PIPA: A NEW PROXIMAL INTERIOR POINT ALGORITHM FOR LARGE-SCALE CONVEX OPTIMIZATION

Marie-Caroline Corbineau[1], Emilie Chouzenoux[1,2], Jean-Christophe Pesquet[1]

[1] University of Paris-Saclay, CentraleSupelec, CVN, Gif-sur-Yvette, France
[2] University of Paris-Est Marne la Vallee, LIGM, UMR CNRS 8049, Champs-sur-Marne, France
{marie-caroline.corbineau, emilie.chouzenoux, jean-christophe.pesquet}@centralesupelec.fr

ABSTRACT

Interior point methods have been known for decades to be useful for the resolution of small to medium size constrained optimization problems. These approaches have the benefit of ensuring feasibility of the iterates through a logarithmic barrier. We propose to incorporate a proximal forward-backward step in the resolution of the barrier subproblem to account for non-necessarily differentiable terms arising in the objective function. The combination of this scheme with a novel line-search strategy gives rise to the so-called Proximal Interior Point Algorithm (PIPA) suitable for the minimization of the sum of a smooth convex function and a non-smooth convex one under general convex constraints. The convergence of PIPA is secured under mild assumptions. As demonstrated by numerical experiments carried out on a large-scale hyperspectral image unmixing application, the proposed method outperforms the state-of-the-art.

Index Terms— Convex optimization, Large-scale problems, Interior point methods, Proximal algorithms, Hyperspectral unimixing.

1. INTRODUCTION

Data processing problems frequently involve the minimization of a cost function F of a variable x belonging to $\mathbb{R}^n$ where n is large, in a feasible domain defined by some convex inequalities: $c_i(x) \leq 0$, $i \in \{1, \ldots, p\}$. One way of reformulating this constrained problem as an unconstrained one is to consider the minimization of $F + B$, where function B is unbounded at the boundary of the feasible domain and thus acts as a *barrier*. Interior point methods [1, 2] rely on this strategy by setting B as a logarithmic term $-\sum_i \ln(-c_i(.))$ weigthed by a so-called *barrier parameter* $\mu > 0$. The unconstrained minimization of this augmented criterion is then carried out for a sequence of barrier parameter values decreasing to 0. Following this general strategy, many approaches have been proposed, the most commonly used being primal-dual algorithms [3, 4, 5] which, under suitable assumptions, converge superlinearly for nonlinear programming [6]. From a numerical viewpoint, interior point methods have demonstrated state-of-the-art performance on several applications, see e.g. [7]. It should be noted that most of interior point algorithms rely on second-order methods and, thus, usually assume that the criterion is at least twice-differentiable.

For various ill-posed problems arising in signal processing, the quality of the solution and its robustness against noise can be significantly improved by adding a non-differentiable regularization term in the objective function. For instance, sparsity and smoothness can be promoted through the ℓ_1-norm and the total variation semi-norm [8], respectively. Hence, it is necessary to design efficient algorithms that can minimize a non-necessarily smooth function, under constraints. One can mention the work in [9, 10, 11] which shows that interior point methods can cope with an ℓ_1 term by introducing additional variables or by using the subdifferential of the ℓ_1-norm. A more efficient way of addressing arbitratry non-differentiable functions is provided by the proximity operator [12], which is the cornerstone of proximal algorithms. Several proximal algorithms [13, 14] like the primal-dual splitting algorithm [15] or the generalized forward-backward splitting algorithm [16] can minimize efficiently a composite objective function under constraints. However, they usually do not ensure the feasibility of each iterate. One can mention Bregman distance approaches which led to entropy-like proximal algorithms, where the Euclidean norm in the definition of the proximity operator is replaced by a divergence measure. Although this measure can be chosen such that constraints are satisfied [17], the computation of the modified proximity operator can be quite intricate.

In this paper we propose a Proximal Interior-Point Algorithm (PIPA) for composite constrained optimization problems. The main idea behind our algorithm is to combine the logarithmic barrier method, which ensures the feasibility of the iterates, with a proximal forward-backward step which can deal with a non-smooth term arising in the objective function. The convergence of PIPA is secured under mild assumptions and the proposed method is shown to exhibit good performance on a large-scale hyperspectral image unmixing application.

This article is organized as follows. The optimization problem is formulated in Section 2, while the algorithm and

convergence results are given in Section 3. Finally, PIPA is compared to several state-of-the-art algorithms in Section 4.

2. PROBLEM FORMULATION

2.1. Notation and definitions

Let $(n, p) \in \mathbb{N}^2$ and let $[\![p]\!] = \{1, ..., p\}$. Let $\Gamma_0(\mathbb{R}^n)$ denote the set of functions which are proper, convex, lower semi-continuous on $\mathbb{R}^n$ and take values in $\mathbb{R} \cup \{+\infty\}$. We denote by $\mathcal{S}_{++}^n$ the set of positive definite matrices in $\mathbb{R}^{n \times n}$.
We also define the proximity operator [12]: for every $\gamma \geq 0$, $A \in \mathcal{S}_{++}^n$, $f \in \Gamma_0(\mathbb{R}^n)$, and $x \in \mathbb{R}^n$,

$$\text{prox}_{\gamma f}^A(x) = \arg\min_{y \in \mathbb{R}^n} \left(\frac{1}{2} \|y - x\|_A^2 + \gamma f(y) \right). \quad (1)$$

2.2. Problem formulation

Our objective is to solve the following composite optimization problem, which is of prominent use in signal/image processing:

$$\mathcal{P}_0 : \quad \underset{x \in \mathbb{R}^n}{\text{minimize}} \, f(x) + g(x)$$
$$\text{s.t.} \; (\forall i \in [\![p]\!]) \, c_i(x) \leq 0 \quad (2)$$

where f and g are convex, f is non-necessarily differentiable, and g is smooth. The constraint functions $(c_i)_{i \in [\![p]\!]}$ are assumed to be convex and twice-differentiable. More precise assumptions under which convergence is secured are detailed in Section 3.1. In the proposed approach, the constrained problem $\mathcal{P}_0$ is solved via a sequence $(\mathcal{P}_{\mu_j})_{j \in \mathbb{N}}$ of unconstrained subproblems parametrized by the barrier coefficient $\mu_j > 0$ and defined as follows:

$$\mathcal{P}_{\mu_j} : \quad \underset{x \in \mathbb{R}^n}{\text{minimize}} \, f(x) + g(x) + \mu_j b(c(x)) \quad (3)$$

where $c(x) = (c_i(x))_{i \in [\![p]\!]}$ and b is the logarithmic barrier: $(\forall z \in \mathbb{R}^p) \, b(z) = -\sum_{i=1}^p \ln(-z_i)$ if $z \in]-\infty, 0[^p$, $+\infty$ otherwise. As the subproblems $(\mathcal{P}_{\mu_j})_{j \in \mathbb{N}}$ do not have an explicit solution, we will need to solve them approximately. This leads to the optimization scheme which is described in the next section.

3. PROPOSED ALGORITHM

The proposed PIPA method along with its related parameters are detailed in Algorithm 1. Let us now comment the main steps of the algorithm.
PIPA comprises two main interlocked loops. Given $j \in \mathbb{N}$, the inner iterations produce an approximate solution to $\mathcal{P}_{\mu_j}$ via a proximal forward-backward step [18]. The latter corresponds to a gradient step on the smooth term $\varphi_{\mu_j} = g(.) + \mu_j b(c(.))$ and a proximal step on the non-differentiable term f. The proximity operator is associated with a variable metric $A_{j,k}$ which should satisfy the following technical condition:

Algorithm 1: Proximal Interior-Point Algorithm (PIPA)

Let $\{(\epsilon_{i,j})_{j \in \mathbb{N}}\}_{i \in [\![4]\!]}$, $(\mu_j)_{j \in \mathbb{N}}$ satisfy Condition 3.2;
Take $0 < \delta < 1$, $0 < \theta < 1$, $0 < \bar{\gamma}$ and $0 < \epsilon$;
Initialize $x_{0,0} \in \mathbb{R}^n$ such that $(\forall i \in [\![p]\!]) \, c_i(x_{0,0}) < 0$;
for $j = 0, 1, \ldots$ **do**
   for $k = 0, 1, \ldots$ **do**
      Choose $A_{j,k}$ according to Condition 3.1;
      for $l = 0, 1, \ldots$ **do**
         $\tilde{x}_{j,k}^l = \text{prox}_{\bar{\gamma}\theta^l f}^{A_{j,k}}(x_{j,k} - \bar{\gamma}\theta^l A_{j,k}^{-1} \nabla \varphi_{\mu_j}(x_{j,k}))$;
         Stop if (4) is satisfied;
      end
      $x_{j,k+1} = \tilde{x}_{j,k}^l$;
      $\gamma_{j,k} = \bar{\gamma}\theta^l$;
      Stop if (5)-(8) are satisfied;
   end
   $x_{j+1,0} = x_{j,k+1}$;
end

Condition 3.1. *For every $j \in \mathbb{N}$ $(\exists (\underline{\nu}, \bar{\nu}) \in \mathbb{R}_{++}^2)$ such that $(\forall k \in \mathbb{N}) \, A_{j,k} \in \mathcal{S}_{++}^n$ and $\underline{\nu} I_n \preceq A_{j,k} \preceq \bar{\nu} I_n$.*

If well-chosen, this variable metric can significantly improve the convergence speed. Note that even if the proximity operator of f has a closed-form expression, the same operator used with a variable metric $\text{prox}_f^{A_{j,k}}$ is not necessarily explicit but can be calculated efficiently using the dual forward-backward algorithm in [19] or its accelerated version [20]. It must be emphasized that, since the barrier is logarithmic, the gradient of φ_{μ_j} is not Lipschitz-continuous. Thus, the forward-backward splitting proximal algorithm [13] needs to be associated with a linesearch to find an appropriate value for the stepsize $\gamma_{j,k}$ [21]. We use the linesearch method investigated in [22] (being a generalization of the one introduced in [23]). If the cost function in $\mathcal{P}_0$ is smooth, then this procedure reduces to an Armijo linesearch along the steepest direction. The backtracking procedure stops if

$$\varphi_{\mu_j}(\tilde{x}_{j,k}^l) - \varphi_{\mu_j}(x_{j,k}) - \langle \tilde{x}_{j,k}^l - x_{j,k}, \nabla \varphi_{\mu_j}(x_{j,k}) \rangle$$
$$\leq \frac{\delta}{\bar{\gamma}\theta^l} \|\tilde{x}_{j,k}^l - x_{j,k}\|_{A_{j,k}}^2. \quad (4)$$

The inner loop stops once a certain accuracy is reached for the minimization of $\mathcal{P}_{\mu_j}$. In addition, the barrier parameter is decreased in the outer loop when the following stopping criteria are met:

$$\|x_{j,k} - x_{j,k+1}\| \leq \epsilon_{1,j} \quad (5)$$

$$\frac{1}{\gamma_{j,k}} \|A_{j,k}(x_{j,k} - x_{j,k+1})\| \leq \epsilon_{2,j} \quad (6)$$

$$\sum_{i=1}^p \left| \frac{c_i(x_{j,k+1})}{c_i(x_{j,k})} - 1 \right| \leq \epsilon_{3,j} \quad (7)$$

$$\mu_j \left\| \sum_{i=1}^{p} \frac{\nabla c_i(x_{j,k}) - \nabla c_i(x_{j,k+1})}{c_i(x_{j,k})} \right\| \leq \epsilon_{4,j}. \qquad (8)$$

As stated below, the sequences $\{(\epsilon_{i,j})_{j \in \mathbb{N}}\}_{i \in [\![4]\!]}$ and the barrier parameter $(\mu_j)_{j \in \mathbb{N}}$ have to be chosen properly to secure the convergence of the sequence of iterates produced by PIPA to a solution to the initial problem $\mathcal{P}_0$.

Condition 3.2. $\{(\epsilon_{i,j})_{j \in \mathbb{N}}\}_{i \in [\![4]\!]}$ and $(\mu_j)_{j \in \mathbb{N}}$ are sequences in $\mathbb{R}_{++}$ converging to 0 such that $(\forall i \in [\![4]\!]) \lim_{j \to \infty} \epsilon_{i,j}/\mu_j = 0$.

3.1. Convergence results

Under the following assumptions, we were able to prove the convergence of PIPA:

Assumption 3.1.

(i) *The set of solutions to $\mathcal{P}_0$ is non empty and bounded ;*

(ii) *f and g are in $\Gamma_0(\mathbb{R}^n)$, bounded from below and $\overline{\mathcal{D}} \subset \mathrm{dom}f \subset \mathrm{dom}g$ where $\mathcal{D} = \{x \in \mathbb{R}^n | c(x) < 0\} \neq \emptyset$;*

(iii) *g is differentiable on its domain and gradient-Lipschitz, and $(\forall i \in [\![p]\!]) \; c_i \in \Gamma_0(\mathbb{R}^n)$ is continuously twice-differentiable on $\mathbb{R}^n$;*

(iv) *$(\forall j \in \mathbb{N}) \; f + \varphi_{\mu_j}$ is a Kurdyka-Łojasiewicz (KL) function.*
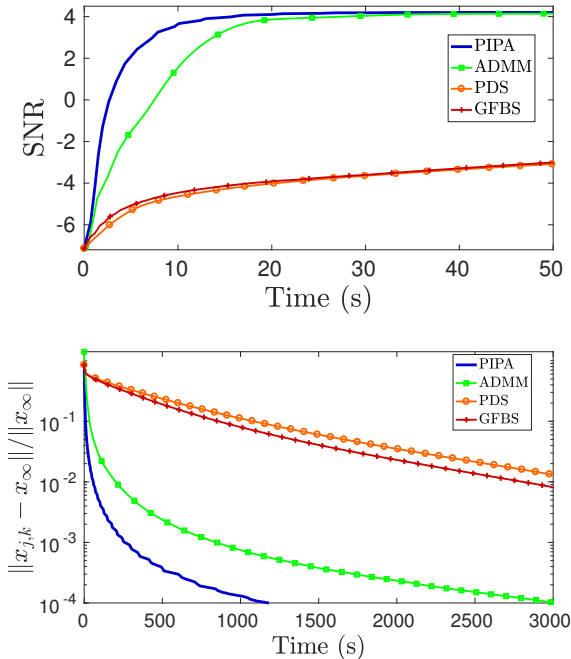
Intuitively, a function satisfies the KL property [24] of Assumption 3.1 (iv) when it can be re-parametrized such that it is sharp around its stationary points. This property is satisfied by a wide range of functions, including the logarithm, all semi-algebraic functions and most of the functions that are commonly used in signal or image processing.

Theorem 3.1. *Under Assumption 3.1 and Condition 3.1,*

(i) *$(\forall j \in \mathbb{N}) \; (x_{j,k})_{k \in \mathbb{N}}$ converges to a solution to $\mathcal{P}_{\mu_j}$;*

(ii) *if in addition Condition 3.2 holds then $(x_{j,0})_{j \in \mathbb{N}}$ is bounded and every limit point is a solution to $\mathcal{P}_0$;*

(iii) *if Condition 3.2 and strict complementarity hold, and if there exists $i \in [\![p]\!]$ such that c_i is strictly convex,[1] then $(x_{j,0})_{j \in \mathbb{N}}$ converges to a solution to $\mathcal{P}_0$.*

For the sake of brevity, proofs are not provided in this paper but some ingredients can be found in [22, 4, 24]. Note that the convergence of $(x_{j,k})_{k \in \mathbb{N}}$ for a fixed barrier parameter μ_j requires to develop a more general framework than in [22] since the variable metrics $(A_{j,k})_{k \in \mathbb{N}}$ are only subject to a boundedness condition. It is also worth noticing that the strict convexity requirement in Theorem 3.1 (iii) can be fullfilled by adding an harmless constraint on the maximal value of $\|x\|^2$.

---

[1] Alternatively, the constraint functions can be linear provided that some full rank property holds.

4. NUMERICAL EXPERIMENTS

4.1. Hyperspectral unmixing model

Imaging spectrometers measure the electromagnetic energy in several spectral bands with high precision but low spatial resolution [25]. Unmixing is then necessary to retrieve the fractional abundances of the pure spectral signatures, also called endmembers. Under the assumption that there is no microscopic interaction between the materials, a linear mixing model can be used to describe the data. The resulting linear unmixing problem can be formulated as the minimization of a penalized least squares criterion under linear constraints [8, 26]:

$$\begin{aligned} \underset{X \in \mathbb{R}^{p \times n}}{\text{minimize}} \quad & \tfrac{1}{2}\|Y - SX\|_2^2 + \kappa \sum_{i=1}^{p} \|WX_i\|_1 \\ s.t. \quad & (\forall j \in [\![n]\!]) \; \sum_{i=1}^{p} X_{i,j} \leq 1 \\ & (\forall i \in [\![p]\!])(\forall j \in [\![n]\!]) \; X_{i,j} \geq 0 \end{aligned} \qquad (9)$$

where p, n and l are respectively the number of endmembers, pixels and spectral bands. $Y \in \mathbb{R}^{l \times n}$ is the observation, $S \in \mathbb{R}^{l \times p}$ is the library that contains the reflectance of the endmembers in every spectral band, $(\forall i \in [\![p]\!]) \; X_i$ is the ith line of the abundance matrix which contains the percentage of the materials in every pixel, and $W \in \mathbb{R}^{n \times n}$ is a wavelet transform operator. In the rest of this section, x denotes the vectorization of X. The linear constraints account for the fact that, for every pixel, the sum of all fractional abundances should be less than one in view of atmospheric absorption.

4.2. Realistic data simulation

We propose to evaluate the performance of PIPA algorithm for the resolution of Problem (9). The observation Y is simulated using the Urban[2] data set. This data set contains the reflectance spectra and abundance maps of $p = 6$ endmembers: asphalt, grass, tree, roof, metal, and dirt. The spectral signatures contain $l = 162$ spectral bands. An artificial atmospheric attenuation is applied on the abundance maps of size 256×256. The result of the linear mixture of the spectral library and attenuated abundance maps is corrupted with a zero-mean white Gaussian noise with variance equal to 2.9×10^{-3}. The regularization weight $\kappa = 8 \times 10^{-3}$ has been set such that it yields the smallest reconstruction error. We choose an orthogonal Daubechies 4 wavelet decomposition performed over 2 resolution levels. The algorithm is implemented with the following parameters:

μ_0	δ	θ	$\bar{\gamma}$	η_1	η_2	η_3	ρ	ζ
0.01	0.99	0.8	0.4	10^5	10^8	10^8	1.5	1.02

The variable metric $A_{j,k}$ is simply chosen equal to $\nabla^2 \varphi_{\mu_j}(x_{j,k})$. The operator $\mathrm{prox}_f^{A_{j,k}}$ is computed numerically thanks to the algorithm in [19] with an accuracy proportional to μ_j.

---

[2] http://www.escience.cn/people/feiyunZHU/Dataset_GT.html

Fig. 1. (top) SNR versus time. (bottom) Distance from the iterates to the solution versus time.

We compare PIPA with three state-of-the-art convex optimization algorithms: the alternating direction of multipliers method splitting algorithm (ADMM) [27] which was applied to hyperspectral unmixing (see e.g. [8]), the primal-dual splitting algorithm (PDS) in [15], and the generalized forward-backward splitting algorithm (GFBS) [16]. The algorithms are implemented on Matlab R2016b and the calculations are performed on a desktop having an Intel Xeon 3.2 GHz processor and 16 GB of RAM.

4.3. Results

The unmixing quality is evaluated based on the signal-to-noise ratio (SNR) where $\bar{x}$ is the ground truth :

$$\text{SNR} = -20 \log_{10}\left(\frac{\|x_{j,k} - \bar{x}\|_2}{\|\bar{x}\|_2}\right). \quad (10)$$

The solution to (9) corresponds to an SNR of 4.18 dB while solving the same optimization problem without the regularization using the approach from [26] gives a lower SNR of 2.98 dB. Figure 1(top) shows that the SNR increases faster with PIPA than with the three other algorithms, which is confirmed by the visual inspection of Figure 2 obtained after running the various algorithms for the same duration (16 s).

The algorithms are also evaluated based on their point-wise convergence. As one can see on Figure 1(bottom), the sequence $(x_{j,k})_{j,k}$ generated by PIPA converges faster to its limit point x_∞ (computed after a very large number of iterations) than the iterates produced by the three other algorithms.

(a)　　　　　　　　　(b)

(c)　　　　　　　　　(d)

(e)　　　　　　　　　(f)
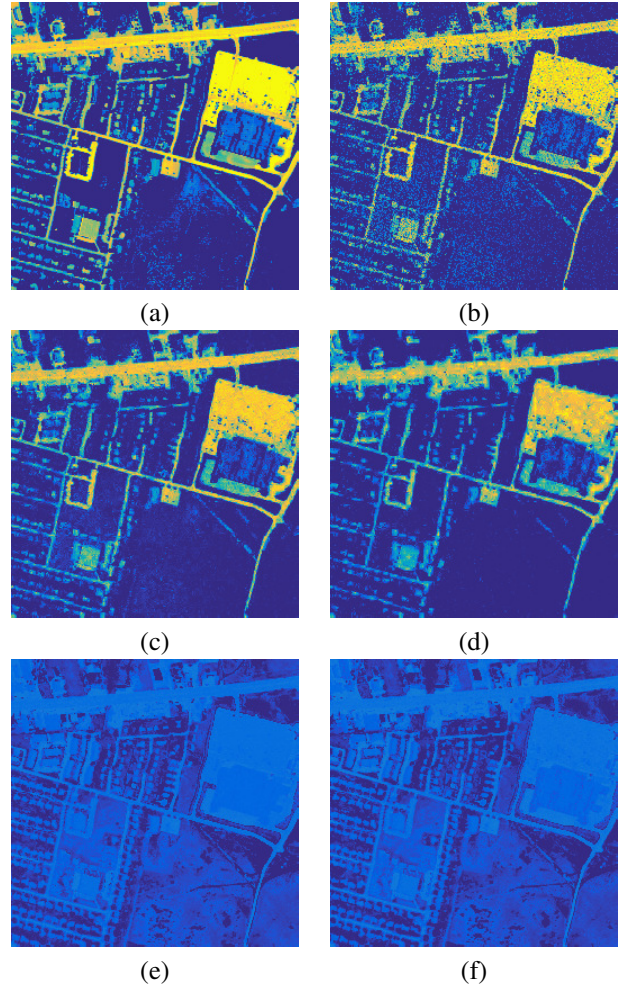
Fig. 2. Abundance map of Asphalt. Ground truth (a), non-regularized solution SNR= 2.98 dB (b). After 16 s with : PIPA SNR= 3.99 dB (c), ADMM SNR= 3.36 dB (d), PDS SNR= -4.22 dB (e), GFBS SNR= -4.01 dB (f).

In this example the proximity operator has been approximated but PIPA appears to be robust to the generated error. Also note that, even if the sequences generated by ADMM, PDS, and GFBS converge asymptotically to a feasible point, after 16 s, the distance between their iterates and the feasible set is respectively 4.4, 0.7 and 0.1.

5. CONCLUSION

This paper has introduced a new method to minimize a sum of two convex functions, one being non-smooth, subject to constraints. This algorithm compares favorably with the state-of-the-art on a large-scale image processing problem. One future direction for further improvements is to investigate theoretically the effect of an inexact proximity operator implementation. Another lead is to remove the convexity assumptions so as to enlarge the application field of the proposed approach.

6. REFERENCES

[1] M. H. Wright, "Interior methods for constrained optimization," *Acta Numerica*, vol. 1, pp. 341–407, 1991.

[2] J. Gondzio, "Interior point methods 25 years later," *European Journal of Operational Research*, vol. 218, no. 3, pp. 587–601, 2012.

[3] A. Forsgren, P. E. Gill, and M. H. Wright, "Interior methods for nonlinear optimization," *SIAM Review*, vol. 44, no. 4, pp. 525–597, 2002.

[4] P. Armand, J. C. Gilbert, and S. Jan-Jégou, "A feasible BFGS interior point algorithm for solving convex minimization problems," *SIAM Journal on Optimization*, vol. 11, no. 1, pp. 199–222, 2000.

[5] S. Bonettini and T. Serafini, "Non-negatively constrained image deblurring with an inexact interior point method," *Journal of Computational and Applied Mathematics*, vol. 231, no. 1, pp. 236–248, 2009.

[6] N. I.M. Gould, D. Orban, A. Sartenaer, and P. L. Toint, "Superlinear convergence of primal-dual interior point algorithms for nonlinear programming," *SIAM Journal on Optimization*, vol. 11, no. 4, pp. 974–1002, 2001.

[7] C. A. Johnson, J. Seidel, and A. Sofer, "Interior-point methodology for 3-D PET reconstruction," *IEEE Transactions on Medical Imaging*, vol. 19, no. 4, pp. 271–285, 2000.

[8] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 11, pp. 4484–4502, 2012.

[9] H. Fu, M. K. Ng, M. Nikolova, and J. L. Barlow, "Efficient minimization methods of mixed ℓ_2-ℓ_1 and ℓ_1-ℓ_1 norms for image restoration," *SIAM Journal on Scientific Computing*, vol. 27, no. 6, pp. 1881–1902, 2006.

[10] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An interior-point method for large-scale ℓ_1-regularized least squares," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 606–617, 2007.

[11] K. Fountoulakis and J. Gondzio, "Performance of first-and second-order methods for ℓ_1-regularized least squares problems," *Computational Optimization and Applications*, vol. 65, no. 3, pp. 605–635, 2016.

[12] H. H. Bauschke and P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*, Springer New York, 2011.

[13] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-point algorithms for inverse problems in science and engineering*, pp. 185–212. Springer New York, 2011.

[14] N. Komodakis and J.-C. Pesquet, "Playing with duality: An overview of recent primal-dual approaches for solving large-scale optimization problems," *IEEE Signal Processing Magazine*, vol. 32, no. 6, pp. 31–54, 2015.

[15] P. L. Combettes, L. Condat, J.-C. Pesquet, and B.C. Vũ, "A forward-backward view of some primal-dual optimization methods in image recovery," in *Proceedings of the IEEE International Conference on Image Processing (ICIP 2014)*, Paris, France, 27-30 October 2014, pp. 4141–4145.

[16] H. Raguet, J. Fadili, and G. Peyré, "A generalized forward-backward splitting," *SIAM Journal on Imaging Sciences*, vol. 6, no. 3, pp. 1199–1226, 2013.

[17] A. N. Iusem, B. F. Svaiter, and M. Teboulle, "Entropy-like proximal methods in convex programming," *Mathematics of Operations Research*, vol. 19, no. 4, pp. 790–814, 1994.

[18] E. Chouzenoux, J.-C. Pesquet, and A. Repetti, "Variable metric forward-backward algorithm for minimizing the sum of a differentiable function and a convex function," *Journal of Optimization Theory and Applications*, vol. 162, no. 1, pp. 107–132, 2014.

[19] P. L. Combettes, D. Dũng, and B. C. Vũ, "Proximity for sums of composite functions," *Journal of Mathematical Analysis and Applications*, vol. 380, no. 2, pp. 680–688, 2011.

[20] F. Abboud, E. Chouzenoux, J.-C. Pesquet, J.-H. Chenot, and L. Laborelli, "Dual block-coordinate forward-backward algorithm with application to deconvolution and deinterlacing of video sequences," *Journal of Mathematical Imaging and Vision*, vol. 59, no. 3, pp. 1–17, 2017.

[21] E. Chouzenoux, S. Moussaoui, and J. Idier, "Majorize–minimize linesearch for inversion methods involving barrier function optimization," *Inverse Problems*, vol. 28, no. 6, pp. 065011, 2012.

[22] S. Salzo, "The variable metric forward-backward splitting algorithm under mild differentiability assumptions," *SIAM Journal on Optimization*, vol. 27, no. 4, 2017.

[23] J. Y. Bello Cruz and T. T.A. Nghia, "On the convergence of the forward-backward splitting method with linesearches," *Optimization Methods and Software*, vol. 31, no. 6, pp. 1209–1238, 2016.

[24] J. Bolte, S. Sabach, and M. Teboulle, "Proximal alternating linearized minimization for nonconvex and nonsmooth problems," *Mathematical Programming*, vol. 146, no. 1-2, pp. 459–494, 2014.

[25] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354–379, 2012.

[26] E. Chouzenoux, M. Legendre, S. Moussaoui, and J. Idier, "Fast constrained least squares spectral unmixing using primal-dual interior-point optimization," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 1, pp. 59–69, 2014.

[27] S. Setzer, G. Steidl, and T. Teuber, "Deblurring Poissonian images by split Bregman techniques," *Journal of Visual Communication and Image Representation*, vol. 21, no. 3, pp. 193–199, 2010.