Reconstruction bayésienne d'images MRFM parcimonieuses

Nicolas DOBIGEON^{1,2}, Alfred O. HERO², Jean-Yves TOURNERET¹

¹Université de Toulouse, IRIT/INP-ENSEEIHT/TéSA, 2 rue Camichel, 31071 Toulouse cedex 7, France

²University of Michigan, Department of EECS, Ann Arbor, MI 48109-2122, USA

{nicolas.dobigeon, jean-yves.tourneret}@enseeiht.fr, hero@umich.edu

Résumé – Cet article présente un algorithme de reconstruction d'images issues d'un microscope à résonance magnétique (*magnetic resonance force microscopy* en anglais, MRFM) qui ont la particularité d'être parcimonieuses, c'est à dire constituées seulement de quelques pixels non nuls. Cet algorithme repose sur un modèle bayésien hierarchique tel que la loi *a priori* des pixels de l'image MRFM est un mélange d'une masse en zéro et d'une loi exponentielle positive. Par ailleurs, les hyperparamètres associés à ce modèle sont estimés de manière non-supervisée à partir des données. La loi *a priori* choisie permet de respecter les contraintes de positivité imposées sur les pixels et fournit un estimateur MAP plus parcimonieux que celui obtenu avec une simple pénalisation ℓ_1 . En outre, l'estimation non supervisée des hyperparamètres rend la reconstruction plus robuste que des méthodes empiriques rencontrées dans la littérature.

Abstract – This paper studies a hierarchical Bayesian model for reconstructing magnetic resonance force microscopy (MRFM) images that consist of only few non-zero pixels. An appropriate prior distribution that promotes sparsity is introduced for the pixel values whereas a standard inverse gamma prior distribution is chosen for the noise variance. The hyperparameters involved in the resulting Bayesian model are also assigned prior distributions, resulting in a hierarchical model. The corresponding posterior distribution of the unknown model parameters and hyperparameters is too complex to derive their MAP estimator. As a consequence, we study a Gibbs sampler that draws samples distributed according this posterior. The generated samples are then used to approximate the MAP estimator of the image pixels. Simulation results show that the proposed Bayesian estimator outperforms previous estimators proposed in the literature.

1 Introduction

Des avancées techniques récentes en matière d'imagerie moléculaire permettent d'obtenir des images dont la résolution atteint aujourd'hui une résolution nanoscopique. Le microscope à résonance magnétique permettra prochainement d'observer des images à un niveau atomique [1]. Les images fournies par cette technologie sont naturellement parcimonieuses dans l'espace d'observation. Nous proposons dans cet article un algorithme bayésien hiérarchique pour reconstruire ce type d'images.

La reconstruction d'images ou de signaux parcimonieux a donné lieu à des travaux divers dans de nombreux champs applicatifs : l'analyse spectrale en astronomie [2], l'analyse de signaux sismiques en géophysique [3]... La stratégie développée dans cet article repose sur un modèle bayésien hiérarchique basé sur le choix d'une loi a priori adéquate pour les pixels de l'image. Cette loi est définie à l'aide d'un mélange d'une loi exponentielle positive et d'une masse en zéro. Elle permet de prendre en compte le fait que peu de pixels de l'image sont non nuls et que ces pixels non-nuls sont positifs. Associer une masse en zéro à une loi à queue lourde a déjà été envisagé dans la littérature puisque ce mélange permet d'accentuer la parcimonie de la loi *a priori*. Ce type de loi de mélange a par exemple été employé dans [4] pour résoudre un problème général de débruitage. Plus récemment, Ting a proposé dans [5] de combiner une masse en zéro avec une loi de Laplace pour reconstruire des images parcimonieuses bruitées. Le principal inconvénient des approches sus-citées est la sensibilité des résultats obtenus aux hyperparamètres du modèle. Des solutions empiriques d'estimation ont été envisagées dans [5] pour résoudre ce problème. Cependant, une grande instabilité des résultats a été constatée, particulièrement à fort rapport signal sur bruit (RSB).

La stratégie étudiée dans cet article permet de dépasser les difficultés énoncées ci-dessus puisque les hyperparamètres associés au modèle bayésien d'intérêt sont estimés directement à partir des images observées. Plus précisément, ces hyperparamètres sont estimés à l'aide de données générées à l'aide d'un échantillonneur de Gibbs [6]. Ces données simulées permettent d'approcher l'estimateur du maximum a posteriori (MAP) des hyperparamètres, ce qui se montre particulièrement bien adapté à la reconstruction d'images parcimonieuses. Dans un cadre bayésien, deux approches sont habituellement considérées pour l'estimation des hyperparamètres à l'aide de méthodes de Monte Carlo par chaînes de Markov (MCMC). La première approche couple une méthode MCMC avec un algorithme d'espérancemaximisation (EM) pour maximiser une fonction de vraisemblance pénalisée [7]. La seconde approche consiste à choisir des lois a priori non-informatives pour les hyperparamètres, introduisant naturellement un deuxième niveau de hiérarchie dans l'inférence bayésienne. Cette seconde stratégie, adoptée dans cet article, a été appliquée avec succès pour des problèmes de segmentation [8] ou d'analyse d'images hyperspectrales [9].

Elle fournit une estimation de la loi *a posteriori* jointe des paramètres et hyperparamètres inconnus du modèle, et non une simple estimation ponctuelle. Cette loi *a posteriori* peut ensuite être utilisée pour calculer des intervalles de confiance associés aux estimateurs, ou même pour calculer la probabilité d'observer une particule dans une zone donnée de l'image [10].

2 Position du problème

Soit X une image inconnue de taille $n_r \times n_c$ à estimer. La matrice observée Y de taille $n_{\rm r} \times n_{\rm c}$ est issue du mélange convolutif bruité $\mathbf{Y} = \boldsymbol{\kappa} \otimes \mathbf{X} + \mathbf{N}$, où \otimes représente l'opérateur 2D de convolution, N est une matrice de bruit de taille $n_{\rm r} \times n_{\rm c}$ et κ est le noyau de convolution modélisant la réponse de l'imageur. Les images X et Y peuvent être vectorisées en un vecteur inconnu $\mathbf{x} \in \mathbb{R}^M$ et un vecteur de mesures $\mathbf{y} \in \mathbb{R}^M$, respectivement, avec $M = n_{\rm r} n_{\rm c}$. Avec ces notations, l'équation de mesure devient y = Hx + n, où H est une matrice Toeplitz de taille $M \times M$ qui décrit l'opération de convolution par la fonction d'étalement κ . Dans l'équation de mesure précédente, n est un bruit additif blanc Gaussien de variance σ^2 inconnue, ce que l'on notera $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_M)$. Le problème considéré dans cet article consiste à estimer le vecteur image x et la variance du bruit σ^2 sous des contraintes de parcimonie et de positivité sur x, à partir des observations y et de la matrice de transformation H.

3 Modèle bayésien hiérarchique

Le modèle d'observation défini précédemment et la loi du bruit **n** conduisent à une vraisemblance Gaussienne de moyenne **Hx** et de variance inconnue σ^2 . La modélisation bayésienne hiérarchique requiert la définition de lois *a priori* pour les paramètres inconnus associés au modèle d'observation précédemment défini. Le vecteur de ces paramètres inconnus est $\theta = \{\mathbf{x}, \sigma^2\}$. Nous détaillons ci-dessous les lois *a priori* choisies pour ces deux paramètres.

3.1 Lois *a priori* des paramètres

Nous choisissons comme loi *a priori* pour le pixel mesuré x_i un mélange d'une masse en zéro et d'une loi exponentielle positive

$$f(x_i|w,a) = (1-w)\delta(x_i) + \frac{w}{a}\exp\left(-\frac{x_i}{a}\right)\mathbf{1}_{\mathbb{R}^*_+}(x_i), \quad (1)$$

où $\delta(\cdot)$ est une masse en zéro ($\delta(x_i) = 0$ si $x_i \neq 0$ et $\delta(0) = 1$) et $\mathbf{1}_{\mathbb{E}}(x)$ est la fonction indicatrice définie sur \mathbb{E} . En supposant que les pixels x_i sont *a priori* indépendants (i = 1, ..., M), on obtient la loi jointe *a priori* suivante pour le vecteur \mathbf{x}

$$f(\mathbf{x}|w,a) = \prod_{i=1}^{M} \left[(1-w)\delta(x_i) + \frac{w}{a} \exp\left(-\frac{x_i}{a}\right) \mathbf{1}_{\mathbb{R}^*_+}(x_i) \right].$$

Cette loi est similaire à la loi LAZE (moyenne pondérée entre une loi de Laplace et une masse en zéro) utilisée par exemple dans [5]. Cependant, elle permet en plus de tenir compte de la positivité des pixels à estimer.

Dans un cadre d'estimation bayésienne, il est courant de choisir une loi inverse-Gamma comme loi *a priori* de la variance σ^2 , i.e., $\sigma^2 | \gamma \sim IG\left(\frac{\nu}{2}, \frac{\gamma}{2}\right)$. Dans cet article, l'hyperparamètre γ est supposé inconnu et sera estimé, tandis que l'hyperparamètre ν sera fixé à $\nu = 2$, comme dans [8], [9].

3.2 Lois *a priori* des hyperparamètres

Il est clair que la qualité de la reconstruction dépend du vecteur d'hyperparamètres $\mathbf{\Phi} = \{a, \gamma, w\}$. Nous proposons d'estimer ces hyperparamètres à partir des données en utilisant un modèle bayésien hiérarchique. Des lois non-informatives sont choisies comme loi *a priori* pour ces hyperparamètres supposés indépendants, i.e., $f(\mathbf{\Phi}) = f(a|\alpha) f(\gamma) f(w)$.

Une loi inverse-gamma est choisie comme loi *a priori* pour l'hyperparamètre $a = \mathcal{TC}(a_1, a_2)$ (2)

$$a|\boldsymbol{\alpha} \sim \mathcal{I}\mathcal{G}\left(\alpha_{0},\alpha_{1}\right), \tag{2}$$

avec $\boldsymbol{\alpha} = [\alpha_0, \alpha_1]^T$. Les hyper-hyperparamètres α_0 et α_1 sont fixés afin d'obtenir une loi vague, c'est à dire avec une grande variance : $\alpha_0 = \alpha_1 = 10^{-10}$.

Une loi *a priori* non-informative de Jeffreys est choisie pour l'hyperparamètre γ 1

$$f(\gamma) \propto \frac{1}{\gamma} \mathbf{1}_{\mathbb{R}_{+}}(\gamma)$$
. (3)

Une loi uniforme sur [0, 1] est choisie comme loi *a priori* pour l'hyperparamètre w

$$w \sim \mathcal{U}\left([0,1]\right). \tag{4}$$

(6)

3.3 Loi a posteriori

avec

La loi *a posteriori* $\{\boldsymbol{\theta}, \boldsymbol{\Phi}\}$ s'écrit :

$$f(\boldsymbol{\theta}, \boldsymbol{\Phi} | \mathbf{y}) \propto f(\mathbf{y} | \boldsymbol{\theta}) f(\boldsymbol{\theta} | \boldsymbol{\Phi}) f(\boldsymbol{\Phi}),$$
 (5)

$$f\left(\boldsymbol{\theta}|\boldsymbol{\Phi}\right) = f\left(\mathbf{x}|a,w\right)f\left(\sigma^{2}|\boldsymbol{\gamma}\right),$$

où $f(\mathbf{y}|\boldsymbol{\theta})$ et $f(\boldsymbol{\Phi})$ ont été définies précédemment. Cette structure hiérarchique permet d'intégrer dans cette loi jointe le paramètre de nuisance σ^2 et le vecteur hyperparamètre $\boldsymbol{\Phi}$, ce qui permet de calculer la loi *a posteriori* marginale de **x**

$$f\left(\mathbf{x}|\mathbf{y}\right) \propto \frac{B\left(n_{1}+1, n_{0}+1\right)}{\left\|\mathbf{y}-\mathbf{Hx}\right\|^{M}} \frac{\Gamma\left(n_{1}+\alpha_{0}\right)}{\left(\left\|\mathbf{x}\right\|_{1}+\alpha_{1}\right)^{n_{1}+\alpha_{0}}}, \quad (7)$$

où $B(\beta_1, \beta_0) = \Gamma(\beta_1) \Gamma(\beta_0) / \Gamma(\beta_1 + \beta_0)$, $\Gamma(.)$ est la fonction gamma, $n_1 = \|\mathbf{x}\|_0$, $n_0 = M - \|\mathbf{x}\|_0$ et $\|\mathbf{x}\|_0$ désigne le nombre d'éléments non nuls de \mathbf{x} .

Malheureusement, le calcul exact de l'estimateur du maximum *a posteriori* de x à partir de cette loi *a posteriori* est difficile, en raison notamment de la nature combinatoire introduite par les quantités n_0 et n_1 dans (7). L'alternative proposée dans cet article consiste à utiliser une méthode MCMC pour générer des échantillons $\mathcal{X} = {\mathbf{x}^{(t)}}_{t=1,...}$ asymptotiquement distribués suivant $f(\mathbf{x}|\mathbf{y})$ [6]. Les échantillons générés permettront ensuite d'approcher l'estimateur MAP

$$\hat{\mathbf{x}}_{\text{MAP}} = \operatorname*{argmax}_{\mathbf{x} \in \mathbb{R}^{M}_{+}} f\left(\mathbf{x} | \mathbf{y}\right) \approx \operatorname*{argmax}_{\mathbf{x} \in \mathcal{X}} f\left(\mathbf{x} | \mathbf{y}\right). \tag{8}$$

Cet algorithme de Gibbs est décrit dans la section suivante.

4 Echantillonneur de Gibbs pour la reconstruction d'images MRFM

4.1 Echantillonnage suivant $f(\mathbf{x} | \sigma^2, \mathbf{y}, \boldsymbol{\alpha})$

Pour générer des échantillons distribués suivant $f(\mathbf{x} | \sigma^2, \mathbf{y}, \boldsymbol{\alpha})$, il est judicieux d'échantillonner suivant la loi jointe

$$f\left(\mathbf{x}, w, a \left| \sigma^{2}, \mathbf{y}, \boldsymbol{\alpha} \right.\right) \tag{9}$$

en utilisant les 3 étapes suivantes.

- Étape 1 : génération suivant $f(w|\mathbf{x})$ Des calculs élémentaires conduisent à

$$f(w|\mathbf{x}) \propto (1-w)^{n_0} w^{n_1},$$
 (10)

où n_0 et n_1 ont été définis dans la section précédente. Ainsi la génération d'échantillons distribués suivant $f(w | \mathbf{x})$ s'effectue selon

$$w | \mathbf{x} \sim \mathcal{B} \left(n_1 + 1, n_0 + 1 \right). \tag{11}$$

– Étape 2 : génération suivant f (a |x, α)
 En examinant la loi *a posteriori* jointe (5), nous obtenons

$$a |\mathbf{x}, \boldsymbol{\alpha} \sim \mathcal{IG}(\|\mathbf{x}\|_0 + \alpha_0, \|\mathbf{x}\|_1 + \alpha_1).$$
 (12)

- Étape 3 : génération suivant $f(\mathbf{x} | w, a, \sigma^2, \mathbf{y})$ La loi *a posteriori* de chaque pixel x_i (i = 1, ..., M)conditionnellement aux autres s'écrit

$$f(x_i|w, a, \sigma^2, \mathbf{x}_{-i}, \mathbf{y}) \propto (1 - w_i)\delta(x_i) + w_i\phi_+(x_i|\mu_i, \eta_i^2),$$
(13)

avec

$$\begin{cases} \mathbf{e}_{i} = \mathbf{y} - \sum_{k \neq i} x_{k} \mathbf{h}_{i}, \\ \eta_{i}^{2} = \frac{\sigma^{2}}{\|\mathbf{h}_{i}\|^{2}}, \\ \mu_{i} = \eta_{i}^{2} \left(\frac{\mathbf{h}_{i}^{T} \mathbf{e}_{i}}{\sigma^{2}} - \frac{1}{a}\right), \\ u_{i} = \frac{w}{a} C\left(\mu_{i}, \eta_{i}^{2}\right) \exp\left(\frac{\mu_{i}^{2}}{2\eta_{i}^{2}}\right), \\ w_{i} = \frac{u_{i}}{u_{i} + (1 - w)}, \end{cases}$$
(14)

où \mathbf{x}_{-i} est le vecteur \mathbf{x} privé de sa *i*ème composante et \mathbf{h}_i est la *i*ème colonne de \mathbf{H} . Dans (13), ϕ_+ (\cdot, m, s^2) est la densité de probabilité de la loi normale de moyenne m et de variance s^2 tronquée sur \mathbb{R}^*_+ . Donc, d'après (13), $x_i | w, a, \sigma^2, \mathbf{x}_{-i}, \mathbf{y}$ suit une loi de Bernoulli-Gauss tronquée sur \mathbb{R}^*_+ avec les paramètres (w_i, μ_i, η_i^2). Pour générer des échantillons suivant cette loi, il est nécessaire de faire appel à une variable de Bernoulli auxiliaire et aux stratégies de simulation présentées dans [11].

4.2 Echantillonnage suivant $f(\sigma^2 | \mathbf{x}, \mathbf{y})$

La loi *a posteriori* de $\sigma^2 | \mathbf{x}, \mathbf{y} |$ est

$$\sigma^2 |\mathbf{x}, \mathbf{y} \sim \mathcal{IG}\left(\frac{M}{2}, \frac{\|\mathbf{y} - \mathbf{Hx}\|^2}{2}\right).$$
 (15)

Paramètre		Vəleur
Description	Nom	valeur
Amplitude du champ magnétique externe	Bext	$9.6 imes 10^3 { m G}$
Valeur de B_{mag} dans la zone de résonnance	Bres	$1.0 imes 10^4 { m ~G}$
Rayon de la sonde	R_0	4.0 nm
Distance entre la sonde et l'échantillon	d	7.6 nm
Moment de la sonde	m	$4.6\times 10^5~{\rm emu}$
Oscillation max. de la sonde	$x_{ m pk}$	0.8 nm
Gradient max. du champ magnétique	G_{\max}	125

TAB. 1 – Paramètres de la réponse MRFM.

5 Résultats de simulation

Nous considérons ici une image synthétique de taille 32×32 générée en utilisant la loi *a priori* introduite au paragraphe 3.1 avec les paramètres a = 1 et w = 0.02 (voir la Fig. 1 (droite)). Sur cette figure, un pixel blanc signifie une valeur nulle. Suivant le modèle physique de l'imageur MRFM décrit dans [12], un noyau de convolution de taille 7×9 , représentée sur la Fig. 1 (gauche), a été généré en utilisant des paramètres physiques réalistes, rapportés dans le tableau 1.



FIG. 1 – Gauche : réponse de l'imageur MRFM. Droite : image parcimonieuse inconnue à estimer.

Les mesures observées sont entachées d'un bruit additif Gaussien avec deux valeurs différentes de variance, correspondant à RSB = 2dB et RSB = 20dB (voir Fig. 2).



FIG. 2 – Gauche (resp. droite) : observations bruités pour un RSB = 2dB (resp. 20dB).

Notre algorithme est appliqué sur ces données avec $N_{\rm MC} = 2000$ itérations de l'échantillonneur de Gibbs dont $N_{\rm bi} = 300$ itérations de chauffage. L'estimateur MAP de l'image à reconstruire est alors calculé selon (8). Ces estimateurs sont représentés sur la Fig. 3 pour les deux niveaux de bruit considérés. Ces résultats sont en accord avec l'image réelle, même à faible RSB.

Par ailleurs, comme l'échantillonneur de Gibbs fournit des échantillons distribués asymptotiquement suivant la loi *a pos*-



FIG. 3 – Gauche (resp. droite) : image estimée pour RSB = 2dB (resp. 20dB).

teriori (7) des paramètres inconnus, ces échantillons peuvent être utilisés pour estimer la loi *a posteriori* de chacun de ces paramètres. Par exemple, les lois *a posteriori* de la probabilité d'avoir un pixel non-nul *w* et de la variance du bruit σ^2 sont représentées sur la figure 4 pour les deux niveaux de bruits. Ces lois sont en accord avec les valeurs théoriques de ces paramètres. De même, les lois *a posteriori* des amplitudes de deux pixels non-nuls sont représentées sur la figure Fig. 5. Ces lois sont aussi en accord avec les valeurs réelles de ces amplitudes, représentées en traits pointillés rouges sur ces figures.



FIG. 4 – Haut, gauche (resp. droite) : loi *a posteriori* de l'hyperparamètre *w* pour RSB = 2dB (resp. RSB = 20dB). Bas, gauche (resp. droite) : loi *a posteriori* du paramètre σ^2 pour RSB = 2dB (resp. RSB = 20dB)



FIG. 5 – Lois *a posteriori* de deux composants non-nuls de x pour RSB = 20dB (les vraies valeurs apparaissent en pointillés rouges).

6 Conclusions et perspectives

Cet article décrit un algorithme bayésien pour la reconstruction d'images MRFM. Afin d'exploiter le caractère parcimonieux de ces images, une loi composée d'une masse en zéro et d'une loi exponentielle est proposée comme loi *a priori* pour les pixels. Grâce à un modèle hiérarchique, l'algorithme proposé permet l'estimation non-supervisée des hyperparamètres associés à cette loi. Un échantillonneur de Gibbs permet alors d'approcher l'estimateur MAP de l'image inconnue. Une perspective à ces travaux consisterait à estimer conjointement la réponse de l'imageur dans le cas où celle-ci n'est pas connue.

Références

- D. Mounce, "Magnetic resonance force microscopy," *IEEE Instr. Meas. Magazine*, vol. 8, no. 2, pp. 20–26, June 2005.
- [2] S. Bourguignon, H. Carfantan, and J. Idier, "A sparsitybased method for the estimation of spectral lines from irregularly sampled data," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, Dec. 2007.
- [3] Q. Cheng, R. Chen, and T.-H. Li, "Simultaneous wavelet estimation and deconvolution of reflection seismic signals," *IEEE Trans. Geosci. and Remote Sensing*, vol. 34, no. 2, pp. 377–384, March 1996.
- [4] I. M. Johnstone and B. W. Silverman, "Needles and straw in haystacks : empirical Bayes estimates of possibly sparse sequences," *Ann. Stat.*, vol. 32, no. 4, pp. 1594– 1649, 2004.
- [5] M. Ting, R. Raich, and A. O. Hero, "Sparse image reconstruction for molecular imaging," *IEEE Trans. Image Process.*, vol. 18, no. 6, pp. 1215–1227, May 2009.
- [6] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed. New York, NY, USA : Springer, 2004.
- [7] E. Kuhn and M. Lavielle, "Coupling a stochastic approximation version of EM with an MCMC procedure," *ESAIM Probab. Statist.*, vol. 8, pp. 115–131, 2004.
- [8] N. Dobigeon, J.-Y. Tourneret, and J. D. Scargle, "Joint segmentation of multivariate astronomical time series : Bayesian sampling with a hierarchical model," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 414–423, Feb. 2007.
- [9] N. Dobigeon, J.-Y. Tourneret, and C.-I Chang, "Semisupervised linear spectral unmixing using a hierarchical Bayesian model for hyperspectral imagery," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 2684–2695, July 2008.
- [10] N. Dobigeon, A. O. Hero, and J.-Y. Tourneret, "Hierarchical Bayesian sparse image reconstruction with application to MRFM," *IEEE Trans. Image Process.*, 2009, à paraître.
- [11] C. P. Robert, "Simulation of truncated normal variables," *Statistics and Computing*, vol. 5, no. 2, pp. 121–125, June 1995.
- [12] J. Mamin, R. Budakian, and D. Rugar, "Point response function of an MRFM tip," IBM Research Division, Tech. Rep., Oct. 2003.