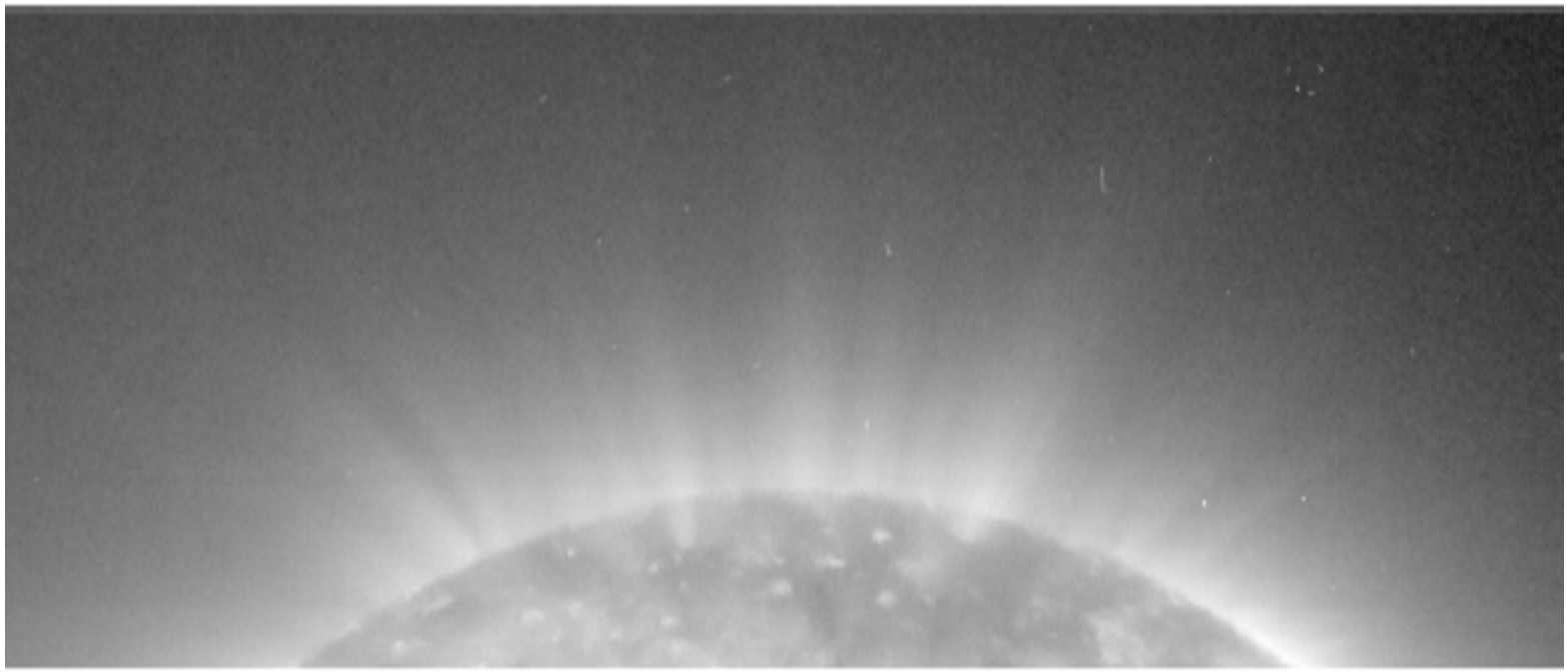
SOLVING A PROBLEM OF SOURCES SEPARATION STEMMING FROM A LINEAR COMBINATION: APPLIED TO THE 3D RECONSTRUCTION OF THE SOLAR ATMOSPHERE

Thomas RODET

Laboratoire des Signaux et Systèmes (CNRS-SUPELEC-UPS)
Groupe Problèmes Inverses
Thomas.Rodet@lss.supelec.fr

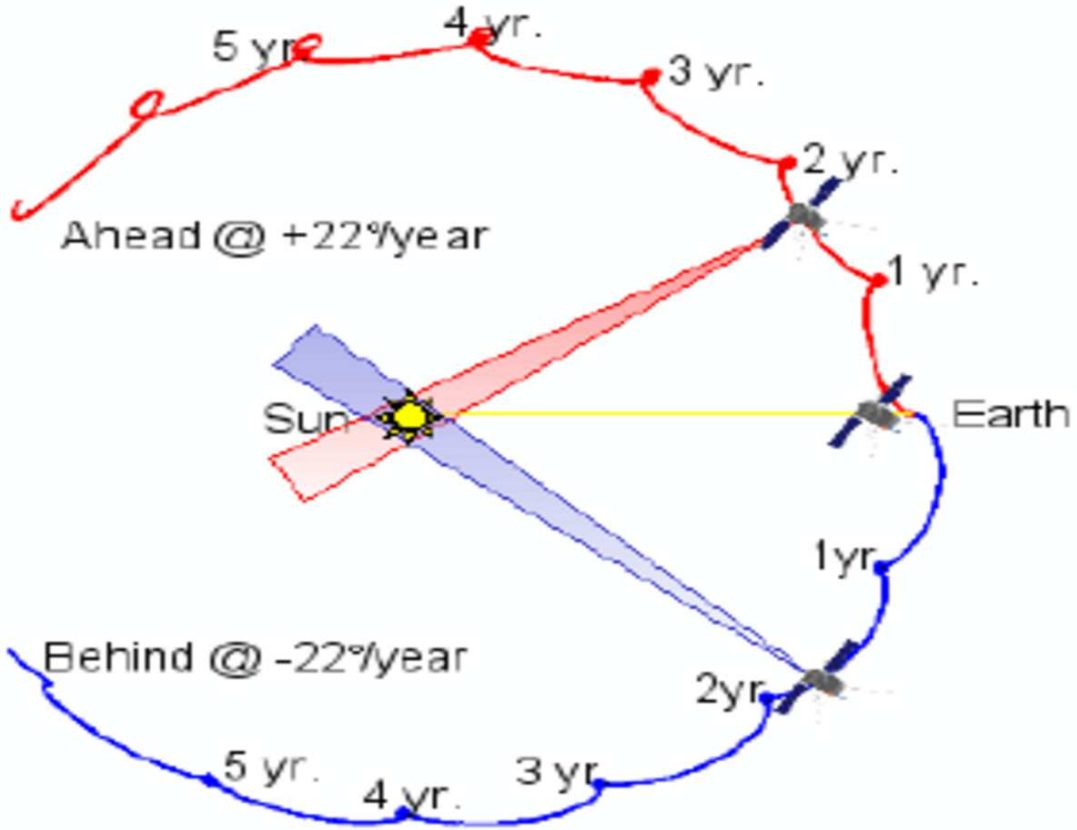
ISSSMA Workshop Medical imaging
8 juin 2012

Contexte scientifique :



- Reconstruire la structure 3D de la couronne solaire.
- Connaître la structure des plumes polaires.
- Utiliser les trois longueurs (171,195,284 angström) par les instruments pour déterminer la densité et la température.

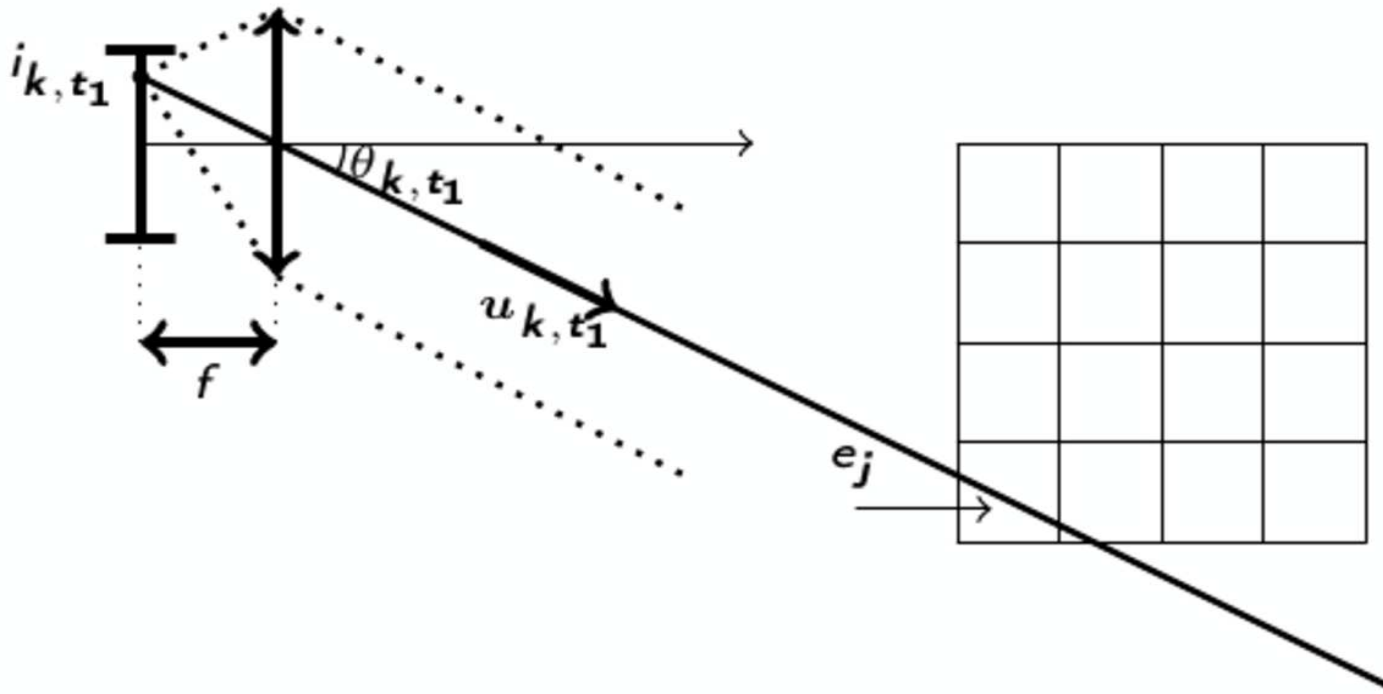
La mission STEREO



- Lancement en octobre 2006.
- Deux satellites identiques qui se décalent chaque année.

La reconstruction de l'atmosphère solaire UV, un problème tomographique

- L'astmosphère est transparente pour cette longueur d'onde.
- On mesure l'intégrale de l'émission selon la ligne de visé.
- Le soleil fait un demi tour en 13 jours.
- Géométrie équivalente à une géométrie conique avec un détecteur sur une calotte sphérique.



Problématiques :

- Le soleil fait une demi rotation en 13 jours.
- À un instant donné, on ne mesure qu'une ou deux projections.
- Les structures ont une évolution trés rapide, elles peuvent être visible uniquement 2 jours.
- Nécéssité de reconstruire une séquence de volumes (3D+T).
- Problème très mal posés 50 fois d'inconnues que de données.
- Obligation d'introduire beaucoup d'information *a priori*.

Plan :

- État de l'art et inférence bayésienne.
- Première approche (thèse N. Barbey)
- Résultats
- Deuxième approche
- Nouvelle méthode bayésienne variationnelle
- Résultats
- Conclusions

État de l'art :

- Pas d'approche analytique.
- B. Inhester 2006 : utilisation d'*a priori* suivant un modèle physique de MHD (MagnétoHydroDynamique)
- R. Frazin 2002 : résolution du problème 3D+T introduction d'un *a priori* temporelle de douceur, et un *a priori* spatial TV.
- M. Butala 2008 : développement d'une approche rapide de filtrage de Kalman stochastique Empirique.

Inférence Bayésienne

Le modèle direct linéaire :

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x}$$

La règle de Bayes :

$$p(\boldsymbol{x}|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x})}{p(\boldsymbol{y})}$$

- $p(\boldsymbol{y}|\boldsymbol{x})$: la vraisemblance (bruit + erreur de modélisation),
- $p(\boldsymbol{x})$: *a priori*,
- $p(\boldsymbol{y})$: Évidence du modèle (facteur de normalisation).
- $p(\boldsymbol{x}|\boldsymbol{y})$: *a posteriori*, densité résumant toute l'information disponible sur notre problème.
- Estimation :
  - MAP : $\hat{\boldsymbol{x}} = \arg\max p(\boldsymbol{x}|\boldsymbol{y})$
  - EAP : $\hat{\boldsymbol{x}} = E\{\boldsymbol{x}\}_{p(\boldsymbol{x}|\boldsymbol{y})}$

Plan :

- État de l'art et inférence bayésienne.
- **Première approche (thèse N. Barbey)**
- Résultats
- Deuxième approche
- Nouvelle méthode bayésienne variationnelle
- Résultats
- Conclusions

Première approche (Thèse N.Barbey)

- Necessité de contraindre fortement la solution.
- Hypothèse que la morphologie des plumes ne varie pas.
- Uniquement l'intensité des plumes varie.

$$\boldsymbol{x}(\vec{r}, t) = e(\vec{r}). \times \boldsymbol{L}\boldsymbol{\gamma}(t)$$

- $\boldsymbol{L}$ matrice correspondant aux zones définissant les plumes (zones possédant la même évolution temporelle).
- $\boldsymbol{\gamma}(t)$ vecteur rassemblant un faible nombre d'évolutions temporelles (4-5).

Cadre bayésien

Vraisemblance :

$$\boldsymbol{y} = \boldsymbol{H}e(\vec{r}). \times \boldsymbol{L}\boldsymbol{\gamma}(t) + \boldsymbol{b}$$

- $\boldsymbol{b}$ bruit blanc gaussien.
- Vraisemblance gaussienne variance σ_b^2.

Informations *a priori*

- *a priori* spatiale introduction d'informations de douceur

$$p(\boldsymbol{e}|\sigma_e) \propto \exp(-\frac{\|\boldsymbol{D}\boldsymbol{e}\|^2}{2\sigma_e^2})$$

- *a priori* temporelle : douceur et rappel vers 1.

$$p(\boldsymbol{\gamma}|\sigma_\gamma, \sigma_1) \propto \exp(-\frac{\|\boldsymbol{D}\boldsymbol{\gamma}\|^2}{2\sigma_\gamma^2} - \frac{\|\boldsymbol{\gamma} - \mathbf{1}\|^2}{2\sigma_1^2})$$

Optimisation

Loi *a posteriori* :

$$p(\boldsymbol{e}, \boldsymbol{\gamma}|\boldsymbol{y}) \propto \exp(-\frac{\|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{e}. \times \boldsymbol{L}\boldsymbol{\gamma}\|}{2\sigma_b^2} - \frac{\|\boldsymbol{D}\boldsymbol{e}\|^2}{2\sigma_e^2} - \frac{\|\boldsymbol{D}\gamma_b\|^2}{2\sigma_\gamma^2} - \frac{\|\boldsymbol{\gamma} - \mathbf{1}\|^2}{2\sigma_1^2})$$

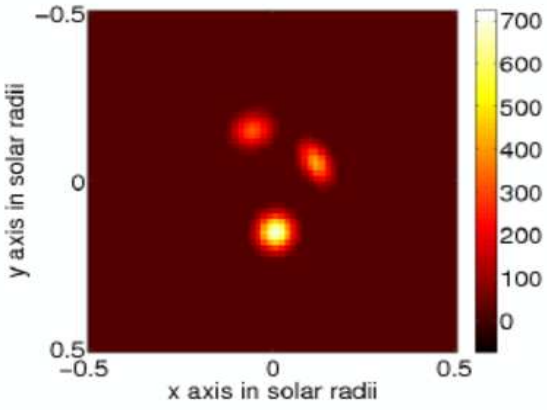
- La loi jointe $p(\boldsymbol{e}, \boldsymbol{\gamma}|\boldsymbol{y})$ n'est pas gaussienne
- Les lois conditionnelles $p(\boldsymbol{e}|\boldsymbol{\gamma}, \boldsymbol{y})$ et$p(\boldsymbol{\gamma}|\boldsymbol{e}, \boldsymbol{y})$ sont gaussiennes.

Optimisation alternée

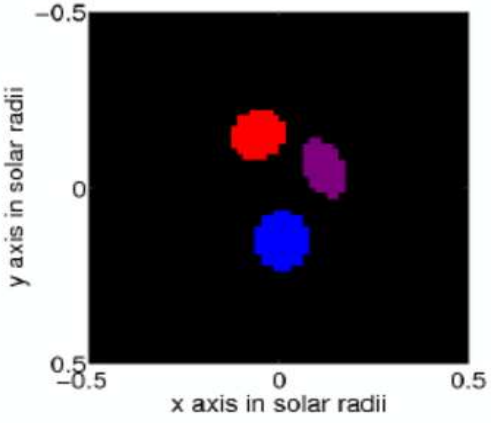
1. Initialisation
2. $\hat{\boldsymbol{e}} = \arg\max p(\boldsymbol{e}|\boldsymbol{\gamma}, \boldsymbol{y})$
3. $\hat{\boldsymbol{\gamma}} = \arg\max p(\boldsymbol{\gamma}|\boldsymbol{e}, \boldsymbol{y})$
4. retour à l'étape 2 jusqu'à convergence.

Résultats sur données simulées

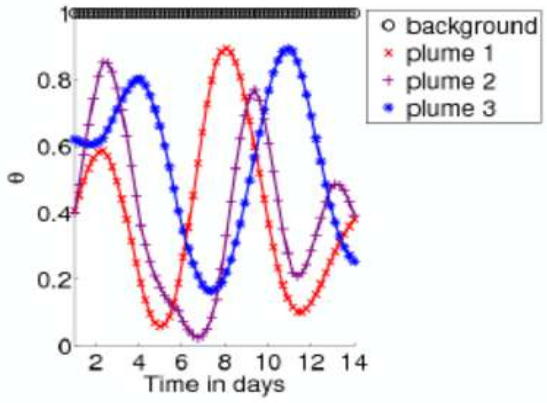
e simulé



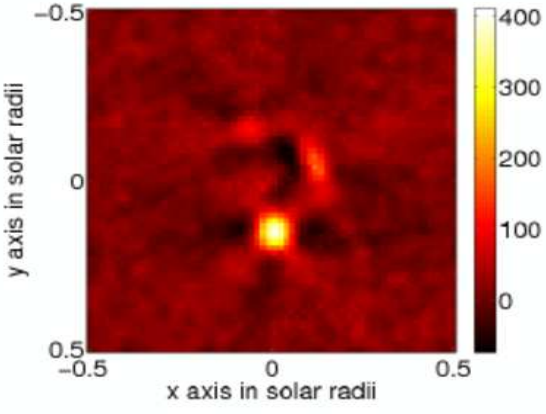
Zones d'évolution L



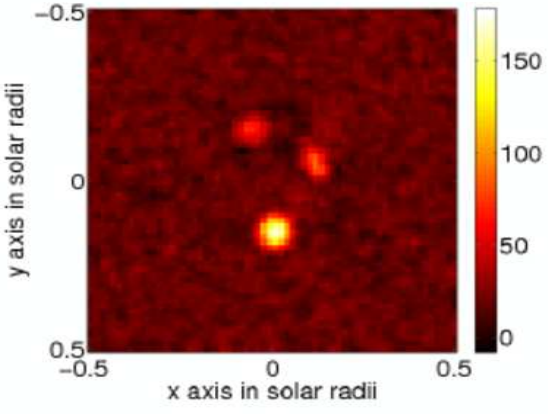
gain γ simulé



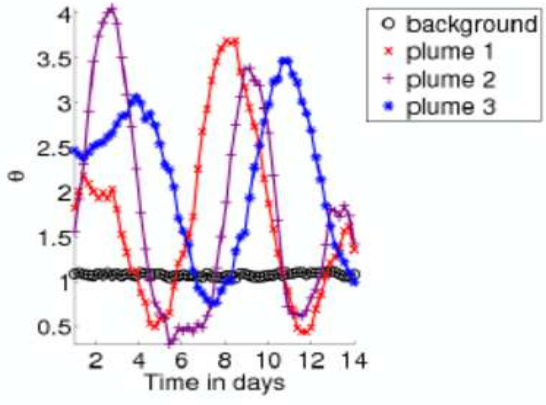
e reconstruit en statique



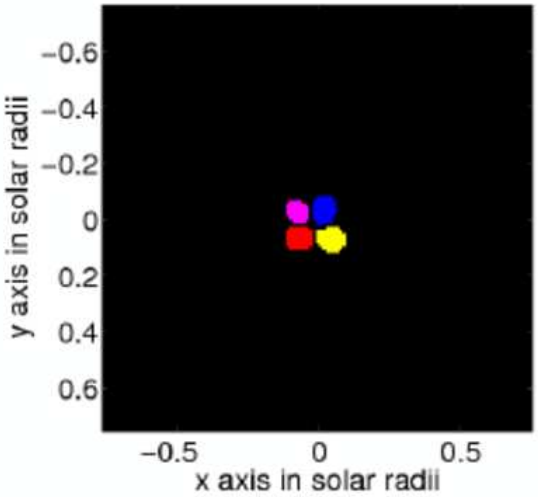
e reconstruit avec évolution temporelle

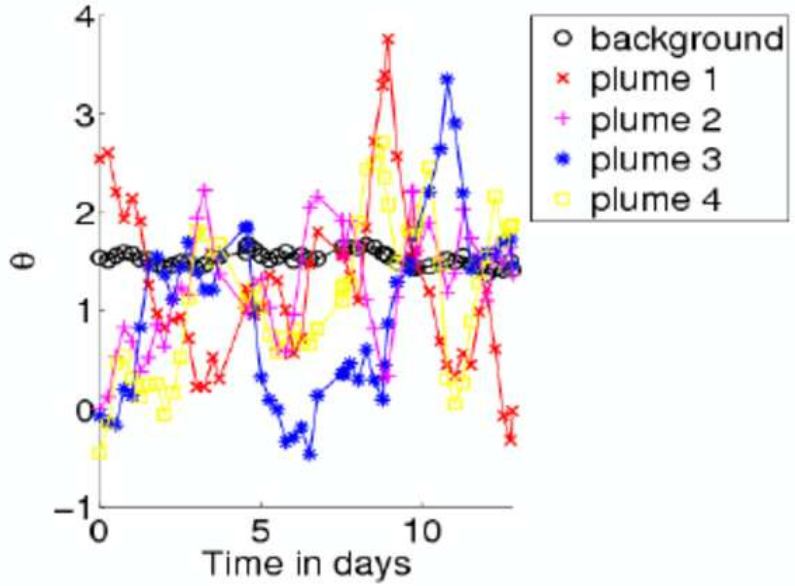


gain γ reconstruit

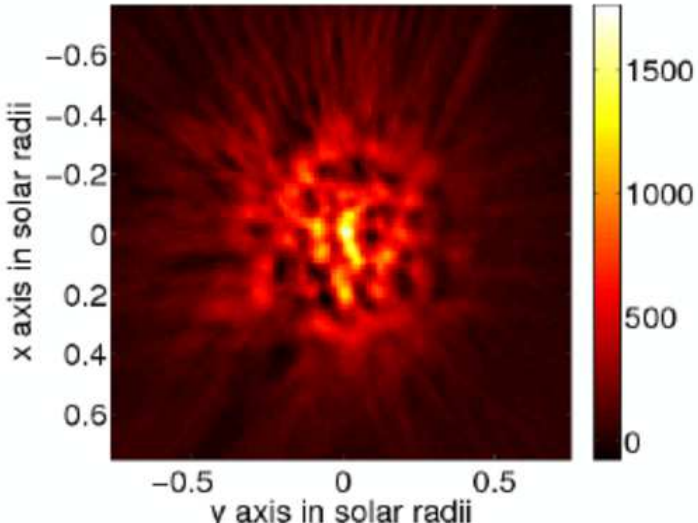


Résultats sur données réelles

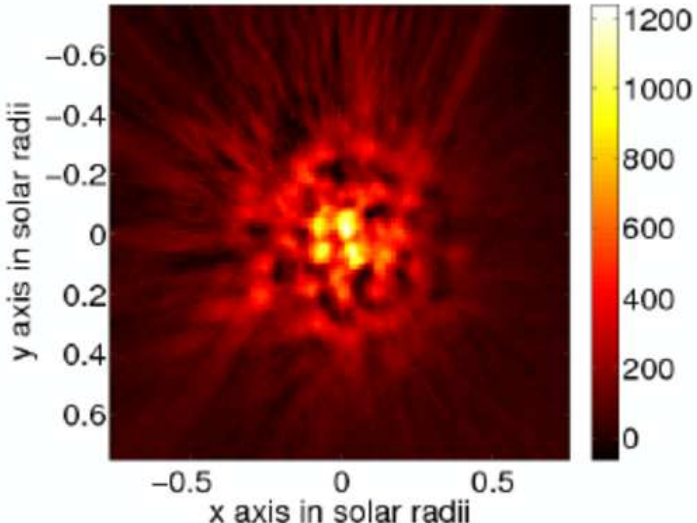




zones d'évolution temporelle L



gain évoluant temporellement γ



Conclusions :

Resultats obtenus :

- Développement d'une paire projecteur, rétroprojecteur efficace (approche de siddon)
- Introduction d'une nouvelle paramétrisation permettant de suivre des evolutions rapides de l'objet

Limitations :

- L'approche est très spécifique à un type d'évolution temporelle.
- On doit connaître les zones d'évolution associées aux plumes.

Plan :

- État de l'art et inférence bayésienne.
- Première approche (thèse N. Barbey)
- Résultats
- **Deuxième approche**
- Nouvelle méthode bayésienne variationnelle
- Résultats
- Conclusions

Deuxième approche (B. Kherrab)

Paramétrisation comme un mélange de sources

- Modèle direct :

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

- On considère que $\boldsymbol{x}$ est un mélange instantané de trois sources $\boldsymbol{e}_1, \boldsymbol{e}_2$ et $\boldsymbol{e}_3$.
- Écriture matricielle :

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{e}$$

avec

$$\boldsymbol{e} = \begin{pmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \\ \boldsymbol{e}_3 \end{pmatrix}, \ \boldsymbol{A} = \begin{pmatrix} 1 & \boldsymbol{\alpha}_1 & \boldsymbol{\gamma}_1 \\ \vdots & \vdots & \vdots \\ 1 & \boldsymbol{\alpha}_T & \boldsymbol{\gamma}_T \end{pmatrix}.$$

- On multiplie le nombre d'inconnues par trois.

Inférence bayésienne

Vraisemblance :

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{A}\boldsymbol{e} + \boldsymbol{b}$$

- $\boldsymbol{b}$ bruit blanc Gaussien.
- Vraisemblance gaussienne.

Informations *a priori*

- *a priori* temporelle de douceur

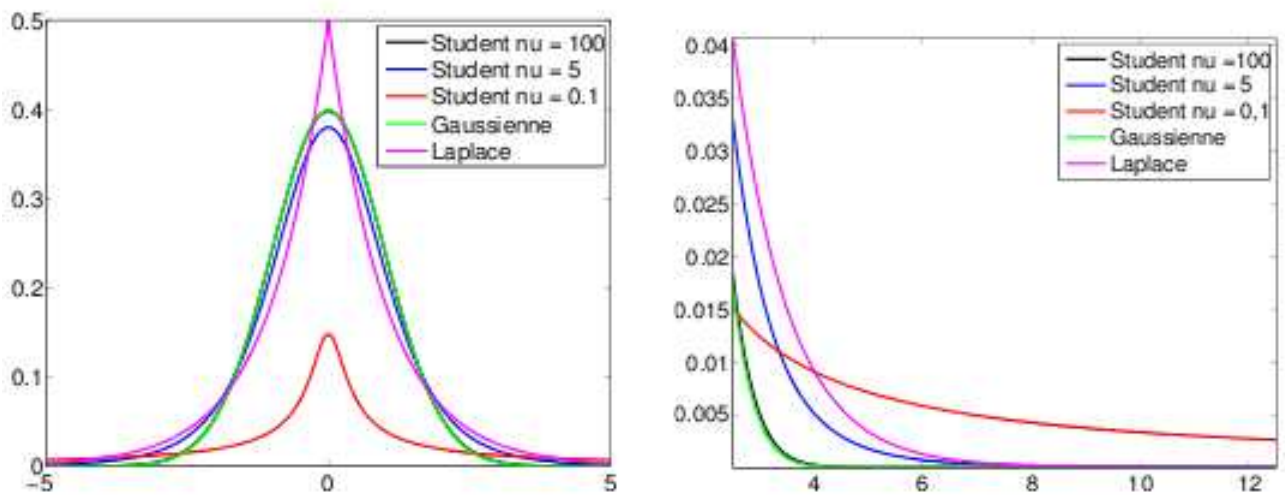
$$p(\boldsymbol{\gamma}, \boldsymbol{\alpha}|\sigma_t) \propto \exp(-\frac{\|\boldsymbol{D}\boldsymbol{\gamma}\|^2 + \|\boldsymbol{D}\boldsymbol{\alpha}\|^2}{2\sigma_t^2})$$

a priori favorisant la parcimonie

Utilisation d'une loi Student-T séparable à lourde queue pour e_2 et e_3.

$$\begin{aligned} p(e_\ell(i)|\lambda,\nu) &= St(e_\ell(i)|\lambda,\nu) \\ &\propto \int \mathcal{N}(e_\ell(i)|0,\beta_\ell(i)*\lambda)\mathcal{G}(\beta_\ell(i)|\frac{\nu}{2},\frac{\nu}{2})d\beta_\ell(i) \end{aligned}$$

Introduction d'un *a priori* de douceur pour e_1 source étendue(interplume).



Loi *a posteriori* jointe

- La loi de Student-t est conjuguée avec la vraisemblance gaussienne.

$$St(e|\lambda,\nu) \propto \int \mathcal{N}(e|0,\beta * \lambda)\mathcal{G}(\beta|\frac{\nu}{2},\frac{\nu}{2})\mathrm{d}\beta$$

- Résolution du problème étendu (Estimation de $\boldsymbol{\beta}$)
- Problème avec 5 fois plus d'inconnues.
- Loi *a posteriori*

$$\begin{aligned} &p(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_2,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\alpha},\boldsymbol{\gamma}|\boldsymbol{y},\sigma_b,\sigma_t,\sigma_e,\lambda,\nu) \\ \propto\ &p(\boldsymbol{y}|\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3,\boldsymbol{\alpha},\boldsymbol{\gamma},\sigma_b)p(\boldsymbol{e}_1|\sigma_e)p(\boldsymbol{e}_2|\boldsymbol{\beta}_2,\lambda)p(\boldsymbol{\beta}_2|\nu) \\ &p(\boldsymbol{e}_3|\boldsymbol{\beta}_3,\lambda)p(\boldsymbol{\beta}_3|\nu)p(\boldsymbol{\alpha}|\sigma_t)p(\boldsymbol{\gamma}|\sigma_t) \end{aligned}$$

Estimation :

- MAP, EAP, MCMC ??

Approche bayésienne variationnelle rapide (Coll. A. Fraysse)

Modèle direct : $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{\epsilon}$

Approches non supervisées (entièrement Bayésiennes)

$$p(\boldsymbol{x}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2 | \boldsymbol{y}, \mathcal{M}) = \frac{p(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{\theta}_1, \mathcal{M}) p(\boldsymbol{x}|\boldsymbol{\theta}_2, \mathcal{M}) p(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2|\mathcal{M})}{p(\boldsymbol{y}|\mathcal{M})}$$

Approcher la loi *a posteriori* p par une loi plus simple q

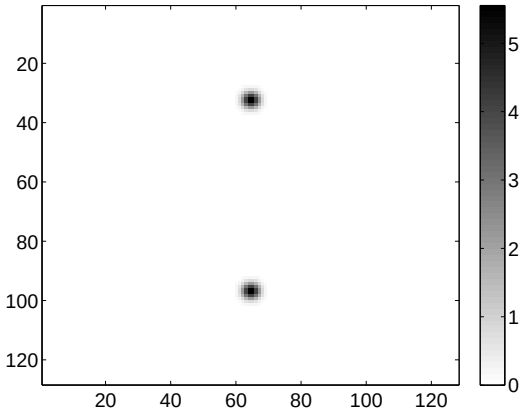
$$q(\boldsymbol{x}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \approx p(\boldsymbol{x}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2 | \boldsymbol{y}, \mathcal{M})$$

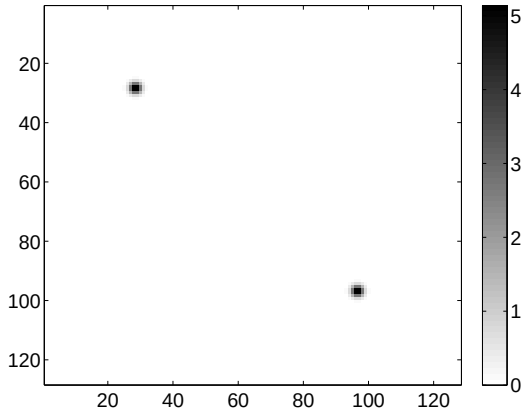
- q minimise la divergence de Kullback-Liebler $\mathrm{KL}\{q\|p\}$ **[MacKay95]**.

Contributions méthodologiques

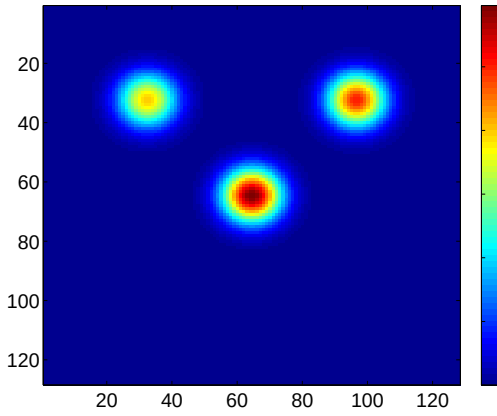
- Nouvel algorithme d'optimisation fonctionnelle basé sur un gradient exponentiel.
- **[Fraysse13]** SIAM Imaging sciences, **[Ayasso12]** dans Inverse

Simulation

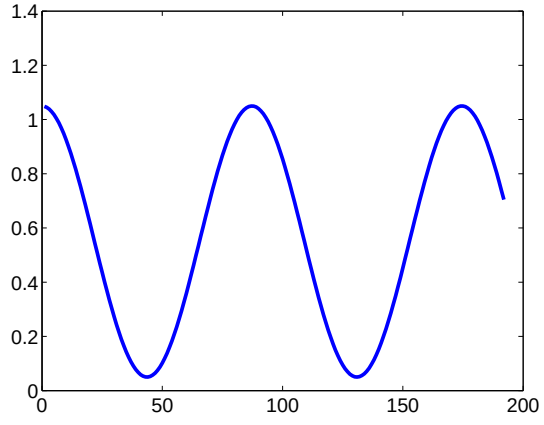


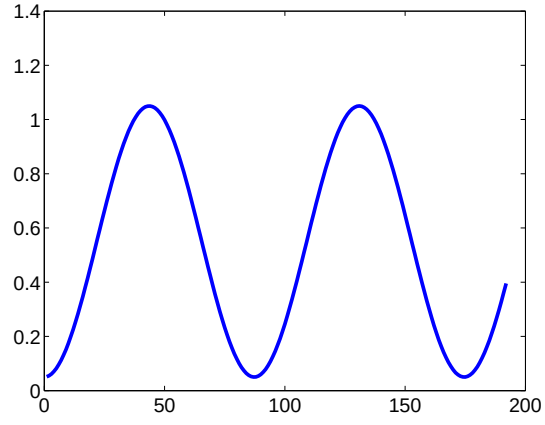


images des plumes (pics)

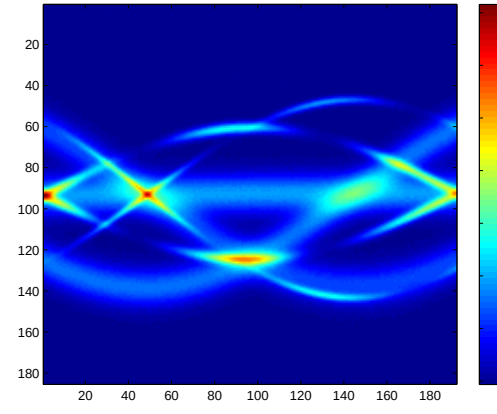


source étendue



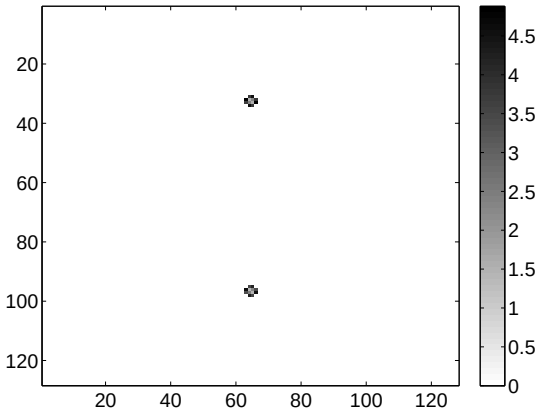


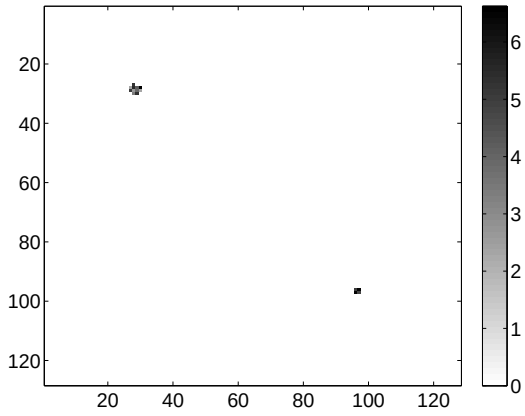
Colonnes de la matrice de mélange $\boldsymbol{A}$



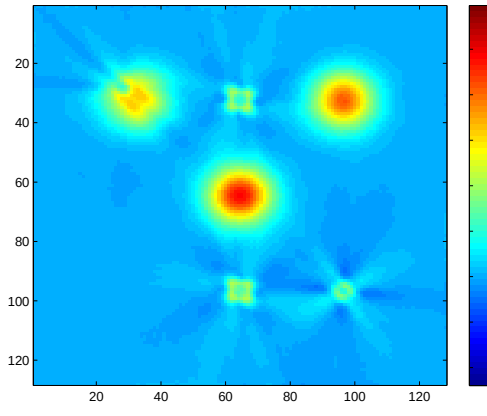
données simulées

Result

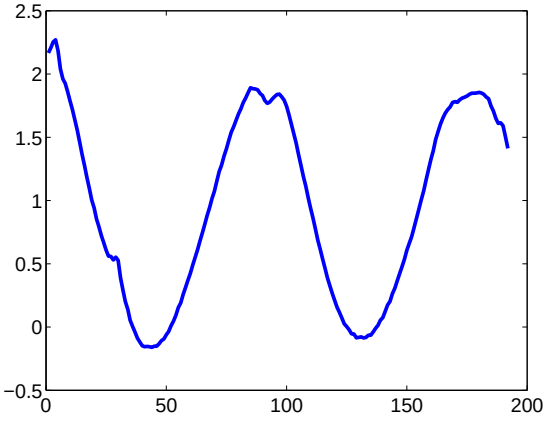


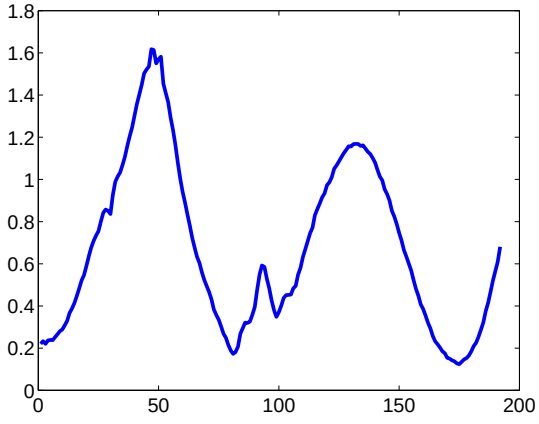


images des plumes (pics)

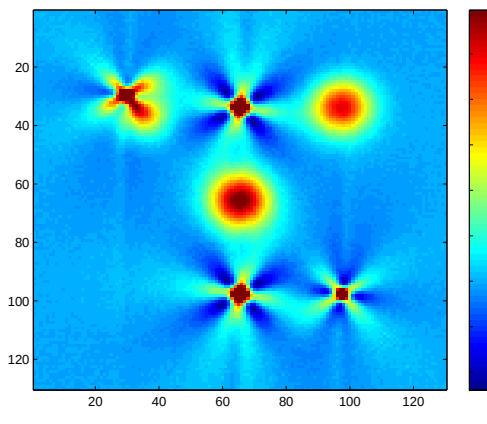


source étendue





Colonnes de la matrice de mélange $\boldsymbol{A}$



données simulées

Conclusions

Avantages :

- Résolution d'un probléme d'inversion séparation.
- Estimation de la fonction densité de probabilités.
- Obtention d'un valeur de l'évidence du modèle (possibilité de choix le nombre de sources),
- Méthode bien plus rapide que les approches MCMC.

Incovénients :

- La fonction n'est pas convexe vis à vis des paramètres.
- On doit utiliser des *a priori* conjugués

Perspectives :

- Faire un traintement au fil de l'eau (filtrage particulaire)
- Appliquer l'approche en imagerie médicale.

D. J. C. MacKay.
Ensemble learning and evidence maximization.
http ://citeseerx.ist.psu.edu/viewdoc/summary ?doi=10.1.1.54.4083, 1995.

A. Fraysse et **T. Rodet**,
« A measure-theoretic variational bayesian algorithm for large dimensional problems. »,
Soumis à *SIAM Imaging Sciences*, pp. 36, 2013.

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 »,
Inverse Problems, vol. 248, n°2, pp. 33, décembre 2012.