



Algorithme EM stochastique pour modèles d'espace d'états linéaires en bruits impulsifs : application à la radio interférométrie

Nawel ARAB[†] Isabelle VIN[†] Mohammed Nabil EL KORSO[‡] Pascal LARZABAL[†]

[†]SATIE, ENS Paris-Saclay, Université Paris-Saclay, 91190 Gif-sur-Yvette, France

[‡]L2S, CentraleSupélec, Université Paris-Saclay, 91190 Gif-sur-Yvette, France

Résumé – Les modèles d'espace d'état constituent un cadre rigoureux pour analyser l'évolution temporelle de processus latents à partir d'observations bruitées. Toutefois, l'hypothèse classique d'un bruit de mesure gaussien se révèle inadéquate lorsque les observations présentent des valeurs aberrantes, comme celles générées par les interférences radiofréquences en radio-interférométrie. Dans cet article, nous modélisons le bruit de mesure à l'aide d'une famille de distributions à queues lourdes, structurée par une hiérarchie à texture latente, conçue pour prendre en compte la nature impulsionnelle des interférences. Nous proposons ensuite une méthode d'estimation fondée sur un algorithme d'espérance-maximisation stochastique, dans lequel l'étape d'espérance est approximée par échantillonnage Monte Carlo de la texture latente, rendant l'inférence tractable. Par ailleurs, une mise à jour régularisée par une pénalisation ℓ_1 est introduite à l'étape M afin de promouvoir la parcimonie des états latents. L'efficacité de la méthode est démontrée sur des données synthétiques d'imagerie radio-interférométrique contaminées par de fortes interférences, montrant des gains significatifs en termes de qualité de reconstruction et de robustesse par rapport aux méthodes de référence basées sur l'EM gaussien ou le lisseur de Kalman.

Abstract – State-space models offer a principled framework for analyzing the temporal evolution of latent processes from noisy observations. However, the standard assumption of Gaussian measurement noise can lead to poor performance in the presence of outliers, such as those induced by radio-frequency interferences (RFI) in radio interferometry. In this work, we propose a robust estimation framework for linear dynamical systems with non-Gaussian, heavy-tailed noise modeled via a compound-Gaussian distribution with latent texture. We develop a Stochastic Expectation–Maximization (SEM) algorithm where the E-step is approximated through Monte Carlo sampling of the latent texture, allowing inference in otherwise intractable posterior distributions. Additionally, we incorporate an ℓ_1 -regularized update in the M-step to enforce sparsity in the latent states. We demonstrate the effectiveness of our method on synthetic radio-interferometric imaging data contaminated with strong outliers, showing substantial improvements in reconstruction quality and robustness compared to Gaussian EM and Kalman-based baselines. This highlights the flexibility and applicability of the SEM framework for robust imaging under impulsive noise.

1 Introduction

Les modèles d'espace d'état offrent un cadre analytique pour décrire l'évolution de variables latentes dans les systèmes dynamiques [11, 16]. Ces modèles permettent de représenter des processus latents évoluant dans le temps, lesquels sont observés de manière indirecte par des mesures bruitées. Traditionnellement, l'hypothèse d'un bruit de mesure gaussien s'est imposée en raison de sa tractabilité et de ses propriétés statistiques bien établies.

Cependant, de nombreux phénomènes réels subissent des perturbations qui s'écartent de cette hypothèse gaussienne. Par exemple, en radio-interférométrie, les observations sont souvent affectées par des sources de bruit non gaussien, telles que les interférences radiofréquences (RFI) [13], issues de signaux d'origine anthropique et introduisant des distorsions significatives dans les mesures astronomiques [12, 2]. Les approches classiques en radio-interférométrie, qui reposent généralement sur l'hypothèse d'un bruit additif gaussien [19], peuvent ainsi conduire à des reconstructions inexactes en omettant de prendre en compte la nature à queues lourdes du bruit réel. La prise en compte du bruit non gaussien a conduit au développement de méthodes avancées telles que le filtrage particulaire [7], l'inférence variationnelle et des variantes non conventionnelles du filtre de Kalman [15, 9, 20]. Les filtres

particulaires, ou méthodes de Monte Carlo séquentielles, sont conçus pour traiter des modèles non linéaires et non gaussiens en représentant la distribution a posteriori des états par un ensemble d'échantillons pondérés (particules). Toutefois, ces méthodes présentent plusieurs inconvénients, notamment la dégénérescence des particules (bien que partiellement atténuable par un rééchantillonnage) et un coût de calcul qui s'accroît de manière exponentielle en grande dimension. L'inférence variationnelle offre une alternative en reformulant le problème d'inférence en une tâche d'optimisation [1]. Elle consiste à approcher des distributions a posteriori complexes par des distributions plus simples et tractables, en minimisant leur divergence, ce qui améliore l'efficacité computationnelle par rapport au filtrage particulaire. Néanmoins, cette approche peut souffrir de minima locaux et d'une sensibilité à la famille de distributions choisie pour l'approximation. Par ailleurs, bien que le filtre de Kalman conventionnel soit généralement utilisé dans le cas gaussien, plusieurs travaux récents ont proposé des variantes adaptées au cas non gaussien [15, 9, 20]. Il est toutefois important de noter que ces approches supposent que les paramètres du modèle sont a priori connus.

Dans ce travail, nous proposons d'estimer conjointement les paramètres du modèle et les états latents dans un cadre d'espace d'état présentant un bruit de mesure composé gaussien

[10, 17, 8]. Plus précisément, nous recourons à un algorithme d'Espérance-Maximisation stochastique (EMS) [14, 5] pour gérer l'intractabilité de la distribution a posteriori. Contrairement à l'EM classique [6], qui requiert le calcul d'espérances complètes sur les variables latentes, la méthode EMS remplace cette étape par un échantillonnage Monte Carlo. Par ailleurs, nous intégrons dans l'étape M une mise à jour par gradient proximal régularisée afin d'imposer des contraintes structurées sur les états latents estimés. Cette régularisation s'avère particulièrement avantageuse dans des applications telles que la radio-interférométrie [3], où des représentations d'images parcimonieuses sont plus réalistes et conduisent à une meilleure précision de reconstruction.

2 Modèle statistique

Nous considérons un modèle linéaire à temps discret d'espace d'état avec un bruit de mesure additif composé gaussien :

$$\mathbf{x}_t = \mathbf{F}_t \mathbf{x}_{t-1} + \mathbf{w}_t, \quad (1)$$

$$\mathbf{y}_t = \mathbf{H}_t \mathbf{x}_t + \tau_t^{-\frac{1}{2}} \mathbf{n}_t, \quad (2)$$

où $\mathbf{y}_t \in \mathbb{C}^m$ représente le vecteur d'observation et $\mathbf{x}_t \in \mathbb{R}^n$ l'état latent, $\mathbf{F}_t \in \mathbb{R}^{n \times n}$ la matrice de transition, et $\mathbf{H}_t \in \mathbb{C}^{m \times n}$ la matrice de mesure. Le bruit d'observation est modélisé par une distribution gaussienne composée [4], définie comme le produit de deux variables aléatoires indépendantes : un bruit complexe gaussien $\mathcal{CN}(\mathbf{0}, \mathbf{R}_t)$ et une texture aléatoire $\tau_t \sim \text{Gamma}(\nu/2, \nu/2)$ avec $\nu > 2$. Les densités de transition et de vraisemblance s'écrivent alors :

$$p(\mathbf{x}_t | \mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \mathbf{F}_t \mathbf{x}_{t-1}, \mathbf{Q}_t), \quad (3)$$

$$p(\mathbf{y}_t | \mathbf{x}_t) = \mathcal{CT}_\nu(\mathbf{y}_t; \mathbf{H}_t \mathbf{x}_t, \mathbf{R}_t), \quad (4)$$

où \mathcal{CT}_ν désigne la distribution de Student- t complexe à ν degrés de liberté. L'état initial suit une loi gaussienne $\mathbf{x}_0 \sim \mathcal{N}(\mathbf{x}_0; \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$.

Choix de modélisation : Deux approches s'offrent à nous pour traiter la texture τ_t . La première consiste à modéliser τ_t explicitement comme variable latente, préservant ainsi une structure conditionnellement gaussienne du modèle. La seconde vise à intégrer analytiquement pour obtenir une vraisemblance marginale de type Student t . Bien que cette dernière offre une formulation plus concise, elle complexifie l'inférence en induisant des non-linéarités structurelles. Notre choix se porte sur la première approche, qui maintient la tractabilité analytique du modèle gaussien sous-jacent. La vraisemblance complète du modèle complet s'écrit alors :

$$p(\mathbf{y}_{1:T}, \mathbf{x}_{0:T}, \boldsymbol{\tau}_{1:T} | \boldsymbol{\theta}) = p(\mathbf{y}_{1:T} | \mathbf{x}_{0:T}, \boldsymbol{\tau}_{1:T}, \boldsymbol{\theta}) p(\mathbf{x}_{0:T} | \boldsymbol{\theta}) p(\boldsymbol{\tau}_{1:T}) \quad (5)$$

où $\boldsymbol{\theta} = \{\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \mathbf{Q}_t, \mathbf{R}_t\}_{t=1}^T$ désigne l'ensemble des paramètres inconnus du modèle. Bien que des méthodes variationnelles ou à particules puissent être envisagées, nous privilégions ici l'algorithme EM pour son efficacité en grande dimension et sa capacité à estimer conjointement états latents et les paramètres du modèle.

3 Cadre de l'Espérance-Maximisation

L'algorithme EM procède par une estimation itérative des états latents et des paramètres du modèle, en alternant :

- **Étape E (espérance)** : calcul de l'espérance du logarithme de la vraisemblance complète, par rapport à la distribution a posteriori des variables latentes conditionnée par les observations et les paramètres du modèle.
- **Étape M (maximisation)** : mise à jour des paramètres en maximisant la fonction obtenue à l'étape E.

3.1 Formulation EM et approximation stochastique

Formellement, l'étape E de l'algorithme EM consiste à évaluer l'espérance de la log-vraisemblance complète, conditionnellement aux observations et aux paramètres courants $\boldsymbol{\theta}^{(k)}$:

$$\mathcal{Q}(\boldsymbol{\theta} | \boldsymbol{\theta}^{(k)}) = \mathbb{E}_{\mathbf{x}_{0:T}, \boldsymbol{\tau}_{1:T} | \mathbf{y}_{1:T}, \boldsymbol{\theta}^{(k)}} [\log p(\mathbf{y}_{1:T}, \mathbf{x}_{0:T}, \boldsymbol{\tau}_{1:T} | \boldsymbol{\theta})] \quad (6)$$

L'inférence exacte de la loi jointe a posteriori $p(\mathbf{x}_{0:T}, \boldsymbol{\tau}_{1:T} | \mathbf{y}_{1:T}, \boldsymbol{\theta}^{(k)})$ dans les modèles d'espace d'état robustes s'avère généralement intractable en raison de l'interdépendance introduite par les observations entre les états \mathbf{x}_t et les textures τ_t . En effet, bien que ces variables soient a priori indépendantes, conditionner par les observations \mathbf{y}_t induit des dépendances a posteriori non triviales. Pour contourner cette intractabilité, nous exploitons une décomposition hiérarchique de l'espérance \mathcal{Q} comme suit :

$$\mathcal{Q}(\boldsymbol{\theta} | \boldsymbol{\theta}^{(k)}) = \mathbb{E}_{\boldsymbol{\tau}_{1:T} | \mathbf{y}_{1:T}, \boldsymbol{\theta}^{(k)}} [\mathbb{E}_{\mathbf{x}_{0:T} | \mathbf{y}_{1:T}, \boldsymbol{\tau}_{1:T}, \boldsymbol{\theta}^{(k)}} [\log p(\mathbf{y}_{1:T}, \mathbf{x}_{0:T}, \boldsymbol{\tau}_{1:T} | \boldsymbol{\theta})]] \quad (7)$$

avec

$$\begin{aligned} \mathbb{E}_{\mathbf{x}_{0:T} | \mathbf{y}_{1:T}, \boldsymbol{\theta}^{(k)}} [\log p(\mathbf{x}_{0:T} | \boldsymbol{\theta})] = & \\ & - \frac{1}{2} \sum_{t=1}^T \left[\text{tr} \left(\mathbf{Q}_t^{-1} \mathbb{E}_{\mathbf{x}_t | \mathbf{y}_{1:T}, \boldsymbol{\tau}_{1:T}, \boldsymbol{\theta}^{(k)}} [\mathbf{x}_t \mathbf{x}_t^\top] \right) \right. \\ & - 2 \text{tr} \left(\mathbf{Q}_t^{-1} \mathbf{F}_t \mathbb{E}_{\mathbf{x}_t | \mathbf{y}_{1:T}, \boldsymbol{\tau}_{1:T}, \boldsymbol{\theta}^{(k)}} [\mathbf{x}_t \mathbf{x}_{t-1}^\top] \right) \\ & \left. + \text{tr} \left(\mathbf{Q}_t^{-1} \mathbf{F}_t \mathbb{E}_{\mathbf{x}_t | \mathbf{y}_{1:T}, \boldsymbol{\tau}_{1:T}, \boldsymbol{\theta}^{(k)}} [\mathbf{x}_{t-1} \mathbf{x}_{t-1}^\top] \mathbf{F}_t^\top \right) \right] \\ & - \frac{1}{2} \log |\mathbf{Q}| - \frac{n}{2} \log(2\pi). \end{aligned} \quad (8)$$

Sous une réalisation fixée de $\boldsymbol{\tau}_{1:T}$, la distribution a posteriori des états $\mathbf{x}_{0:T}$ est gaussienne et peut être estimée de manière récursive via le lisseur de Rauch–Tung–Striebel (RTS), couramment appelé lisseur de Kalman [16]. Cela permet d'obtenir, pour chaque instant t , les moments conditionnels suivants :

$$\hat{\boldsymbol{\mu}}_t := \mathbb{E}_{\mathbf{x}_t | \mathbf{y}_{1:T}, \boldsymbol{\tau}_{1:T}, \boldsymbol{\theta}^{(k)}} [\mathbf{x}_t], \quad (9a)$$

$$\hat{\boldsymbol{\Sigma}}_t := \mathbb{E}_{\mathbf{x}_t | \mathbf{y}_{1:T}, \boldsymbol{\tau}_{1:T}, \boldsymbol{\theta}^{(k)}} [\mathbf{x}_t \mathbf{x}_t^\top], \quad (9b)$$

$$\hat{\boldsymbol{\Sigma}}_{t,t-1} := \mathbb{E}_{\mathbf{x}_t | \mathbf{y}_{1:T}, \boldsymbol{\tau}_{1:T}, \boldsymbol{\theta}^{(k)}} [\mathbf{x}_t \mathbf{x}_{t-1}^\top]. \quad (9c)$$

En substituant ces estimateurs dans l'expression de \mathcal{Q} et en ne retenant que les termes dépendants de $\boldsymbol{\tau}_{1:T}$, on obtient après calcul :

$$\begin{aligned} \mathcal{Q}(\boldsymbol{\theta} | \boldsymbol{\theta}^{(k)}) = & - \sum_{t=1}^T \mathbb{E}_{\boldsymbol{\tau}_t | \mathbf{y}_t, \boldsymbol{\theta}^{(k)}} \left[m \log \tau_t + \tau_t (\mathbf{y}_t^\dagger \mathbf{y}_t - 2\Re(\mathbf{y}_t^\dagger \mathbf{H} \hat{\boldsymbol{\mu}}_t) \right. \\ & \left. + \text{tr}(\mathbf{H} \hat{\boldsymbol{\Sigma}}_t \mathbf{H}^\dagger) \right] + \sum_{t=1}^T \mathbb{E}_{\boldsymbol{\tau}_t | \mathbf{y}_t, \boldsymbol{\theta}^{(k)}} \left[\left[\left(\frac{\nu}{2} - 1 \right) \log \tau_t - \frac{\nu \tau_t}{2} \right] \right] + C, \end{aligned}$$

où C regroupe les termes constants et les probabilités de transition. L'évaluation analytique de cette espérance reste impossible car la distribution a posteriori $p(\tau_t | \mathbf{y}_{1:T}, \boldsymbol{\theta}^{(k)})$ n'admet pas de forme close. Pour pallier cette difficulté, nous recourons à une variante stochastique de l'algorithme EM (EMS) [14, 5] où l'espérance est approchée par échantillonnage de Monte Carlo. À chaque itération k , l'algorithme EMS met à jour l'estimation de la fonction \mathcal{Q}_s en approximant les espérances requises par des réalisations stochastiques $\tau_{1:T}^s$ générées selon $p(\tau_t | \mathbf{y}_{1:T}, \boldsymbol{\theta}^{(k)})$ par Metropolis-Hastings :

$$\begin{aligned} \mathcal{Q}_s(\boldsymbol{\theta} | \boldsymbol{\theta}^{(k)}) = & -m \sum_{t=1}^T \log \tau_t^s + \sum_{t=1}^T \tau_t^s \left[\mathbf{y}_t^\dagger \mathbf{y}_t - 2\Re(\mathbf{y}_t^\dagger \mathbf{H} \hat{\boldsymbol{\mu}}_t^s) \right. \\ & \left. + \text{tr}(\mathbf{H} \hat{\boldsymbol{\Sigma}}_t^s \mathbf{H}^\dagger) \right] + \sum_{t=1}^T \left[\left(\frac{\nu}{2} - 1 \right) \log \tau_t^s - \frac{\nu \tau_t^s}{2} \right] + C, \end{aligned}$$

où les quantités $\hat{\boldsymbol{\mu}}_t^s = \mathbb{E}_{\mathbf{x}_t | \mathbf{y}_{1:T}, \tau_{1:T}^s, \boldsymbol{\theta}^{(k)}}[\mathbf{x}_t]$ et $\hat{\boldsymbol{\Sigma}}_t^s = \mathbb{E}_{\mathbf{x}_t | \mathbf{y}_{1:T}, \tau_{1:T}^s, \boldsymbol{\theta}^{(k)}}[\mathbf{x}_t \mathbf{x}_t^\top]$ sont estimées par le lisseur de Kalman.

3.2 Estimation par vraisemblance pénalisée

Dans de nombreuses applications, il est souhaitable d'imposer de la parcimonie sur les états latents $\mathbf{x}_{0:T}$. Pour ce faire, nous augmentons l'étape M standard en incluant une régularisation sur la séquence $\mathbf{x}_{0:T}$. Dans ce contexte, $\mathcal{Q}_s(\mathbf{x}_{0:T}, \boldsymbol{\theta} | \boldsymbol{\theta}^{(k)})$ représente le terme d'attache aux données de la fonction objectif à maximiser. Plus précisément, nous considérons les états $\mathbf{x}_{0:T}$ comme des variables d'optimisation *libres* et résolvons :

$$\mathbf{x}_{0:T}^{(k+1)} = \arg \max_{\mathbf{x}_{0:T}} \left\{ \mathcal{Q}_s(\mathbf{x}_{0:T}, \boldsymbol{\theta} | \boldsymbol{\theta}^{(k)}) - \lambda \sum_{t=0}^T \mathcal{R}(\mathbf{x}_t) \right\}, \quad (10)$$

où $\lambda > 0$ est un paramètre de régularisation, et $\mathcal{Q}_s(\mathbf{x}_{0:T}, \boldsymbol{\theta} | \boldsymbol{\theta}^{(k)})$ est l'approximation EMS de la log-vraisemblance complète, en fonction des états $\mathbf{x}_{0:T}$. Nous maintenons les états latents comme variables libres, plutôt que de substituer directement par les estimations obtenues à l'étape E, ces dernières servant uniquement d'initialisation pour le sous-problème d'optimisation à l'étape M. Il convient de noter que cette stratégie dite de « vraisemblance pénalisée » ne coïncide pas entièrement avec l'EM standard pour l'estimation du Maximum A Posteriori (MAP), dans la mesure où les états $\mathbf{x}_{0:T}$ ont été initialement intégrés dans l'étape E. Bien que cette approche induise une estimation d'état avec une pénalisation favorisant la parcimonie, elle ne constitue pas une solution MAP stricte pour l'ensemble du modèle bayésien. Néanmoins, elle représente, d'un point de vue pratique, une méthode efficace et modulaire pour intégrer la parcimonie à chaque itération.

Pour résoudre l'équation (10), nous calculons le gradient de $\mathcal{Q}_s(\mathbf{x}_{0:T}, \boldsymbol{\theta} | \boldsymbol{\theta}^{(k)})$ par rapport à \mathbf{x}_t . Celui-ci s'écrit :

$$\begin{aligned} \nabla_{\mathbf{x}_t} \mathcal{Q}_s = & 2 \tau_t^s (\mathbf{H}_t^\top \mathbf{y}_t - \mathbf{H}_t^\top \mathbf{H}_t \mathbf{x}_t) \\ & + \begin{cases} -\Sigma_0^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_0) + \mathbf{F}_2^\top \mathbf{Q}_2^{-1} (\mathbf{x}_2 - \mathbf{F}_2 \mathbf{x}_1), & t = 0, \\ -\mathbf{Q}_t^{-1} (\mathbf{x}_t - \mathbf{F}_t \mathbf{x}_{t-1}) + \mathbf{F}_{t+1}^\top \mathbf{Q}_{t+1}^{-1} (\mathbf{x}_{t+1} - \mathbf{F}_{t+1} \mathbf{x}_t), & 1 \leq t < T, \\ -\mathbf{Q}_T^{-1} (\mathbf{x}_T - \mathbf{F}_T \mathbf{x}_{T-1}), & t = T. \end{cases} \end{aligned}$$

En raison de la nature non différentiable du terme de pénalisation, la mise à jour des états repose sur un gradient proximal :

$$\mathbf{x}_t^{(k+1)} = \text{prox}_{\gamma \lambda \|\cdot\|_1} \left(\mathbf{x}_t^{(k)} + \gamma \nabla_{\mathbf{x}_t} \mathcal{Q}_s(\mathbf{x}_{0:T}, \boldsymbol{\theta} | \boldsymbol{\theta}^{(k)}) \right), \quad (11)$$

où l'opérateur proximal associé à la pénalisation ℓ_1 est défini, composante par composante, par la fonction de seuillage $\mathcal{T}_{\gamma \lambda} : \mathbb{R}^n \rightarrow \mathbb{R}^n$, définie par $\mathcal{T}_\alpha(x)_i = \text{sgn}(x_i) \max(|x_i| - \gamma \lambda, 0)$.

4 Application à la reconstruction d'images radio interférométriques

4.1 Configuration expérimentale

Pour évaluer la performance de l'approche EMS robuste proposée, nous nous intéressons à une application en imagerie radio-interférométrique, dont l'objectif est de reconstruire des images du ciel à partir de mesures de visibilité incomplètes et bruitées [18]. Les interféromètres radio mesurent des composantes de Fourier de la distribution d'intensité du ciel, et la restauration de l'image originale exploite la relation de transformation de Fourier entre l'image (du ciel) et les visibilité, par la relation de Van-Cittert-Zernicke [18]. La méthode proposée se révèle particulièrement adaptée à ce contexte en raison de la présence de bruit non gaussien, pouvant provenir par exemple des RFI. Nous fixons l'horizon temporel à $T = 10$, et considérons une séquence d'images astronomiques $\mathbf{x}_{0:T}$ de dimension $n = 64 \times 64$. L'image initiale \mathbf{x}_0 est modélisée par des distributions gaussiennes bivariées, puis les états ultérieurs sont générés selon une marche aléatoire gaussienne, $\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_t$, avec $\mathbf{w}_t \sim \mathcal{N}(0, \alpha \mathbf{I}_n)$ et un pas de variance $\alpha = 10^{-4}$. À chaque instant t , les visibilité $\mathbf{y}_{1:T}$ sont générées à partir du modèle direct (2), où $\mathbf{H} \in \mathbb{C}^{m \times n}$ est l'opérateur de mesure correspondant au schéma d'échantillonnage de Fourier incomplet du radio interféromètre. L'expression analytique de \mathbf{H} est donnée par

$$\mathbf{H} = \exp \left(\frac{-j2\pi}{\lambda_w} \boldsymbol{\Delta}_r \mathbf{L}^T \right), \quad (12)$$

où λ_w est la longueur d'onde d'observation, $\boldsymbol{\Delta}_r$ est la matrice des lignes de base (baselines) de dimension $m \times 2$ contenant les différences de positions en coordonnées cartésiennes entre chaque paire d'antennes, et \mathbf{L} est la matrice de dimension $n \times 2$ contenant les vecteurs directeurs (l, m) pour chacun des n points sources (pixels). Dans un contexte pratique d'imagerie radio-interférométrique, les fréquences spatiales échantillonnées correspondent directement aux lignes de base du réseau d'antennes (normalisées par la longueur d'onde des signaux considérés). Pour reproduire un scénario réaliste, nous synthétisons les visibilité en nous basant sur le plan uv du Very Large Array (VLA), constitué de $p = 27$ antennes et opérant à une fréquence de 3,8 GHz. À chaque instant t , $m = 351$ visibilité sont mesurées.

Initialisation : L'estimation initiale de l'état $\boldsymbol{\mu}_0$ est obtenue à partir d'une *dirty image*, calculée en appliquant la transformée de Fourier inverse aux visibilité mesurées. La matrice de covariance $\Sigma_0 = 10^{-3} \mathbf{I}_m$ est choisie empiriquement et le paramètre de régularisation λ est sélectionné par validation croisée. Le paramètre de degré de liberté $\nu = 2,5$, supposé connu, contrôle la lourdeur des queues de la distribution : plus ν est petit, plus la distribution est à queues lourdes, ce qui renforce la robustesse aux valeurs aberrantes.

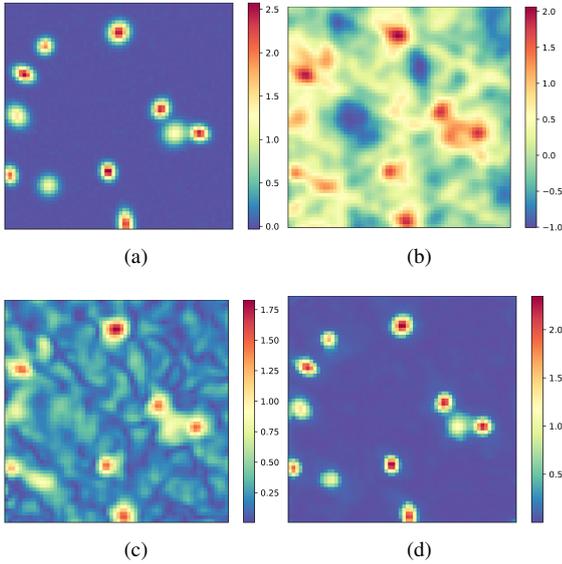


FIGURE 1 : Résultats de reconstruction à l’instant T : (a) image de référence, (b) initialisation (dirty image), (c) image estimée par lisseuse de Kalman, (d) image estimée par EMS robuste.

TABLE 1 : Comparaison des performances de reconstruction

Méthode	MSE ↓	SSIM ↑	PSNR ↑
EM gaussien	0.327	0.638	28.801 dB
Lisseuse de Kalman	0.467	0.196	21.632 dB
EMS proposé	0.019	0.804	35.09 dB

4.2 Résultats numériques

La robustesse de la méthode proposée a été évaluée sur des données radio-interférométriques contaminées par des RFI simulées selon le modèle décrit dans [13, 2]. Celles-ci, ajoutées de manière additive aux visibilitées synthétiques, possèdent une puissance 4 fois supérieure à celle du signal astronomique. Nous avons comparé l’algorithme proposé à deux approches de référence : (1) un algorithme EM gaussien standard tel que présenté dans [17], conçu pour l’estimation d’espaces d’état linéaires dans un cadre strictement gaussien, et (2) un lisseuse de Kalman dit *idéal*, supposant les paramètres du modèle connus. Le Tableau 1 synthétise les performances obtenues. L’approche proposée réduit l’erreur moyenne quadratique (MSE) d’un facteur 17 et améliore rapport signal sur bruit instantané (PSNR) de 7 dB comparé aux méthodes de référence. La mesure de similarité structurelle (SSIM) élevée (0.804) atteste de la préservation des structures spatiales fines.

Ces améliorations quantitatives sont corroborées par une inspection visuelle des images reconstruites à l’instant T (Fig. 1), qui révèle une fidélité d’image nettement améliorée et une qualité perceptuelle supérieure en présence de bruit induit par des RFI.

5 Conclusion

Ce travail introduit une méthode d’estimation robuste pour les modèles d’espace d’état linéaires affectés par un bruit impulsif, en combinant un échantillonnage stochastique des textures latentes avec une régularisation favorisant la parcimonie des

états. Cette approche permet de dépasser les limitations des méthodes classiques fondées sur l’hypothèse gaussienne, souvent inadéquate en présence de valeurs aberrantes. Les résultats obtenus en radio-interférométrie montrent une nette amélioration de la qualité de reconstruction. Ces avancées ouvrent des perspectives pour l’analyse de données réelles issues de grands réseaux d’antennes, ainsi que pour l’extension vers des modèles non linéaires ou multi-échelles adaptés à des scénarios d’observation complexes.

Références

- [1] G. AGAMENNONI, J. I. NIETO et E. M. NEBOT : Approximate inference in state-space models with heavy-tailed noise. *IEEE Transactions on Signal Processing*, 60(10):5024–5037, 2012.
- [2] N. ARAB, Y. MHIRI, I. VIN, M. N. EL KORSO et P. LARZABAL : Unrolled expectation maximization algorithm for radio interferometric imaging in presence of non gaussian interferences. *Signal Processing*, 2025.
- [3] R. E. CARRILLO, J. D. MCEWEN et Y. WIAUX : Sparsity averaging reweighted analysis (sara) : A novel algorithm for radio interferometric imaging. *Monthly Notices of the Royal Astronomical Society*, 439(4): 3591–3604, 2013.
- [4] J.-P. DELMAS, M. N. EL KORSO, F. PASCAL et S. FORTUNATI : *Elliptically Symmetric Distributions in Signal Processing and Machine Learning*. Springer Nature, 2024.
- [5] B. DELYON, M. LAVIELLE et E. MOULINES : Convergence of a stochastic approximation version of the em algorithm. *Annals of Statistics*, 27(1):94–128, 1999.
- [6] A. P. DEMPSTER, N. M. LAIRD et D. B. RUBIN : Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society, Series B (Methodological)*, 39(1):1–38, 1977.
- [7] A. DOUCET, A. M. JOHANSEN *et al.* : A tutorial on particle filtering and smoothing : Fifteen years later. In *Handbook of Nonlinear Filtering*, volume 12, pages 656–704. 2009.
- [8] V. ELVIRA et É. CHOUZENOUX : Graphical inference in linear-gaussian state-space models. *IEEE Transactions on Signal Processing*, 70:4757–4771, 2022.
- [9] M. A. GANDHI et L. MILI : Robust kalman filter based on a generalized maximum-likelihood-type estimator. *IEEE Transactions on Signal Processing*, 58(5):2509–2520, 2010.
- [10] Z. GHAHRAMANI et G. E. HINTON : Parameter estimation for linear dynamical systems. 1996.
- [11] J. D. HAMILTON : State-space models. In *Handbook of Econometrics*, volume 4, pages 3039–3080. 1994.
- [12] A. LESHEM et A.-J. van der VEEN : Introduction to interference mitigation techniques in radio astronomy. In *Perspectives on Radio Astronomy : Technologies for Large Antenna Arrays*, page 201. 2000.
- [13] A. LESHEM et A.-J. van der VEEN : Radio astronomical imaging in the presence of strong radio interference. *IEEE Transactions on Information Theory*, 2000.
- [14] G. J. MCLACHLAN et T. KRISHNAN : *The EM Algorithm and Extensions*. Wiley, 2 édition, 2008.
- [15] A. K. ROONIZI : Kalman filtering in non-gaussian model errors : A new perspective [tips & tricks]. *IEEE Signal Processing Magazine*, 39(3):105–114, May 2022.
- [16] S. SÄRKÄ : *Bayesian Filtering and Smoothing*. Cambridge University Press, 3 édition, 2013.
- [17] R. H. SHUMWAY et D. S. STOFFER : An approach to time series smoothing and forecasting using the em algorithm. *Journal of Time Series Analysis*, 3(4):253–264, 1982.
- [18] A. R. THOMPSON, J. M. MORAN et G. W. SWENSON : *Interferometry and Synthesis in Radio Astronomy*. Springer Nature, 2017.
- [19] Y. WIAUX, L. JACQUES, G. PUY, A. M. SCAIFE et P. VANDERGHEYNST : Compressed sensing imaging techniques for radio interferometry. *Monthly Notices of the Royal Astronomical Society*, 395(3):1733–1742, 2009.
- [20] J. ZHAO, M. NETTO et L. MILI : A robust iterated extended kalman filter for power system dynamic state estimation. *IEEE Transactions on Power Systems*, 32(4):3205–3216, 2017.