



HAL
open science

Mixture of noises and sampling non log-concave distributions

Pierre Palud, Franck Le Petit, Pierre Chainais, Emeric Bron, Pierre-Antoine Thouvenin

► **To cite this version:**

Pierre Palud, Franck Le Petit, Pierre Chainais, Emeric Bron, Pierre-Antoine Thouvenin. Mixture of noises and sampling non log-concave distributions. XXVIIIème Colloque Francophone de Traitement du Signal et des Images (GRETSI 2022), Sep 2022, Nancy, France. . hal-03952453

HAL Id: hal-03952453

<https://hal.science/hal-03952453>

Submitted on 23 Jan 2023

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.



Distributed under a Creative Commons Attribution 4.0 International License

1/ Aims and contributions

Real life inverse problems: **multiple noises**, **non-linear** models, need for **uncertainty** quantification.

Bayesian inference with

- Mixture of noises
- Non-log-concave posterior (multimodal)
- Non-gradient-Lipschitz log-posterior

- ★ Likelihood approximation for a mixture of noises,
- ★ New kernel to efficiently sample from posterior,
- ★ Application to astrophysical data.

2/ Observation model

→ **Forward model** $f : \Theta \in \mathbb{R}^{ND} \mapsto \mathbf{Y} \in \mathbb{R}^{NL}$, strictly positive, twice differentiable, covers multiple decades.

→ **Observation model:**

$$y_{n,\ell} = \max \left\{ \omega, \epsilon_{n,\ell}^{(m)} f_{\ell}(\theta_n) + \epsilon_{n,\ell}^{(a)} \right\}$$

with $\begin{cases} \omega > 0 \\ \epsilon_{n,\ell}^{(a)} \sim \mathcal{N}(0, \sigma_a^2) \text{ i.i.d.} \\ \epsilon_{n,\ell}^{(m)} \sim \log \mathcal{N}(0, \sigma_m^2) \text{ i.i.d.} \end{cases}$

⇒ untractable likelihood $\pi(\mathbf{Y} | \Theta)$.

→ Literature often neglects one noise [1]. But when f covers many decades: dominant noise depends on θ_n .
⇒ need to address full mixture model.

How to **approximate the likelihood** with **controlled error**? → See 4/

3/ Bayesian approach & sampling

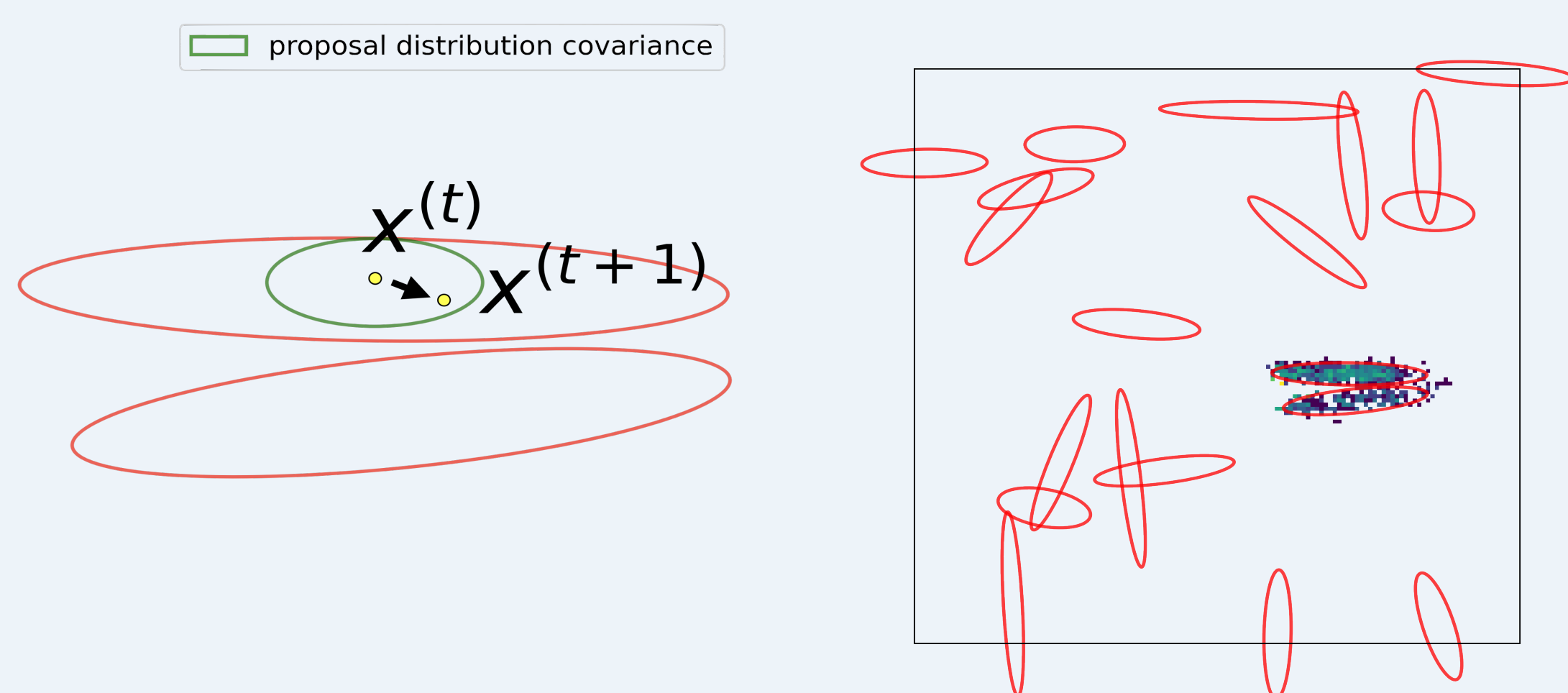
→ Combine likelihood with informative **prior** $\pi(\Theta)$: yields **Posterior** distribution:

$$\pi(\Theta | \mathbf{Y}) \propto \pi(\mathbf{Y} | \Theta) \pi(\Theta)$$

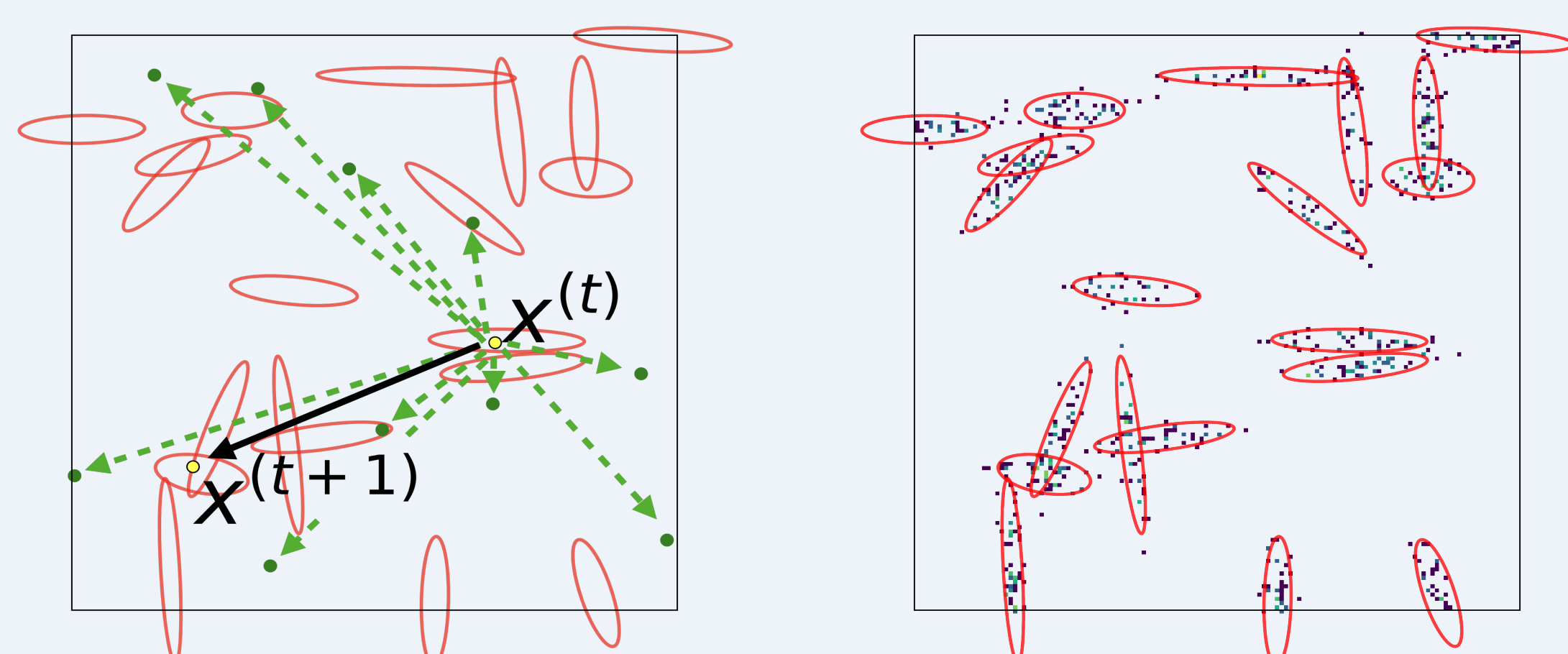
→ Sampling: MCMC
⇒ uncertainty quantification

→ Model: Non-linear
⇒ **multimodal** distribution

✓ Preconditioned-MALA [2]: local exploration



✓ MTM [3]: jumps between minima.



4/ Mixture of noises: Likelihood approximation

1. Build 2 likelihood approximations with moment matching:

Gaussian approx $\pi^{(a)}$	lognormal approx $\pi^{(m)}$
$y_{n,\ell} \simeq f_{\ell}(\theta_n) + e_{n,\ell}^{(a)}$	$y_{n,\ell} \simeq e_{n,\ell}^{(m)} f_{\ell}(\theta_n)$
$e_{n,\ell}^{(a)} \sim \mathcal{N}(m_{a,n,\ell}, s_{a,n,\ell}^2)$	$e_{n,\ell}^{(m)} \sim \log \mathcal{N}(m_{m,n,\ell}, s_{m,n,\ell}^2)$

2. Combine approximations with weight function $\lambda_{\ell}^{a_{\ell}}$:

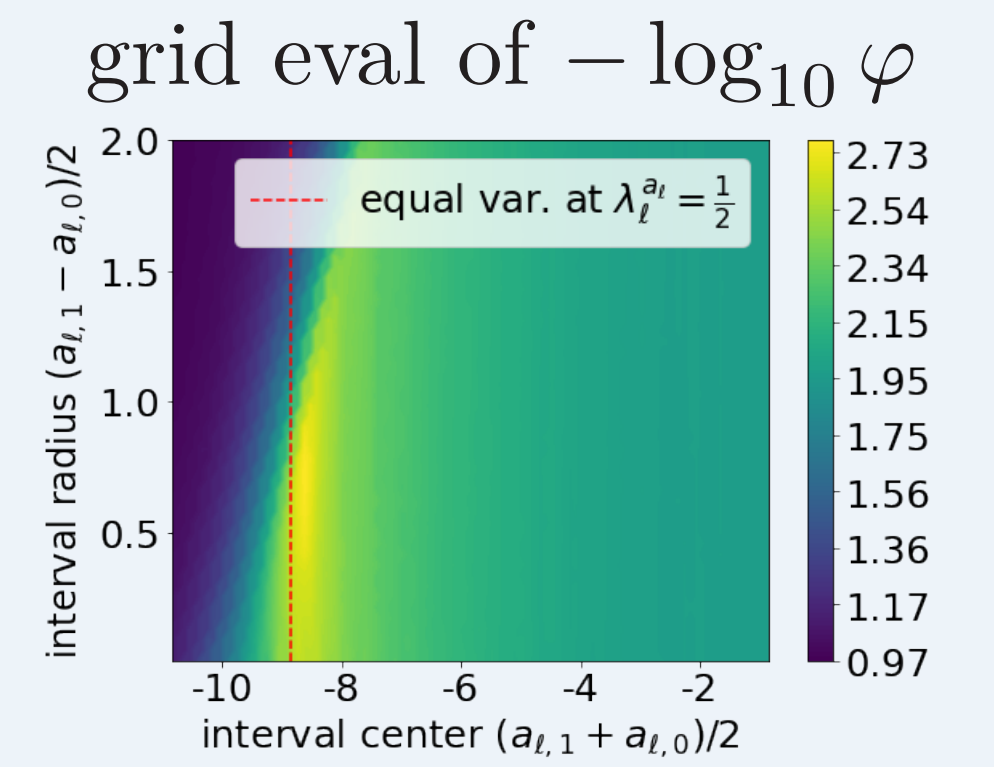
$$\tilde{\pi}_{a_{\ell}}(y_{n,\ell} | \theta_n) \propto \underbrace{\pi^{(a)}(y_{n,\ell} | \theta_n)}_{\text{Gaussian approx}}^{1 - \lambda_{\ell}^{a_{\ell}}(\theta_n)} \underbrace{\pi^{(m)}(y_{n,\ell} | \theta_n)}_{\text{lognormal approx}}^{\lambda_{\ell}^{a_{\ell}}(\theta_n)}$$

with $\lambda_{\ell}^{a_{\ell}} \in [0, 1]$, twice differentiable, parametrized with $\mathbf{a}_{\ell} \in \mathbb{R}^2$
⇒ negative log of this approx: easy to work with.

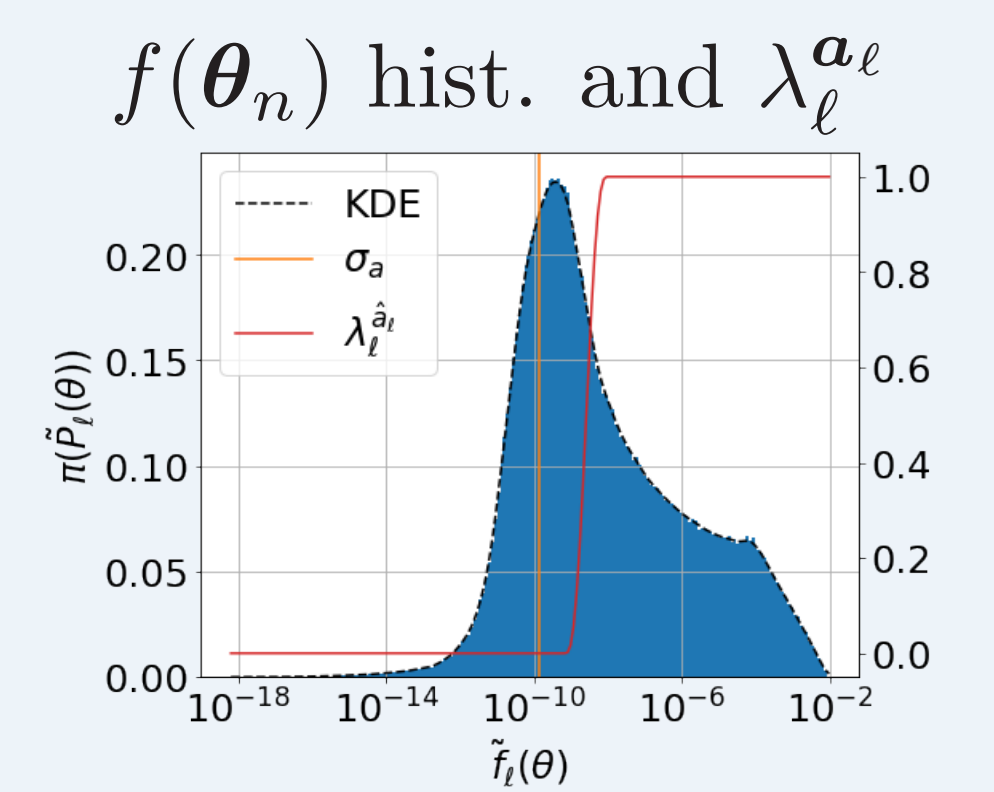
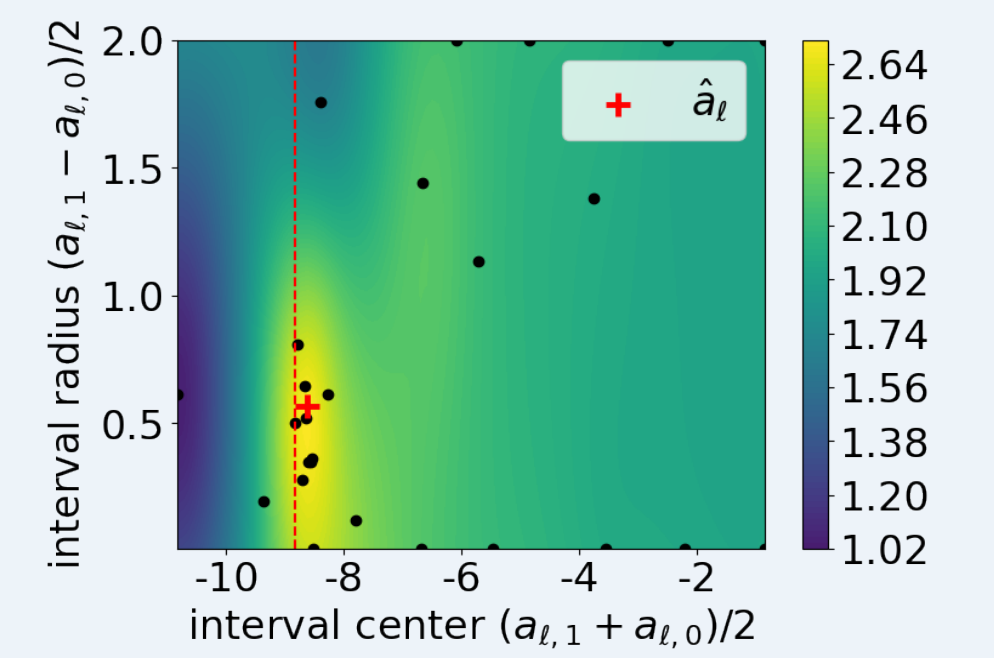
Tuning \mathbf{a}_{ℓ} :

Approx error = Kolmogorov-Smirnov-based metric $\varphi(\mathbf{a}_{\ell})$.
 φ minimized with Bayesian Optimization (BO).

→ From approx error: $\tilde{\pi}_{a_{\ell}}$ **better than** $\pi^{(a)}$ **or** $\pi^{(m)}$ **alone.**



BO, 25 points



5/ Astrophysics synthetic dataset

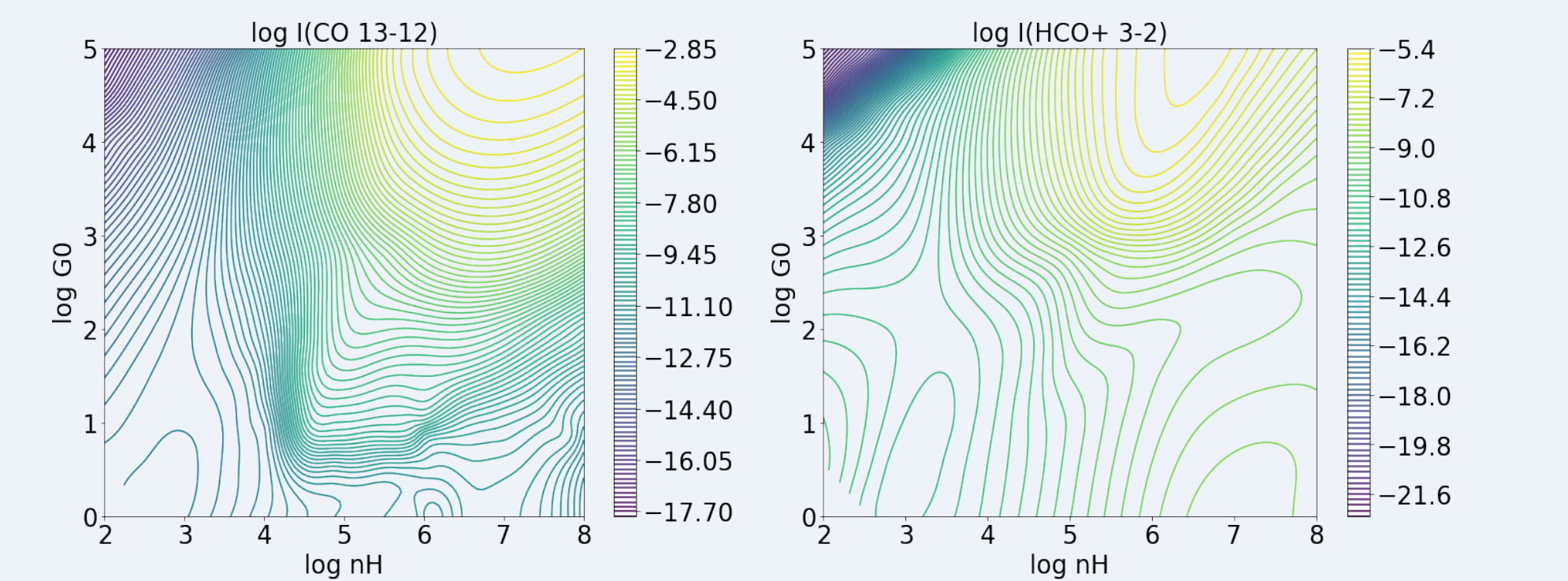
→ $\Theta \in \mathbb{R}^{900 \times 4}$: high dimensional

→ $\pi(\Theta)$: spatial (Laplacian L_2 -norm)

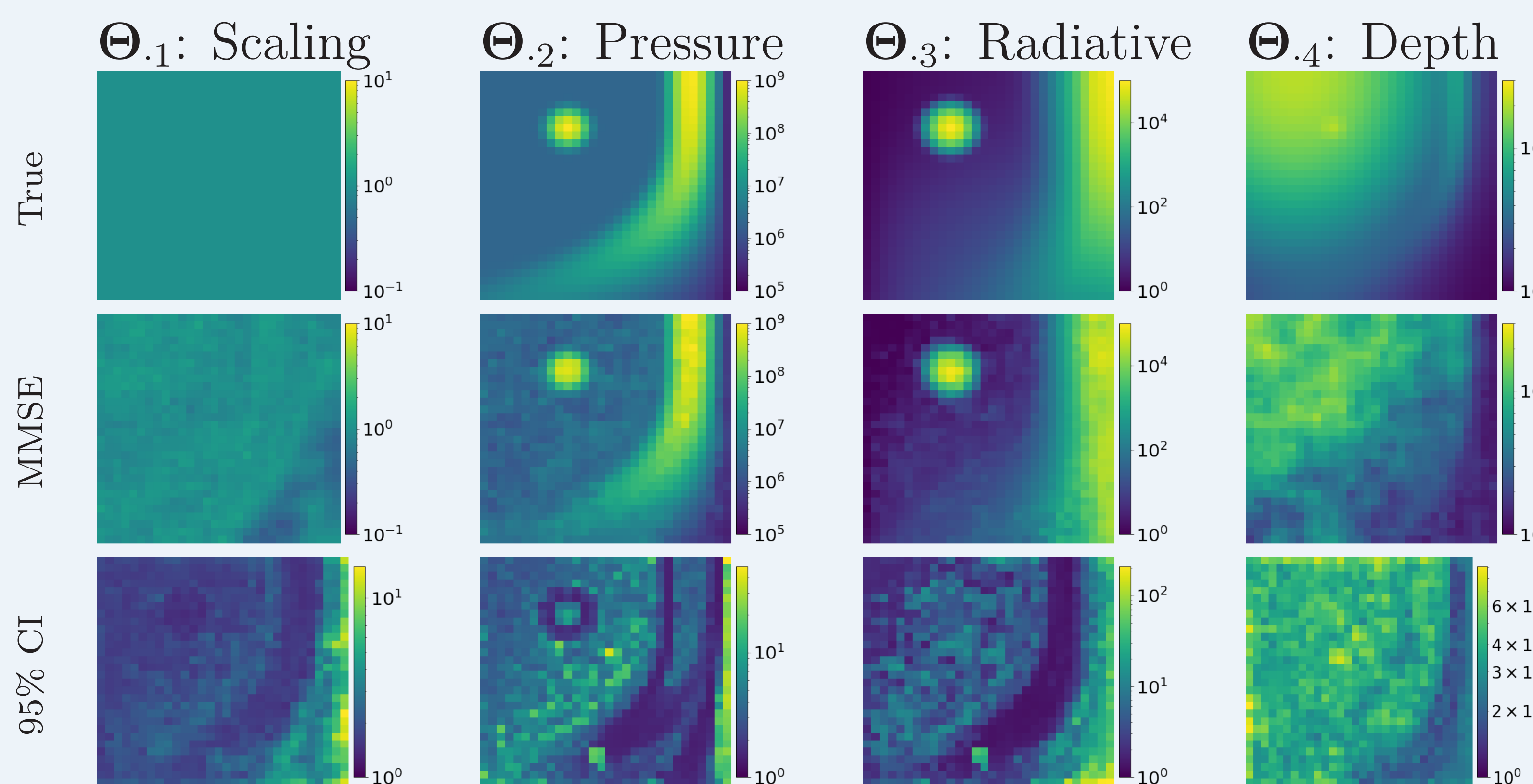
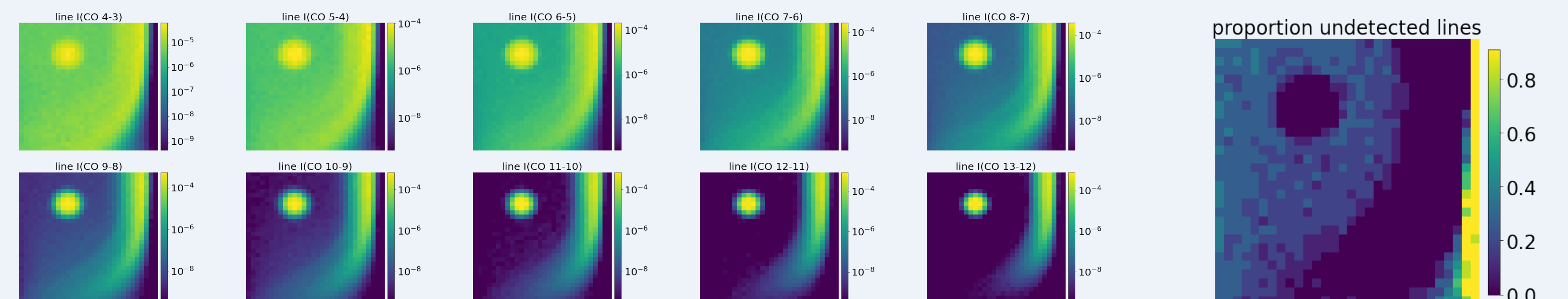
Forward model: Meudon PDR code [4]

✗ non-linear

✗ non-gradient-Lipschitz

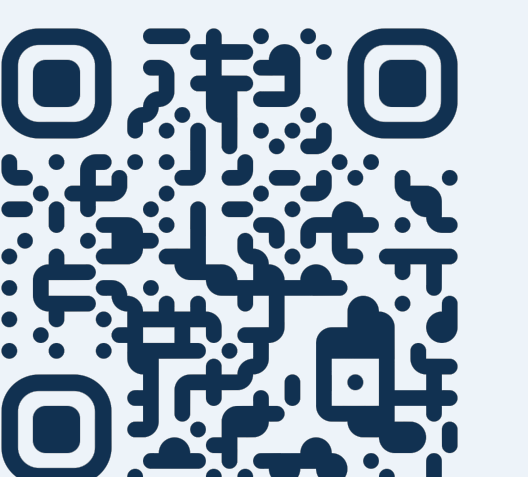


Synthetic observations $\mathbf{Y} \in \mathbb{R}^{900 \times 10}$: integrated intensities of excited lines of CO



6/ Conclusion

- ✓ Mixture of noises ⇒ likelihood approx.
- ✓ Non-log-concave posterior ⇒ MTM kernel.
- ✓ Non-gradient-Lipschitz log-posterior ⇒ P-MALA kernel.
- ✓ Application on astrophysical inverse problem.
- Application to Orion-B data, James Webb Spatial Telescope, etc.



References

- [1] Nicholson and Kaipio, *An additive Approximation to Multiplicative Noise*, 2018
- [2] Xifara et al., *Langevin diffusions and the Metropolis-adjusted Langevin algorithm*, 2014
- [3] Martino, *A review of multiple try MCMC algorithms for signal processing*, 2018
- [4] Le Petit et al., *A Model for Atomic and Molecular Interstellar Gas: The Meudon PDR Code*, 2006