

## I/ Information theory in the Bayesian framework

Within the Bayesian framework, the prior distribution is updated by the observed statistics to provide a posterior distribution thanks to Bayes' theorem.

$$\text{Dataset } \mathbf{y} = (y_i)_{i=1}^k \in \mathcal{Y}^k \longrightarrow \text{likelihood } \ell_k(\mathbf{y}|\theta) \longrightarrow \text{posterior } p(\theta|\mathbf{y}) \propto \ell_k(\mathbf{y}|\theta)\pi(\theta)$$

$$\text{Parameter } \theta \in \Theta \longrightarrow \text{prior } \pi(\theta)$$

In practical studies, the influence of the prior distribution on the posterior distribution has to be assessed. We propose to use sensitivity analysis tools for achieving this task.

### Links between reference priors and sensitivity analysis

**What?** A revisit of the reference prior theory.

**Why?** To interpret the information issued from the prior, and to determine an objective one.

**How?** With a refinement of the mutual information definition with sensitivity indices.

**Definition:** Mutual information as a sensitivity index

$$I_D(\pi|k) = \mathbb{E} \left[ D(\mathbb{P}_{\mathbf{y}} || \mathbb{P}_{\mathbf{y}|\theta}) \right]$$

Reference prior theory principle: maximize  $I_D(\pi|k)$  w.r.t.  $\pi$ , i.e., maximize the impact of  $\theta$  onto the distribution of the data.

#### Our main result

We take  $D = D_f$  an  $f$ -divergence with  $f = \gamma x^\delta + o(x^\delta)$ ,  $\delta < 1$ .

**Theorem:** Sensitivity index reference prior (Van Biesbroeck 2023)

Consider  $\mathcal{P}$ : the class of priors admitting a continuous and positive p.d.f. on  $\Theta$ .

For any prior in  $\mathcal{P}$  with density  $\pi$ :  $\lim_{k \rightarrow \infty} k^{d\delta/2} I_{D_f}(\pi|k) = \gamma C_\delta \int_{\Theta} \pi(\theta)^{1+\delta} J(\theta)^{-\delta} d\theta$ .

Moreover, if  $\gamma(\delta + 1) > 0$ , then:  $\lim_{k \rightarrow \infty} k^{d\delta/2} (I_{D_f}(J|k) - I_{D_f}(\pi|k)) \geq 0$ ,

where  $J(\theta)$  is the Jeffreys prior p.d.f. The equality stands iff  $\pi = J$ .

► This theorem states that the Jeffreys prior is the optimal reference prior under our sensitivity index based mutual information.

### Proper reference priors under generalized moment constraints

**What?** A suggestion of a framework to derive priors simultaneously objective and proper.

**Why?** Usual reference priors often lead to improper priors, jeopardizing their implementation.

**How?** Relying on the sensitivity-index-reference-prior theory, and considering the decay rates of the Jeffreys prior.

**Theorem:** Proper reference prior (Van Biesbroeck 2024b)

Assume  $g: \Theta \rightarrow \mathbb{R}_+$  be such that

$$\int_{\Theta} J(\theta) g(\theta)^{1+1/\delta} d\theta < \infty, \quad \int_{\Theta} J(\theta) g(\theta)^{1/\delta} d\theta < \infty.$$

Then the proper prior with p.d.f.  $\pi^* \propto J \cdot g^{1/\delta}$  is the sensitivity-index-reference-prior over the ones such that their p.d.f.  $\pi$  verify  $\pi \cdot g \in L^1(\Theta)$ .

► This theorem states that the knowledge of the decay rates of  $J$  allows to build a proper reference prior.

### Sensitivity Indices for design of experiments

**What?** A robust design of experiments for input-output models:

$\mathbf{y} = (\mathbf{a}, \mathbf{z})$ , with  $\mathbf{z} = h(\mathbf{a}, \theta)$  for a stochastic function  $h$ .

**Why?** For models of this type, careful choice of input data would improve the estimation.

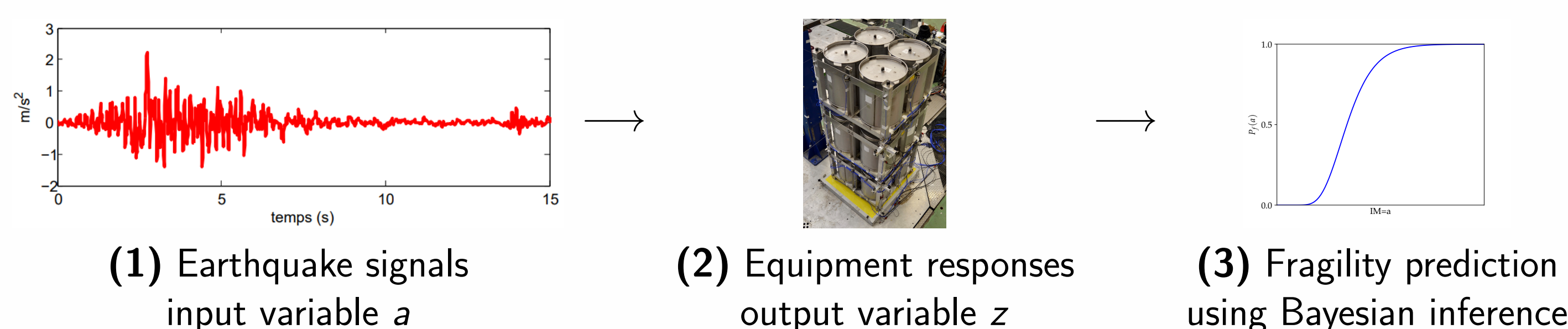
**How?** With sensitivity indices, we maximize the information brought by the data to the posterior.

We follow the idea of (Da Veiga 2015):

maximize w.r.t.  $\mathbf{a}_{k+1}$ :  $S(\mathbf{a}_{k+1}) = \mathbb{E}[D(\mathbb{P}_{\theta|\mathbf{a},\mathbf{z}} || \mathbb{P}_{\theta|\mathbf{a},\mathbf{z},\mathbf{a}_{k+1},\mathbf{z}_{k+1}})]$

## II/ Application to seismic fragility curves estimation

We seek to apply the above to the Bayesian estimation of seismic fragility curves, from a set of simulated or experimental data.



(1) The ground-motion is summarized in a scalar: the Intensity Measure (IM). This is the **input data**  $\mathbf{a}$ .

(2) The equipment fails or succeeds. The **output data** is **binary**:  $\mathbf{z} \in \{0, 1\}$ .

(3) The log-normal model expresses  $p(\mathbf{z}|\mathbf{a}, \theta) = \Phi\left(\frac{\log \mathbf{a} - \log \alpha}{\beta}\right)$ .

with  $\theta = (\alpha, \beta)$ , supposed to be a r.v. with distribution  $\pi$ : the prior.

The likelihood is given by  $\ell_k(\mathbf{z}, \mathbf{a}|\theta) = \prod_{i=1}^k p(\mathbf{z}_i|\mathbf{a}_i, \theta)p(\mathbf{a}_i)$ ,  $(\mathbf{z}, \mathbf{a}) = (\mathbf{z}_i, \mathbf{a}_i)_{i=1}^k$ .

### Problem 1: Choice of the prior

The data are limited due to the simulation cost of the studied equipment, therefore, the prior influences the estimations.

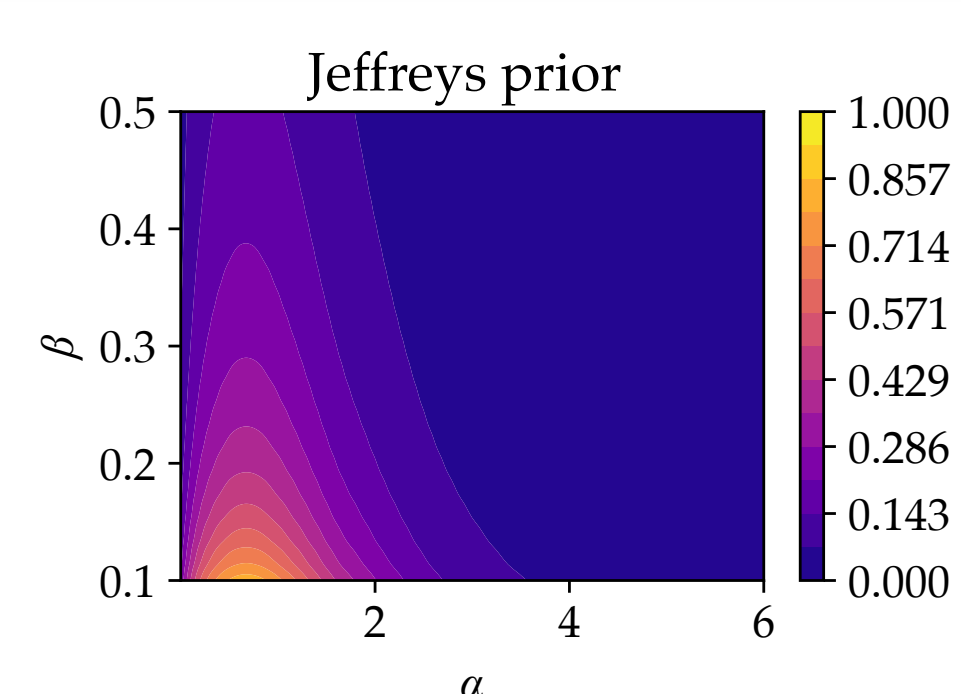
Also, nuclear safety requires a full auditability of the method, and so of all the choices involved.

The prior choice, thus, must be objective!

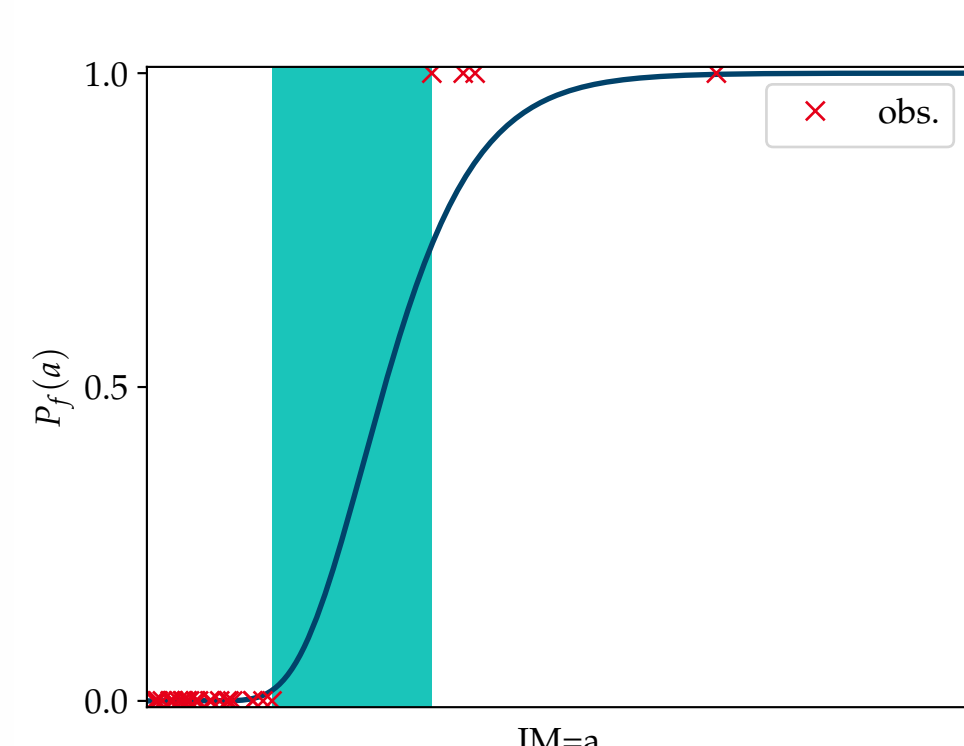
#### Our solution

► We computed the Jeffreys prior in the context of fragility curves (Van Biesbroeck et al. 2024a). We derive from its decay rates the proper reference prior based on sensitivity indices.

► The prior is approximated from its decay rates in the same fashion as in (Gu et al. 2018) for fast computation and fast generation of *a posteriori* estimates.



### Problem 2: Degenerate samples of the data



When the observed data are such that failures and non-failures are discriminated w.r.t. their IM, the **likelihood** decays slowly as a function of  $\beta$ , leading to improper posterior with Jeffreys prior. We say the data are “degenerate”.

$$\beta \rightarrow 0 \longrightarrow \beta \rightarrow \infty$$

$$\frac{1}{\beta} \rightarrow \infty \longrightarrow \frac{1}{\beta} \rightarrow 0$$

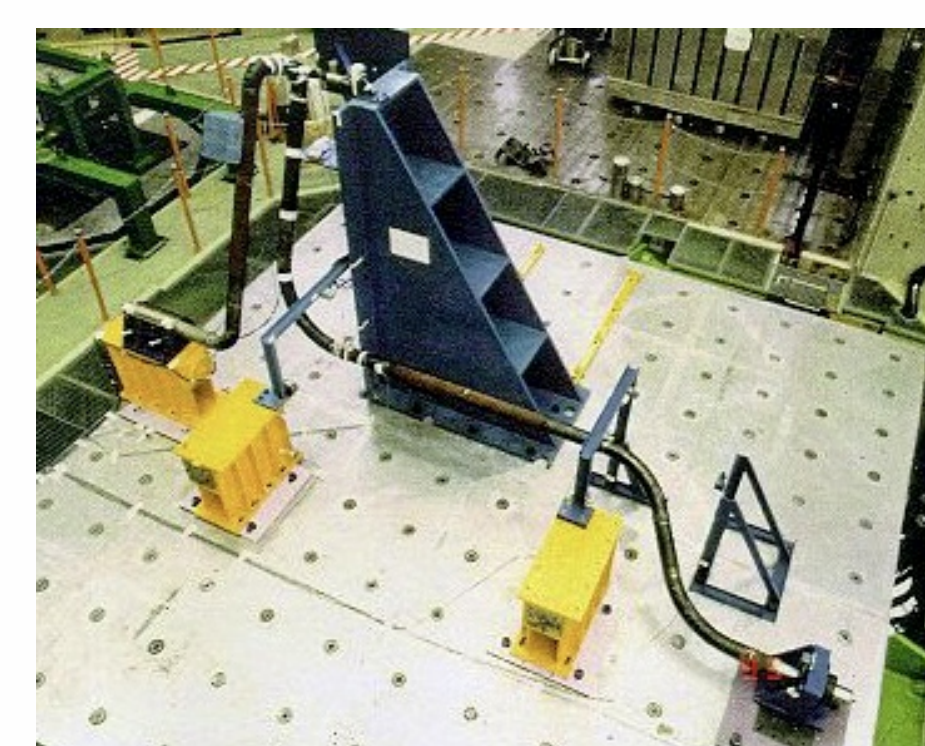
Above are the decay rates of the **likelihood** and of the Jeffreys prior w.r.t.  $\beta$  when the data are degenerate.

► This phenomenon happens more likely when the data is sparse, when the correlation between the IM and the structure's response is high, or when failures are rare.

#### Our solution

We carry out experimental design based on the maximization of sensitivity indices, in order to choose the right IM for robust estimation.

### Case study : equipment tested on the AZALEE shaking table

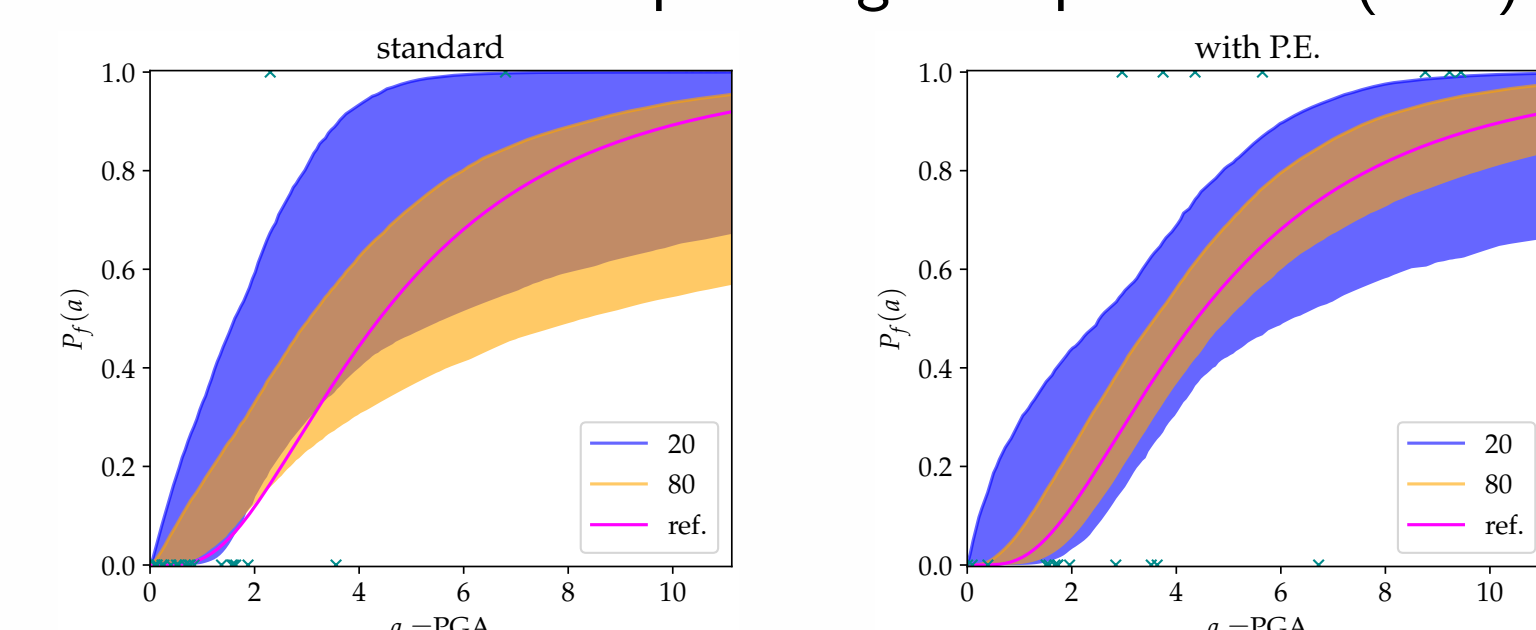


#### Piping system mock-up of a nuclear reactor

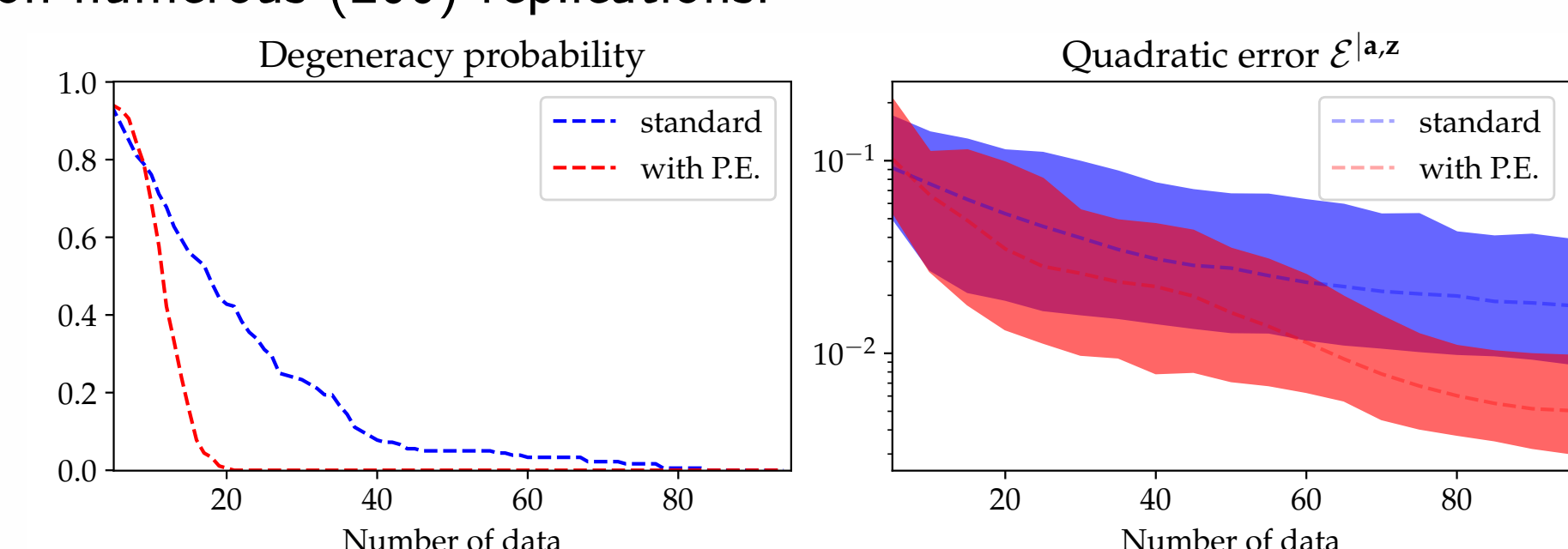
- The piping system was studied experimentally and numerically.
- The numerical model was validated by the experimental campaign.
- $10^4$  calculations were performed as a reference.
- Failure defined as excessive rotation of the first elbow.

### Results

► Example of estimations without and with planning of experiments (P.E.):



► Performances on numerous (200) replications:



The results demonstrate the robustness and the reliability of the method.

### References

- S. Da Veiga (2015). “Global sensitivity analysis with dependence measures”. In: *Journal of Statistical Computation and Simulation* 85.7, pp. 1283–1305.
- M. Gu, X. Wang and J. O. Berger (2018). “Robust Gaussian stochastic process emulation”. In: *The Annals of Statistics* 46.6A, pp. 3038–3066.
- A. Van Biesbroeck (2023). “Generalized mutual information and their reference prior under Csizar f-divergence”. In: *arXiv* 2310.10530.
- A. Van Biesbroeck, C. Gauchy, C. Feau, and J. Garnier (2024a). “Reference prior for Bayesian estimation of seismic fragility curves”. In: *Prob. Eng. Mech.* 76, pp. 103622
- A. Van Biesbroeck (2024b). “Properly constrained reference priors decay rates for efficient and robust posterior inference”. In: *arXiv* 2409.13041.

