

Un modèle bayésien exact pour les problèmes inverses non-linéaires en présence de bruits divers

Nicolas GOEMAN Pierre-Antoine THOUVENIN Pierre CHAINAIS

Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRIStAL, F-59000 Lille, France

Résumé – Les problèmes inverses en astronomie sont généralement très mal posés, en raison de la non-linéarité du modèle direct reliant les paramètres physiques d’intérêt aux observations, la présence de multiples sources de bruit et la censure éventuelle des données en raison de la sensibilité limitée des capteurs. En l’absence de vérité terrain, les algorithmes de Monte-Carlo par chaînes de Markov permettent de résoudre ces problèmes en assortissant les estimateurs d’incertitudes quantifiées. Afin de simplifier l’inférence de paramètres dans des modèles aussi complexes, des méthodes approchées sont souvent proposées, ignorant ou simplifiant certaines des difficultés mentionnées. Ce travail propose une approche exacte basée sur un modèle bayésien hiérarchique prenant en compte tous les aspects évoqués. Cette approche s’avère plus simple à utiliser qu’un modèle approché proposé dans la littérature, nécessitant de calibrer plusieurs hyperparamètres en amont. Les résultats obtenus sur des observations astronomiques synthétiques du médium interstellaire sont également meilleurs.

Abstract – Inverse problems in astronomy are generally very ill-posed. This is due to the non-linearity of the forward model linking the physical parameters of interest to the observations, the presence of multiple sources of noise and data censorship possibly resulting from a limited sensitivity of the sensors used to acquire the observations. In absence of ground truth data, Markov chain Monte Carlo algorithms allow these problems to be solved, forming estimates with quantified uncertainties. To facilitate the inference process with such intractable models, approximate methods are generally proposed by ignoring or simplifying some of the constraints mentioned. This work proposes an exact approach based on a hierarchical Bayesian model accounting for all the difficulties of the problem. The results obtained on synthetic astronomical data show better quality estimates, using a model which does not require any prior calibration phase compared to an approximate model previously proposed in the literature.

1 Introduction

La résolution de problèmes inverses mal posés en astronomie présente de nombreux défis liés à la complexité des modèles physiques considérés et des imperfections des instruments de mesure. En raison d’erreurs de calibration et de la sensibilité du capteur, il est naturel de considérer conjointement des sources de bruit additif et multiplicatif ainsi qu’une censure à bas régime signal sur bruit. Il est en particulier primordial de tenir compte de ces multiples sources de bruit lorsque l’amplitude des observations couvre plusieurs décades, avec des dynamiques très différentes parmi les observations disponibles. De plus, le modèle direct implique très souvent un modèle physique coûteux à évaluer et complexe. Celui-ci peut être remplacé par un modèle réduit de type boîte noire plus rapide à évaluer, mais qui reste fortement non-linéaire [6]. Enfin, la quantification des incertitudes d’estimation est cruciale pour les applications scientifiques où il n’est pas possible de se confronter à une vérité terrain, comme en astronomie. Les algorithmes de Monte-Carlo par chaînes de Markov (MCMC) [9] permettent d’obtenir des estimations et d’en quantifier les incertitudes. L’estimateur d’erreur quadratique moyenne minimale (MMSE) est un choix courant.

La présence de deux sources de bruit conduit à une loi a posteriori non standard, difficile à échantillonner efficacement

Ce travail est soutenu par la Chaire IA Sherlock ANR-20-CHIA-0031-01 portée par P. Chainais, ainsi que par le programme national d’investissement d’avenir ANR-16-IDEX-0004 ULNE et la Région Hauts-de-France. Ce travail a bénéficié de moyens de calculs du Mésocentre de l’Université de Lille. Nous remercions Pierre Palud pour les riches discussions ayant facilité le développement de cette méthode hiérarchique.

à l’aide d’algorithmes classiques. En pratique, plusieurs travaux proposent de négliger l’une des sources de bruit [1, 2]. L’approche proposée dans [7] permet de prendre en compte les difficultés ci-dessus en considérant les deux sources de bruit à l’aide d’une approximation de la loi a posteriori. Dans cette approche, une combinaison d’approximations est proposée allant d’un modèle de bruit additif à bas rapport signal sur bruit vers un modèle de bruit multiplicatif à haut rapport signal sur bruit. L’interpolation progressive entre ces deux régimes de bruit est gérée par une fonction de poids dans une loi de vraisemblance approchée. Ce modèle approché nécessite de calibrer en amont plusieurs hyperparamètres, parfois difficiles à optimiser.

Cet article propose une alternative à l’approche [7] basée sur un modèle bayésien hiérarchique exact. Cette approche décrite en partie 2 ne requiert pas de calibration d’hyperparamètres, ce qui facilite son adoption par des astronomes. D’autre part, la loi de vraisemblance hiérarchique permet d’utiliser un échantillonneur de Gibbs avec de bonnes propriétés détaillé en partie 3. Les lois conditionnelles ont chacune des contraintes qui leur sont propres et qui nécessitent des noyaux adaptés. En partie 4, l’approche proposée est évaluée sur données synthétiques par rapport à [7] en termes de qualité de reconstruction et d’intervalles de crédibilité.

2 Modèle proposé

2.1 Modèle d’observation

Les observations considérées sont typiquement des cartes de N pixels acquises en L canaux spectraux $\mathbf{Y} = (y_{n,\ell})_{n,\ell=1}^{N \times L} \in$

$(\mathbb{R}_+^*)^{N \times L}$. Chaque observation $y_{n,\ell}$ est supposée reliée aux paramètres d'intérêt $\theta_n \in \mathbb{R}^D$ par un modèle direct coûteux à évaluer, remplacé en général par un modèle réduit $f_\ell : \mathbb{R}^D \rightarrow \mathbb{R}_+^*$ moins coûteux et tel que $\mathbf{f} = (f_1, \dots, f_L)$ (voir par exemple [6]). Les observations sont perturbées à la fois par un bruit additif $\epsilon_{n,\ell} \sim \mathcal{N}(0, \sigma_{a,n,\ell}^2)$ et par un bruit multiplicatif $\xi_{n,\ell} \sim \text{Log-}\mathcal{N}(-\sigma_m^2/2, \sigma_m^2)$ paramétré de telle sorte que $\mathbb{E}[\xi_{n,\ell}] = 1$. Les paramètres $\sigma_a = (\sigma_{a,n,\ell})_{n,\ell}$ et σ_m sont supposés connus et les sources de bruit indépendantes. Enfin, une censure par borne inférieure $\omega_{n,\ell} > 0$ est appliquée sur chaque observation, représentant un seuil de sensibilité du matériel d'acquisition. Le modèle direct d'observation est défini par

$$y_{n,\ell} = \max\left\{\omega_{n,\ell}, \xi_{n,\ell} f_\ell(\theta_n) + \epsilon_{n,\ell}\right\}. \quad (1)$$

La présence de deux sources de bruit, additif et multiplicatif, rend la vraisemblance compliquée à manipuler pour l'inférence. Une possibilité proposée dans [7] consiste à construire une approximation de la vraisemblance via une interpolation entre deux vraisemblances respectivement associées à un bruit additif gaussien et un bruit multiplicatif log-normal. À faible régime de rapport signal sur bruit (SNR), le bruit multiplicatif $\xi_{n,\ell}$ peut raisonnablement être négligé et la vraisemblance est alors approximée par une loi gaussienne [2]. À haut SNR, le bruit additif $\epsilon_{n,\ell}$ peut être négligé et la vraisemblance est approchée par une loi log-normale. La transition entre ces deux régimes est modulée par une fonction de poids $\lambda \in \mathcal{C}^2(\mathbb{R}^D, [0, 1])$ qui dépend de l'amplitude du signal.

Cette approche a permis d'obtenir de très bonnes estimations des paramètres d'intérêt. Néanmoins, elle nécessite un ajustement fin des hyperparamètres qui définissent λ pour assurer la qualité du modèle approché. Ces hyperparamètres sont directement liés aux paramètres σ_a et σ_m , et doivent par conséquent être ajustés au cas par cas pour chaque problème traité. Cela représente une difficulté supplémentaire.

2.2 Modèle hiérarchique de vraisemblance

Nous proposons d'étudier une formulation alternative de (1) sous la forme hiérarchique

$$\begin{cases} u_{n,\ell} = \xi_{n,\ell} f_\ell(\theta_n) \sim \text{Log-}\mathcal{N}(\ln f_\ell(\theta_n) - \sigma_m^2/2, \sigma_m^2), \\ y_{n,\ell} = \max\left\{\omega_{n,\ell}, u_{n,\ell} + \epsilon_{n,\ell}\right\}. \end{cases} \quad (2)$$

Ce modèle introduit une variable auxiliaire $\mathbf{U} \in (\mathbb{R}_+^*)^{N \times L}$. Celle-ci doit être inférée conjointement avec les paramètres d'intérêt $\Theta = (\theta_1, \dots, \theta_N) \in \mathbb{R}^{N \times D}$. En notant $c_{n,\ell} = \mathbf{1}_{\{y_{n,\ell} = \omega_{n,\ell}\}}$ connus par les observations, la fonction de vraisemblance s'écrit sous la forme

$$\pi(\mathbf{Y}|\mathbf{U}) = \prod_{n=1}^N \prod_{\ell=1}^L \pi(y_{n,\ell}|u_{n,\ell}), \quad (3)$$

avec

$$\begin{aligned} \pi(y_{n,\ell}|u_{n,\ell}) &= \phi(y_{n,\ell}|u_{n,\ell}, \sigma_{a,n,\ell}^2)^{(1-c_{n,\ell})} \\ &\quad \times \Phi(\omega_{n,\ell}|u_{n,\ell}, \sigma_{a,n,\ell}^2)^{c_{n,\ell}}, \end{aligned} \quad (4)$$

où $\phi(\cdot|\mu, \sigma^2)$ est la densité de probabilité d'une loi gaussienne de paramètres μ et σ^2 et $\Phi(\cdot|\mu, \sigma^2)$ sa fonction de répartition.

La censure concentre ici toute la masse de $\phi(y_{n,\ell}|u_{n,\ell}, \sigma_{a,n,\ell}^2)$ sur $[-\infty, \omega_{n,\ell}[$ sous la forme d'un Dirac au point $\omega_{n,\ell}$. Le modèle hiérarchique ne nécessite pas d'ajustement précis d'hyperparamètres et se prête bien à un échantillonnage de Gibbs, alternant le tirage d'échantillons des paramètres Θ et \mathbf{U} .

2.3 Lois a priori et conditionnelles

Avec l'introduction de la variable auxiliaire \mathbf{U} , la loi a posteriori augmentée $\pi(\Theta, \mathbf{U}|\mathbf{Y})$ s'exprime comme suit :

$$\pi(\Theta, \mathbf{U}|\mathbf{Y}) \propto \pi(\mathbf{Y}|\mathbf{U}) \pi(\mathbf{U}|\Theta) \pi(\Theta). \quad (5)$$

Nous considérons le même a priori que [7], qui combine de deux régularisations distinctes

$$\pi(\Theta) \propto \exp\left[-\sum_{d=1}^D \tau_d h(\theta_{\cdot,d}) - \delta \sum_{n=1}^N \tilde{\tau}_C(\theta_n)\right], \quad (6)$$

où $\theta_{\cdot,d} = (\theta_{n,d})_{1 \leq n \leq N}$, $\theta_n = (\theta_{n,d})_{1 \leq d \leq D}$, et la fonction h est une régularisation spatiale basée sur un opérateur de gradient discret 2D. La fonction $\tilde{\tau}_C$ est une approximation différentiable de la fonction indicatrice sur un compact $C = [l_1, u_1] \times \dots \times [l_D, u_D] \subset \mathbb{R}^D$ [5], permettant de restreindre les paramètres à un domaine de validité et de satisfaire les hypothèses de différentiabilité requises par les noyaux de transition mentionnés en partie 3. Les hyperparamètres $\tau = (\tau_d)_{1 \leq d \leq D} \in]0, +\infty[^D$ et $\delta > 0$ contrôlent le poids de l'a priori par rapport à la vraisemblance. Un algorithme d'échantillonnage de Gibbs peut être appliqué pour générer des échantillons asymptotiquement distribués suivant la loi $\pi(\Theta, \mathbf{U}|\mathbf{Y})$. C'est le choix qui a été fait, comme détaillé en partie 3. Les conditionnelles de Θ et \mathbf{U} sont données par :

$$\begin{cases} \pi(\Theta|\mathbf{Y}, \mathbf{U}) \propto \prod_{n=1}^N \pi(\mathbf{u}_n|\theta_n) \pi(\theta_n), & (7a) \\ \pi(\mathbf{U}|\mathbf{Y}, \Theta) \propto \prod_{n=1}^N \prod_{\ell=1}^L \pi(y_{n,\ell}|u_{n,\ell}) \pi(u_{n,\ell}|\theta_n). & (7b) \end{cases}$$

Cette approche a l'avantage de séparer les difficultés. Le modèle direct \mathbf{f} n'apparaît en effet que dans la conditionnelle sur Θ dans (7a). L'équation (7b) montre que la prise en compte simultanée des deux sources de bruit, additif et multiplicatif, n'intervient que dans la loi conditionnelle de \mathbf{U} . En particulier, la densité de $\pi(u_{n,\ell}|\theta_n)$ est log concave sur l'intervalle $]0, f_\ell(\theta_n) \exp(1 - \frac{3}{2}\sigma_m^2)]$. Pour une valeur $\sigma_m = \ln(1.3)$ considérée comme grande pour l'application traitée en partie 4, ce support contient plus de 99.9% de la masse de $\pi(u_{n,\ell}|\theta_n)$. La conditionnelle sur \mathbf{U} est alors très régulière et son échantillonnage plus facile à gérer.

3 Échantillonneur de Gibbs

La partie 3 présente les stratégies d'échantillonnage utilisées pour tirer des échantillons suivant les lois conditionnelles de Θ et \mathbf{U} dans le cadre d'un échantillonneur de Gibbs.

3.1 Échantillonnage de $\Theta|\mathbf{Y}, \mathbf{U}$

La non-linéarité et la multimodalité potentiellement introduites par le modèle direct \mathbf{f} sont toujours présentes dans

$\pi(\Theta|Y, U)$ et sont traitées en utilisant des noyaux issus de [7]. Un premier noyau *Preconditioned Metropolis-Adjusted Langevin Algorithm* (PMALA) [10] est utilisé pour permettre une exploration locale de la loi conditionnelle à l'aide d'une matrice de préconditionnement $G^{(t)} \in \mathbb{R}^{ND \times ND}$ définie positive. Nous renvoyons à [7] pour plus de détails sur le choix du préconditionneur.

Le second noyau utilisé est un *Independant-Multiple Try Metropolis* (I-MTM) [3, 4] qui permet d'explorer les différents modes de la loi conditionnelle. Le noyau sélectionne un candidat parmi K candidats générés par une loi de proposition. Cette méthode produit des candidats dans des régions et donc des modes divers. Nous utilisons ici la même loi de proposition que [7, équations (30) et (31)].

3.2 Échantillonnage de $U|Y, \Theta$

D'après (7b), les $u_{n,\ell}$ sont tous indépendants les uns des autres conditionnellement à Θ et Y . Ainsi, un échantillonneur de Gibbs chromatique peut être utilisé : celui-ci consiste à proposer et accepter ou rejeter chaque $u_{n,\ell}$ en parallèle et de façon indépendante, accélérant ainsi l'échantillonnage de la variable auxiliaire. Il convient donc finalement de trouver un noyau de transition efficace pour $u_{n,\ell}|y_{n,\ell}, \theta_n$. La conditionnelle $\pi(U|Y, \Theta)$ est soumise à la présence des deux sources de bruit dont les contributions respectives dépendent du régime de bruit. Il est ainsi nécessaire de construire un noyau de transition qui considère ces différentes dynamiques.

Nous optons pour une marche aléatoire gaussienne de variance fixée $\sigma_{a,n,\ell}^2$ pour l'ensemble des paramètres auxiliaires associés aux observations non-censurées, telles que $c_{n,\ell} = 0$. Dans le produit $\pi(y_{n,\ell}|u_{n,\ell}) \times \pi(u_{n,\ell}|\theta_n)$ de (7b), le poids respectif de chacun des termes dépend du SNR. La variance du terme associé au bruit multiplicatif varie avec l'amplitude du modèle direct f_ℓ , tandis que la variance du bruit additif reste constante. À haut SNR, la variance de $y_{n,\ell}|u_{n,\ell}$ est très inférieure à celle de $u_{n,\ell}|\theta_n$. La vraisemblance $\pi(y_{n,\ell}|u_{n,\ell})$ a donc un poids plus important dans la densité de la loi conditionnelle $\pi(u_{n,\ell}|y_{n,\ell}, \theta_n)$ que $\pi(u_{n,\ell}|\theta_n)$. Un candidat proposé dans un tel régime doit par conséquent être raisonnablement proche de $y_{n,\ell}$ et l'initialisation de U aux observations Y est alors très importante pour rapidement explorer des zones de forte probabilité dans l'espace des paramètres.

Les observations censurées, telles que $c_{n,\ell} = 1$, sont très peu informatives pour les paramètres auxiliaires associés. Le modèle direct f à l'état courant est alors un meilleur choix pour guider la proposition. Dans ce régime de faible SNR, le rapport de poids entre $\pi(u_{n,\ell}|f_\ell(\theta_n))$ et $\pi(y_{n,\ell}|u_{n,\ell})$ est cette fois inversé. Il est alors plus adapté de perturber $f_\ell(\theta_n)$ par un bruit multiplicatif. La loi de proposition est alors définie par

$$\begin{cases} u_{n,\ell}^{(t+1)} \sim u_{n,\ell}^{(t)} + \mathcal{N}(0, \sigma_{a,n,\ell}^2) \text{ si } c_{n,\ell} = 0, \\ u_{n,\ell}^{(t+1)} \sim f_\ell(\theta_n^{(t+1)}) \text{Log-}\mathcal{N}(-\sigma_m^2/2, \sigma_m^2) \text{ sinon.} \end{cases} \quad (8)$$

Cette proposition est très générique et ne nécessite pas de paramétrisation précise basée sur les niveaux de bruits $\sigma_{a,n,\ell}$ et σ_m ainsi que sur les seuils de censure $\omega_{n,\ell}$. Elle peut a priori être facilement adaptée à d'autres problèmes inverses dont la structure est similaire.

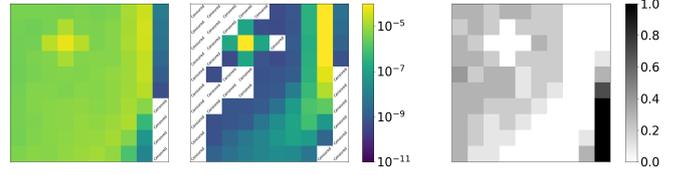


FIGURE 1 : Exemples de cartes d'observations avec les pixels censurés en blanc (à gauche et au centre) et proportion de raies spectrales censurées par pixel (à droite).

4 Résultats

4.1 Conditions expérimentales

Nous évaluons l'approche proposée dans le même cadre expérimental que [7]. Les observations $Y \in (\mathbb{R}_+^*)^{N \times L}$ sont des cartes d'observations hyperspectrales dans le domaine radio. Chaque carte comporte $N = 100$ pixels. Pour chaque pixel n nous cherchons à inférer $D = 4$ paramètres physiques θ_n à partir de $L = 10$ lignes d'émissions. Les paramètres physiques $\theta_n = (\kappa, P_{th}, G_0, A_V)$ représentent respectivement un paramètre de calibration du système d'acquisition, la pression thermique, l'intensité du flux radiatif UV et l'extinction visuelle. Le modèle direct réduit utilisé est un réseau de neurones profond. Les détails concernant l'architecture, l'entraînement et l'évaluation de ce réseau sont donnés dans [6]. Les observations synthétiques Y sont générées à partir de cartes de paramètres Θ^* suivant (2) avec $\sigma_{a,n,\ell} = 1.38715 \times 10^{-10}$ pour tout $(n, \ell) \in \llbracket 1, N \rrbracket \times \llbracket 1, L \rrbracket$, $\sigma_m = \ln(1.1)$ et $\omega_{n,\ell} = 3\sigma_{a,n,\ell}$. Deux cartes d'observations sont représentées en figure 1 avec la proportion de lignes d'observation censurées par pixel.

L'approche hiérarchique proposée a été implémentée dans la librairie Python BEETROOTS [8] avec laquelle les expériences de [7] ont également été réalisées. Des corrections sur la régularisation spatiale ainsi que la proposition I-MTM de [7] ont par ailleurs été apportées à la librairie BEETROOTS dans le cadre de ce travail, et en ont amélioré les résultats. L'approche hiérarchique est ici comparée à l'implémentation révisée de [7] pour $T = 5000$ échantillons et une période de chauffe $T_{BI} = 500$. Les paramètres de régularisation sont fixés à $\tau = 2D$ et $\delta = 1$ pour les deux approches. Les noyaux de transition utilisés pour échantillonner la loi conditionnelle de Θ sont les mêmes que ceux utilisés pour la méthode approchée [7] avec les paramètres $\epsilon = 5 \times 10^{-4}$, $\alpha = 0.99$, $\eta = 1 \times 10^{-5}$ pour PMALA et $K = 100$ pour I-MTM. La qualité d'un estimateur $\hat{\Theta}$ est évaluée à l'aide du SNR de reconstruction $(R\text{-SNR}) 20 \log_{10} \left(\frac{\|\Theta^*\|}{\|\hat{\Theta} - \Theta^*\|} \right)$ et des intervalles de crédibilité à 95%. Nous utilisons dans ces expériences l'estimateur MMSE $\hat{\Theta} = \sum_{t=T_{BI}+1}^T \Theta^{(t)} / (T - T_{BI})$. L'amplitude des paramètres θ_d , pour $d \in \llbracket 2, D \rrbracket$, couvre des décades très différentes d'un paramètre à l'autre. Afin d'obtenir des échelles comparables, chaque carte de paramètres θ_d pour $d \in \llbracket 2, D \rrbracket$ et la variable Θ complète sont évaluées en échelle logarithmique normalisée par dimension. Le paramètre κ est évalué individuellement en échelle linéaire, car il ne couvre pas plusieurs décades. Tous les paramètres Θ sont échantillonnés en échelle logarithmique normalisée comme dans [7]. Les expériences ont été menées sur un ordinateur équipé d'un processeur Intel Xeon Silver 4314 et de 16 Go de mémoire vive.

TABLE 1 : Rapport signal sur bruit de reconstruction (R-SNR) en dB pour les estimations MMSE $\hat{\Theta}$ des deux méthodes.

	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	$\hat{\Theta}$
Approchée [7]	16.47	12.33	16.89	11.70	12.95
Hiéarchique	18.60	13.61	17.10	11.07	13.18

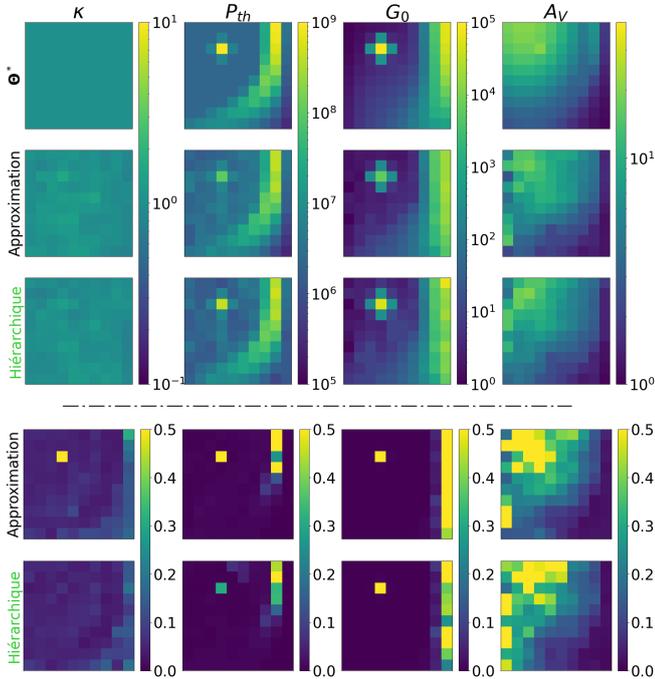


FIGURE 2 : Estimations MMSE pour Θ : vérité terrain Θ^* (1^{re} ligne), méthode approchée [7] (2^e ligne), méthode hiéarchique (3^e ligne). Taille relative de l'intervalle de crédibilité à 95% par rapport à la taille de l'intervalle de validité C_d : méthode approchée (4^e ligne), méthode hiéarchique (5^e ligne).

4.2 Résultats et discussions

La figure 2 présente les estimations MMSE obtenues pour les deux méthodes par rapport à la vérité terrain Θ^* . L'approche hiéarchique fournit de très bons résultats, reproduisant fidèlement les comportements physiques observés sur la vérité terrain Θ^* . Le tableau 1 rassemble les performances des deux approches, et montre que le modèle hiéarchique donne un estimateur de qualité globalement meilleure en termes de R-SNR que le modèle approché [7]. Les méthodes approchée et hiéarchique ont des temps d'exécution similaires, respectivement 319 et 345 millisecondes par itération en moyenne.

Les intervalles de crédibilité à 95% en figure 2 sont semblables pour les deux méthodes, ce qui montre que l'approche hiéarchique explore tout aussi bien l'espace des paramètres Θ . Nous pouvons remarquer une baisse du niveau d'incertitude sur le pixel de forte intensité présent dans le quart supérieur gauche de la carte. Le modèle hiéarchique incite fortement la variable auxiliaire à être proche des observations dans cette zone à haut régime du SNR, où la vraisemblance a un poids plus important. L'intensité de ce pixel est donc moins lissée par la régularisation spatiale que dans la méthode approchée, ce qui peut être observé sur les estimations MMSE en figure 2.

5 Conclusion

Cet article propose une approche alternative exacte aux méthodes habituelles approchées [1, 2, 7] utilisées dans la résolution de problèmes inverses en présence de multiples sources de bruit, par exemple en astronomie. L'approche proposée considère simultanément les difficultés induites par la non-linéarité de modèle direct, la multiplicité des sources de bruit et la censure des observations. L'approche hiéarchique proposée est en ce sens exacte (en dehors de l'émulateur f), plus générique et plus simple à utiliser que l'approche [7] tout en produisant de meilleurs résultats. En particulier, les détails dans les cartes de paramètres physiques sont mieux retranscrits dans les estimations issues de l'approche proposée. Les intervalles de crédibilité à 95% montrent que cette approche explore l'espace des paramètres aussi bien que la méthode approchée [7]. La variable auxiliaire introduite par cette nouvelle approche est relativement simple à échantillonner, et n'entraîne pas de coûts de calcul supplémentaires notables. Enfin, cette famille de modèles étant très générique, la méthode d'inférence proposée présente un intérêt potentiel pour d'autres applications, comportant tout ou partie des difficultés traitées dans ce travail.

Références

- [1] Sylvain DURAND, Jalal FADILI et Mila NIKOLOVA : Multiplicative noise removal using L1 fidelity on frame coefficients, 2010.
- [2] Yumei HUANG, Michael NG et Tiejong ZENG : The convex relaxation method on deconvolution model with multiplicative noise. *Commun. Comput. Phys.*, page 27, 2013.
- [3] Jun S LIU, Faming LIANG et Wing Hung WONG : The multiple-try method and local optimization in metropolis sampling. *J. Am. Stat. Assoc.*, 95(449):121–134, 2000.
- [4] Luca MARTINO : A review of multiple try MCMC algorithms for signal processing. *Digital Signal Processing*, 75:134–152, avril 2018.
- [5] Jorge NOCEDAL et Stephen J. WRIGHT : *Numerical optimization*. Springer, 2006.
- [6] Pierre PALUD, Lucas EINIG *et al.* : Neural network-based emulation of interstellar medium models. *Astronomy & Astrophysics*, 678:A198, octobre 2023.
- [7] Pierre PALUD *et al.* : Efficient sampling of non log-concave posterior distributions with mixture of noises. *IEEE Trans. Signal Process.*, 71:2491–2501, 2023.
- [8] Pierre PALUD *et al.* : Beetroots (Bayesian inference with spaTial Regularization of nOisy multi-line ObservaTion mapS), janvier 2024.
- [9] Christian P. ROBERT et George CASELLA : *Monte Carlo Statistical Methods*. Springer Texts in Statistics. Springer, New York, NY, 2004.
- [10] T. XIFARA *et al.* : Langevin diffusions and the Metropolis-adjusted Langevin algorithm. *Statistics & Probability Letters*, 91:14–19, août 2014.