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]CVN, CentraleSupélec, Université Paris-Saclay, France
[2]LIGM, UMR CNRS 8049, Université Paris-Est Marne la Vallée, France

19 April 2018
Calgary, ICASSP 2018

In collaboration with

E. Chouzenoux

J.-C. Pesquet

Interior Point Methods

Many problems in signal/image processing (image restoration, enhancement, denoising/deblurring, spectral unmixing) can be formulated as **constrained minimization problems** $\rightarrow$ need efficient methods for solving those.

Interior Point Methods

Many problems in signal/image processing (image restoration, enhancement, denoising/deblurring, spectral unmixing) can be formulated as **constrained minimization problems** $\rightarrow$ need efficient methods for solving those.

Constrained Problem

$$\mathcal{P}_0 : \quad \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x) \quad \text{s.t.} \quad (\forall i \in \{1, \ldots, p\}) \;\; c_i(x) \leq 0$$

where

- $f : \mathbb{R}^n \mapsto]-\infty, +\infty]$ convex
- $(\forall i \in \{1, \ldots, p\})$ $c_i : \mathbb{R}^n \mapsto]-\infty, +\infty]$ convex, smooth

Interior Point Methods

Many problems in signal/image processing (image restoration, enhancement, denoising/deblurring, spectral unmixing) can be formulated as **constrained minimization problems** $\rightarrow$ need efficient methods for solving those.

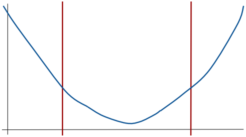
Constrained Problem

$$\mathcal{P}_0 : \quad \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x) \quad \text{s.t.} \quad (\forall i \in \{1, \ldots, p\}) \;\; c_i(x) \leq 0$$

where

- $f : \mathbb{R}^n \mapsto]-\infty, +\infty]$ convex
- $(\forall i \in \{1, \ldots, p\})$ $c_i : \mathbb{R}^n \mapsto]-\infty, +\infty]$ convex, smooth

How to minimize f while ensuring that every iterate is feasible ?



Interior Point Methods

Many problems in signal/image processing (image restoration, enhancement, denoising/deblurring, spectral unmixing) can be formulated as **constrained minimization problems** → need efficient methods for solving those.

Constrained Problem

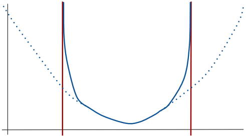
$$\mathcal{P}_0 : \quad \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad f(x) \qquad s.t. \quad (\forall i \in \{1,\dots,p\}) \;\; c_i(x) \leq 0$$

where

- $f : \mathbb{R}^n \mapsto]-\infty, +\infty]$ convex
- $(\forall i \in \{1,\dots,p\})$ $c_i : \mathbb{R}^n \mapsto]-\infty, +\infty]$ convex, smooth

How to minimize f while ensuring that every iterate is feasible ?

→ Add a barrier function



Logarithmic Barrier

Constrained Problem

$$\mathcal{P}_0 : \quad \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x) \quad \text{s.t.} \quad (\forall i \in \{1, \dots, p\}) \;\; c_i(x) \leq 0$$

$$\Downarrow$$

Unconstrained Subproblem

$$\mathcal{P}_\mu : \quad \underset{x \in \mathbb{R}^n}{\text{minimize}} \;\; f(x) \;\; \underbrace{- \mu \sum_{i=1}^{p} \ln(-c_i(x))}_{\to +\infty \text{ as } c_i(x) \to 0^-}$$

Where $\mu > 0$ is the barrier parameter.

$\mathcal{P}_0$ is replaced by a sequence of subproblems $(\mathcal{P}_{\mu_j})_{j \in \mathbb{N}}$.

Logarithmic Barrier

Constrained Problem

$$\mathcal{P}_0 : \quad \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x) \quad \text{s.t.} \quad (\forall i \in \{1, \dots, p\}) \;\; c_i(x) \leq 0$$

$\Downarrow$

Unconstrained Subproblem

$$\mathcal{P}_\mu : \quad \underset{x \in \mathbb{R}^n}{\text{minimize}} \;\; f(x) \quad \underbrace{- \mu \sum_{i=1}^{p} \ln(-c_i(x))}_{\to +\infty \text{ as } c_i(x) \to 0^-}$$

Where $\mu > 0$ is the barrier parameter.

$\mathcal{P}_0$ is replaced by a sequence of subproblems $(\mathcal{P}_{\mu_j})_{j \in \mathbb{N}}$.

- Subproblems are solved approximately for a sequence $\mu_j \to 0$.
- Main advantage : every iterate is feasible.
- Primal-dual algorithm : superlinear convergence for NLP. **[Gould *et al.*, 2001]**

Logarithmic Barrier

Constrained Problem

$$\mathcal{P}_0 : \quad \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad f(x) \qquad \text{s.t.} \quad (\forall i \in \{1,\ldots,p\}) \ \ c_i(x) \leq 0$$

$\Downarrow$

Unconstrained Subproblem

$$\mathcal{P}_\mu : \quad \underset{x\in\mathbb{R}^n}{\text{minimize}} \ f(x) \quad \underbrace{- \boldsymbol{\mu} \sum_{i=1}^{p} \ln\left(-c_i(x)\right)}_{\to +\infty \text{ as } c_i(x) \to 0^-}$$

Where $\mu > 0$ is the barrier parameter.

$\mathcal{P}_0$ is replaced by a sequence of subproblems $(\mathcal{P}_{\mu_j})_{j\in\mathbb{N}}$.

- Subproblems are solved approximately for a sequence $\mu_j \to 0$.
- Main advantage : every iterate is feasible.
- Primal-dual algorithm : superlinear convergence for NLP. **[Gould *et al.*, 2001]**

✗ Require the inversion of an $n \times n$ matrix at each step : medium size applications.
✗ First or second order methods : limited to smooth functions. **[Armand *et al.*, 2000]**

Problem of Interest

Quality of the solution and robustness against noise can be improved by adding a **non-differentiable term** (ℓ_1, TV, ...).

Composite Constrained Problem

$$\mathcal{P}_0 : \quad \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x) + g(x) \quad \text{s.t.} \quad (\forall i \in \{1, \dots, p\}) \;\; c_i(x) \leq 0$$

where

- $f : \mathbb{R}^n \mapsto]-\infty, +\infty]$ convex, **non-differentiable**
- $g : \mathbb{R}^n \mapsto]-\infty, +\infty]$ convex, smooth
- $(\forall i \in \{1, \dots, p\})$ $c_i : \mathbb{R}^n \mapsto]-\infty, +\infty]$ convex, smooth

Problem of Interest

Quality of the solution and robustness against noise can be improved by adding a **non-differentiable term** (ℓ_1, TV, ...).

Composite Constrained Problem

$$\mathcal{P}_0 : \quad \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x) + g(x)$$
$$\text{s.t.} \quad (\forall i \in \{1, \dots, p\}) \quad c_i(x) \leq 0$$

where

- $f : \mathbb{R}^n \mapsto]-\infty, +\infty]$ convex, **non-differentiable**
- $g : \mathbb{R}^n \mapsto]-\infty, +\infty]$ convex, smooth
- $(\forall i \in \{1, \dots, p\})$ $c_i : \mathbb{R}^n \mapsto]-\infty, +\infty]$ convex, smooth

How to address the non-smooth term while ensuring that every iterate is feasible ?

Problem of Interest

Quality of the solution and robustness against noise can be improved by adding a **non-differentiable term** (ℓ_1, TV, ...).

Composite Constrained Problem

$$\mathcal{P}_0 : \quad \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x) + g(x) \qquad \text{s.t.} \quad (\forall i \in \{1, \dots, p\}) \;\; c_i(x) \leq 0$$

where

- $f : \mathbb{R}^n \mapsto]-\infty, +\infty]$ convex, **non-differentiable**
- $g : \mathbb{R}^n \mapsto]-\infty, +\infty]$ convex, smooth
- $(\forall i \in \{1, \dots, p\})$ $c_i : \mathbb{R}^n \mapsto]-\infty, +\infty]$ convex, smooth

How to address the non-smooth term while ensuring that every iterate is feasible ?

→ Combine the logarithmic barrier method with proximal tools.

Notation and Definitions

- Let $\mathcal{S}^+(\mathbb{R}^N)$ be the set of symmetric positive definite matrices of $\mathbb{R}^{N\times N}$.
- The weighted norm induced by $U \in \mathcal{S}^+(\mathbb{R}^N)$ is $\|.\|_U = \sqrt{\langle . \mid U . \rangle}$.
- Let $\Gamma_0(\mathbb{R}^N)$ denote the set of proper lower semicontinuous convex functions from $\mathbb{R}^N$ to $]-\infty, +\infty]$.

Proximity Operator

The **proximal operator** [a] $\mathrm{prox}_f^U(x)$ of $f \in \Gamma_0(\mathbb{R}^N)$ at $x \in \mathbb{R}^N$ relative to the metric induced by $U \in \mathcal{S}^+(\mathbb{R}^N)$ is the unique vector $\widehat{y} \in \mathbb{R}^N$ such that

$$f(\widehat{y}) + \frac{1}{2}\|\widehat{y} - x\|_U^2 = \inf_{y\in\mathbb{R}^N} f(y) + \frac{1}{2}\langle y - x \mid U(y-x)\rangle.$$

a. http ://proximity-operator.net/

Example :

- Indicator function : projection.
- ℓ_1 norm : soft-thresholding.

Proposed Approach

$\mathcal{P}_0$ is replaced by a sequence of subproblems $(\mathcal{P}_{\mu_j})_{j\in\mathbb{N}}$.

Unconstrained Subproblems

$$\mathcal{P}_{\mu_j}: \quad \underset{x\in\mathbb{R}^n}{\text{minimize}} \ f(x) + \underbrace{g(x) - \boldsymbol{\mu_j} \sum_{\boldsymbol{i=1}}^{\boldsymbol{p}} \ln\left(-\boldsymbol{c_i}(\boldsymbol{x})\right)}_{\textit{smooth}}$$

Our algorithm comprises two interlocked loops.

- Given $\mu_j > 0$, $(x_{j,k})_k$ is produced via several forward-backward (proximal gradient) steps.
- Once $x_{j,k}$ is close enough to the solution of $\mathcal{P}_{\mu_j}$, the barrier parameter μ_j is updated.

Iteration Scheme

Forward-Backward Step

For j fixed,

$$x_{j,k+1} = \mathrm{prox}^{A_{j,k}}_{\gamma_{j,k} f}(x_{j,k} - \gamma_{j,k} A^{-1}_{j,k} \nabla \varphi_{\mu_j}(x_{j,k}))$$

where $\varphi_{\mu_j}(x) = g(x) - \mu_j \sum_{i=1}^{p} \ln(-c_i(x))$.

Iteration Scheme

Forward-Backward Step

For j fixed,

$$x_{j,k+1} = \text{prox}_{\gamma_{j,k} f}^{A_{j,k}} (x_{j,k} - \gamma_{j,k} A_{j,k}^{-1} \boldsymbol{\nabla \varphi_{\mu_j}(x_{j,k})})$$

where $\varphi_{\mu_j}(x) = g(x) - \mu_j \sum_{i=1}^{p} \ln(-c_i(x))$.

- Gradient step on the smooth term ;
-
-
-

Iteration Scheme

Forward-Backward Step

For j fixed,

$$x_{j,k+1} = \mathbf{prox}^{A_{j,k}}_{\gamma_{j,k} f}(x_{j,k} - \gamma_{j,k} A^{-1}_{j,k} \nabla \varphi_{\mu_j}(x_{j,k}))$$

where $\varphi_{\mu_j}(x) = g(x) - \mu_j \sum_{i=1}^{p} \ln(-c_i(x))$.

- Gradient step on the smooth term ;
- Proximal step on the non-differentiable function f ;
-
-

Iteration Scheme

Forward-Backward Step

For j fixed,

$$x_{j,k+1} = \mathrm{prox}_{\gamma_{j,k} f}^{\boldsymbol{A}_{j,k}} (x_{j,k} - \gamma_{j,k} \boldsymbol{A}_{j,k}^{-1} \nabla \varphi_{\mu_j}(x_{j,k}))$$

where $\varphi_{\mu_j}(x) = g(x) - \mu_j \sum_{i=1}^{p} \ln(-c_i(x))$.

- Gradient step on the smooth term ;
- Proximal step on the non-differentiable function f ;
- Preconditioner $A_{j,k}$ for acceleration **[Chouzenoux *et al.*, 2016]** ;
-

Iteration Scheme

Forward-Backward Step

For j fixed,

$$x_{j,k+1} = \mathrm{prox}^{A_{j,k}}_{\boldsymbol{\gamma_{j,k}} f}(x_{j,k} - \boldsymbol{\gamma_{j,k}} A^{-1}_{j,k} \nabla \varphi_{\mu_j}(x_{j,k}))$$

where $\varphi_{\mu_j}(x) = g(x) - \mu_j \sum_{i=1}^{p} \ln(-c_i(x))$.

- Gradient step on the smooth term ;
- Proximal step on the non-differentiable function f ;
- Preconditioner $A_{j,k}$ for acceleration **[Chouzenoux *et al.*, 2016]** ;
- Stepsize $\gamma_{j,k} > 0$ found using a backtracking strategy **[Salzo, 2017]** since φ_{μ_j} is not Lipschitz-differentiable.

Algorithm

Proximal Interior point Algorithm (PIPA)

Initialization

Let $\bar{\gamma} > 0$, $(\delta, \theta) \in]0,1[^2$, $\mu_0 > 0$;
Initialize $x_{0,0}$ such that $(\forall i \in \{1,\dots,p\})\ c_i(x_{0,0}) < 0$;

For $j = 0, 1, \dots$

- For $k = 0, 1, \dots$
  - **Choose $\boldsymbol{A_{j,k}}$** satisfying a boundedness condition ;
  - For $l = 0, 1, \dots$
    - $\tilde{x}^l_{j,k} = \operatorname{prox}^{A_{j,k}}_{\bar{\gamma}\theta^l f}(x_{j,k} - \bar{\gamma}\theta^l A^{-1}_{j,k} \nabla \varphi_{\mu_j}(x_{j,k}))$;
    - Stop if the **backtracking condition** is met ;
  - $x_{j,k+1} = \tilde{x}^l_{j,k}$;
  - $\gamma_{j,k} = \bar{\gamma}\theta^l$;
  - Stop if **precision conditions** are met ;
- $x_{j+1,0} = x_{j,k+1}$;
- **Update $\boldsymbol{\mu_j}$**;

Theoretical Results

Assumptions

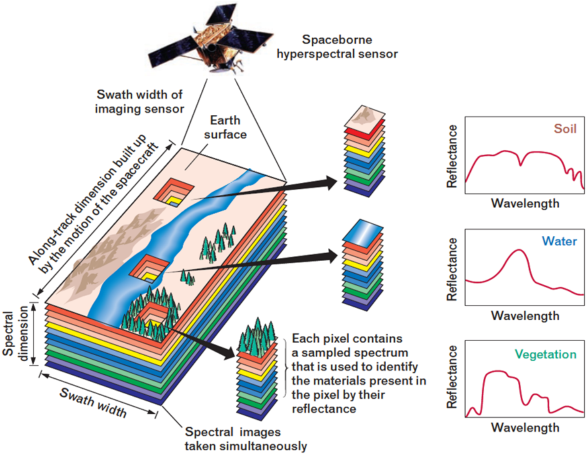
- The set of solutions to $\mathcal{P}_0$ is nonemtpy and bounded ;
- f, g and the constraints are convex, g is Lipschitz-differentiable and the constraints are continuously twice-differentiable ;
- The strict interior of the feasible set is nonempty ;
- $(\forall j \in \mathbb{N})$ $f + \varphi_{\mu_j}$ is a Kurdyka-Lojasiewicz (KL) function ;
- $(\forall j \in \mathbb{N})$ $(A_{j,k})_k$ are bounded from above and from below ;
- $\lim_{j\to\infty} \mu_j = 0$ and $(\forall i \in \{1,\ldots,4\})$ $\lim_{j\to\infty} \epsilon_{i,j}/\mu_j = 0$.

Convergence

Under some mild technical assumptions :

- for all $j \in \mathbb{N}$, $(x_{j,k})_{k\in\mathbb{N}}$ **converges to a solution to** $\mathcal{P}_{\mu_j}$;
- $(x_{j,0})_{j\in\mathbb{N}}$ **is bounded and every cluster point of it is a solution to** $\mathcal{P}_0$;
- if in addition strict complementarity holds, and if there exists $i \in \{1,\ldots,p\}$ such that c_i is strictly convex (or alternatively, for linear constraints, if some full rank property is satisfied) then $(x_{j,0})_{j\in\mathbb{N}}$ **converges to a solution to** $\mathcal{P}_0$.

Hyperspectral Unmixing Problem



Hyperspectral Unmixing Model

Optimization Problem

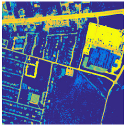
$$\begin{aligned}
\underset{X\in\mathbb{R}^{p\times n}}{\text{minimize}} \quad & \frac{1}{2}\|Y - SX\|_2^2 + \kappa \sum_{i=1}^{p} \|(WX_i)_d\|_1 \\
s.t. \quad & (\forall j \in \{1,\ldots,n\}) \ \sum_{i=1}^{p} X_{i,j} \leq 1 \\
& (\forall i \in \{1,\ldots,p\})(\forall j \in \{1,\ldots,n\}) \ X_{i,j}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

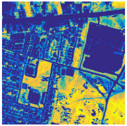
- p, n, s : number of endmembers, pixels, spectral bands
- $Y \in \mathbb{R}^{s\times n}$: observation
- $S \in \mathbb{R}^{s\times p}$: library
- $X \in \mathbb{R}^{p\times n}$: abundance matrix
- $W \in \mathbb{R}^{n\times n}$: orthogonal wavelet basis
- $\|(.)_d\|_1$: ℓ_1 norm of the detail wavelet coefficients

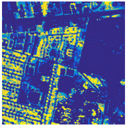
Experimental Setting

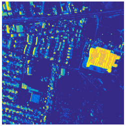
Realistic Data Simulation

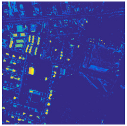
- Urban [1] data set : $p = 6$ endmembers (known spectral signatures), $s = 162$ spectral bands, $n = 256 \times 256$ pixels
- Gaussian noise : $\sigma^2 = 4.1 \times 10^{-3}$.

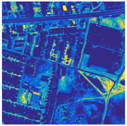

Asphalt Road


Grass


Tree


Roof


Metal

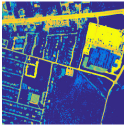

Dirt

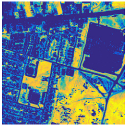
1. http ://www.escience.cn/people/feiyunZHU/Dataset_GT.html

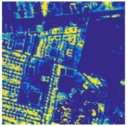
Experimental Setting

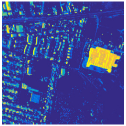
Realistic Data Simulation

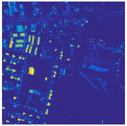
- Urban[1] data set : $p = 6$ endmembers (known spectral signatures), $s = 162$ spectral bands, $n = 256 \times 256$ pixels $\rightarrow$ ***large-scale pb : $> 3.9 \times 10^5$ variables***.
- Gaussian noise : $\sigma^2 = 4.1 \times 10^{-3}$.

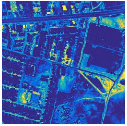

Asphalt Road


Grass


Tree


Roof


Metal


Dirt

1. http ://www.escience.cn/people/feiyunZHU/Dataset_GT.html

Experimental Setting

Realistic Data Simulation

- Urban [1] data set : $p = 6$ endmembers (known spectral signatures), $s = 162$ spectral bands, $n = 256 \times 256$ pixels $\rightarrow$ *large-scale pb :* $> 3.9 \times 10^5$ *variables.*
- Gaussian noise : $\sigma^2 = 4.1 \times 10^{-3}$.

Reconstruction Model

- Regularization weight : $\kappa = 10^{-2}$.
- W : orthogonal Daubechies 4 wavelet decomposition over 2 resolution levels.

1. http ://www.escience.cn/people/feiyunZHU/Dataset_GT.html

Experimental Setting

Realistic Data Simulation

- Urban [1] data set : $p = 6$ endmembers (known spectral signatures), $s = 162$ spectral bands, $n = 256 \times 256$ pixels $\rightarrow$ *large-scale pb : $> 3.9 \times 10^5$ variables.*
- Gaussian noise : $\sigma^2 = 4.1 \times 10^{-3}$.

Reconstruction Model

- Regularization weight : $\kappa = 10^{-2}$.
- W : orthogonal Daubechies 4 wavelet decomposition over 2 resolution levels.

Algorithm Parameters

- Variable metric : $A_{j,k} := \nabla^2 \varphi_{\mu_j}(x_{j,k})$ **[Becker *et al.*, 2012]**.
- The barrier parameter $(\mu_j)_{j \in \mathbb{N}}$ and the stopping criteria $\{\epsilon_{i,j}/\mu_j\}_{i \in \{1,\ldots,4\}}$ follow a geometric decrease.

1. http ://www.escience.cn/people/feiyunZHU/Dataset_GT.html

Experimental Setting

Realistic Data Simulation

- Urban [1] data set : $p = 6$ endmembers (known spectral signatures), $s = 162$ spectral bands, $n = 256 \times 256$ pixels $\rightarrow$ *large-scale pb :* $> 3.9 \times 10^5$ *variables.*
- Gaussian noise : $\sigma^2 = 4.1 \times 10^{-3}$.

Reconstruction Model

- Regularization weight : $\kappa = 10^{-2}$.
- W : orthogonal Daubechies 4 wavelet decomposition over 2 resolution levels.

Algorithm Parameters

- Variable metric : $A_{j,k} := \nabla^2 \varphi_{\mu_j}(x_{j,k})$ **[Becker *et al.*, 2012]**.
- The barrier parameter $(\mu_j)_{j \in \mathbb{N}}$ and the stopping criteria $\{\epsilon_{i,j}/\mu_j\}_{i \in \{1,\ldots,4\}}$ follow a geometric decrease.

Implementation

- Matlab R2016b, Intel Xeon 3.2 GHz processor and 16 GB of RAM.
- Code will be available soon on `https://github.com/mccorbineau`.

1. http ://www.escience.cn/people/feiyunZHU/Dataset_GT.html

Comparison

State-of-the-Art Algorithms

No reg : interior point least squares algorithm without regularization **[Chouzenoux *et al.*, 2014]**

ADMM : alternating direction of multipliers method **[Setzer *et al.*, 2010]**

PDS : primal-dual splitting algorithm **[Combettes *et al.*, 2014]**

GFBS : generalized forward-backward splitting algorithm **[Raguet *et al.*, 2013]**

Evaluation Metric

Signal-to-Noise Ratio

$$\mathrm{SNR} = -10\log_{10}\left(\sum_{i=1}^{p}\frac{\|X_i-\bar{X}_i\|_2^2}{\|\bar{X}_i\|_2^2}\right) \quad ; \quad \mathrm{SNR}_i = -10\log_{10}\left(\frac{\|X_i-\bar{X}_i\|_2^2}{\|\bar{X}_i\|_2^2}\right)$$

where $\bar{X}_i$ is the true abundance map of the i^{th} endmember.

Distance from the Iterates to the Solution

$$\frac{\|x_{j,k}-x_\infty\|_2}{\|x_\infty\|_2}$$

where x is the vectorization of X and x_∞ is obtained after a very large number of iterations.

Quantitative Results

- No reg : SNR = 1.96 dB / With regularization : SNR = 3.65 dB

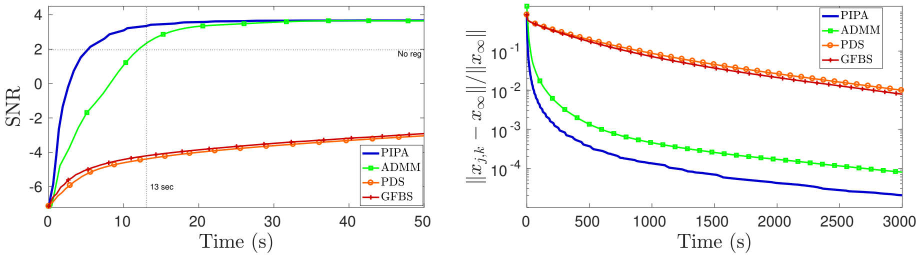


FIGURE – Left : global SNR versus time. Right : distance from the iterates to the solution versus time.

	Asphalt Road	Grass	Tree	Roof	Metal	Dirt
No reg	10.12	11.21	11.86	14.91	4.90	13.68
ADMM	6.75	11.47	12.56	14.66	**7.57**	11.47
PDS	2.06	3.33	4.73	6.63	-0.08	10.27
GFBS	2.17	3.57	4.76	7.66	0.05	10.31
PIPA	**10.98**	**11.70**	**12.73**	**15.19**	7.06	**14.57**

TABLE – SNR (dB) for all endmembers after 13 sec.

Asphalt Road

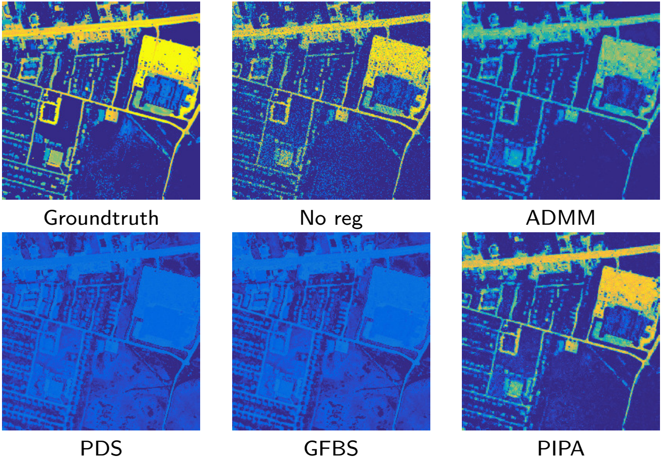


FIGURE – Abundance map of asphalt road after 13 sec.

No reg	10.12
ADMM	6.75
PDS	2.06
GFBS	2.17
PIPA	**10.98**

TABLE – SNR (dB)

Grass

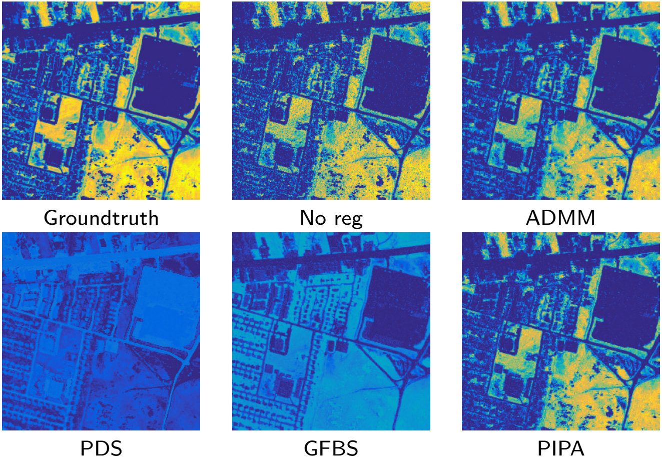


FIGURE – Abundance map of grass after 13 sec.

No reg	11.21
ADMM	11.47
PDS	3.33
GFBS	3.57
PIPA	**11.70**

TABLE – SNR (dB)

Dirt

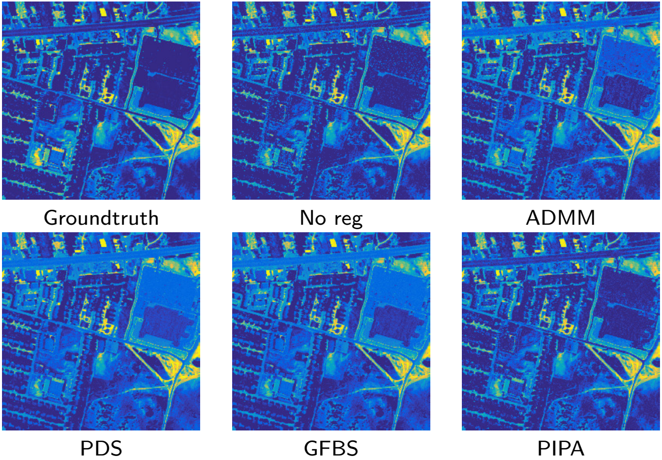


FIGURE – Abundance map of dirt after 13 sec.

No reg	13.68
ADMM	11.47
PDS	10.27
GFBS	10.31
PIPA	**14.57**

TABLE – SNR (dB)

Conclusion

Application of a new proximal interior point algorithm to hyperspectral unmixing with a non-differentiable regularization.

- Convergence guaranteed under mild assumptions.
- Possibility to include an arbitrary preconditioner.
- Good performance in the context of a large-scale image recovery application.

→ Extension of the convergence proof to inexact proximity operator.

→ Other applications.

References

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.

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.

E. Chouzenoux, J.-C. Pesquet and A. Repetti.
A block coordinate variable metric forward-backward algorithm
Journal of Global Optimization, Vol. 66, No. 3, pp 457–485, 2016.

S. Salzo.
The variable metric forward-backward splitting algorithm under mild differentiability assumptions
SIAM Journal on Optimization, Vol. 27, No. 4, pp 2153–2181, 2017.

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.

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.

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
Proceedings of the IEEE International Conference on Image Processing (ICIP 2014), pp 4141–4145, 2014.

H. Raguet, J. Fadili, G. and Peyré.
A generalized forward-backward splitting
SIAM Journal on Imaging Sciences, Vol. 6, No. 3, pp 1199–1226, 2013.

S. Becker and J. Fadili.
A forward-backward view of some primal-dual optimization methods in image recovery
Proceedings of the 25th Advances in Neural Information Processing Systems Conference (NIPS 2012), pp 2618–2626, 2012.

Thank you !

Stopping Criteria

Backtracking [Salzo, 2017]

For j and k fixed, the backtracking procedure stops if :

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

If $f := 0$, Armijo linesearch along the steepest direction.

Stopping Criteria

Backtracking **[Salzo, 2017]**

For j and k fixed, the backtracking procedure stops if :

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

If $f := 0$, Armijo linesearch along the steepest direction.

Accuracy for Solving $\mathcal{P}_{\mu_j}$

The barrier parameter is decreased as soon as the following criteria are met :

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

$$\sum_{i=1}^{p}\left|\frac{c_i(x_{j,k+1})}{c_i(x_{j,k})} - 1\right| \leq \epsilon_{3,j} \qquad \mu_j\left\|\sum_{i=1}^{p}\frac{\nabla c_i(x_{j,k}) - \nabla c_i(x_{j,k+1})}{c_i(x_{j,k})}\right\| \leq \epsilon_{4,j}$$

where $\{(\epsilon_{i,j})_{j\in\mathbb{N}}\}_{i\in\{1,\ldots,4\}}$ and $(\mu_j)_{j\in\mathbb{N}}$ are strictly positive sequences converging to 0 such that $(\forall i \in \{1,\ldots,4\})\ \lim_{j\to\infty} \epsilon_{i,j}/\mu_j = 0$.

Tree

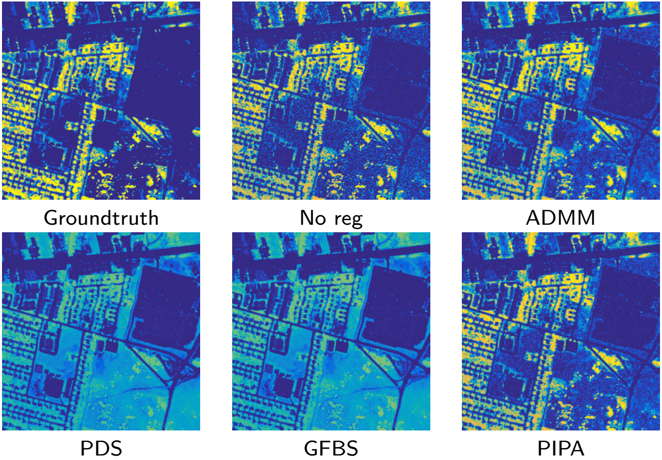
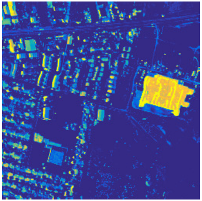


FIGURE – Abundance map of tree after 13 sec.

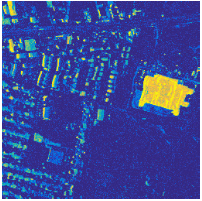
No reg	11.86
ADMM	12.56
PDS	4.73
GFBS	4.76
PIPA	**12.73**

TABLE – SNR (dB)

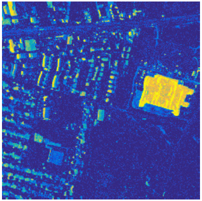
Roof



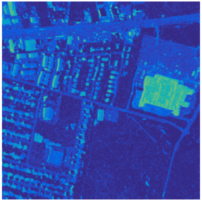
Groundtruth



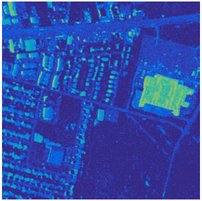
No reg



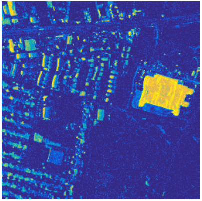
ADMM



PDS



GFBS



PIPA

FIGURE – Abundance map of roof after 13 sec.

No reg	14.91
ADMM	14.66
PDS	6.63
GFBS	7.66
PIPA	**15.19**

TABLE – SNR (dB)

Metal

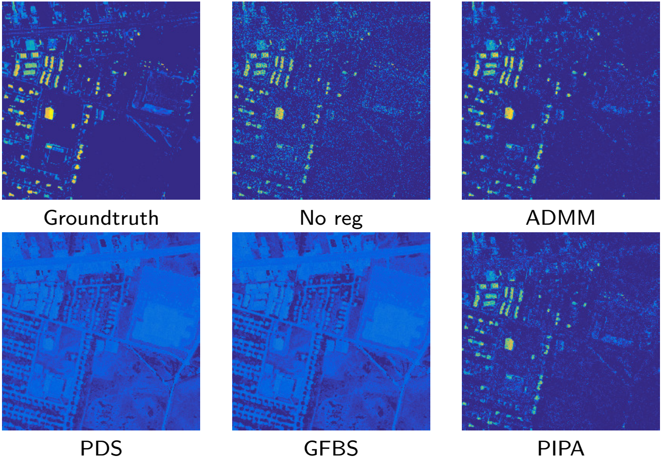


FIGURE – Abundance map of metal after 13 sec.

No reg	4.90
ADMM	**7.57**
PDS	-0.08
GFBS	0.05
PIPA	7.06

TABLE – SNR (dB)