

Reconstruction d'un spectre RMN 2D par maximum d'entropie

Emilie CHOUZENOUX¹, Saïd MOUSSAOUI¹, Jérôme IDIER¹, François MARIETTE^{2,3}

¹IRCCYN, CNRS UMR 6597, 1, rue de la Noë, BP 92101, F-44321 Nantes Cedex 03, France
phone: + 33 2 40 37 69 24, fax: + 33 2 40 37 69 30
{emilie.chouzenoux, said.moussaoui, jerome.idier}@irccyn.ec-nantes.fr

²Cemagref, UR TERE CS 64427, 17 avenue de Cucillé, 35044 Rennes Cedex, France
phone: + 33 2 23 48 21 21, fax: + 33 2 23 48 21 15
francois.mariette@cemagref.fr

³Université Européenne de Bretagne, France

Résumé – La résonance magnétique nucléaire (RMN) est une méthode moderne de spectroscopie utilisée pour l'analyse de la composition de produits biologiques. Nous nous intéressons dans cet article à l'estimation d'un spectre corrélation $T_1 - T_2$ à partir de mesures RMN. Les difficultés de l'estimation sont liées au caractère mal-posé de ce problème inverse et à la taille importante des données à traiter. La méthode d'estimation est fondée sur le Maximum d'Entropie et nous proposons deux algorithmes de reconstruction itérative ; le premier est fondé sur l'algorithme de Bryan et Skilling et le second utilise le gradient conjugué non-linéaire. Par ailleurs, la structure du modèle d'observation est avantageusement exploitée pour alléger le coût de calcul sans employer les approximations proposées récemment par Vankataramanan *et coll.*. De plus, nous proposons le rajout d'une étape de recherche de pas adaptée à la fonction entropique afin d'assurer une décroissance du critère. Les algorithmes sont évalués sur un exemple synthétique et leur applicabilité est illustrée sur des données réelles.

Abstract – Nuclear magnetic resonance (NMR) is a modern spectroscopy tool for the study of biological structures. This article deals with the reconstruction of $T_1 - T_2$ correlation spectrum from 2D NMR measures. The ill-posed character of this inverse problem and the large size of data constitute the main difficulties of the reconstruction. The Maximum Entropy method is used for estimation and two optimization algorithms are compared, the first is based on the classical Skilling and Bryan method and the second uses the nonlinear conjugate gradient method. In order to reduce computational cost, we exploit the specific structure of the observation model to avoid the approximations recently proposed in Vankataramanan *et coll.*. Furthermore, an original line search strategy suitable for the entropy function is used to ensure the decrease of the criterion at each iteration. These proposed algorithms are tested on a synthetic example and their applicability is illustrated through real NMR data processing.

1 Introduction

La spectroscopie par résonance magnétique nucléaire permet d'accéder à des informations structurales aux échelles microscopiques ou moléculaires. Par exemple, dans le domaine agro-alimentaire cette technique est employée pour déterminer la teneur en eau ou en matière grasse d'un échantillon biologique. D'une façon générale, les mesures sont basées sur la réalisation d'enregistrements des temps de relaxation, T_1 et T_2 de façon indépendante. Des travaux ont démontré que la mesure couplée de ces paramètres permettait un accroissement de la robustesse du calcul mais aussi une mise en évidence de couplages entre T_1 et T_2 aidant notablement à leur interprétation [2, 11]. Toutefois la réalisation de mesures couplées implique de disposer d'une méthode de traitement des données pour l'estimation des spectres à 2 dimensions. Cet article présente une méthode basée sur le Maximum d'Entropie pour l'estimation d'un spectre de corrélation $T_1 - T_2$.

Le signal mesuré dépend à la fois de la durée t_1 corres-

pondant au temps laissé au spin pour évoluer vers sa position d'équilibre (relaxation T_1) et de l'instant t_2 d'enregistrement de l'amplitude de l'écho RMN à t_1 fixé. Le modèle physique [11] relie ce signal RMN 2D, noté $X(t_1, t_2)$, à la distribution des populations de molécules en fonction des temps de relaxation T_1 et T_2 , notée $S(T_1, T_2)$. Ces temps de relaxation caractérisent la dynamique des molécules en fonction de t_1 et t_2 , respectivement. Ce modèle s'exprime par

$$X(t_1, t_2) = \iint k_1(t_1, T_1)S(T_1, T_2)k_2(t_2, T_2)dT_1dT_2 \quad (1)$$

avec des noyaux

$$\begin{cases} k_1(t_1, T_1) &= 1 - e^{-t_1/T_1} \\ k_2(t_2, T_2) &= e^{-t_2/T_2} \end{cases}$$

En pratique les signaux RMN, notés $Y(t_1, t_2)$, sont affectés d'un bruit additif $E(t_1, t_2)$, supposé i.i.d. gaussien centré. Après discrétisation, le modèle s'écrit sous la forme matricielle

$$\mathbf{Y} = \mathbf{K}_1 \mathbf{S} \mathbf{K}_2^t + \mathbf{E} \quad (2)$$

$$\mathbf{K}_1 \in \mathbb{R}^{m_1 \times N_1}, \mathbf{K}_2 \in \mathbb{R}^{m_2 \times N_2}, \mathbf{Y} \in \mathbb{R}^{m_1 \times m_2}, \mathbf{S} \in \mathbb{R}^{N_1 \times N_2}$$

où m_1, m_2 représentent le nombre de points de mesure en t_1, t_2 et N_1, N_2 la taille de la grille T_1, T_2 .

L'objectif du traitement est d'estimer le spectre \mathbf{S} à partir de \mathbf{Y} en utilisant le modèle (2) et sous la contrainte $\mathbf{S} \geq 0$ (au sens où tous les éléments de \mathbf{S} sont positifs). Ce problème de reconstruction est équivalent à une inversion numérique d'une transformée de Laplace 2D, qui est très instable en raison du mauvais conditionnement des matrices \mathbf{K}_1 et \mathbf{K}_2 [11].

2 Méthodes de résolution

2.1 Position du problème

Afin d'obtenir une solution régularisée, les méthodes existantes minimisent un critère d'adéquation aux données $C(\mathbf{S})$ augmenté d'un terme de pénalisation $R(\mathbf{S})$, sous contrainte de positivité

$$\min_{\mathbf{S} \geq 0} L(\mathbf{S}) = C(\mathbf{S}) + \beta R(\mathbf{S}). \quad (3)$$

Le bruit étant supposé gaussien, $C(\mathbf{S})$ s'exprime par

$$C(\mathbf{S}) = \frac{1}{2} \|\mathbf{Y} - \mathbf{K}_1 \mathbf{S} \mathbf{K}_2^t\|_F^2 \quad (4)$$

où $\|\cdot\|_F$ représente la norme de Frobenius.

Comme le spectre RMN 2D correspond à une distribution de populations de molécules, une pénalisation de type maximum d'entropie [6] nous paraît adaptée. Une méthode de reconstruction fondée sur le maximum d'entropie a d'ailleurs été précédemment proposée en RMN 1D [5]. L'entropie de Shannon sera utilisée

$$R(\mathbf{S}) = \sum_{T_1, T_2} S(T_1, T_2) \log S(T_1, T_2) \quad (5)$$

Une autre difficulté concerne la dimension du problème [12]. En effet, les méthodes existantes requièrent la réécriture du modèle direct sous la forme

$$\mathbf{y} = \mathbf{K} \mathbf{s} + \mathbf{e}$$

avec $\mathbf{y} = \text{vect}(\mathbf{Y})$, $\mathbf{s} = \text{vect}(\mathbf{S})$, $\mathbf{e} = \text{vect}(\mathbf{E})$ et $\mathbf{K} = \mathbf{K}_1 \otimes \mathbf{K}_2$. L'opérateur vect correspond à la transformation matrice-vecteur dans l'ordre lexicographique et \otimes le produit de Kronecker. 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. Pour contourner ce problème, [11] applique une troncature de la décomposition en valeurs singulières des matrices \mathbf{K}_1 et \mathbf{K}_2 pour réduire la taille de la matrice \mathbf{K} . Or d'une part cette troncature est source de perte d'informations et d'autre part comme il sera montré dans le paragraphe 2.4, il n'est pas nécessaire de calculer cette matrice. En effet, l'évaluation du gradient et du hessien du critère peut être réalisée avec une complexité réduite en tirant profit de la forme séparable du noyau de convolution dans le modèle d'observation.

2.2 Méthode de Bryan et Skilling

La méthode proposée par Skilling et Bryan [10] se fonde sur une approximation quadratique du critère dans un sous-espace \mathcal{D}_r de dimension r engendré à chaque itération k par les vecteurs

$$\{\mathbf{d}_k^{(1)} \dots \mathbf{d}_k^{(r)}\} \quad (6)$$

décrits par les vecteurs de base suivants

$$\begin{cases} \mathbf{d}_k^{(1)} &= \phi_k(\nabla C(\mathbf{s}_k)) \\ \mathbf{d}_k^{(2)} &= \phi_k(\nabla R(\mathbf{s}_k)) \\ \mathbf{d}_k^{(j)} &= \phi_k(\nabla^2 C(\mathbf{s}_k)) \left(\frac{\mathbf{d}_k^{(j-2)}}{\|\mathbf{d}_k^{(j-2)}\|_{\mathbf{s}_k}} - \frac{\mathbf{d}_k^{(j-1)}}{\|\mathbf{d}_k^{(j-1)}\|_{\mathbf{s}_k}} \right) \end{cases} \quad (7)$$

en définissant les coefficients de normalisation

$$\|\mathbf{d}^{(j)}\|_{\mathbf{s}} = \left(\sum_n s_n (\mathbf{d}_n^{(j)})^2 \right)^{1/2} \quad (8)$$

et la fonction d'échelle ϕ_k correspondant à une multiplication terme à terme par \mathbf{s}_k . D'après [8], cette fonction permet de prendre en compte la positivité. L'équation de mise à jour de \mathbf{s} est alors donnée par

$$\mathbf{s}_{k+1} = \mathbf{s}_k + \sum_{j=1}^r \xi_k^{(j)} \mathbf{d}_k^{(j)} = \mathbf{s}_k + \mathbf{D}_k \boldsymbol{\xi}_k, \quad (9)$$

où \mathbf{D}_k est la matrice dont les colonnes contiennent les vecteurs de la base et les poids $\xi_k^{(j)}$ résultent de la minimisation d'une approximation quadratique du critère $L(\mathbf{s})$ au point \mathbf{s}_k sur le sous espace \mathcal{D}_r . Ces poids sont donc solution de l'équation normale

$$(\mathbf{D}_k^t \nabla^2 L(\mathbf{s}_k) \mathbf{D}_k) \boldsymbol{\xi}_k = -\mathbf{D}_k^t \nabla L(\mathbf{s}_k). \quad (10)$$

Cet algorithme peut être vu comme un algorithme de descente itérative $\mathbf{s}_{k+1} = \mathbf{s}_k + \alpha_k \boldsymbol{\delta}_k$ dans lequel la direction est $\boldsymbol{\delta}_k = \mathbf{D}_k \boldsymbol{\xi}_k$ et le pas α_k est fixé à 1. Cependant, le choix d'un pas unitaire ne garantit pas la décroissance du critère, ni même le respect de la contrainte de positivité. En pratique, si l'application de (9) fournit des composantes négatives, elles sont remplacées par une petite valeur positive ϵ . Cependant, cette projection peut provoquer une croissance du critère.

Afin de garantir la décroissance du critère à chaque itération, nous proposons la modification suivante : Un test portant sur la décroissance du critère est effectué, et une valeur de pas assurant la décroissance est adoptée dans la direction $\boldsymbol{\delta}_k$ si ce test est négatif ; la formule de pas utilisée est issue de travaux récents détaillés dans [1].

D'autre part, nous suggérons de sélectionner à chaque itération la valeur de r la plus favorable en terme de décroissance du critère, plutôt que de fixer r à une valeur constante. En effet, nous avons constaté que le comportement et l'efficacité de la méthode de Skilling et Bryan dépendait de façon critique du choix de r , et qu'il était difficile d'anticiper le meilleur choix. La stratégie que nous proposons n'entraîne pas d'augmentation du coût de calcul par itération, à condition de tirer parti de l'emboîtement des matrices normales du système linéaire (10).

2.3 Méthode du gradient conjugué non linéaire

D'après [9, p. 1022], la convergence mathématique de l'approche de Skilling et Bryan reste à prouver. Pour cette raison, nous proposons une autre famille d'algorithmes construits sur une base théorique plus solide, et plus spécifiquement un algorithme de gradient conjugué non-linéaire, utilisant une direction de descente donnée par :

$$\mathbf{d}_k = -\mathbf{c}_k \text{sign}(\mathbf{g}_k^t \mathbf{c}_k) \quad \text{avec} \quad \mathbf{c}_k = -\mathbf{g}_k + \beta_k \mathbf{d}_{k-1} \quad (11)$$

où β_k est le coefficient de conjugaison. La mise à jour $\mathbf{s}_{k+1} = \mathbf{s}_k + \alpha_k \mathbf{d}_k$ nécessite la recherche d'un pas α_k s'approchant du minimiseur de la fonction scalaire $\ell(\alpha) = L(\mathbf{s}_k + \alpha \mathbf{d}_k)$. Dans le cas du maximum d'entropie, $\ell(\alpha)$ (ou l'opposé de sa dérivée) tend vers l'infini lorsque α se rapproche de $\bar{\alpha}$, plus petite valeur du pas pour laquelle une composante du vecteur $\mathbf{s}_k + \bar{\alpha} \mathbf{d}_k$ s'annule. Or la présence de cette barrière nuit à l'efficacité des méthodes classiques de recherche de pas, telles que le backtracking, la dichotomie et l'interpolation cubique [7]. C'est pourquoi nous avons appliqué une stratégie de pas fondée sur la théorie des algorithmes Maximisation-Minimisation (MM) [3, 4]. La minimisation de la fonction ℓ est obtenue en effectuant des minimisations successives de fonctions majorantes de ℓ . La fonction $h(\alpha, \alpha')$ est dite tangente majorante de ℓ en α' si

$$h(\alpha', \alpha') = \ell(\alpha') \quad \text{et} \quad h(\alpha, \alpha') \geq \ell(\alpha), \quad \forall \alpha \in [0, \bar{\alpha}] \quad (12)$$

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

$$\alpha^{j+1} = \arg \min_{\alpha} h(\alpha, \alpha^j). \quad (13)$$

Dans le cas d'une barrière logarithmique, la minimisation d'une fonction majorante de la forme :

$$h_k(\alpha, \alpha') = \ell(\alpha') + (\alpha - \alpha') \dot{\ell}(\alpha') + \frac{1}{2} m_k (\alpha - \alpha')^2 + \gamma_k \left[(\bar{\alpha} - \alpha') \log \left(\frac{\bar{\alpha} - \alpha'}{\bar{\alpha} - \alpha} \right) - \alpha + \alpha' \right] \quad (14)$$

est étudiée dans [1]. $h_k(\alpha, \alpha')$ est strictement convexe et possède une unique minimiseur dont la forme est explicite

$$\alpha' + \frac{-A_2 + \sqrt{A_2^2 - 4A_1 A_3}}{2A_1}, \quad (15)$$

avec $A_1 = -m_k$, $A_2 = \gamma_k - \dot{\ell}(\alpha') + m_k(\bar{\alpha} - \alpha')$ et $A_3 = (\bar{\alpha} - \alpha') \dot{\ell}(\alpha')$. Une méthode de calcul des paramètres $(m_k, \gamma_k) > 0$ est présentée dans [1]. Les résultats montrent la convergence de l'algorithme GCNL vers la solution de (3) pour différentes formes de conjugaisons.

2.4 Mise en œuvre

Les algorithmes proposés peuvent être appliqués à la reconstruction d'un spectre RMN 2D sans la réduction de dimension suggérée par [11]. En effet, à l'aide d'opérations matricielles ne nécessitant pas la matrice \mathbf{K} , le gradient du critère

d'adéquation aux données et le produit hessien-vecteur s'expriment par :

$$\begin{aligned} \nabla C(\mathbf{s}) &= \text{vect}(-\mathbf{K}_1^t (\mathbf{Y} - \mathbf{K}_1 \mathbf{S} \mathbf{K}_2^t) \mathbf{K}_2) \\ \nabla^2 C(\mathbf{s}) \cdot \mathbf{v} &= \text{vect}(\mathbf{K}_1^t \mathbf{K}_1 \mathbf{V} \mathbf{K}_2^t \mathbf{K}_2) \end{aligned}$$

où \mathbf{V} est la transformation vecteur-matrice de \mathbf{v} . D'autre part, nous proposons d'utiliser le critère d'arrêt

$$\frac{\|\mathbf{s}_{k+1} - \mathbf{s}_k\|_2}{\|\mathbf{s}_k\|_2} < 10^{-4} \quad (16)$$

3 Résultats expérimentaux

Exemple synthétique : Afin de tester les méthodes d'optimisation proposées, nous comparons les algorithmes suivants

- Méthode de Skilling et Bryan (SB)
- Variante n°1 : Ajout de la recherche de pas (SB1)
- Variante n°2 : Recherche de r entre 1 et r_m (SB2)
- Gradient conjugué non linéaire (GCNL)

Les mesures sont simulées à l'aide du modèle direct (2) avec un RSB de 30 dB (Fig. 1.a) et $m_1 = 200$, $m_2 = 500$. La reconstruction a été réalisée pour $N_1 = 100$, $N_2 = 100$ avec $\beta = 10^{-7}$.

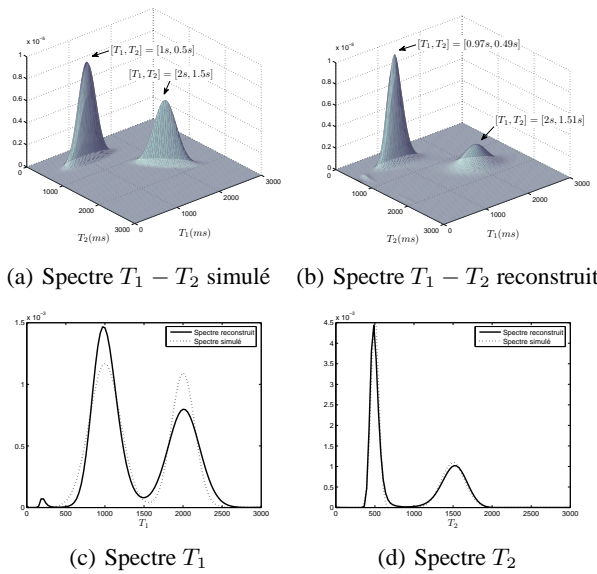


FIGURE 1 – Exemple synthétique de reconstruction RMN 2D

Nous pouvons constater que l'algorithme SB et ses deux variantes présentent des temps de convergence similaires (Tab. 1). Néanmoins, ces variantes permettent d'assurer la décroissance du critère, donc un comportement plus régulier. Enfin, la méthode SB2 permet de se prévaloir du réglage de la dimension tout en restant aussi rapide que SB1 avec le meilleur choix de r . D'autre part, l'algorithme GCNL requiert un plus grand nombre d'itérations mais le coût de calcul par itération étant faible, il reste compétitif avec la méthode de Skilling et Bryan. En outre, l'analyse visuelle des résultats révèle la bonne estimation de

| | | <i>Iter</i> | <i>Temps(s)</i> | $\ s - s^o\ _2 / \ s^o\ _2$ |
|------------------|-----------|-------------|-----------------|-----------------------------|
| GCNL | | 1042 | 98, 17 | 0, 31 |
| SB (<i>r</i>) | 4 | 2063 | 195, 77 | 0, 311 |
| | 5 | 963 | 104, 8 | 0, 2969 |
| | 7 | 675 | 86, 406 | 0, 2974 |
| SB1 (<i>r</i>) | 4 | 2063 | 195, 77 | 0, 311 |
| | 5 | 963 | 104, 13 | 0, 2969 |
| | 7 | 795 | 106, 48 | 0, 30 |
| SB2 | $r_m = 5$ | 783 | 155, 16 | 0, 299 |

TABLE 1 – Comparaison entre les algorithmes SB et GCNL pour la reconstruction de données synthétiques. Les calculs sont effectués sur un Intel Pentium 4 3.2 GHz, 3 GB RAM

la position des modes du spectre (Fig. 1). Cependant, la sous-estimation de l'amplitude de l'un des deux modes s'explique par le faible nombre de points de mesure en t_1 .

Données expérimentales : Nous présentons également les résultats d'estimation d'un spectre $T_1 - T_2$ d'un échantillon de pomme. Les mesures ont été effectuées pour $m_1 = 50$, $m_2 = 10000$ et la reconstruction a été réalisée pour $N_1 = 200$, $N_2 = 200$ et $\beta = 10^{-8}$. Les algorithmes SB2 et GCNL ont été appliqués. Le spectre reconstruit (Fig. 2) présente 3 modes dont

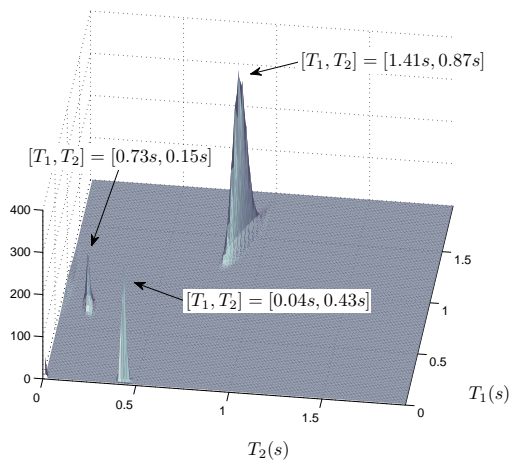


FIGURE 2 – Reconstruction de données expérimentales RMN 2D

les positions sont conformes à l'analyse théorique. De plus, le mode situé en $[T_1; T_2] = [1, 41s, 0, 87s]$ met en évidence une dépendance $T_1 - T_2$ et par conséquent illustre l'avantage de l'analyse 2D car cette information n'apparaît pas sur les spectres T_1 et T_2 séparément.

En ce qui concerne les méthodes d'optimisation, la méthode SB2 converge en 1794 itérations (2700 s) tandis que l'algorithme GCNL requiert 7993 itérations (4900 s). D'une façon générale, la méthode de Skilling et Bryan est sensiblement plus rapide que l'algorithme GCNL. De notre point de vue, cette efficacité est due à la construction de la direction à partir d'un sous-espace défini par plusieurs directions.

4 Conclusion

Nous avons présenté une méthode de reconstruction d'un spectre RMN 2D par maximum d'entropie. Cette méthode présente l'avantage d'un coût de calcul raisonnable et d'assurer la décroissance monotone du critère à minimiser. Bien que cette méthode s'avère efficace sur les données simulées et expérimentales, l'absence de résultats théoriques de convergence de la méthode SB nous oriente vers d'autres méthodes à base de sous-espaces ou encore vers l'accélération de la méthode GCNL par l'utilisation d'un préconditionneur adéquat.

Références

- [1] E. Chouzenoux, S. Moussaoui, and J. Idier. A majorize-minimize line search algorithm for barrier function. Technical report, IRCCyN, 2009. <http://hal.archives-ouvertes.fr/IRCCYN-ADTSI>.
- [2] A. E. English, K. P. Whittall, M. L. G. Joy, and R. M. Henkelman. Quantitative two-dimensional time correlation relaxometry. *Magn. Reson. Med.*, 22 :425–434, 1991.
- [3] D. R. Hunter and K. L. A tutorial on MM algorithms. *Amer. Statist.*, 58(1) :30–37, février 2004.
- [4] C. Labat and J. Idier. Convergence of conjugate gradient methods with a closed-form stepsize formula. *J. Optim. Theory Appl.*, 136(1) :43–60, janvier 2008.
- [5] F. Mariette, J. P. Guillemet, C. Tellier, and P. Marchal. Continuous relaxation time distribution decomposition by MEM. *Signal Treat. and Signal Anal. in NMR*, pages 218–234, 1996.
- [6] A. Mohammad-Djafari and G. Demoment. Maximum entropy image reconstruction in X-ray and diffraction tomography. *IEEE Trans. Medical Imaging*, 7(4) :345–354, 1988.
- [7] W. Murray and M. H. Wright. Line search procedures for the logarithmic barrier function. *SIAM J. Optimization*, 4(2) :229–246, 1994.
- [8] C. Pichon and E. Thiébaud. Non-parametric reconstruction of distribution functions from observed galactic discs. *Month. Not. Roy. Astr. Soc.*, 301(2) :419–434, 1998.
- [9] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes : The Art of Scientific Computing*. Cambridge Univ. Press, New York, 3ème édition, 1992.
- [10] J. Skilling and R. K. Bryan. Maximum entropy image reconstruction : General algorithm. *Month. Not. Roy. Astr. Soc.*, 211 :111–124, 1984.
- [11] Y. Q. Song, L. Venkataramanan, M. D. Hürlimann, M. Flaum, P. Frulla, and C. Straley. T1-T2 correlation spectra obtained using a fast two-dimensional Laplace inversion. *J. Magn. Reson.*, 154 :261–268, 2002.
- [12] L. Venkataramanan, Y. Q. Song, and M. D. Hürlimann. Solving Fredholm integrals of the first kind with tensor product structure in 2 and 2.5 dimensions. *IEEE Trans. Signal Processing*, 50(5) :1017–1026, 2002.