

Soutenance d'Habilitation à diriger des Recherches

Inversion Bayésienne

Thomas Rodet

Maître de conférences à l'Université Paris-Sud 11
Laboratoire des Signaux et Systèmes (CNRS-SUPELEC-UPS)
Groupe Problèmes Inverses

20 novembre 2012

Expériences de Recherche

- 1999 Thèse sur la reconstruction tomographique rapide (3D + T)
- 2002 Études post-doctorales :
- Algorithme de reconstruction tomographique à partir de données hélicoïdales.
 - Propriétés de l'algorithme FDK.
 - Condition d'échantillonnage en TEP.
- 2003 Maître de Conférences à l'Université Paris Sud 11 : poste 61-34
- Reconstruction sur-résolue d'images astrophysiques infra-rouge.
 - Reconstruction tomographique 3D +T de la couronne solaire.
 - Approche MCMC pour la résolution de problèmes de déconvolution non supervisée.
- 2008 Prise de la responsabilité du GPI :
- Reconstruction des données SPIRE/Herschel.
 - Développement d'approches bayésiennes variationnelles accélérées.
 - Inversion de données tomographiques polychromatiques.
- 2011 Demi CRCT LIGM : Bayésien variationnel et ondelettes.

Encadrements de thèses

- N. Barbey Détermination de la structure tridimensionnelle de la couronne solaire à partir d'images des missions spatiales SoHO et STEREO. Soutenue le 19/11/2008 et encadrée à 40 % avec F. Auchère et J.-C. Vial.
- F. Orieux Inversion bayésienne myope et non-supervisée pour l'imagerie sur-résolue. Application à l'instrument SPIRE de l'observatoire spatial Herschel. Soutenue le 16/11/2009 et encadrée à 50 % avec A. Abergel et J.-F. Giovannelli.
- C. Cai Approches bayésiennes pour la reconstruction des fractions d'os et d'eau à partir de données tomographiques polychromatiques. Soutenance 01/2013, encadrée à 50 % avec A. Djafari et S. Legoupil.
- L. Chen Approches statistiques dédiées à la reconstruction d'images de la mâchoire. Soutenance 12/2014, encadrée à 60 % avec N. Gac et C. Maury.
- Y. Zheng Nouvelles méthodes bayésiennes variationnelles. Soutenance 10/2014, encadrée à 60 % avec A. Fraysse.

Plan :

- Curriculum Vitae
- Illustration de la démarche suivie lors de la résolution d'un problème inverse.
- Nouvelle approche bayésienne variationnelle.
- Illustration sur un problème astrophysique
- Perspectives de recherche.

Modèle direct :

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \epsilon$$

- Choix : loi du bruit ϵ .

Règle de Bayes :

$$p(\mathbf{x}|\mathbf{y}, \mathcal{M}) = \frac{p(\mathbf{y}|\mathbf{x}, \mathcal{M})p(\mathbf{x}|\mathcal{M})}{p(\mathbf{y}|\mathcal{M})}$$

Informations *a priori*

- Introduction d'une information diffuse connue avant de faire les mesures.

Estimation

- MAP : $\hat{\mathbf{x}} = \arg \max p(\mathbf{x}|\mathbf{y}, \mathcal{M})$
- EAP : $\hat{\mathbf{x}} = E\{\mathbf{x}\}_{p(\mathbf{x}|\mathbf{y}, \mathcal{M})}$

Démarche suivie lors de la résolution d'un problème inverse

Paramétrisation de l'objet inconnu

- Modèle paramétrique
 - faible nombre de paramètres.
 - Décomposition sur une famille ou une base de fonctions
- Non paramétrique

Modèle direct

- Temps de calcul
- Précision

Vraisemblance

- Bruit iid, blanc ou corrélé.

Informations *a priori*

- Information de douceur.
- Régularité par morceaux, parcimonie.

Inversion des données IRS (Sptizer) [Rodet08]

Contexte

F. Orieux (Collaboration avec l'IAS A. Abergel)

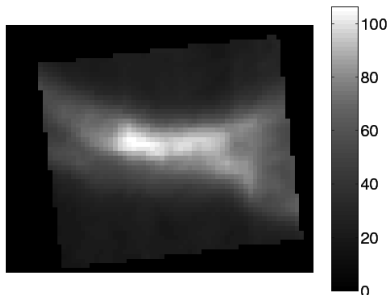
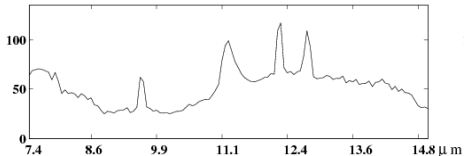
Problématique

- Inversion sur-résolue d'un cube (α, β, λ) .
- Largeur de la RI de l'optique dépend de λ .

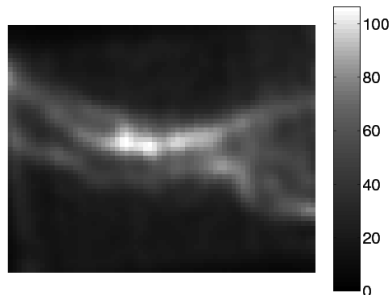
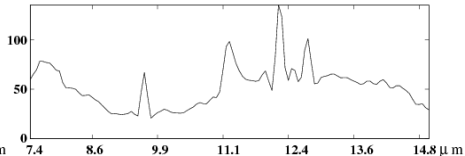
Contributions

- Décomposition d'objets inconnus sur une famille de gaussiennes.
- Modèle direct linéaire $\mathbf{y} = \mathbf{H}\mathbf{x} + \epsilon$,
- Approximation de la RI de l'optique par une gaussienne, utilisation de la stratégie d'acquisition.
- Calcul analytique (rapide) des éléments de la matrice \mathbf{H} .

Inversion de données réelles : la tête de cheval

Pipeline NASA à $11.27 \mu\text{m}$ 

Spectre du point central (NASA)

Approche proposée à $11.27 \mu\text{m}$ 

Spectre du point central (proposée)

Inversion des données EIT (SoHO, STEREO) [Kherrab10]

Contexte

N. Barbey, B. Kherrab (Collaboration avec l'IAS F. Auchère et J.-C. Vial)

Problématique

- Reconstruction 3D + T de la couronne solaire à partir de 1 à 2 projections par instant de reconstruction (Tomographie dynamique).
- Hypothèse : La morphologie des plumes est invariante, seule l'amplitude change.

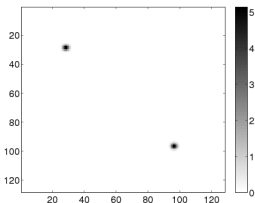
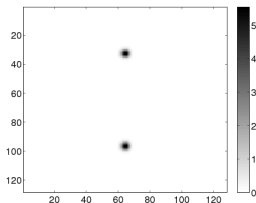
Contributions

- Formulation comme un problème de séparation de sources :

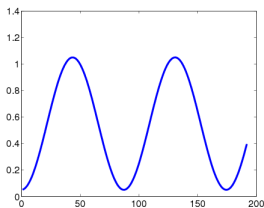
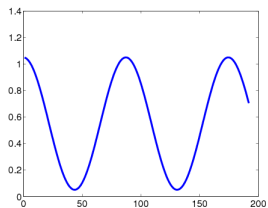
$$\mathbf{x}(\vec{r}, t) = \mathbf{e}_1(\vec{r}) + \alpha(t)\mathbf{e}_2(\vec{r}) + \gamma(t)\mathbf{e}_3(\vec{r})$$

- Introduction d'information de parcimonie.
- Approche bayésienne variationnelle.

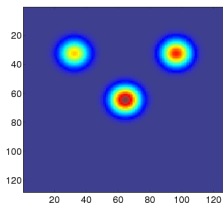
Simulation d'une coupe de la calotte polaire du soleil



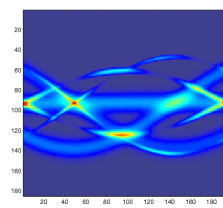
images des plumes (pics)



Dynamique d'évolution de chaque image de pics

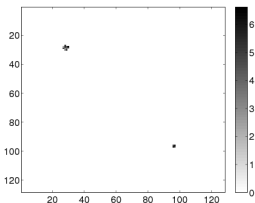
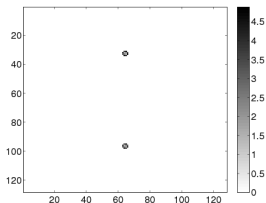


source étendue

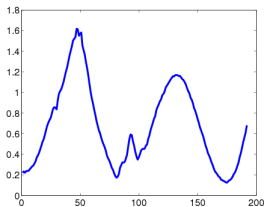
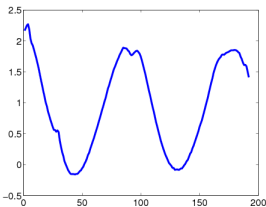


données simulées

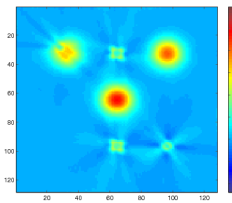
Reconstruction



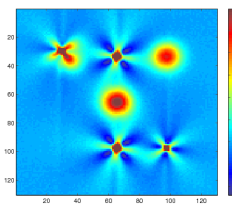
images des plumes (pics)



Dynamique d'évolution de chaque image de pics



source étendue



Rec. statique (FBP)

Reconstruction/Séparation de matériaux à partir de données polychromatiques issues d'un scanner à rayons X [Cai12]

Contexte

C. Cai (Collaboration avec le CEA S. Legoupil)

Problématique

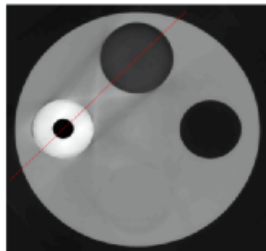
- Réduction du problème de durcissement de spectre.
- Détermination des abondances d'eau et d'os à partir de mesures à 2 tensions de sources de rayons X.

Contributions

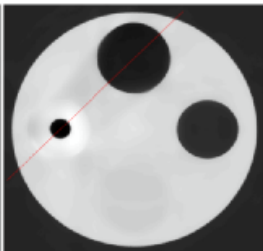
- Paramétrisation en 2 matériaux.
- Bruit gaussien de variance inconnue.
- Maximisation alternée.

Méthode de référence

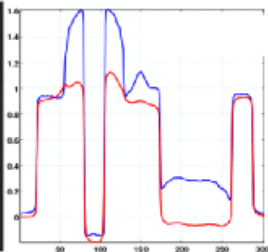
Approche proposée



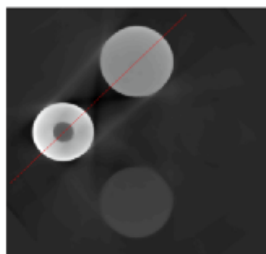
Abondance d'eau



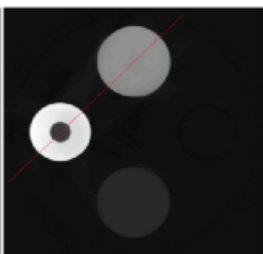
Abondance d'eau



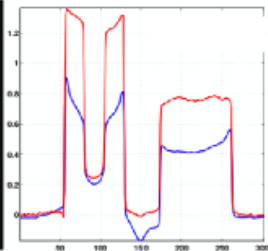
Profils



Abondance d'os



Abondance d'os



Profils

Plan :

- Curriculum Vitae
- Illustration de la démarche suivie lors de la résolution d'un problème inverse.
- **Nouvelle approche bayésienne variationnelle.**
- Illustration sur un problème astrophysique.
- Perspectives de recherche.

Problèmes inverses de grande dimension

Motivations

- Développement de la micro-électronique.
- Augmentation de la taille des données (ex : CCD $256 \times 256 \rightarrow 2048 \times 2048$).

Approches non supervisées (entièrement Bayésiennes)

$$p(\mathbf{x}, \theta_1, \theta_2 | \mathbf{y}, \mathcal{M}) = \frac{p(\mathbf{y} | \mathbf{x}, \theta_1, \mathcal{M}) p(\mathbf{x} | \theta_2, \mathcal{M}) p(\theta_1, \theta_2 | \mathcal{M})}{p(\mathbf{y} | \mathcal{M})}$$

Estimation \Rightarrow approximation de la loi *a posteriori*

- Approximation par un point : Maximum *a posteriori* (optimisation).
- Approximations stochastiques : MCMC
- Approximations analytiques : Laplace, approche bayésienne variationnelle.

État de l'art

Limitation des approches MCMC lorsque \mathbf{x} est de grande dimension

- Approche Metropolis-Hastings [**Robert97**].
 - Problème du choix de la loi de proposition \Rightarrow taux d'acceptation.
- Échantillonneur de Gibbs $\mathbf{x}^k \sim p(\mathbf{x}|\theta_1^k, \theta_2^k \mathbf{y}, \mathcal{M})$.
 - Nécessité d'inverser la matrice de covariance de \mathbf{x} à chaque itération.

Principe de l'approche bayésienne variationnelle

- Approcher la loi *a posteriori* $p(\mathbf{x}, \theta_1, \theta_2 | \mathbf{y}, \mathcal{M})$ par une loi plus simple (séparable) $q(\mathbf{x}, \theta_1, \theta_2)$ [**Hinton93**][**MacKay95**] comparable aux approches en champ moyen [**Parisi88**].
- La loi approchante q minimise la divergence de Kullback-Liebler $KL\{q||p\}$

Approche bayésienne variationnelle classique

Évidence du modèle :

- Simplifions les notations $\mathbf{w} = (\mathbf{x}, \theta_1, \theta_2)$

$$\begin{aligned}
 \log p(\mathbf{y}|\mathcal{M}) &= \log \frac{p(\mathbf{w}, \mathbf{y}|\mathcal{M})}{p(\mathbf{w}|\mathbf{y}, \mathcal{M})} \\
 &= \log \frac{q(\mathbf{w})p(\mathbf{w}, \mathbf{y}|\mathcal{M})}{q(\mathbf{w})p(\mathbf{w}|\mathbf{y}, \mathcal{M})} \\
 &= \int \log \frac{q(\mathbf{w})p(\mathbf{w}, \mathbf{y}|\mathcal{M})}{q(\mathbf{w})p(\mathbf{w}|\mathbf{y}, \mathcal{M})} q(\mathbf{w}) d\mathbf{w} \\
 &= \int q(\mathbf{w}) \log \left(\frac{p(\mathbf{w}, \mathbf{y}|\mathcal{M})}{q(\mathbf{w})} \right) d\mathbf{w} \\
 &+ \int q(\mathbf{w}) \log \left(\frac{q(\mathbf{w})}{p(\mathbf{w}|\mathbf{y}, \mathcal{M})} \right) d\mathbf{w} \\
 &= \mathcal{F}(q, \mathbf{y}) + \text{KL}\{q(\mathbf{w})\|p(\mathbf{w}|\mathbf{y}, \mathcal{M})\}
 \end{aligned}$$

- Minimiser $\text{KL}\{q(\mathbf{w})\|p(\mathbf{w}|\mathbf{y}, \mathcal{M})\} \equiv$ maximiser $\mathcal{F}(q, \mathbf{y})$ appelée énergie libre.

Approche bayésienne variationnelle classique

Loi approchante séparable $q(\mathbf{w}) = \prod_i q_i(w_i)$

Maximisation

$$\begin{aligned} q_i(w_i) &= \arg \max_{q_i(w_i)} \mathcal{F}(q, \mathbf{y}) \quad \text{s.c.} \quad \int q_i(w_i) dw_i = 1 \\ &= \frac{1}{Z_i} \exp \left[\langle \log p(\mathbf{y}, \mathbf{w} | \mathcal{M}) \rangle_{\prod_{j \neq i} q_j(w_j)} \right], \end{aligned}$$

avec $\langle h(\mathbf{w}) \rangle_{\prod_{j \neq i} q_j^{(k)}(w_j)} = \int h(\mathbf{w}) \prod_{j \neq i} q_j^{(k)}(w_j) dw_j$.

Algorithme de point fixe (maximisation alternée)

$$q_i^{(k+1)}(w_i) = \frac{1}{Z_i} \exp \left[\langle \log p(\mathbf{y}, \mathbf{w} | \mathcal{M}) \rangle_{\prod_{j < i} q_j^{(k+1)}(w_j) \prod_{j > i} q_j^{(k)}(w_j)} \right]$$

Lois conjuguées

- Loi *a posteriori* dans une famille connue \Rightarrow mise à jour uniquement des paramètres des lois.

Nouvelle approche bayésienne variationnelle adaptée aux problèmes de grande dimension

Limitation de l'approche bayésienne variationnelle

Difficulté pour les problèmes de grandes dimension :

- faible séparabilité \rightarrow inversion d'une grande matrice de covariance.
- forte séparabilité \rightarrow approches type Gauss-Seidel (lent).

Idées :

- Utiliser une approche de type gradient.
- Déterminer un pas optimal.

Difficultés :

- L'espace des fonctions à densité de probabilité n'a pas de structure d'espace de Hilbert.
- Difficultés d'appliquer la contrainte de positivité dans un espace de dimension infinie (KKT pas applicable).

Approche gradient pour le bayésien variationnel

[Frayse12]

Mise à jour exponentielle

- En dimension finie [Kivinen97].
- Mise à jour des densités de probabilité (dimension infinie) :

$$\begin{aligned}
 q^{k+1}(\mathbf{w}) &= K(\alpha_k) \exp(\alpha_k df(q^k, \mathbf{w})) q^k(\mathbf{w}) \\
 &= \tilde{K}(\alpha_k) q^k(\mathbf{w}) \left(\prod_i \frac{q_i^r(w_i)}{q_i^k(w_i)} \right)^{\alpha_k}
 \end{aligned}$$

avec

$$q_i^r(w_i) = \exp \left(\langle \log p(\mathbf{y}, \mathbf{w} | \mathcal{M}) \rangle_{\prod_{j \neq i} q_j^k(w_j)} \right)$$

Détermination de α_k

$$\alpha_k = \arg \max_{\alpha} \mathcal{F}(q^{\alpha}, \mathbf{y})$$

Algorithmme

- 1 Initialisation $q^0(\boldsymbol{w})$
- 2 Calcul de q^r
- 3 Calcul d'un pas sous optimum (Approximation de Taylor de $\mathcal{F}(q^\alpha)$ en q^k)
- 4 Mise à jour de la loi :

$$q^{k+1} = \tilde{K}(\alpha_k) q^k(\boldsymbol{w}) \left(\prod_i \frac{q_i^r(w_i)}{q_i^k(w_i)} \right)^{\alpha_k}$$

- 5 Retourner à l'étape 2 jusqu'à convergence.

Plan :

- Curriculum Vitae
- Illustration de la démarche suivie lors de la résolution d'un problème inverse.
- Nouvelle approche bayésienne variationnelle.
- **Illustration sur un problème astrophysique.**
- Perspectives de recherche.

Application à la séparation de composantes astrophysiques

[Ayasso12]

Paramétrisation

$$\mathbf{x} = \mathbf{s} + \mathbf{p}.$$

- \mathbf{s} composante étendue (ex : fond diffus cosmologique).
- \mathbf{p} composante impulsionnelle.

Vraisemblance

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \epsilon = \mathbf{H}(\mathbf{s} + \mathbf{p}) + \epsilon$$

- ϵ blanc bruit Gaussien de moyenne \mathbf{o} (offset) et d'inverse variance γ_b .

$$p(\mathbf{y}|\mathbf{p}, \mathbf{s}, \mathbf{o}, \gamma_b, \mathcal{M}) \propto \exp\left[-\frac{1}{2}\gamma_b\|\mathbf{y} - \mathbf{H}\mathbf{s} - \mathbf{H}\mathbf{p} - \mathbf{o}\|^2\right]$$

Loi *a priori* et *a posteriori*

Informations *a priori*

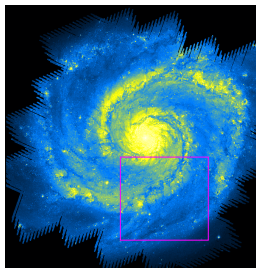
- *a priori* gaussien iid sur les offsets \mathbf{o} de moyenne, \mathbf{o}_m et de précision γ_o .
- *a priori* gaussien corrélé sur \mathbf{s} (information de douceur).
- *a priori* Student-t iid sur \mathbf{p} .

$$\text{Student}(p_i | 0, \lambda, \nu) = \int_{\mathbb{R}} \mathcal{N}(0, (\lambda a_i)^{-1}) \text{Gamma}\left(\frac{\nu}{2}, \frac{\nu}{2}\right) da_i$$

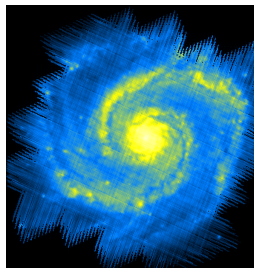
loi *a posteriori*

$$\begin{aligned} p(\mathbf{s}, \mathbf{p}, \mathbf{a}, \mathbf{o} | \mathbf{y}, \gamma_b, \gamma_s, \nu, \lambda, \mathbf{o}_m, \gamma_o, \mathcal{M}) \propto & \exp -\frac{1}{2} [\gamma_b \|\mathbf{y} - \mathbf{H}\mathbf{s} - \mathbf{H}\mathbf{p} - \mathbf{R}\mathbf{o}\|^2 \\ & + \gamma_s \|\mathbf{D}\mathbf{s}\|^2 + \gamma_o \|\mathbf{o} - \mathbf{o}_m\|^2] \prod_i \sqrt{a_i} \exp \left[-\frac{1}{2} p_i^2 a_i \lambda \right] a_i^{\frac{\nu}{2} - 1} \exp \left[-\frac{1}{2} \nu a_i \right] \end{aligned}$$

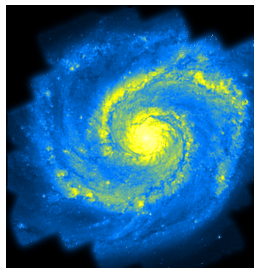
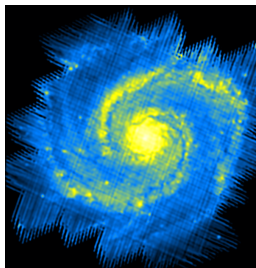
Simulation



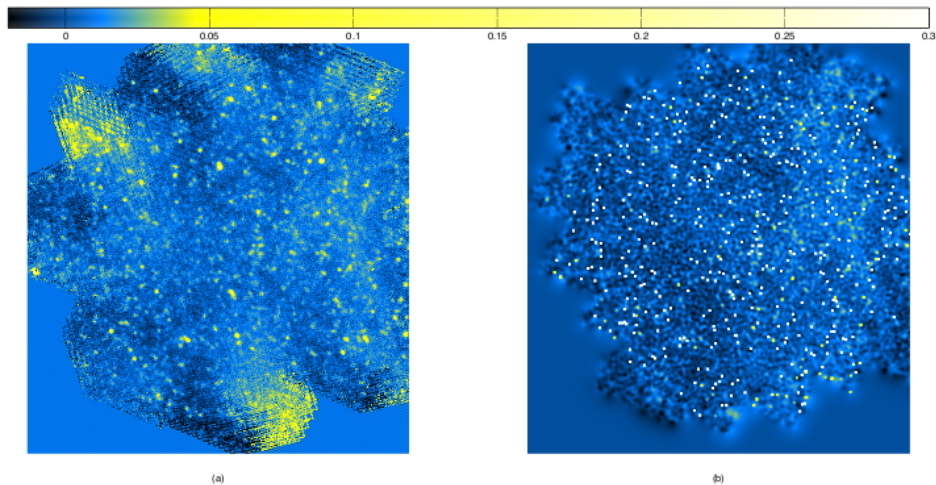
(a)



(b)



Reconstruction du champ 17p7323



Plan :

- Curriculum Vitae
- Illustration de la démarche suivie lors de la résolution d'un problème inverse.
- Nouvelle approche bayésienne variationnelle.
- Illustration sur un problème astrophysique.
- **Perspectives de recherche.**

Perspectives de recherche

- Inversion de données astrophysiques
- Inversion de données tomographiques
- Méthodologie bayésienne variationnelle
- Sélection de modèles

Inversion de données astrophysiques

Inversion de données issues d'un FTS

- Spectromètre à transformée de Fourier.
- Interféromètre de Michelson.
- Besoin de modélisation.

Échantillonnage adaptatif

- « Experimental Design » [Seeger11].

$$\begin{aligned} \mathbf{y} &= \mathbf{H}\mathbf{x} + \epsilon \\ \mathbf{y}^* &= \mathbf{H}^*\mathbf{x} \end{aligned}$$

- Déterminer \mathbf{H}^* le plus informatif sur \mathbf{x} .
- Calcul de $p(\mathbf{y}^* | \mathbf{H}^*, \mathbf{H}, \mathbf{y}) = \int p(\mathbf{y}^* | \mathbf{H}^*, \mathbf{x}) p(\mathbf{x} | \mathbf{H}, \mathbf{y}) d\mathbf{x}$.
- Besoin d'une approche bayésienne variationnelle rapide.

Inversion de données tomographiques

Séparation de matériaux en utilisant le rayonnement diffusé

- Interaction rayonnement matière (effet photoélectrique, diffusion Compton).
- Utilisation du rayonnement diffusé pour mieux discriminer les matériaux.

Traitement récursif de données tomographiques 3D +T

- Le traitement récursif \Rightarrow propagation de l'information [**Smidl08**].

$$p(\mathbf{x}_n | \mathbf{y}_{0:n-1}) = \int p(\mathbf{x}_n | \mathbf{x}_{n-1}) p(\mathbf{x}_{n-1} | \mathbf{y}_{0:n-1}) d\mathbf{x}_{n-1}$$

$$p(\mathbf{x}_n | \mathbf{y}_{0:n}) \propto p(\mathbf{y}_n | \mathbf{x}_n) p(\mathbf{x}_n | \mathbf{y}_{0:n-1})$$

- Utilisation de nouvelles approches bayésiennes variationnelles.
- Premier résultat dans le cas d'un *a posteriori* gaussien [**AitElFquih11**].
- Alternative au filtrage particulaire.

Méthodologie bayésienne variationnelle

Amélioration des approches de descente de gradient

- Amélioration de la direction de descente (thèse Y. Zheng).
- Pas de notion de conjugaison dans l'espace de densités de probabilité.

Applications aux dictionnaires à base d'ondelettes (collaboration avec J.-C. Pesquet et C. Chaux)

- *a priori* de parcimonie sur les coefficients du dictionnaire.
- Degré de parcimonie différent pour chaque type de coefficient.
- Estimation conjointe d'un grand nombre d'hyperparamètres.

Sélection de modèles

- Choix entre un faible nombre de modèles.
- Approche non paramétrique.

Bibliographie



G. Hinton and D. van Camp.

Keeping neural networks simple by minimizing the description length of the weights.
In Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory, pp. 5–13.
ACM Press, New York, NY, 1993.



J. Kivinen and M. Warmuth.

Exponentiated gradient versus gradient descent for linear predictors.
Information and Computation, 132(1) :1–63, 1997.



D. J. C. MacKay.

Ensemble learning and evidence maximization.
<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.54.4083>, 1995.



G. Parisi.

Statistical field theory.
Addison Wesley, Redwood City, CA, 1988.



C. Robert.

Simulations par la méthode MCMC.
Economica, Paris, 1997.



M. Seeger and Nickisch H.

Large scale bayesian inference and experimental design for sparse linear models.
SIAM Journal on Imaging Sciences, 4(1) :166–199, 2011.



V. Smidl and A. Quinn.

Variational bayesian filtering.
IEEE Transactions on Signal Processing, 56(10) :5020–5030, octobre 2008.

Articles en rapport avec la présentation



A. Fraysse et T. Rodet,

« A measure-theoretic variational bayesian algorithm for large dimensional problems. »,
Soumis à *SIAM Imaging Sciences*, 36 pages, 2012.



C. Cai, T. Rodet, S. Legoupil et A. Mohammad-Djafari,

« Bayesian reconstruction of water and bone fractions in polychromatic multi-energy computed tomography »,
en révision *Medical Physics*, 23 pages ,2012.



H. Ayasso, T. Rodet et A. Abergel,

« A variational bayesian approach for unsupervised super-resolution using mixture models of point and smooth sources applied to astrophysical mapmaking »,
accepté dans *Inverse Problems*, 33 pages, 2012.



B. Ait el Fquih et T. Rodet,

« Variational bayesian Kalman filtering in dynamical tomography »,
in *Proceedings of the International Conference on Acoustic, Speech and Signal Processing*, Prague,
mai 2011, n°ld 2548, pp. 4004 - 4007.



B. Kherrab, T. Rodet et J. Idier,

« Solving a problem of sources separation stemming from a linear combination : Applied to the 3D reconstruction of the solar atmosphere »,
in *Proceedings of the International Conference on Acoustic, Speech and Signal Processing*, Dallas,
USA, 2010, pp. 1338 – 1341.



T. Rodet, F. Orieux, J.-F. Giovannelli et A. Abergel,

« Data inversion for over-resolved spectral imaging in astronomy »,
IEEE Journal of Selected Topics in Signal Processing, vol. 2, n°5, pp. 802–811, octobre 2008.