A PENALIZED WEIGHTED LEAST SQUARES APPROACH FOR RESTORING DATA CORRUPTED WITH SIGNAL-DEPENDENT NOISE

Audrey Repetti, Emilie Chouzenoux and Jean-Christophe Pesquet

Université Paris-Est, Lab. d'Informatique Gaspard Monge, UMR CNRS 8049,
Champs-sur-Marne, 77454 Marne-la-Vallée, France
{first.last}@univ-paris-est.fr

ABSTRACT

This paper addresses the problem of recovering an image degraded by a linear operator and corrupted with an additive Gaussian noise with a signal-dependent variance. The considered observation model arises in several digital imaging devices. To solve this problem, a variational approach is adopted relying on a weighted least squares criterion which is penalized by a non-smooth function. In this context, the choice of an efficient optimization algorithm remains a challenging task. We propose here to extend a recent primal-dual proximal splitting approach by introducing a preconditioning strategy that is shown to significantly speed up the algorithm convergence. The good performance of the proposed method is illustrated through image restoration examples.

Index Terms— preconditioning, convex optimization, image restoration, denoising, deconvolution, total variation, regularization.

1. INTRODUCTION

In many image processing tasks such as denoising, inpainting and deblurring, one has to solve an inverse problem where an image $\hat{x}$ needs to be estimated from observed data y related to the original image $\overline{x}$ through a linear model and some noise corruption process. An efficient strategy to address this problem is to define $\hat{x}$ as a minimizer of an appropriate cost function. More specifically, we will focus on the following formulation:

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad (f(x) = h(x) + g(x)), \qquad (1)$$

where $\mathcal{H}$ is the image space, h is the so-called data fidelity term, and g is a regularization function incorporating *a priori* information so as to guarantee the stability of the solution w.r.t. the observation noise. In this paper, $\mathcal{H}$ is assumed to be a real separable Hilbert space, h is a convex and differentiable function from $\mathcal{H}$ to $\mathbb{R}$, and g belongs to the class $\Gamma_0(\mathcal{H})$ of lower semicontinuous, proper, convex functions from $\mathcal{H}$ to $(-\infty, +\infty]$.

Iterative optimization methods must generally be employed to solve (1). Starting from an initial guess x_0, these algorithms generate a sequence of updated estimates $(x_k)_{k>0}$ until a sufficient accuracy is reached. In the case of large scale optimization problems such as those encountered in image restoration, one major concern is to find an optimization algorithm able to deliver reliable numerical solutions in a reasonable time.

The primal-dual proximal splitting algorithm which was recently proposed in [1] allows us to solve (1) for any Lipschitz differentiable function h and for arbitrary linear operators arising in the expression of g. This algorithm can deal with a non-smooth function g, while not requiring any inversion of the involved linear operators. It is thus characterized by a low computational cost per iteration. However, this advantage is sometimes counterbalanced by a slow convergence. In the context of descent methods for smooth convex optimization, an efficient way to accelerate the convergence is to transform the space of original variables into a space in which the Hessian of the cost function has more clustered eigenvalues. This transformation is performed by multiplying, at each iteration, the search direction by a so-called preconditioning matrix that approximates the inverse of the Hessian [2, 3]. Recently, similar preconditioning techniques for non-smooth convex optimization have been considered in the case of the algorithm in [4], leading to significant acceleration of its convergence [5].

In this paper, we introduce a new preconditioned primal-dual splitting algorithm, which generalizes the algorithm in [1], and we give conditions under which its convergence is guaranteed. The proposed algorithm has the ability to solve (1) under weaker assumptions on h and g than those in [5]. Moreover, we show that this results in an efficient minimization method for a penalized weighted least squares problem arising in the context of image restoration when the data are corrupted with an additive Gaussian noise with signal-dependent variance. The target application is of particular interest since this kind of noise is encountered in real-world digital imaging devices having CMOS or CCD sensors [6, 7, 8].

The rest of this paper is as follows: In Section 2, we describe the proposed preconditioned primal-dual splitting algorithm and investigate its convergence properties. Section 3 addresses the application of this algorithm to the restoration of images degraded by a blur and an additive white Gaussian noise, the variance of which is an affine function of the signal. Numerical results illustrating the superiority of the proposed preconditioned algorithm with respect to its non preconditioned version, are provided in Section 4. Finally, some conclusions are drawn in Section 5.

2. PROPOSED OPTIMIZATION METHOD

2.1. Convex optimization tools

Let us introduce some notations which will be useful in the sequel. First, we define the weighted norm

$$(\forall \boldsymbol{x} \in \mathcal{H}) \qquad \|\boldsymbol{x}\|_{\boldsymbol{R}} = \langle \boldsymbol{x}|\boldsymbol{R}\boldsymbol{x}\rangle^{1/2}, \qquad (2)$$

where $\boldsymbol{R}$ is some positive definite self-adjoint linear operator from $\mathcal{H}$ to $\mathcal{H}$. The associated scalar product reads:

$$\left(\forall (\boldsymbol{x}, \boldsymbol{x}') \in \mathcal{H}^2\right) \qquad \langle \boldsymbol{x}|\boldsymbol{x}'\rangle_{\boldsymbol{R}} = \langle \boldsymbol{x}|\boldsymbol{R}\boldsymbol{x}'\rangle. \qquad (3)$$

The proposed proximal splitting method makes use of the concepts of conjugate function and proximity operator which are recalled below.

Definition 1. *Let ψ be a function in $\Gamma_0(\mathcal{H})$.*

(i) The conjugate *of ψ in the Hilbert space $(\mathcal{H}, \|.\|_{\boldsymbol{R}})$ is defined as*

$$\psi_{\boldsymbol{R}}^* : \mathcal{H} \to [-\infty, +\infty] : \boldsymbol{v} \mapsto \sup_{\boldsymbol{\xi} \in \mathcal{H}} \left(\langle \boldsymbol{\xi}|\boldsymbol{v}\rangle_{\boldsymbol{R}} - \psi(\boldsymbol{\xi})\right).$$

(ii) For every $\boldsymbol{\xi} \in \mathcal{H}$, the minimization problem

$$\underset{\boldsymbol{\pi} \in \mathcal{H}}{\text{minimize}} \;\; \psi(\boldsymbol{\pi}) + \frac{1}{2}\|\boldsymbol{\xi} - \boldsymbol{\pi}\|_{\boldsymbol{R}}^2,$$

admits a unique solution, which is denoted by $\text{prox}_{\boldsymbol{R},\psi}(\boldsymbol{\xi})$. The so-defined function $\text{prox}_{\boldsymbol{R},\psi} : \mathcal{H} \to \mathcal{H}$ is the proximity operator *of ψ in the Hilbert space $(\mathcal{H}, \|.\|_{\boldsymbol{R}})$.*

In the particular case when $\boldsymbol{R}$ reduces to the identity operator, the usual notation ψ^* (resp. prox_ψ) will be used for the conjugate (resp. the proximity operator) of ψ.

2.2. Minimization problem

Let $(\mathcal{G}_j)_{1 \leq j \leq J}$ be some real Hilbert spaces. We consider convex optimization problems of the form (1) where h is convex and differentiable with a μ-Lipschitzian gradient ∇h for some $\mu \in (0, +\infty)$ and g takes the form

$$g(\boldsymbol{x}) = g_0(\boldsymbol{x}) + \sum_{j=1}^{J} g_j(\boldsymbol{L}_j\boldsymbol{x}), \qquad (4)$$

with $g_0 \in \Gamma_0(\mathcal{H})$, and, for all $j \in \{1, \dots, J\}$, $g_j \in \Gamma_0(\mathcal{G}_j)$ and $\boldsymbol{L}_j : \mathcal{H} \to \mathcal{G}_j$ is a non-zero bounded linear operator. Moreover, we make the following assumption:

Assumption 1. *Problem (1) has at least one solution and one of the following statements holds.*

1. For every $j \in \{1, \dots, J\}$, g_j is real-valued.

2. $\mathcal{H}$ and $(\mathcal{G}_j)_{1 \leq j \leq J}$ are finite-dimensional, and there exists $\boldsymbol{x} \in \text{ri dom } g_0$ such that, for every $j \in \{1, \dots, J\}$,

$$\boldsymbol{L}_j\boldsymbol{x} \in \text{ri dom } g_j, \qquad (5)$$

where $\text{ri dom } g_j$ denotes the relative interior of the domain of function g_j.

2.3. Preconditioned primal-dual splitting algorithm

We propose the following preconditioned primal-dual proximal splitting method to solve the considered optimization problem.

Algorithm 1 Preconditioned M+L FBF algorithm.

Let $(\gamma_k)_{k \in \mathbb{N}}$ be a sequence of $(0, +\infty)$.
Initialization:
Let $\boldsymbol{x}_0 \in \mathcal{H}$, and, for every $j \in \{1, \dots, J\}$, let $\boldsymbol{v}_{j,0} \in \mathcal{G}_j$
Iterations:
For $k = 0, \dots$

$\quad\Big|\; \boldsymbol{y}_{1,k} = \boldsymbol{x}_k - \gamma_k \boldsymbol{Q}\big(\nabla h(\boldsymbol{x}_k) + \sum_{j=1}^{J} \boldsymbol{L}_j^* \boldsymbol{v}_{j,k}\big)$
$\quad\Big|\; \boldsymbol{p}_{1,k} = \text{prox}_{\boldsymbol{Q}^{-1}, \gamma_k g_0}(\boldsymbol{y}_{1,k})$
$\quad\Big|\; \text{For } j = 1, \dots, J$
$\quad\Big|\;\; \Big|\; \boldsymbol{y}_{2,j,k} = \boldsymbol{v}_{j,k} + \gamma_k \boldsymbol{R}_j \boldsymbol{L}_j \boldsymbol{x}_k$
$\quad\Big|\;\; \Big|\; \boldsymbol{p}_{2,j,k} = \text{prox}_{\boldsymbol{R}_j^{-1}, \gamma_k g_j^*}(\boldsymbol{y}_{2,j,k})$
$\quad\Big|\;\; \Big|\; \boldsymbol{q}_{2,j,k} = \boldsymbol{p}_{2,j,k} + \gamma_k \boldsymbol{R}_j \boldsymbol{L}_j \boldsymbol{p}_{1,k}$
$\quad\Big|\;\; \Big|\; \boldsymbol{v}_{j,k+1} = \boldsymbol{v}_{j,k} - \boldsymbol{y}_{2,j,k} + \boldsymbol{q}_{2,j,k}.$
$\quad\Big|\; \boldsymbol{q}_{1,k} = \boldsymbol{p}_{1,k} - \gamma_k \boldsymbol{Q}\big(\nabla h(\boldsymbol{p}_{1,k}) + \sum_{j=1}^{J} \boldsymbol{L}_j^* \boldsymbol{p}_{2,j,k}\big)$
$\quad\Big|\; \boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \boldsymbol{y}_{1,k} + \boldsymbol{q}_{1,k}.$

Hereabove, $\boldsymbol{Q}$ is a positive definite self-adjoint linear operator from $\mathcal{H}$ to $\mathcal{H}$, and, for every $j \in \{1, \dots, J\}$, $\boldsymbol{R}_j$ is a positive definite self-adjoint linear operator from $\mathcal{G}_j$ to $\mathcal{G}_j$. When $\boldsymbol{Q}$ and $\boldsymbol{R}_j$ are identity operators, we recover the Monotone+Lipschitz Forward Backward Forward (M+L FBF) algorithm, the convergence properties of which are analysed in [1]. The convergence of Algorithm 1 for an arbitrary choice of preconditioning operators is guaranteed by the following result:

Theorem 1. *For every $k \in \mathbb{N}$, let the step-size γ_k be chosen in $[\varepsilon, (1-\varepsilon)/\tau]$ where $\varepsilon \in (0, 1/(\tau+1))$,*

$$\tau = \mu^{(\boldsymbol{Q})} + \sqrt{\sum_{j=1}^{J} \|\boldsymbol{R}_j^{1/2}\boldsymbol{L}_j\boldsymbol{Q}^{1/2}\|^2}, \qquad (6)$$

and $\mu^{(\boldsymbol{Q})}$ is a Lipschitz constant of $\nabla\left(h \circ \boldsymbol{Q}^{1/2}\right)$. Let $(\boldsymbol{x}_k)_{k \in \mathbb{N}}$ be a sequence generated by Algorithm 1. Under Assumption 1, there exists a solution $\hat{\boldsymbol{x}}$ to (1) such that $(\boldsymbol{x}_k)_{k \in \mathbb{N}}$ and $(\boldsymbol{p}_{1,k})_{k \in \mathbb{N}}$ converge weakly to $\hat{\boldsymbol{x}}$. Moreover, if g_0 or h is uniformly convex at $\hat{\boldsymbol{x}}$, then $(\boldsymbol{x}_k)_{k \in \mathbb{N}}$ and $(\boldsymbol{p}_{1,k})_{k \in \mathbb{N}}$ converge strongly to $\hat{\boldsymbol{x}}$.

3. APPLICATION TO DATA RECOVERY IN THE PRESENCE OF SIGNAL-DEPENDENT NOISE

In this section, we consider an observed image $\boldsymbol{y} \in \mathbb{R}^N$ related to an original signal $\overline{\boldsymbol{x}} \in [0, +\infty)^N$ through the generic signal-dependent noise observation model:

$$(\forall n \in \{1, \dots, N\}) \qquad y_n = z_n + \sigma_n(z_n)w_n, \qquad (7)$$

where, $(z_n)_{1 \leq n \leq N} = \boldsymbol{H}\overline{\boldsymbol{x}}$, $\boldsymbol{H} \in \mathbb{R}^{N \times N}$ is a matrix with non-negative elements, $(w_n)_{1 \leq n \leq N}$ is a realization of a random vector $\boldsymbol{W} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_N)$ ($\boldsymbol{I}_N$ denotes the identity matrix of $\mathbb{R}^{N \times N}$), and

$$(\forall n \in \{1, \dots, N\}) \quad \sigma_n : [0, +\infty) \to (0, +\infty)$$
$$z_n \mapsto \sqrt{\alpha_n z_n + \beta_n} \qquad (8)$$

with $\alpha_n \geqslant 0$, $\beta_n > 0$. The observation model (7)-(8) arises in a number of real digital imaging devices [6, 7, 8] where the acquired images are contaminated by signal-dependent Photon shot noise and by independent electrical or thermal noise. Gaussian-dependent noise can also be viewed as a second order approximation of Gaussian-Poissonian noise which is frequently encountered in astronomy, medicine and biology [9].

In the context of inverse problems, the original signal can be retrieved by solving (1) where $\mathcal{H}$ corresponds to the Euclidean space $\mathbb{R}^N$ and the data fidelity term is the following weighted least squares criterion [10, 11]:

$$(\forall \boldsymbol{x} \in [0, +\infty)^N) \quad h(\boldsymbol{x}) = \frac{1}{2} \sum_{n=1}^{N} \frac{(y_n - [\boldsymbol{H}\boldsymbol{x}]_n)^2}{\alpha_n [\boldsymbol{H}\boldsymbol{x}]_n + \beta_n} \qquad (9)$$

(for every $n \in \{1, \dots, N\}$, $[\boldsymbol{H}\boldsymbol{x}]_n$ denotes the n-th component of vector $\boldsymbol{H}\boldsymbol{x}$). In addition, g is a regularization function taking the form (4) which aims at enforcing some desirable properties of the target image. Here, we will set, for every $j \in \{1, \dots, J\}$, $\mathcal{G}_j = \mathbb{R}^{N_j}$. For instance, a sparsity prior in an analysis frame with frame operator $\boldsymbol{L}_j$ is introduced by taking g_j equal to $\lambda_j \|\cdot\|_1$ with $\lambda_j > 0$. Block sparsity measures [12] can also be easily addressed in the proposed framework. Another popular choice in image restoration is the total variation penalization [13]. The above penalties can be considered individually ($J = 1$) or combined in a hybrid manner ($J > 1$) [14]. Finally, to take into account the expected dynamic range of the output image, a possible choice is to take g_0 equal to $\iota_{A^N} : (x_n)_{1 \leq n \leq N} \mapsto \sum_{n=1}^{N} \iota_A(x_n)$, where ι_A designates the indicator function of a closed convex subset A of $[0, +\infty)$, i.e. $\iota_A(x_n) = 0$ if $x_n \in A$, $+\infty$ otherwise.

3.1. Properties of the weighted least squares criterion

We now mention some properties of the considered data fidelity term.

Proposition 1. *For every symmetric positive matrix $\boldsymbol{Q}$ in $\mathbb{R}^{N \times N}$, the gradient of $h \circ \boldsymbol{Q}^{1/2}$ is $\mu^{(\boldsymbol{Q})}$-Lipschitz with*

$$\mu^{(\boldsymbol{Q})} = \|\boldsymbol{M}^{1/2} \boldsymbol{H} \boldsymbol{Q}^{1/2}\|^2, \qquad (10)$$

where $\boldsymbol{M} = \mathrm{Diag}(\mu_1, \dots, \mu_N)$ and, for every $n \in \{1, \dots, N\}$,

$$\mu_n = \frac{(\alpha_n y_n + \beta_n)^2}{\beta_n^3}. \qquad (11)$$

Note that h is only defined for non-negative values of its arguments. It can be extended to the whole space $\mathbb{R}^N$ by setting

$$(\forall \boldsymbol{x} \in \mathbb{R}^N) \qquad h(\boldsymbol{x}) = \sum_{n=1}^{N} h^{(n)}([\boldsymbol{H}\boldsymbol{x}]_n), \qquad (12)$$

where, for every $n \in \{1, \dots, N\}$, $h^{(n)}$ is a convex, twice-differentiable function, whose expression is readily derived from (9) for non-negative values of its arguments, and which takes a quadratic form on $(-\infty, 0]$. By appropriately choosing the quadratic form, h is a convex differentiable function with a Lipschitzian gradient on the whole space $\mathbb{R}^N$. Finally, let us emphasize that the proximity operator of h does not have a simple expression, so that the algorithm proposed in [4] does not appear appropriate in this context.

3.2. Implementation of the proximity operators

Here, we provide some additional results concerning the computation of the proximity operators in Algorithm 1.

Proposition 2. *Assume that the preconditioning matrices are such that*

$$(\forall j \in \{1, \dots, J\}) \quad \boldsymbol{R}_j = \rho_j \boldsymbol{I}_{N_j}, \qquad (13)$$
$$\boldsymbol{Q} = \mathrm{Diag}(\kappa_1, \dots, \kappa_N), \qquad (14)$$

with, for every $j \in \{1, \dots, J\}$, $\rho_j \in (0, +\infty)$, and, for every $n \in \{1, \dots, N\}$, $\kappa_n \in (0, +\infty)$. Then, for every $j \in \{1, \dots, J\}$ and for every $\boldsymbol{y}_{2,j} \in \mathbb{R}^{N_j}$,

$$\mathrm{prox}_{\boldsymbol{R}_j^{-1}, \gamma_k g_j^*}(\boldsymbol{y}_{2,j}) =$$
$$\boldsymbol{y}_{2,j} - \gamma_k \rho_j \, \mathrm{prox}_{\gamma_k^{-1} \rho_j^{-1} g_j}(\gamma_k^{-1} \rho_j^{-1} \boldsymbol{y}_{2,j}). \qquad (15)$$

Moreover, if g_0 is separable, i.e.

$$(\forall \boldsymbol{x} = (x_n)_{1 \leq n \leq N} \in \mathbb{R}^N) \quad g_0(\boldsymbol{x}) = \sum_{n=1}^{N} g_0^{(n)}(x_n), \qquad (16)$$

then, for all $\boldsymbol{y}_1 = (y_{1,n})_{1 \leq n \leq N} \in \mathbb{R}^N$,

$$\mathrm{prox}_{\boldsymbol{Q}^{-1}, \gamma_k g_0}(\boldsymbol{y}_1) = (\mathrm{prox}_{\gamma_k \kappa_n g_0^{(n)}}(y_{1,n}))_{1 \leq n \leq N}. \qquad (17)$$

Therefore, provided that diagonal preconditioning matrices of the form (13)-(14) are used, our algorithm still benefits from the low computational cost of its original non preconditioned version.

4. SIMULATION EXAMPLES

We now demonstrate the practical performance of our method on three image restoration scenarios. In our experiments, we use the standard House, Peppers and JetPlane images of size 256×256 from http://sipi.usc.edu/database/. To generate the observed images $\boldsymbol{y}$, we degraded the original images with a blur operator $\boldsymbol{H}$. For each image, a different convolution kernel was considered, namely a truncated Gaussian of standard deviation 1 and size 7×7, a uniform kernel of size 5×5 and a uniform motion of length 7 and angle $20°$. The images were further corrupted with the considered signal-dependent additive noise with $\alpha = 0.1$ and $\beta = 50$ (see (8)). The restoration is performed by solving (1) where function h is given by (9) and $g = \iota_{A^N} + \lambda \operatorname{tv}$ where $A = [0, 255]$, tv denotes the total variation semi-norm and $\lambda > 0$ is the regularization parameter. The quality of the results is evaluated in terms of Signal to Noise Ratio (SNR) and Mean Structural SIMilarity index (MSSIM [15]). For each experiment, parameter λ was adjusted to maximize the SNR between the original and reconstructed images. In Figs. 1, 2 and 3, we display the degraded and reconstructed images and indicate the associated SNR and MSSIM values.

Our choice of regularization term matches with (4) by defining $\mathcal{H} = \mathbb{R}^N$, $g_0 = \iota_{A^N}$, $J = 1$, $N_1 = 2N$ and $\boldsymbol{L}_1 = \left[(\boldsymbol{\Delta}^{\mathrm{h}})^\top \ (\boldsymbol{\Delta}^{\mathrm{v}})^\top \right]^\top$, where $\boldsymbol{\Delta}^{\mathrm{h}} \in \mathbb{R}^{N \times N}$ (resp. $\boldsymbol{\Delta}^{\mathrm{v}} \in \mathbb{R}^{N \times N}$) corresponds to a horizontal (resp. vertical) gradient operator, and, for every $\boldsymbol{x} \in \mathbb{R}^N$, $g_1(\boldsymbol{L}_1 \boldsymbol{x}) = \lambda \sum_{n=1}^N \left(([\boldsymbol{\Delta}^{\mathrm{h}} \boldsymbol{x}]_n)^2 + ([\boldsymbol{\Delta}^{\mathrm{v}} \boldsymbol{x}]_n)^2 \right)^{1/2}$. Since g_0 is the indicator function of a separable closed convex set, its proximity operator reduces to a set of projections onto real intervals. Moreover, the proximity operator of g_1 is easily deduced from the proximity operator of the ℓ_2 norm given in [16]. There remains to set the positive constants $(\kappa_n)_{1 \leqslant n \leqslant N}$ and ρ_1 arising in the expressions (13)-(14) of the preconditioning matrices. A good performance was obtained with the empirical tuning

$$\begin{cases} \kappa_n &= \frac{1}{2\mu_n} \quad (\forall n \in \{1, \ldots, N\}), \\ \rho_1 &= \frac{\|\boldsymbol{L}_1\|^2}{\|\boldsymbol{L}_1 \boldsymbol{Q}^{1/2}\|^2}. \end{cases} \tag{18}$$

For such a choice, (6) becomes

$$\tau = \frac{1}{2} \|\boldsymbol{M}^{1/2} \boldsymbol{H} \boldsymbol{M}^{-1/2}\|^2 + \|\boldsymbol{L}_1\|. \tag{19}$$

Let us emphasize that the spectral norms $\|\boldsymbol{L}_1 \boldsymbol{Q}^{1/2}\|$ and $\|\boldsymbol{M}^{1/2} \boldsymbol{H} \boldsymbol{M}^{-1/2}\|$ can be efficiently computed using the iterative algorithm in [17, Alg. 4].

Fig. 4 illustrates the variations of $\left(f(\boldsymbol{x}_k) - f(\hat{\boldsymbol{x}}) \right)_k$ and $\left(\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\| \right)_k$ using the proposed preconditioned algorithm and its original non preconditioned version, where the optimal solution $\hat{\boldsymbol{x}}$ is precomputed using a large number of iterations. Note that, since $\boldsymbol{Q}$ and $(\boldsymbol{R}_j)_{1 \leq j \leq J}$ have been chosen diagonal, both algorithms require a similar computational time per iteration. For the three experiments, the preconditioning strategy leads to a significant acceleration in terms of decay of both criterion and residual error norm values. Typically, when performing tests on an Intel Xeon, E5440 @ 2.83GHz using a Matlab 7 implementation, a stabilized value of the SNR is obtained after 85 s by the proposed method.

5. CONCLUSION

In the present paper, we have proposed a general preconditioning strategy to improve the performance of the primal-dual proximal splitting algorithm in [1]. We have established some conditions on the algorithm parameters under which the convergence is ensured. Moreover, we have shown that the computational cost of the algorithm can be maintained low by choosing diagonal preconditioners. These theoretical contributions have led to the proposal of an efficient penalized weighted least squares method for restoring an image degraded by a linear operator and a signal-dependent noise. Note that similar preconditioning strategies could be applied to the proximal algorithms in [18, 19].

6. REFERENCES

[1] P. L. Combettes and J.-C. Pesquet, "Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators," *Set-Valued Var. Anal.*, vol. 20, no. 2, pp. 307–330, Jun. 2012.

[2] K. Chen, *Matrix preconditioning techniques and applications*, vol. 19 of *Cambridge Monogr. Appl. Comput. Math.*, Cambridge University Press, 2005.

[3] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer-Verlag, New York, NY, 1999.

[4] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imaging Vision*, vol. 40, no. 1, pp. 120–145, 2010.

[5] T. Pock and A. Chambolle, "Diagonal preconditioning for first order primal-dual algorithms in convex optimization," in *IEEE International Conference on Computer Vision (ICCV)*, Barcelona, Spain, Nov. 6-13, 2011, pp. 1762–1769.

[6] G. E. Healey and R. Kondepudy, "Radiometric CCD camera calibration and noise estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 3, pp. 267–276, Mar. 1994.

[7] H. Tian, B. Fowler, and A. E. Gamal, "Analysis of temporal noise in CMOS photodiode active pixel sensor," *IEEE J. Solid-State Circuits*, vol. 36, no. 1, pp. 92–101, Jan. 2001.

[8] J. R. Janesick, *Photon transfer*, vol. PM170, SPIE Press Book, 2007.

[9] A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian, "Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data," *IEEE Trans. Image Process.*, vol. 17, no. 10, pp. 1737–1754, Oct. 2008.

Fig. 1. Noisy blurred image with Gaussian blur, $\alpha = 0.1$ and $\beta = 50$, SNR=22 dB, MSSIM=0.525 (left) and deblurred image (right) with the proposed algorithm, SNR=27.1 dB, MSSIM=0.835.
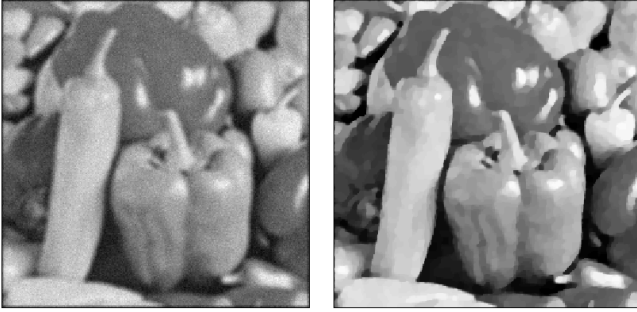

Fig. 2. Noisy blurred image with uniform blur, $\alpha = 0.1$ and $\beta = 50$, SNR=19.15 dB, MSSIM=0.644 (left) and deblurred image (right) with the proposed algorithm, SNR=24.14 dB, MSSIM=0.887.
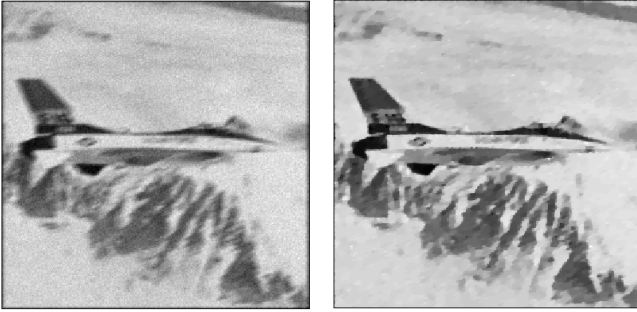

Fig. 3. Noisy blurred image with motion blur, $\alpha = 0.1$ and $\beta = 50$, SNR=20.16 dB, MSSIM=0.569 (left) and deblurred image (right) with the proposed algorithm, SNR=25.11 dB, MSSIM=0.854.

Fig. 4. Comparison of non preconditioned (dashed line) and preconditioned (solid line) versions of our algorithm, for the restoration of House (a,b), Peppers (c,d) and JetPlane (e,f).

[10] J. M. M. Anderson, B. A. Mair, M. Rao, and C.-H. Wu, "Weighted least-squares reconstruction methods for Positron Emission Tomography," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 159–165, Apr. 1997.

[11] C. Chaux, J.-C. Pesquet, and N. Pustelnik, "Nested iterative algorithms for convex constrained image recovery problems," *SIAM J. Imaging Sci.*, vol. 2, no. 2, pp. 730–762, Jun. 2009.

[12] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3042 –3054, Jun. 2010.

[13] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259–268, 1992.

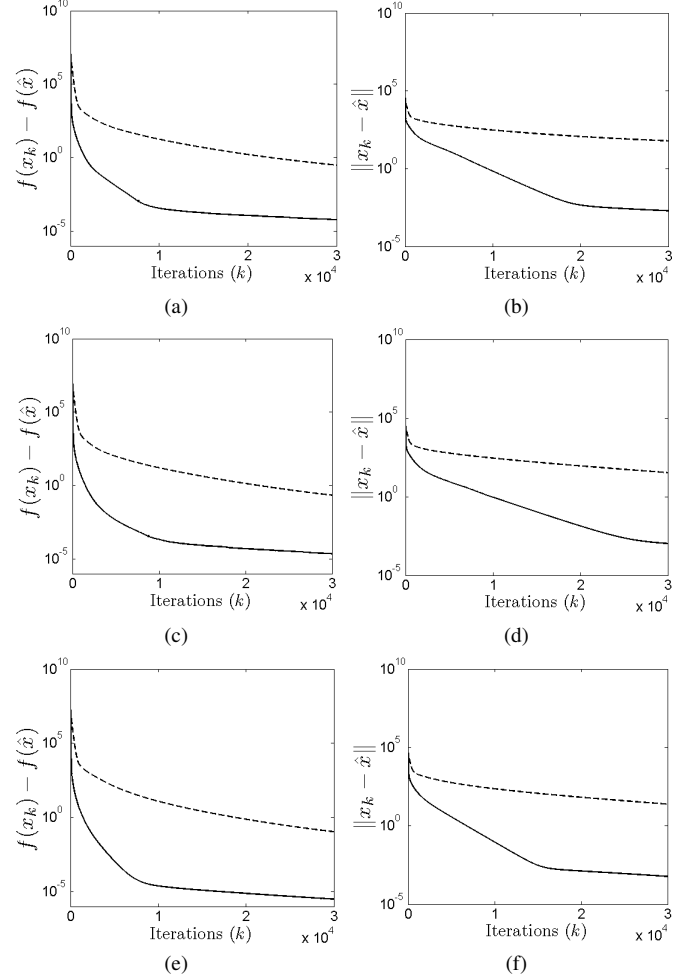[14] N. Pustelnik, C. Chaux, and J.-C. Pesquet, "Parallel proximal algorithm for image restoration using hybrid regularization," *IEEE Trans. Image Process.*, vol. 20, no. 9, pp. 2450–2462, Sep. 2011.

[15] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600 –612, Apr. 2004.

[16] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, H. H. Bauschke, R. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, Eds., pp. 185–212. Springer-Verlag, New York, 2010.

[17] L. Chaâri, N. Pustelnik, C. Chaux, and J.-C. Pesquet, "Solving inverse problems with overcomplete transforms and convex optimization techniques," in *SPIE, Wavelets XIII*, San Diego, CA, USA, Aug. 2-6, 2009, vol. 7446, 14 p.

[18] L. Condat, "A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms," Tech. Rep., 2011, http://hal.archives-ouvertes.fr/hal-00609728.

[19] H. Raguet, J. Fadili, and G. Peyré, "Generalized Forward-Backward splitting," Tech. Rep., 2011, http://hal.archives-ouvertes.fr/hal-00613637.