DTU Compute
Department of Applied Mathematics and Computer Science

UQ for a Posterior with Besov Priors

Andreas Horst[1,*], Babak Afkham[1] , Yiqiu Dong[1], Jakob Lemvig[1]

[1] Technical University of Denmark, DTU Compute, CUQI.

1 Introduction

Besov priors are a family of wavelet based priors used in Bayesian inversion as a model that promotes Besov space function regularity, for example, a Besov prior can promote discontinuous functions. **Besov priors** are important since:

- they are **discretization invariant**
- they can **preserve edges** in imaging applications
- they can promote **sparsity** in a wavelet basis

2 The problem

Consider a discrete linear inverse problem

$$y = Af + \epsilon, \tag{1}$$

- $A \in \mathbb{R}^{m\times n}$ is the linear forward operator.
- $f \in \mathbb{R}^n$ is the unknown assumed to follow a **Besov prior**.
- $y \in \mathbb{R}^m$ is the observed data.
- $\epsilon \sim \mathcal{N}(0, \sigma^2 I_m)$ is the noise with $\sigma > 0$.

We seek the solution to eq. (1) in terms of the posterior distribution

Posterior

$$\pi(f|y) \propto \exp(-\tfrac{1}{2\sigma^2}\|Af - y\|_2^2 - \lambda\|Bf\|_p^p), \tag{2}$$

- $B \in \mathbb{R}^{n\times n}$ is a **weighted wavelet transform**

$$Bf = \left[c_0, w_{0,0}, \cdots, 2^{j(s-1/2-1/p)}w_{j,k}, \cdots\right]^T. \tag{3}$$

- $\{c_0, w_{j,k}\}$ are discrete wavelet coefficients associated with a suitable wavelet basis $\{\psi_{j,k}\}$.
- $\lambda, s > 0$, and $p \geq 1$ are prior parameters.

Goal

Sample from the posterior in eq. (2).

3 Sampling Method

The randomize-then-optimize method [1] (RTO) draws samples from a target posterior

$$\pi_{\text{trg}}(f|y) \propto \exp(-\frac{1}{2}\|F(f)\|_2^2), \tag{4}$$

where $F(f) : \mathbb{R}^n \to \mathbb{R}^{n+m}$.

RTO procedure

1. Draw proposal sample by solving

$$f_{\text{prop}} = \arg\min_f \frac{1}{2}\|Q^T F(f) - \xi\|_2^2, \tag{5}$$

where $Q \in \mathbb{R}^{(m+n)\times n}$ is an orthogonal projection and $\xi \sim \mathcal{N}(0, I_n)$.

2. Correct the proposal sample to the target with the Metropolis-Hastings algorithm.

RTO properties

- Optimization based.
- Independence sampler.

Challenge

RTO only works for posteriors on the form eq. (4).

We introduce a **prior transform** from Besov f to standard Gaussian $\tilde{f}$ given by

$$\tilde{f} = g^{-1}(Bf), \tag{6}$$

where $g : \mathbb{R}^n \to \mathbb{R}^n$ is an invertible mapping from a p-Gaussian to a standard Gaussian constructed by an **inverse CDF method**. Changing variables using the prior transformation yields a **transformed posterior** on the form

$$\pi(\tilde{f}|y) \propto \exp\left(-\tfrac{1}{2\sigma^2}\|AB^{-1}g(\tilde{f}) - y\|_2^2 - \tfrac{1}{2}\|\tilde{f}\|_2^2\right). \tag{7}$$

4 Numerical Test

For the numerical test, we will use an **inpainting problem**. The inpainting problem is to recover features of a signal in regions of missing measurements. The test scenario is shown in fig. 1.

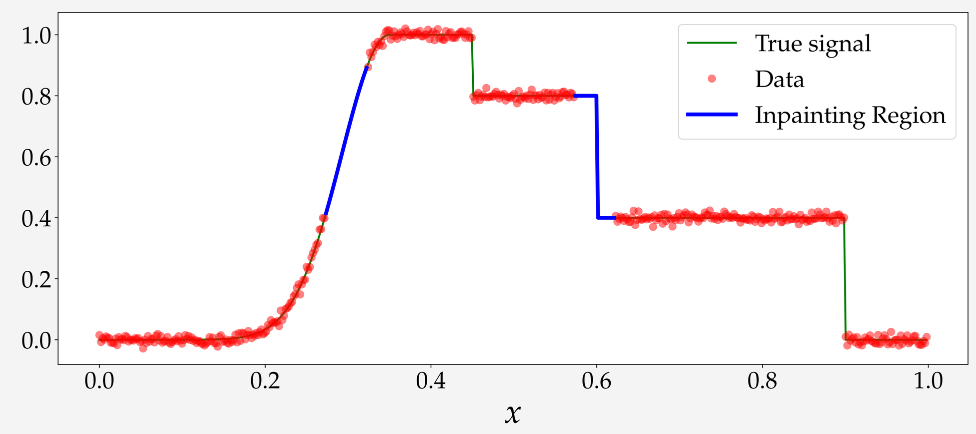

Figure 1: True signal, noisy data, and inpainting region.

We conduct 2 numerical tests with the following setup:

- For both tests, we draw 10000 samples from the posterior using RTO and compute the posterior mean and the posterior 95% credible interval (CI)
- Case 1:
  - Fixing $s = 1.2$, $p = 1.5$, and $\lambda = 0.025$.
  - Varying the wavelet basis $\{\psi_{j,k}\}$ between Daubechies-1 (discontinuous) and Daubechies-8 (smooth) wavelets [2].

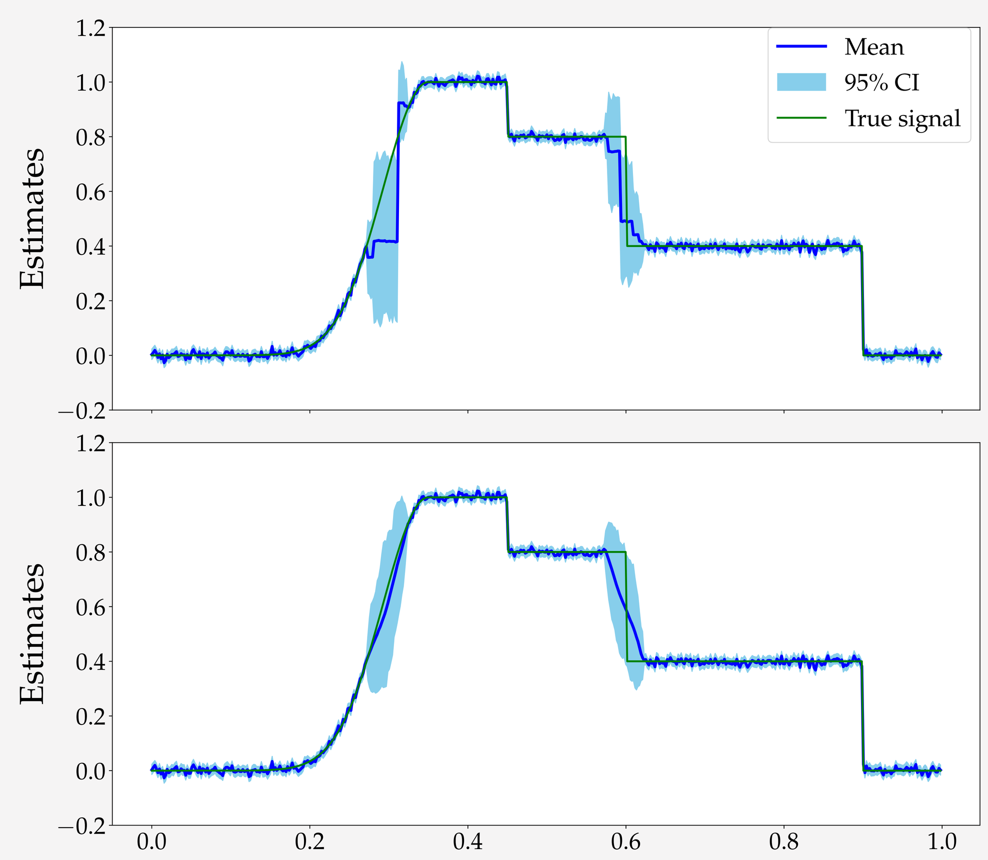

Figure 2: Case 1: Posterior estimates for the Daubechies-1 wavelet case (top) and for the Daubechies-8 wavelet case (bottom).

Notice that the Daubechies-1 wavelet prior fills the inpainting region with a jump while the Daubechies-8 wavelet prior fills the inpainting region with a smooth transition, as seen in fig. 2.

- Case 2:
  - Fixing the wavelet basis $\{\psi_{j,k}\}$ to Daubechies-8 and $\lambda = 0.025$.
  - Varying the parameters between $s \in \{0.8, 1.4\}$ and $p \in \{1.5, 2.0\}$.

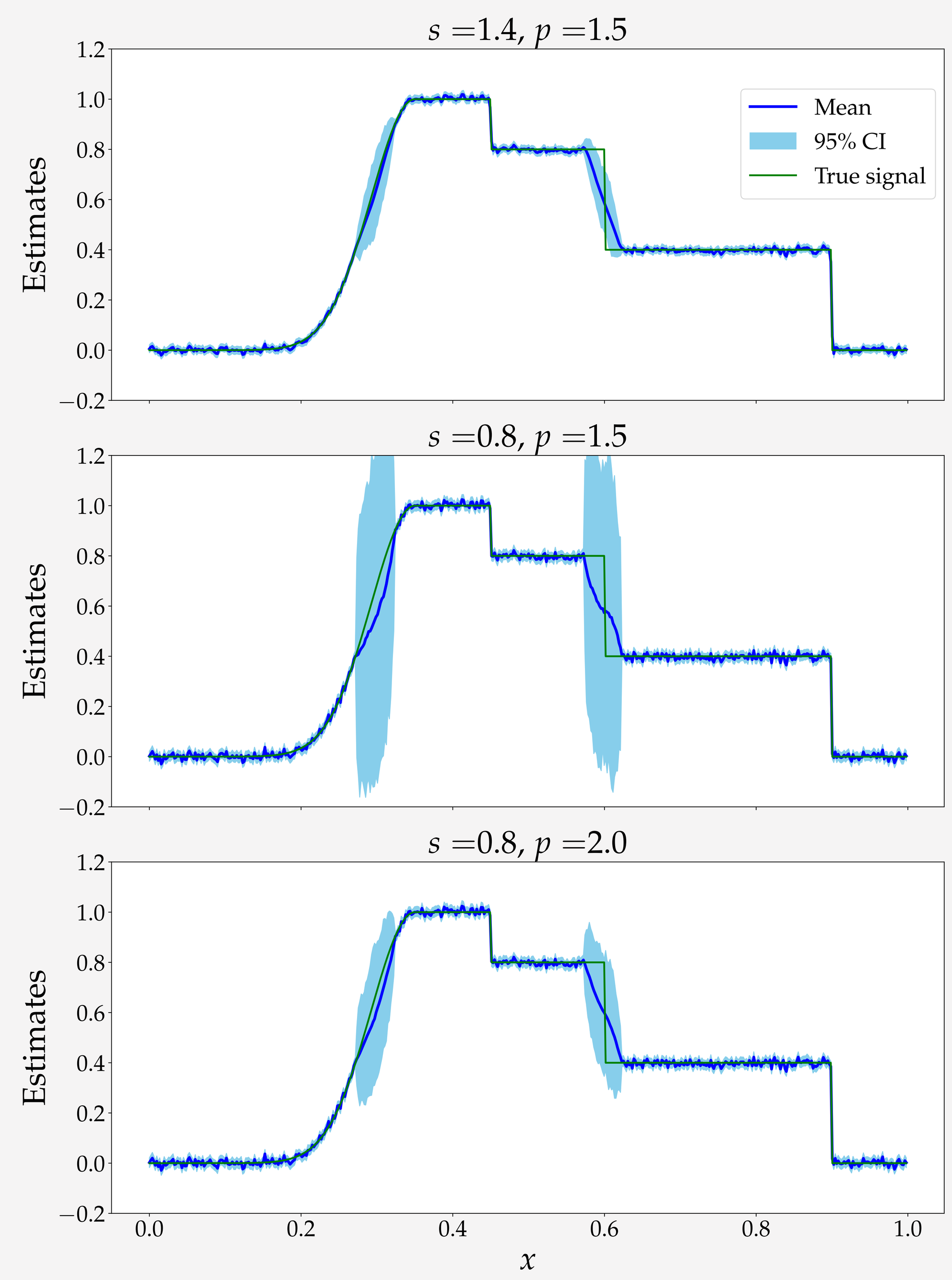

Figure 3: Case 2: Posterior estimates where the base case ($s = 0.8$, $p = 1.5$) is in the middle and s increases in the top while p increases in the bottom with respect to the base case.

Notice that increasing s and p reduces the width of the 95% CI and makes the posterior mean get closer to the true signal in the smooth part of the inpainting region, as seen in fig. 3.

5 Conclusion

The numerical results suggest that:

- The smoothness of the wavelet basis determines the smoothness of the posterior mean estimate.
- The prior parameters regularizes the posterior estimates and concentrates the posterior around the mean.

References

[1] J. M. Bardsley. "Randomize-Then-Optimize: A Method for Sampling from Posterior Distributions in Nonlinear Inverse Problems". In: *SIAM Journal on Scientific Computing* 36.4 (2014), A1895–A1910.

[2] I. Daubechies. "Orthonormal bases of compactly supported wavelets". In: *Communications on Pure and Applied Mathematics* 41.7 (1988), pp. 909–996.