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
Mixture of noises and sampling non-log-concave distributions

Pierre Palud1,2, F. Le Petit3, P. Chainais4, E. Bron1, P.-A. Thouvenin2, ORIONB consortium
1 LERMA, Obs. de Paris, PSL Research Univ., CNRS, Meudon
2 Univ. Lille, CNRS, Centrale Lille, UMR 9189 - CRISiL, France

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 \( f : \Theta \in \mathbb{R}^{ND} \mapsto Y \in \mathbb{R}^{NL} \); strictly positive, twice differentiable, covers multiple decades.
→ Observation model:

\[
y_{n,t} = \max \{ \omega, \epsilon_{n,t}^{(a)} f_r(\theta_n) + \epsilon_{n,t}^{(s)} \}
\]

with
\[
\begin{align*}
\omega & > 0, \\
\epsilon_{n,t}^{(a)} & \sim \mathcal{N}(0, \sigma_n^2) \quad \text{i.i.d.}
\end{align*}
\]
\[
\begin{align*}
\epsilon_{n,t}^{(s)} & \sim \log \mathcal{N}(0, \sigma_n^2) \quad \text{i.i.d.}
\end{align*}
\]

⇒ untractable likelihood \( \pi(Y | \Theta) \).
→ Literature often neglects one noise [1]. But when \( f \) covers many decades: dominant noise depends on \( \theta_n \).
⇒ 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:

\[
\pi(\Theta | Y) \propto \pi(Y | \Theta) \pi(\Theta)
\]

→ 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:

\[
\begin{align*}
\text{Gaussian approx } \pi^{(a)}(y_{n,t} | \theta_n) & \sim f_r(\theta_n) + \epsilon_{n,t}^{(a)} \\
\text{lognormal approx } \pi^{(m)}(y_{n,t} | \theta_n) & \sim \log \mathcal{N}(f_r(\theta_n), \sigma_{n,t}^2)
\end{align*}
\]

2. Combine approximations with weight function \( \lambda_n^\ell_\ell \):

\[
\tilde{\pi}_n(y_{n,t} | \theta_n) \propto \pi^{(a)}(y_{n,t} | \theta_n)^{1-\lambda_n^\ell_\ell} \pi^{(m)}(y_{n,t} | \theta_n)^{\lambda_n^\ell_\ell}
\]

with \( \lambda_n^\ell_\ell \in [0,1] \), twice differentiable, parametrized with \( a_\ell \in \mathbb{R}^2 \)
⇒ negative log of this approx: easy to work with.

Tuning \( a_\ell \):
Approx error = Kolmogorov-Smirnov-based metric \( \varphi(a_\ell) \).
\( \varphi \) minimized with Bayesian Optimization (BO).

→ From approx error: \( \tilde{\pi}_n \) better than \( \pi^{(a)} \) or \( \pi^{(m)} \) alone.

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]
\( \times \) non-linear
\( \times \) non-gradient-Lipschitz

Synthetic observations \( Y \in \mathbb{R}^{900 \times 10} \); integrated intensities of excited lines of CO

6/ Conclusion
✓ Mixture of noises \( \Rightarrow \) likelihood approx.
✓ Non-log-concave posterior \( \Rightarrow \) MTM kernel.
✓ Non-gradient-Lipschitz \( \Rightarrow \) P-MALA kernel.
✓ Application on astrophysical inverse problem.
→ Application to Orion-B data, James Webb Spatial Telescope, etc.

References