



**HAL**  
open science

## Estimation des paramètres d'un prior Bernoulli-Gaussien en problèmes inverses

Pierre Barbault, Matthieu Kowalski, Charles Soussen

► **To cite this version:**

Pierre Barbault, Matthieu Kowalski, Charles Soussen. Estimation des paramètres d'un prior Bernoulli-Gaussien en problèmes inverses. GRETSI 2022 - XXIIIXème Colloque Francophone de Traitement du Signal et des Images, Sep 2022, Nancy, France. hal-03718426

**HAL Id: hal-03718426**

**<https://centralesupelec.hal.science/hal-03718426v1>**

Submitted on 8 Jul 2022

**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.

# Estimation des paramètres d'un prior Bernoulli-Gaussien en problèmes inverses

Pierre BARBAULT, Matthieu KOWALSKI, Charles SOUSSEN

Laboratoire des Signaux et Systèmes  
3 rue Joliot Curie, 91190 Gif-sur-Yvette, France

pierre.barbault@centralesupelec.fr, matthieu.kowalski@universite-paris-saclay.fr  
Charles.SousSEN@centralesupelec.fr

**Résumé** – Le modèle Bernoulli-Gaussien (BG) est un modèle classique des signaux parcimonieux. Nous proposons une méthode d'estimation automatique des paramètres BG dans le cadre des problèmes inverses. L'originalité vient de la reformulation du problème en deux problèmes emboîtés de déconvolution (sans contrainte) et de débruitage parcimonieux, via l'introduction d'un signal latent. La démarche adoptée repose sur une méthode d'espérance-minimisation pour le débruitage parcimonieux non supervisé, puis son extension à des problèmes inverses. L'algorithme proposé est évalué sur des problèmes de déconvolution, et comparé à des approches supervisées.

**Abstract** – The Bernoulli-Gaussian (BG) model is a classical model for sparse signals. We propose a method for automatic estimation of BG parameters in the context of inverse problems. The originality comes from the reformulation of the problem in two nested problems of deconvolution (without constraint) and sparse denoising, via the introduction of a latent signal. The approach adopted is based on an expectation-minimization method for unsupervised parsimonious denoising, and on its extension to more involved inverse problems. The proposed algorithm is evaluated on deconvolution problems and compared to supervised approaches.

## 1 Introduction

L'analyse parcimonieuse est devenue un incontournable pour résoudre des problèmes inverses linéaires de la forme

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{e} \quad (1)$$

où  $\mathbf{y} \in \mathbb{R}^{N_y}$  est le signal observé,  $\mathbf{x} \in \mathbb{R}^{N_x}$  le signal source,  $\mathbf{H} \in \mathbb{R}^{N_y \times N_x}$  un opérateur linéaire, et  $\mathbf{e} \in \mathbb{R}^{N_y}$  est un bruit supposé blanc et gaussien. La reconstruction parcimonieuse est souvent abordée par la minimisation d'un critère du type :

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \lambda \mathcal{R}(\mathbf{x}) \quad (2)$$

où  $\|\cdot\|$  est la norme euclidienne, le terme de régularisation  $\mathcal{R}$  est une mesure de parcimonie (par exemple, les "normes"  $\ell_1$  ou  $\ell_0$ ), et  $\lambda > 0$  un hyperparamètre fixé. La minimisation de (2) peut se faire grâce aux méthodes de descente proximale [6].

Les approches par régularisation nécessitent de choisir  $\lambda$ . Pour la minimisation  $\ell_1$ , l'approche SURE (Stein Unbiased Risk Estimator) [3] proposée pour des problèmes de débruitage, a ensuite été généralisée à des problèmes inverses plus généraux [11]. SURE ne s'appliquant pas au cadre  $\ell_0$ , SCORE a été proposée [7] dans le cadre du débruitage.

L'approche bayésienne offre une alternative aux méthodes de type optimisation, la parcimonie étant introduite à l'aide d'un modèle *a priori* spécifique. Le modèle Bernoulli-Gaussien (BG) est un choix naturel pour décrire des signaux parcimonieux. De plus, la formulation (2) avec une pénalisation  $\ell_0$  s'interprète comme un cas limite (quand la variance du processus

gaussien tend vers l'infini) d'estimation au sens du *maximum a posteriori* (MAP) [17]. Dans [5], les auteurs proposent un algorithme EM stochastique, où le support de  $\mathbf{x}$  joue le rôle de variables cachées, pour estimer à la fois les paramètres et  $\mathbf{x}$  dans un problème de déconvolution. [14] propose une méthode itérative qui permet d'estimer les paramètres d'un modèle BG lorsque  $\mathbf{H}$  est l'union de dictionnaires unitaires. Cependant, cette méthode produit un estimateur non-convergent même dans le cas du débruitage comme montré dans [9]. Enfin, malgré des avancées pour les rendre plus efficaces [1], les approches par échantillonnage du type MCMC dans un cadre BG passent difficilement à l'échelle [10].

Cet article traite de l'estimation non supervisée des paramètres d'un modèle BG. La section 2 décrit un algorithme EM dans le cas du débruitage. La section 3 étend cette approche aux problèmes inverses linéaires grâce à un découplage des étapes d'inversion de données et de débruitage, inspiré de [8]. Ces méthodes sont évaluées sur des problèmes de débruitage et de déconvolution dans la section 4.

## 2 Débruitage de signaux BG

On considère d'abord le problème de débruitage  $\mathbf{y} = \mathbf{x} + \mathbf{e}$  ( $\mathbf{H} = \mathbf{I}_N$ , avec  $N = N_y = N_x$ ). On suppose que  $\mathbf{x}$  suit une loi BG,  $\mathbf{x} \sim \mathcal{BN}(p, 0, \sigma_x^2 \mathbf{I}_N)$ , i.e. pour tout  $n \in \{1, \dots, N\}$ ,

$$p(x_n | p, \sigma_x^2) = (1 - p)\delta(x_n) + p \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left(-\frac{x_n^2}{2\sigma_x^2}\right) \quad (3)$$

où  $\delta$  est la distribution de Dirac. Les paramètres  $p \in (0, 1)$  et  $\sigma_x^2$  contrôlent respectivement le degré de parcimonie du signal et la dispersion des valeurs non nulles. Le bruit est supposé blanc, gaussien et indépendant de  $\mathbf{x}$  :  $\mathbf{e} \sim \mathcal{N}(0, \sigma_e^2 \mathbf{I}_N)$ .

## 2.1 Algorithme EM

Le but est d'estimer  $\theta = (\sigma_e^2, p, \sigma_x^2)$  en maximisant la vraisemblance des données  $p(\mathbf{y}|\theta)$ . En considérant  $\mathbf{x}$  comme une variable cachée, on obtient l'algorithme EM :

$$\theta^{(t+1)} = \underset{\theta}{\operatorname{argmin}} \mathbb{E}_{\mathbf{x}|\mathbf{y}, \theta^{(t)}} [\mathcal{L}(\mathbf{y}, \mathbf{x}|\theta)]. \quad (4)$$

où  $\mathcal{L}(\mathbf{y}, \mathbf{x}|\theta) = -\log p(\mathbf{y}, \mathbf{x}|\theta)$ . Les propositions 2.1 et 2.2, démontrées dans notre article [2], résument les étapes E et M, avec la notation  $\mathcal{N}_x(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ .

**Proposition 2.1** (Etape E pour le débruitage de signaux BG). Soient  $\theta^{(t)} = \{\sigma_e^{2(t)}, p^{(t)}, \sigma_x^{2(t)}\}$  les paramètres du modèle à l'itération  $t$ . On pose

$$\phi_n^{(t)} = \frac{p^{(t)} \mathcal{N}_{y_n}(0, \sigma_x^{2(t)} + \sigma_e^{2(t)})}{p^{(t)} \mathcal{N}_{y_n}(0, \sigma_x^{2(t)} + \sigma_e^{2(t)}) + (1-p^{(t)}) \mathcal{N}_{y_n}(0, \sigma_e^{2(t)})}$$

et

$$\mu^{(t)} = \frac{\sigma_x^{2(t)}}{\sigma_x^{2(t)} + \sigma_e^{2(t)}}, \quad \nu^{(t)} = \frac{\sigma_x^{2(t)} \sigma_e^{2(t)}}{\sigma_x^{2(t)} + \sigma_e^{2(t)}}.$$

Alors,

$$\begin{aligned} \mathbb{E}_{\mathbf{x}|\mathbf{y}, \theta^{(t)}} [\mathcal{L}(\mathbf{y}, \mathbf{x}|\theta)] &= C + \frac{1}{2\sigma_e^2} \sum_{n=1}^N y_n^2 (1 - \phi_n^{(t)}) \\ &+ N \log \left( \frac{\sigma_e}{1-p} \right) + \frac{1}{2\sigma_e^2} \sum_{n=1}^N \left( (\mu^{(t)} - 1)^2 y_n^2 + \nu^{(t)} \right) \phi_n^{(t)} \\ &+ \frac{1}{2\sigma_x^2} \sum_{n=1}^N \left( (\mu^{(t)} y_n)^2 + \nu^{(t)} \right) \phi_n^{(t)} + \log \left( \frac{1-p}{p} \sigma_x \right) \sum_{n=1}^N \phi_n^{(t)} \end{aligned}$$

avec  $C$  une constante qui ne dépend pas de  $\theta$ .

Notons que d'après la loi de Bayes,  $\phi_n^{(t)}$  est égal à la probabilité *a posteriori*  $p(x_n \neq 0 | y_n, \theta^{(t)})$  du support de  $\mathbf{x}$ .

**Proposition 2.2** (Etape M pour le débruitage de signaux BG). Les paramètres estimés à l'itération  $t+1$  s'écrivent :

$$p^{(t+1)} = \frac{1}{N} \sum_{n=1}^N \phi_n^{(t)}, \quad \sigma_x^{2(t+1)} = \nu^{(t)} + \frac{\mu^{(t)2}}{N p^{(t+1)}} \sum_{n=1}^N y_n^2 \phi_n^{(t)}$$

$$\sigma_e^{2(t+1)} = \frac{1}{N} \sum_{n=1}^N y_n^2 - \frac{2\mu^{(t)}}{N} \sum_{n=1}^N y_n^2 \phi_n^{(t)} + p^{(t+1)} \sigma_x^{2(t+1)}$$

## 2.2 Méthode des moments

L'algorithme EM présente des garanties de convergence vers un maximum local de  $p(\mathbf{y}|\theta)$ . Pour l'initialiser, on exploite le fait que  $\mathbf{y}$  est un mélange de gaussiennes centrées de covariance  $\sigma_e^2 \mathbf{I}_N$  et  $(\sigma_x^2 + \sigma_e^2) \mathbf{I}_N$ , avec les poids  $1-p$  et  $p$ . Les paramètres  $\theta$  sont estimés à partir des moments d'ordre pair de  $\mathbf{y}$ .

**Proposition 2.3.** L'estimateur des moments pour le débruitage de signaux BG est donné par

$$p = \frac{M_2 - \sigma_e^2}{\sigma_x^2}, \quad \sigma_x^2 = \frac{M_4/3 - M_2\sigma_e^2 - \sigma_e^2}{M_2 - \sigma_e^2}$$

$$(\sigma_e^2)^2 + \frac{M_6/5 - M_4M_2}{3M_2^2 - M_4} \sigma_e^2 = \frac{M_6/5M_2 - M_4/3}{3M_2^2 - M_4}$$

avec  $M_k = \frac{1}{N} \sum_{n=1}^N y_n^k$

*Preuve.* On exprime les moments  $\mathbb{E}_y [y^k]$  d'ordre  $k = 2, 4, 6$  en fonction de  $\theta$  et résout le système. Les moments  $\mathbb{E}_y [y^k]$  sont remplacés par les moments empiriques  $M_k$ .  $\square$

Une fois les paramètres  $\theta$  estimés par l'algorithme EM, on s'intéresse aux estimateurs (supervisés) de  $\mathbf{x}$ , i.e., le *MAP marginal* (MMAP) et la moyenne *a posteriori* (MMSE) [15]. Le MMAP estime en premier lieu le support de  $\mathbf{x}$  par marginalisation des amplitudes, et le MMSE effectue une moyenne des estimés de  $\mathbf{x}$  sur l'ensemble des supports possibles. L'estimateur du MMAP s'écrit [15] :

$$x_n^{MMAP} = \begin{cases} \frac{\sigma_x^2}{\sigma_x^2 + \sigma_e^2} y_n & \text{si } |y_n| > \lambda^{MMAP} \\ 0 & \text{sinon} \end{cases} \quad (5)$$

avec

$$\lambda^{MMAP} = \sqrt{2\sigma_e^2 \frac{\sigma_x^2 + \sigma_e^2}{\sigma_x^2} \log \left( \frac{1-p}{p} \sqrt{1 + \frac{\sigma_x^2}{\sigma_e^2}} \right)}. \quad (6)$$

L'estimateur de la moyenne *a posteriori* est donné par

$$x_n^{MMSE} = \phi_n^{(t)} y_n \quad (7)$$

## 3 Extension aux problèmes inverses

L'algorithme EM de la section 2 conduit à des calculs insolubles pour des problèmes inverses plus généraux que le débruitage, car l'étape E nécessite de calculer une espérance par rapport à tous les supports possibles pour  $\mathbf{x}$  [5]. Nous proposons de reformuler le problème (1) pour faire apparaître un sous-problème de débruitage parcimonieux dans l'esprit des travaux [8] dédiés à la restauration d'un signal dans une base d'ondelettes. On introduit le signal latent  $\mathbf{z} \in \mathbb{R}^{N \times}$  tel que :

$$\mathbf{y} = \mathbf{H}\mathbf{z} + \mathbf{n}, \quad \mathbf{z} = \mathbf{x} + \mathbf{b} \quad (8)$$

où  $\mathbf{b} \sim \mathcal{N}(0, \sigma_b^2 \mathbf{I}_{N_x})$  et  $\mathbf{n} \sim \mathcal{N}(0, \sigma_e^2 \mathbf{I}_{N_y} - \sigma_b^2 \mathbf{H}\mathbf{H}^T)$  sont indépendants, avec  $\sigma_b^2 = \frac{\sigma_e^2}{\|\mathbf{H}\mathbf{H}^T\|}$ .

Les hyperparamètres du problème sont  $\theta = (\sigma_b^2, p, \sigma_x^2)$ , où  $\sigma_e$  se déduit de  $\sigma_b$  comme indiqué ci-dessus. En considérant  $\mathbf{x}$  comme variable cachée, l'algorithme EM s'écrit :

$$\min_{\mathbf{z}, \theta} \mathbb{E}_{\mathbf{x}|\mathbf{y}, \mathbf{z}^{(t)}, \theta^{(t)}} [-\log p(\mathbf{y}, \mathbf{x}, \mathbf{z}|\theta)]. \quad (9)$$

Comme  $p(\mathbf{y}, \mathbf{x}, \mathbf{z}|\theta) = p(\mathbf{y}|\mathbf{z}, \theta)p(\mathbf{x}, \mathbf{z}|\theta)$ , (9) devient :

$$\min_{\mathbf{z}} \left\{ -\log p(\mathbf{y}|\mathbf{z}, (\sigma_e^2)^{(t)}) + \min_{\theta} (\mathbb{E}_{\mathbf{x}|\mathbf{z}^{(t)}, \theta^{(t)}} [\mathcal{L}(\mathbf{x}, \mathbf{z}|\theta)]) \right\}$$

où  $(\sigma_e^2)^{(t)} = \|\mathbf{H}\mathbf{H}^T\|(\sigma_b^2)^{(t)}$  et  $\min_{\theta}(\mathbb{E}_{\mathbf{x}|\mathbf{z}^{(t)},\theta^{(t)}}[\mathcal{L}(\mathbf{z},\mathbf{x}|\theta)])$  est le problème de débruitage en  $\mathbf{z}^{(t)}$ , voir section 2. Soit  $\theta^{(t+1)}$  l'estimée courante de  $\theta$  et  $\hat{\mathbf{x}}^{(t)} = \mathbb{E}_{\mathbf{x}|\mathbf{z}^{(t)},\theta^{(t)}}[\mathbf{x}]$  l'estimateur de la moyenne *a posteriori* du problème de débruitage d'un signal BG à partir de  $\mathbf{z}^{(t)}$ , donné par (7). En exploitant la linéarité de l'espérance, la minimisation en  $\mathbf{z}$  devient simplement

$$\min_{\mathbf{z}} -\log p(\mathbf{z}|\mathbf{y}, \hat{\mathbf{x}}^{(t)}, \theta^{(t+1)}). \quad (10)$$

$p(\mathbf{z}|\mathbf{y}, \hat{\mathbf{x}}^{(t)}, \theta^{(t+1)})$  étant gaussien, le minimiseur, noté  $\mathbf{z}^{(t+1)}$ , peut s'écrire [13]

$$\mathbf{z}^{(t+1)} = \hat{\mathbf{x}}^{(t)} + \frac{1}{\|\mathbf{H}\mathbf{H}^T\|} \mathbf{H}^T (\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}^{(t)}) \quad (11)$$

L'Algorithme 1 résume l'algorithme EM obtenu.

---

**Algorithme 1:** Algorithme EM pour les problèmes inverses avec un *a priori* BG.

---

**Résultat:** Estimation des paramètres du modèle.

$t = 0, \mathbf{z}^{(t)} = \mathbf{0}, \mathbf{x}^{(t)} = \mathbf{0};$

**Tant que**  $t < N_t$  **faire**

$\mathbf{z}^{(t+1)} = \mathbf{x}^{(t)} + \frac{1}{\|\mathbf{H}\mathbf{H}^T\|} \mathbf{H}^T (\mathbf{y} - \mathbf{H}\mathbf{x}^{(t)});$   
 Estimer  $\theta^{(t+1)}$  en utilisant prop.2.2 avec  $\mathbf{y} = \mathbf{z}^{(t)};$   
 $\mathbf{x}^{(t+1)} = \mathbb{E}_{\mathbf{x}|\mathbf{z}^{(t+1)},\theta^{(t+1)}}[\mathbf{x}];$   
 $t = t + 1;$

**fin**

---

Cet algorithme a pour but de fournir une estimation des paramètres  $p$  et  $\sigma_x^2$  du prior BG ainsi que de la variance  $\sigma_e^2 = \sigma_b^2 \|\mathbf{H}\mathbf{H}^T\|$  du bruit. De plus, par construction, l'algorithme fournit une estimation de  $\mathbf{x}$  au sens de la moyenne *a posteriori*,  $\mathbb{E}[\mathbf{x}|\mathbf{z}^{(t)}]$ . On peut aisément calculer le seuil du MAP marginal (6) obtenu pour le débruitage de  $\mathbf{z}^{(t)}$ , noté ensuite  $\lambda^{MAP}$ .

## 4 Résultats expérimentaux

On évalue la méthode proposée sur des problèmes de débruitage de signaux et de déconvolution avec des images BG simulées et des images naturelles. On simule à chaque expérience 100 réalisations d'images de taille  $256^2$  et/ou de bruit avec  $p \in (0, 1)$ ,  $\sigma_x^2 = 1$  et  $\sigma_e^2$  choisi afin de fixer un SNR cible (*input SNR*)  $i\text{SNR} = 10 \log_{10}(\|\mathbf{H}\mathbf{x}\|^2/\|\mathbf{e}\|^2)$ . Pour les images naturelles,  $\mathbf{x}$  rassemble les coefficients en ondelettes, supposés parcimonieux. La transformée est de Daubechies-4 avec redondance par *cycle spinning*. L'algorithme 1 est initialisé par l'image nulle. Les résultats sont évalués en termes de SNR en sortie,  $o\text{SNR} = 10 \log_{10}(\|\mathbf{x}\|^2/\|\mathbf{x} - \hat{\mathbf{x}}\|^2)$ , où  $\hat{\mathbf{x}}$  est l'estimée de  $\mathbf{x}$ . Les expériences ont été faites en Python sur un Intel Core i7 CPU à 1.8 GHz avec 16 GB de RAM.

### 4.1 Débruitage

La Fig. 1 illustre la capacité d'estimation des paramètres de signaux simulés de la méthode donnée à la Prop. 2.2 dans le

cas du débruitage. La mesure affichée est l'erreur relative, *i.e.*  $|\hat{\theta}_i - \theta_i|/\theta_i$ , avec  $\hat{\theta}_i$  le paramètre estimé.

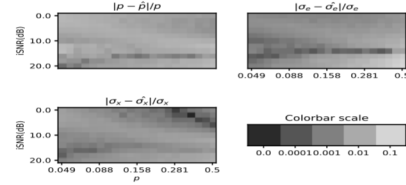


FIGURE 1 – Estimation des paramètres en débruitage.

La Fig. 2 représente les performances en débruitage de signaux simulés pour  $p = 0.01$  avec trois *iSNR* différents. A *iSNR* fixé, l'approche proposée est comparée à son pendant supervisé (courbes en trait plein), c'est à dire aux résultats de seuillage dur obtenus après convergence pour différentes valeurs de l'hyperparamètre  $\lambda$ , cf (2). Ces résultats sont comparés à l'estimateur de la moyenne *a posteriori* en sortie d'EM (traits horizontaux pointillés). Enfin la valeur du seuil MAP estimé en sortie d'EM est matérialisé par une étoile. Remarquablement, les performances obtenues par débruitage non supervisé (en terme de *oSNR*) sont égales ou proches des meilleurs résultats possibles par débruitage supervisé.

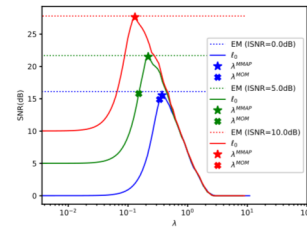


FIGURE 2 – Comparaison en débruitage de signaux simulés.

En débruitage d'images naturelles, la méthode proposée permet d'atteindre automatiquement un *oSNR* proche du résultat optimal obtenu par seuillage dur, voire supérieur, soit par l'estimateur obtenu en sortie d'algorithme, soit par sélection automatique du seuil.

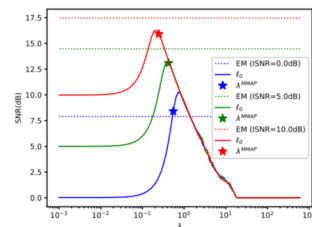


FIGURE 3 – Comparaison en débruitage d'images.



## 4.2 Déconvolution

Dans cette partie, l'opérateur  $\mathbf{H}$  résulte de la combinaison d'un filtre passe-bas de la forme  $K(px, py) = \int_{(px,py)-0.5}^{(px,py)+0.5} \mathcal{N}_{x,y}(0, \eta) dx dy$ , où  $\eta$  est un paramètre d'étalement, et de l'opérateur de synthèse de la transformée en ondelettes redondante pour les images. Les images sont simulées comme précédemment. La Fig. 4 montre l'erreur relative vis-à-vis de chaque paramètre. On peut noter que l'estimation se dégrade pour  $\eta > 0.3$ .

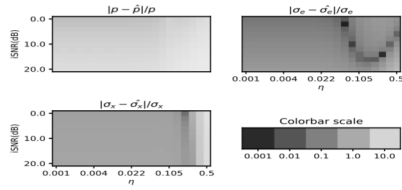


FIGURE 4 – Estimation des paramètres en déconvolution.

La Fig. 5 compare l'approche proposée avec la minimisation  $\ell_0$  faite par algorithme IHT [12]. Pour des noyaux larges, le oSNR donné par l'estimateur MMAP en minimisation  $\ell_0$  avec  $\lambda$  donne de meilleurs résultats que l'estimateur de la moyenne *a posteriori*. L'EM non supervisé fournit une valeur de  $\lambda$  pertinente pour égaliser le meilleur oSNR de la minimisation  $\ell_0$ .

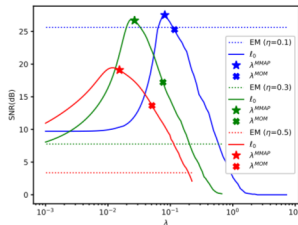


FIGURE 5 – Déconvolution de signaux simulés.

Pour les images, on peut voir sur la Fig. 6 que les estimations en sortie d'algorithme ou par sélection automatique de  $\lambda$  sont comparables. Contrairement aux signaux simulés, les résultats sont moins sensibles aux erreurs d'estimations des cartes d'activations des coefficients, le SNR étant calculé dans le domaine image et non dans le domaine ondelette.

## 5 Conclusion

Cet article présente un algorithme entièrement non supervisé pour estimer les paramètres d'un modèle BG en problème inverse. En débruitage, l'estimateur de la moyenne *a posteriori* atteint le meilleur oSNR comme attendu. En déconvolution, un algorithme itératif de type "ISTA" est obtenu avec une évaluation automatique du seuil  $\lambda$  utilisée dans le seuillage dur ité-

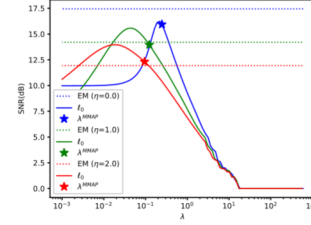


FIGURE 6 – Déconvolution d'images.

ratif. Par la suite, une comparaison avec des méthodes bayésiennes [4, 16] sera menée.

## Références

- [1] M. C. Amrouche, H. Carfantan, and J. Idier. A partially collapsed Gibbs sampler for unsupervised nonnegative sparse signal restoration. In *Proc. IEEE ICASSP*, Toronto, Canada, Jun. 2021.
- [2] P. Barbault, M. Kowalski, and C. Soussen. Parameter estimation in sparse inverse problems using Bernoulli-Gaussian prior. In *ICASSP*, 2022.
- [3] T. Blu and F. Luisier. The SURE-LET approach to image denoising. *IEEE Trans. Image Process.*, 16(11), 2007.
- [4] L. Chaari, J.-Y. Tourneret, and C. Chau. Sparse signal recovery using a Bernoulli generalized Gaussian prior. In *EUSIPCO*, 2015.
- [5] F. Champagnat, Y. Goussard, and J. Idier. Unsupervised deconvolution of sparse spike trains using stochastic approximation. *IEEE Trans. Signal Process.*, 44(12), Dec. 1996.
- [6] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *SIAM J. Multiscale Model. Simul.*, 4(4), 2005.
- [7] C.-A. Deledalle, G. Peyré, and J. Fadili. Stein consistent risk estimator (score) for hard thresholding. *arXiv preprint arXiv:1301.5874*, 2013.
- [8] M. Figueiredo and R. Nowak. An EM algorithm for wavelet-based image restoration. *IEEE Trans. Image Process.*, 12(8), 2003.
- [9] E. Gassiat, F. Monfront, and Y. Goussard. On simultaneous signal estimation and parameter identification using a generalized likelihood approach. *IEEE Trans. Inf. Theory*, 38(1), 1992.
- [10] D. Ge, J. Idier, and E. Le Carpentier. A new MCMC algorithm for blind Bernoulli-Gaussian deconvolution. In *Proc. 16th Eur. Sig. Proc. Conf.*, 2008.
- [11] R. Giryes, M. Elad, and Y. C. Eldar. The projected GSURE for automatic parameter tuning in iterative shrinkage methods. *Applied and Computational Harmonic Analysis*, 30(3), 2011.
- [12] E. T. Hale, W. Yin, and Y. Zhang. Fixed-point continuation for  $\ell_1$ -minimization: Methodology and convergence. *SIAM J. Optimization*, 19(3), 2008.
- [13] S. M. Kay. *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall, Inc., 1993.
- [14] M. Kowalski and T. Rodet. An unsupervised algorithm for hybrid/morphological signal decomposition. In *ICASSP*, 2011.
- [15] M. Protter, I. Yavneh, and M. Elad. Closed-form MMSE estimation for signal denoising under sparse representation modeling over a unitary dictionary. *IEEE Trans. Signal Process.*, 58(7), 2010.
- [16] J. G Serra, J. Mateos, R. Molina, and A. K Katsaggelos. Variational EM method for blur estimation using the spike-and-slab image prior. *Digital Signal Processing*, 88, 2019.
- [17] C. Soussen, J. Idier, D. Brie, and J. Duan. From Bernoulli-Gaussian deconvolution to sparse signal restoration. *IEEE Trans. Signal Process.*, 59(10), 2011.