



HAL
open science

Optimisation globale pour la résolution de problèmes parcimonieux en norme l_0

Sébastien Bourguignon, Jordan Ninin, Hervé Carfantan, Marcel Mongeau

► **To cite this version:**

Sébastien Bourguignon, Jordan Ninin, Hervé Carfantan, Marcel Mongeau. Optimisation globale pour la résolution de problèmes parcimonieux en norme l_0 . Assemblée Générale du GDR ISIS, Mar 2015, Lyon, France. . hal-01194776

HAL Id: hal-01194776

<https://hal.science/hal-01194776v1>

Submitted on 9 Sep 2015

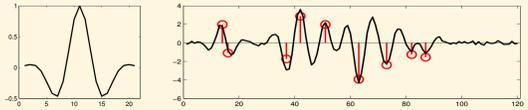
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

APPROXIMATION PARCIMONIEUSE

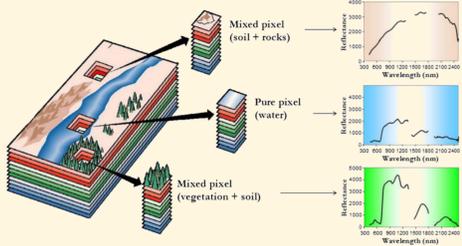
Problèmes inverses parcimonieux :
on cherche x avec peu de composantes non nulles tel que $y \simeq Hx$.

- **Déconvolution de trains d'impulsions**
(sismologie, contrôle non destructif par ultrasons)



filtre h ○ x , $-y = h * x + \text{bruit}$
 H = matrice de convolution
 $x \rightsquigarrow$ détection de discontinuités / défauts

- **Démélange de spectres** (imagerie hyperspectrale)



H = bibliothèque de spectres purs
 $x \rightsquigarrow$ composantes présentes + proportions

- **Estimation de modèles, sélection de variables, ...**

PROGRAMMATION MIXTE EN NOMBRES ENTIERS (MIP)

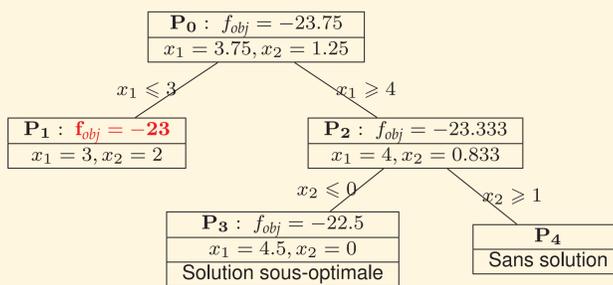
MIP quadratique à contraintes quadratiques :

$$\min_{v \in \mathbb{R}^J} v^T F v + c^T v, \text{ t.q. } \begin{cases} v^T B v + d^T v \leq e, & l_b \leq v \leq u_b, \\ A_{in} v \leq b_{in}, & A_{eq} v = b_{eq}, \\ v_j \in \mathbb{Z}, \forall j \in \mathcal{I}. \end{cases}$$

- **Méthodes de résolution exactes:** *Branch and Cut*

– **Branch and Bound :** énumération implicite des solutions

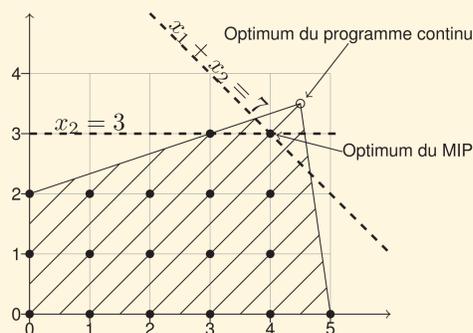
$$\min_{x_1, x_2} f_{obj} = -5x_1 - 4x_2 \text{ t.q. } \begin{cases} x_1 + x_2 \leq 5, & x_1 \geq 0, \\ 10x_1 + 6x_2 \leq 45, & x_2 \geq 0 \end{cases}$$



– **Méthodes de coupe (Gomory, ...)** :

\rightsquigarrow relaxations continues successives (variables binaires)

\rightsquigarrow raffinement de la solution du programme continu pour converger vers une solution entière



\Rightarrow **convergence avec preuve d'optimalité**

- **Choix de la mesure d'attache aux données**

– quadratique \rightsquigarrow MIP quadratiques

– linéaire par morceaux \rightsquigarrow MIP linéaires

$\rightsquigarrow \|y - Hx\|_1$ (bruit laplacien)

$\rightsquigarrow \|y - Hx\|_\infty$ (bruit uniforme)

REFORMULATIONS DE CRITÈRES PARCIMONIEUX

- **Reformulation de la norme ℓ_0**

– Variables binaires b_q telles que $x_q = 0 \Leftrightarrow b_q = 0$.

$$\rightsquigarrow \|x\|_0 = \sum_{q=1}^Q b_q.$$

– **Hypothèse de borne :** $\forall q, -M < x_q < M$. Alors :

$$\min_{x \in \mathcal{F}} \|x\|_0 \Leftrightarrow \min_{x \in \mathcal{F}} \sum_{q=1}^Q b_q \text{ t.q. } -Mb \leq x \leq Mb$$

\rightsquigarrow reformulation de $\mathcal{P}_{0/2}$ et \mathcal{P}_{0+2}

$$\|x\|_0 \leq K \Leftrightarrow \exists b \in \{0, 1\}^Q \text{ t.q. } \begin{cases} \sum_q b_q \leq K \\ -Mb \leq x \leq Mb \end{cases}$$

\rightsquigarrow reformulation de $\mathcal{P}_{2/0}$

- **Reformulation des attaches aux données ℓ_1 et ℓ_∞**

$$\min_{x \in \mathcal{F}} \|y - Hx\|_1 \Leftrightarrow \min_{x \in \mathcal{F}} \sum_{n=1}^n w_n \text{ t.q. } -w \leq y - Hx \leq w$$

$$\min_{x \in \mathcal{F}} \|y - Hx\|_\infty \Leftrightarrow \min_{x \in \mathcal{F}} t \text{ t.q. } -t1_N \leq y - Hx \leq t1_N$$

\rightsquigarrow **MIP linéaires à contraintes linéaires**

- **Un exemple :** $\min_{x \in \mathbb{R}^Q} \|y - Hx\|_1 \text{ t.q. } \|x\|_0 \leq K$

$$\Leftrightarrow \min_{\substack{x \in \mathbb{R}^Q \\ b \in \{0, 1\}^Q \\ w \in \mathbb{R}^N}} \sum_n w_n \text{ t.q. } \begin{cases} \sum_q b_q \leq K \\ -w \leq y - Hx \leq w \\ -Mb \leq x \leq Mb \end{cases}$$

COÛTS DE CALCUL

- **Problèmes de déconvolution parcimonieuse**

$x \in \mathbb{R}^{100}$, $K = 9$ composantes non nulles

Mise en œuvre sous IBM ILOG CPLEX

- **Réglage heuristique de la borne M :**

i) Initialisation $M_0 = \|H^T y\|_\infty / \|h\|_2^2$

ii) Si la borne est atteinte en la solution ($\hat{x}_q = M_k$), la reformulation de la norme ℓ_0 n'est pas valide. L'optimisation est alors relancée avec $M_{k+1} = 1.1M_k$.

| RSB | $\mathcal{P}_{0/2}$ | $\mathcal{P}_{0/1}$ | $\mathcal{P}_{0/\infty}$ | $\mathcal{P}_{2/0}$ | $\mathcal{P}_{1/0}$ | $\mathcal{P}_{\infty/0}$ | \mathcal{P}_{0+2} | \mathcal{P}_{0+1} | $\mathcal{P}_{0+\infty}$ |
|-------|---------------------|---------------------|--------------------------|---------------------|---------------------|--------------------------|---------------------|---------------------|--------------------------|
| 30 dB | 122 | 3 | 2.1 | 2.8 | 9.1 | 21 | 1.2 | 4.9 | 5.2 |
| 20 dB | 243 ⁽³⁾ | 4.7 | 2.6 | 14 | 28 | 45 ⁽¹⁾ | 64 | 17 ⁽²⁾ | 121 |
| 10 dB | 513 ⁽¹²⁾ | 21 | 4.5 | 35 ⁽¹⁾ | 182 ⁽¹⁾ | 206 ⁽⁴⁾ | 121 ⁽⁵⁾ | 234 ⁽²⁾ | 141 ⁽⁶⁾ |

Temps de calcul (s) pour l'optimisation des neuf problèmes reformulés en MIP (moyennes sur 20 réalisations).

Entre parenthèses : nombre de réalisations n'ayant pas fourni de solution optimale en 1000 s.

Rq : exploration combinatoire complète évaluée à 1.5 an.

\rightsquigarrow **Optimisation possible ... dans une certaine mesure**

\rightsquigarrow **Coût \nearrow avec le bruit et le nombre de composantes**

\rightsquigarrow **Attache aux données ℓ_2 : formulation $\mathcal{P}_{2/0}$**

ℓ_1 ou ℓ_∞ : formulation $\mathcal{P}_{0/p}$

- Plus de résultats dans [5]...

OPTIMISATION EN NORME ℓ_0

Problème d'optimisation bi-objectif :

$$\min_x \{ \|y - Hx\|_2^2, \|x\|_0 \}, \quad \|x\|_0 := \text{Card} \{q | x_q \neq 0\}$$

$$\rightarrow \min_x \|x\|_0 \text{ t.q. } \|y - Hx\|_2^2 \leq \epsilon \quad (\mathcal{P}_{0/2})$$

$$\rightarrow \min_x \|y - Hx\|_2^2 \text{ t.q. } \|x\|_0 \leq K \quad (\mathcal{P}_{2/0})$$

$$\rightarrow \min_x \|y - Hx\|_2^2 + \mu \|x\|_0 \quad (\mathcal{P}_{p+0})$$

- **Norme ℓ_0 à valeurs discrètes \rightsquigarrow pb combinatoire**

- **Méthodes approchées peu coûteuses :** [1]

– relaxation convexe : $\|x\|_0 \rightsquigarrow \|x\|_1$

– méthodes gloutonnes : exploration combinatoire partielle

... **mais ne résolvent pas le problème initial**

– H très corrélé \rightsquigarrow conditions théoriques inapplicables

– résultats empiriques défavorables

- **Travaux en norme ℓ_0 [2, 3, 4] : convergence locale**

- **Pourtant : l'optimisation exacte en norme ℓ_0 ne requiert pas forcément une recherche exhaustive!**

RÉFÉRENCES

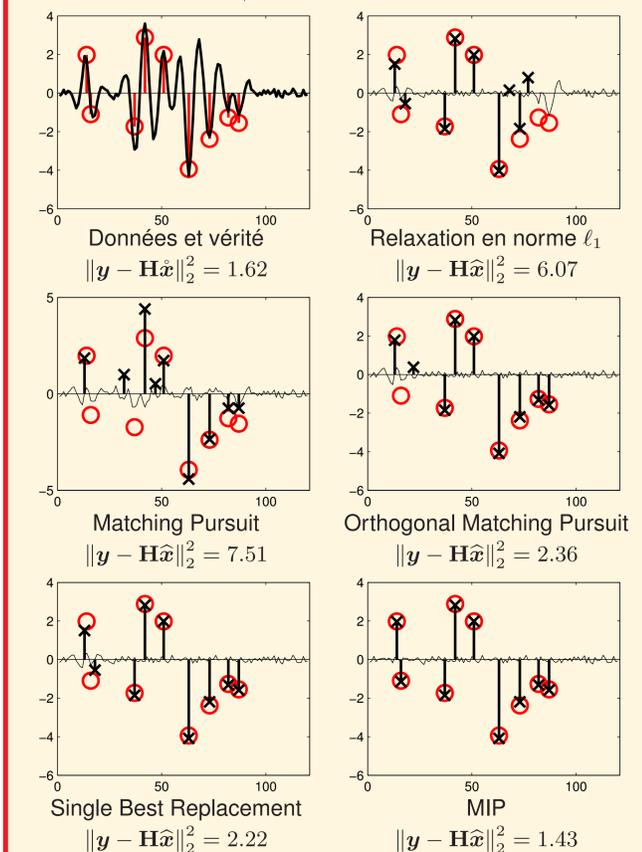
- [1] JA Tropp and SJ Wright, "Computational methods for sparse solution of linear inverse problems," *Proc. IEEE*, vol. 98, no. 6, pp. 948–958, 2010.
- [2] H Mohimani, M Babaie-Zadeh, and C Jutten, "A fast approach for overcomplete sparse decomposition based on smoothed ℓ_0 norm," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 289–301, 2009.
- [3] T Blumensath and ME Davies, "Iterative thresholding for sparse approximations," *J. Fourier Anal. Appl.*, vol. 14, no. 5–6, pp. 629–654, 2008.
- [4] Z Lu and Y Zhang, "Sparse approximation via penalty decomposition methods," *SIAM J. Optimization*, vol. 23, no. 4, pp. 2448–2478, 2013.
- [5] S Bourguignon, J Ninin, H Carfantan, and M Mongeau, "Exact resolution of sparse approximation problems via mixed-integer programming," *submitted to IEEE Trans. Signal Process.*, 2015.
- [6] I Tomic and S Drewes, "Learning joint intensity-depth sparse representations," *IEEE Trans. Image Process.*, vol. 23, no. 5, 2014.

REMERCIEMENTS

Ce travail a été réalisé dans le cadre d'un projet Jeunes Chercheurs 2013 du GdR ISIS (SB-JN). MM a été partiellement financé par l'Agence Nationale de la Recherche, projet JCJC ATOMIC n°12-JS02-009-01.

SOLUTIONS OBTENUES

- **Exemple pour $\mathcal{P}_{2/0}$ (RSB = 20 dB) :**



- **Erreurs d'estimation entre \hat{x} (vérité) et \hat{x} (estimé) :**

i) convolution des impulsions par un noyau gaussien ($\sigma = 1$)

ii) $d(\hat{x}, \hat{x}) :=$ erreur ℓ_2 entre les deux signaux convolués

| RSB | Relax. ℓ_1 | MP | OMP | SBR | MIP |
|-------|-----------------|-----|-----|-----|-----|
| 30 dB | 9.1 | 3.9 | 3.3 | 2.3 | 0.1 |
| 20 dB | 8.6 | 3.9 | 3.3 | 2.2 | 0.3 |
| 10 dB | 7.5 | 4.1 | 3.6 | 2.8 | 1.6 |

Distances $d(\hat{x}, \hat{x})$ moyennées sur 100 réalisations

PERSPECTIVES

- **Réduction du coût de calcul**

– Stratégies « intelligentes » de branchement basées sur la connaissance du problème

– Autres types de résolutions exactes : *quadratic convex reformulation* (QCR)

\rightsquigarrow approximations successives du MIP par des programmes d'optimisation semi-définie positive (SDP)

- Prise en compte de **contraintes supplémentaires** (positivité, somme à 1)

- Problèmes avec **parcimonie structurée** [6]

- **Optimisation multi-objectifs**