HAL
open science

Robustesse au bruit des régularisations polyhédrales

Samuel Vaiter, Gabriel Peyré, Jalal M. Fadili

To cite this version:

Samuel Vaiter, Gabriel Peyré, Jalal M. Fadili. Robustesse au bruit des régularisations polyhédrales. 24th GRETSI Symposium on Signal and Image Processing, Sep 2013, Brest, France. pp.ID130. hal-00927075

HAL Id: hal-00927075

https://hal.science/hal-00927075v1

Submitted on 13 Jan 2014

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.

Robustesse au bruit des régularisations polyhédrales

Samuel VAITER[1], Gabriel PEYRÉ[1], Jalal FADILI[2]

[1]CNRS, CEREMADE, Université Paris-Dauphine,
Place du Maréchal De Lattre De Tassigny, 75775 Paris Cedex 16, France

[2]GREYC, CNRS-ENSICAEN-Université de Caen,
6, Bd du Maréchal Juin, 14050 Caen Cedex, France

vaiter@ceremade.dauphine.fr, peyre@ceremade.dauphine.fr,
jalal.fadili@greyc.ensicaen.fr.

Résumé – Cet article traite de la robustesse au bruit d'une régularisation polyhédrale pour la résolution de problèmes inverses linéaires. Ce travail démontre le premier résultat de stabilité des régularisations de type polyhédrale permettant, entre autre, une analyse unifiée des régularisations parcimonieuses, parcimonieuses de type analyse et anti-parcimonieuses. Nous explicitons une condition qui assure que la face polyhédrale supportée par le vecteur d'entrée est égale à celle du vecteur retrouvé par régularisation polyhédrale dans le cadre d'une observation bruitée. Cette condition implique également que l'erreur ℓ^2 est proportionelle au niveau du bruit.

Abstract – In this paper, we establish robustness to noise perturbations of polyhedral regularization of linear inverse problems. We provide a sufficient condition that ensures that the polyhedral face associated to the true vector is equal to that of the recovered one. This criterion also implies that the ℓ^2 recovery error is proportional to the noise level for a range of parameter. Our criterion is expressed in terms of the hyperplanes supporting the faces of the unit polyhedral ball of the regularization. This generalizes to an arbitrary polyhedral regularization results that are known to hold for sparse synthesis and analysis ℓ^1 regularization which are encompassed in this framework. As a byproduct, we obtain recovery guarantees for ℓ^∞ and $\ell^1 - \ell^\infty$ regularization.

1 Introduction

1.1 Régularisation polyèdrale

Dans cet article, nous considérons le cadre des problèmes inverses linéaires prenant la forme suivante $y = \Phi x_0 + w$ où $y \in \mathbb{R}^Q$ est le vecteur d'observations, $x_0 \in \mathbb{R}^N$ le vecteur inconnu que l'on cherche à retrouver, w le bruit et Φ un opérateur linéaire du domaine signal $\mathbb{R}^N$ dans le domaine des observations $\mathbb{R}^Q$. Ce modèle permet de représenter par exemple des problèmes de convolution, tomographie et d'échantillonnage compressé.

Un polyhèdre $\mathcal{P}$ est un sous-ensemble de $\mathbb{R}^N$ tel que $\mathcal{P} = \{x \in \mathbb{R}^N \setminus Ax \leqslant b\}$ pour une certaine matrice $A \in \mathbb{R}^{N \times N_H}$ et $b \in \mathbb{R}^{N_H}$. L'inégalité $Ax \leqslant b$ doit se lire composante par composante. Nous considérons à partir de maintenant une fonctionnelle polyhédrale de la forme

$$J_H(x) = \max_{1 \leqslant i \leqslant N_H} \langle x, h_i \rangle,$$

où $H = (h_i)_{i=1}^{N_H}$ est une matrice de $\mathbb{R}^{N \times N_H}$. Ainsi, J_H est la jauge de l'ensemble convexe polyhédral étoilé compact $\mathcal{P}_H = \{x \in \mathbb{R}^N \setminus J_H(x) \leqslant 1\}$. Elle est ainsi continue, positive, bornée, et sous-linéaire [8]. Notons que en général, il ne s'agit pas d'une norme dans le cas où J_H n'est pas symétrique.

Afin de résoudre le problème inverse, nous introduisons la régularisation variationnelle suivante

$$x^\star \in \operatorname*{argmin}_{x \in \mathbb{R}^N} \frac{1}{2}\|y - \Phi x\|^2 + \lambda J_H(x), \qquad (P_\lambda(y))$$

avec $\lambda > 0$ le paramètre de régularisation. L'ensemble des minimiseurs est alors un ensemble compact convexe et non-vide par coercivité de la fonctionelle J_H.

Si le bruit est nul, $w = 0$, nous utiliserons plutôt la version contrainte du problème

$$x^\star \in \operatorname*{argmin}_{\Phi x = y} J_H(x). \qquad (P_0(y))$$

1.2 Lien avec la parcimonie et l'antiparcimonie

Nous donnons ici quelques exemples de régularisations connues qui se trouvent être polyhédrales. La norme ℓ^1 est définie comme

$$J_{H_1}(x) = \|x\|_1 = \sum_{i=1}^N |x_i|.$$

En terme de régularisation polyhédrale, il s'agit de choisir $H_1 \in \mathbb{R}^{N \times 2^N}$ tel que les colonnes de H_1 enumèrent tous les signes possibles de longueur N, i.e $\{-1, 1\}^N$. Le problème $(P_\lambda(y))$ devient alors le *Lasso*, introduit dans [9], ou *Basis Pursuit Denoising* [3], utilisé pour les modèles de données parcimonieux. Plus généralement, nous considérons les régularisation de type ℓ^1 analyse de la forme

$$J_H(x) = \|Lx\|_1,$$

où $L \in \mathbb{R}^{P\times N}$ est un dictionnaire d'analyse, éventuellement redondant. Cela correspond à prendre $H = L^* H_1$ où * correspond à la matrice adjointe. L'exemple le plus typique d'apriori analyse est la variation totale anisotrope.

La norme ℓ^∞ s'exprime comme

$$J_{H_\infty}(x) = \|x\|_\infty = \max_{1\leqslant i\leqslant N} |x_i|.$$

En terme de régularisation polyhédrale, le choix de la matrice H est $H_\infty = [\mathrm{Id}_N, -\mathrm{Id}_N] \in \mathbb{R}^{N\times 2N}$. On parle de régularisation anti-parcimonieuse, utilisé par exemple pour le calcul approché de plus proches voisins [6].

Il est possible [1] d'imposer une structure par blocs aux entrées non nulles induite par la parcimonie en définissant une norme $\ell^1 - \ell^\infty$. Soit $\mathcal{B}$ une partition de $\{1, \ldots, N\}$. La norme associée à cette structure de blocs est

$$J_{H^\infty_{\mathcal{B}}}(x) = \sum_{b\in\mathcal{B}} \|x_b\|_\infty,$$

correspondant à la matrice $H^\infty_{\mathcal{B}} \in \mathbb{R}^{N\times\prod_{b\in\mathcal{B}} 2^{|b|}}$ enumérant les signes possibles de chaque blocs. Si pour tout tous les blocs sont de taille 1, on retrouve la norme ℓ^1 tandis que si la structure de blocs est composée par un seul élément, on obtient la norme ℓ^∞.

2 Contributions

Définition 1. *On définit le H-*support $\mathrm{supp}_H(x)$ *d'un vecteur $x \in \mathbb{R}^N$ comme l'ensemble*

$$\mathrm{supp}_H(x) = \{i \in \{1, \ldots, N\} \setminus \langle x_i,\, h_i\rangle = J_H(x)\}\,.$$

Dans cet article, nous donnons une condition suffisante, dependant du H-support de x_0, sous laquelle le problème $(P_\lambda(y))$ admet une solution unique, et que si le rapport signal sur bruit est suffisamment grand, le H-support de cette solution coincide avec celui de x_0. De plus, si λ est choisi en proportion du bruit, l'écart ℓ^2 entre cette unique solution et x_0 est de l'ordre du niveau du bruit.

Définition 2. *Un H-support I satisfait la* condition d'injectivité restreinte *si*

$$\mathrm{Ker}\,\Phi \cap \mathrm{Ker}\, H_I^* = \{0\}, \qquad (\mathcal{C}_I)$$

où H_I désigne la matrice dont les colonnes sont celles de H indexées par I.

Lorsque celle-ci est satisfaite, on définit la projection orthogonale $\Gamma_I^\perp$ sur $\Phi(\mathrm{Ker}\, H_I^*)^\perp$ par les opérateurs suivants :

$$M_I = (U^*\Phi^*\Phi U)^{-1} \quad , \quad \Gamma_I = \Phi U M_I U^* \Phi^*$$
$$\text{et} \quad \Gamma_I^\perp = \mathrm{Id} - \Gamma_I.$$

où U est une base quelconque de $\mathrm{Ker}\, H_I^*$. La forme bilinéaire symétrique induite par $\Gamma_I^\perp$ sur $\mathbb{R}^N$ sera noté $\langle\cdot,\,\cdot\rangle_{\Gamma_I^\perp}$ et sa forme quadratique $\|\cdot\|^2_{\Gamma_I^\perp}$.

Définition 3. *Soit I un H-support satisfaisant $(\mathcal{C}_I)$. Le* critère d'identifiabilité *de I est*

$$\mathbf{IC}_H(I) = \max_{z_I\in\mathrm{Ker}\, H_I} \min_{i\in I} (\tilde{\Phi}_I^* \Gamma_I^\perp \tilde{\Phi}_I \mathbb{I}_I + z_I)_i$$

où $\mathbb{I}_I \in \mathbb{R}^{|I|}$ est composé uniquement de 1 et $\tilde{\Phi}_I = \Phi H_I^{+,*} \in \mathbb{R}^{Q\times|I|}$ avec $^+$ dénotant la pseudo-inverse de Moore–Penrose .

Le calcul de $\mathbf{IC}_H(I)$ nécessite la résolution d'un problème d'optimisation convexe, et peut se ramener à une programmation linéaire :

$$\mathbf{IC}_H(I) = \max_{(r,z_I)\in\mathbb{R}\times\mathbb{R}^{|I|}} r$$

sous la contrainte linéaire

$$\begin{cases} \forall i \in I, r \leqslant (\tilde{\Phi}_I^* \Gamma_I^\perp \tilde{\Phi}_I \mathbb{I}_I + z_I)_i \\ H_I z_I = 0. \end{cases}$$

2.1 Robustesse au bruit

Notre contribution principale est la suivante.

Théorème 1. *Soient $x_0 \in \mathbb{R}^N \setminus \{0\}$ et I son H-support satisfaisant $(\mathcal{C}_I)$. Soit $y = \Phi x_0 + w$. Supposons que $\tilde{\Phi}_I \mathbb{I}_I \neq 0$ et $\mathbf{IC}_H(I) > 0$. Alors ils existent deux constantes $c_I, \tilde{c}_I$ satisfaisant*

$$\frac{\|w\|_2}{T} < \frac{\tilde{c}_I}{c_I} \quad \text{où} \quad T = \min_{j\in I^c} J_H(x_0) - \langle x_0,\, h_j\rangle > 0,$$

tel que si λ vérifie la condition $c_I\|w\|_2 < \lambda < T\tilde{c}_I$, le vecteur $x^\star \in \mathbb{R}^N$ définit par

$$x^\star = \mu H_I^{+,*}\mathbb{I}_I + U M_I U^* \Phi^*(y - \mu\tilde{\Phi}_I\mathbb{I}_I)$$
$$\text{et} \quad 0 < \mu = J_H(x_0) + \frac{\langle \tilde{\Phi}_I \mathbb{I}_I,\, w\rangle_{\Gamma_I^\perp} - \lambda}{\|\tilde{\Phi}_I\mathbb{I}_I\|^2_{\Gamma_I^\perp}}$$

est l'unique solution de $(P_\lambda(y))$. De plus, il y a robustesse du support : $\mathrm{supp}_H(x^\star) = \mathrm{supp}_H(x_0)$. *Si λ est choisi proportionnellement au niveau du bruit, alors $\|x^\star - x_0\|_2 = O(\|w\|_2)$.*

Ce théorème exprime que pour une certaine classe de signaux, ceux qui vérifie la condition $\mathbf{IC}_H(I) > 0$ jointe sur les vecteurs actifs de H et les colonnes de Φ, la solution du problème $(P_\lambda(y))$ est unique lorsque le niveau du bruit $\|w\|_2$ n'est pas trop grand devant le niveau du signal T. Pour que cette analyse soit valide, il faut que le débruitage soit suffisant ($c_I\|w\|_2 < \lambda$) mais conserve le H-support ($\lambda < \tilde{c}_I T$). En pratique x_0 n'est pas connu, ainsi la formule de $x^\star$ est principalement d'un intérêt théorique afin de mesure l'erreur ℓ^2 entre x_0 et le vecteur retrouvé $x^\star$.

2.2 Identifiabilité

En l'absence de bruit $w = 0$, le résultat suivant montre que la condition $\mathbf{IC}_H(I) > 0$ implique également l'identifiabilité du signal.

Théorème 2. *Soient $x_0 \in \mathbb{R}^N \setminus \{0\}$ et I son H-support satisfaisant $(\mathcal{C}_I)$. Si $\mathbf{IC}_H(I) > 0$, alors le vecteur x_0 est l'unique solution de $(P_0(y))$.*

3 Relations à des travaux antérieurs

Dans le cas des normes ℓ^1 et ℓ^1 analyse, notre condition est équivalente à celles définies dans [5] et [10]. À notre connaissance, il n'existe pas de résultat de robustesse sur un bruit générique dans le cas de la norme ℓ^∞, mais [1] étudie la robustesse d'une sous-classe de norme polyhédrale obtenue par relaxation convexe de fonction de perte combinatoire. Une étude numérique de quelques régularisations polyhédrales peut se trouver dans [7]. Dans [4] une étude est faite de $(P_0(y))$ dans le cas d'une matrice aléatoire (échantillonage compressé). Dans ce même contexte, [2] étudie les épaisseurs gaussiennes d'une certaine sous-classe de norme polyhédrales dans le cas d'une régularisation sous contrainte.

4 Esquisse de preuve

Dans cette section, nous proposons une idée de la preuve du théorème 1.

Les conditions du premier ordre de minimalité du problème de minimisation $(P_\lambda(y))$ sont résumés dans le lemme suivant, conséquence de la forme de la sous-différentielle de $J_H(x^\star)$.

Lemme 1. *Un vecteur $x^\star$ est solution de $(P_\lambda(y))$ si, et seulement si, il existe $v_I \in \Sigma_I$ tel que*

$$\Phi^*(\Phi x - y) + \lambda H_I v_I = 0,$$

avec $I = \mathrm{supp}_H(x)$.

Ces conditions peuvent être raffinés par la connaissance du support de $x^\star$, permettant ainsi d'obtenir une condition suffisante d'unicité. Nous omettons ici la démonstration de ce lemme par manque de place.

Lemme 2. *Soit $x^\star \in \mathbb{R}^N$, $I = \mathrm{supp}_H(x^\star)$ et $\mu = J_H(x^\star)$. Supposons que $(\mathcal{C}_I)$ soit satisfait. Soit U une base de $\mathrm{Ker}\, H_I^*$. Il existe $z \in \mathrm{Ker}\, H_I$ tel que*

$$v_I = z_I + \frac{1}{\lambda}\tilde{\Phi}_I^* \Gamma_I^\perp (y - \mu \tilde{\Phi}_I \mathbb{I}_I) \in \Sigma_I,$$

si, et seulement si, $x^\star$ est une solution de $(P_\lambda(y))$. De plus, si $v_I \in \mathrm{ri}\,\Sigma_I$, alors $x^\star$ est l'unique solution de $(P_\lambda(y))$.

Rappelons que $\mathrm{ri}\,\Sigma_I$ représente l'intérieur relatif de Σ_I, i.e l'intérieur de Σ_I par rapport à la topologie de son enveloppe affine.

Nous pouvons dès lors démontrer le théorème 1. Soit I le H-support de x_0. Considérons la restriction de $(P_\lambda(y))$ au H-support I.

$$x^\star = \underset{\substack{x \in \mathbb{R}^N \\ \mathrm{supp}_H(x) \subseteq I}}{\mathrm{argmax}} \ \frac{1}{2}\|y - \Phi x\|_2^2 + J_H(x). \qquad (\mathcal{P}_\lambda(y)_I)$$

Du fait de $(\mathcal{C}_I)$, la fonction objectif est fortement convexe sur l'ensemble des signaux de H-support I. Ainsi, $x^\star$ est défini de façon unique.

La preuve est divisée en quatre parties. Nous explicitions (**1.**) une forme implicite de $x^\star$. Puis (**2.**), nous vérifions que le H-support de $x^\star$ est le même que celui de x_0 ainsi que la compatibilité (**3.**) de $J_H(x^\star)$. En utilisant le lemme 2, nous prouvons (**4.**) que $x^\star$ est l'unique minimiseur de $(P_\lambda(y))$.

1. Expression de $x^\star$. Du fait que $x^\star = \mu H_I^{+,*}\mathbb{I}_I + U\alpha$ avec $\mu = J_H(x^\star)$, nous avons

$$U^*\Phi^*(\Phi x - y) = \mu U^*\Phi^*\Phi H_I^{+,*}\mathbb{I}_I + (U^*\Phi^*\Phi U)\alpha - U^*\Phi^* y = 0.$$

Ainsi,

$$U\alpha = U M_I U^* \Phi^*(y - \mu \Phi H_I^{+,*}\mathbb{I}_I).$$

Dès lors, comme $y = \Phi x_0 + w$ avec $\mathrm{supp}_H(x_0) = I$, $x^\star$ s'exprime comme

$$\begin{aligned} x^\star &= \mu H_I^{+,*}\mathbb{I}_I + U M_I U^*\Phi^*(y - \mu\Phi H_I^{+,*}\mathbb{I}_I) \\ &= \mu H_I^{+,*}\mathbb{I}_I + U M_I U^*\Phi^*((\mu_0 - \mu)\Phi H_I^{+,*}\mathbb{I}_I + w) + U\alpha_0 \\ &= x_0 - (\mu_0 - \mu) H_I^{+,*}\mathbb{I}_I \\ &\quad + U M_I U^*\Phi^*((\mu_0 - \mu)\Phi H_I^{+,*}\mathbb{I}_I + w), \end{aligned}$$

avec $\mu_0 = J_H(x_0)$. Ainsi, $x^\star$ satisfait l'équation implicite suivante.

$$x^\star = x_0 + (\mu_0 - \mu)[U M_I U^*\Phi^*\Phi - \mathrm{Id}] H_I^{+,*}\mathbb{I}_I + U M_I U^*\Phi^* w. \qquad (1)$$

2. Vérification que le H-support de $x^\star$ est I. Pour certifier que le H-support de $x^\star$ est I, il faut imposer que

$$\begin{aligned} \forall i \in I, &\quad \langle h_i,\, x^\star\rangle = J_H(x^\star) = \mu \\ \forall j \in I^c, &\quad \langle h_j,\, x^\star\rangle < J_H(x^\star) = \mu. \end{aligned}$$

Les composantes de $x^\star$ sur I sont telles que $H_I^* x^\star = \mu \mathbb{I}_I$. Comme J_H est une fonctionelle sous-additive, nous bornons les composantes sur I^c par une inégalité triangulaire sur (1) pour obtenir

$$\begin{aligned} \max_{j \in I^c}\langle h_j,\, x^\star\rangle \leqslant &\max_{j \in I^c}\langle h_j,\, x_0\rangle \\ &+ (\mu_0 - \mu)\|H_{I^c}^*[U M_I U^*\Phi^*\Phi - \mathrm{Id}] H_I^{+,*}\mathbb{I}_I\|_\infty \\ &+ \|H_{I^c}^* U M_I U^*\Phi^* w\|_\infty. \end{aligned}$$

En notant

$$\begin{aligned} C_1 &= \|H_{I^c}^*[U M_I U^*\Phi^*\Phi - \mathrm{Id}] H_I^{+,*}\mathbb{I}_I\|_\infty, \\ C_2 &= \|H_{I^c}^* U M_I U^*\Phi^*\|_{2,\infty}, \\ T &= \mu_0 - \max_{j \in I^c}\langle h_j,\, x_0\rangle, \end{aligned}$$

nous bornons les corrélations en dehors du H-support par

$$\max_{j \in I^c}\langle h_j,\, x^\star\rangle \leqslant \mu_0 - T + (\mu_0 - \mu)C_1 + C_2\|w\|.$$

Il existe des constantes c_1, c_2 satisfaisant $c_1\|w\| < c_2 T + \lambda$ tel que

$$0 \leqslant \mu_0 - T + (\mu_0 - \mu)C_1 + C_2\|w\| < \mu \qquad (2)$$

Sous cette condition de faisabilité, nous obtenons que

$$\max_{j \in I^c}\langle h_j,\, x^\star\rangle < \mu,$$

ce qui prouve que le H-support de $x^\star$ est égal à I.

3. Valeur de $\mu = J_H(x^\star)$. En utilisant le lemme 2 avec $H = U^*H$, alors $x^\star$ est solution de $(\mathcal{P}_\lambda(y)_I)$. De plus, il existe $z_I \in \operatorname{Ker} H_I$ tel que

$$v_I = z_I + \frac{1}{\lambda}\tilde{\Phi}_I^*\Gamma_I^\perp(y - \mu\tilde{\Phi}_I\mathbb{I}_I) \in \Sigma_I. \tag{3}$$

On décompose x_0 comme

$$x_0 = \mu_0 H_I^{+,*}\mathbb{I}_I + U\alpha_0.$$

Du fait que $y = \Phi x_0 + w$, nous avons

$$\Gamma_I^\perp y = \Gamma_I^\perp(\mu_0\tilde{\Phi}_I\mathbb{I}_I + \Phi U\alpha_0 + w).$$

De plus, comme

$$\Gamma_I\Phi U\alpha_0 = \Phi U(U^*\Phi^*\Phi U)^{-1}U^*\Phi^*\Phi U\alpha_0 = \Phi U\alpha_0,$$

nous obtenons que

$$\Gamma_I^\perp y = \mu_0\Gamma_I^\perp\tilde{\Phi}_I\mathbb{I}_I + \Gamma_I^\perp w.$$

Dès lors, l'équation (3) peut se réécrire de façon équivalente comme

$$v_I = z_I + \frac{1}{\lambda}\tilde{\Phi}_I^*\Gamma_I^\perp\left((\mu_0 - \mu)\tilde{\Phi}_I\mathbb{I}_I + w\right).$$

En particulier, $\langle v_I, \mathbb{I}_I\rangle = \lambda$. Ainsi,

$$\lambda = \langle\lambda v_I, \mathbb{I}_I\rangle = \langle\lambda\tilde{z}_I, \mathbb{I}_I\rangle + \langle\tilde{\Phi}_I^*\Gamma_I^\perp((\mu_0 - \mu)\tilde{\Phi}_I\mathbb{I}_I + w, \mathbb{I}_I\rangle.$$

Du fait que $\tilde{z}_I \in \operatorname{Ker} H_I$, nous avons $\langle z_I, \mathbb{I}_I\rangle = 0$. Ainsi μ satisfait l'équation suivante :

$$\begin{aligned}\lambda &= \langle\tilde{\Phi}_I^*\Gamma_I^\perp((\mu_0 - \mu)\tilde{\Phi}_I\mathbb{I}_I + w, \mathbb{I}_I\rangle \\ &= (\mu_0 - \mu)\|\tilde{\Phi}_I\mathbb{I}_I\|^2_{\Gamma_I^\perp} + \langle\tilde{\Phi}_I\mathbb{I}_I, w\rangle_{\Gamma_I^\perp}.\end{aligned}$$

Ainsi la valeur de μ est donnée par

$$\mu = \mu_0 + \frac{\langle\tilde{\Phi}_I\mathbb{I}_I, w\rangle_{\Gamma_I^\perp} - \lambda}{\|\tilde{\Phi}_I\mathbb{I}_I\|^2_{\Gamma_I^\perp}} > 0. \tag{4}$$

4. Conditions du lemme 2. Nous allons construire un certificat $\tilde{v}_I$ grâce au contrôle imposé par $\mathbf{IC}_H(I) < 1$. Considérons maintenant le vecteur $\tilde{v}_I$ défini par

$$\tilde{v}_I = \tilde{z}_I + \frac{1}{\lambda}\tilde{\Phi}_I^*\Gamma_I^\perp\left((\mu_0 - \mu)\tilde{\Phi}_I\mathbb{I}_I + w\right),$$

avec

$$\tilde{z}_I = \frac{1}{\mu - \mu_0}\left(\operatorname*{argmax}_{z_I \in \operatorname{Ker} H_I} \min_{i\in I}(\tilde{\Phi}_I^*\Gamma_I^\perp\tilde{\Phi}_I\mathbb{I}_I + z_I)_i\right)$$

Quand la condition (2) est vérifiée, le H-support de $x^\star$ est I, donc nous avons simplement à vérifier que $\tilde{v}_I$ est un élément de $\operatorname{ri}\Sigma_I$. Comme $\langle\tilde{z}_I, \mathbb{I}_I\rangle = 0$, nous calculons

$$\begin{aligned}&\langle\tilde{v}_I, \mathbb{I}_I\rangle \\ =&\langle z_I + \frac{1}{\lambda}\tilde{\Phi}_I^*\Gamma_I^\perp\left((\mu_0 - \mu)\tilde{\Phi}_I\mathbb{I}_I + w\right), \mathbb{I}_I\rangle + \langle\tilde{z}_I - z_I, \mathbb{I}_I\rangle \\ =&\langle v_I, \mathbb{I}_I\rangle + 0 \\ =&\lambda.\end{aligned}$$

En reportant l'expression (4) de $(\mu_0 - \mu)$ dans la définition de $\tilde{v}_I$, nous obtenons l'expression

$$\tilde{v}_I = \tilde{z}_I + \frac{1}{\lambda}\left(\tilde{\Phi}_I^*\Gamma_I^\perp w + \frac{\langle\tilde{\Phi}_I\mathbb{I}_I, w\rangle_{\Gamma_I^\perp} - \lambda}{\|\tilde{\Phi}_I\mathbb{I}_I\|^2_{\Gamma_I^\perp}}\tilde{\Phi}_I^*\Gamma_I^\perp\tilde{\Phi}_I\mathbb{I}_I\right).$$

Pour une certaine constante c_3 tel que $c_3\|w\| - \mathbf{IC}_H(I)\cdot\lambda > 0$, nous avons

$$\forall i \in I, \quad v_i > 0.$$

En combinant cette remarque avec le fait que $\langle\tilde{v}_I, \mathbb{I}_I\rangle = \lambda$, ceci prouve que $\tilde{v}_I$ est un élément de l'intérieur relatif de Σ_I. D'après le lemme 2, $x^\star$ est ainsi l'unique minimiseur de $(P_\lambda(y))$, ce qui conclut la preuve du théorème.

Références

[1] F. Bach. Structured sparsity-inducing norms through submodular functions. *Advances in Neural Information Processing Systems*, 2010.

[2] V. Chandrasekaran, B. Recht, P.A. Parrilo, and A.S. Willsky. The Convex Geometry of Linear Inverse Problems. *Foundations of Computational Mathematics*, 12(6) :805–849, 2012.

[3] S.S. Chen, D.L. Donoho, and M.A. Saunders. Atomic decomposition by basis pursuit. *SIAM journal on scientific computing*, 20(1) :33–61, 1998.

[4] D.L. Donoho and J. Tanner. Counting the faces of randomly-projected hypercubes and orthants, with applications. *Discrete & computational geometry*, 43(3) :522–541, 2010.

[5] J.J. Fuchs. On sparse representations in arbitrary redundant bases. *Information Theory, IEEE Transactions on*, 50(6) :1341–1344, 2004.

[6] H. Jégou, T. Furon, and J.J. Fuchs. Anti-sparse coding for approximate nearest neighbor search. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 2029–2032. IEEE, 2012.

[7] S. Petry and G. Tutz. Shrinkage and variable selection by polytopes. *Journal of Statistical Planning and Inference*, 142(1) :48–64, 2012.

[8] R.T. Rockafellar. *Convex analysis*, volume 28. Princeton University Press, 1996.

[9] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.

[10] S. Vaiter, G. Peyré, C. Dossal, and J. Fadili. Robust sparse analysis regularization. *to appear in IEEE Transactions on Information Theory*, 2012.