right)} \quad \frac{1}{2} < p \leq 1, \quad q > 1. \quad (33)$$

Moreover, assume that  $\delta_k$  converges to zero at least as  $\frac{1}{\tau_k}$ . If the set of the solutions of (1) is nonempty and  $\text{diam}(\text{dom}(f_1^*))$  is finite, then, the sequence  $\{x^{(k)}\}$  converges to a solution of (1).

*Proof.* As observed above, we have  $u^{(k)} = w^{(k)} + A^T y^{(k+1)} \in \partial_{\epsilon_k} f(x^{(k)})$ , where  $\epsilon_k = \delta_k + \sigma_k$ . Since  $\text{diam}(\text{dom}(f_1^*))$  is finite, we can apply Lemma 4.1 obtaining  $\sigma_k \leq (2\tau_k)^{-1} D^2$ . By the assumption (33) on  $\tau_k$  and on  $\delta_k$  we obtain that  $\epsilon_k = \mathcal{O}\left(\frac{1}{k^p}\right)$  and, as a consequence,  $\alpha_k \epsilon_k = \mathcal{O}\left(\frac{1}{k^{2p}}\right)$ . Since  $\frac{1}{2} < p \leq 1$  and  $q > 1$ , all assumptions (13)–(16) of Theorem 2.1 are satisfied and we obtain the result.  $\square$

On the other side, the SSL procedure for dynamically computing the primal stepsize  $\alpha_k$  can be also implemented. In this case, we have to provide a sequence  $\{L_k = \sqrt{1 + \gamma_k}\}$  such that  $\sum \gamma_k < \infty$ . For sake of simplicity we assume  $\delta_k = 0$ .

In this case the value  $\epsilon_k = f_1(Ax^{(k)}) + f_1^*(y^{(k+1)}) - y^{(k+1)T} Ax^{(k)}$  is controlled by the dual stepsize  $\tau_k$  (see Lemma 4.1). Then it is possible to compute  $\alpha_k$  by Steps 2–5 of Algorithm 1 and the next iterate  $x^{(k+1)}$ .

Finally, Theorem 3.1 allows to conclude that the sequence  $\{f(x^{(k)})\}$  generated by SPDHG method combined with SSL algorithm converges to the minimum of  $f(x)$  in (30), as stated in the following Corollary under the same boundedness assumptions as in Corollary 4.1.

**COROLLARY 4.2** *Let  $\{x^{(k)}\}$  be the sequence generated by Algorithm 1. Here  $u^{(k)} = w^{(k)} + A^T y^{(k+1)}$  in (25),  $w^{(k)} \in \partial f_0(x^{(k)})$  and  $y^{(k+1)}$  is computed as in (31). Assume that  $\lim_{k \rightarrow \infty} \tau_k = \infty$ ,  $L_k = \sqrt{1 + \mathcal{O}\left(\frac{1}{k^q}\right)}$ , with  $q > 1$ , and that there exists  $\rho > 0$  such that  $\|w^{(k)}\| \leq \rho$ . If the set of the solutions of (1) is nonempty and  $\text{diam}(\text{dom}(f_1^*))$  is finite, then the sequence  $\{f(x^{(k)})\}$  converges to  $f(x^*)$ , with  $x^* \in X^*$ .*

*Proof.* Since  $w^{(k)} \in \partial f_0(x^{(k)})$ , by Lemma 4.1 we have  $u^{(k)} \in \partial_{\epsilon_k} f(x^{(k)})$ , where  $\epsilon_k = f_1(Ax^{(k)}) + f_1^*(y^{(k+1)}) - y^{(k+1)T} Ax^{(k)}$ . Since  $\text{diam}(\text{dom}(f_1^*))$  is finite, we can apply the second part of Lemma 4.1 obtaining  $\epsilon_k \leq (2\tau_k)^{-1} D^2$ , for a positive constant  $D$  such that  $\text{diam}(\text{dom}(f_1^*)) \leq D$ . Since  $\lim_{k \rightarrow \infty} \tau_k = \infty$ , we have  $\lim_{k \rightarrow \infty} \epsilon_k = 0$  and by Theorem 3.1 we obtain the result.  $\square$



## 5 Application: edge preserving deblurring of Poisson images

In this section we further specialize the SPDHG method, by focusing on a specific application in the image restoration context. Our aim is to suggest a strategy to compute a suitable scaling matrix  $D_k$ , fully defining the algorithm; as observed by several authors, this choice should be driven according to the specific problem features, such as the structure of the constraints and objective function [16, 36, 53].

For these reasons, we describe first some details of the image reconstruction problems which, in the Bayesian framework, can be formulated as a constrained convex minimization problems of the form (30). For these problems, the function  $f_0(x)$  measures the data discrepancy and should be chosen according to the noise statistics: in particular, when the data suffer from Poisson noise, the Maximum Likelihood principle leads to the generalized Kullback–Leibler divergence

$$f_0(x) = \sum_{i=1}^n g_i \log \frac{g_i}{(Hx)_i + b} + (Hx)_i + b - g_i \quad (34)$$

where  $g \in \mathbb{R}^n$  is the observed image,  $H \in \mathbb{R}^{n \times n}$  represents the blurring operator while  $b \in \mathbb{R}$  is a nonnegative background term. Standard assumptions on  $H$  are that it has nonnegative entries and  $H^T e > 0$ , where  $e \in \mathbb{R}^n$  is the vector of all ones.

Since the entries of the unknown vector  $x$  represents the image pixels, a meaningful solution is obtained by defining the constraint set as the nonnegative orthant, i.e.  $X = \{x \in \mathbb{R}^n : x_i \geq 0\}$ . On the other side,  $f_1(Ax)$  plays the role of a regularization term enforcing suitable properties on the solution of (30). Typically, to preserve the edges in the solutions of (30),  $f_1(Ax)$  can be chosen as

$$f_1(Ax) = \beta TV(x), \quad TV(x) = \sum_{i=1}^n \|A_i x\|, \quad A_i \in \mathbb{R}^{2 \times n} \quad (35)$$

where  $TV(x)$  is the discrete, nonsmooth, Total Variation (TV) functional,  $\beta$  is a positive regularization parameter and  $A_i \in \mathbb{R}^{2 \times n}$  is defined such that  $A_i x$  represents the discrete gradient of the image  $x$  at the pixel  $i$ . In these settings, the matrix  $A$  is defined by blocks as  $A = (A_1^T \ A_2^T \ \cdots \ A_n^T)^T \in \mathbb{R}^{2n \times n}$ .

In order to simplify the notation, we assume that  $x \in \mathbb{R}^n$  is a  $N \times N$  image, i.e.  $n = N^2$  and we will indicate the component  $x_\ell$ ,  $\ell = 1, \dots, n$ , also as  $x_{i,j}$ ,  $i, j = 1, \dots, N$ , with the correspondence  $j = \lfloor (\ell - 1)/N \rfloor + 1$ ,  $i = \ell - \lfloor (\ell - 1)/N \rfloor \cdot N$ , where  $\lfloor \cdot \rfloor$  denotes the integer quotient. With this notation, the  $\ell$ -th discrete gradient of the image  $x$  can be written as

$$A_\ell x = \begin{pmatrix} x_{i+1,j} - x_{i,j} \\ x_{i,j+1} - x_{i,j} \end{pmatrix}$$

where some boundary conditions are assumed.

In this case, since  $f_0$  is differentiable, we define  $w^{(k)} = \nabla f_0(x^{(k)})$  in (32), so that  $\delta_k = 0$  for all  $k$  in Corollary 4.1. Moreover, the resolvent operator in (31) consists in a simple projection onto the set  $B \times B \times \cdots \times B \subset \mathbb{R}^{2n}$ , where  $B = \{z \in \mathbb{R}^2 : \|z\| \leq 1\}$ .

In order to devise a suitable scaling matrix  $D_k$  for SPDHG, we adapt to our case the split gradient strategy proposed in [9, 36] for nonnegatively constrained differentiable problems, which demonstrated to be very effective in several applications [8, 13, 44, 50, 53].

The key point of this approach consists in finding a subgradient decomposition of the form  $u^{(k)} = V(x^{(k)}) - U(x^{(k)})$  with  $V(x^{(k)}) > 0$  and  $U(x^{(k)}) \geq 0$  for all  $k$  and then defining  $D_k$  in



(32) as a diagonal scaling matrix whose entries are the projection of  $x_i^{(k)}/V_i(x^{(k)})$  onto the set  $[1/L_k, L_k]$ .

This strategy has the advantage to agree with the nonnegativity constraints and strongly depends on the form of the subgradient  $u^{(k)}$ .

For a practical implementation of this strategy, we have to find a decomposition of the vector  $u^{(k)} = \nabla f_0(x^{(k)}) + \beta A^T y^{(k+1)}$  as the difference of two nonnegative terms.

As concerns the first term, the gradient of  $f_0$  has the natural decomposition  $\nabla f_0(x) = H^T e - H^T v(x)$ , where  $v(x)$  denotes the vector with entries  $v_i(x) = g_i/(Hx + b)_i$ ; by the assumptions on  $H$ , we have  $H^T e > 0$  and  $H^T v(x) \geq 0$  for all  $x \geq 0$ .

Thus, it remains to find a decomposition of the vector  $A^T y^{(k+1)}$  in (32). To this end, we compute the explicit expression of it as a function of  $x^{(j)}$ ,  $j = 0, \dots, k$ . We first observe that, if the dual variable is partitioned as

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad y_i \in \mathbb{R}^2,$$

the updating rule (31) can be written as

$$\begin{aligned} \tilde{y}^{(k)} &= y^{(k)} + \tau_k \beta A x^{(k)} \\ y^{(k+1)} &= S_k \tilde{y}^{(k)} \end{aligned}$$

where  $S_k$  is a diagonal  $2n \times 2n$  matrix with the following diagonal entries

$$(S_k)_{2i-1, 2i-1} = (S_k)_{2i, 2i} = \frac{1}{\max\{1, \|\tilde{y}_i^{(k)}\|\}}, \quad i = 1, \dots, n \quad (36)$$

If the method is initialized with  $y^{(0)} = 0$ , the dual variable can be written as

$$\begin{aligned} y^{(0)} &= 0 \\ y^{(1)} &= \beta \tau_0 S_0 A x^{(0)} \\ y^{(2)} &= \beta S_1 (\tau_0 S_0 A x^{(0)} + \tau_1 A x^{(1)}) \\ y^{(3)} &= \beta S_2 (\tau_0 S_1 S_0 A x^{(0)} + \tau_1 S_1 A x^{(1)} + \tau_2 A x^{(2)}) \\ &\vdots \\ y^{(k+1)} &= \beta \sum_{j=0}^k \tau_j \tilde{S}_j^k A x^{(j)} \end{aligned}$$

where

$$\tilde{S}_j^k = \prod_{i=j}^k S_i$$

As a consequence, the  $\epsilon$ -subgradient of  $f_1 \circ A$  employed in (32) can be expressed as

$$\beta A^T y^{(k+1)} = \beta^2 \sum_{j=0}^k \tau_j A^T \tilde{S}_j^k A x^{(j)} \quad (37)$$

The following simple lemma, which directly follows from the definition of  $A$ , indicates a possible decomposition of each term in the summation at the right hand side of (37) as the difference between a positive and a nonnegative term.

LEMMA 5.1 *Every matrix–vector product of the form  $A^T S A x$  where  $S$  is a  $2n \times 2n$  diagonal matrix with positive entries such that  $S_{2\ell, 2\ell} = S_{2\ell-1, 2\ell-1} = s_\ell$ ,  $\ell = 1, \dots, n$ ,  $x \geq 0$ , can be decomposed as*

$$A^T S A x = V_S x - U_S x$$

where

$$\begin{aligned} (V_S x)_{i,j} &= (2s_{i,j} + s_{i,j-1} + s_{i-1,j})x_{i,j} \geq 0 \\ (U_S x)_{i,j} &= s_{i,j}(x_{i+1,j} + x_{i,j+1}) + s_{i,j-1}x_{i,j-1} + s_{i-1,j}x_{i-1,j} \geq 0 \end{aligned}$$

with the correspondence  $s_\ell \equiv s_{i,j}$ ,  $j = \lfloor (\ell - 1)/N \rfloor + 1$ ,  $i = \ell - \lfloor (\ell - 1)/N \rfloor \cdot N$ . The  $\epsilon$ -subgradient of  $f$  in (32) can be decomposed as

$$u^{(k)} = \nabla f_0(x^{(k)}) + \beta A^T y^{(k+1)} = V(x^{(k)}) - U(x^{(k)})$$

where

$$V(x^{(k)}) = H^T e + \beta^2 \sum_{j=0}^k \tau_j V_{\tilde{S}_j^k} x^{(j)}. \quad (38)$$

Even if it seems quite complicate, the term

$$V^R(x^{(k)}) = \beta^2 \sum_{j=0}^k \tau_j V_{\tilde{S}_j^k} x^{(j)}$$

can be easily computed in a recursive way, by introducing three auxiliary vectors, as described in Algorithm 2. We also notice that, since the scaling matrix  $D_k$  is diagonal and the constraint set  $X$  is the nonnegative orthant, the projection  $P_{X, D_k^{-1}}(\cdot)$  reduces to the usual Euclidean projection  $P_{\geq 0}(\cdot)$  (see Step 7).

By induction it can be shown that the computation of  $V(x^{(k)})$  in (42) actually gives (38). For sake of simplicity, we limit ourselves to show that this is true for  $k = 0, 1$ . Indeed, from Lemma 5.1 and from (39)–(41) we have

$$\begin{aligned} V^R(x^{(0)})_{i,j} &= \beta^2 \tau_0 (V_{S_0} x^{(0)})_{i,j} \\ &= \beta^2 \tau_0 (2s_{i,j}^{(0)} + s_{i-1,j}^{(0)} + s_{i,j-1}^{(0)}) x_{i,j}^{(0)} \\ &= 2p_{i,j}^{(0)} + q_{i,j}^{(0)} + r_{i,j}^{(0)} \end{aligned}$$

**Algorithm 2** Scaled Primal–Dual Hybrid Gradient (SPDHG)

Choose the starting point  $x^{(0)} \in X$  and set  $y^{(0)} = 0$ ,  $p^{(-1)} = q^{(-1)} = r^{(-1)} = 0$ . Choose the sequences  $\{\alpha_k\}$ ,  $\{\tau_k\}$ ,  $\{\gamma_k\}$ .

FOR  $k = 0, 1, 2, \dots$  DO THE FOLLOWING STEPS:

STEP 1. Compute  $\tilde{y}^{(k)} = y^{(k)} + \beta\tau_k Ax^{(k)}$ ;

STEP 2. Compute  $s_\ell^{(k)} = \frac{1}{\max\{1, \|\tilde{y}_\ell^{(k)}\|\}}$ ,  $\ell = 1, \dots, n$  and define  $S_k$  as in (36);

STEP 3. Dual update:  $y^{(k+1)} = S_k \tilde{y}^{(k)}$ ;

STEP 4. Auxiliary vectors update for the decomposition:

$$p_{i,j}^{(k)} = (p_{i,j}^{(k-1)} + \beta^2 \tau_k x_{i,j}^{(k)}) s_{i,j}^{(k)} \quad (39)$$

$$q_{i,j}^{(k)} = (q_{i,j}^{(k-1)} + \beta^2 \tau_k x_{i,j}^{(k)}) s_{i-1,j}^{(k)} \quad (40)$$

$$r_{i,j}^{(k)} = (r_{i,j}^{(k-1)} + \beta^2 \tau_k x_{i,j}^{(k)}) s_{i,j-1}^{(k)} \quad (41)$$

for  $i, j = 1, \dots, N$

STEP 5. Compute the positive part of the decomposition:

$$V(x^{(k)}) = H^T e + (2p^{(k)} + q^{(k)} + r^{(k)}) \quad (42)$$

STEP 6. Compute the scaling matrix:

$$L_k = \sqrt{1 + \gamma_k}$$

$$(D_k)_{\ell,\ell} = \min \left\{ L_k, \max \left\{ L_k^{-1}, \frac{x_\ell^{(k)}}{V(x^{(k)})_\ell} \right\} \right\}$$

STEP 7. Primal update:  $x^{(k+1)} = P_{\geq 0}(x^{(k)} - \alpha_k D_k(\nabla f_0(x^{(k)}) + \beta A^T y^{(k+1)}))$ .

END

$$\begin{aligned} V^R(x^{(1)})_{i,j} &= \beta^2 \tau_0 (V_{S_0 S_1} x^{(0)})_{i,j} + \beta^2 \tau_1 (V_{S_1} x^{(1)})_{i,j} \\ &= \beta^2 \tau_0 (2s_{i,j}^{(0)} s_{i,j}^{(1)} + s_{i-1,j}^{(0)} s_{i-1,j}^{(1)} + s_{i,j-1}^{(0)} s_{i,j-1}^{(1)}) x_{i,j}^{(0)} + \\ &\quad + \beta^2 \tau_1 (2s_{i,j}^{(1)} + s_{i-1,j}^{(1)} + s_{i,j-1}^{(1)}) x_{i,j}^{(1)} \\ &= 2(\beta^2 \tau_0 s_{i,j}^{(0)} x_{i,j}^{(0)} + \beta^2 \tau_1 x_{i,j}^{(1)}) s_{i,j}^{(1)} + \\ &\quad + (\beta^2 \tau_0 s_{i-1,j}^{(0)} x_{i,j}^{(0)} + \beta^2 \tau_1 x_{i,j}^{(1)}) s_{i-1,j}^{(1)} + \\ &\quad + (\beta^2 \tau_0 s_{i,j-1}^{(0)} x_{i,j}^{(0)} + \beta^2 \tau_1 x_{i,j}^{(1)}) s_{i,j-1}^{(1)} \\ &= 2(p_{i,j}^{(0)} + \beta^2 \tau_1 x_{i,j}^{(1)}) s_{i,j}^{(1)} + (q_{i,j}^{(0)} + \beta^2 \tau_1 x_{i,j}^{(1)}) s_{i-1,j}^{(1)} + (r_{i,j}^{(0)} + \beta^2 \tau_1 x_{i,j}^{(1)}) s_{i,j-1}^{(1)} \\ &= 2p_{i,j}^{(1)} + q_{i,j}^{(1)} + r_{i,j}^{(1)} \end{aligned}$$

Algorithm 2 can be adapted for both the stepsize selection strategies described in Sections 2 and 3. In the first case, three prefixed sequences  $\{\alpha_k\}$ ,  $\{\tau_k\}$ , and  $\{\gamma_k\}$  satisfying the assumptions of

Corollary 4.1 have to be provided.

In the other case, the SSL procedure for dynamically compute the primal stepsize  $\alpha_k$  can be included in Algorithm 2. Here, only the sequences  $\{\tau_k\}$ , and  $\{\gamma_k\}$  should be given such that  $\lim_{k \rightarrow \infty} \tau_k = \infty$  and  $\sum \gamma_k < \infty$ .

## 6 Numerical experience

The aim of our numerical experience is twofold: first, we are interested in evaluating the effect of the scaling on the convergence behaviour of the  $\epsilon$ -subgradient method. Secondly, we compare the two steplength selection strategies presented in Sections 2-3.

To this end we consider four different versions of the method (31)–(32):

**PDHG** corresponds to the choices  $D_k = I$  and  $\alpha_k$  chosen as an *a priori* diminishing, divergent series, square summable sequence in (32). It actually consists in the method in [15];

**SPDHG** is Algorithm 2 with  $\alpha_k$  chosen as an *a priori* diminishing, divergent series, square summable sequence;

**SL** is the  $\epsilon$ -subgradient level method given in Algorithm 1 with  $D_k = I$  and  $u^{(k)} = \nabla f_0(x^{(k)}) + \beta A^T y^{(k+1)}$ , where  $y^{(k+1)}$  is updated as in (31);

**SSL** is the same as above but with the scaling matrix  $D_k$  defined as at the Step 6 of Algorithm 2.

The numerical experiments described in this section have been performed in the MATLAB environment (R2012b) on a PC equipped with an Intel Core i7-3517U processor 1.9 GHz, 8 GB RAM.

In our experiments, we consider problem (30) where  $f_0, f_1$  are defined in (34)–(35) and  $H$  represents the convolution operator with a given point spread function (psf). Thus, assuming periodic boundary conditions, the matrix–vector products involving  $H$  can be computed by the Fast Fourier Transform (FFT) and, by a simple normalization of the psf, we also have  $H^T e = H e = e$ ; moreover, in our experiments,  $H$  is nonsingular, although very ill conditioned, and  $g > 0$ , so that problem (30) has a unique solution [28].

We consider a set of three test problems generated as in [47] where the data  $g$  is obtained with the following procedure: the selected original image  $\bar{x}$  is rescaled so that the maximum pixel intensity is a specified value  $I_{max}$ . Then, the rescaled image is convolved with the psf and the background  $b$  is added. Finally, Poisson noise is introduced by the Matlab `imnoise` function and the simulated data  $g$  is obtained after scaling back again by  $I_{max}$ .

For each test problem, the regularization parameter  $\beta$  has been empirically selected by computing the solution of (30) for different values of  $\beta$  and choosing that for which we observed the minimum  $l_2$  relative distance with respect to  $\bar{x}$ .

The features of each test problem are specified below.

**cameraman**: the  $256 \times 256$  original image is the ‘cameraman’ available in the Matlab package, while the psf is a Gaussian function, with standard deviation 1.3, truncated at the  $9 \times 9$  central pixels. The other parameters are  $I_{max} = 1000$ ,  $b = 0$ ,  $\beta = 0.005$ ; the  $l_2$  relative distance between  $\bar{x}$  and  $g$  is 0.1209, while  $g_i \in [4, 250]$ .

	PDHG			SL	
	$\tau_k$	$\alpha_k$	$\gamma_k$	$\tau_k$	$\gamma_k$
<i>cameraman</i>	$0.9 + 10^{-2}k$	$(0.04 + 10^{-5}k)^{-1}$	-	$0.5 + 5 \cdot 10^{-2}k$	-
<i>micro</i>	$0.9 + 10^{-3}k$	$(0.04 + 10^{-4}k)^{-1}$	-	$0.9 + 10^{-1}k$	-
<i>phantom</i>	$0.9 + 10^{-3}k$	$(0.2 + 10^{-5}k)^{-1}$	-	$0.9 + 10^{-2}k$	-
	SPDHG			SSL	
	$\tau_k$	$\alpha_k$	$\gamma_k$	$\tau_k$	$\gamma_k$
<i>cameraman</i>	$0.5 + 5 \cdot 10^{-3}k$	$(0.5 + 5 \cdot 10^{-5}k)^{-1}$	$10^{13}k^{-2}$	$0.7 + 5 \cdot 10^{-2}k$	$10^{13}k^{-2}$
<i>micro</i>	$0.4 + 10^{-5}k$	$(0.4 + 10^{-5}k)^{-1}$	$10^{13}k^{-2}$	$0.9 + 10^{-2}k$	$10^{13}k^{-2}$
<i>phantom</i>	$0.5 + 10^{-4}k$	$(0.5 + 10^{-5}k)^{-1}$	$10^{13}k^{-2}$	$0.9 + 10^{-2}k$	$10^{13}k^{-2}$

Table 1: Parameter settings.

**micro:** the original image is the confocal microscopy phantom of size  $128 \times 128$  described in [52], scaled by 10; the psf is the one in [52] truncated at the  $9 \times 9$  central pixels. Here we set  $I_{max} = 1$ ,  $b = 0$ ,  $\beta = 0.0477$ ; the original image pixels are in the range  $[10, 690]$ , the  $l_2$  relative distance between  $\bar{x}$  and  $g$  is 0.1442, while  $g_i \in [1, 778]$ .

**phantom:** the original image  $\bar{x}$  is the  $256 \times 256$  Shepp-Logan phantom, generated by the Matlab function `phantom`, scaled by a factor 1000, while the psf is a Gaussian function, with standard deviation 3, truncated at the  $9 \times 9$  central pixels. In this case we set  $I_{max} = 1$ ,  $b = 10$  and  $\beta = 0.00526$ . The values of the original image are in the range  $[0, 1000]$ , the  $l_2$  relative distance between  $\bar{x}$  and  $g$  is 0.4643, while  $g_i \in [1, 934]$ .

For all test problems we compute the solution  $x^*$  of the minimization problem (30) by running 50000 iterations of PIDSplit method [47]. Then, we evaluate the progress toward this solution at each iteration in terms of the  $l_2$  relative error from the minimum point and the relative difference from the optimal value

$$e^k = \frac{\|x^{(k)} - x^*\|}{\|x^*\|} \quad f^k = \frac{f(x^{(k)}) - f(x^*)}{f(x^*)}.$$

Following the assumptions of Corollaries 4.1 and 4.2, we choose the sequences of parameters as follows

$$\tau_k = t_1 + t_2 k \quad \alpha_k = \frac{1}{t_3 + t_4 k} \quad \gamma_k = \frac{t_5}{k^{1+t_6}}$$

In order to illustrate the effectiveness of the methods, the values  $t_i$  have been manually optimized for each test problem to obtain a faster decrease of  $e^k$  (see Table 1).

Moreover, for the initialization of both SL and SSL, we adopt the rule

$$\delta_0 = 0.9 f(x^{(0)})$$

while the other parameters are  $\nu_1 = \nu_2 = 0.5$ ,  $B = 0.9 \|u^{(0)}\| \|D_0\|_\infty^{\frac{1}{2}}$ .

The plots in Figure 1 have been obtained by running 3000 iterations of the algorithms, reporting the errors  $e^k$ ,  $f^k$  with respect to the computational time in seconds.

From the numerical experience we observe that the presence of the scaling can help to accelerate the progress towards the solution, with both stepsize selection strategies. As concerns the scaling matrix bounds, the best results are obtained by selecting large initial values for  $\gamma_k$

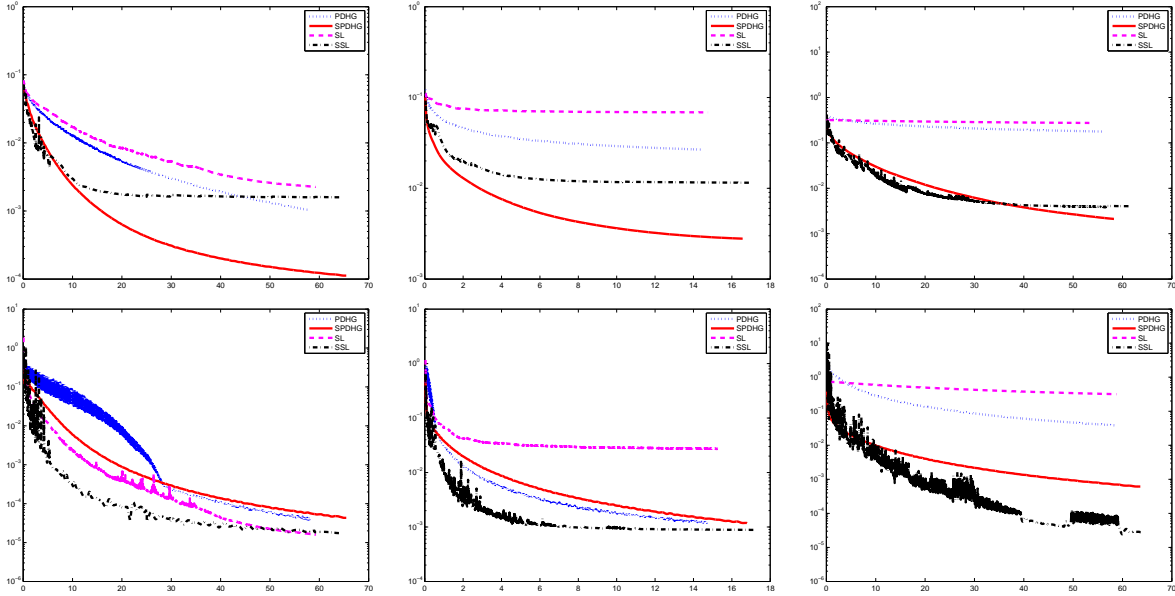


Figure 1: Image deblurring results. Upper row: plots of the relative minimization error  $e^k$ . Lower row: plots of the relative difference from the optimal function value  $f^k$ . Left column: *cameraman*. Middle column: *micro*. Right column: *phantom*. All plots are with respect to the computational time in seconds and use a logarithmic scale on the vertical axis.

(see Table 1) and, thus, for  $L_k$ , allowing more freedom to choose the scaling matrix especially at the first iterations.

It is also interesting to observe that the adaptive computation of  $\alpha_k$  combined with the proposed scaling technique seems to work quite well, leading to performances that are, in some cases, close to the ‘best’ ones obtainable by manually tuning the stepsize sequence. Indeed, the PDHG and SPDHG performances are sensitive to the choice of  $\tau_k$ , and especially of  $\alpha_k$ , making difficult to devise a general rule to select these sequences.

Comparing the first and the second row in Figure 1, we also observe that a faster approaching to the solution  $x^*$  does not always correspond to a faster decrease of the objective function: we suppose that this phenomenon is due to the ill conditioning of problem (30). For completeness, we experimentally observed that too large initial values of the primal stepsize  $\alpha_k$  in PDHG and SPDHG may produce an unbounded sequence  $\{u^{(k)}\}$  and, as a consequence, the algorithms fail to converge: this indicates that the assumption on the  $\epsilon$ -subgradient boundedness is crucial.

## 7 Conclusions

In this paper we proposed a generalization of the  $\epsilon$ -subgradient projection method with variable scaling matrix for nonsmooth, convex, constrained optimization, developing the related convergence analysis when the stepsize parameters are either provided as *a priori* selected sequences or dynamically computed by an adaptive procedure. Exploiting the duality principle, we described a special case of the proposed method which applies to the minimization of the sum of two convex functions. For a specific problem of this form in the image restoration framework,

we fully detailed the algorithm, also suggesting a strategy to compute the scaling matrix. The numerical experience shows that the presence of a suitable variable scaling matrix can accelerate the progress of the iterates towards the solution. Moreover, the results obtained combining the variable scaling with the adaptive procedure for the computation of the stepsize parameter are encouraging.

Future work will be addressed to further investigate dynamic choices of the stepsize and of the scaling matrix, with the aim to devise effective ‘black-box’ algorithms which are able to handle practical applications with a minimum of user supplied parameters.

## References

- [1] Y. ALBER, A. IUSEM, AND M. SOLODOV, *On the projected subgradient method for nonsmooth convex optimization in a Hilbert space*, Math. Prog., 81 (1998), pp. 23–35.
- [2] R. D. ASMUNDIS, D. D. SERAFINO, F. RICCIO, AND G. TORALDO, *On spectral properties of steepest descent methods*, IMA Journal of Numerical Analysis, 33 (2013), pp. 1416–1435.
- [3] A. AUSLANDER AND M. TEBoulLE, *Projected subgradient methods with non-euclidean distances for non-differentiable convex minimization and variational equalities*, Math. Program. Ser. B, 120 (2009).
- [4] J. BARDSLEY AND A. LUTTMAN, *Total Variation-Penalized Poisson Likelihood Estimation for Ill-Posed Problems*, Adv. Comput. Math., 31 (2009), pp. 35–59.
- [5] J. BARDSLEY AND J. NAGY, *Covariance-preconditioned iterative methods for nonnegatively constrained astronomical imaging*, SIAM J. Matrix Anal. A., 27 (2006), pp. 1184–1198.
- [6] J. BARZILAI AND J. M. BORWEIN, *Two-point step size gradient methods*, IMA J. Numer. Anal., 8 (1988), pp. 141–148.
- [7] H. BAUSCHKE AND P. COMBETTES, *Convex analysis and monotone operator theory in Hilbert spaces*, CMS Books on Mathematics, Springer, 2011.
- [8] M. BERTERO, P. BOCCACCI, G. DESIDER, AND G. VICIDOMINI, *Image deblurring with Poisson data: From cells to galaxies*, Inverse Problems, 25 (2009).
- [9] M. BERTERO, H. LANTERI, AND L. ZANNI, *Iterative image reconstruction: a point of view*, in Mathematical Methods in Biomedical Imaging and Intensity-Modulated Radiation Therapy (IMRT), Y. Censor, M. Jiang, and A. K. Louis, eds., Birkhauser-Verlag, Pisa, Italy, 2008, pp. 37–63.
- [10] D. BERTSEKAS, *Convex Optimization Theory*, Athena Scientific, Belmont, Massachusetts, 5 november 2012 ed., June 2009, ch. 6 on Convex Optimization Algorithms, pp. 251–489.
- [11] D. P. BERTSEKAS AND T. J. N., *Parallel and distributed computation*, Prentice Hall International, Englewood Cliffs, New Jersey, 1989.
- [12] S. BONETTINI, G. LANDI, E. L. PICCOLOMINI, AND L. ZANNI, *Scaling techniques for gradient projection-type methods in astronomical image deblurring*, Int. J. Comput. Math., 90 (2013), pp. 9–29.

- [13] S. BONETTINI AND M. PRATO, *Nonnegative image reconstruction from sparse Fourier data: a new deconvolution algorithm*, Inverse Probl., 26 (2010), p. 095001.
- [14] S. BONETTINI AND V. RUGGIERO, *An alternating extragradient method for total variation based image restoration from Poisson data*, Inverse Problems, 27 (2011), p. 095001.
- [15] ———, *On the convergence of primal-dual hybrid gradient algorithms for total variation image restoration*, J. Math. Imaging Vis., 44 (2012), pp. 236–253.
- [16] S. BONETTINI, R. ZANELLA, AND L. ZANNI, *A scaled gradient projection method for constrained image deblurring*, Inverse Probl., 25 (2009), p. 015002.
- [17] U. BRÄNNLUND, K. KIWIEL, AND P. LINDBERG, *A descent proximal level bundle method for convex nondifferentiable optimization*, Oper. Res. Lett., 17 (1995), pp. 121–126.
- [18] P. COMBETTES, *Quasi-Féjerian analysis of some optimization algorithms*, Studies in Computational Mathematics, 8 (2001), pp. 115–152.
- [19] P. COMBETTES AND B. VÛ, *Variable metric quasi-Féjer monotonicity*, Nonlinear Analysis: Theory, Methods, and Applications, 78 (2013).
- [20] ———, *Variable metric forward-backward splitting with applications to monotone inclusions in duality*, Optimization, 63 (2014), pp. 1289–1318.
- [21] L. CONDAT, *A primal-dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms*, Journal of Optimization Theory and Applications, 158 (2013), pp. 460–479.
- [22] R. CORREA AND C. LEMARÉCHAL, *Convergence of some algorithms for convex minimization*, Mathematical Programming, 62 (1993), pp. 261–275.
- [23] Y. H. DAI, W. W. HAGER, K. SCHITTKOWSKI, AND H. ZHANG, *The cyclic Barzilai-Borwein method for unconstrained optimization*, IMA J. Numer. Anal., 26 (2006), pp. 604–627.
- [24] F.-X. DUPE, M. FADILI, AND J.-L. STARCK, *Inverse problems with Poisson noise: primal and primal-dual splitting*, in Image Processing (ICIP), 2011 18th IEEE International Conference on, Brussels, 2011, IEEE, pp. 1901–1904.
- [25] ———, *Linear inverse problems with various noise models and mixed regularizations*, in Proceedings of the 5th International ICST Conference on Performance Evaluation Methodologies and Tools, Brussels, 2011, ICST, pp. 621–626.
- [26] Y. ERMOLIEV, *Stochastic quasigradient methods and their application to system optimization*, Stochastics, 9 (1983), pp. 1–36.
- [27] E. ESSER, X. ZHANG, AND T. CHAN, *A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science*, SIAM J. Imaging Sci., 3 (2010), pp. 1015–1046.
- [28] M. FIGUEIREDO AND J. BIOUSCAS-DIAS, *Restoration of poissonian images using alternating direction optimization*, IEEE Transactions on Image Processing, 19 (2010), pp. 3133–3145.



- [29] R. FLETCHER, *A limited memory steepest descent method*, Math. Program., 135 (2012), pp. 413–436.
- [30] G. FRASSOLDATI, G. ZANGHIRATI, AND L. ZANNI, *New adaptive stepsize selections in gradient methods*, J. Ind. Manage. Optim., 4 (2008), pp. 299–312.
- [31] S. GEMAN AND D. GEMAN, *Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images*, IEEE Trans. Pattern Anal. Mach. Intell., 6 (1984), pp. 721–741.
- [32] J. GOFFIN, *On convergence rates of subgradient optimization methods*, Mathematical Programming, 13 (1977), pp. 329–347.
- [33] J. GOFFIN AND K. KIWIEL, *Convergence of a simple subgradient level method*, Mathematical Programming, 85 (1999), pp. 207–211.
- [34] W. W. HAGER, B. A. MAIR, AND H. ZHANG, *An affine-scaling interior-point CBB method for box-constrained optimization*, Math. Program., 119 (2009), pp. 1–32.
- [35] K. KIWIEL, *Convergence of approximate and incremental subgradient methods for convex optimization*, SIAM J. Optim., 14 (2004), pp. 807–840.
- [36] H. LANTÉRI, M. ROCHE, AND C. AIME, *Penalized maximum likelihood image restoration with positivity constraints: multiplicative algorithms*, Inverse Probl., 18 (2002), pp. 1397–1419.
- [37] T. LARSSON, M. PATRIKSSON, AND A.-B. STRÖMBERG, *On the convergence of conditional  $\epsilon$ -subgradient methods for convex programs and convex-concave saddle-point problems*, European Journal of Operational Research, 151 (2003), pp. 461–473.
- [38] D. LORENZ AND T. POCK, *An inertial forward-backward algorithm for monotone inclusions*, Journal of Mathematical Imaging and Vision, (2014).
- [39] A. NEDIĆ, *Subgradient methods for convex minimization*, PhD thesis, Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science, 2002.
- [40] A. NEDIĆ AND D. P. BERTSEKAS, *Incremental subgradient methods for nondifferentiable optimization*, SIAM J. on Optim., 12 (2001), pp. 109–138.
- [41] Y. NESTEROV, *Introductory Lectures on Convex Optimization*, 2004.
- [42] E. H. N. A. D. PIERRO, *Incremental subgradients for constrained convex optimization: a unified framework and new methods*, Math. Program., 103 (2005).
- [43] B. POLYAK, *Introduction to Optimization*, Optimization Software - Inc., Publication Division, N.Y., 1987.
- [44] M. PRATO, A. L. CAMERA, S. BONETTINI, AND M. BERTERO, *A convergent blind deconvolution method for post-adaptive-optics astronomical imaging*, Inverse Probl., 29 (2013), p. 065017.
- [45] S. ROBINSON, *Linear convergence of epsilon-subgradient descent methods for a class of convex functions*, Math. Program., Ser. A, 86 (1999).

- [46] R. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
- [47] S. SETZER, G. STEIDL, AND T. TEUBER, *Deblurring Poissonian images by split Bregman techniques*, J. Vis. Commun. Image R., 21 (2010), pp. 193–199.
- [48] N. SHOR, *Minimization Methods for Nondifferentiable Functions*, Springer-Verlag, Berlin, 1985.
- [49] M. SOLODOV AND S. ZAVIEV, *Error stability properties of generalized gradient-type algorithms*, J. Opt. Theory and Appl., 98 (1998), pp. 663–680.
- [50] C. THEYS, N. DOBIGEON, J.-Y. TOURNERET, AND H. LANTÉRI, *Linear unmixing of hyperspectral images using a scaled gradient method*, in Proceedings of the IEEE/SP 15th Workshop on Statistical Signal Processing, Aug. 2009, pp. 729–732.
- [51] S. VILLA, S. SALZO, L. BALDASSARRE, AND A. VERRI, *Accelerated and inexact forward-backward algorithms*, SIAM Journal on Optimization, 23 (2013), pp. 1607–1633.
- [52] R. WILLETT AND R. NOWAK, *Platelets: A multiscale approach for recovering edges and surfaces in photon limited medical imaging*, IEEE Transactions on Medical Imaging, 22 (2003), pp. 332–350.
- [53] R. ZANELLA, P. BOCCACCI, L. ZANNI, AND M. BERTERO, *Efficient gradient projection methods for edge-preserving removal of Poisson noise*, Inverse Probl., 25 (2009), p. 045010.