



**HAL**  
open science

## **OPTIMED : Optimisation itérative pour la résolution de problèmes inverses de grande taille**

Lotfi Chaâri, Emilie Chouzenoux, Nelly Pustelnik, Caroline Chaux, Saïd Moussaoui

► **To cite this version:**

Lotfi Chaâri, Emilie Chouzenoux, Nelly Pustelnik, Caroline Chaux, Saïd Moussaoui. OPTIMED : Optimisation itérative pour la résolution de problèmes inverses de grande taille. *Traitement du Signal*, 2011, 28 (3-4), pp.329-374. hal-00733443

**HAL Id: hal-00733443**

**<https://hal.science/hal-00733443>**

Submitted on 18 Sep 2012

**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.

---

# OPTIMED : Optimisation itérative pour la résolution de problèmes inverses de grande taille

Lotfi Chaari <sup>1</sup>, Emilie Chouzenoux <sup>2</sup>, Nelly Pustelnik <sup>3</sup>, Caroline Chaux <sup>2</sup>, Saïd Moussaoui <sup>4</sup>

<sup>1</sup> INRIA Rhône-Alpes, Projet MISTIS

<sup>2</sup> Université Paris-Est, IGM (UMR CNRS 8049)

<sup>3</sup> Laboratoire de Physique, ENS Lyon (UMR CNRS 5672)

<sup>4</sup> LUNAM Université, Ecole Centrale Nantes, IRCCyN (UMR CNRS 6597)

---

*RÉSUMÉ.* Cet article s'intéresse au développement d'algorithmes itératifs pour la reconstruction d'images. La particularité de ces problèmes réside dans la taille élevée des images et la quantité importante de données générée par le processus de mesure. Outre le fait de surmonter la difficulté liée à la grande taille du problème, les algorithmes proposés se fondent sur des résultats de convergence théoriquement établis. Nous illustrons ces développements à l'aide de trois modalités d'imagerie distinctes : la tomographie par émission de positrons (TEP), l'imagerie par résonance magnétique nucléaire (IRM) parallèle et la résonance magnétique nucléaire bidimensionnelle (RMN 2D).

*ABSTRACT.* This paper deals with image reconstruction using iterative algorithms. The framework under consideration involves large amount of data and/or large observation operator. We proposed here algorithms for which we were able to state convergence proofs. The reliability and accuracy of the proposed approaches are demonstrated via three applications: the positron emission tomography (PET), the parallel magnetic resonance imaging (MRI) and two-dimensional nuclear magnetic resonance (2D NMR).

*MOTS-CLÉS :* Optimisation convexe, algorithmes itératifs, TEP, IRM, RMN-2D

*KEYWORDS:* Convex optimization, iterative algorithms, PET, MRI, 2D-NMR

---

## 1. Introduction

Le projet ANR OPTIMED (Algorithmes d'optimisation décomposés pour les problèmes d'imagerie médicale de grande taille) fait partie de l'action de recherche MDMSA (Masse de Données : Modélisation, Simulation, Applications). L'objectif de ce projet était, d'une part, de proposer des solutions méthodologiques pour la résolution de problèmes de reconstruction d'images de grande taille et, d'autre part, de les appliquer à des problèmes spécifiques en imagerie médicale. Deux modalités d'imagerie médicale ont été étudiées : la tomographie par émission de positrons (TEP) et l'imagerie par résonance magnétique (IRM). Par ailleurs, l'application a été étendue au cas de la résonance magnétique nucléaire bidimensionnelle (RMN-2D), qui est une modalité de mesure ayant un besoin avéré d'une méthode de reconstruction efficace.

Sur le plan méthodologique, l'enjeu était d'apporter des solutions à des problèmes inverses de grande taille (3D, 3D+t,...) en résolvant des problèmes d'optimisation où la critère à minimiser reflète le modèle physique d'acquisition des données ainsi que des connaissances *a priori* dont on peut disposer sur l'objet imagé. Deux aspects ont été particulièrement pris en compte dans le développement des méthodes :

– l'amélioration de la qualité des images reconstruites par un choix judicieux du critère à optimiser, en accord avec les propriétés statistiques des bruits rencontrés (poissonien ou gaussien) et en introduisant des informations *a priori* pertinentes. Parmi les modèles *a priori* envisageables, ceux liés au caractère creux ou parcimonieux des données utiles, après une certaine représentation linéaire (e.g. une décomposition sur une trame d'ondelettes), sont attractifs pour des données volumineuses présentant une régularité spatiale et temporelle. Par ailleurs, la formulation lagrangienne des contraintes que doit respecter la solution nécessite le rajout d'une fonction barrière dans le critère. Ainsi, certains choix de modèles de bruit (e.g. poissonien) et d'*a priori* peuvent conduire à un problème d'optimisation de fonctions convexes non différentiables, voire non finies. Or, au début de ce projet, les méthodes d'optimisation permettant d'aborder de tels problèmes étaient assez limitées.

– le temps de calcul est aussi un aspect crucial dans un contexte de diagnostic médical. Il est donc nécessaire de développer et implémenter des algorithmes itératifs ayant une vitesse de convergence aussi rapide que possible. Ainsi, un compromis entre vitesse de convergence et coût de calcul par itération doit être réalisé. Les méthodes de descente itérative gradient reliées permettent d'atteindre un tel objectif à condition de disposer de méthodes de préconditionnement et de recherche de pas peu coûteuses. Il convient, par ailleurs, d'exploiter au mieux les moyens de calculs existant en tirant, par exemple, profit du parallélisme offert par les architectures multi-coeurs modernes. Ceci nous a amené à concevoir des algorithmes d'optimisation décomposés adaptés à cette parallélisation.

En s'appuyant sur trois problèmes de reconstruction d'images de grande taille, les développements méthodologiques réalisés pour, d'une part, prendre en compte les spécificités de chaque modalité de mesure et, d'autre part, proposer des algorithmes d'optimisation efficaces et présentant un coût de calcul raisonnable, sont présentés

ici. La section 2 récapitule les deux grandes familles d’algorithmes d’optimisation utilisées ; les approches proximales et les méthodes dites de majoration-minimisation (MM), ainsi que le lien entre ces deux approches. Les sections 3, 4 et 5 présentent respectivement les trois applications traitées. Enfin, quelques conclusions issues de l’étude menée au sein de ce projet ANR sont développées dans la Section 6.

## 2. Quelques schémas itératifs pour la résolution du problème d’optimisation

### 2.1. Approches proximales

De nombreux travaux ont porté sur la minimisation d’un critère convexe composé d’une somme de deux fonctions non-nécessairement différentiables. On peut par exemple citer, l’algorithme EM (Expectation-Maximization) introduit dans (Figueiredo *et al.*, 2003), l’algorithme Landweber seuillé (Bect *et al.*, 2004; Daubechies *et al.*, 2004) également connu sous le nom d’algorithme explicite-implicite dans une forme plus générale (Combettes *et al.*, 2005) et l’algorithme de Douglas-Rachford (Combettes *et al.*, 2007a). Puis dans (Combettes *et al.*, 2008), les auteurs ont proposé une extension de l’algorithme de Douglas-Rachford dans un espace produit (PPXA) permettant de résoudre la minimisation d’une somme de plus de deux fonctions non-nécessairement différentiables. Plus récemment d’autres algorithmes proximaux ont vu le jour tels que PPXA+ et ses dérivées (Pesquet *et al.*, 2011; Afonso *et al.*, 2011b; Setzer *et al.*, 2010) ou encore les méthodes primales-duales (Chen *et al.*, 1994; Esser *et al.*, 2010; Chambolle *et al.*, 2011; Combettes *et al.*, 2011; Briceño-Arias *et al.*, 2011).

L’ensemble de ces algorithmes fait appel à la notion d’opérateur proximal introduit par Moreau dans (Moreau, 1962; Moreau, 1965) et possèdent des garanties de convergence. Nous définirons dans un premier temps cet opérateur et certaines de ses propriétés qui nous seront utiles dans la suite de l’article. Nous présenterons ensuite les itérations et les propriétés de convergence de l’algorithme explicite-implicite, de l’algorithme de Douglas-Rachford et de PPXA qui sont les algorithmes proximaux mis en œuvre dans les sections 3 et 4. On pourra se référer à l’ouvrage de Bauschke et Combettes (Bauschke *et al.*, 2011) pour plus de détails sur l’analyse convexe et la théorie des opérateurs monotones.

#### 2.1.1. Opérateur proximal

**Definition 2.1** Soit  $\mathcal{H}$  un espace de Hilbert. L’opérateur proximal de  $f \in \Gamma_0(\mathcal{H})$ <sup>1</sup> (noté  $\text{prox}_f$ ) associe à tout point  $u \in \mathcal{H}$ , l’unique minimiseur de la fonction  $f + \|\cdot - u\|^2$ , i.e.

$$\text{prox}_f : \mathcal{H} \rightarrow \mathcal{H} : u \mapsto \arg \min_{v \in \mathcal{H}} \frac{1}{2} \|v - u\|^2 + f(v). \quad [1]$$

1.  $\Gamma_0(\mathcal{H})$  désigne l’ensemble des fonctions de  $\mathcal{H}$  dans  $] - \infty, +\infty[$ , convexes, semi-continues inférieurement et propres.

Cet opérateur généralise la notion de projection sur un ensemble convexe fermé non vide  $C \subset \mathcal{H}$  telle que  $\text{prox}_{\iota_C} = P_C$ .<sup>2</sup> Une forme bien connue de cet opérateur proximal est celle associée à une norme  $\ell_1$  (antilog vraisemblance d'une loi de Laplace).

Des formes explicites d'opérateurs proximaux pour des distributions gaussienne généralisée, exponentielle, gamma, chi, uniforme et, plus généralement, pour des densités de probabilité log-concaves sont données dans (Combettes *et al.*, 2005; Chaux *et al.*, 2007; Combettes *et al.*, 2010).

La propriété suivante concerne le calcul de l'opérateur proximal d'une fonction  $f$  de  $\Gamma_0(\mathcal{H})$  composée avec un opérateur linéaire borné.

**Proposition 2.1** (Combettes *et al.*, 2007a, proposition 11) *Soit  $\mathcal{G}$  un espace de Hilbert réel, soit  $f$  une fonction de  $\Gamma_0(\mathcal{G})$  et soit  $L: \mathcal{H} \rightarrow \mathcal{G}$  un opérateur linéaire borné. Supposons que la composition de  $L$  et  $L^*$  satisfait  $L \circ L^* = \chi \text{Id}$ , pour  $\chi \in ]0, +\infty[$ . On a alors :  $f \circ L \in \Gamma_0(\mathcal{H})$  et*

$$\text{prox}_{f \circ L} = \text{Id} + \chi^{-1} L^* \circ (\text{prox}_{\chi f} - \text{Id}) \circ L. \quad [2]$$

Une généralisation de cette proposition est faite dans (Pustelnik *et al.*, 2011, Proposition 3.4) lorsque  $f$  est séparable et  $L$  est un isomorphisme diagonalisable par  $(o_m)_{m \in \mathbb{K} \subset \mathbb{N}}$  une base orthonormale de  $\mathcal{G}$ .

Nous allons à présent introduire quelques algorithmes proximaux. Ces derniers sont robustes aux erreurs commises dans le calcul des gradients et des opérateurs proximaux. Afin de quantifier l'erreur qui est tolérée, on définit pour deux suites  $(x_n)_{n \in \mathbb{N}}$  et  $(y_n)_{n \in \mathbb{N}}$  de  $\mathcal{H}$ ,

$$(\forall n \in \mathbb{N}) \quad x_n \approx y_n \Leftrightarrow \sum_{n \in \mathbb{N}} \|x_n - y_n\| < +\infty. \quad [3]$$

### 2.1.2. Algorithme explicite-implicite (FB)

Soient  $f_1$  et  $f_2$  deux fonctions de  $\Gamma_0(\mathcal{H})$  telles que  $\text{Argmin } f_1 + f_2 \neq \emptyset$  où  $f_2$  est une fonction différentiable sur  $\mathcal{H}$ , de gradient  $\beta$ -Lipschitz. On souhaite trouver

$$\hat{x} \in \underset{x \in \mathcal{H}}{\text{Argmin}} f_1(x) + f_2(x). \quad [4]$$

Soit  $x_0 \in \mathcal{H}$  la valeur initiale. L'algorithme construit la suite  $(x_\ell)_{\ell \in \mathbb{N}}$  en prenant, pour tout  $\ell \in \mathbb{N}$ ,

$$x_{\ell+1} \approx x_\ell + \lambda_\ell (\text{prox}_{\gamma_\ell f_1}(x_\ell - \gamma_\ell \nabla f_2(x_\ell)) - x_\ell) \quad [5]$$

où  $\gamma_\ell > 0$  désigne le « pas » de l'algorithme et  $\lambda_n > 0$  le paramètre de relaxation. La convergence faible de  $(x_\ell)_{\ell \in \mathbb{N}}$  vers une solution du problème [4] a été démontrée dans (Combettes *et al.*, 2005) lorsque les hypothèses suivantes sont vérifiées :

---

2. projecteur sur l'ensemble convexe fermé non vide  $C \subset \mathcal{H}$

- 1)  $0 < \underline{\gamma} \leq \bar{\gamma} < 2\beta^{-1}$  où  $\underline{\gamma} = \inf_{\ell \in \mathbb{N}} \gamma_\ell$  et  $\bar{\gamma} = \sup_{\ell \in \mathbb{N}} \gamma_\ell$ ,
- 2)  $0 < \underline{\lambda} \leq \lambda_\ell \leq 1, \forall \ell \in \mathbb{N}$

La convergence forte de cet algorithme est démontrée dans (Combettes *et al.*, 2005, Théorème 3.4) et (Combettes *et al.*, 2007b, Théorème 4.5) sous certaines hypothèses supplémentaires. Des résultats de convergence linéaire sont disponibles dans la littérature, par exemple dans (Chen *et al.*, 1997; Bredies *et al.*, 2008; Chaux *et al.*, 2009).

### 2.1.3. Algorithme de Douglas-Rachford (DR)

Pour cet algorithme, la contrainte de différentiabilité au sens de Lipschitz formulée dans le paragraphe précédent est relâchée. De plus, nous considérons le problème de minimisation où  $f_1$  et  $f_2$  sont deux fonctions de  $\Gamma_0(\mathcal{H})$  telles que  $\text{Argmin}(f_1 + f_2) \neq \emptyset$  et on suppose que  $\text{int}(\text{dom } f_1) \cap \text{dom } f_2 \neq \emptyset$ .<sup>3</sup> On souhaite trouver

$$\hat{x} \in \underset{x \in \mathcal{H}}{\text{Argmin}} f_1(x) + f_2(x). \quad [6]$$

Soit  $x_0 \in \mathcal{H}$  la valeur initiale. L'algorithme construit la suite  $(x_\ell)_{\ell \geq 0}$  telle que pour tout  $\ell \in \mathbb{N}$ ,

$$\begin{cases} x_{\ell+\frac{1}{2}} \approx \text{prox}_{\gamma f_2} x_\ell \\ x_{\ell+1} \approx x_\ell + \lambda_\ell (\text{prox}_{\gamma f_1}(2x_{\ell+\frac{1}{2}} - x_\ell) - x_{\ell+\frac{1}{2}}) \end{cases} \quad [7]$$

où  $\gamma > 0$  désigne le « pas » de l'algorithme,  $(\lambda_\ell)_{\ell \in \mathbb{N}}$  est une suite de réels positifs associée au paramètre de relaxation. La convergence faible de  $(x_\ell)_{\ell \in \mathbb{N}}$  vers  $\tilde{x} \in \mathcal{H}$ , telle que  $\hat{x} = \text{prox}_{\gamma f_2} \tilde{x}$  est une solution du problème [6], est garantie sous l'hypothèse suivante :

- 1)  $\lambda_\ell \in ]0, 2[$  et  $\sum_{\ell \in \mathbb{N}} \lambda_\ell (2 - \lambda_\ell) = +\infty, \forall \ell \in \mathbb{N}$ ,

Des conditions de convergence faible de la suite  $(x_{\ell+\frac{1}{2}})_{\ell \in \mathbb{N}}$  ont récemment été données dans (Svaiter, 2011). Par ailleurs, des conditions de convergence forte sont données dans (Chaux *et al.*, 2009; Combettes, 2009).

### 2.1.4. Algorithme parallèle proximal (PPXA)

Dans (Combettes *et al.*, 2008), les auteurs se sont intéressés à la formulation de l'algorithme de Douglas-Rachford dans l'espace produit de façon à proposer un algorithme permettant de trouver une solution au problème d'optimisation convexe :

$$\hat{x} \in \underset{x \in \mathcal{H}}{\text{Argmin}} \sum_{j=1}^J f_j(x_j). \quad [8]$$

où, pour tout  $i \in \{1, \dots, J\}$ ,  $f_i$  est une fonction de  $\Gamma_0(\mathcal{H})$ . Soient  $(u_{j,0})_{1 \leq j \leq J} \in (\mathcal{H})^J$ ,  $(\omega_j)_{1 \leq j \leq J} \in ]0, 1]^J$  tel que  $\sum_{j=1}^J \omega_j = 1$  et  $x_0 = \sum_{j=1}^J \omega_j u_{j,0}$  la valeur initiale. L'algorithme PPXA construit la suite  $(x_\ell)_{\ell \geq 0}$  telle que pour tout  $\ell \in \mathbb{N}$ ,

3.  $\text{dom } f$  : domaine de la fonction  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

**Pour**  $j = 1, \dots, J$  **faire**  
  |  $p_{j,\ell} \approx \text{prox}_{\gamma f_j / \omega_j} u_{j,\ell}$   
**Fin Pour**  
 $p_\ell = \sum_{j=1}^J \omega_j p_{j,\ell}$   
Choisir  $\lambda_\ell \in ]0, 2[$ .  
**Pour**  $j = 1, \dots, J$  **faire**  
  |  $u_{j,\ell+1} = u_{j,\ell} + \lambda_\ell (2 p_\ell - x_\ell - p_{j,\ell})$   
**Fin Pour**  
 $x_{\ell+1} = x_\ell + \lambda_\ell (p_\ell - x_\ell)$

Algorithme 1: Algorithme PPXA.

où  $\gamma > 0$  désigne le « pas » de l'algorithme,  $(\lambda_\ell)_{\ell \in \mathbb{N}}$  est une suite de réels positifs associée au paramètre de relaxation. La suite  $(x_\ell)_{\ell \geq 1}$  générée par l'algorithme PPXA converge vers une solution du problème [8]. La convergence est démontrée dans (Combettes *et al.*, 2008) sous les hypothèses suivantes :

- 1)  $\lim_{\|x\| \rightarrow +\infty} f_1(x) + \dots + f_J(x) = +\infty$ .
- 2)  $\bigcap_{j=1}^J \text{rint dom } f_j \neq \emptyset$ .
- 3)  $\sum_{\ell \in \mathbb{N}} \lambda_\ell (2 - \lambda_\ell) = +\infty$ .

L'hypothèse 1 traduit la coercivité de la somme de fonctions à minimiser. Les hypothèses 1 et 2 assurent l'existence d'une solution.

## 2.2. Approches par Majoration-Minimisation

### 2.2.1. Schéma général

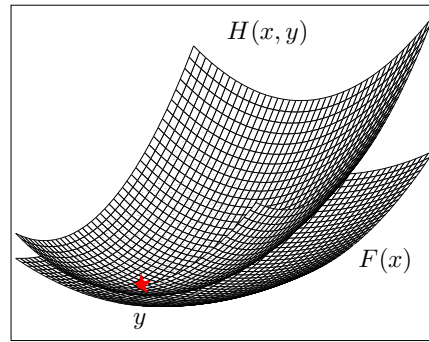
Le principe de Majoration-Minimisation (MM) apparaît pour la première fois dans l'ouvrage (Ortega *et al.*, 1970). Il consiste à substituer une séquence de problèmes plus simples au problème initial d'optimisation

$$\min_{\mathbf{x} \in \mathcal{D}_f} F(\mathbf{x}), \quad [9]$$

en approchant le critère  $F$  par une *approximation tangente majorante*, c'est-à-dire une fonction tangente au critère en un point et supérieure ou égale au critère en tous les autres points :

$$\begin{cases} H(\mathbf{x}, \mathbf{y}) \geq F(\mathbf{x}), & \forall \mathbf{x} \in \mathcal{D}_f \\ H(\mathbf{y}, \mathbf{y}) = F(\mathbf{y}) \end{cases} \quad [10]$$

La figure 1 illustre le principe de la majoration sur un exemple bi-dimensionnel.



**Figure 1.** Approximation tangente majorante  $H(\mathbf{x}, \mathbf{y})$  de  $F$  en  $\mathbf{y}$ . La fonction  $H$  est supérieure à  $F$  en tout point et est tangente à  $F$  en  $\mathbf{y}$ .

La minimisation du critère  $F$  par Majoration-Minoration consiste à résoudre le problème initial [9] par la règle de mise à jour MM

$$\mathbf{x}_{k+1} = \underset{\mathbf{x} \in \mathcal{D}_f}{\operatorname{argmin}} H(\mathbf{x}, \mathbf{x}_k). \quad [11]$$

De nombreuses contributions algorithmiques dans le cadre de l'inversion pénalisée en traitement du signal et de l'image s'identifient avec un schéma de minimisation MM. C'est le cas notamment des algorithmes EM (Dempster *et al.*, 1977) et proximaux Forward-Backward (Combettes *et al.*, 2005). La stratégie MM a été utilisée également comme principe constructif d'algorithmes, dans (Böhning *et al.*, 1988) ainsi que dans les algorithmes développés par Fessler *et al.* (Erdogan *et al.*, 1999; Fessler *et al.*, 1998; Sothivirat *et al.*, 2002) en reconstruction tomographique. Les propriétés de l'algorithme [11] sont détaillées dans (Jacobson *et al.*, 2007). En particulier, la récurrence MM [11] entraîne la décroissance de la suite  $\{F(\mathbf{x}_k)\}$ . Ceci est une conséquence directe des relations [10]. En effet,

$$F(\mathbf{x}_{k+1}) \underset{(1)}{\leq} H(\mathbf{x}_{k+1}, \mathbf{x}_k) \underset{(2)}{\leq} H(\mathbf{x}_k, \mathbf{x}_k) \underset{(3)}{=} F(\mathbf{x}_k), \quad [12]$$

où les relations (1) à (3) proviennent des propriétés de

- (1) Majoration
- (2) Minimisation
- (3) Tangence



### 2.2.2. Majoration-Minimisation quadratique

L'intérêt de la méthode MM est de transposer le problème d'optimisation [9] en une série de problèmes plus simples. Le choix d'une fonction  $H$  facile à minimiser est donc primordial. Afin d'assurer l'existence et l'unicité des itérés, il est naturel de choisir une fonction majorante strictement convexe. La Majoration-Minoration quadratique (MMQ) se base sur la construction d'approximations tangentes majorantes quadratiques du critère. Une approximation tangente majorante quadratique de  $F$  s'écrit

$$H(\mathbf{x}, \mathbf{y}) = F(\mathbf{y}) + \nabla F(\mathbf{y})^T (\mathbf{x} - \mathbf{y}) + \frac{1}{2} (\mathbf{x} - \mathbf{y})^T \mathbf{A}(\mathbf{y}) (\mathbf{x} - \mathbf{y}), \quad [13]$$

où  $\mathbf{A}(\mathbf{y}) \in \mathbb{R}^{n \times n}$  est une matrice définie positive assurant que [13] vérifie les conditions [10] pour tout  $\mathbf{y}$ . De façon générale, toute matrice  $\mathbf{A}(\mathbf{y})$  telle que  $\mathbf{A}(\mathbf{y}) - \nabla^2 F(\mathbf{y})$  soit définie positive entraîne que [13] est une approximation majorante de  $F$  (Hunter *et al.*, 2004). Cependant, ce résultat ne permet pas de définir une méthode de construction pour les matrices  $\mathbf{A}(\mathbf{y})$ . En pratique, ces matrices sont construites à partir de propriétés sur la structure analytique du critère. Deux techniques de constructions d'approximations majorantes quadratiques se distinguent. La stratégie de Geman & Yang (GY) (Geman *et al.*, 1995) utilise la propriété de Lipschitz-différentiabilité du critère à majorer. La stratégie de Geman & Reynolds (GR) (Geman *et al.*, 1992) est basée sur la propriété de concavité de la composition du critère à minimiser avec la fonction racine carrée. Les constructions GY et GR sont à la base des algorithmes semi-quadratiques (SQ) utilisés en reconstruction d'image (Charbonnier *et al.*, 1997; Idier, 2001). D'autres techniques de construction sont présentées dans (Erdogan *et al.*, 1999; Hunter *et al.*, 2004; Zheng *et al.*, 2004; Lange *et al.*, 1995; De Pierro, 1995).

Lorsque l'approximation majorante est quadratique, son minimiseur s'obtient de façon analytique. Il en résulte le schéma de minimisation très simple, que nous notons MMQ :

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{A}(\mathbf{x}_k)^{-1} \nabla F(\mathbf{x}_k). \quad [14]$$

Lorsque le problème est de grande taille, l'inversion de  $\mathbf{A}(\mathbf{x}_k)$  peut être coûteuse. Pour pallier ce problème, la matrice  $\mathbf{A}(\mathbf{x}_k)^{-1}$  peut être approchée à l'aide d'itérations de gradient conjugué linéaire (Labat *et al.*, 2007). Une autre possibilité est d'utiliser un schéma de minimisation par blocs (Erdogan *et al.*, 1999; Sotthivirat *et al.*, 2002).

### 2.2.3. Recherche de pas par Majoration-Minimisation

Dans le cadre de l'optimisation différentiable, une famille fondamentale d'algorithmes est basée sur la stratégie de descente itérative

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k \quad [15]$$

alternant calcul d'une direction de descente et détermination d'un pas scalaire quantifiant l'avancée le long de la direction choisie. Une stratégie pertinente de calcul du

pas doit tirer profit de la structure du critère pénalisé, ce qui est le cas des algorithmes de pas par Majoration-Minoration (MM) qui font l'objet de ce paragraphe.

### 2.2.3.1. Recherche de pas MM quadratique

La fonction majorante quadratique multi-dimensionnelle [13] permet de construire la fonction scalaire  $h_k(\alpha, \alpha_k^j)$ , approximation majorante quadratique de  $f_k(\alpha) = F(\mathbf{x}_k + \alpha \mathbf{d}_k)$  en un pas  $\alpha_k^j$

$$h_k(\alpha, \alpha_k^j) = f_k(\alpha_k^j) + (\alpha - \alpha_k^j) \dot{f}_k(\alpha_k^j) + \frac{1}{2} m_k^j (\alpha - \alpha_k^j)^2 \quad [16]$$

avec

$$m_k^j = \mathbf{d}_k^T \mathbf{A}_k^j \mathbf{d}_k. \quad [17]$$

où

$$\mathbf{A}_k^j = \mathbf{A}(\mathbf{x}_k + \alpha_k^j \mathbf{d}_k) \quad [18]$$

est la courbure de la quadratique  $H(\mathbf{x}, \mathbf{x}_k + \alpha_k^j \mathbf{d}_k)$ . Ce constat amène naturellement à utiliser comme recherche de pas un algorithme MMQ scalaire. Cette approche, proposée dans (Fessler *et al.*, 1999; Rivera *et al.*, 2003; Labat *et al.*, 2008), a pour intérêt de fournir de manière très simple une suite  $\alpha_k^1, \dots, \alpha_k^j$  qui converge vers un minimum local de  $f_k$  (et donc vers le minimiseur global lorsque  $F$  est convexe). La recherche de pas MMQ est une récurrence MM appliquée aux fonctions majorantes  $h_k(\cdot, \alpha_k^j)$ . La récurrence est initialisée avec  $\alpha_k^0 = 0$  et s'écrit :

$$\begin{cases} \alpha_k^0 & = 0, \\ \alpha_k^{j+1} & = \operatorname{argmin}_{\alpha} h_k(\alpha, \alpha_k^j), j < J, \\ \alpha_k & = \alpha_k^J. \end{cases} \quad [19]$$

Notons que le schéma itératif [19] généralise les stratégies de pas constant proposées dans (Bertsekas, 1999, Chap.1) et (Sun *et al.*, 2001; Chen *et al.*, 2002).

Une accélération significative de l'algorithme de descente itérative [15] est apportée en utilisant une stratégie de sous-espaces (Miele *et al.*, 1969; Conn *et al.*, 1994; Narkiss *et al.*, 2005; Yuan, 2007). Cette procédure consiste à explorer le critère non plus le long de directions de descente mais à l'intérieur de sous-espaces engendrés par plusieurs directions. Le problème de minimisation de  $F$  est résolu par le schéma itératif

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{D}_k \mathbf{s}_k \quad [20]$$

où  $\mathbf{D}_k = [\mathbf{d}_k^1, \dots, \mathbf{d}_k^r]$  est un ensemble de directions et  $\mathbf{s}_k$  est un pas multi-dimensionnel ayant pour objectif de minimiser  $f_k(\mathbf{s}) = F(\mathbf{x}_k + \mathbf{D}_k \mathbf{s})$ . Nous avons étendu dans (Chouzenoux *et al.*, 2011a) la recherche de pas MMQ scalaire en une recherche de pas MMQ multi-dimensionnelle pour les algorithmes de type sous-espace. Nous avons montré que cette nouvelle stratégie de pas (et par conséquent, son homologue scalaire) assure la convergence du schéma [20] sous des hypothèses faibles sur les directions  $\mathbf{d}_k^i$  (Chouzenoux *et al.*, 2011a; Chouzenoux *et al.*, 2011b).

### 2.2.3.2. Recherche de pas MM log-quadratique

L'élaboration des procédures de pas MMQ se basent sur l'hypothèse de l'existence de fonctions quadratiques majorantes du critère à minimiser. Cependant, certains critères pénalisés ne vérifient pas cette hypothèse. C'est le cas en particulier des critères contenant un terme de *barrière*, c'est à dire un terme à gradient non borné. Cette singularité se rencontre fréquemment dans le contexte du traitement du signal par exemple dans les critères de maximum d'entropie (Mohammad-Djafari *et al.*, 1988), de vraisemblance Poissonienne (Bertero *et al.*, 2009) et est également utilisée comme outil d'optimisation sous contrainte, par exemple dans les algorithmes de points intérieurs (Fiacco *et al.*, 1967).

Lorsqu'un algorithme de descente itérative est utilisé pour la minimisation d'un tel critère, la présence de la barrière rend inefficace les procédures standards de recherche de pas (Murray *et al.*, 1994; Wright, 1995). De plus, la forme singulière de la barrière rend impossible l'utilisation de la recherche de pas MM quadratique. Nous avons présenté dans (Chouzenoux *et al.*, 2009a; Chouzenoux *et al.*, 2010) une nouvelle stratégie de recherche de pas MM qui tient compte de la singularité de la barrière en se basant sur des approximations majorantes contenant un terme logarithmique. Le pas est toujours obtenu par une récurrence MM très simple, et bénéficie de propriétés de convergence fortes (Chouzenoux *et al.*, 2009a). Une application de cette stratégie à la résolution de problèmes de maximum d'entropie sera présentée en section 5.

### 2.2.4. Lien entre approche MM et l'algorithme proximal Forward-Backward

Considérons le problème de la minimisation de  $F(\mathbf{x}) = f_1(\mathbf{x}) + f_2(\mathbf{x})$ , où  $f_1$  est une fonction convexe et  $f_2$  est une fonction différentiable de gradient  $\beta$ -Lipschitz. Dans sa formulation exacte (i.e., sans ajout de termes d'erreur), l'algorithme FB s'écrit

$$\mathbf{x}_{k+1} = \lambda_k \hat{\mathbf{x}}_k^{\gamma_k} + (1 - \lambda_k) \mathbf{x}_k, \quad [21]$$

où

$$\hat{\mathbf{x}}_k^{\gamma_k} = \text{prox}_{\gamma_k f_1}(\mathbf{x}_k - \gamma_k \nabla f_2(\mathbf{x}_k)). \quad [22]$$

D'après la définition de l'opérateur proximal  $\text{prox}$ ,  $\hat{\mathbf{x}}_k^{\gamma_k}$  s'écrit également

$$\hat{\mathbf{x}}_k^{\gamma_k} = \underset{\mathbf{x}}{\text{argmin}} h_{\gamma_k}(\mathbf{x}, \mathbf{x}_k) \quad [23]$$

où  $h_{\gamma}(\mathbf{x}, \mathbf{x}_k)$  est définie, pour tout  $\gamma > 0$ , par

$$h_{\gamma}(\mathbf{x}, \mathbf{x}_k) = f_2(\mathbf{x}_k) + (\mathbf{x} - \mathbf{x}_k)^T \nabla f_2(\mathbf{x}_k) + \frac{1}{2\gamma} \|\mathbf{x} - \mathbf{x}_k\|^2 + f_1(\mathbf{x}). \quad [24]$$

A l'aide du formalisme des approches par Majoration-Minoration, nous établissons le résultat suivant.

**Propriété 2.2** *Si pour tout  $k$ ,  $\lambda_k \in ]0, 1]$  et  $\gamma_k \leq \frac{1}{\beta \lambda_k}$ , alors la suite  $(F(\mathbf{x}_k))$  générée par l'algorithme FB [21] est décroissante.*

**Démonstration 2.3** *D'après la définition de  $\hat{\mathbf{x}}_k^{\gamma_k}$ ,*

$$h_{\gamma_k}(\hat{\mathbf{x}}_k^{\gamma_k}, \mathbf{x}_k) \leq h_{\gamma_k}(\mathbf{x}_k, \mathbf{x}_k) = F(\mathbf{x}_k), \quad [25]$$

soit

$$(\hat{\mathbf{x}}_k^{\gamma_k} - \mathbf{x}_k)^T \nabla f_2(\mathbf{x}_k) + \frac{1}{2\gamma_k} \|\hat{\mathbf{x}}_k^{\gamma_k} - \mathbf{x}_k\|^2 + f_1(\hat{\mathbf{x}}_k^{\gamma_k}) - f_1(\mathbf{x}_k) \leq 0. \quad [26]$$

*D'autre part, la propriété de Lipschitz-différentiabilité de  $f_2$  entraîne que la fonction  $h_{\beta}(\mathbf{x}, \mathbf{x}_k)$  est une approximation tangente majorante de  $F(\mathbf{x}) : \forall \mathbf{x}$ ,*

$$F(\mathbf{x}) \leq h_{\beta^{-1}}(\mathbf{x}, \mathbf{x}_k) = f_2(\mathbf{x}_k) + (\mathbf{x} - \mathbf{x}_k)^T \nabla f_2(\mathbf{x}_k) + \frac{\beta}{2} \|\mathbf{x} - \mathbf{x}_k\|^2 + f_1(\mathbf{x}). \quad [27]$$

*Pour établir la décroissance de  $(F(\mathbf{x}_k))$ , il suffit de montrer la négativité de*

$$\begin{aligned} & h_{\beta^{-1}}(\mathbf{x}_{k+1}, \mathbf{x}_k) - F(\mathbf{x}_k) \\ &= \lambda_k \left( (\hat{\mathbf{x}}_k^{\gamma_k} - \mathbf{x}_k)^T \nabla f_2(\mathbf{x}_k) + \frac{\lambda_k \beta}{2} \|\hat{\mathbf{x}}_k^{\gamma_k} - \mathbf{x}_k\|^2 \right. \\ & \quad \left. + \frac{1}{\lambda_k} (f_1(\lambda_k \hat{\mathbf{x}}_k^{\gamma_k} + (1 - \lambda_k) \mathbf{x}_k) - f_1(\mathbf{x}_k)) \right). \end{aligned} \quad [28]$$

*D'après [24] et [28]*

$$\begin{aligned} & \frac{1}{\lambda_k} (h_{\beta^{-1}}(\mathbf{x}_{k+1}, \mathbf{x}_k) - F(\mathbf{x}_k)) - h_{\gamma_k}(\hat{\mathbf{x}}_k^{\gamma_k}, \mathbf{x}_k) - F(\mathbf{x}_k) \\ &= \left( \frac{\lambda_k \beta}{2} - \frac{1}{2\gamma_k} \right) \|\hat{\mathbf{x}}_k^{\gamma_k} - \mathbf{x}_k\|^2 + f_1(\mathbf{x}_k) - f_1(\hat{\mathbf{x}}_k^{\gamma_k}) \\ & \quad + \frac{1}{\lambda_k} (f_1(\lambda_k \hat{\mathbf{x}}_k^{\gamma_k} + (1 - \lambda_k) \mathbf{x}_k) - f_1(\mathbf{x}_k)) \end{aligned} \quad [29]$$

*La condition  $\lambda_k \in ]0, 1]$  et la convexité de  $f_1$  impliquent*

$$f_1(\lambda_k \hat{\mathbf{x}}_k^{\gamma_k} + (1 - \lambda_k) \mathbf{x}_k) - f_1(\mathbf{x}_k) \leq \lambda_k f_1(\hat{\mathbf{x}}_k^{\gamma_k}) + (1 - \lambda_k) f_1(\mathbf{x}_k). \quad [30]$$

*Donc*

$$\begin{aligned} & \frac{1}{\lambda_k} (h_{\beta^{-1}}(\mathbf{x}_{k+1}, \mathbf{x}_k) - F(\mathbf{x}_k)) - h_{\gamma_k}(\hat{\mathbf{x}}_k^{\gamma_k}, \mathbf{x}_k) - F(\mathbf{x}_k) \\ & \leq \left( \frac{\lambda_k}{2\beta} - \frac{1}{2\gamma_k} \right) \|\hat{\mathbf{x}}_k^{\gamma_k} - \mathbf{x}_k\|^2 + f_1(\mathbf{x}_k) - f_1(\hat{\mathbf{x}}_k^{\gamma_k}) \\ & \quad + \frac{1}{\lambda_k} (\lambda_k f_1(\hat{\mathbf{x}}_k^{\gamma_k}) + (1 - \lambda_k) f_1(\mathbf{x}_k) - f_1(\mathbf{x}_k)) \end{aligned} \quad [31]$$

$$= \left( \frac{\lambda_k \beta}{2} - \frac{1}{2\gamma_k} \right) \|\hat{\mathbf{x}}_k^{\gamma_k} - \mathbf{x}_k\|^2 \quad [32]$$

La condition  $\gamma_k \leq \frac{1}{\lambda_k \beta}$  entraîne la négativité du facteur  $\left(\frac{\lambda_k \beta}{2} - \frac{1}{2\gamma_k}\right)$ . Par conséquent,

$$h_{\beta^{-1}}(\mathbf{x}_{k+1}, \mathbf{x}_k) - F(\mathbf{x}_k) \leq \lambda_k (h_{\gamma}(\hat{\mathbf{x}}_k^{\gamma_k}, \mathbf{x}_k) - F(\mathbf{x}_k)) \leq 0, \quad [33]$$

soit  $h_{\beta^{-1}}(\mathbf{x}_{k+1}, \mathbf{x}_k) \leq F(\mathbf{x}_k)$ .

Lorsque  $\lambda_k = 1$ , l'algorithme FB s'écrit sous la forme [11] avec  $H(\mathbf{x}, \mathbf{x}_k) = h_{\gamma_k}(\mathbf{x}, \mathbf{x}_k)$ . D'après la propriété 2.2, la décroissance de  $(F(\mathbf{x}_k))$  est assurée si  $\gamma_k \leq \beta^{-1}$ , ce qui est simplement la condition nécessaire et suffisante pour que  $h_{\gamma_k}(\mathbf{x}, \mathbf{x}_k)$  soit une fonction majorante de  $F(\mathbf{x})$ . Ce lien entre algorithme FB non relaxé et algorithme MM est souligné dans (Beck *et al.*, 2009).

### 3. Reconstruction d'images en tomographie par émission de positrons

#### 3.1. Principe de la TEP

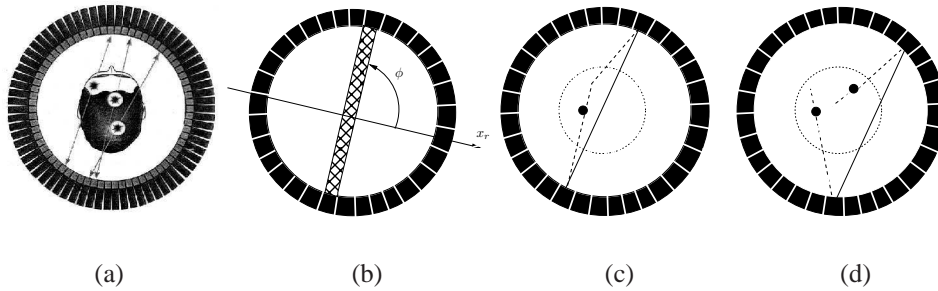
La TEP permet de mesurer l'activité métabolique d'un organe par le biais d'un traceur radioactif émetteur de positrons. Un exemple de traceur est le Fluorodesoxyglucose marqué au Fluor 18 ( $^{18}\text{F}$ -FDG). Il se caractérise par un marqueur  $^{18}\text{F}$  incorporé dans un analogue du glucose. Ce traceur se fixe sur les tissus ayant un métabolisme de glucose élevé, comme les tissus cancéreux. Dans la matière, les positrons émis par le traceur vont perdre leur énergie cinétique et, au repos, chacun des positrons va s'annihiler avec un électron. La masse de ces deux particules est convertie en énergie sous forme de deux photons gamma. Les deux photons issus de l'annihilation d'un positron avec un électron sont émis à  $180^\circ$  l'un de l'autre. Voyageant à la vitesse de la lumière, ils vont être détectés par une couronne de détecteurs, située autour du patient, presque simultanément ( $\pm$  quelques nanosecondes). Le parcours des photons est illustré sur la figure 2(a) pour trois annihilations distinctes.

On considère que deux photons proviennent d'une même annihilation s'ils sont détectés pendant un intervalle de temps de l'ordre de quelques nanosecondes, appelé fenêtre de coïncidences. Les deux détecteurs forment un tube de réponse (TOR) illustré figure 2(b) qui permet une localisation de la zone d'annihilation quelque part à l'intérieur du tube. À ce stade, on peut remarquer que pour une injection de  $2.10^8$  Bq (*i.e.* désintégrations par seconde), des centaines de milliards de désintégrations se produisent dans tout l'organisme mais seulement quelques centaines de millions sont détectées. En particulier, des photons émis dans la direction opposée du plan de coupe de la figure 2(a) ne seront pas détectés. En réalité, des dizaines d'anneaux (couronnes de détecteurs) sont empilées. Les données sont enregistrées sous forme de sinogramme. Il s'agit d'un histogramme représentant le nombre de détectations pour chaque tube de réponse (angle azimutal  $\phi$  + position radiale  $x_r$ ).

Lorsque les photons traversent la matière, des perturbations interviennent (Valk *et al.*, 2004). Il s'agit de phénomènes distincts :

- l'atténuation apparaît lorsque les photons traversent la matière et perdent de l'énergie. Les photons peuvent être absorbés et ne jamais sortir de l'organisme.
- Les coïncidences diffusées, illustrées figure 2(c), apparaissent lorsque le passage par la matière modifie le parcours des photons. Leurs trajets n'étant plus rectilignes, un biais peut être introduit et le lieu de l'annihilation ne se situe plus le long du TOR.
- Les coïncidences multiples se produisent dans le cas de figure où plus de deux photons sont détectés dans une fenêtre de coïncidences.
- Les coïncidences fortuites, illustrées figure 2(d), se produisent lorsque les photons sont détectés dans la même fenêtre de coïncidences mais proviennent de deux annihilations différentes.
- l'incertitude engendrée par acolinéarité. Nous avons mentionné qu'en théorie les photons partent dans des directions opposées lorsque l'annihilation se produit à stabilité du positron et de l'électron. En pratique, la cinétique restante du positron modifie les directions prises par les photons, l'angle formé est alors légèrement inférieur à  $180^\circ$ . Cet effet peut être modélisé par une convolution avec une gaussienne.

Les doses d'activité injectées sont très faibles, ce qui se traduit par un faible nombre d'évènements détectés et donc des données très bruitées.



**Figure 2.** (a) Principe de détection des coïncidences, (b) tube de réponse, (c) coïncidences diffusées et (d) coïncidences fortuites.

### 3.1.1. Modèle de dégradation spatio-temporel

Sous l'hypothèse que le nombre d'évènements suit une loi de Poisson et que le processus de détection est poissonien, nous pouvons considérer le modèle direct de dégradation suivant :

$$(\forall t \in \{1, \dots, T\}) \quad z_t = \mathcal{P}(A\bar{y}_t + c_t + d_t), \quad [34]$$

où l'effet du bruit de Poisson est représenté par  $\mathcal{P}$ .  $\bar{y}_t = (\bar{y}_{i,t})_{1 \leq i \leq N}$  représente une paramétrisation finie, composée de  $N$  éléments nommés « voxels », de la distribution spatiale de la concentration radioactive pour une trame temporelle  $t$ . Dans la suite nous nous référerons à cette distribution en parlant d'activité spatiale dans les tissus. Son unité de mesure est le Bq/cc. La paramétrisation finie  $z_t = (z_{j,t})_{1 \leq j \leq M}$

modélise les données TEP dynamiques et correspond au nombre de coïncidences détectées par chaque tube de réponse durant une trame-temporelle  $t$ . Notons que dans ce modèle,  $(z_t)_{t \in \{1, \dots, T\}}$  désigne un ensemble de sinogrammes, chacun composé de  $M$  éléments, nommés « bins », ( $M = \text{nb. d'angles azimutaux} \times \text{nb. de positions radiales}$  en 2D et  $M = \text{nb. d'angles azimutaux} \times \text{nb. de positions radiales} \times (\text{nb. d'anneaux})^2$  en 3D). L'effet des coïncidences diffusées et fortuites est modélisé par des termes additifs notés, respectivement,  $(d_t)_{t \in \{1, \dots, T\}}$  et  $(c_t)_{t \in \{1, \dots, T\}}$ . Notons que ces effets seront uniquement pris en compte lors du traitement de données réelles. La matrice  $A = (A_{j,i})_{1 \leq j \leq M, 1 \leq i \leq N}$  désigne la matrice système ; elle correspond à l'implantation numérique de l'opération de projection. Chaque élément  $A_{j,i}$  représente la probabilité qu'une coïncidence produite au niveau du voxel  $i$  soit détectée par le tube de réponse  $j$ . D'autres calculs envisageables sont, par exemple, basés sur la ligne d'intersection ou une mesure d'angle solide (Valk *et al.*, 2004). Dans le cas de données réelles, la matrice  $A$  inclut les facteurs d'atténuation des photons dans la matière.

Dans ce contexte, l'objectif est de trouver une paramétrisation finie  $(\hat{y}_t)_{t \in \{1, \dots, T\}}$  au plus proche de la paramétrisation originale  $(\bar{y}_t)_{t \in \{1, \dots, T\}}$ , en utilisant l'information contenue dans l'ensemble des sinogrammes temporels  $(z_t)_{t \in \{1, \dots, T\}}$ .

### 3.1.2. Extraction de paramètres physiologiques

Parmi les enjeux liés à la reconstruction dynamique TEP, on peut citer l'extraction de paramètres physiologiques. Pour cela il est nécessaire de connaître le modèle compartimental faisant le lien entre la mesure dans les tissus  $(\hat{y}_t)_{t \in \{1, \dots, T\}}$  et les paramètres physiologiques. Le modèle existant dans le cas d'un examen [ $^{18}\text{F}$ ]-FDG en imagerie cérébrale est détaillé dans (Valk *et al.*, 2004; Gjedde, 1982). Nous nous contenterons à ce stade de remarquer que l'intérêt du médecin porte principalement sur l'extraction des valeurs des constantes caractéristiques de chaque compartiment. Ces dernières ne peuvent pas être connues explicitement mais peuvent dépendre du métabolisme de glucose, noté  $\text{rCMRglu}$  qui se déduit de l'activité mesurée  $(\hat{y}_t)_{t \in \{1, \dots, T\}}$  par l'approche PATLAK (Gjedde, 1982; Patlak *et al.*, 1983). Par exemple, la constante  $K_1$  qui traduit la qualité du transfert des molécules à travers la BBB (Blood Brain Barrier) dépend linéairement du  $\text{rCMRglu}$ . Par conséquent une valeur de  $\text{rCMRglu}$  faible traduira un mauvais transfert des molécules à travers la BBB. L'enjeu de cette étude consiste à se rapprocher au plus de l'activité mesurée dans les tissus à un instant  $t$  pour ensuite obtenir une valeur de  $\text{rCMRglu}$  la plus précise possible tant au niveau de la valeur que pour la finesse de la résolution.

Étant donné le niveau de bruit important au niveau du voxel en TEP dynamique, on calcule généralement la valeur de  $\text{rCMRglu}$  en la moyennant sur un organe. L'objectif recherché consiste alors à améliorer le rapport signal sur bruit à l'échelle du voxel de façon à déterminer le paramètre d'intérêt au niveau du voxel et non de l'organe. C'est ce qu'on appelle l'imagerie paramétrique.

### 3.2. Etat de l'art sur les méthodes de reconstruction

La première application des algorithmes proximaux en TEP dynamique a été suggérée par Verhaeghe *et al.* (Verhaeghe *et al.*, 2008) en considérant les itérations de l'algorithme explicite-implicite (FB) (Daubechies *et al.*, 2004) pour minimiser un critère composé d'un terme de fidélité quadratique et d'une régularisation par des ondelettes splines. Dans (Sureau *et al.*, 2008), les auteurs effectuent le débruitage des courbes d'activité temporelle (CAT) comme pré-traitement à la reconstruction. L'étape de débruitage est basée sur l'algorithme de Douglas-Rachford (DR) qui permet, rappelons-le, contrairement au FB, de minimiser un critère incluant une fonction n'étant pas de gradient Lipschitz, telle qu'une divergence de Kullback-Leibler généralisée. Les algorithmes FB et DR ne sont pas directement applicables pour effectuer simultanément les tâches de débruitage et de reconstruction en utilisant un critère non-différentiable, une régularisation basée sur des trames d'ondelettes, et un terme de fidélité aux données adapté au bruit de Poisson. Dans (Chaux *et al.*, 2009; Pustelnik *et al.*, 2009), nous avons proposé une classe d'algorithmes imbriqués permettant de trouver la solution de tels critères. Le fait que la divergence de Kullback-Leibler généralisée ne soit pas de gradient Lipschitz requiert l'utilisation de l'extension quadratique. Cette méthode permet d'améliorer la qualité de reconstruction mais elle possède deux principaux inconvénients : un temps de calcul important dû à la boucle interne et la présence d'artefacts introduits par la régularisation ondelettes.

Pour contourner les problèmes rencontrés en combinant la classe des algorithmes imbriqués et l'extension quadratique, nous proposons d'utiliser l'algorithme proximal parallèle (PPXA) (Combettes *et al.*, 2008). Nous verrons que cette approche est particulièrement bien adaptée à l'imagerie TEP dynamique grâce à sa structure parallèle et à la possibilité de gérer une régularisation hybride telle que la combinaison d'une pénalisation par trames d'ondelettes et une mesure de variation totale (Rudin *et al.*, 1992). Cette méthode sera dans un premier temps évaluée sur des données simulées puis elle sera appliquée sur des données réelles.

### 3.3. Formulation du critère et choix de de l'algorithme de minimisation

Dans un cadre stochastique, nous considérons d'une part que l'activité spatio-temporelle dans les tissus  $y = (y_{i,t})_{(i,t) \in N_T}$ , où  $N_T = \{1, \dots, N\} \times \{1, \dots, T\}$ , est une réalisation d'un vecteur aléatoire à valeurs réelles  $Y = (Y_{i,t})_{(i,t) \in N_T}$ . Nous supposons que le vecteur aléatoire  $Y$  possède des composantes indépendantes qui ont une loi de probabilité notée  $\mu_{Y_{i,t}}(y_{i,t})$  et que ces fonctions sont log-concaves. D'autre part, nous supposons que l'activité projetée dans le domaine sinogramme, notée,

$$u = \left( u_{j,t} = \sum_{i=1}^N A_{ji} y_{i,t} \right)_{(j,t) \in M_T} \quad \text{avec} \quad M_T = \{1, \dots, M\} \times \{1, \dots, T\}$$

et les observations  $z = (z_{j,t})_{(j,t) \in M_T}$  sont respectivement des réalisations des vecteurs aléatoires  $U = (U_{j,t})_{(j,t) \in M_T}$  et  $Z = (Z_{j,t})_{(j,t) \in M_T}$ . Le vecteur aléatoire  $Z$



est supposé avoir des composantes indépendantes discrètes de probabilité conditionnelle  $\mu_{Z_{j,t}|U_{j,t}=u_{j,t}}(z_{j,t})$  qui correspond à la vraisemblance poissonnienne. Par conséquent, l'activité estimée  $\hat{y}$  par l'approche MAP s'exprime par :

$$\hat{y} \in \underset{y \in \mathbb{R}^{NT}}{\text{Argmax}} \prod_{t=1}^T \prod_{j=1}^M \mu_{Z_{j,t}|U_{j,t}=\sum_{i=1}^N A_{ji}y_{i,t}}(z_{j,t}) \prod_{t=1}^T \prod_{i=1}^N \mu_{Y_{i,t}}(y_{i,t}). \quad [35]$$

Nous précisons ci-dessous la définition de la vraisemblance poissonnienne pour des données spatio-temporelles,

$$\mu_{Z_{j,t}|U_{j,t}=u_{j,t}}(z_{j,t}) = \frac{(u_{j,t})^{z_{j,t}}}{z_{j,t}!} \exp(-u_{j,t}).$$

La reformulation du critère [35] par passage à l'anti-logarithme, nous amène au problème d'optimisation convexe qui suit :

$$\hat{y} \in \underset{y \in \mathbb{R}^{NT}}{\text{Argmin}} \tilde{g}(y) + \tilde{f}(y) \quad [36]$$

lorsque  $\tilde{f}$  est une fonction convexe de  $\mathbb{R}^{NT}$  dans  $\mathbb{R}$ . La fonction  $\tilde{g}$  désigne le terme de fidélité aux données, il s'agit d'une mesure assurant la proximité de l'activité estimée par rapport aux observations. Cette fonctionnelle correspond à la divergence de Kullback-Leibler généralisée. La principale difficulté réside dans le choix de la fonction  $\tilde{f}$  qui désigne le terme de régularisation. L'intérêt de régulariser dans le domaine des trames d'ondelettes a été démontré dans de récents travaux (Verhaeghe *et al.*, 2008; Sureau *et al.*, 2008; Starck *et al.*, 2010), nous reformulons donc le problème de minimisation [36] dans le domaine des trames.

On désigne par  $F \in \mathbb{R}^{K \times NT}$  (tel que  $K \geq NT$ ) la matrice associée à l'opérateur d'analyse d'une trame d'ondelettes spatio-temporelles. Dans la suite, nous considérerons des trames ajustées telles que  $F^* \circ F = \nu \text{Id}$  avec  $\nu \in ]0, +\infty[$ . On désigne par  $x \in \mathbb{R}^K$  les coefficients associés à l'activité  $y \in \mathbb{R}^{NT}$ . Par conséquent l'activité originale mesurée dans les tissus  $\bar{y}$  peut s'écrire  $\bar{y} = F^* \bar{x} = (F_t^* \bar{x})_{1 \leq t \leq T}$  où, pour tout  $t \in \{1, \dots, T\}$ ,  $F_t^* \in \mathbb{R}^{N \times K}$  représente la restriction de l'opérateur de synthèse de trame aux données reconstruites à l'instant  $t$ . Le modèle direct [34] devient :

$$(\forall t \in \{1, \dots, T\}) \quad z_t = \mathcal{D}(AF_t^* \bar{x} + c_t + d_t), \quad [37]$$

où  $\bar{x}$  représente le vecteur inconnu des coefficients de trame d'ondelettes associés à l'activité originale dans les tissus. Dans ce contexte, le problème de minimisation [36] peut être reformulé comme :

$$\hat{x} \in \underset{x \in \mathbb{R}^K}{\text{Argmin}} g(x) + f(x) \quad [38]$$

où l'activité reconstruite vaut  $\hat{y} = F^* \hat{x}$ . Le terme de fidélité aux données est la fonction  $g$  telle que :

$$g = \sum_{t=1}^T \Psi_t \circ A \circ F_t^*$$

où, pour tout  $u_t = (u_{j,t})_{1 \leq j \leq M} \in \mathbb{R}^M$ ,

$$\Psi_t(u_t) = \sum_{j=1}^M \psi_{j,t}(u_{j,t}) \quad \text{avec} \quad \psi_{j,t}(u_{j,t}) = -\log \mu_{Z_{j,t}|U_{j,t}=u_{j,t}}(z_{j,t}). \quad [39]$$

Pour le terme de régularisation  $f$ , des études ont démontré l'avantage d'utiliser une norme  $\ell_1$  (Verhaeghe *et al.*, 2008). Plus récemment, des régularisations de type hybrides ( $\ell_1$  + variation totale) ont été proposées et ont démontrées leur intérêt (Pustelnik *et al.*, 2011; Afonso *et al.*, 2011a). En choisissant ce type de régularisation, la fonction  $f$  peut s'écrire de la forme :

$$(\forall x = (x_k)_{1 \leq k \leq K}) \quad f(x) = \sum_{k=1}^K \chi_k |x_k| + \vartheta \sum_{t=1}^T \text{tv}(c_t^* x) \quad [40]$$

avec, pour tout  $k \in \{1, \dots, K\}$ ,  $\chi_k > 0$  et  $\vartheta > 0$ . Remarquons à ce stade que les paramètres  $\chi_k$  peuvent avoir des valeurs différentes pour chaque coefficient de trame. Dans un souci de simplifier l'estimation de ces paramètres, nous nous contenterons de les optimiser pour chaque sous-bande. Nous ajoutons dans le critère une contrainte sur la dynamique des données. On note  $C$  un ensemble convexe fermé de  $\mathbb{R}^K$  tel que  $C^* = F^* C = [0, V_{\max}]^{NT}$  où  $V_{\max}$  dépend de la quantité de traceur injecté. Le problème [38] se reformule alors :

$$\hat{x} \in \underset{x \in \mathbb{R}^K}{\text{Argmin}} \sum_{t=1}^T \left( \Psi_t(A F_t^* x) + \vartheta \text{tv}(F_t^* x) \right) + \sum_{k=1}^K \chi_k |x_k| + \iota_C(x), \quad [41]$$

Pour trouver un minimiseur de [41], un algorithme bien adapté est l'algorithme PPXA (Combettes *et al.*, 2008). Les itérations de PPXA pour résoudre [41] sont présentées par l'algorithme 2.

Soit  $\gamma \in ]0, +\infty[$  et  $\lambda = 1$ .  
 Pour tout  $q \in \{1, \dots, 4\}$ , soit  $(\omega_q)_{1 \leq q \leq 4} \in ]0, 1]^4$  tel que  $\sum_{q=1}^4 \omega_q = 1$ .  
 Soient  $\ell = 0$ ,  $(u_{q,0})_{1 \leq q \leq 4} \in (\mathbb{R}^K)^4$  et  $x_0 = \sum_{q=1}^4 \omega_q u_{q,0}$ .  
 Pour  $\ell = 0, \dots$

Pour $t = 1, \dots, T$	$p_{1,t,\ell} = \text{prox}_{\frac{\gamma}{\omega_1} \Psi_t(A F_t^*)} u_{1,\ell}$
	$p_{2,t,\ell} = \text{prox}_{\frac{\gamma \vartheta}{\omega_2} \text{tv}(F_t^*)} u_{2,\ell}$
	$p_{3,\ell} = \left( \text{prox}_{\frac{\gamma \chi_k}{\omega_3}  \cdot } u_{3,\ell,k} \right)_{1 \leq k \leq K}$
	$p_{4,\ell} = P_C u_{4,\ell}$
	$p_\ell = \omega_1 p_{1,\ell} + \omega_2 p_{2,\ell} + \omega_3 p_{3,\ell} + \omega_4 p_{4,\ell}$
	Pour $q = 1, \dots, 4$ $u_{q,\ell+1} = u_{q,\ell} + \lambda (2 p_\ell - x_\ell - p_{q,\ell})$
	$x_{\ell+1} = x_\ell + \lambda (p_\ell - x_\ell)$

Algorithme 2: PPXA pour minimiser le critère [41]

Le calcul de  $P_C$  se base sur la proposition 2.1,

$$(\forall x \in \mathbb{R}^K) \quad P_C(x) = x + \frac{1}{\nu} F(P_{C^*}(F^*x) - F^*x) \quad [42]$$

L'opérateur proximal de la norme  $\ell_1$  est une étape de seuillage doux des coefficients de trame. Les deux façons de procéder au calcul de l'opérateur proximal associé à la variation totale sont (i) l'utilisation de l'approche itérative proposée dans (Chambolle, 2004) et (ii) la décomposition du terme de variation totale en une somme de fonctions dont on sait calculer l'opérateur proximal. On se reportera à (Combettes *et al.*, 2008) pour plus de détails sur cette dernière façon de procéder. Il en résulte que  $\text{tv}$  peut être décomposée en une somme de 4 fonctions quand des filtres de Roberts sont considérés. Dans la suite nous adopterons cette deuxième solution et la fonction  $\text{tv}$  sera écrite comme une somme de 4 fonctions  $(\text{tv}_s)_{1 \leq s \leq 4}$ . Le forme explicite de l'opérateur proximal de  $\text{tv}_s \circ F_t^*$  résulte ensuite de la proposition 2.1.

La principale difficulté de cet algorithme réside dans le calcul de l'opérateur proximal de  $\Psi_t \circ A$ . On se réfèrera à (Pustelnik *et al.*, 2011, Proposition 3.4) pour mieux comprendre la limitation de ce calcul dans le cas où  $A$  désigne l'opérateur associé à la matrice système. Il apparaît clairement que pour le  $A$  considéré dans cette étude, la propriété sur  $A^* \circ A$  n'est pas vérifiée. Nous adaptions alors une technique de décomposition de  $A$  pour résoudre ce problème.

Nous supposons que pour tout  $r \in \{1, \dots, R\}$ ,  $\mathbb{I}_r$  est une partition de  $\{1, \dots, M\}$  en ensembles non vides et nous désignons par  $A^{(r)} = (A_{j,i})_{j \in \mathbb{I}_r, 1 \leq i \leq N}$  la matrice constituée de lignes de  $A$  sans recouvrement et donc orthogonales entre elles. Contrairement au cadre de la déconvolution, il est difficile de représenter la matrice système en TEP. En pratique, ces sous-ensembles désignent les lignes de la matrice sys-

tème associées aux bins des TOR parallèles n'ayant pas de pixels communs. De plus, pour tout  $t \in \{1, \dots, T\}$ , nous pouvons écrire chaque  $\Psi_t$  comme une somme de  $R$  fonctions notées  $\Psi_t^{(r)}: \mathbb{R}^{\text{Card}\{\mathbb{I}_r\}} \rightarrow ]-\infty, +\infty]$ , qui, pour tout  $r \in \{1, \dots, R\}$ , sont telles que  $\Psi_t^{(r)} = \sum_{j \in \mathbb{I}_r} \Psi_{j,t}$  et qui opèrent sur les bins de sinogrammes appartenant à  $\mathbb{I}_r$ . Il en résulte que  $\Psi_t \circ A = \sum_{r=1}^R \Psi_t^{(r)} \circ A^{(r)}$ . Cette décomposition permet d'avoir une forme explicite pour le calcul de l'opérateur proximal associé à  $\Psi_t^{(r)} \circ A^{(r)}$ .

La méthode qui découle de l'ensemble des remarques précédentes est résumée par l'algorithme 3. La convergence de la suite  $(x_\ell)_{\ell \in \mathbb{N}}$  vers un minimiseur de [41] est assurée par la coercivité de  $f_1$  et par le fait que  $\left( \bigcap_{t=1}^T \left( \bigcap_{r=1}^R (A^{(r)})^{-1}(\text{rint dom } \Psi^{(r)}) \right) \right) \cap \text{rint } C^* \neq \emptyset$ . Cette dernière condition est vérifiée car  $]0, +\infty[ \subset \text{dom } \Psi^{(r)}$ ,  $C^* = [0, V_{\max}]^N$  et puisque pour tout  $r \in \{1, \dots, R\}$ ,  $A^{(r)}$  est supposée positive, à valeurs réelles et composée de lignes étant toutes nulles.

On choisit  $\gamma = 10$  et  $\lambda = 1$ .

Soit  $\omega_1 = \dots = \omega_R = \frac{1}{4R}$ ,  $\omega_{R+1} = \dots = \omega_{R+4} = \frac{1}{16}$  et  $\omega_{R+5} = \omega_{R+6} = \frac{1}{4}$ .

Soit  $\ell = 0$  et, pour tout  $q \in \{1, \dots, R+6\}$ ,  $u_{q,0} = F z$ .

Soit  $x_0 = \sum_{q=1}^{R+6} \omega_q u_{q,0}$ .

Pour  $\ell = 0, \dots$

Pour $t = 1, \dots, T$	
	Pour $r = 1, \dots, R$ $p_{r,t,\ell} = \text{prox}_{\frac{\gamma}{\omega_q} \Psi_t^{(r)} \circ A^{(r)} \circ F_t^*}(u_{1,\ell})$
	Pour $s = 1, \dots, 4$ $p_{R+s,t,\ell} = \text{prox}_{\frac{\gamma}{\omega_{R+s}} \text{tv}_s(F_t^* \cdot)} u_{R+s,\ell}$
	$p_{R+5,\ell} = \text{prox}_{\frac{\gamma}{\omega_{R+5}} \ \cdot\ _1} u_{R+5,\ell}$
	$p_{R+6,\ell} = P_C(u_{R+6,\ell})$
	$p_\ell = \sum_{q=1}^{R+6} \omega_q p_{q,\ell}$
	Pour $q = 1, \dots, R+6$ $u_{q,\ell+1} = u_{q,\ell} + \lambda (2 p_\ell - x_\ell - p_{q,\ell})$
	$x_{\ell+1} = x_\ell + \lambda (p_\ell - x_\ell)$

Algorithme 3: PPXA pour minimiser critère [41].

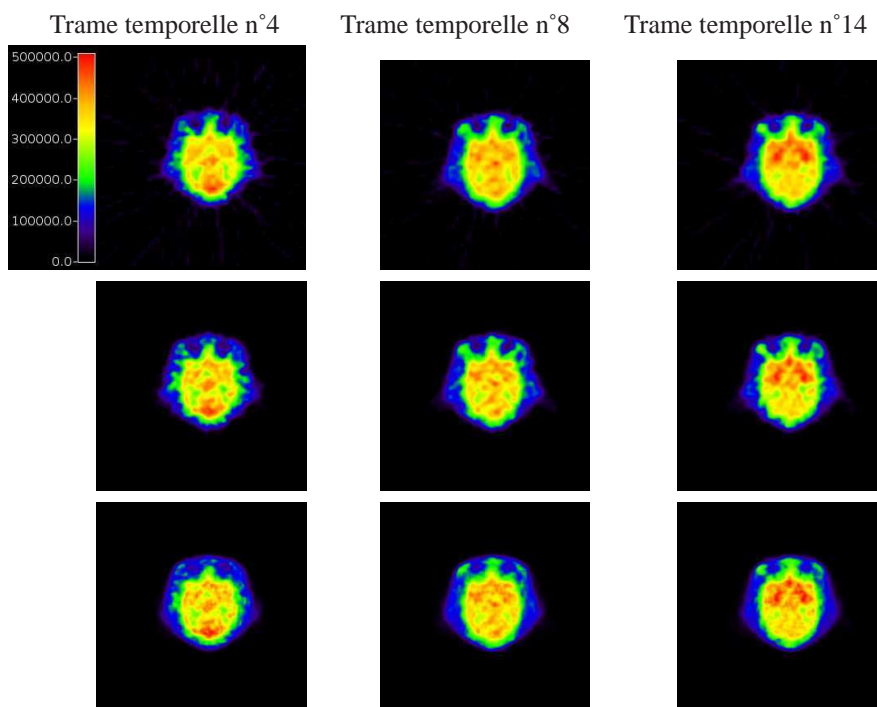
### 3.4. Résultats expérimentaux sur des données réelles

Nous avons traité des données provenant d'un examen [ $^{18}\text{F}$ ]-FDG, où une activité de 284 MBq a été injectée à un babouin de 4.5 kg. Une acquisition en mode (2D+t), avec les *septa*<sup>4</sup> sortis, a été réalisée. Les coïncidences vraies et retardées ont été ac-

4. Les septa permettent d'éliminer les photons faisant un angle avec le plan de coupe.

quises séparément. Nous avons acquis 128 trames d'une durée variant de 10 secondes à 30 secondes. La durée totale de l'examen est de 54 minutes. La transmission a été réalisée avant injection par une source rotative de germanium 57 ; cette procédure permettra de corriger les effets d'atténuation durant la reconstruction.

Nous avons procédé à la reconstruction de ces données réelles par rétroprojection filtrée (logiciel du constructeur à savoir SIEMENS), par la méthode SIEVES et par notre approche (PPXA). Pour chacune d'entre elles, les effets d'atténuations, de normalisation et les coïncidences fortuites sont pris en compte lors de la reconstruction. Cependant, nous ne tenons pas compte des coïncidences diffusées (ce choix est justifié par le fait qu'il s'agit d'une acquisition 2D+t). La matrice système utilisée dans les reconstructions SIEVES et PPXA modélise uniquement les effets géométriques<sup>5</sup>. Nous présentons sur la figure 3 les images obtenues par les différentes méthodes aux instants  $t = 4$ ,  $t = 8$  et  $t = 14$ .



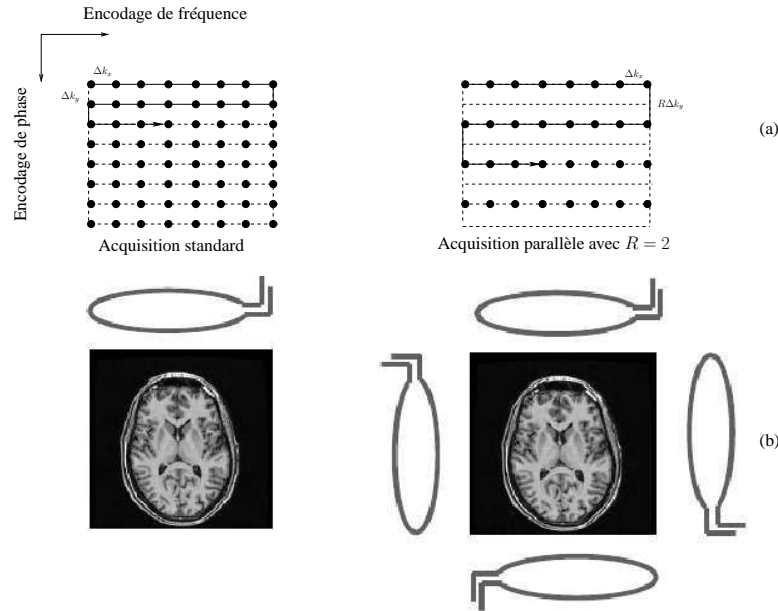
**Figure 3.** Comparaison des activités reconstruites (en Bq/cc) pour la coupe 24 de l'acquisition de données sur babouin. 1ère ligne : SIEMENS, 2ème ligne : SIEVES, 3ème ligne : PPXA.

5. Les effets dégradants liés à résolution des détecteurs ne sont pas considérés.

## 4. Reconstruction régularisée en IRM parallèle

### 4.1. L'IRM parallèle

L'imagerie par résonance magnétique parallèle (IRMp) est une technique d'imagerie rapide qui permet l'acquisition d'images de haute résolution spatiale et/ou temporelle. Elle repose sur le sous-échantillonnage de l'espace de Fourier ( $k$ -space) d'un facteur  $R$  appelé facteur de réduction, ainsi que l'utilisation de plusieurs antennes réceptrices ayant des profils de sensibilité spatiale différents, mais complémentaires. Ces  $L$  antennes réceptrices sont utilisées pour mesurer la densité des spins à l'intérieur de l'objet à imager, appelée  $\bar{\rho}$ <sup>6</sup>. Les images acquises par chaque antenne (images à champ de vue réduit) sont ensuite utilisées ensemble pour reconstruire des images à champ de vue complet. Le principe de sous-échantillonnage de l'espace de Fourier et de l'utilisation de plusieurs antennes réceptrices est illustré respectivement dans les figures 4.1[a] et 4.1[b].



**Figure 4.** Différence entre l'acquisition standard (gauche) et l'acquisition parallèle (droite) en IRM : (a) : échantillonnage du  $k$ -space ;(b) : configuration des antennes réceptrices ( $L=4$ ).

Concrètement, le signal  $\tilde{d}_\ell$  reçu par chaque antenne ( $1 \leq \ell \leq L$ ) n'est autre que la transformée de Fourier du champ 2D  $\bar{\rho} \in \mathbb{R}^{X \times Y}$  que l'on cherche à estimer, pondéré

6. La barre horizontale permet de distinguer la "vraie" donnée de la variable.

par le profil de sensibilité  $s_\ell$  et évalué au niveau de la position spatiale  $\mathbf{k}_r = (k_x, k_y)^\top$  du  $k$ -space<sup>7</sup> :

$$\tilde{d}_\ell(\mathbf{k}_r) = \int \bar{\rho}(\mathbf{r}) s_\ell(\mathbf{r}) e^{-i2\pi\mathbf{k}_r^\top \mathbf{r}} d\mathbf{r} + \tilde{n}_\ell(\mathbf{k}_r), \quad [43]$$

où  $\tilde{n}_\ell(\mathbf{k}_r)$  est un bruit gaussien centré indépendant et identiquement distribué (iid) dans le  $k$ -space, mais dépendant de l'antenne, et  $\mathbf{r} = (x, y)^\top \in X \times Y$  est la position spatiale dans le domaine image.

Pour reconstruire des images à champ de vue complet, deux grandes familles de méthodes existent : celles opérant directement dans l'espace de Fourier comme la méthode GRAPPA (Griswold *et al.*, 2002), et celles opérant dans le domaine spatial telles que la méthode SENSE (Pruessmann *et al.*, 1999). Dans la suite, nous allons nous concentrer sur cette dernière famille de méthodes qui permet de reconstruire une image à champ de vue complet à partir des images à champ de vue réduit acquises par chaque antenne, mais après application d'une transformée de Fourier inverse à chacune d'entre elles pour passer dans le domaine spatial.

Soit  $\Delta y = \frac{Y}{R}$  la période de recouvrement,  $y$  et  $x$  respectivement les positions dans le domaine spatial le long de la direction d'encodage de phase et de fréquence. En tenant compte du sous-échantillonnage du  $k$ -space d'un facteur  $R$ , la transformée de Fourier inverse permet d'obtenir la formulation spatiale de l'équation [43] sous une forme matricielle :

$$\mathbf{d}(\mathbf{r}) = \mathbf{S}(\mathbf{r})\bar{\rho}(\mathbf{r}) + \mathbf{n}(\mathbf{r}), \quad [44]$$

où

$$\mathbf{S}(\mathbf{r}) \triangleq \begin{bmatrix} s_1(x, y) & \dots & s_1(x, y + (R-1)\Delta y) \\ \vdots & \vdots & \vdots \\ s_L(x, y) & \dots & s_L(x, y + (R-1)\Delta y) \end{bmatrix}, \quad \mathbf{n}(\mathbf{r}) \triangleq \begin{bmatrix} n_1(x, y) \\ n_2(x, y) \\ \vdots \\ n_L(x, y) \end{bmatrix}$$

$$\bar{\rho}(\mathbf{r}) \triangleq \begin{bmatrix} \bar{\rho}(x, y) \\ \bar{\rho}(x, y + \Delta y) \\ \vdots \\ \bar{\rho}(x, y + (R-1)\Delta y) \end{bmatrix} \quad \text{et} \quad \mathbf{d}(\mathbf{r}) \triangleq \begin{bmatrix} d_1(x, y) \\ d_2(x, y) \\ \vdots \\ d_L(x, y) \end{bmatrix}.$$

En se basant sur ce modèle, l'étape de reconstruction consiste à résoudre le système [44] pour retrouver  $\bar{\rho}(\mathbf{r})$  à partir de  $\mathbf{d}(\mathbf{r})$  et une estimation préalable de  $\mathbf{S}(\mathbf{r})$  au niveau de chaque position  $\mathbf{r} = (x, y)^\top$ . Il est à noter que les images  $(d_\ell)_{1 \leq \ell \leq L}$ , ainsi que  $\bar{\rho}$  sont à valeurs complexes, mais seulement le module  $|\bar{\rho}|$  est utilisé pour la visualisation.

7.  $\cdot^\top$  est l'opérateur transposé

## 4.2. Régularisation spatio-temporelle

### 4.2.1. Reconstruction SENSE

Dans sa forme la plus simple, la reconstruction SENSE revient à résoudre un problème d'inversion uni-dimensionnel grâce à la séparabilité de la transformée de Fourier. L'algorithme de reconstruction SENSE minimise un critère des moindres carrés pondérés (MQP)  $\mathcal{J}_{\text{MQP}}$ , donné par :

$$\mathcal{J}_{\text{MQP}}(\rho) = \sum_{\mathbf{r} \in \{1, \dots, X\} \times \{1, \dots, Y/R\}} \| \mathbf{d}(\mathbf{r}) - \mathbf{S}(\mathbf{r})\rho(\mathbf{r}) \|_{\Psi^{-1}}^2, \quad [45]$$

où  $\| \cdot \|_{\Psi^{-1}} = \sqrt{(\cdot)^{\text{H}} \Psi^{-1} (\cdot)}$ . La matrice de covariance du bruit  $\Psi$  est généralement estimée au préalable à partir d'une acquisition de référence. Ainsi, le problème de reconstruction SENSE n'est autre qu'une estimation au sens du maximum de vraisemblance sous des hypothèses de bruit gaussien. Elle admet donc la forme suivante à chaque position spatiale  $\mathbf{r}$  :

$$\hat{\rho}_{\text{MQP}}(\mathbf{r}) = (\mathbf{S}^{\text{H}}(\mathbf{r})\Psi^{-1}\mathbf{S}(\mathbf{r}))^{\#} \mathbf{S}^{\text{H}}(\mathbf{r})\Psi^{-1}\mathbf{d}(\mathbf{r}), \quad [46]$$

avec  $(\cdot)^{\text{H}}$  (resp.  $(\cdot)^{\#}$ ) dénote le transposé complexe conjugué (resp. pseudo-inverse). Il faut noter que l'approche de reconstruction décrite ci-dessus a été conçue pour reconstruire une seule coupe (image 2D). Itérer sur les coupes est donc nécessaire pour reconstruire un volume de données 3D ou 4D.

En pratique, les performances de la méthode SENSE sont limitées par *i*) la présence de distorsions au niveau des mesures  $\mathbf{d}(\mathbf{r})$ , *ii*) le mauvais conditionnement de  $\mathbf{S}(\mathbf{r})$ , notamment au niveau du centre de l'image et *iii*) la présence d'erreurs d'estimation de  $\mathbf{S}(\mathbf{r})$  au niveau des interfaces cerveau/air. Pour atteindre une inversion plus robuste du problème inverse mal posé en question, une régularisation est généralement nécessaire. Dans ce contexte, des méthodes de régularisation non lisse ont été largement utilisées dans la littérature pour pallier les inconvénients de la régularisation classique (Tikhonov, 1963). à titre d'exemple, on peut citer la régularisation au sens de la variation totale et avec des représentations en ondelettes qui ont été introduites dans la littérature de reconstruction en IRMp respectivement par *Liu et al.* (*Liu et al.*, 2007) et *Chaari et al.* (*Chaari et al.*, 2008) en 2007 et 2008. Dans ce qui suit, nous allons nous focaliser sur la régularisation avec des représentations en ondelettes pour décrire cette régularisation spatio-temporelle pour la reconstruction en IRMp.

### 4.2.2. Reconstruction régularisée

Comme dans (*Chaari et al.*, 2011b) où des représentations en ondelettes séparables 2D sont utilisées pour la régularisation, toutes les méthodes de reconstruction de la littérature en IRMp opèrent coupe par coupe. L'inconvénient de cette stratégie réside dans le fait que la continuité spatiale est brisée entre les coupes adjacentes. D'un autre côté, une étude en IRM fonctionnelle (IRMf) nécessite l'acquisition du volume



cérébral 3D plusieurs fois dans le temps. Itérer sur les volumes 3D est donc nécessaire afin de reconstruire les volumes 4D ( $3D+t$ ) correspondant à une session IRMf. Les volumes 3D sont par conséquent considérés indépendants bien que les séries temporelles en IRMf soient corrélées dans le temps pour deux simples raisons : (i) le signal BOLD (*Blood Oxygenated Level-Dependent*) (Ogawa *et al.*, 1990) lui même est une version filtrée passe-bas du signal de l'activité neuronale ; (ii) les artefacts physiologiques rendent les points temporels du signal BOLD très dépendants. Pour ces raisons, modéliser la dépendance temporelle à travers les acquisitions temporelles au niveau de l'étape de reconstruction pourrait profondément influencer l'analyse statistique à mener. Nous nous sommes donc intéressés à l'extension de l'approche proposée dans (Chaari *et al.*, 2011b) afin de :

- tenir compte des dépendance spatiales 3D entre les coupes adjacentes en utilisant de représentations en ondelettes 3D,
- exploiter les dépendances temporelles entre les volumes 3D acquis en appliquant une régularisation supplémentaire le long de la dimension temporelle des données 4D.

Cette régularisation supplémentaire permettra d'améliorer le Rapport Signal sur Bruit (RSB) à travers les volumes acquis, et par conséquent la fiabilité de l'analyse statistique en IRMf.

La nouvelle reconstruction 4D régularisée des  $N_r$  volumes 3D acquis nécessite donc la réécriture du modèle d'observation dans [44] comme suit :

$$\mathbf{d}^t(\mathbf{r}) = \mathbf{S}(\mathbf{r})\boldsymbol{\rho}^t(\mathbf{r}) + \mathbf{n}^t(\mathbf{r}), \quad [47]$$

où  $t \in \{1, \dots, N_r\}$  représente le temps d'acquisition et  $\mathbf{r} = (x, y, z)$  la position spatiale 3D,  $z \in \{1, \dots, Z\}$  étant la position le long de la troisième direction (sélection de coupe). à un instant d'acquisition  $t$  donné, l'image à champ de vue complet et à valeurs complexes  $\bar{\rho}^t$  de taille  $X \times Y \times Z$  peut être vue comme un élément de l'espace euclidien  $\mathbb{C}^K$  avec  $K = X \times Y \times Z$  engendré par le produit scalaire standard  $\langle \cdot | \cdot \rangle$  et la norme  $\| \cdot \|$ . Nous utiliserons un opérateur de décomposition en ondelettes 3D dyadique  $F$  sur  $j_{\max}$  niveaux de résolutions.

Le champ de coefficients résultant de la décomposition de l'image à estimer  $\rho^t$  est défini comme  $\zeta^t = (\zeta_a^t, (\zeta_{o,j}^t)_{o \in \mathbb{O}, 1 \leq j \leq j_{\max}})$  avec  $o \in \mathbb{O} = \{0, 1\}^3 \setminus \{(0, 0, 0)\}$ ,  $\zeta_a^t = (\zeta_{a,k}^t)_{1 \leq k \leq K_{j_{\max}}}$  et  $\zeta_{o,j}^t = (\zeta_{o,j,k}^t)_{1 \leq k \leq K_j}$  où  $K_j = K2^{-3j}$  est le nombre de coefficients d'ondelette dans une sous-bande donnée au niveau de résolution  $j$  (en supposant que  $X, Y$  et  $Z$  sont multiple de  $2^{j_{\max}}$ ).

En adoptant cette notation, les coefficients d'ondelette ont été ré-indexés de telle sorte que  $\zeta_a^t$  désigne le vecteur des coefficients d'approximation au niveau de résolution  $j_{\max}$ , tandis que  $\zeta_{o,j}^t$  désigne le vecteur de coefficients à l'orientation  $o$  et au niveau de résolution  $j$ . L'utilisation d'une décomposition en ondelettes 3D permet aussi de lisser les artefacts de reconstruction le long de la direction de sélection de coupe, ce qui reste impossible avec les approches opérant coupe par coupe. L'approche de régularisation proposée repose donc sur l'utilisation de deux types de pénalisation :

la première décrit la connaissance spatiale *a priori* sur les coefficients d'ondelette de l'image 3D à estimer et s'écrit comme suit :

$$g(\zeta) = \sum_{t=1}^{N_r} \left[ \sum_{k=1}^{K_{j_{\max}}} \Phi_a(\zeta_{a,k}^t) + \sum_{o \in \mathbb{O}} \sum_{j=1}^{j_{\max}} \sum_{k=1}^{K_j} \Phi_{o,j}(\zeta_{o,j,k}^t) \right], \quad [48]$$

où  $\zeta = (\zeta^1, \zeta^2, \dots, \zeta^{N_r})$  et on a  $\forall o \in \mathbb{O}$  et  $j \in \{1, \dots, j_{\max}\}$ ,

$$\forall \xi \in \mathbb{C}, \quad \Phi_{o,j}(\xi) = \Phi_{o,j}^{\text{Re}}(\xi) + \Phi_{o,j}^{\text{Im}}(\xi) \quad [49]$$

où  $\Phi_{o,j}^{\text{Re}}(\xi) = \alpha_{o,j}^{\text{Re}} |\text{Re}(\xi - \mu_{o,j})| + \frac{\beta_{o,j}^{\text{Re}}}{2} |\text{Re}(\xi - \mu_{o,j})|^2$  et  $\Phi_{o,j}^{\text{Im}}(\xi) = \alpha_{o,j}^{\text{Im}} |\text{Im}(\xi - \mu_{o,j})| + \frac{\beta_{o,j}^{\text{Im}}}{2} |\text{Im}(\xi - \mu_{o,j})|^2$  avec  $\mu_{o,j} = \mu_{o,j}^{\text{Re}} + \imath \mu_{o,j}^{\text{Im}} \in \mathbb{C}$ , et  $\alpha_{o,j}^{\text{Re}}, \beta_{o,j}^{\text{Re}}, \alpha_{o,j}^{\text{Im}}, \beta_{o,j}^{\text{Im}}$  sont des constantes positives. Ci-dessus,  $\text{Re}(\cdot)$  et  $\text{Im}(\cdot)$  (ou bien  $\cdot^{\text{Re}}$  et  $\cdot^{\text{Im}}$ ) sont respectivement les parties réelle et imaginaire. Un modèle similaire est adopté pour les coefficients d'approximation.

Le deuxième terme de régularisation pénalise la variation temporelle entre les volumes 3D successifs :

$$h(\zeta) = \kappa \sum_{t=2}^{N_r} \|F^* \zeta^t - F^* \zeta^{t-1}\|_p^p \quad [50]$$

où  $F^*$  est l'opérateur de reconstruction en ondelettes 3D. Les paramètres  $\alpha_{o,j} = (\alpha_{o,j}^{\text{Re}}, \alpha_{o,j}^{\text{Im}})$ ,  $\beta_{o,j} = (\beta_{o,j}^{\text{Re}}, \beta_{o,j}^{\text{Im}})$ ,  $\mu_{o,j} = (\mu_{o,j}^{\text{Re}}, \mu_{o,j}^{\text{Im}})$ ,  $\kappa \in \mathbb{R}_+$  et  $p \in [1, +\infty[$  sont inconnus et doivent être estimés (Chaari *et al.*, 2011a).

Afin d'obtenir le volume 3D reconstruit  $\rho^t$  relatif au temps d'acquisition  $t$ , l'opérateur  $F^*$  est ensuite appliqué à chaque composante  $\zeta^t$  de  $\zeta$ .

#### 4.2.3. Procédure d'optimisation

En se basant sur la formulation de la section 4.2.2, le critère à optimiser peut être réécrit comme suit :

$$\mathcal{J}_{\text{ST}}(\zeta) = \mathcal{J}_{\text{MQPT}}(\zeta) + g(\zeta) + h(\zeta) \quad [51]$$

où  $\mathcal{J}_{\text{MQPT}}(\zeta) = \sum_{t=1}^{N_r} \mathcal{J}_{\text{MQP}}(F^* \zeta^t)$ . Pour minimiser  $\mathcal{J}_{\text{ST}}$ , nous allons recourir au concept des opérateurs proximaux (Moreau, 1965). Comme les données traitées ici sont à valeurs complexes, nous étendons la définition de ces opérateurs aux fonctions à variable complexe. Pour la fonction

$$\Phi: \mathbb{C}^K \rightarrow ]-\infty, +\infty]; \quad x \mapsto \phi^{\text{Re}}(\text{Re}(x)) + \phi^{\text{Im}}(\text{Im}(x)), \quad [52]$$

où  $\phi^{\text{Re}}$  et  $\phi^{\text{Im}}$  sont des fonctions dans  $\Gamma_0(\mathbb{R}^K)$  et  $\text{Re}(x)$  (resp.  $\text{Im}(x)$ ) est le vecteur des parties réelles (resp. imaginaires) des composantes de  $x \in \mathbb{C}^K$ , l'opérateur proximal est défini par :

$$\text{prox}_{\Phi}: \mathbb{C}^K \rightarrow \mathbb{C}^K; \quad x \mapsto \text{prox}_{\phi^{\text{Re}}}(\text{Re}(x)) + \imath \text{prox}_{\phi^{\text{Im}}}(\text{Im}(x)). \quad [53]$$

## 4.2.3.1. Opérateur proximal du terme d'attache aux données

Selon des règles standards de calcul des opérateurs proximaux (Combettes *et al.*, 2010, Table 1.1), l'opérateur proximal du terme d'attache aux données  $\mathcal{J}_{\text{MQP}}$  est donné pour chaque vecteur de coefficients  $\zeta^t$  (avec  $t \in \{1, \dots, N_r\}$ ) par  $\text{prox}_{\mathcal{J}_{\text{MQP}}}(\zeta^t) = F u^t$ , où l'image  $u^t$  est de telle sorte que  $\forall \mathbf{r} \in \{1, \dots, X\} \times \{1, \dots, Y/R\} \times \{1, \dots, Z\}$ ,

$$\mathbf{u}^t(\mathbf{r}) = (\mathbf{I}_R + 2\mathbf{S}^H(\mathbf{r})\mathbf{\Psi}^{-1}\mathbf{S}(\mathbf{r}))^{-1}(\boldsymbol{\rho}^t(\mathbf{r}) + 2\mathbf{S}^H(\mathbf{r})\mathbf{\Psi}^{-1}\mathbf{d}^t(\mathbf{r})), \quad [54]$$

où  $\boldsymbol{\rho}^t = F^* \zeta^t$ .

## 4.2.3.2. Opérateur proximal de la fonction de régularisation spatiale

Selon (Chaari *et al.*, 2011b), et pour chaque niveau de résolution  $j$  et orientation  $o$ , l'opérateur proximal de la fonction de régularisation spatiale  $\Phi_{o,j}$  est donné par :

$$\begin{aligned} \forall \xi \in \mathbb{C}, \quad \text{prox}_{\Phi_{o,j}} \xi = & \frac{\text{signe}(\text{Re}(\xi - \mu_{o,j}))}{\beta_{o,j}^{\text{Re}} + 1} \max\{|\text{Re}(\xi - \mu_{o,j})| - \alpha_{o,j}^{\text{Re}}, 0\} \\ & + i \frac{\text{sign}(\text{Im}(\xi - \mu_{o,j}))}{\beta_{o,j}^{\text{Im}} + 1} \max\{|\text{Im}(\xi - \mu_{o,j})| - \alpha_{o,j}^{\text{Im}}, 0\} + \mu_{o,j} \end{aligned} \quad [55]$$

où la fonction signe est définie comme  $\forall \xi \in \mathbb{R}$ ,  $\text{sign}(\xi) = +1$  si  $\xi \geq 0$  et  $-1$  sinon

## 4.2.3.3. Opérateur proximal de la fonction de régularisation temporelle

Comme  $h$  n'est pas séparable par rapport à la variable temporelle  $t$ , aucune expression simple de son opérateur proximal n'est disponible. Pour contourner ce problème et pouvoir utiliser l'algorithme PPXA, nous proposons une décomposition en temps d'acquisitions pairs/impairs. Ce terme de régularisation est ensuite décomposé en une somme de deux fonctions  $h_1$  et  $h_2$  plus simples à traiter :

$$h(\zeta) = h_1(\zeta) + h_2(\zeta) \quad [56]$$

où  $h_1(\zeta) = \kappa \sum_{t=1}^{N_r/2} \|F^* \zeta^{2t} - F^* \zeta^{2t-1}\|_p^p$  et  $h_2(\zeta) = \kappa \sum_{t=1}^{N_r/2-1} \|F^* \zeta^{2t+1} - F^* \zeta^{2t}\|_p^p$ .

Comme  $h_1$  (resp.  $h_2$ ) est séparable par rapport à la variable temporelle  $t$ , son opérateur proximal peut être facilement calculé. En effet, considérons la fonction suivante :

$$\Psi : \mathbb{C}^K \times \mathbb{C}^K \longrightarrow \mathbb{R}; (\zeta^t, \zeta^{t-1}) \mapsto \kappa \|F^* \zeta^t - F^* \zeta^{t-1}\|_p^p = \psi \circ H(\zeta^t, \zeta^{t-1}), \quad [57]$$

où  $\psi(\cdot) = \kappa \|F^* \cdot\|_p^p$  et  $H$  est l'opérateur linéaire défini comme

$$H : \mathbb{C}^K \times \mathbb{C}^K \longrightarrow \mathbb{C}^K; (a, b) \mapsto a - b. \quad [58]$$

L'opérateur adjoint associé  $H^*$  est donc donné par

$$H^* : \mathbb{C}^K \longrightarrow \mathbb{C}^K \times \mathbb{C}^K; a \mapsto (a, -a). \quad [59]$$

Comme  $HH^* = 2\text{Id}$ , l'opérateur proximal de  $\Psi$  peut être facilement calculé en utilisant (Combettes *et al.*, 2007a, Prop. 11) :

$$\text{prox}_\Psi = \text{prox}_{\psi \circ H} = \text{Id} + \frac{1}{2}H^* \circ (\text{prox}_{2\psi} - \text{Id}) \circ H. \quad [60]$$

Pour plus de détails sur le calcul de l'opérateur proximal de  $\text{prox}_{2\psi}$ , le lecteur peut se référer à (Chaux *et al.*, 2007).

#### 4.2.3.4. L'algorithme PPXA pour la reconstruction régularisée en IRMp

Le critère à minimiser peut être ré-écrit comme suit :

$$\mathcal{J}_{\text{ST}}(\zeta) = \mathcal{J}_{\text{MQPT}}(\zeta) + g(\zeta^t) + h_1(\zeta^t) + h_2(\zeta^t). \quad [61]$$

Comme  $\mathcal{J}_{\text{ST}}$  est composé de plus de deux fonctions convexes mais non-nécessairement différentiables, l'utilisation de l'algorithme PPXA représente une solution adéquate pour la minimisation de ce critère. Il est à noter que cet algorithme ne nécessite pas l'utilisation d'une sous-itération comme ce qui est proposé dans (Chaari *et al.*, 2011b). L'algorithme proposé est résumé dans Algorithme 4. Dans cet algorithme, les poids  $\omega_i$  ont été fixés à  $1/4$  pour chaque  $i \in \{1, \dots, 4\}$ . Le paramètre  $\gamma$  a été fixé à 200 puisque cette valeur semble donner la meilleure vitesse de convergence en pratique. L'algorithme ci-dessous converge généralement après 50 itérations en fixant le seuil d'arrêt  $\varepsilon$  à  $10^{-4}$ . La validation sur des données réelles de l'approche de reconstruction ainsi proposée est présentée dans la section 4.3.

Soit  $(\gamma, \varepsilon) \in ]0, +\infty[^2$ ,  $(\omega_i)_{1 \leq i \leq 4} \in ]0, 1[^4$  de telle sorte que  $\sum_{i=1}^4 \omega_i = 1$ ,  $(\zeta_i^{(n)})_{1 \leq i \leq 4} \in (\mathbb{C}^{K \times N_r})^4$  où  $\zeta_i^{(n)} = (\zeta_i^{1,(n)}, \zeta_i^{2,(n)}, \dots, \zeta_i^{N_r,(n)})$ ,  $n = 0$ , et  $\zeta_i^{t,(n)} = ((\zeta_{i,a}^{t,(n)}), ((\zeta_{i,o,j}^{t,(n)}))_{o \in \mathbb{O}, 1 \leq j \leq j_{\max}})$  pour chaque  $i \in \{1, \dots, 4\}$  et  $t \in \{1, \dots, N_r\}$ . Prendre aussi  $\zeta^{(n)} = \sum_{i=1}^4 \omega_i \zeta_i^{(n)}$  et  $\mathcal{J}^{(n)} = 0$ .

**Répéter**

**Prendre**  $p_4^{1,(n)} = \zeta_4^{1,(n)}$ .

**Pour**  $t = 1$  à  $N_r$  **faire**

**Calculer**  $p_1^{t,(n)} = \text{prox}_{\gamma \mathcal{J}_{\text{MQP}} / \omega_1}(\zeta_1^{t,(n)})$ .

**Calculer**  $p_2^{t,(n)} = (\text{prox}_{\gamma \Phi_a / \omega_2}(\zeta_{2,a}^{t,(n)}), (\text{prox}_{\gamma \Phi_{o,j} / \omega_2}(\zeta_{2,o,j}^{t,(n)}))_{o \in \mathbb{O}, 1 \leq j \leq j_{\max}})$ .

**Si**  $t$  **est pair** **alors**

**calculer**  $(p_3^{t,(n)}, p_3^{t-1,(n)}) = \text{prox}_{\gamma \Psi / \omega_3}(\zeta_3^{t,(n)}, \zeta_3^{t-1,(n)})$

**Si**  $t$  **est impair** **et**  $t > 1$  **alors**

**calculer**  $(p_4^{t,(n)}, p_4^{t-1,(n)}) = \text{prox}_{\gamma \Psi / \omega_4}(\zeta_4^{t,(n)}, \zeta_4^{t-1,(n)})$ .

**fin si**

**Si**  $t > 1$  **alors**

**Prendre**  $P^{t-1,(n)} = \sum_{i=1}^4 \omega_i p_i^{t-1,(n)}$ .

**fin si**

**fin pour**

**Prendre**  $p_4^{N_r,(n)} = \zeta_4^{N_r,(n)}$ .

**Prendre**  $P^{N_r,(n)} = \sum_{i=1}^4 \omega_i p_i^{N_r,(n)}$ .

**Prendre**  $P^{(n)} = (P^{1,(n)}, P^{2,(n)}, \dots, P^{N_r,(n)})$ .

**Prendre**  $\lambda_n \in [0, 2]$ .

**Pour**  $i = 1$  à  $4$  **faire**

**Prendre**  $p_i^{(n)} = (p_i^{1,(n)}, p_i^{2,(n)}, \dots, p_i^{N_r,(n)})$ .

$\zeta_i^{(n)} = \zeta_i^{(n)} + \lambda_n (2P^{(n)} - \zeta^{(n)} - p_i^{(n)})$ .

**fin pour**

$\zeta^{(n+1)} = \zeta^{(n)} + \lambda_n (P^{(n)} - \zeta^{(n)})$ .

$n \leftarrow n + 1$ .

**jusqu'à**  $|\mathcal{J}_{\text{ST}}(\zeta^{(n)}) - \mathcal{J}_{\text{ST}}(\zeta^{(n-1)})| \leq \varepsilon \mathcal{J}_{\text{ST}}(\zeta^{(n-1)})$ .

**Prendre**  $\hat{\zeta} = \zeta^{(n)}$ .

**Retourner**  $\hat{\rho}^t = F^* \hat{\zeta}^t$  **pour** chaque  $t \in \{1, \dots, N_r\}$ .

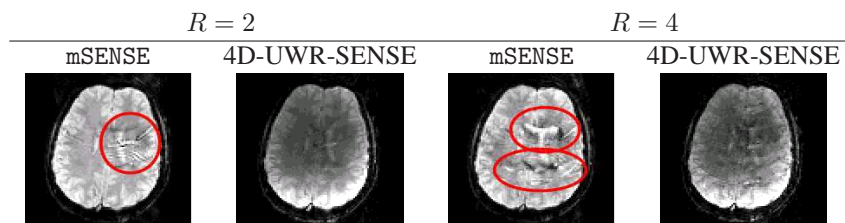
Algorithme 4: Algorithme PPXA pour l'optimisation du critère [61] : **4D-UWR-SENSE**

### 4.3. Validation sur des données réelles

Pour la validation, des données d'IRM fonctionnelle ont été acquises avec une séquence en écho de gradient EPI (*Gradient-Echo EPI*) (Temps d'Echo= 30 ms, Temps de Répétition= 2400 ms, épaisseur de coupe = 3 mm, orientation transver-

sale,  $FOV = 192 \text{ mm}^2$ ) sur un scanner *3 T Siemens Trio* et durant un protocole cognitif *localizer* (Pinel *et al.*, 2007). Cette expérience a été conçue pour cartographier les fonctions auditives, visuelles et motrices du cerveau, ainsi que des fonctions cognitives de plus haut niveau comme les fonctions de calcul et de compréhension du langage. Ces données correspondent à une seule session IRMf comprenant  $N_r = 128$  répétitions. L'antenne utilisée lors de l'acquisition contenait 32 canaux ayant des profils de sensibilité spatiale différents. Pour chaque sujet, les données IRMf ont été acquises à la résolution spatiale  $2 \times 2 \text{ mm}^2$  dans le plan en utilisant deux facteurs de réduction différents ( $R = 2$  or  $R = 4$ ).

Ces données ont ensuite été reconstruites en utilisant la méthode proposée, la méthode UWR-SENSE proposée dans (Chari *et al.*, 2011b), ainsi que l'algorithme mSENSE (implémentation SENSE de Siemens) qui est compatible avec les contraintes temps-réel. Cependant, à cause de la complexité opératoire, l'approche proposée est appliquée hors ligne pour reconstruire les données. à titre d'exemple, une coupe EPI est reconstruite en 4 s en utilisant l'algorithme UWR-SENSE sur une architecture Intel Xeon CPU@ 2.67GHz. En mettant en œuvre des techniques de calcul parallèle multi-processus (bibliothèque OMP), un volume EPI composé de 40 coupes est reconstruit en 22 s, ce qui rend toute la série temporelle des 128 volumes disponible après 47 min de calcul. En utilisant l'approche de régularisation spatio-temporelle proposée, 40 min suffisent pour reconstruire toute la série temporelle, mais elle nécessite plus d'espace mémoire puisque tous les volumes sont traités simultanément. Des coupes axiales sont illustrées dans la figure 5 afin de comparer l'approche proposée (4D-UWR-SENSE) et l'algorithme mSENSE. Cette figure montre que cette approche permet d'éliminer les artefacts de reconstruction affectant les images fournies par le scanner (mSENSE). Il est à noter que la différence d'homogénéité du signal est due à un filtre de post-traitement appliqué par le scanner aux images mSENSE.



**Figure 5.** Images reconstruites en utilisant les algorithmes mSENSE et 4D-UWR-SENSE pour  $R = 2$  et  $R = 4$ . Les ellipsoïdes en rouge indiquent les positions des artefacts de reconstruction obtenus avec l'algorithme mSENSE.

Après reconstruction, un Modèle Linéaire Général (MLG) (Friston *et al.*, 1994) a été utilisé pour détecter les réponses BOLD aux stimuli présentés. Nous nous sommes ensuite concentrés sur le contraste **Clic gauche vs. Clic droit** (Cg-Cd) pour lequel on s'attend à avoir des activations induites au niveau du cortex moteur. Pour les données acquises avec  $R = 2$ , la figure 6[haut] montre que toutes les méthodes de reconstruc-

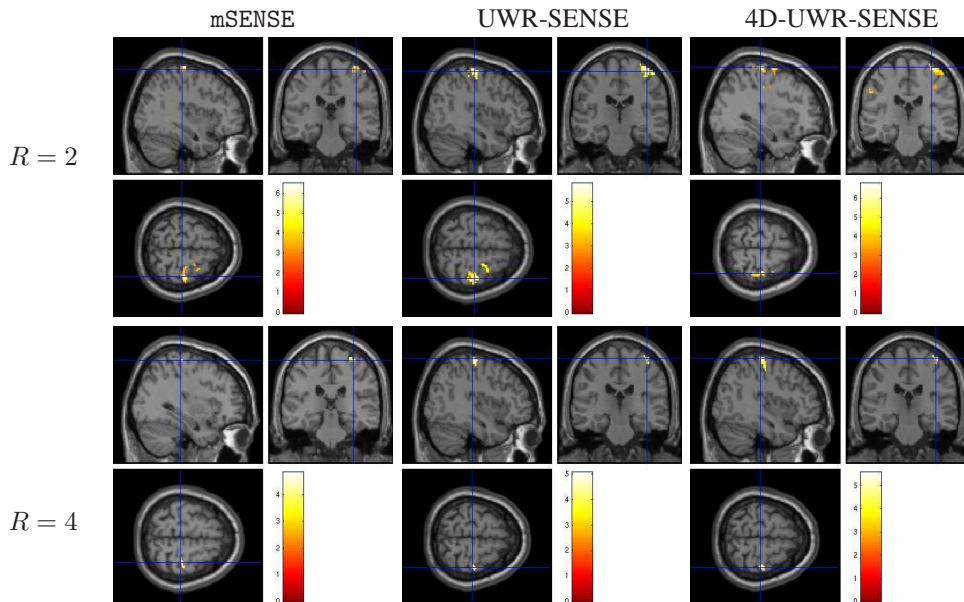
tion utilisées permettent de détecter les activations attendues dans le gyrus précentral droit. Cependant, les résultats statistiques illustrés dans le Tableau 1 montrent que l'approche de régularisation spatio-temporelle permet de détecter plus de *clusters* que les autres approches. Les résultats qualitatifs (figure 6[bas]) sur les données acquises avec  $R = 4$ , montrent que cette approche permet aussi de détecter plus d'activations.

D'un point de vue qualitatif, le Tableau 1 confirme la supériorité de l'approche 4D-UWR-SENSE proposée dans le sens que, aussi bien pour  $R = 2$  que  $R = 4$ , elle permet d'avoir des zones d'activation plus larges avec des scores statistiques plus élevés.

## 5. Reconstruction de spectres en relaxométrie RMN bidimensionnelle

### 5.1. Principe de la reconstruction RMN

La résonance magnétique nucléaire (RMN) est une modalité de spectroscopie permettant d'accéder à des informations structurales aux échelles microscopiques ou moléculaires de produits biologiques, en se basant sur des propriétés quantiques des noyaux (Canet *et al.*, 2002). Plus précisément, la RMN quantifie les phénomènes de relaxation qui interviennent au sein du spin des noyaux d'un échantillon lorsque celui-ci est soumis à un champ magnétique. Un échantillon biologique homogène est carac-



**Figure 6.** Cartes Student-*t* superposées aux images anatomiques pour le contraste *Cg-Cd*.

**Tableau 1.** Les résultats les plus significatifs pour le contraste Cg-Cd (corrigés pour les comparaisons multiples avec une p-valeur  $p = 0.05$ ).

		niveau cluster		niveau voxel		
		p-valeur	Taille	p-valeur	Score T	Position
$R = 2$	mSENSE	$< 10^{-3}$	79	$< 10^{-3}$	6.49	38 -26 66
	UWR-SENSE	$< 10^{-3}$	144	0.004	5.82	40 -22 63
		0.03	21	0.064	4.19	24 -8 63
	4D-UWR-SENSE	$< 10^{-3}$	<b>172</b>	0.001	<b>6.78</b>	34 -24 69
$< 10^{-3}$		79	0.001	6.49	38 -26 66	
$R = 4$	mSENSE	0.006	21	0.295	4.82	34 -28 63
	UWR-SENSE	$< 10^{-3}$	33	0.120	5.06	40 -24 66
	4D-UWR-SENSE	$< 10^{-3}$	<b>51</b>	0.006	<b>5.57</b>	40 -24 66

térisé par ses constantes de temps de relaxation longitudinale et transverse  $T_1$  et  $T_2$ . Ces paramètres physiques sont fondamentaux. En effet, leur détermination permet de préciser l'environnement électronique et donc chimique du noyau étudié. Leur connaissance permet des applications nombreuses : En médecine, les temps de relaxation sont à l'origine du contraste des images IRM. En agro-alimentaire, on peut utiliser le fait que, dans un système complexe, le temps de relaxation des protons de l'eau soit différent du temps de relaxation des protons de la matière grasse, afin d'en déterminer la teneur en eau et en matière grasse (Mariette, 2004).

Les composants hétérogènes sont composés d'un mélange de groupes d'atomes individuels possédant des couples  $(T_1, T_2)$  différents. Ce mélange est caractérisé globalement par une distribution  $X(T_1, T_2)$  permettant d'identifier et de quantifier les différents modes de relaxation présents dans le composant. Cette cartographie bi-dimensionnelle est privilégiée par rapport aux spectres indépendants  $X(T_1)$  et  $X(T_2)$  car elle permet une mise en évidence de couplages entre les paramètres de relaxation  $T_1$  et  $T_2$  aidant notablement à leur interprétation (English *et al.*, 1991; Song *et al.*, 2002). La technique permettant de déterminer la distribution couplée  $X(T_1, T_2)$ , nommée *RMN bi-dimensionnelle* ou *RMN 2D*, fait l'objet de notre étude.

### 5.1.1. Le modèle d'acquisition des données

L'expérience RMN 2D consiste à réaliser une série de mesures caractérisées par trois paramètres : le temps de répétition  $\tau_1$  et l'angle de bascule  $\Phi$  agissant sur la relaxation longitudinale  $T_1$ , et le temps d'écho  $\tau_2$  agissant sur la relaxation transversale  $T_2$ . La valeur mesurée  $Y(\tau_1, \tau_2, \Phi)$  est liée à la distribution  $X(T_1, T_2)$  par une intégrale de Fredholm du premier type (Hürlimann *et al.*, 2002; Song *et al.*, 2002)

$$Y(\tau_1, \tau_2, \Phi) = \int \int k_1(\tau_1, T_1, \Phi) X(T_1, T_2) k_2(\tau_2, T_2) dT_1 dT_2 \quad [62]$$



avec les noyaux

$$\begin{cases} k_1(\tau_1, T_1, \Phi) &= 1 - \gamma e^{-\tau_1/T_1} \\ k_2(\tau_2, T_2) &= e^{-\tau_2/T_2} \end{cases}, \quad \gamma = 1 - \cos \Phi.$$

Le spectre  $X(T_1, T_2)$  est déterminé en effectuant plusieurs expériences RMN avec des paramètres d'acquisition différents. Plus précisément, l'angle  $\Phi$  est maintenu constant (typiquement  $\Phi = 90^\circ$  ou  $180^\circ$ ) et les mesures expérimentales sont collectées sur une grille de  $m_1 \times m_2$  valeurs discrètes du couple  $(\tau_1, \tau_2)$ , ce qui fournit une matrice de données  $\mathbf{Y} \in \mathbb{R}^{m_1 \times m_2}$ . L'objectif du traitement est d'estimer le spectre  $X(T_1, T_2)$  à partir des mesures  $\mathbf{Y}$  en utilisant le modèle [62] pour une valeur du paramètre  $\Phi$  donnée. La quantité à estimer représentant une distribution, elle doit respecter la contrainte  $X \geq 0$ . Ce problème de reconstruction, équivalent à une inversion numérique d'une transformée de Laplace 2D, est un problème inverse mal posé (Sternin, 2008).

### 5.1.2. Le problème de reconstruction

Une approche standard consiste à reconstruire la distribution  $X(T_1, T_2)$  sur une grille régulière en  $T_1$  et  $T_2$  de taille  $n_1 \times n_2$ . Pour cela, les noyaux  $k_1$  et  $k_2$  sont mis sous forme discrète  $\mathbf{K}_1 \in \mathbb{R}^{m_1 \times n_1}$  et  $\mathbf{K}_2 \in \mathbb{R}^{m_2 \times n_2}$ . Cette modélisation se traduit par l'équation linéaire

$$\mathbf{Y} = \mathbf{K}_1 \mathbf{X} \mathbf{K}_2^t + \mathbf{E} \quad [63]$$

où  $\mathbf{E}$  concentre l'erreur de modélisation et également le bruit lié à la mesure. Le problème inverse résultant de ce modèle linéaire est l'estimation de  $\mathbf{X}$  à partir de  $\mathbf{Y}$  sous la contrainte  $\mathbf{X} \geq 0$  (au sens où tous les éléments de  $\mathbf{X}$  sont positifs).

Une difficulté majeure concerne la dimension du problème. En effet, les méthodes standard d'inversion requièrent la réécriture du modèle direct sous la forme

$$\mathbf{y} = \mathbf{K} \mathbf{x} + \boldsymbol{\epsilon}$$

avec  $\mathbf{y} = \text{vect}(\mathbf{Y})$ ,  $\mathbf{x} = \text{vect}(\mathbf{X})$ ,  $\boldsymbol{\epsilon} = \text{vect}(\mathbf{E})$  et  $\mathbf{K} = \mathbf{K}_1 \otimes \mathbf{K}_2$  où l'opérateur  $\text{vect}$  correspond à la transformation matrice-vecteur dans l'ordre lexicographique (c'est-à-dire par concaténation des éléments de la matrice parcourue ligne par ligne) et  $\otimes$  le produit de Kronecker (Van Loan, 2000). Néanmoins, ces méthodes ne sont plus applicables dès lors que la taille du problème augmente. En effet, une configuration typique est  $m_1 = 50$ ,  $m_2 = 10^4$  et  $n_1 \times n_2 = 200 \times 200$ , ce qui conduit à une matrice  $\mathbf{K}$  de  $2 \cdot 10^{10}$  éléments dont la construction et le stockage sont presque impossibles.

Une autre difficulté est liée au mauvais conditionnement du modèle à inverser. En effet, à cause de la forme exponentielle des noyaux, les matrices  $\mathbf{K}_1$  et  $\mathbf{K}_2$  sont de rang déficient et très mal conditionnées (Butler *et al.*, 1981). Par conséquent, une approche régularisée doit être envisagée.

## 5.2. Méthodes d'inversion pénalisée existantes

Nous présentons deux stratégies qui ont été employées pour résoudre le problème inverse de la détermination de  $\mathbf{X}$  connaissant  $\mathbf{Y}$ ,  $\mathbf{K}_1$  et  $\mathbf{K}_2$ . Chacune basée sur son propre principe de régularisation et d'optimisation, a comme objectif de chercher  $\mathbf{X}$  comme le minimiseur d'un critère d'adéquation aux données des moindres carrés augmenté d'un terme de pénalisation  $R(\mathbf{X})$ , sous contrainte de positivité

$$\min_{\mathbf{X} \geq \mathbf{0}} \left( F(\mathbf{X}) = \frac{1}{2} \|\mathbf{Y} - \mathbf{K}_1 \mathbf{X} \mathbf{K}_2^t\|_F^2 + \lambda R(\mathbf{X}) \right), \quad [64]$$

où  $\|\cdot\|_F$  représente la norme de Frobenius. La vérification de la contrainte de positivité est désirable du point de vue de l'interprétation physique, mais a aussi un effet favorable sur la résolution du spectre reconstruit. En particulier, la contrainte de positivité permet de limiter la présence d'oscillations, notamment autour des pics reconstruits (Chiang *et al.*, 2005).

- La méthode employée dans (Butler *et al.*, 1981; Venkataramanan *et al.*, 2002; Lamanna, 2005) repose sur l'utilisation de la pénalisation de Tikhonov

$$R(\mathbf{X}) = \frac{1}{2} \|\mathbf{x}\|^2 = \frac{1}{2} \|\mathbf{X}\|_F^2. \quad [65]$$

La minimisation sous contrainte de  $F$  est effectuée par l'algorithme itératif de Butler-Reeds-Dawson (BRD) (Butler *et al.*, 1981). Il s'agit d'un algorithme de Newton appliqué à la fonction duale du problème contraint au sens de la dualité Lagrange. L'algorithme BRD s'interprète également comme la minimisation de la fonction duale de  $F$ , au sens de la dualité Legendre-Fenchel (Chouzenoux *et al.*, 2010, App.A). Un inconvénient majeur de cet algorithme est que son implémentation nécessite l'inversion d'un système de taille  $m \times m$  à chaque itération, où  $m$  est la taille des mesures. Dans le cas de la RMN 2D,  $m = m_1 m_2$ , et les valeurs usuelles de  $m_1$  et de  $m_2$  entraînent un coût prohibitif de la méthode. Afin de pallier ce problème, une étape de compression des données préalable à l'application de l'algorithme BRD est nécessaire (Venkataramanan *et al.*, 2002), avec le risque de perdre des informations.

- La pénalisation par maximum d'entropie (ME) (Skilling *et al.*, 1984; Eggermont, 1993), reposant sur l'entropie de Shannon, est reconnue comme une stratégie efficace pour la reconstruction RMN 1D (Laue *et al.*, 1985; Mariette *et al.*, 1996). Dans le cas 2D, elle s'écrit

$$R(\mathbf{X}) = \sum_{ij} X_{ij} \log X_{ij}. \quad [66]$$

Elle présente l'avantage d'assurer la positivité de la solution de façon implicite, la norme du gradient de  $R$  étant infinie sur le bord de l'orthant positif. La minimisation du critère pénalisé peut donc être effectuée sans tenir compte de la contrainte. Cependant, à cause de la singularité de la pénalisation entropique, les algorithmes classiques d'optimisation non linéaire s'avèrent extrêmement lents et ont été rapidement remplacés par des méthodes dédiées à la forme spécifique du critère telle que

la méthode proposée par Skilling et Bryan (Skilling *et al.*, 1984). Cette méthode se fonde sur la minimisation successive d'approximations quadratiques du critère dans des sous-espaces construits de telle manière à favoriser la positivité des itérés (Pichon *et al.*, 1998). En pratique, si des composantes négatives apparaissent, elles sont remplacées par une petite valeur positive. D'après (Press *et al.*, 1992, p. 1022), la convergence de cet algorithme n'est pas établie. En particulier, l'opération de projection peut entraîner une croissance du critère entre deux itérés consécutifs (Chouzenoux *et al.*, 2009b). Pour cette raison, nous avons proposé un nouvel algorithme construit sur une base théorique plus solide, pour résoudre le problème de minimisation du critère pénalisé ME.

### 5.3. Algorithme de minimisation proposé

Soit le critère à minimiser

$$F(\mathbf{x}) = \|\mathbf{y} - \mathbf{K}\mathbf{x}\|^2 + \lambda \sum_i x_i \log x_i. \quad [67]$$

La méthode proposée est un algorithme de Newton tronqué (TN) (Dembo *et al.*, 1982; Nash, 2000). Cet algorithme est basé sur le schéma de descente itérative

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k, \quad [68]$$

où la direction de recherche  $\mathbf{d}_k$  est obtenue par résolution approchée des équations de Newton :

$$\nabla^2 F(\mathbf{x})_k \mathbf{d}_k = -\nabla F(\mathbf{x}_k). \quad [69]$$

En pratique, la direction  $\mathbf{d}_k$  résulte d'itérations de gradient conjugué sur le système [69] arrêtées avant convergence. D'autre part, le pas  $\alpha_k$  est obtenu par une méthode de recherche linéaire appliquée à la fonction scalaire  $f_k(\alpha) = F(\mathbf{x}_k + \alpha \mathbf{d}_k)$ . Nous avons adapté cet algorithme à la minimisation de [67] pour la RMN 2D à travers les modifications suivantes : La présence de la barrière entropique est prise en compte par le choix d'une stratégie de pas de type MM non quadratique. L'effet du mauvais conditionnement de la matrice  $\mathbf{K}$  et du Hessien du terme entropique est réduit par l'ajout d'un préconditionneur. Enfin, le coût de calcul de l'algorithme lié à la grande taille du problème est réduit par une stratégie d'implémentation originale.

#### 5.3.1. Recherche de pas

Dans le cas de la pénalisation ME, la dérivée de  $f_k$  tend vers  $-\infty$  lorsque l'une des composantes du vecteur  $\mathbf{x}_k + \alpha \mathbf{d}_k$  s'annule, c'est à dire lorsque  $\alpha$  est égal à l'une des deux valeurs limites

$$\alpha_k^- = \max_{i, d_{k,i} > 0} \left( \frac{-x_i}{d_{k,i}} \right), \quad \alpha_k^+ = \min_{i, d_{k,i} < 0} \left( \frac{-x_i}{d_{k,i}} \right). \quad [70]$$

Cette singularité nuit à l'efficacité des méthodes classiques de recherche de pas, telles que le backtracking, la dichotomie et l'interpolation cubique (Murray *et al.*, 1994).

C'est pourquoi nous avons proposé la stratégie de pas fondée sur la théorie des algorithmes Majoration-Minimisation (MM), présentée en Section 2.2. La minimisation de la fonction  $f_k$  est obtenue en effectuant des minimisations successives de fonctions majorantes de  $f_k$ .

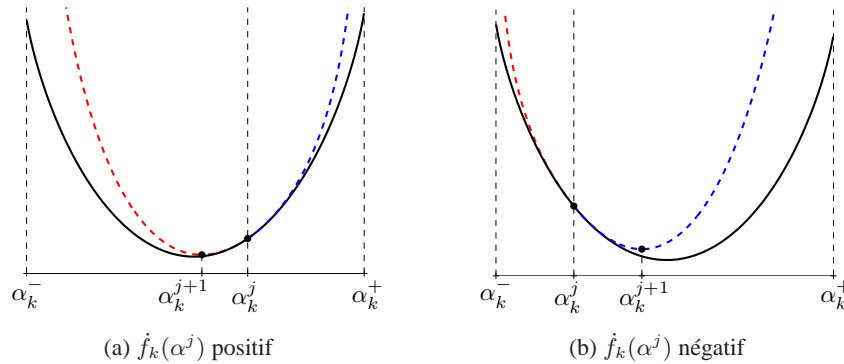
Le problème de minimisation initial de  $f_k(\alpha)$  est remplacé par une séquence de sous-problèmes plus simples donnant lieu à la règle de mise à jour MM :

$$\alpha_k^0 = 0, \quad \alpha_k^{j+1} = \operatorname{argmin}_{\alpha} h_k(\alpha, \alpha_k^j), \quad j \in \{1, \dots, J-1\} \quad [71]$$

où  $h_k(\cdot, \alpha_k^j)$  est une fonction tangente majorante de  $f_k$  en  $\alpha_k^j$ . La forme singulière de la pénalisation entropique est prise en compte à travers une fonction majorante non quadratique de la forme :

$$h_k(\alpha, \alpha_k^j) = \begin{cases} p_0^- + p_1^- \alpha + p_2^- \alpha^2 - p_3^- \log(\alpha - \alpha_k^-) & \text{for all } \alpha \in (\alpha_k^-, \alpha_k^j] \\ p_0^+ + p_1^+ \alpha + p_2^+ \alpha^2 - p_3^+ \log(\alpha_k^+ - \alpha) & \text{for all } \alpha \in [\alpha_k^j, \alpha_k^+) \end{cases} \quad [72]$$

Nous avons développé une stratégie de calcul des paramètres  $p_n^\pm$ ,  $n = 0 \dots 3$  permettant de garantir la propriété de majoration de  $h_k(\alpha, \alpha_k^j)$  (Chouzenoux *et al.*, 2009a).



**Figure 7.** Principe de la recherche de pas MM pour le critère du maximum d'entropie. La fonction tangente majorante  $h_k(\cdot, \alpha^j)$  (pointillés) pour  $f_k$  (ligne continue) en  $\alpha_j$  est définie par morceaux sur  $(\alpha_k^-, \alpha_k^j]$  et  $[\alpha_k^j, \alpha_k^+)$ . Son minimiseur fournit la prochaine itérée  $\alpha_k^{j+1}$ . Il s'obtient en minimisant la portion gauche (Cas (a)) ou droite (Cas (b)) suivant le signe de  $\dot{f}_k(\alpha_k^j)$

De plus, nous avons montré que la stratégie de pas proposée assure la convergence de l'algorithme TN vers le minimiseur de [67] quelle que soit la valeur du paramètre  $J$ . En pratique, les meilleures performances de l'algorithme en terme de nombre d'itérations et de temps de calcul sont réalisées lorsque l'on fait une seule itération de recherche de pas MM ( $J = 1$ ).

### 5.3.2. Préconditionnement

Comme le souligne (Hansen, 1998), le Hessien du terme entropique est très mal conditionné pour des valeurs proches de la frontière de l'orthant positif, car certaines de ces valeurs propres tendent vers l'infini. De plus, les termes exponentiels dans les noyaux  $k_1$  et  $k_2$  impliquent que  $\mathbf{K}_1$  et  $\mathbf{K}_2$  sont aussi très mal conditionnés. L'utilisation d'un préconditionneur  $M_k$  approchant l'inverse de la matrice  $\nabla^2 F(\mathbf{x}_k)$  dans l'étape de gradient conjugué est nécessaire pour avoir une convergence rapide de l'algorithme. La méthode de préconditionnement que nous avons proposé se base sur la décomposition en valeurs singulières tronquée de  $\mathbf{K}$  (Chouzenoux *et al.*, 2010, Sec.III-D). La troncature de la SVD est obtenue par une technique inspirée de (Venkataramanan *et al.*, 2002) : On calcule tout d'abord les TSVD de  $\mathbf{K}_1$  et  $\mathbf{K}_2$  de rangs respectifs  $v_1$  et  $v_2$  puis on approche  $\mathbf{K}$  par le produit de kronecker de ces deux TSVD. Le réglage des paramètres  $v_1, v_2$  dépend de la répartition des valeurs propres des matrices  $\mathbf{K}_i$ . En pratique, un bon compromis est réalisé pour  $4 \leq v_i \leq 8$ .

### 5.3.3. Implémentation

L'algorithme proposé peut être appliqué à la reconstruction d'un spectre RMN 2D sans la réduction de dimension suggérée par (Song *et al.*, 2002). En effet, le gradient du critère d'adéquation aux données ainsi que les produits hessien-vecteur et préconditionneur-vecteur s'expriment à l'aide d'opérations matricielles ne faisant pas intervenir la matrice  $\mathbf{K}$ , à condition d'exploiter la structure factorisée du modèle d'observation [63] (Chouzenoux *et al.*, 2010, Sec.III-E).

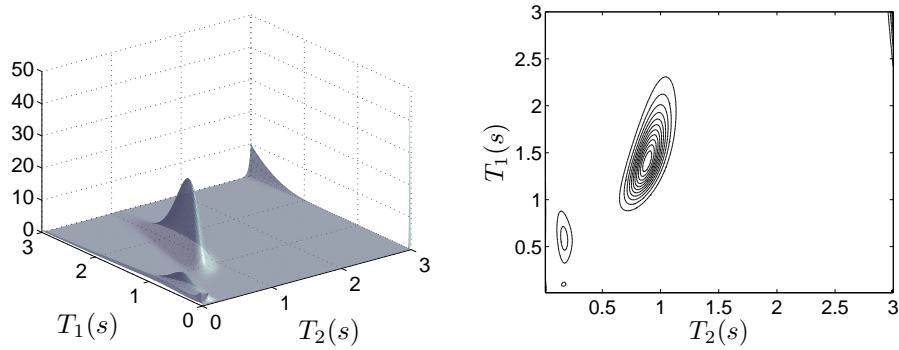
L'application de l'algorithme à des données expérimentales nécessite de trouver une façon de régler automatiquement le paramètre de régularisation  $\lambda$ . Pour ce faire, nous avons proposé une méthode basé sur la règle du Chi-carré (ou Chi-deux) (Chouzenoux *et al.*, 2010, Sec.IV-A-5). Une deuxième difficulté provient de la différence entre l'angle de bascule  $\Phi$  réellement engendré lors de l'acquisition et sa valeur désirée  $\hat{\Phi}$  (Song *et al.*, 2002). Par conséquent, le paramètre  $\gamma = 1 - \cos \Phi$  du modèle d'observation est erroné, ce qui a une influence néfaste sur les résultats de reconstruction, indépendamment de l'algorithme d'optimisation utilisé (Song *et al.*, 2002; Chouzenoux *et al.*, 2010). Nous avons développé une stratégie d'estimation de ce paramètre qui ne nécessite pas de modifier le modèle ni de pré-traiter les données (Chouzenoux *et al.*, 2010, Sec.IV-A-6).

## 5.4. Résultats expérimentaux

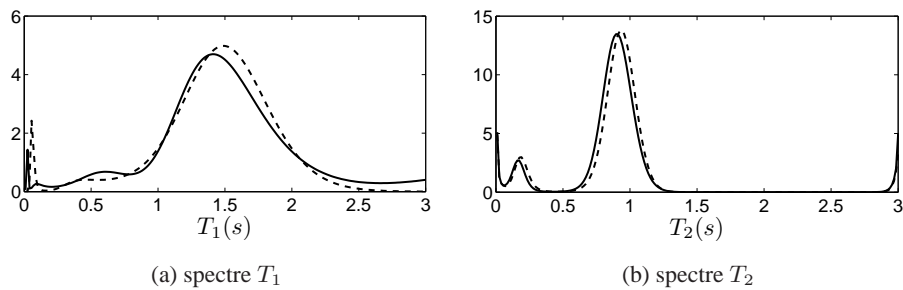
Nous présentons les résultats d'estimation d'un spectre  $T_1 - T_2$  d'un échantillon de pomme. Les mesures ont été effectuées pour  $m_1 = 50$  valeurs de  $\tau_1$ , réparties non uniformément entre 30 ms et 12 s. Dans tous les cas, le moment magnétique résultant est mesuré en  $m_2 = 10000$  valeurs du temps d'écho réparties uniformément entre 600  $\mu$ s et 8 s. L'algorithme TN ainsi développé a été utilisé pour reconstruire le spectre en  $N_1 = N_2 = 200$  valeurs des temps de relaxation  $T_1$  et  $T_2$ , réparties

uniformément entre 25 ms et 3 s. Les données ont été obtenues à partir d'un champ de radio-fréquence d'amplitude et de durée réglées afin de produire un angle de bascule de  $\hat{\Phi} = 90^\circ$  soit  $\hat{\gamma} = 1$ .

La figure 8 présente le spectre 2D reconstruit pour cette valeur estimée du paramètre  $\gamma$ . D'après notre stratégie d'estimation de  $\gamma$ , l'angle de bascule effectif est  $\Phi \approx 85^\circ$  (i.e.,  $\gamma = 0.92$ ), différent de  $\hat{\Phi}$ . On remarque que la distribution présente une corrélation entre  $T_1$  et  $T_2$ , par exemple au niveau du pic localisé autour de  $[T_1 = 1.4 \text{ s}, T_2 = 0.9 \text{ s}]$ . Cette information est très utile pour l'analyse de la structure moléculaire de l'échantillon (Hürlimann *et al.*, 2002). La similarité entre les spectres 1D  $T_1$  et  $T_2$  reconstruits et les spectres 1D déduits des marginalisations du spectre 2D mise en évidence sur la figure 9 illustre la qualité de la reconstruction 2D.



**Figure 8.** Spectre RMN 2D reconstruit par la méthode proposée à partir des données réelles.



**Figure 9.** Distributions 1D obtenues par marginalisation du spectre 2D reconstruit par la méthode proposée (trait plein) ou par inversion (trait pointillé).

## 6. Conclusion

Le projet ANR OPTIMED a offert une opportunité à l'utilisation d'outils modernes de l'analyse convexe et de la théorie des opérateurs monotones pour résoudre des problèmes appliqués de grande taille. En particulier, l'opérateur proximal d'une fonction convexe, qui généralise l'opérateur de projection pour les ensembles convexes, a joué un rôle clef dans ces développements et a permis de ramener des problèmes complexes d'optimisation à des problèmes de recherche d'un point fixe ayant une structure souple et propice à leur décomposition. Les méthodes proximales dans le cas de régularisation de type  $\ell_1$  sont étroitement liées à l'application itérative de techniques de seuillage doux, dont l'efficacité a été démontrée pour des analyses en ondelettes et pour des variantes de ces dernières. Un ingrédient essentiel dans l'amélioration des temps de convergence des méthodes d'optimisation proposées a été également la mise au point d'algorithmes de recherche de pas dédiés aux critères à barrières. Ces développements méthodologiques ont eu un impact direct sur les techniques de reconstruction d'images TEP, IRMp et RMN-2D.

## 7. Partenaires

Les résultats de ce papier ont été obtenus dans le cadre du projet ANR (programme ARA - MDMSA - 2005) OPTIMED, <http://optimized.univ-mlv.fr>, qui s'est déroulé de mars 2006 à septembre 2009, et dont les partenaires sont : l'Université Paris Est Marne-la-Vallée (J.-C. Pesquet, coordinateur), l'Université Pierre et Marie Curie Paris 6 (P. L. Combettes), le Commissariat à l'Energie Atomique - Centre de Saclay (NeuroSpin : Ph. Ciuciu et Service Hospitalier Frédéric Joliot (SHFJ) : C. Comtat) et l'Institut de Recherche en Communications et Cybernétique de Nantes (IRCCyN) (J. Idier).

## 8. Bibliographie

- Afonso M., Bioucas-Dias J., Figueiredo M., « Hybrid synthesis-analysis frame-based regularization : a criterion and an algorithm », *Signal processing with adaptive sparse structured representations - SPARS11*, Edinburgh, Scotland, 27-30 Jun., 2011a.
- Afonso M. V., Bioucas-Dias J. M., Figueiredo M. A. T., « An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems », *IEEE Trans. Image Process.*, vol. 20, n° 3, p. 681-695, Mar., 2011b.
- Bauschke H. H., Combettes P. L., *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics, Springer, 2011.
- Beck A., Teboulle M., « Gradient-based algorithms with applications to signal-recovery problems », in Y. D. Palomar (eds), *Convex Optimization in Signal Processing and Communications*, Cambridge University Press, 2009.

- Bect J., Blanc-Féraud L., Aubert G., Chambolle A., « A  $\ell_1$ -unified variational framework for image restoration », in T. Pajdla, J. Matas (eds), *Proc. European Conference on Computer Vision*, vol. LNCS 3024, Springer, Prague, Czech Republic, p. 1-13, May, 2004.
- Bertero M., Boccacci P., Desidera G., Vicidomini G., « Image deblurring with Poisson data : from cells to galaxies », *IEEE Trans. Image Process.*, 2009.
- Bertsekas D. P., *Nonlinear Programming*, 2ème edn, Athena Scientific, Belmont, MA, USA, 1999.
- Böhning D., Lindsay B. G., « Monotonicity of quadratic-approximation algorithms », *Ann. Inst. Statist. Math.*, vol. 40, n° 4, p. 641-663, 1988.
- Bredies K., Lorenz D. A., « Linear convergence of iterative soft-thresholding », *J. Fourier Anal. Appl.*, vol. 14, n° 5-6, p. 813-837, Dec., 2008.
- Briceño-Arias L. M., Combettes P. L., « A Monotone + Skew Splitting Model for Composite Monotone Inclusions in Duality », *SIAM J. Optim.*, 2011. à paraître.
- Butler J. P., Reeds J. A., Dawson S. V., « Estimating Solutions of First Kind Integral Equations with Nonnegative Constraints and Optimal Smoothing », *SIAM J. Numer. Anal.*, vol. 18, n° 3, p. 381-397, juin, 1981.
- Canet D., Boubel J.-C., Canet-Soulas E., *La RMN 2-ième édition : concepts, méthodes et applications*, 2ème edn, Dunod, Paris, 2002.
- Chari L., Mériaux S., Badillo S., Pesquet J.-C., Ciuciu P., « Multidimensional Wavelet-based Regularized Reconstruction for Parallel Acquisition in neuroimaging », 2011a, submitted to IEEE Trans. on Medical Imaging.
- Chari L., Pesquet J.-C., Benazza-Benyahia A., Ciuciu P., « Autocalibrated Parallel MRI Reconstruction in the Wavelet Domain », *Proc. IEEE ISBI*, Paris, France, p. 756-759, May 14-17, 2008.
- Chari L., Pesquet J.-C., Benazza-Benyahia A., Ciuciu P., « A wavelet-based regularized reconstruction algorithm for SENSE parallel MRI with applications to neuroimaging », *Medical Image Analysis*, vol. 15, n° 2, p. 185-201, Nov., 2011b.
- Chambolle A., « An algorithm for total variation minimization and applications », *J. Math. Imaging Vision*, vol. 20, n° 1-2, p. 89-97, Jan., 2004.
- Chambolle A., Pock T., « A first-order primal-dual algorithm for convex problems with applications to imaging », *J. Math. Imaging Vision*, vol. 40, n° 1, p. 120-145, 2011.
- Charbonnier P., Blanc-Féraud L., Aubert G., Barlaud M., « Deterministic edge-preserving regularization in computed imaging », *IEEE Trans. Image Process.*, vol. 6, p. 298-311, 1997.
- Chaux C., Combettes P. L., Pesquet J.-C., Wajs V. R., « A variational formulation for frame-based inverse problems », *Inverse Problems*, vol. 23, n° 4, p. 1495-1518, Aug., 2007.
- Chaux C., Pesquet J.-C., Pustelnik N., « Nested Iterative algorithms for convex constrained image recovery problems », *SIAM J. Imaging Sci.*, vol. 2, n° 2, p. 730-762, June, 2009.
- Chen G. H.-G., Rockafellar R. T., « Convergence rates in forward-backward splitting », *SIAM J. Optim.*, vol. 7, p. 421-444, 1997.
- Chen G., Teboulle M., « A proximal-based decomposition method for convex minimization problems », *Math. Program.*, vol. 64, p. 81-101, 1994.
- Chen X., Sun J., « Global Convergence of Two-parameter Family of Conjugate Gradient Methods without Line Search », *J. Comput. Appl. Math.*, vol. 146, p. 37-45, 2002.



- Chiang Y.-W., Borbat P. P., Freed J. H., « Maximum entropy : A complement to Tikhonov regularization for determination of pair distance distributions by pulsed ESR », *J. Magn. Reson.*, vol. 177, n° 2, p. 184-196, déc., 2005.
- Chouzenoux E., Idier J., Moussaoui S., « A Majorize-Minimize Strategy for Subspace Optimization Applied to Image Restoration », *IEEE Trans. Image Process.*, vol. 20, n° 18, p. 1517-1528, juin, 2011a.
- Chouzenoux E., Moussaoui S., Idier J., A Majorize-Minimize line search algorithm for barrier functions, Technical report, IRCCyN, nov., 2009a. <http://hal.archives-ouvertes.fr/hal-00362304>.
- Chouzenoux E., Moussaoui S., Idier J., Mariette F., « Reconstruction d'un spectre RMN 2D par maximum d'entropie », *Actes 22e coll. GRETSI*, Dijon, sep., 2009b.
- Chouzenoux E., Moussaoui S., Idier J., Mariette F., « Efficient Maximum Entropy Reconstruction of Nuclear Magnetic Resonance T1-T2 Spectra », *IEEE Trans. Signal Process.*, vol. 58, n° 12, p. 6040-6051, déc., 2010.
- Chouzenoux E., Pesquet J.-C., Talbot H., Jezierska A., « A Memory Gradient Algorithm for l2-l0 Regularization with Applications to Image Restoration », *Proc. IEEE ICIP*, Brussels-Belgique, sep., 2011b.
- Combettes P. L., « Iterative construction of the resolvent of a sum of maximal monotone operators », *J. Convex Anal.*, vol. 16, n° 4, p. 727-748, 2009.
- Combettes P. L., Đinh Dũng, Vũ B. C., « Proximity for sums of composite functions », *J. Math. Anal. Appl.*, vol. 380, n° 2, p. 680-688, Aug., 2011.
- Combettes P. L., Pesquet J.-C., « A Douglas-Rachford Splitting Approach to Nonsmooth Convex Variational Signal Recovery », *IEEE J. Selected Topics Signal Process.*, vol. 1, n° 4, p. 564-574, Dec., 2007a.
- Combettes P. L., Pesquet J.-C., « Proximal thresholding algorithm for minimization over orthonormal bases », *SIAM J. Optim.*, vol. 18, n° 4, p. 1351-1376, Nov., 2007b.
- Combettes P. L., Pesquet J.-C., « A proximal decomposition method for solving convex variational inverse problems », *Inverse Problems*, vol. 24, n° 6, p. x+27, Dec., 2008.
- Combettes P. L., Pesquet J.-C., « Proximal splitting methods in signal processing », in H. H. Bauschke, R. Burachik, P. L. Combettes, V. Elser, D. R. Luke, H. Wolkowicz (eds), *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, Springer-Verlag, New York, 2010.
- Combettes P. L., Wajs V. R., « Signal Recovery by proximal forward-backward splitting », *Multiscale Model. and Simul.*, vol. 4, n° 4, p. 1168-1200, Nov., 2005.
- Conn A. R., Gould N., Sartenaer A., Toint P. L., On iterated-subspace minimization methods for nonlinear optimization, Technical Report n° 94-069, Rutherford Appleton Laboratory, Oxfordshire, UK, mai, 1994. <ftp://130.246.8.32/pub/reports/cgstRAL94069.ps.Z>.
- Daubechies I., Defrise M., De Mol C., « An iterative thresholding algorithm for linear inverse problems with a sparsity constraint », *Comm. Pure Applied Math.*, vol. 57, n° 11, p. 1413-1457, Nov., 2004.
- De Pierro A. R., « A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography », *IEEE Trans. Medical Imaging*, vol. 14, n° 1, p. 132-137, mars, 1995.

- Dembo R. S., Eisenstat S. C., Steihaug S., « Inexact Newton Methods », *SIAM J. Numer. Anal.*, vol. 19, n° 2, p. 400-408, avr., 1982.
- Dempster A. P., Laird N. M., Rubin D. B., « Maximum likelihood from incomplete data via the EM algorithm », *J. R. Statist. Soc. B*, vol. 39, p. 1-38, 1977.
- Eggermont P. P. B., « Maximum entropy regularization for Fredholm integral equations of the first kind », *SIAM J. Math. Anal.*, vol. 24, n° 6, p. 1557-1576, 1993.
- English A. E., Whittall K. P., Joy M. L. G., Henkelman R. M., « Quantitative two-dimensional time correlation relaxometry », *Magn. Reson. Med.*, vol. 22, p. 425-434, 1991.
- Erdogan H., Fessler J. A., « Monotonic Algorithms for Transmission Tomography », *IEEE Trans. Medical Imaging*, vol. 18, n° 9, p. 801-814, sep., 1999.
- Esser E., Zhang X., Chan T., « A General Framework for a Class of First Order Primal-Dual Algorithms for Convex Optimization in Imaging Science », *SIAM J. Imaging Sci.*, vol. 3, n° 4, p. 1015-1046, 2010.
- Fessler J. A., Booth S. D., « Conjugate-Gradient Preconditioning Methods for Shift-Variant PET Image Reconstruction », *IEEE Trans. Image Process.*, vol. 8, n° 5, p. 688-699, mai, 1999.
- Fessler J. A., Erdogan H., « A paraboloidal surrogates algorithm for convergent penalized-likelihood emission image reconstruction », *IEEE Nuclear Science Symposium*, vol. 2, p. 1132-1135, 1998.
- Fiacco A. V., McCormick G. P., « The sequential unconstrained minimization technique (SUMT) without parameters », *Operations Research*, vol. 15, n° 5, p. 820-827, sep., 1967.
- Figueiredo M. A. T., Nowak R. D., « An EM algorithm for wavelet-based image restoration », *IEEE Trans. Image Process.*, vol. 12, n° 8, p. 906-916, Aug., 2003.
- Friston K. J., Holmes A. P., Worsley K. J., Poline J. P., Frith C. D., Frackowiak R. S. J., « Statistical parametric map in functional imaging : A general linear approach », *Human Brain Mapping*, vol. 2, n° 4, p. 189-210, Apr., 1994.
- Geman D., Reynolds G., « Constrained restoration and the recovery of discontinuities », *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, n° 3, p. 367-383, mars, 1992.
- Geman D., Yang C., « Nonlinear image recovery with half-quadratic regularization », *IEEE Trans. Image Process.*, vol. 4, n° 7, p. 932-946, juil., 1995.
- Gjedde A., « Calculation of cerebral glucose phosphorylation from brain uptake of glucose analogs in vivo : a re-examination », *Brain Res.*, vol. 257, n° 2, p. 237-274, 1982.
- Griswold M. A., Jakob P. M., Heidemann R. M., Nittka M., Jellus V., Wang J., Kiefer B., Haase A., « Generalized autocalibrating partially parallel acquisitions GRAPPA », *Magnetic Resonance in Medicine*, vol. 47, n° 6, p. 1202-1210, Jun., 2002.
- Hansen P. C., *Rank-deficient and discrete ill-posed problems : numerical aspects of linear inversion*, SIAM, Philadelphia, PA, USA, 1998.
- Hunter D. R., Kenneth L., « A Tutorial on MM Algorithms », *Amer. Statist.*, vol. 58, n° 1, p. 30-37, fév., 2004.
- Hürlimann M. D., Venkataramanan L., « Quantitative Measurement of Two-Dimensional Distribution Functions of Diffusion and Relaxation in Grossly Inhomogeneous Fields », *J. Magn. Reson.*, vol. 157, p. 31-42, 2002.
- Idier J., « Convex Half-Quadratic Criteria and Interacting Auxiliary Variables for Image Restoration », *IEEE Trans. Image Process.*, vol. 10, n° 7, p. 1001-1009, juil., 2001.

- Jacobson M. W., Fessler J. A., « An Expanded Theoretical Treatment of Iteration-Dependent Majorize-Minimize Algorithms », *IEEE Trans. Image Process.*, vol. 16, n° 10, p. 2411-2422, oct., 2007.
- Labat C., Idier J., Convergence of truncated half-quadratic and Newton algorithms, with application to image restoration, Technical report, IRCCyN, juin, 2007. <http://www.irccyn.ec-nantes.fr/~idier/pub/labat07b.pdf>.
- Labat C., Idier J., « Convergence of conjugate gradient methods with a closed-form stepsize formula », *J. Optim. Theory Appl.*, vol. 136, n° 1, p. 43-60, jan., 2008.
- Lamanna R., « On the inversion of multicomponent NMR relaxation and diffusion decays in heterogeneous systems », *Concepts Magn. Reson. Part A*, vol. 26A, n° 2, p. 78-90, 2005.
- Lange K., Fessler J. A., « Globally convergent algorithms for maximum a posteriori transmission tomography », *IEEE Trans. Image Process.*, vol. 4, n° 10, p. 1430-1438, oct., 1995.
- Laue E. D., Skilling J., Staunton J., Sibisi S., Brereton R. G., « Maximum entropy method in nuclear magnetic resonance spectroscopy », *J. Magn. Reson.*, vol. 62, n° 3, p. 437-452, 1985.
- Liu B., Ying L., Steckner M., Jun X., Sheng J., « Regularized SENSE Reconstruction using iteratively refined total variation method », *IEEE International Symposium on Biomedical Imaging (ISBI)*, p. 121-124, Washington DC, Apr. 12-15, 2007.
- Mariette F., « Relaxation RMN et IRM : un couplage indispensable pour l'étude des produits alimentaires », *C. R. Acad. Sci.*, vol. Chimie 7, p. 221-232, 2004.
- Mariette F., Guillement J. P., Tellier C., Marchal P., « Continuous relaxation time distribution decomposition by MEM », *Signal Treat. and Signal Anal. in NMR*, vol. 10, p. 218-234, 1996.
- Miele A., Cantrell J. W., « Study on a memory gradient method for the minimization of functions », *J. Optim. Theory Appl.*, vol. 3, n° 6, p. 459-470, 1969.
- Mohammad-Djafari A., Demoment G., « Utilisation de l'Entropie dans les problèmes de restauration et de reconstruction d'images », *IEEE Trans. Signal Process.*, vol. 5, n° 4, p. 235-248, 1988.
- Moreau J. J., « Fonctions convexes duales et points proximaux dans un espace hilbertien », *C. R. Acad. Sci.*, vol. 255, p. 2897-2899, 1962.
- Moreau J. J., « Proximité et dualité dans un espace hilbertien », *Bull. Soc. Math. France*, vol. 93, p. 273-299, 1965.
- Murray W., Wright M. H., « Line Search Procedures for the Logarithmic Barrier Function », *SIAM J. Optim.*, vol. 4, n° 2, p. 229-246, 1994.
- Narkiss G., Zibulevsky M., Sequential Subspace Optimization Method for Large-Scale Unconstrained Problems, Technical Report n° 559, Israel Institute of Technology, oct., 2005.
- Nash S. G., « A survey of truncated-Newton methods », *J. Comput. Appl. Math.*, vol. 124, p. 45-59, 2000.
- Ogawa S., Lee T. M., Kay A. R., Tank D. W., « Brain magnetic resonance imaging with contrast dependent on blood oxygenation », *Proceedings of the National Academy of Sciences of the United States of America*, vol. 87, p. 9868-9872, Dec., 1990.
- Ortega J. M., Rheinboldt W. C., *Iterative solution of nonlinear equations in several variables*, Academic Press, New York, USA, 1970.

- Patlak C. S., Blasberg R. G., Fenstermacher J. D., « Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data », *J. Cereb. Blood Flow Metab.*, vol. 3, n° 1, p. 1-7, Mar., 1983.
- Pesquet J.-C., Pustelnik N., « A parallel inertial proximal optimization method », 2011. [http://www.optimization-online.org/DB\\_HTML/2010/11/2825.html](http://www.optimization-online.org/DB_HTML/2010/11/2825.html).
- Pichon C., Thiébaud E., « Non-Parametric reconstruction of distribution functions from observed galactic discs », *Month. Not. Roy. Astr. Soc.*, vol. 301, n° 2, p. 419-434, 1998.
- Pinel P., Thirion B., Mériaux S., Jobert A., Serres J. and Le Bihan D., Poline J.-B., Dehaene S., « Fast reproducible identification and large-scale databasing of individual functional cognitive networks », *BMC Neurosci.*, vol. 8, n° 1, p. 91, Oct., 2007.
- Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., *Numerical Recipes : The Art of Scientific Computing*, 3ème edn, Cambridge Univ. Press, New York, 1992.
- Pruessmann K. P., Weiger M., Scheidegger M. B., Boesiger P., « SENSE : sensitivity encoding for fast MRI », *Magnetic Resonance in Medicine*, vol. 42, n° 5, p. 952-962, Jul., 1999.
- Pustelnik N., Chaux C., Pesquet J.-C., « Parallel Proximal Algorithm for Image Restoration Using Hybrid Regularization », *IEEE Trans. Image Process.*, vol. 20, n° 9, p. 2450-2462, Sep., 2011.
- Pustelnik N., Chaux C., Pesquet J.-C., Sureau F. C., Dusch E., C C., « Adapted Convex Optimization Algorithm for Wavelet-Based Dynamic PET Reconstruction », *International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine (Fully3D)*, Beijing, Chine, 5-10 Sep., 2009.
- Rivera M., Marroquin J., « Efficient half-quadratic regularization with granularity control », *Image and Vision Computing*, vol. 21, n° 4, p. 345-357, avr., 2003.
- Rudin L., Osher S., Fatemi E., « Nonlinear total variation based noise removal algorithms », *Physica D*, vol. 60, n° 1-4, p. 259-268, Nov., 1992.
- Setzer S., Steidl G., Teuber T., « Deblurring Poissonian images by split Bregman techniques », *J. Visual Communication and Image Representation*, vol. 21, n° 3, p. 193-199, 2010.
- Skilling J., Bryan R. K., « Maximum entropy image reconstruction : General algorithm », *Month. Not. Roy. Astr. Soc.*, vol. 211, p. 111-124, 1984.
- Song Y. Q., Venkataramanan L., Hürlimann M. D., Flaum M., Frulla P., Straley C., « T1-T2 correlation spectra obtained using a fast two-dimensional Laplace inversion », *J. Magn. Reson.*, vol. 154, p. 261-268, 2002.
- Soththivirat S., Fessler J. A., « Image recovery using partitioned-separable paraboloidal surrogate coordinate ascent algorithms », *IEEE Trans. Image Process.*, vol. 11, n° 3, p. 306-317, mars, 2002.
- Starck J.-L., Murtagh F., Fadili M. J., *Sparse Image and Signal Processing : Wavelets, Curvelets, Morphological Diversity*, Cambridge University Press, 2010.
- Sternin E., « Use of Inverse Theory Algorithms in the Analysis of Biomembrane NMR Data », *Methods in Membrane Lipids*, vol. 400, Humana Press, p. 103-125, 2008.
- Sun J., Zhang J., « Global Convergence of Conjugate Gradient Methods without Line Search », *Ann. Oper. Res.*, vol. 103, p. 161-173, mars, 2001.
- Sureau F. C., Pesquet J.-C., Chaux C., Pustelnik N., Reader A. J., Comtat C., Trebossen R., « Temporal wavelet denoising of PET sinograms and images », *Proc. Nuclear Science Symposium Conference Record*, Dresden, Germany, 19-25 Oct., 2008.

- Svaiter B. F., « Weak convergence on Douglas-Rachford method », *SIAM Journal on Control and Optimization*, vol. 49, n° 1, p. 280-287, 2011.
- Tikhonov A., « Tikhonov Regularization of incorrectly posed problems », *Soviet Mathematics Doklady*, vol. 4, p. 1624-1627, 1963.
- Valk P., Bailey D. L., Townsend D. W., Maisey M. N., *Positron Emission Tomography. Basic science and clinical practice*, Springer, 2004.
- Van Loan C. F., « The ubiquitous Kronecker product », *J. Comput. Appl. Math.*, vol. 123, p. 85-100, 2000.
- Venkataramanan L., Song Y., Hürlimann M. D., « Solving Fredholm integrals of the first kind with tensor product structure in 2 and 2.5 dimensions », *IEEE Trans. Signal Process.*, vol. 50, n° 5, p. 1017-1026, 2002.
- Verhaeghe J., Van De Ville D., Khalidov I., D'Asseler Y., Lemahieu I., Unser M., « Dynamic PET Reconstruction Using Wavelet Regularization with Adapted Basis Functions », *IEEE Trans. Med. Imag.*, vol. 27, n° 7, p. 943-959, July, 2008.
- Wright M. H., « Why a pure primal Newton barrier step may be infeasible », *SIAM J. Optim.*, vol. 5, p. 1-12, 1995.
- Yuan Y., « Subspace techniques for nonlinear optimization », *Some Topics in Industrial and Applied Mathematics*, vol. 8, Series on Concrete and Applicable Mathematics, p. 206-218, 2007.
- Zheng J., Saquib S. S., Sauer K., Bouman C. A., « Parallelizable Bayesian Tomography Algorithms with rapid, guaranteed convergence », *SIAM J. Optim.*, vol. 14, n° 4, p. 1043-1056, 2004.