Proximal methods for constrained cosparse modelling

N. Pustelnik
in collaboration with G. Cherchia, J.-C. Pesquet, and B. Pesquet-Popescu

Journée SMAI-SIGMA
18 novembre 2011

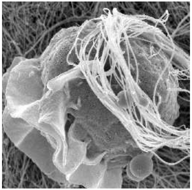
Outline

1. Problem : image recovery.
2. Regularized approach versus Constrained approach.
3. Proposed solution to the general constrained minimization problem.
4. Experimental results.
5. Conclusions.

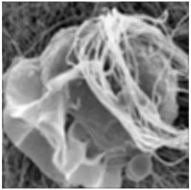
Degradation model

$$\boxed{z = \mathcal{P}_\alpha(T\overline{x})}$$

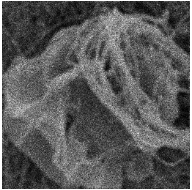
- $\overline{x}$ **: original image** in the Hilbert space $\mathcal{H}$ which is assumed to be sparse after some appropriate transform,
- T : a linear operator from $\mathcal{H}$ to $\mathbb{R}^K$,
- $\mathcal{P}_\alpha$: effect of noise where α is the **scaling parameter,**
- z **: degraded image** of size K.



Original ($\overline{x}$)



Convolved ($T\overline{x}$)

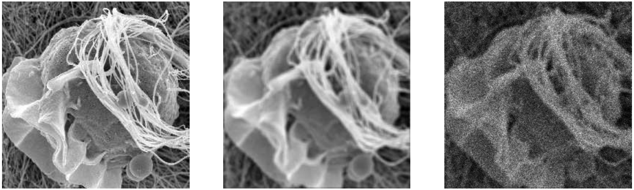


Degraded (z)

Degradation model

$$\boxed{z = \mathcal{P}_\alpha(T\overline{x})}$$

- $\overline{x}$ **: original image** in the Hilbert space $\mathcal{H}$ which is assumed to be sparse after some appropriate transform,
- T : a linear operator from $\mathcal{H}$ to $\mathbb{R}^K$,
- $\mathcal{P}_\alpha$: effect of noise where α is the **scaling parameter,**
- z **: degraded image** of size K.



Original ($\overline{x}$) Convolved ($T\overline{x}$) Degraded (z)

- **Question** : How can we recover $\overline{x}$ from the observations z.

Existing works : Gaussian noise

Regularized approach	**Constrained approach**
$$\min_{x\in\mathcal{H}} \|Tx - z\|^2 + \lambda f(x)$$	$$\min_{\|Tx-z\|^2\leq\eta} f(x)$$
[Tikhonov, 1963]	[Combettes, Trussell, 1991]

Existing works : Gaussian noise

Regularized approach	**Constrained approach**
$$\min_{x \in \mathcal{H}} \|Tx - z\|^2 + \lambda f(x)$$	$$\min_{\|Tx-z\|^2 \leq \eta} f(x)$$
[Tikhonov, 1963]	[Combettes, Trussell, 1991]

If $f = \|F \cdot\|^2$

→ Gradient-based methods	→ POCS [Trussell, Civanlar, 1984] → Subgradient projections [Luo, Combettes, 1999]

Existing works : Gaussian noise

Regularized approach	**Constrained approach**
$\min_{x \in \mathcal{H}} \|Tx - z\|^2 + \lambda f(x)$	$\min_{\|Tx-z\|^2 \leq \eta} f(x)$
[Tikhonov, 1963]	[Combettes, Trussell, 1991]

If $f = \|F \cdot\|^2$

→ Gradient-based methods

→ POCS [Trussell, Civanlar, 1984]

→ Subgradient projections [Luo, Combettes, 1999]

If $f(x) = \sum_i |(Fx)^{(i)}|_1$

(F: a wavelet transform, an analysis frame) [Elad et al,2007][Nam et al.,2011]

→ Proximal methods [Combettes, Pesquet, 2011]

→ Proximal methods [Combettes, Pesquet, 2011]

Existing works : Poisson noise

Regularized approach	**Constrained approach**
$$\min_{x \in \mathcal{H}} D_{KL}(Tx, z) + \lambda f(x)$$	$$\min_{D_{KL}(Tx,z) \leq \eta} f(x)$$

Existing works : Poisson noise

Regularized approach	**Constrained approach**
$\min_{x \in \mathcal{H}} D_{KL}(Tx, z) + \lambda f(x)$	$\min_{D_{KL}(Tx,z) \leq \eta} f(x)$

If $f = \|F \cdot\|^2$

Regularized approach:

→ Cross-Entropy minimization

[Byrne, 1993]

→ Barrier function optimization

[Chouzenoux *et al.*, 2011]

Constrained approach:

→ **?**

Existing works : Poisson noise

Regularized approach	**Constrained approach**
$\min_{x\in\mathcal{H}} D_{KL}(Tx,z)+\lambda f(x)$	$\min_{D_{KL}(Tx,z)\le\eta} f(x)$

If $f=\|F\cdot\|^2$

→ Cross-Entropy minimization [Byrne, 1993] → ?

→ Barrier function optimization [Chouzenoux *et al.*, 2011]

If $f(x)=\sum_i |(Fx)^{(i)}|_1$

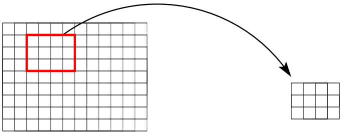
(where F can denote a gradient filter, a wavelet transform, a frame)

→ Proximal methods [Combettes, Pesquet, 2011] → ?

Considered problem

$$\underset{x\in\mathcal{H}}{\text{minimize}}\ \|Fx\|_{2,1} \quad \text{subject to} \quad \begin{cases} x \in C \\ g(Tx,z) \leq \eta. \end{cases}$$

- $C \subset \mathcal{H}$: nonempty closed convex set, models the data range dynamic,
- $g(\cdot,z) \in \Gamma_0(\mathbb{R}^K)$ – the class of convex, l.s.c, and proper functions,
- $\eta \in \mathbb{R}$,
- F : bounded linear operator from $\mathcal{H}$ to $\ell^2(\mathbb{K})$,
- $\|\cdot\|_{2,1} = \sum_{b\in\mathbb{L}} \|B_b \cdot\|$: a block sparsity measure,
- for every $b \in \mathbb{L} \subset \mathbb{K}$, B_b is some **block selection transform**.
  A linear transform B from $\ell^2(\mathbb{K})$ to $\mathbb{R}^L$ will be said to be a block selection transform if it allows us to select a block of L data from its input vector.

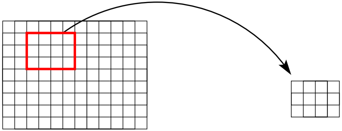


Considered problem

$$\underset{x \in \mathcal{H}}{\text{minimize}} \sum_{b \in \mathbb{L}} \|B_b F x\| \quad \text{subject to} \quad \begin{cases} x \in C \\ g(Tx, z) \leq \eta. \end{cases}$$

- $C \subset \mathcal{H}$: nonempty closed convex set, models the data range dynamic,
- $g(\cdot, z) \in \Gamma_0(\mathbb{R}^K)$ – the class of convex, l.s.c, and proper functions,
- $\eta \in \mathbb{R}$,
- F : bounded linear operator from $\mathcal{H}$ to $\ell^2(\mathbb{K})$,
- $\| \cdot \|_{2,1} = \sum_{b \in \mathbb{L}} \|B_b \cdot \|$: a block sparsity measure,
- for every $b \in \mathbb{L} \subset \mathbb{K}$, B_b is some **block selection transform**.

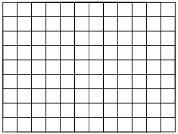
  A linear transform B from $\ell^2(\mathbb{K})$ to $\mathbb{R}^L$ will be said to be a block selection transform if it allows us to select a block of L data from its input vector.



Considered problem

$$\underset{x\in\mathcal{H}}{\text{minimize}} \sum_{b\in\mathbb{L}} \|B_b F x\| \quad \text{subject to} \quad \begin{cases} x \in C \\ g(Tx, z) \leq \eta. \end{cases}$$

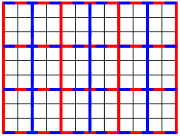
- For computational reasons, it will be assumed that there exists a partition of $\mathbb{L}$ in S subsets $(\mathbb{L}_s)_{1\leq s\leq S}$ of non-overlapping blocks : $\sum_{b\in\mathbb{L}} \|B_b \cdot \| = \sum_{s=1}^{S} \sum_{b\in\mathbb{L}_s} \|B_b \cdot \|$.



Considered problem

$$\underset{x\in\mathcal{H}}{\text{minimize}} \sum_{b\in\mathbb{L}} \|B_b F x\| \quad \text{subject to} \quad \begin{cases} x \in C \\ g(Tx, z) \leq \eta. \end{cases}$$

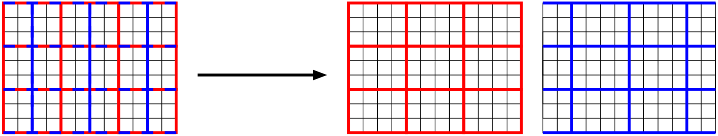
- For computational reasons, it will be assumed that there exists a partition of $\mathbb{L}$ in S subsets $(\mathbb{L}_s)_{1\leq s\leq S}$ of non-overlapping blocks : $\sum_{b\in\mathbb{L}} \|B_b \cdot\| = \sum_{s=1}^{S} \sum_{b\in\mathbb{L}_s} \|B_b \cdot\|$.



Considered problem

$$\underset{x\in\mathcal{H}}{\text{minimize}} \sum_{b\in\mathbb{L}} \|B_b F x\| \quad \text{subject to} \quad \begin{cases} x \in C \\ g(Tx, z) \leq \eta. \end{cases}$$

- For computational reasons, it will be assumed that there exists a partition of $\mathbb{L}$ in S subsets $(\mathbb{L}_s)_{1\leq s\leq S}$ of non-overlapping blocks : $\sum_{b\in\mathbb{L}} \|B_b \cdot\| = \sum_{s=1}^{S} \sum_{b\in\mathbb{L}_s} \|B_b \cdot\|$.

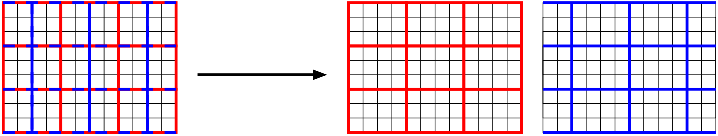


$(S = 2)$

Considered problem

$$\underset{x\in\mathcal{H}}{\text{minimize}} \sum_{b\in\mathbb{L}} \|B_b F x\| \quad \text{subject to} \quad \begin{cases} x \in C \\ g(Tx, z) \leq \eta. \end{cases}$$

- For computational reasons, it will be assumed that there exists a partition of $\mathbb{L}$ in S subsets $(\mathbb{L}_s)_{1\leq s\leq S}$ of non-overlapping blocks : $\sum_{b\in\mathbb{L}} \|B_b \cdot\| = \sum_{s=1}^{S} \sum_{b\in\mathbb{L}_s} \|B_b \cdot\|$.

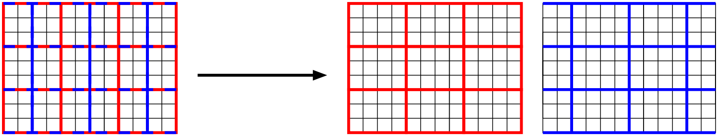


$(S = 2)$

Considered problem

$$\underset{x\in\mathcal{H}}{\text{minimize}} \sum_{s=1}^{S} \sum_{b\in\mathbb{L}_s} \|B_b F x\| \quad \text{subject to} \quad \begin{cases} x \in C \\ g(Tx, z) \leq \eta. \end{cases}$$

- For computational reasons, it will be assumed that there exists a partition of $\mathbb{L}$ in S subsets $(\mathbb{L}_s)_{1\leq s\leq S}$ of non-overlapping blocks : $\sum_{b\in\mathbb{L}} \|B_b \cdot\| = \sum_{s=1}^{S} \sum_{b\in\mathbb{L}_s} \|B_b \cdot\|$.

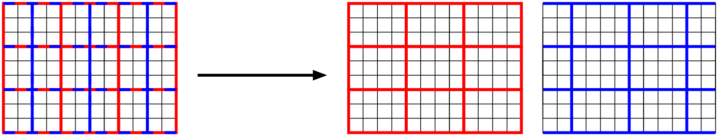


$(S = 2)$

Considered problem

$$\underset{x\in\mathcal{H}}{\text{minimize}} \sum_{s=1}^{S} \sum_{b\in\mathbb{L}_s} \|B_b F x\| \quad \text{subject to} \quad \begin{cases} x \in C \\ g(Tx, z) \le \eta. \end{cases}$$

- For computational reasons, it will be assumed that there exists a partition of $\mathbb{L}$ in S subsets $(\mathbb{L}_s)_{1\le s\le S}$ of non-overlapping blocks : $\sum_{b\in\mathbb{L}} \|B_b \cdot\| = \sum_{s=1}^{S} \sum_{b\in\mathbb{L}_s} \|B_b \cdot\|$.



$(S = 2)$

- *<u>Particular case</u> : $S = 1$, $\mathbb{L} = \mathbb{L}_1 = \mathbb{K}$ and, for every $b \in \mathbb{L}$, B_b selects one element (i.e. one pixel)* $\rightarrow$ **the classical ℓ^1-norm is obtained**.

Considered problem

$$\underset{x \in \mathcal{H}}{\text{minimize}} \sum_{s=1}^{S} f_s(Fx) \quad \text{subject to} \quad \begin{cases} x \in C \\ g(Tx, z) \leq \eta. \end{cases}$$

- $f_s = \sum_{b \in \mathbb{L}_s} \|B_b \cdot\|$,
- for every $b \in \mathbb{L}_s$, $B_b : \ell^2(\mathbb{K}) \to \mathbb{R}^{L_b}$ is a block selection operator,
- $(\mathbb{L}_s)_{1 \leq s \leq S}$ is a partition of $\mathbb{L}$.

Considered problem

$$\underset{x\in\mathcal{H}}{\text{minimize}} \sum_{s=1}^{S} f_s(Fx) \quad \text{subject to} \quad \begin{cases} x \in C \\ g(Tx, z) \le \eta. \end{cases}$$

- $f_s = \sum_{b\in\mathbb{L}_s} \|B_b \cdot\|$,
- for every $b \in \mathbb{L}_s$, $B_b : \ell^2(\mathbb{K}) \to \mathbb{R}^{L_b}$ is a block selection operator,
- $(\mathbb{L}_s)_{1\le s\le S}$ is a partition of $\mathbb{L}$.

The criterion can be rewritten :

$$\underset{x\in\mathcal{H}}{\text{minimize}} \sum_{s=1}^{S} f_s(Fx) + \iota_C(x) + \iota_D(Tx)$$

- ι_C : indicator function (is equal to 0 on C and $+\infty$ on $\mathcal{H} \setminus C$),
- $D = \{u \in \mathbb{R}^K \mid g(u,z) \le \eta\} = \operatorname{lev}_{\le\eta} g(\cdot, z)$.

Algorithms to minimize $\sum_{s=1}^{S} f_s(F\cdot) + \iota_C(\cdot) + \iota_D(T\cdot)$

Proximal algorithms :

- To solve $\min_{x\in\mathcal{H}} \sum_{i=1}^{n} f_i(L_i x)$ where L_i denotes a bounded linear operator and f_i denotes a convex, l.s.c., and proper function.
- Based on proximal tools : $\mathrm{prox}_f x = \underset{p\in\mathcal{H}}{\mathrm{argmin}} \frac{1}{2}\|p - x\|^2 + f(p)$.

Proximity operator :

- Generalization of projection onto a closed convex set : $\mathrm{prox}_{\iota_C} = P_C$.
- Numerous closed form (ℓ_p-norm, gamma,...) [Chaux et al., 2007].

Existing algorithms :

- Primal : FB [Combettes, Wajs, 2005], DR [Combettes, Pesquet, 2007], PPXA+ [Pesquet, Pustelnik, 2011].
- Primal-dual : M+SFBF [Briceño-Arias, Combettes, 2011], M+LFBF [Combettes, Pesquet, 2011], Generalized FB [Raguet et al., 2011], [Condat, 2011], [Vu, 2011].

Primal algorithm : PPXA+ [Pesquet, Pustelnik, 2011]

Initialization

$(\epsilon_i)_{1\leq i\leq n} \in [0,1[^n, (\omega_i)_{1\leq i\leq n} \in \,]0,+\infty[^n,$

$(\lambda_\ell)_{\ell\in\mathbb{N}}$ be a sequence of reals,

$(v_{i,0})_{1\leq i\leq n} \in (\mathcal{H})^n, (p_{i,-1})_{1\leq i\leq n} \in (\mathcal{H})^n,$

$u_0 = \arg\min_{u\in\mathcal{H}} \sum_{i=1}^n \omega_i \|L_i u - v_{i,0}\|^2$

For every $i \in \{1,\ldots,n\}$, $(a_{i,\ell})_{\ell\in\mathbb{N}}$ be a sequence of reals,

For $\ell = 0,1,\ldots$

For $i = 1,\ldots,n$

$p_{i,\ell} = \mathrm{prox}_{\frac{(1-\epsilon_i)f_i}{\omega_i}}\big((1-\epsilon_i)v_{i,\ell} + \epsilon_i p_{i,\ell-1}\big) + a_{i,\ell}$

$c_\ell = \arg\min_{u\in\mathcal{H}} \sum_{i=1}^n \omega_i \|L_i u - p_{i,\ell}\|^2$

For $i = 1,\ldots,n$

$v_{i,\ell+1} = v_{i,\ell} + \lambda_\ell\big(L_i(2c_\ell - u_\ell) - p_{i,\ell}\big)$

$u_{\ell+1} = u_\ell + \lambda_\ell(c_\ell - u_\ell)$

Primal algorithm : PPXA+ [Pesquet, Pustelnik, 2011]

The weak convergence of the sequence $(u_\ell)_{\ell\in\mathbb{N}}$ to a minimizer of $\sum_{i=1}^n f_i \circ L_i$ is established under the following assumptions :

1. $\mathbf{0} \in \operatorname{sri}\{(L_1 v - w, \dots, L_n v - w) \mid v \in \mathcal{H}, x_1 \in \operatorname{dom} f_1, \dots, x_n \in \operatorname{dom} f_n\}$,
2. There exists $\underline{\lambda} \in]0,2[$ such that $(\forall \ell \in \mathbb{N})$, $\underline{\lambda} \leq \lambda_{\ell+1} \leq \lambda_\ell$,
3. For every $i \in \{1,\dots,n\}$, $a_{i,\ell}$ are absolutely summable sequences in $\mathcal{H}$.
4. $\sum_{i=1}^n \omega_i L_i^* L_i$ is an isomorphism. (PPXA+ iterations can be slightly modified to avoid this assumption)

Primal-Dual algorithm : M+SFBF [Briceño-Arias, Combettes, 2011]

Initialization

For every $i \in \{1,\dots,n\}$, $\omega_i \in]0,1]$ such that $\sum_{i=1}^m \omega_i = 1$
For every $i \in \{1,\dots,n\}$, $v_{i,0} \in \mathcal{G}_i$ and $u_{i,0} \in \mathcal{H}$,
$\beta = \max_{1\leq i\leq m} \|L_i\|$, let $\epsilon \in]0, 1/(\beta+1)[$, let $(\gamma_\ell)_{\ell\leq 0}$ in $[\epsilon,(1-\epsilon)/\beta]$.

For $\ell = 0, 1, \dots$

$u_\ell = \sum_{i=1}^n \omega_i u_{i,\ell}$
For $i = 1,\dots,n$
$y_{1,i,\ell} = u_{1,i,\ell} - \gamma_\ell(L_i^* v_{i,\ell} + a_{1,i,\ell})$
$y_{2,i,\ell} = v_{1,i,\ell} + \gamma_\ell(L_i u_{i,\ell} + a_{2,i,\ell})$
$p_{1,\ell} = \sum_{i=1}^n \omega_i y_{1,i,\ell}$
For $i = 1,\dots,n$
$p_{2,i,\ell} = \mathrm{prox}_{\gamma_\ell f_i^*} y_{2,i,\ell} + b_{i,\ell}$
$q_{1,i,\ell} = p_{1,\ell} - \gamma_\ell(L_i^* p_{2,i,\ell} + c_{1,i,\ell})$
$q_{2,i,\ell} = p_{2,\ell} + \gamma_\ell(L_i p_{1,\ell} + c_{2,i,\ell})$
$u_{i,\ell+1} = u_{i,\ell} - y_{1,i,\ell} + q_{1,i,\ell}$
$v_{i,\ell+1} = v_{i,\ell} - y_{2,i,\ell} + q_{2,i,\ell}$

Primal-Dual algorithm : M+SFBF [Briceño-Arias, Combettes, 2011]

The weak convergence of the sequence $(u_\ell)_{\ell\in\mathbb{N}}$ to a minimizer of $\sum_{i=1}^n \omega_i f_i \circ L_i$ is established under the following assumptions :

1. $\mathbf{0} \in \mathrm{ran}\, \sum_{i=1}^n \omega_i L_i^* \circ (\partial f_i) \circ L_i$,
2. For every $i \in \{1,\ldots,n\}$, $(a_{1,i,\ell})_{\ell\in\mathbb{N}}$ and $(c_{1,i,\ell})_{\ell\in\mathbb{N}}$ are absolutely summable sequences in $\mathcal{H}$,
3. For every $i \in \{1,\ldots,n\}$, $(a_{2,i,\ell})_{\ell\in\mathbb{N}}$, $(b_{i,\ell})_{\ell\in\mathbb{N}}$, and $(c_{2,i,\ell})_{\ell\in\mathbb{N}}$ are absolutely summable sequences in $\mathcal{G}_i$.

Proximity operators to compute

$$\underset{x \in \mathcal{H}}{\text{minimize}} \sum_{s=1}^{S} f_s(Fx) + \iota_C(x) + \iota_D(Tx)$$

Computation of the proximity operators :

- $f_s = \sum_{b \in \mathbb{L}_s} \|B_b \cdot \|$,
- ι_C with $C \subset \mathcal{H}$: nonempty closed convex set, models data dynamic,
- ι_D with $D = \left\{ u \in \mathbb{R}^K \;\middle|\; g(u,z) \leq \eta \right\}$

Proximity operators to compute

$$\underset{x \in \mathcal{H}}{\text{minimize}} \sum_{s=1}^{S} f_s(Fx) + \iota_C(x) + \iota_D(Tx)$$

Computation of the proximity operators :

- $f_s = \sum_{b \in \mathbb{L}_s} \|B_b \cdot \|$,
  $\rightarrow$ Closed form [Peyré, Fadili, 2011].

- ι_C with $C \subset \mathcal{H}$: nonempty closed convex set, models data dynamic,
  $\rightarrow$ Closed form : projection onto a hypercube [Rockafellar, 1969].

- ι_D with $D = \left\{ u \in \mathbb{R}^K \;\middle|\; g(u, z) \leq \eta \right\}$
  $\rightarrow$ Closed form if $g(\cdot, z) = \| \cdot - z\|^2$ [Rockafellar, 1969].

Proximity operators to compute

$$\underset{x\in\mathcal{H}}{\text{minimize}} \sum_{s=1}^{S} f_s(Fx) + \iota_C(x) + \iota_D(Tx)$$

Computation of the proximity operators :

- $f_s = \sum_{b\in\mathbb{L}_s} \|B_b \cdot\|$,
  $\longrightarrow$ Closed form [Peyré, Fadili, 2011].

- ι_C with $C \subset \mathcal{H}$: nonempty closed convex set, models data dynamic,
  $\longrightarrow$ Closed form : projection onto a hypercube [Rockafellar, 1969].

- ι_D with $D = \left\{u \in \mathbb{R}^K \;\middle|\; g(u,z) \leq \eta\right\}$
  $\longrightarrow$ Closed form if $g(\cdot,z) = \|\cdot - z\|^2$ [Rockafellar, 1969].
  $\longrightarrow$ NO closed form in a general context.

Epigraphical projection

How to handle a convex constraint $\widetilde{D}$ of the form

$$\widetilde{D} = \{ v \in \mathbb{R}^{KM} \mid h(v) \leq \eta \} \quad ?$$

where

- the generic vector v has been decomposed into K blocks of coordinates as follows

$$v^\top = [\underbrace{(\mathrm{v}^{(1)})^\top}_{\text{size } M}, \dots, \underbrace{(\mathrm{v}^{(K)})^\top}_{\text{size } M}],$$

- $(\forall v \in \mathbb{R}^{KM})$, $h(v) = \sum_{r=1}^{K} h_r(\mathrm{v}^{(r)})$,
- For every r, h_r is a function in $\Gamma_0(\mathbb{R})$.

Epigraphical projection

How to handle a convex constraint $\widetilde{D}$ of the form

$$\widetilde{D} = \left\{v \in \mathbb{R}^{KM} \mid h(v) \leq \eta\right\} = \left\{v \in \mathbb{R}^{KM} \mid \sum_{r=1}^{K} h_r(\mathrm{v}^{(r)}) \leq \eta\right\} \quad ?$$

where

- the generic vector v has been decomposed into K blocks of coordinates as follows

$$v^\top = [\underbrace{(\mathrm{v}^{(1)})^\top}_{\text{size } M}, \dots, \underbrace{(\mathrm{v}^{(K)})^\top}_{\text{size } M}],$$

- $(\forall v \in \mathbb{R}^{KM})$, $h(v) = \sum_{r=1}^{K} h_r(\mathrm{v}^{(r)})$,
- For every r, h_r is a function in $\Gamma_0(\mathbb{R})$.

Epigraphical projection

How to handle a convex constraint $\widetilde{D}$ of the form

$$\widetilde{D} = \Big\{ v = (\mathrm{v}^{(r)})_{1 \le r \le K} \in \mathbb{R}^{KM} \mid \sum_{r=1}^{K} h_r(\mathrm{v}^{(r)}) \le \eta \Big\} \quad ?$$

Solution : Define an auxiliary vector $\zeta = \big(\zeta^{(r)}\big)_{1 \le r \le K} \in \mathbb{R}^K$.

$\Rightarrow$ the inequality in $\widetilde{D}$ can be equivalently rewritten as

$$\begin{cases} \sum_{r=1}^{K} \zeta^{(r)} \le \eta \\ (\forall r \in \{1, \dots, K\}), \;\; h_r(\mathrm{v}^{(r)}) \le \zeta^{(r)}. \end{cases}$$

Epigraphical projection

How to handle a convex constraint $\widetilde{D}$ of the form

$$\widetilde{D} = \left\{ v = (\mathrm{v}^{(r)})_{1 \le r \le K} \in \mathbb{R}^{KM} \mid \sum_{r=1}^{K} h_r(\mathrm{v}^{(r)}) \le \eta \right\} \quad ?$$

Solution : Define an auxiliary vector $\zeta = \left(\zeta^{(r)}\right)_{1 \le r \le K} \in \mathbb{R}^K$.

$\Rightarrow$ the inequality in $\widetilde{D}$ can be equivalently rewritten as

$$\begin{cases} \sum_{r=1}^{K} \zeta^{(r)} \le \eta \\ (\forall r \in \{1, \dots, K\}), \;\; h_r(\mathrm{v}^{(r)}) \le \zeta^{(r)}. \end{cases} \quad \Leftrightarrow \begin{cases} \zeta \in V \\ (v, \zeta) \in E \end{cases}$$

where

$$\begin{cases} V = \left\{ \zeta \in \mathbb{R}^K \mid 1_K^\top \zeta \le \eta \right\} \\ E = \{ (v, \zeta) \in \mathbb{R}^{KM} \times \mathbb{R}^K \mid (\forall r \in \{1, \dots, K\}) \; (\mathrm{v}^{(r)}, \zeta^{(r)}) \in \operatorname{epi} h_r \}. \end{cases}$$

Epigraphical projection

- $$\boxed{V = \{\zeta \in \mathbb{R}^K \mid 1_K^\top \zeta \leq \eta\}},$$

- $$\boxed{E = \{(v, \zeta) \in \mathbb{R}^{KM} \times \mathbb{R}^K \mid (\forall r \in \{1, \ldots, K\})\ (\mathrm{v}^{(r)}, \zeta^{(r)}) \in \operatorname{epi} h_r\}}$$

Epigraphical projection

- $\boxed{V = \{\zeta \in \mathbb{R}^K \mid 1_K^\top \zeta \leq \eta\}}$,
  → The projection operator is simply given by

$$(\forall \zeta \in \mathbb{R}^K) \qquad P_V(\zeta) = \begin{cases} \zeta & \text{if } 1_K^\top \zeta \leq \eta \\ \zeta + \frac{\eta - 1_K^\top \zeta}{K} 1_K & \text{otherwise,} \end{cases}$$

- $\boxed{E = \{(v, \zeta) \in \mathbb{R}^{KM} \times \mathbb{R}^K \mid (\forall r \in \{1, \ldots, K\})\ (\mathrm{v}^{(r)}, \zeta^{(r)}) \in \operatorname{epi} h_r\}}$

Epigraphical projection

- $\boxed{V = \{\zeta \in \mathbb{R}^K \mid 1_K^\top \zeta \leq \eta\}}$,
  → The projection operator is simply given by

$$(\forall \zeta \in \mathbb{R}^K) \qquad P_V(\zeta) = \begin{cases} \zeta & \text{if } 1_K^\top \zeta \leq \eta \\ \zeta + \frac{\eta - 1_K^\top \zeta}{K} 1_K & \text{otherwise,} \end{cases}$$

- $\boxed{E = \{(v, \zeta) \in \mathbb{R}^{KM} \times \mathbb{R}^K \mid (\forall r \in \{1, \ldots, K\}) \, (\mathrm{v}^{(r)}, \zeta^{(r)}) \in \operatorname{epi} h_r\}}$
  → The projection onto E [Bauschke, Combettes, 2011] is given by

$$(\forall (v, \zeta) \in \mathbb{R}^{KM} \times \mathbb{R}^K) \qquad P_E(v, \zeta) = (p, \theta)$$

$$\text{where } \begin{cases} \theta = (\theta^{(1)}, \ldots \theta^{(K)})^\top \quad \text{and} \quad p^\top = ((\mathrm{p}^{(1)})^\top, \ldots, (\mathrm{p}^{(K)})^\top), \\ (\forall r \in \{1, \ldots, K\}) \quad (\mathrm{p}^{(r)}, \theta^{(r)}) = P_{\operatorname{epi} h_r}(\mathrm{v}^{(r)}, \zeta^{(r)}). \end{cases}$$

Epigraphical projection

- $\boxed{V = \{\zeta \in \mathbb{R}^K \mid 1_K^\top \zeta \leq \eta\}}$,

  $\rightarrow$ The projection operator is simply given by

$$(\forall \zeta \in \mathbb{R}^K) \qquad P_V(\zeta) = \begin{cases} \zeta & \text{if } 1_K^\top \zeta \leq \eta \\ \zeta + \frac{\eta - 1_K^\top \zeta}{K} 1_K & \text{otherwise,} \end{cases}$$

- $\boxed{E = \{(v, \zeta) \in \mathbb{R}^{KM} \times \mathbb{R}^K \mid (\forall r \in \{1, \dots, K\})\ (\mathrm{v}^{(r)}, \zeta^{(r)}) \in \operatorname{epi} h_r\}}$

  $\rightarrow$ The projection onto E [Bauschke, Combettes, 2011] is given by

$$(\forall (v, \zeta) \in \mathbb{R}^{KM} \times \mathbb{R}^K) \qquad P_E(v, \zeta) = (p, \theta)$$

where $\begin{cases} \theta = (\theta^{(1)}, \dots \theta^{(K)})^\top \quad \text{and} \quad p^\top = ((\mathrm{p}^{(1)})^\top, \dots, (\mathrm{p}^{(K)})^\top), \\ (\forall r \in \{1, \dots, K\}) \quad (\mathrm{p}^{(r)}, \theta^{(r)}) = P_{\operatorname{epi} h_r}(\mathrm{v}^{(r)}, \zeta^{(r)}). \end{cases}$

$\Rightarrow$ Lower-dimensional problem of the determination of the projection onto the convex subset epi h_r for each $r \in \{1, \dots, K\}$. These projections have a closed form expression in a number of cases.

Epigraphical projection with a closed form

Explicit form of the projection operator associated with :

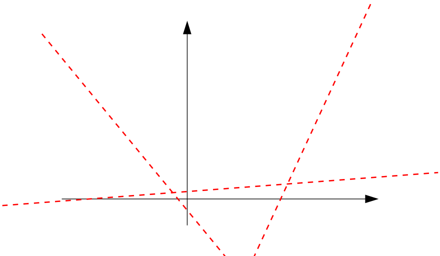
$$\boxed{h_r(\mathrm{v}^{(r)}) = \max\{\mathrm{v}^{(r,j)} + \eta^{(r,j)} \mid 1 \leq j \leq M\}}$$

where

$\rightarrow$ $\mathrm{v}^{(r)} = (\mathrm{v}^{(r,1)}, \ldots, \mathrm{v}^{(r,M)})^\top \in \mathbb{R}^M$

$\rightarrow$ $r \in \{1, \ldots, R\}$ and $(\eta^{(r,1)}, \ldots, \eta^{(r,M)})^\top \in \mathbb{R}^M$

Example for $R = 1$ and $M = 3$:



Epigraphical projection with a closed form

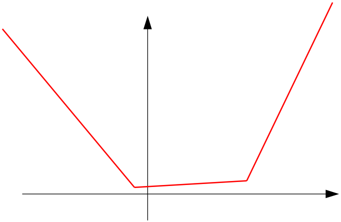
Explicit form of the projection operator associated with :

$$h_r(\mathrm{v}^{(r)}) = \max\{\mathrm{v}^{(r,j)} + \eta^{(r,j)} \mid 1 \leq j \leq M\}$$

where

- $\rightarrow$ $\mathrm{v}^{(r)} = (\mathrm{v}^{(r,1)}, \dots, \mathrm{v}^{(r,M)})^\top \in \mathbb{R}^M$
- $\rightarrow$ $r \in \{1, \dots, R\}$ and $(\eta^{(r,1)}, \dots, \eta^{(r,M)})^\top \in \mathbb{R}^M$

Example for $R = 1$ and $M = 3$:



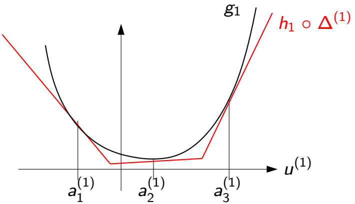
Algorithmic solution

$$g(u,z) = \sum_{r=1}^{K} g_r(u^{(r)}, z^{(r)}) \simeq \sum_{r=1}^{K} h_r(\Delta^{(r)} u^{(r)})$$

where

- $h_r(\mathrm{v}^{(r)}) = \max\{\mathrm{v}^{(r,j)} + \eta^{(r,j)} \mid 1 \leq j \leq M\}$,
- $\eta^{(r,j)} = g_r(a_j^{(r)}, z^{(r)}) - \delta_j^{(r)} a_j^{(r)}$,
- $\delta_j^{(r)} \in \mathbb{R}$ is any subgradient of $g_r(\cdot, z^{(r)})$ at $a_j^{(r)}$,
- $\Delta^{(r)} = [\delta_1^{(r)}, \ldots, \delta_M^{(r)}]^\top$.

→ The approximation can be as close as desired by choosing M large enough.



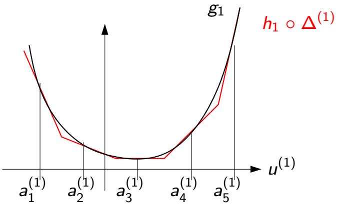
Algorithmic solution

$$g(u,z) = \sum_{r=1}^{K} g_r(u^{(r)}, z^{(r)}) \simeq \sum_{r=1}^{K} h_r(\Delta^{(r)} u^{(r)})$$

where

- $h_r(\mathrm{v}^{(r)}) = \max\{\mathrm{v}^{(r,j)} + \eta^{(r,j)} \mid 1 \leq j \leq M\}$,
- $\eta^{(r,j)} = g_r(a_j^{(r)}, z^{(r)}) - \delta_j^{(r)} a_j^{(r)}$,
- $\delta_j^{(r)} \in \mathbb{R}$ is any subgradient of $g_r(\cdot, z^{(r)})$ at $a_j^{(r)}$,
- $\Delta^{(r)} = [\delta_1^{(r)}, \ldots, \delta_M^{(r)}]^\top$.

$\rightarrow$ The approximation can be as close as desired by choosing M large enough.



Algorithmic solution

$$\underset{x\in\mathcal{H}}{\text{minimize}} \sum_{s=1}^{S} f_s(Fx) + \iota_C(x) + \iota_D(Tx)$$

$\Rightarrow$ Approximated criterion :

$$\underset{(x,\zeta)\in\mathcal{H}\times\mathbb{R}^K}{\text{minimize}} \sum_{s=1}^{S} f_s(Fx) + \iota_C(x) + \iota_V(\zeta) + \iota_E(\Delta Tx, \zeta)$$

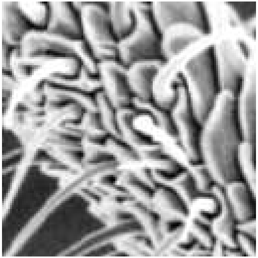
where

- $D = \{u \in \mathbb{R}^K \mid g(u,z) \leq \eta\}$,
- $V = \{\zeta \in \mathbb{R}^K \mid 1_K^\top \zeta \leq \eta\}$,
- $E = \{(v,\zeta) \in \mathbb{R}^{KM} \times \mathbb{R}^K \mid (\forall r \in \{1,\ldots,K\})\ (\mathrm{v}^{(r)}, \zeta^{(r)}) \in \operatorname{epi} h_r\}$,
- For every $u \in \mathbb{R}^K$, $g(u,z) = \sum_{r=1}^K g_r(u^{(r)}, z^{(r)}) \simeq \sum_{r=1}^K h_r(\Delta^{(r)} u^{(r)})$.

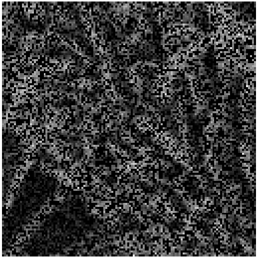
Experimental results

≪ **Compressed sensing** ≫ **experiment in the presence of Poisson noise** :

- Electron microscopy image of size $N = 128 \times 128$ ($\mathcal{H} = \mathbb{R}^N$),
- T denotes a randomly decimated blur : uniform blur of size 3×3 and approximately 60% of missing data, that leads to $K = 9834$,
- Poisson noise with scaling parameter 0.5.



Original

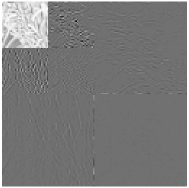


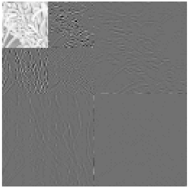
Degraded

Experimental results

Choice of the criterion : $\sum_{s=1}^{S} f_s(F\cdot) + \iota_C(\cdot) + \iota_D(T\cdot)$

- Data fidelity : approximation of the Poisson likelihood,
  - Influence of M,
- $C = [0, 255]^N$,
- F : Dual-Tree Transform (DTT) – symmlet 6, 2 levels,
- Blocks :
  - ℓ_1-reg : Classical ℓ_1 cost function,

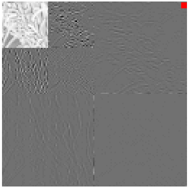


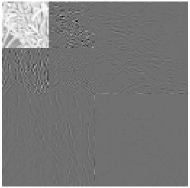


Experimental results

Choice of the criterion : $\sum_{s=1}^{S} f_s(F\cdot) + \iota_C(\cdot) + \iota_D(T\cdot)$

- Data fidelity : approximation of the Poisson likelihood,
  - Influence of M,
- $C = [0, 255]^N$,
- F : Dual-Tree Transform (DTT) – symmlet 6, 2 levels,
- Blocks :
  - ℓ_1-reg : Classical ℓ_1 cost function,

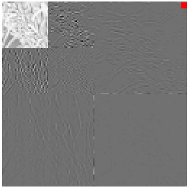


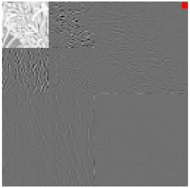


Experimental results

Choice of the criterion : $\sum_{s=1}^{S} f_s(F\cdot) + \iota_C(\cdot) + \iota_D(T\cdot)$

- Data fidelity : approximation of the Poisson likelihood,
  - Influence of M,
- $C = [0, 255]^N$,
- F : Dual-Tree Transform (DTT) – symmlet 6, 2 levels,
- Blocks :
  - ℓ_1-reg : Classical ℓ_1 cost function,
  - Block_PrimalDual : Blocks gathering primal and dual DTT coefficients,

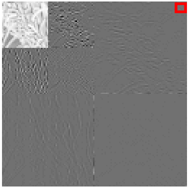


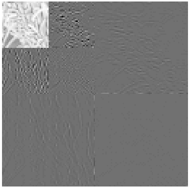


Experimental results

Choice of the criterion : $\sum_{s=1}^{S} f_s(F\cdot) + \iota_C(\cdot) + \iota_D(T\cdot)$

- Data fidelity : approximation of the Poisson likelihood,
  - Influence of M,
- $C = [0, 255]^N$,
- F : Dual-Tree Transform (DTT) – symmlet 6, 2 levels,
- Blocks :
  - ℓ_1-reg : Classical ℓ_1 cost function,
  - Block_PrimalDual : Blocks gathering primal and dual DTT coefficients,
  - Block_4Pixel_overlap : spatially overlapping blocks of size 2×2 are employed for each tree (primal or dual) separately.

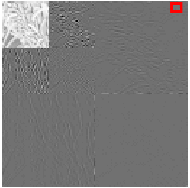


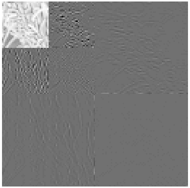


Experimental results

Choice of the criterion : $\sum_{s=1}^{S} f_s(F\cdot) + \iota_C(\cdot) + \iota_D(T\cdot)$

- Data fidelity : approximation of the Poisson likelihood,
  - Influence of M,
- $C = [0, 255]^N$,
- F : Dual-Tree Transform (DTT) – symmlet 6, 2 levels,
- Blocks :
  - ℓ_1-reg : Classical ℓ_1 cost function,
  - Block_PrimalDual : Blocks gathering primal and dual DTT coefficients,
  - Block_4Pixel_overlap : spatially overlapping blocks of size 2×2 are employed for each tree (primal or dual) separately.

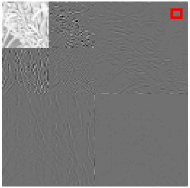


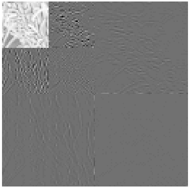


Experimental results

Choice of the criterion : $\sum_{s=1}^{S} f_s(F\cdot) + \iota_C(\cdot) + \iota_D(T\cdot)$

- Data fidelity : approximation of the Poisson likelihood,
  - Influence of M,
- $C = [0, 255]^N$,
- F : Dual-Tree Transform (DTT) – symmlet 6, 2 levels,
- Blocks :
  - ℓ_1-reg : Classical ℓ_1 cost function,
  - Block_PrimalDual : Blocks gathering primal and dual DTT coefficients,
  - Block_4Pixel_overlap : spatially overlapping blocks of size 2×2 are employed for each tree (primal or dual) separately.

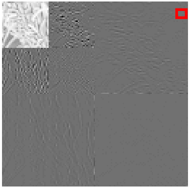


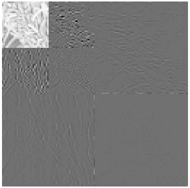


Experimental results

Choice of the criterion : $\sum_{s=1}^{S} f_s(F\cdot) + \iota_C(\cdot) + \iota_D(T\cdot)$

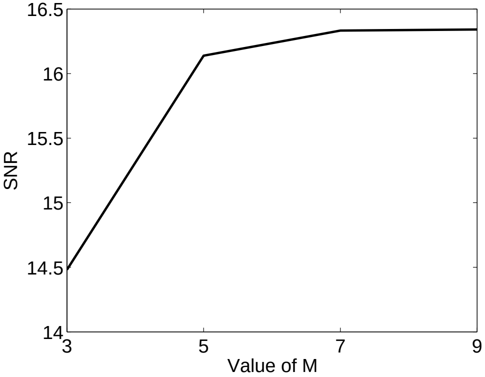
- Data fidelity : approximation of the Poisson likelihood,
  - Influence of M,
- $C = [0, 255]^N$,
- F : Dual-Tree Transform (DTT) – symmlet 6, 2 levels,
- Blocks :
  - ℓ_1-reg : Classical ℓ_1 cost function,
  - Block_PrimalDual : Blocks gathering primal and dual DTT coefficients,
  - Block_4Pixel_overlap : spatially overlapping blocks of size 2×2 are employed for each tree (primal or dual) separately.





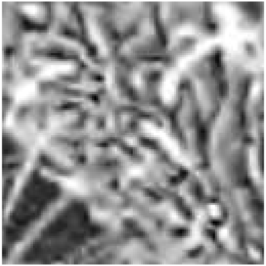
Experimental results

- Impact of M,
- Results for ℓ_1-reg,

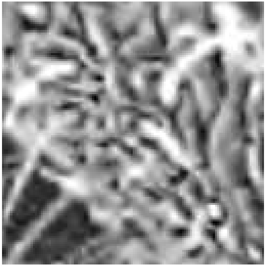


Experimental results

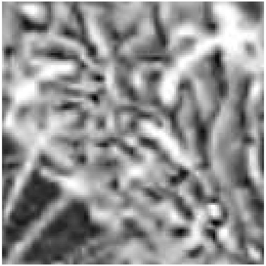
- Impact of M,
- Results for ℓ_1-reg,



$M = 3$
SNR = 14.5 dB



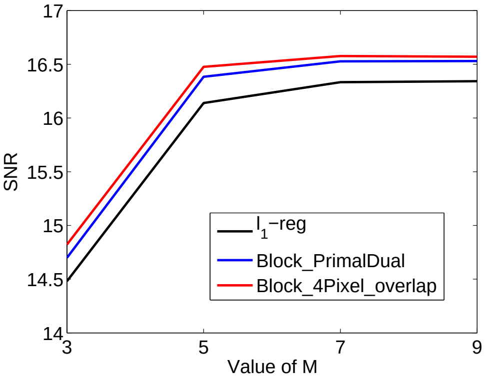
$M = 5$
SNR = 16.1 dB



$M = 7$
SNR = 16.3 dB

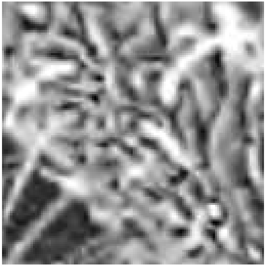
Experimental results

- Impact of M,
- Impact of the regularization term.

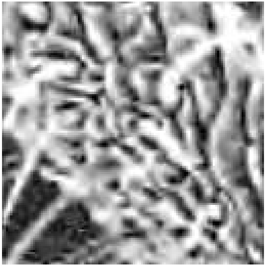


Experimental results

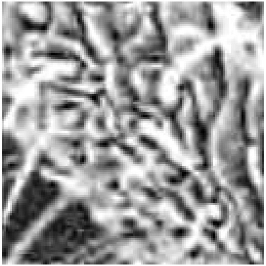
- $M = 7$,
- Impact of the regularization term.



ℓ_1-reg
SNR = 16.3 dB



Block_PrimalDual
SNR = 16.5 dB



Block_4Pixel_overlap
SNR = 16.6 dB

Conclusion and future works

- Convex optimization approach for solving cosparse modelling problems under flexible convex constraints.
- Use of recent proximal algorithms combined with a novel **epigraphical projection technique**.
- Approach applied to a reconstruction problem involving data corrupted with Poisson noise.

Thank you for your attention.