Force reconstruction

A Bayesian perspective

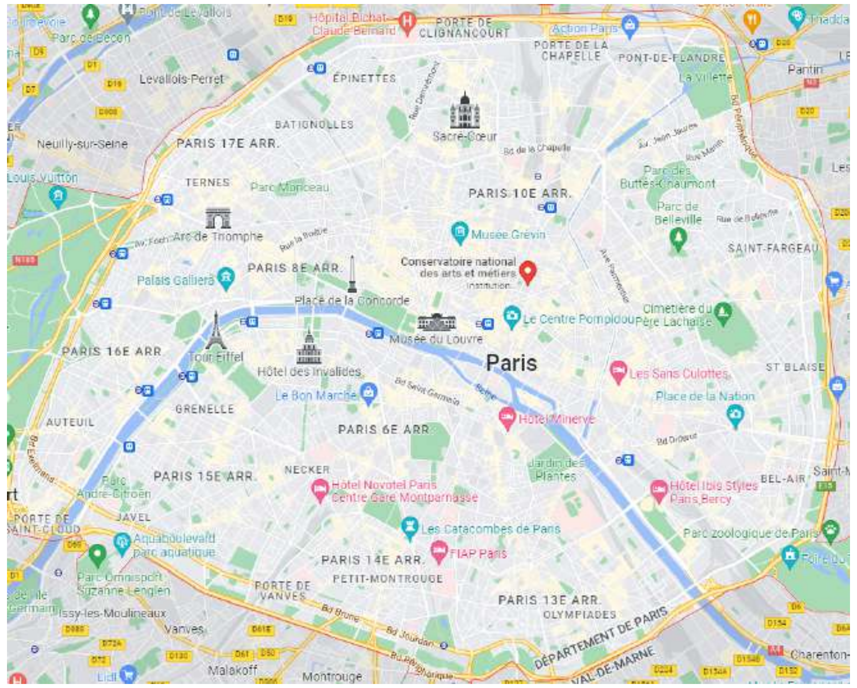
Mathieu AUCEJO

Thursday 13th October 2022

Who am I ?

It's me !

- Associate professor
- @ Le Cnam



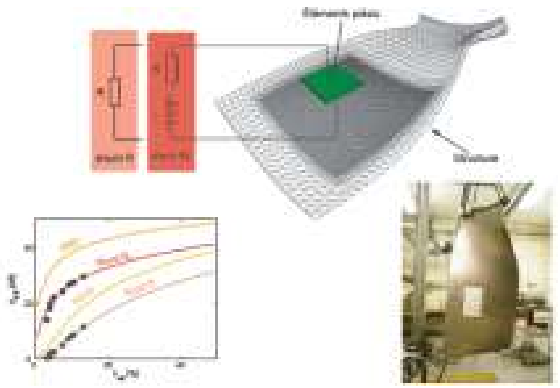
Who am I ?

It's me !

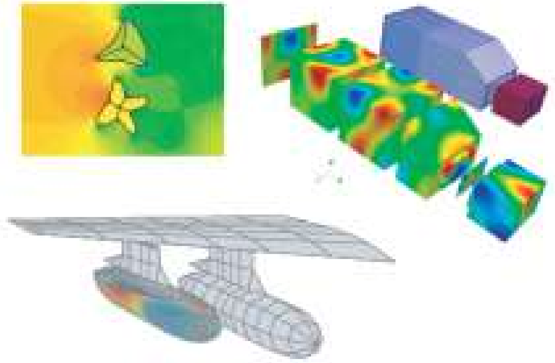
- Associate professor
- @ Le Cnam
- @ LMSSC

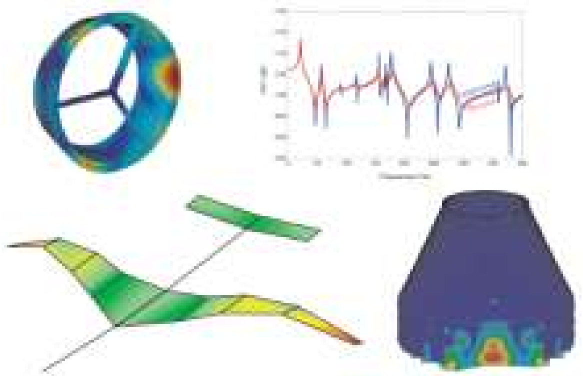
✉ mathieu.aucejo@lecnam.net

Smart structures
and adaptive interfaces

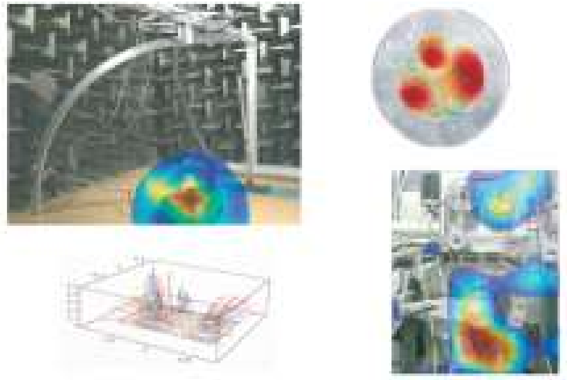


Fluid-structure interaction
and vibro-acoustics





Linear and nonlinear
structural dynamics



Source identification
and noise control

Outline

Outline

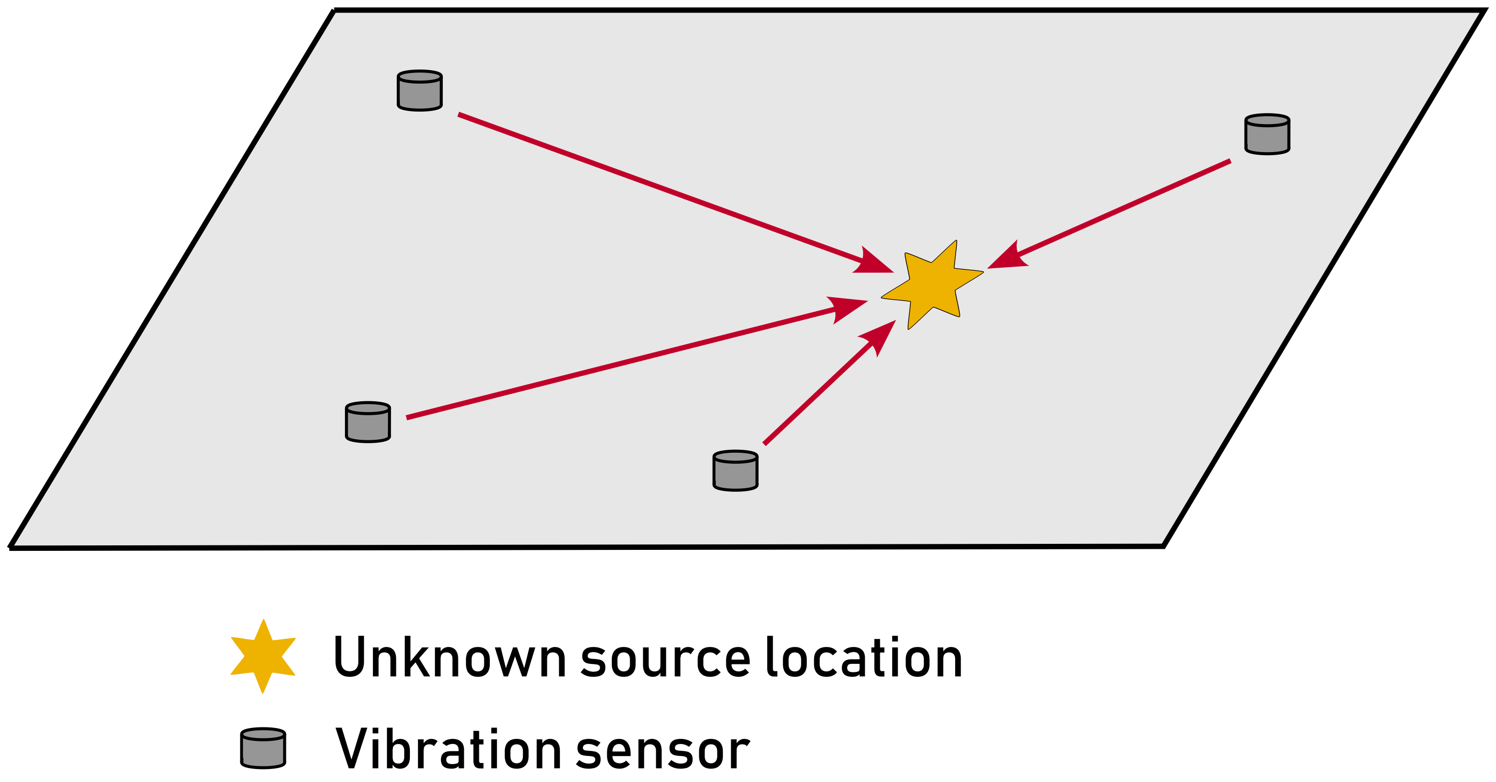
1. Generalities
2. State of the art
3. Bayesian Force regularization
4. Extensions

Definition

Force identification is an inverse problem aiming at characterizing some features of the sources exciting a mechanical structure

Types of problems

1. Localization

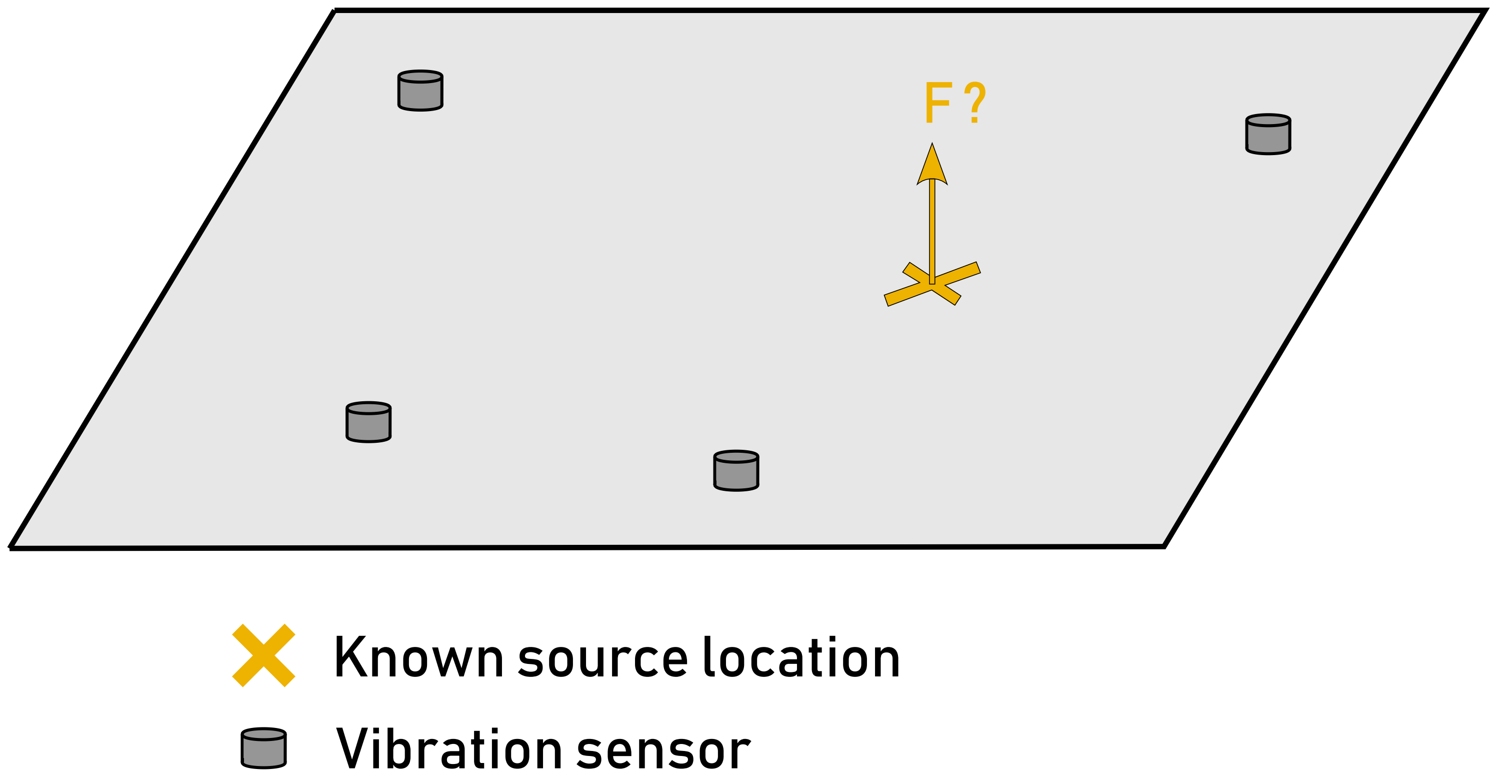


Definition

Force identification is an inverse problem aiming at characterizing some features of the sources exciting a mechanical structure

Types of problems

1. Localization
2. Quantification

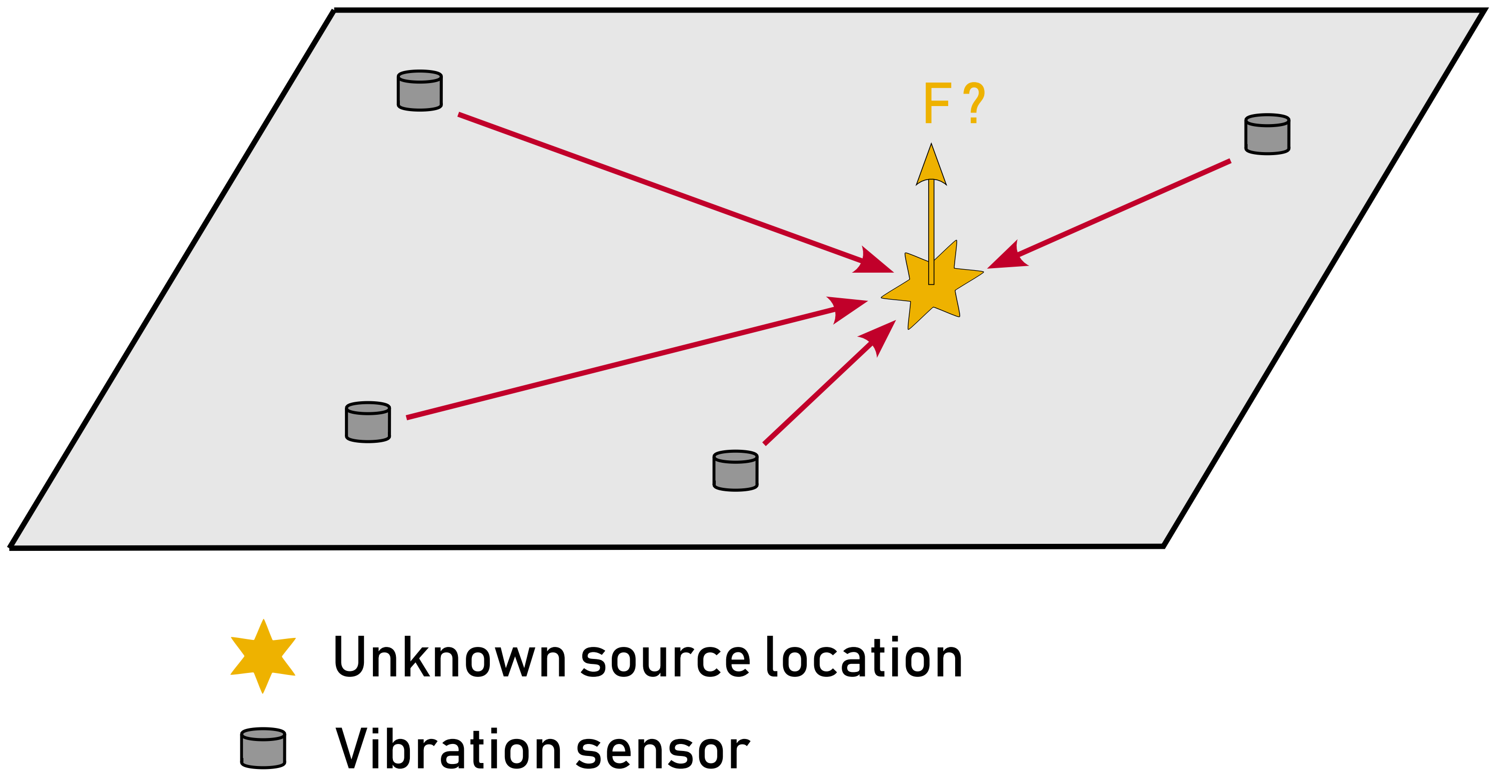


Definition

Force identification is an inverse problem aiming at characterizing some features of the sources exciting a mechanical structure

Types of problems

1. Localization
2. Quantification
3. Reconstruction

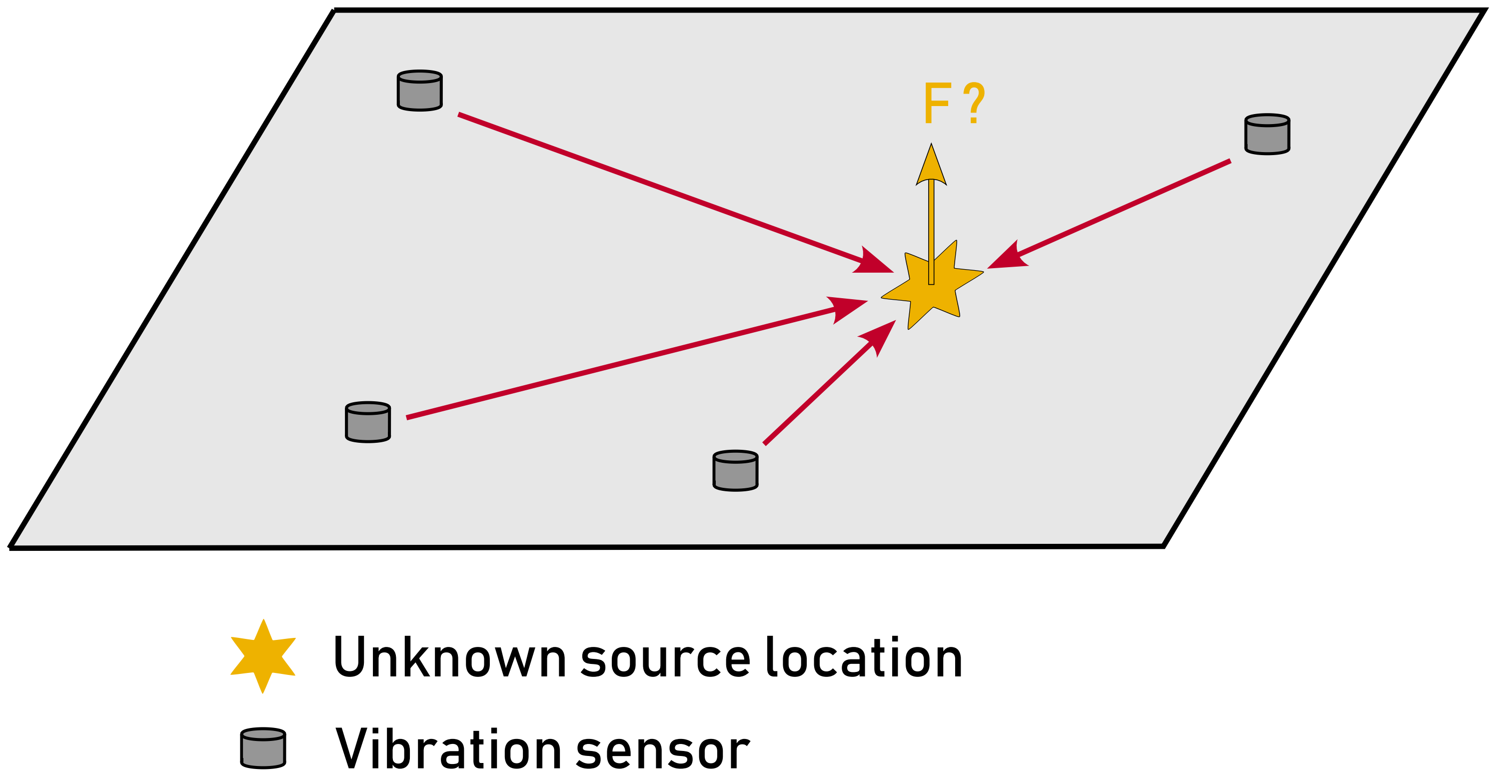


Definition

Force identification is an inverse problem aiming at characterizing some features of the sources exciting a mechanical structure

Types of problems

1. Localization
2. Quantification
3. Reconstruction
4. Separation / Classification

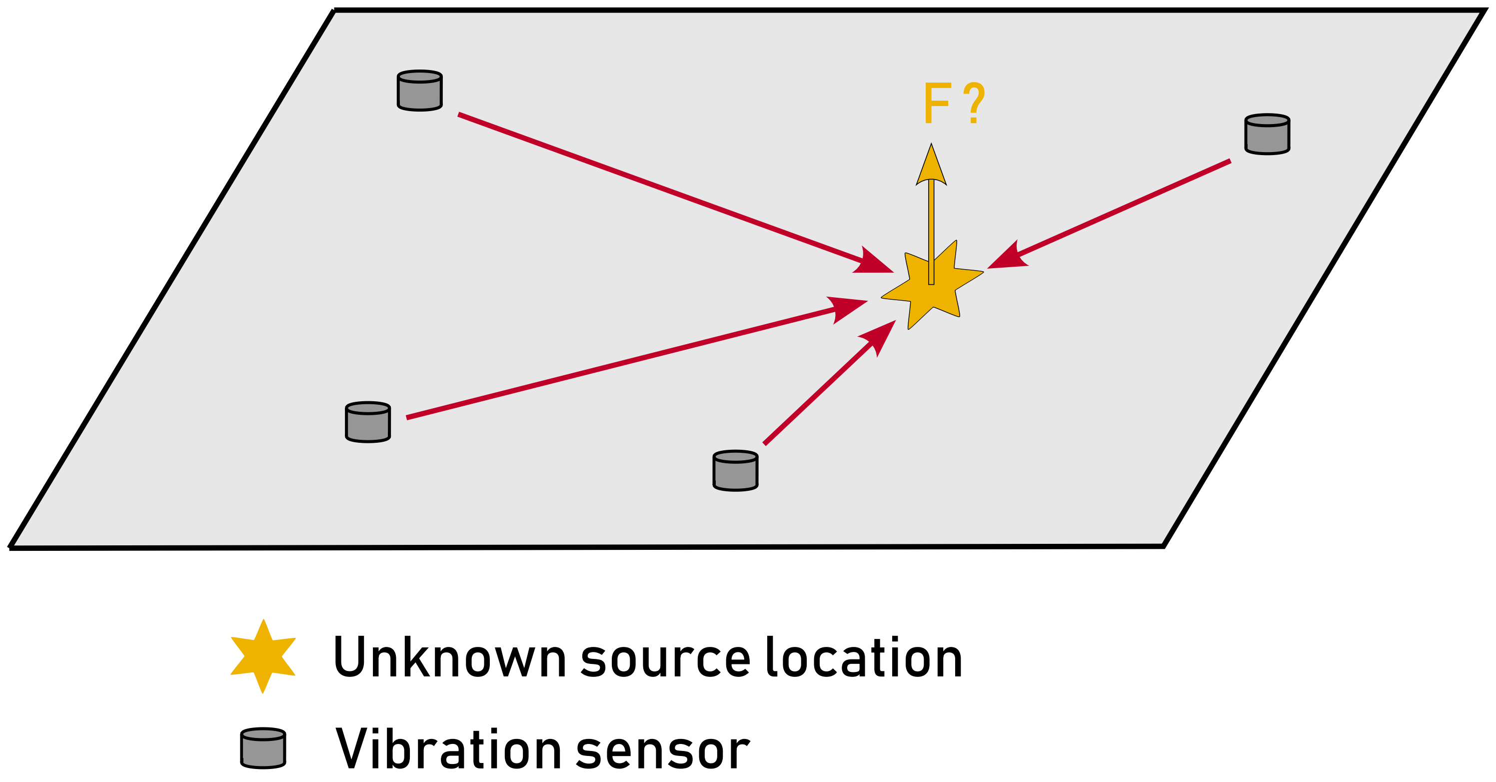


Definition

Force identification is an inverse problem aiming at characterizing some features of the sources exciting a mechanical structure

Types of problems

1. Localization
2. Quantification
3. **Reconstruction**
4. Separation / Classification



Restriction

In this lecture, we restrict ourselves to reconstruction problems expressed as a linear system

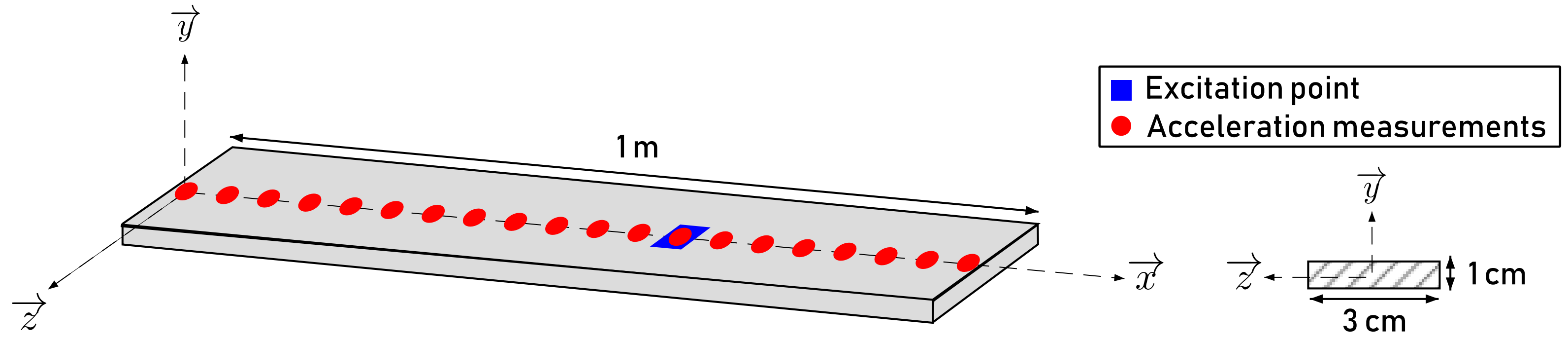
$$\mathbf{X} = \mathbf{H}\mathbf{F} + \mathbf{N}$$

- $\mathbf{X}$ is the measured vibration field
- $\mathbf{H}$ describes the dynamic behavior of the structure (LTI assumption)
- $\mathbf{F}$ is the excitation field to reconstruct
- $\mathbf{N}$ is the noise corrupting the vibration data

➥ This talk will not cover methods such as Kalman Filters, Neural Networks, Virtual Fields, ...

Outline

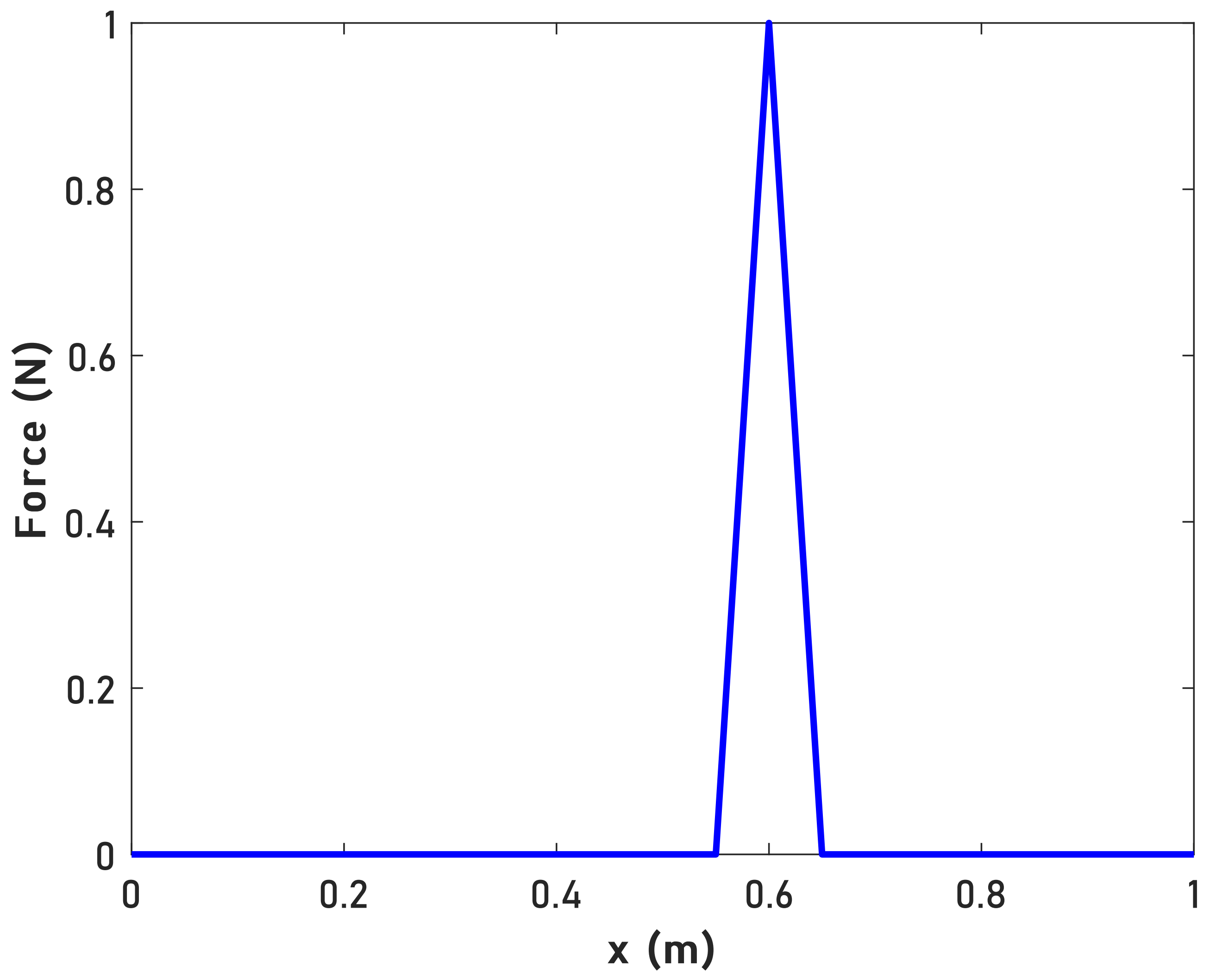
Leading example Free-free steel beam in the frequency domain



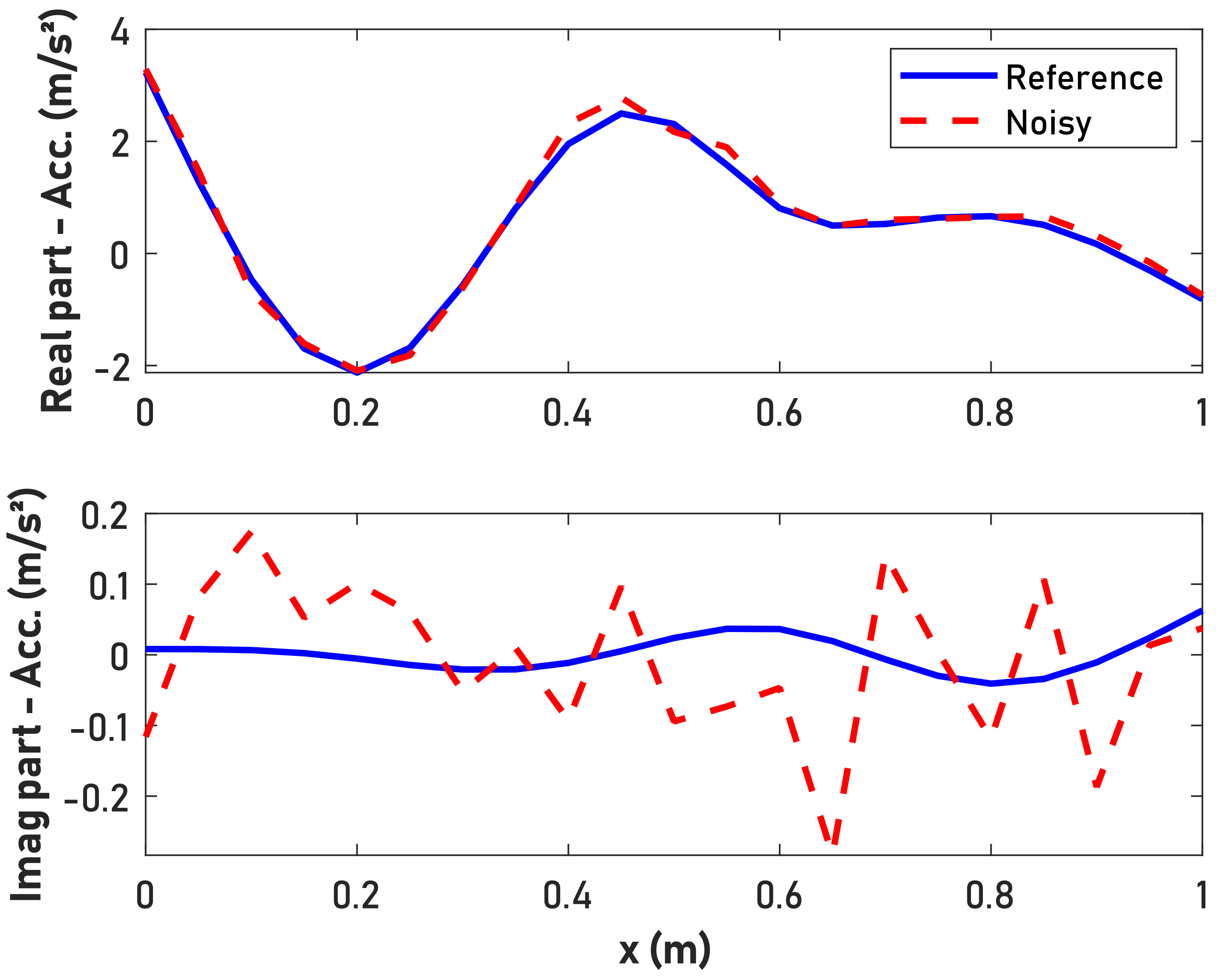
- Unit harmonic point force @ 350 Hz
- Measurement noise level - 20 dB
- Data generation - ⚠ Inverse crime
  - $\mathbf{X}$ - Modal expansion (8 modes, $f_8 \approx 992$ Hz)
  - $\mathbf{H}$ - FEM (20 beam elements)
  - Colocated reconstruction configuration
  - Equal-determined inverse problem

Main objective

Reconstruct

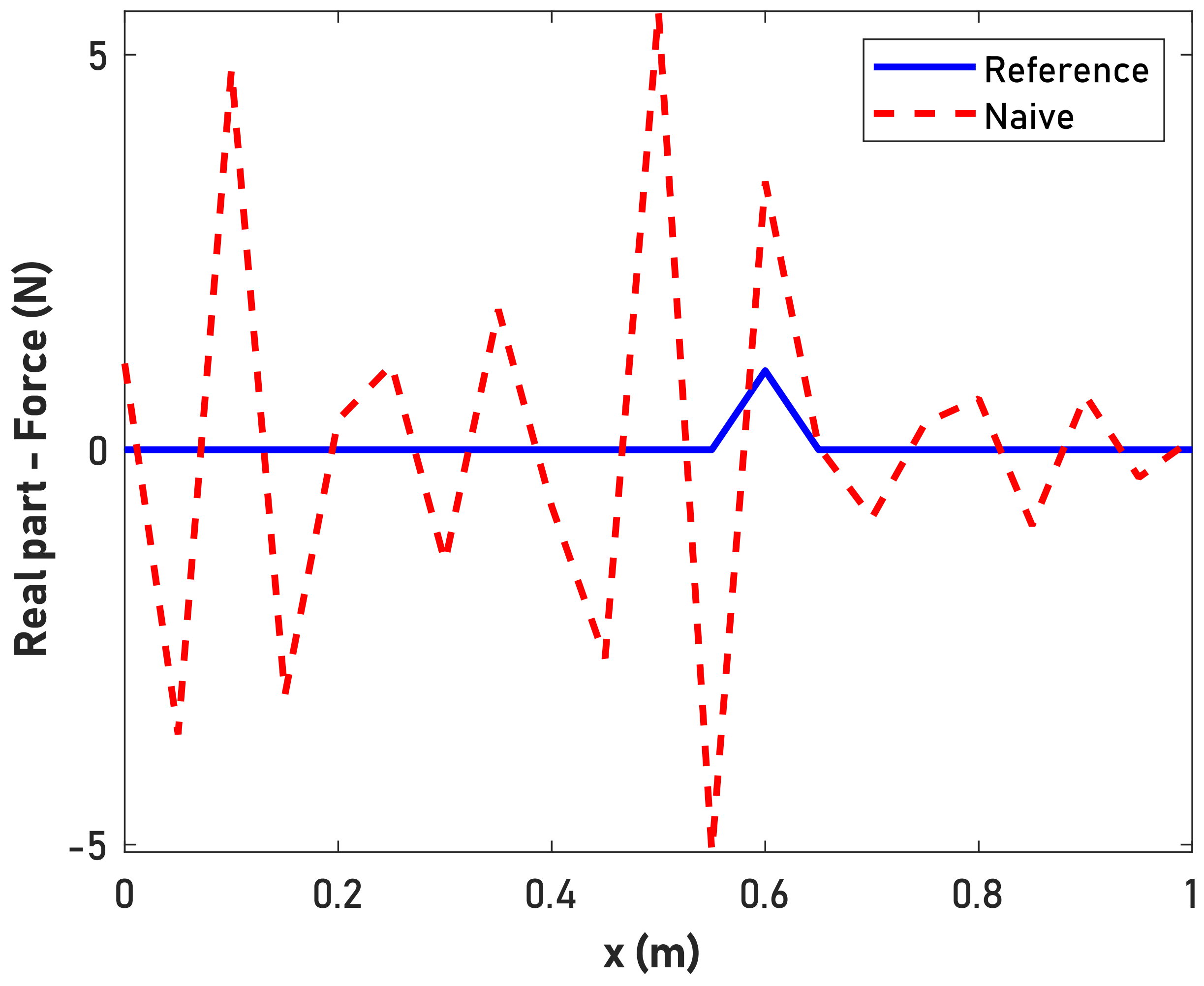


From



Naive reconstruction

$$\widehat{\mathbf{F}} = \mathbf{H}^{-1}\mathbf{X}$$



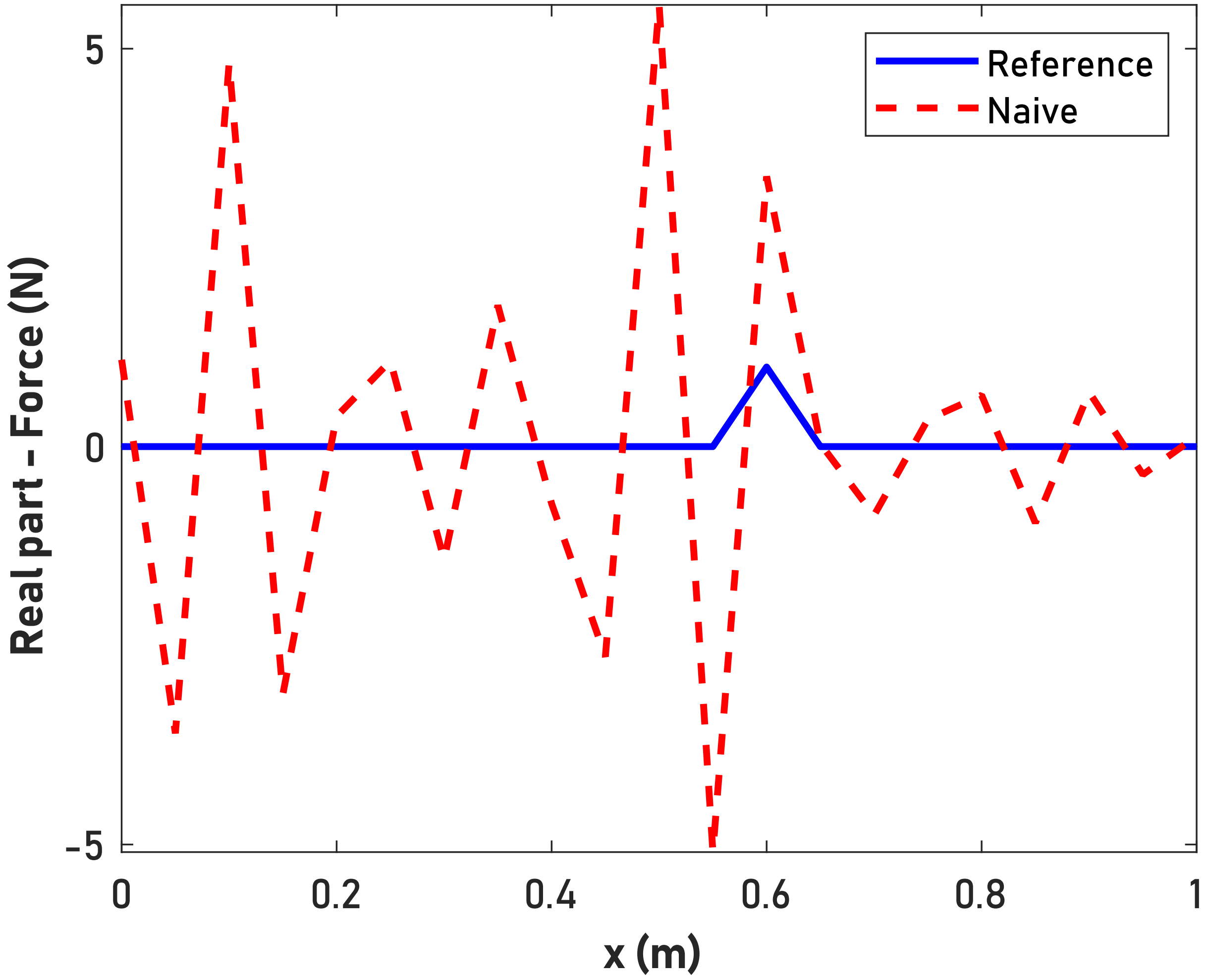
What's wrong ?

- Formally, one has:

$$\widehat{\mathbf{F}} = \sum_{i=1}^{21} \frac{\mathbf{v}_i \mathbf{u}_i^H \mathbf{X}}{\sigma_i}$$

Naive reconstruction

$$\widehat{\mathbf{F}} = \mathbf{H}^{-1}\mathbf{X}$$



What's wrong ?

- Formally, one has:

$$\widehat{\mathbf{F}} = \mathbf{F}_{\text{true}} + \sum_{i=1}^{21} \frac{\mathbf{v}_i \mathbf{u}_i^H \mathbf{N}}{\sigma_i}$$

- But $\mathbf{H}$ is ill-conditioned - $\kappa(\mathbf{H}) \approx 1300$

  Here $\sigma_{21} \approx 2.5 \cdot 10^{-2}$

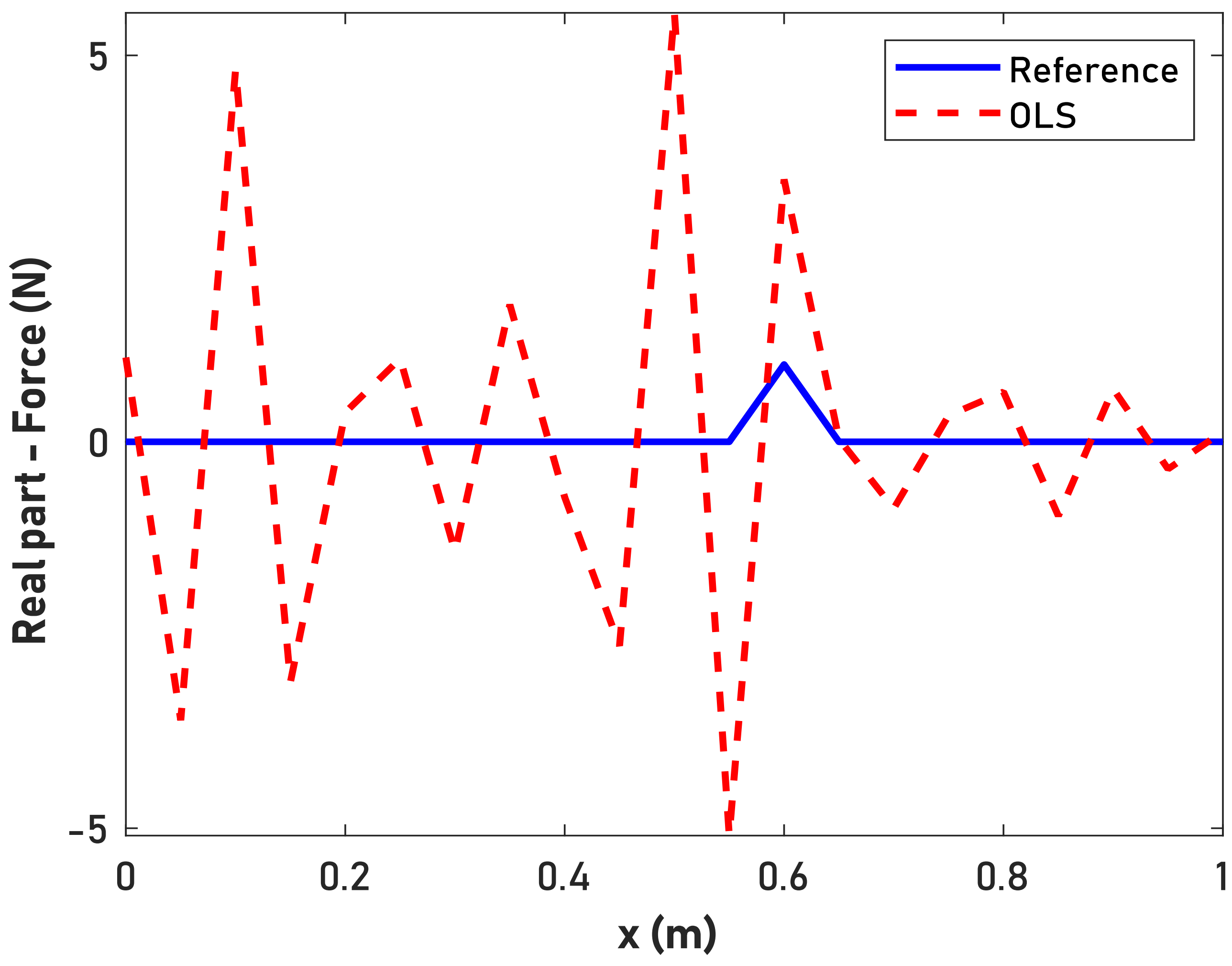
➥ The noise is amplified by the smallest singular values

➥ Ill-posed inverse problems in Hadamard sense

Ordinary Least Squares (OLS)

Idea **Find $\widehat{\mathbf{F}}$ minimizing the sum of the squared errors**

$$\widehat{\mathbf{F}} = \underset{\mathbf{F}}{\operatorname{argmin}} \ \|\mathbf{X} - \mathbf{H}\mathbf{F}\|_2^2$$



What's wrong ?

- Formally, one has:

$$\widehat{\mathbf{F}} = (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H\mathbf{X}$$

- But using the SVD

$$\widehat{\mathbf{F}} = \sum_{i=1}^{21} \frac{\mathbf{v}_i\mathbf{u}_i^H\mathbf{X}}{\sigma_i}$$

➡ Same as the naive approach ! (equal-det. problems)

➡ Useful for over/under-determined problems

Truncated SVD

Idea Filter the smallest singular values of $\mathbf{H}$

In practice Retain the first $\mathbf{M}$ singular values $(\mathbf{M} < \mathbf{21})$ such that

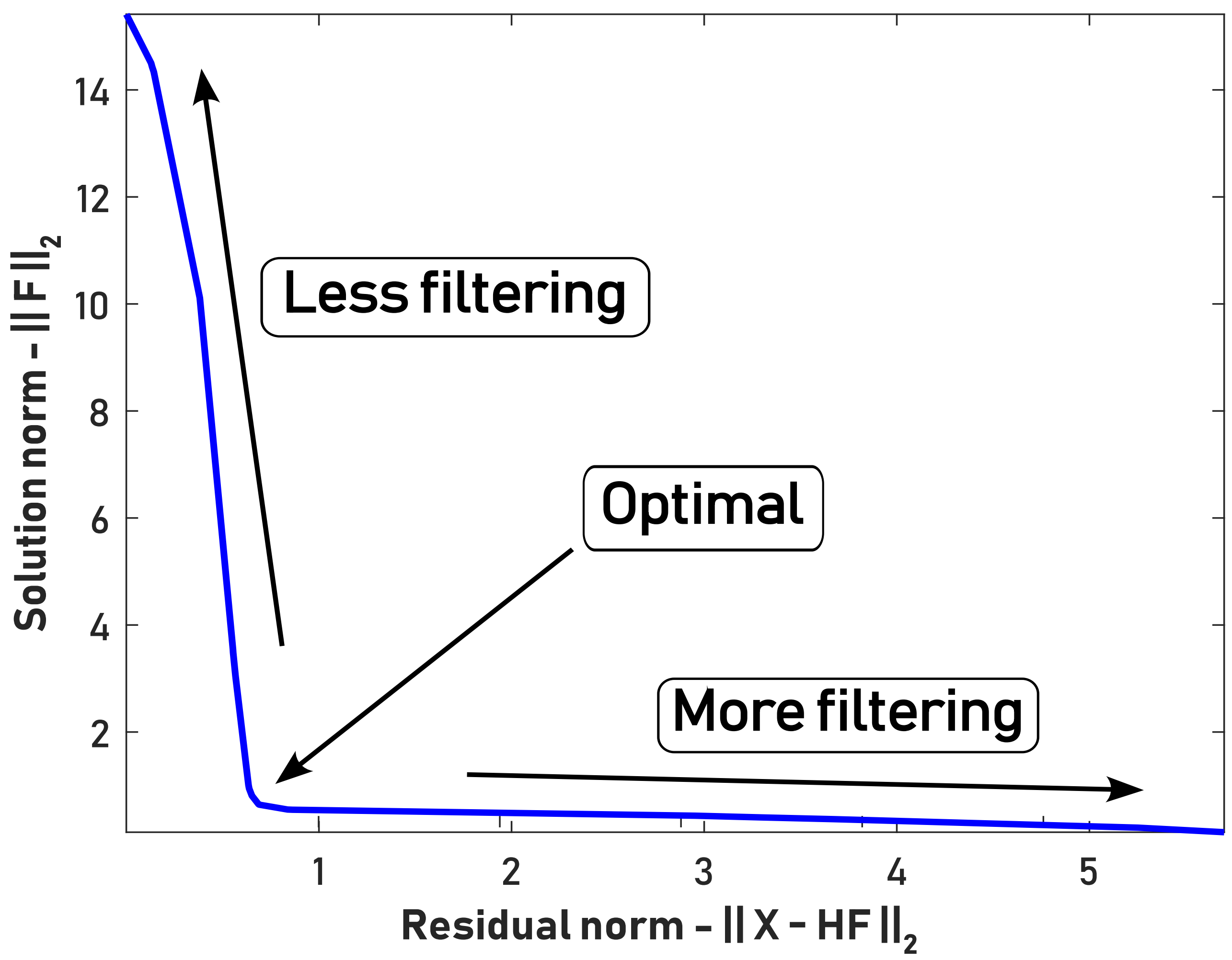
$$\widehat{\mathbf{F}} = \sum_{i=1}^{M} \frac{\mathbf{v}_i \mathbf{u}_i^H \mathbf{X}}{\sigma_i}$$

How to select M ?

One possible solution L-curve principle

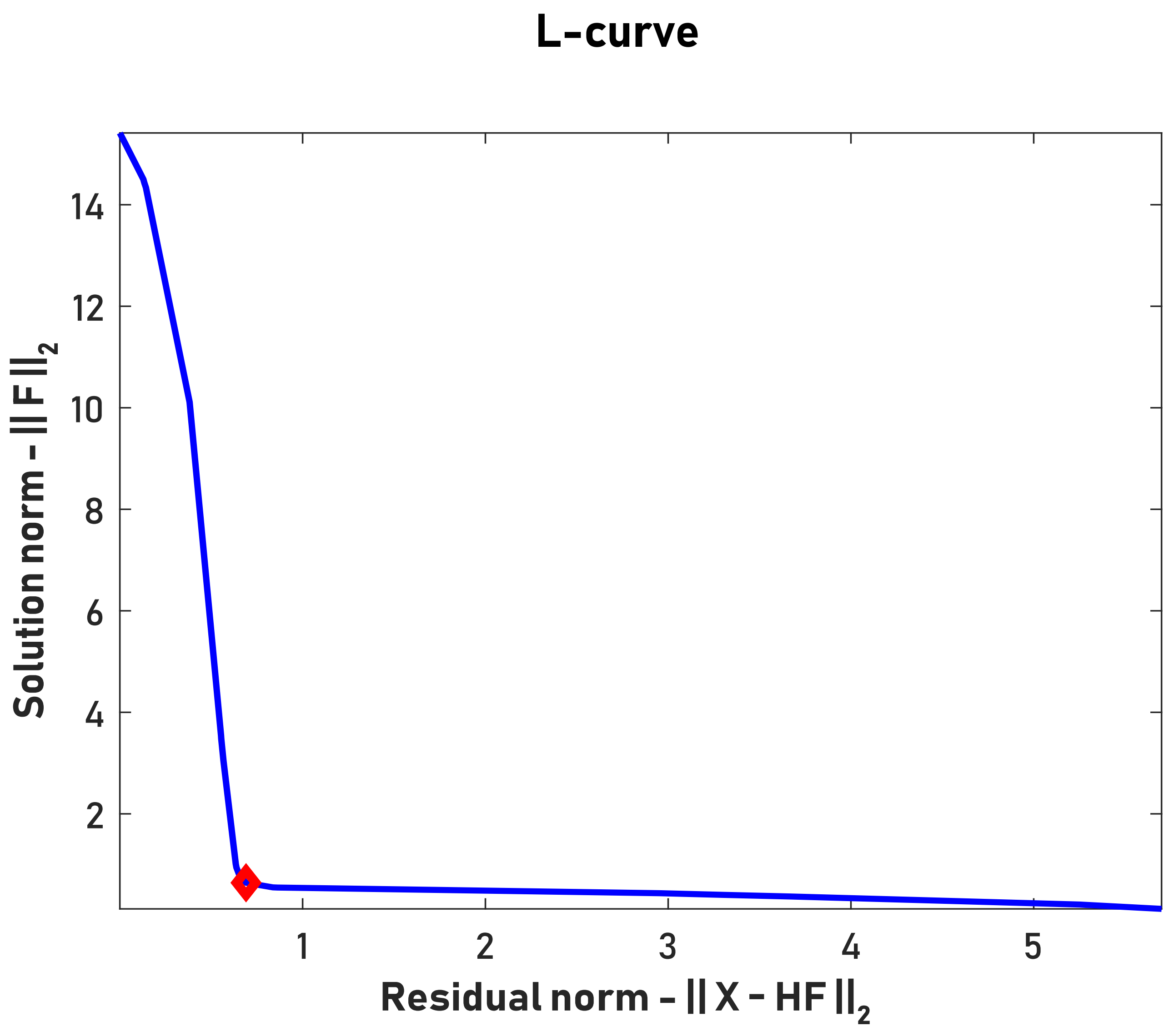
$$L_c(M) = \left(\left\| \mathbf{X} - \mathbf{H}(M)\widehat{\mathbf{F}} \right\|_2, \left\| \widehat{\mathbf{F}} \right\|_2 \right) \text{ with } \mathbf{H}(M) = \sum_{i=1}^{M} \sigma_i \mathbf{u}_i \mathbf{v}_i^H$$

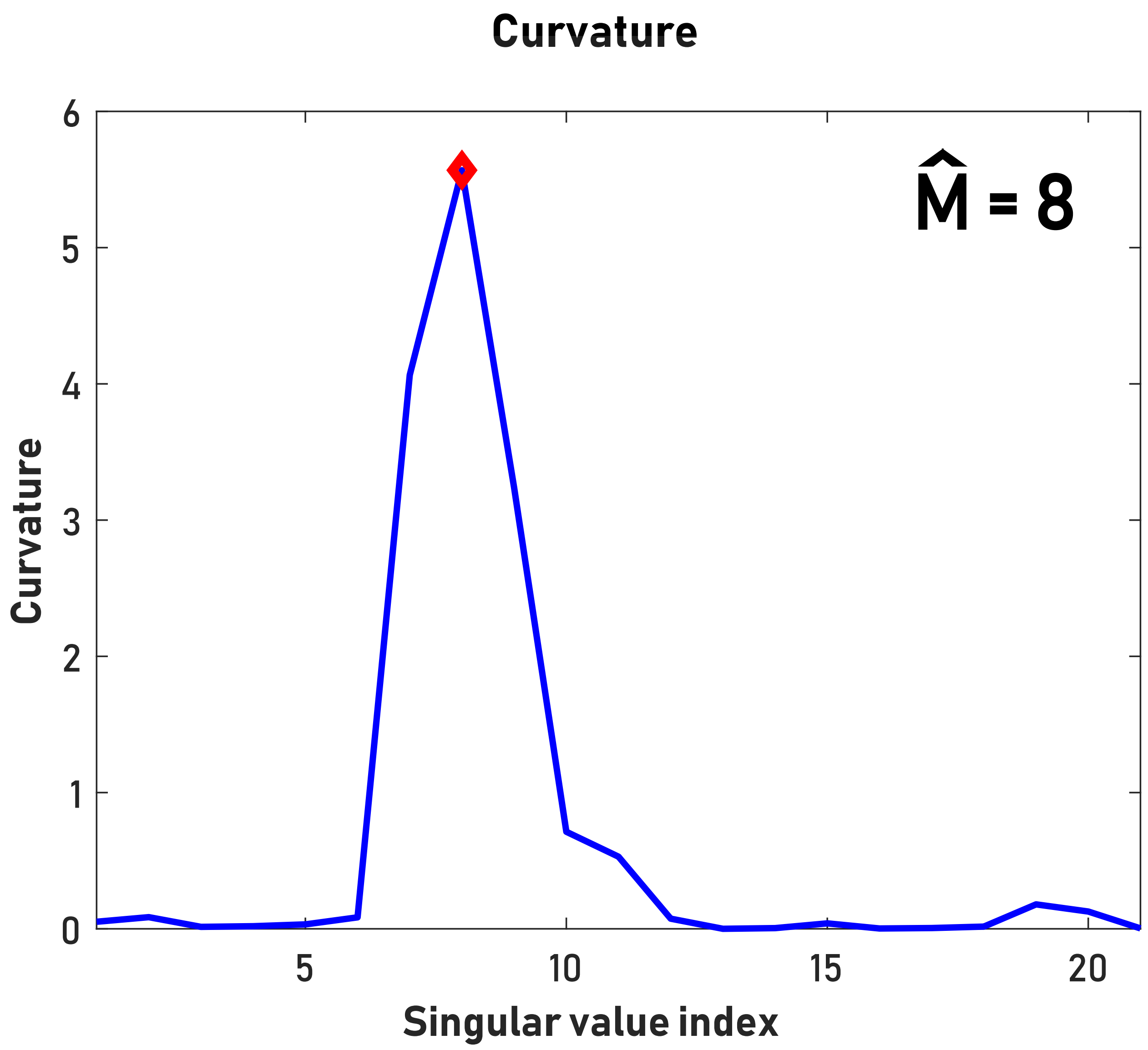
L-curve



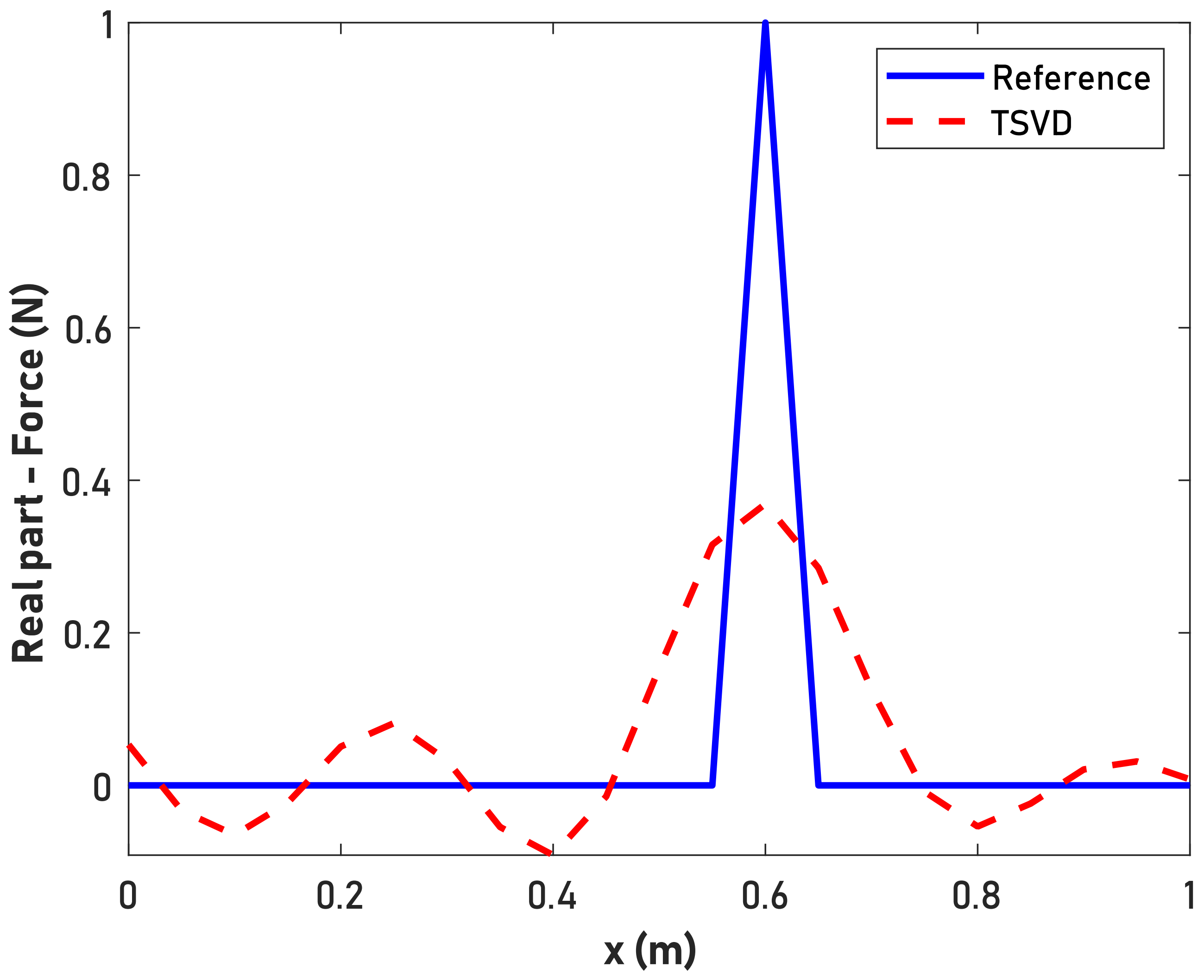
One possible solution L-curve principle

$$\widehat{M} = \underset{M}{\operatorname{argmax}} \; K[L(M)]$$





Application



- Low pass filtering effect ⇒ Smooth solution

➡ Not adapted to sparse sources

What to do ?

Constrain the space of admissible solutions !

ℓ_2-regularization Tikhonov regularization

$$\widehat{\mathbf{F}} = \underset{\mathbf{F}}{\operatorname{argmin}} \; \|\mathbf{X} - \mathbf{HF}\|_2^2 \; \text{ subject to } \; \|\mathbf{F}\|_2^2 \leq \tau$$

ℓ_2-regularization Tikhonov regularization

$$\widehat{\mathbf{F}} = \underset{\mathbf{F}}{\operatorname{argmin}} \ \|\mathbf{X} - \mathbf{HF}\|_2^2 + \lambda \|\mathbf{F}\|_2^2$$

How to select λ ?

In practice Many methods are available

- Morozov's discrepancy principle
- Generalized Cross Validation (GCV)
- Reginska's method
- Bayesian Estimator
- L-curve principle
- ....

ℓ_2-regularization Tikhonov regularization

$$\widehat{\mathbf{F}} = \underset{\mathbf{F}}{\operatorname{argmin}} \; \|\mathbf{X} - \mathbf{HF}\|_2^2 + \lambda \|\mathbf{F}\|_2^2$$

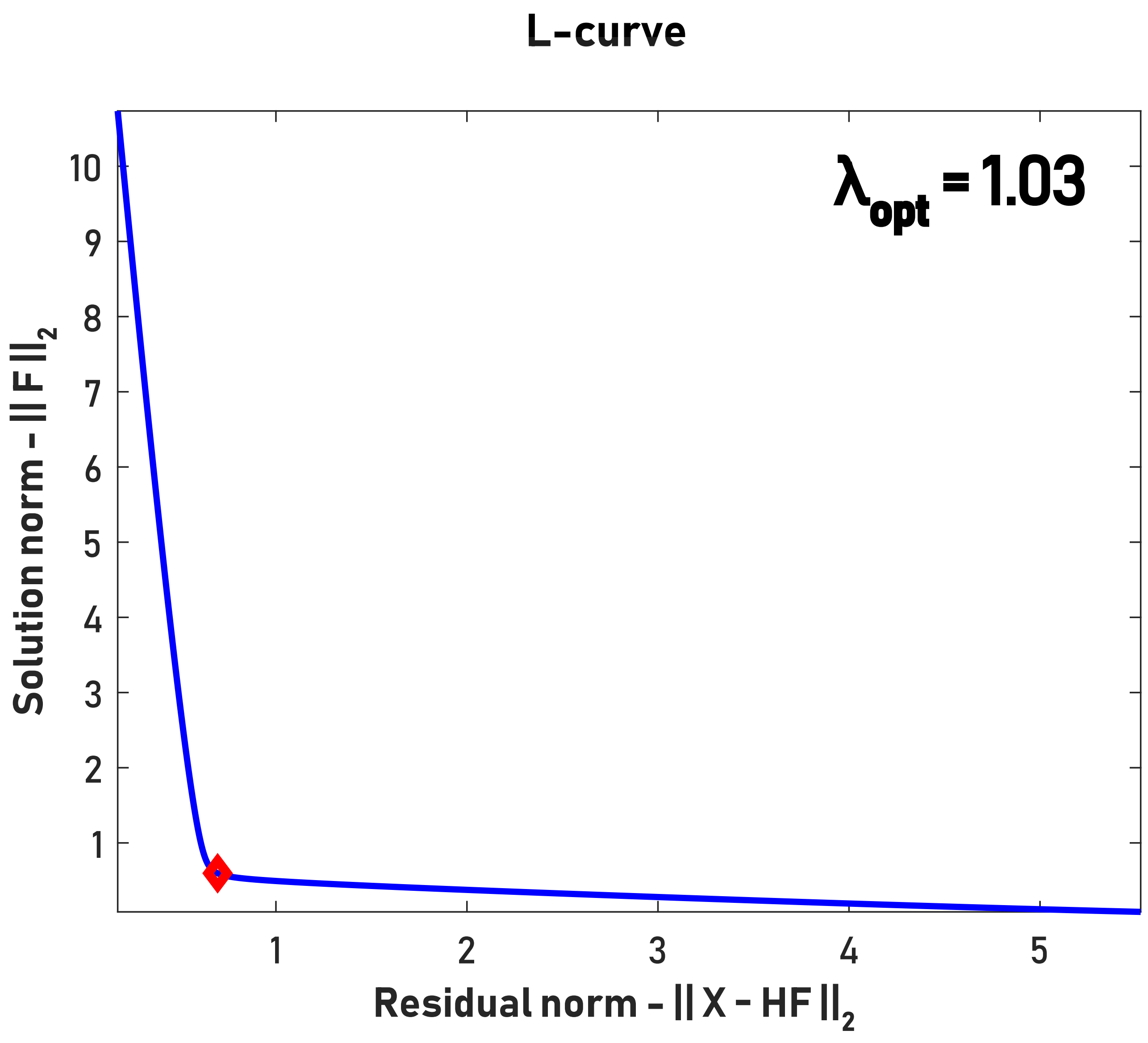
How to select λ ?

In practice Many methods are available

- Morozov's discrepancy principle
- Generalized Cross Validation (GCV)
- Reginska's method
- Bayesian Estimator
- **L-curve principle**
- ....

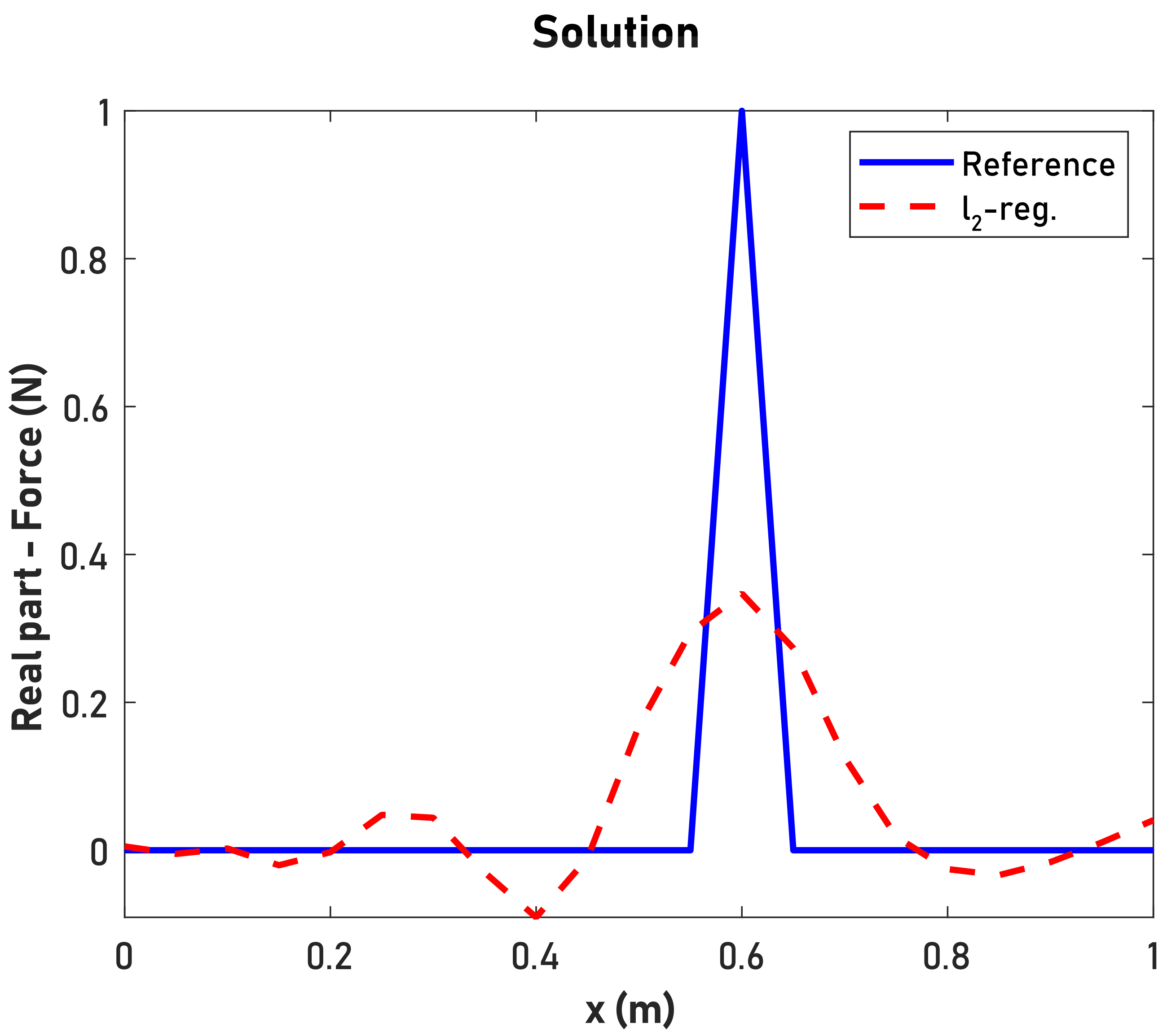
ℓ_2-regularization Application

$$\widehat{\mathbf{F}} = (\mathbf{H}^H\mathbf{H} + \lambda\mathbf{I})^{-1}\mathbf{H}^H\mathbf{X}$$



ℓ_2-regularization Application

$$\widehat{\mathbf{F}} = (\mathbf{H}^H\mathbf{H} + \lambda\mathbf{I})^{-1}\mathbf{H}^H\mathbf{X}$$



- Low pass filtering effect ⇒ Smooth solution

➡ Not adapted to sparse sources

How to explain this result ?

Filter factors Basics

$$\widehat{\mathbf{F}} = \sum_{i=1}^{21} f_i \frac{\mathbf{v}_i \mathbf{u}_i^H \mathbf{X}}{\sigma_i}$$

where f_i is the filter factor defined such that

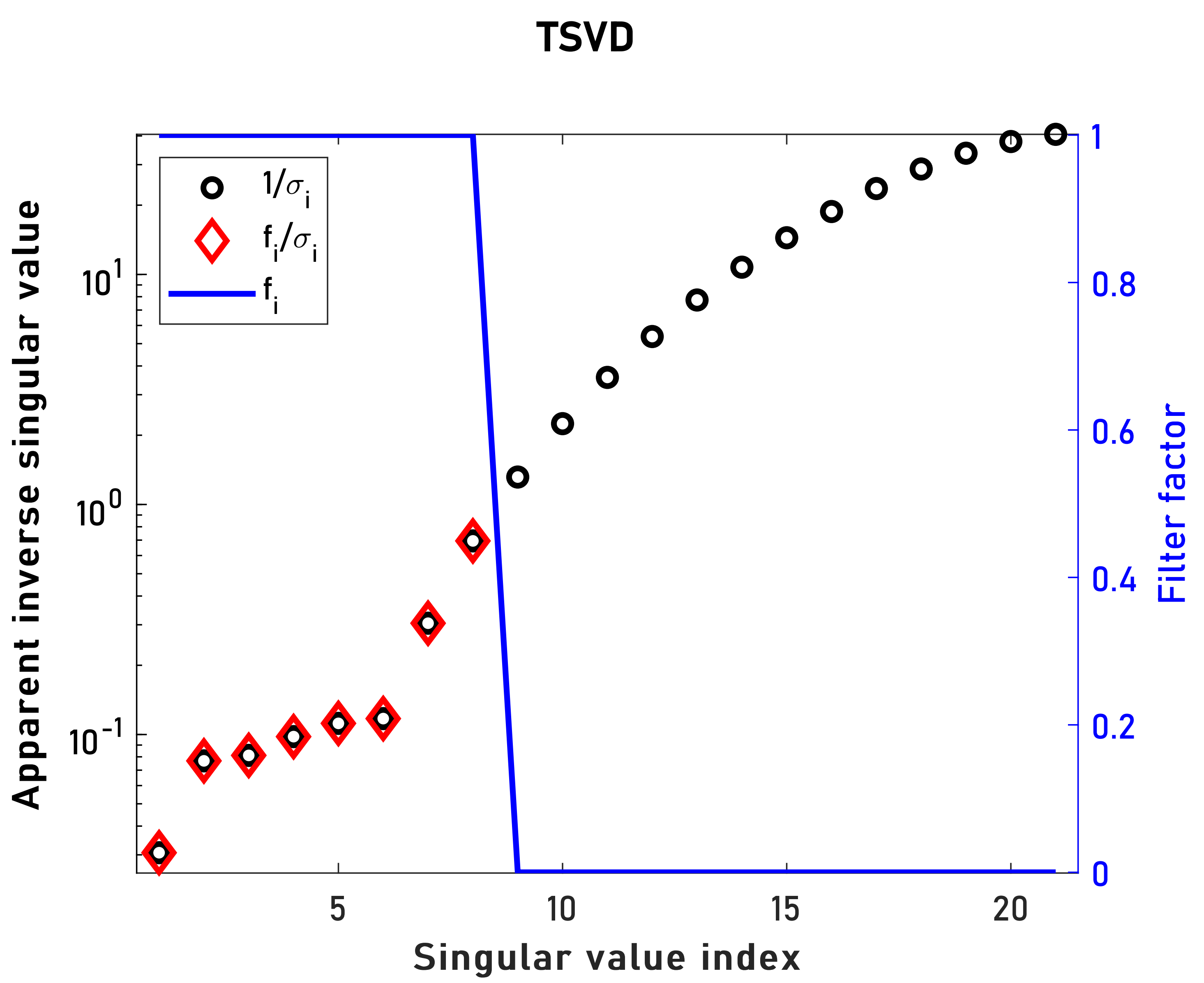
TSVD

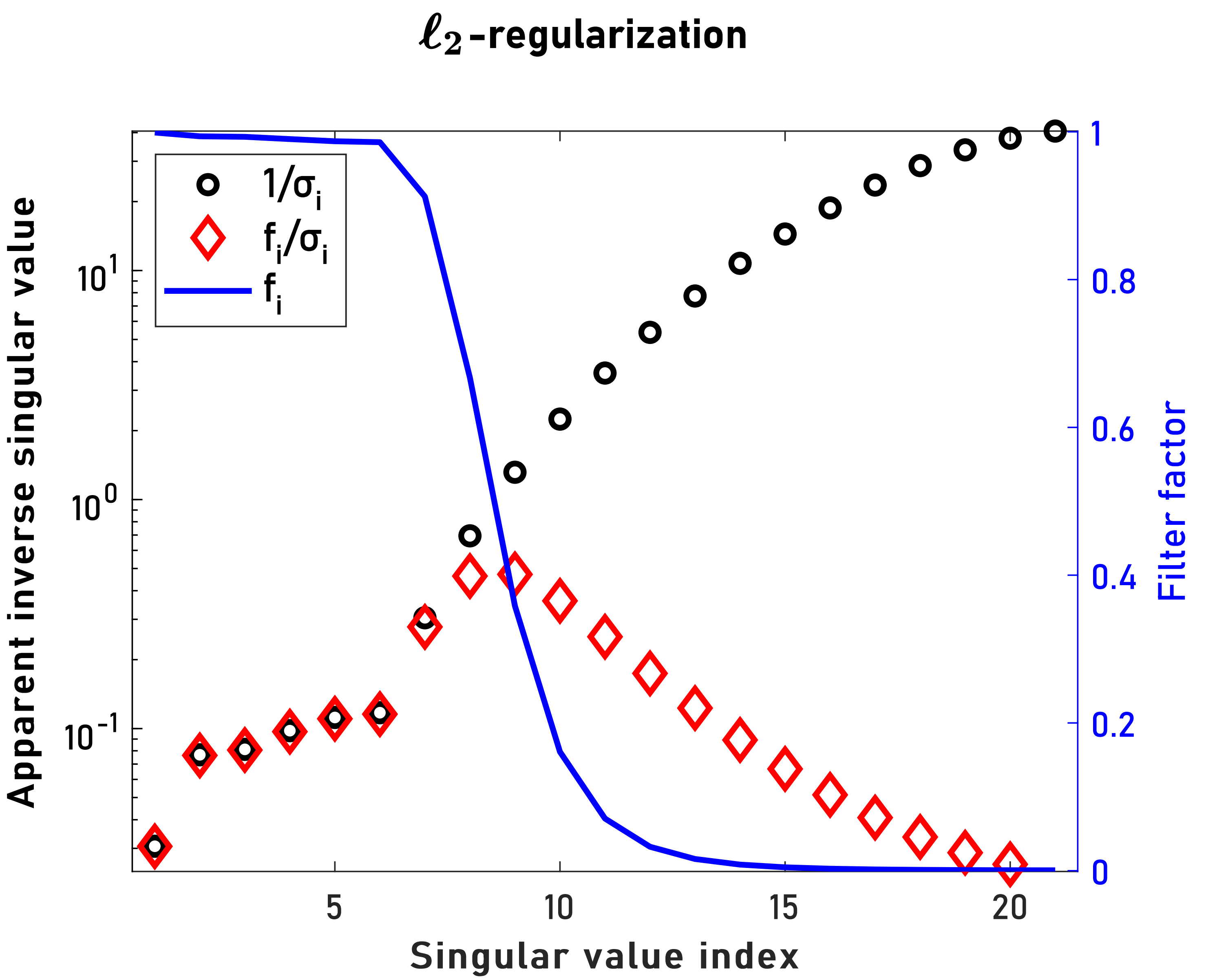
$$f_i = \begin{cases} 1 & \text{for } i \leq M \\ 0 & \text{otherwise} \end{cases}$$

ℓ_2-regularization

$$f_i = \frac{\sigma_i^2}{\sigma_i^2 + \lambda}$$

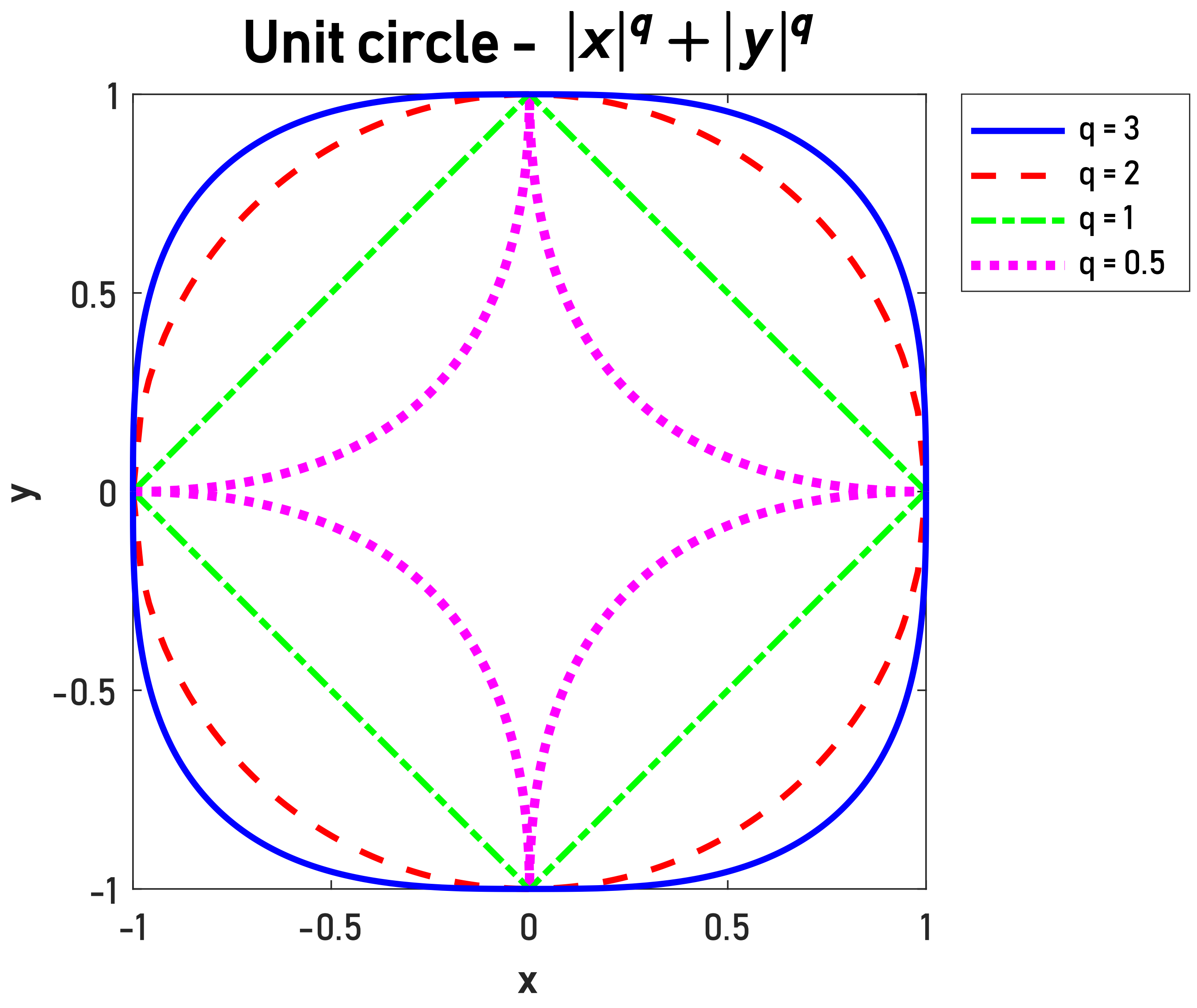
Filter factors In action





ℓ_q-regularization Generalities

$$\widehat{\mathbf{F}} = \underset{\mathbf{F}}{\operatorname{argmin}} \ \|\mathbf{X} - \mathbf{HF}\|_2^2 + \lambda \|\mathbf{F}\|_q^q$$



- The smaller q is, the larger is the weight on small values of $\mathbf{F}$
- For large values of $\mathbf{F}$, the smaller q is, the smaller is the weight on these values

➡ $q \geq 2$ - Smooth solution

➡ $q \leq 1$ - Sparse solution

⚠ Non-convex minimization problem when $q < 1$

ℓ_q-regularization Numerical resolution

The first-order optimality condition for the ℓ_q-regularization leads to

$$\widehat{\mathbf{F}} = \left(\mathbf{H}^H\mathbf{H} + \lambda\mathbf{W}(\widehat{\mathbf{F}})\right)^{-1}\mathbf{H}^H\mathbf{X} \quad \text{with} \quad w_{ii} = \frac{q}{2}\left|\widehat{F}_i\right|^{q-2}$$

➡ Implementation of an iterative process

$$\widehat{\mathbf{F}}^{(k)} = \left(\mathbf{H}^H\mathbf{H} + \lambda^{(k)}\mathbf{W}\left(\widehat{\mathbf{F}}^{(k-1)}\right)\right)^{-1}\mathbf{H}^H\mathbf{X}$$

ℓ_q-regularization Numerical resolution

The first-order optimality condition for the ℓ_q-regularization leads to

$$\widehat{\mathbf{F}} = \left(\mathbf{H}^H\mathbf{H} + \lambda\mathbf{W}(\widehat{\mathbf{F}})\right)^{-1}\mathbf{H}^H\mathbf{X} \quad \text{with} \quad w_{ii} = \frac{q}{2}\left|\widehat{F}_i\right|^{q-2}$$

➡ Implementation of an iterative process

$$\widehat{\mathbf{F}}^{(k)} = \underset{\mathbf{F}}{\operatorname{argmin}} \|\mathbf{X} - \mathbf{H}\mathbf{F}\|_2^2 + \lambda^{(k)}\|\mathbf{L}\mathbf{F}\|_2^2 \quad \text{with} \quad \mathbf{W}\left(\widehat{\mathbf{F}}^{(k-1)}\right) = \mathbf{L}^H\mathbf{L}$$

where $\lambda^{(k)}$ is selected from the following L-curve

$$L_c(\lambda^{(k)}) = \left(\|\mathbf{X} - \mathbf{H}\mathbf{F}(\lambda^{(k)})\|_2, \|\mathbf{L}\mathbf{F}(\lambda^{(k)})\|_2\right)$$

When the iterative process has converged, one has

$$\|\mathbf{L}\widehat{\mathbf{F}}\|_2^2 \approx \|\widehat{\mathbf{F}}\|_q^q$$

ℓ_q-regularization Practical implementation

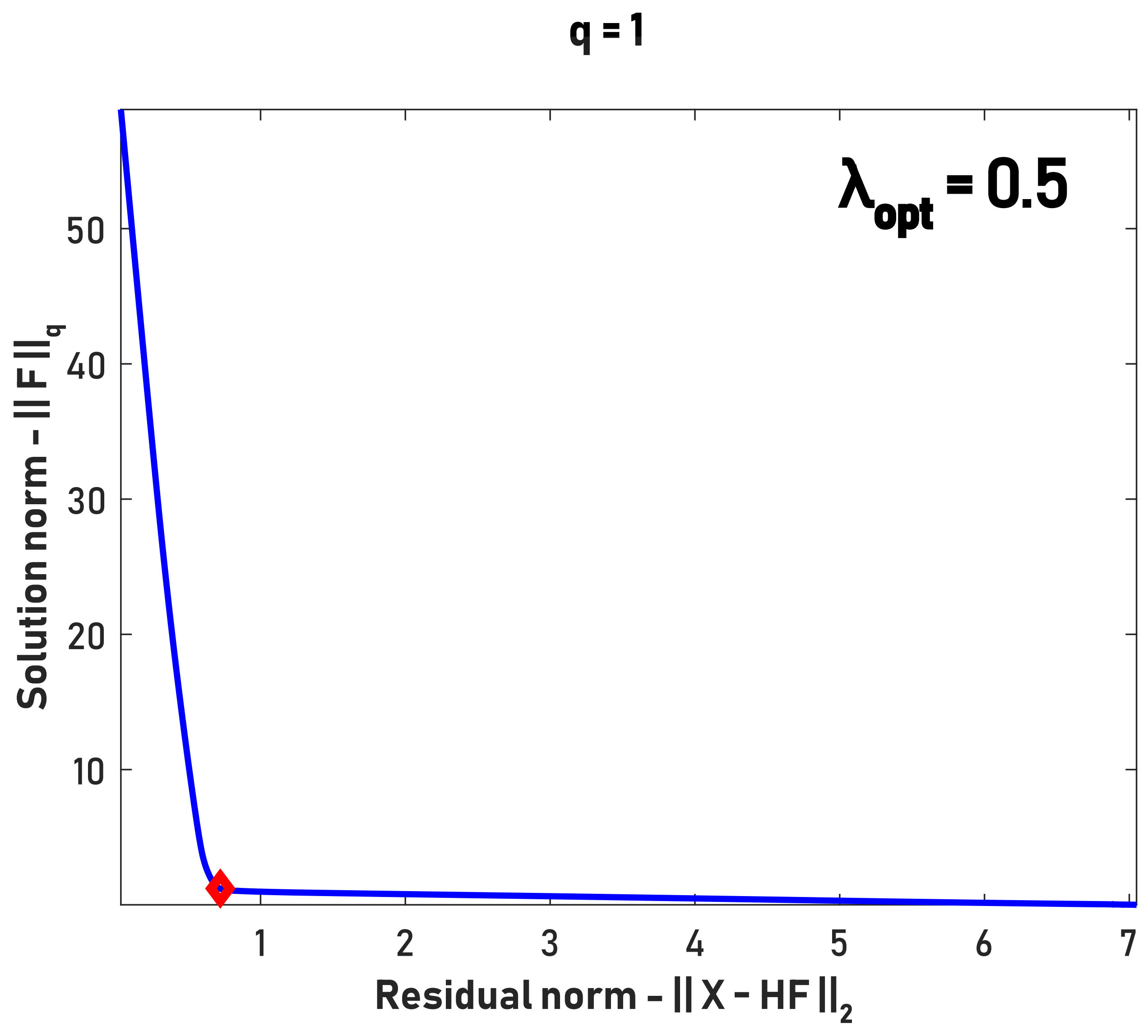
Matlab

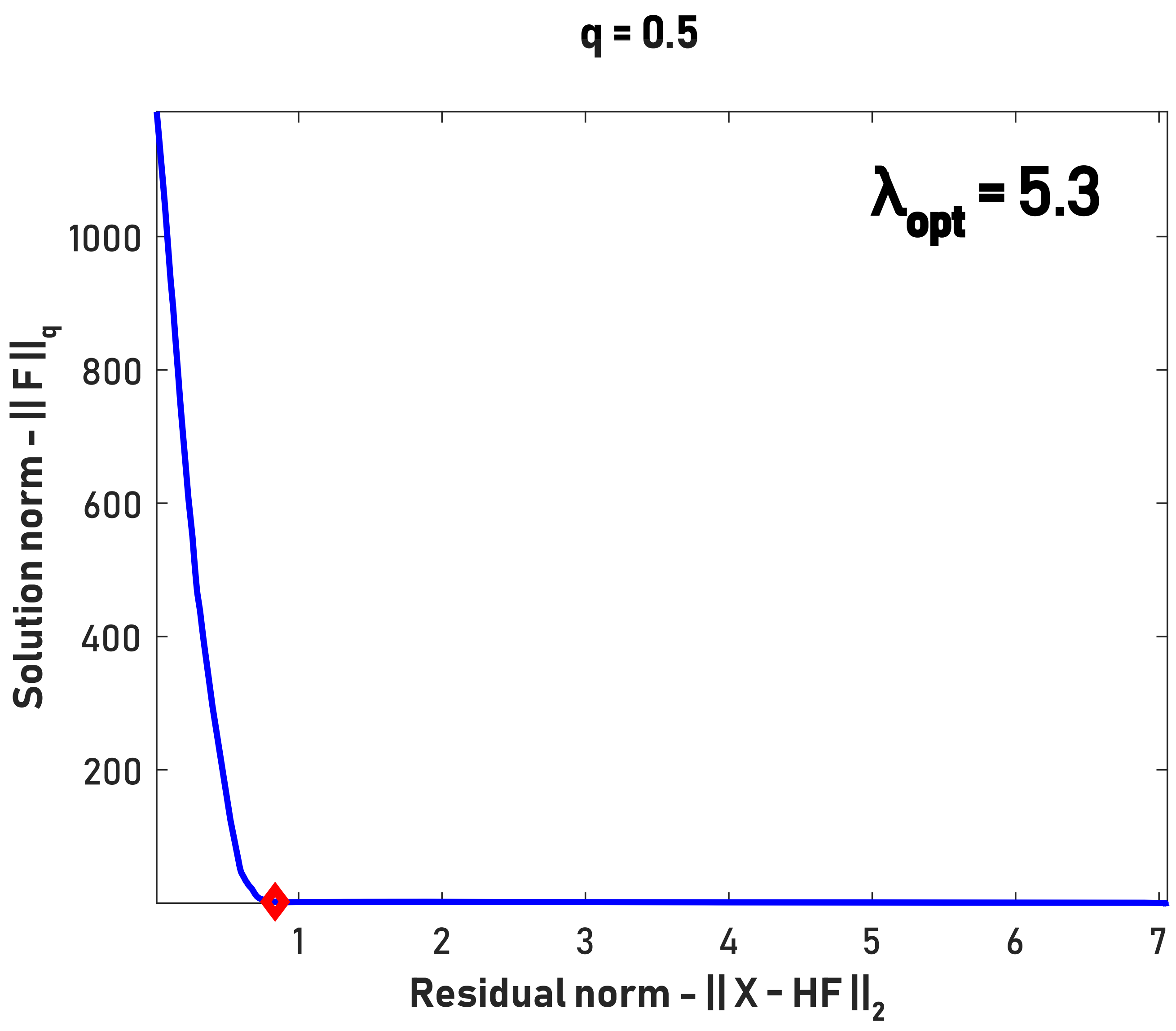
```matlab
function [F, lamb] = lp_reg(H, X, q, tol)

% Initialization
N = size(H, 2)
Hh = H'*H; % For speed
Hx = H'*X;

L = eye(N)
lamb = lcurve(H, L, X);
F = (Hh + lamb*L)\(Hx);
F0 = F; % For convergence monitoring

% Iteration
crit = 1; % Convergence criterion
while crit > tol
        W = weight(F, q);
        L = sqrt(W) % W = L'*L;
        lamb = lcurve(H, L, X);
        F = (Hh + lamb*W)\Hx;

        % Convergence monitoring
        crit = norm(F - F0, 1)/norm(F0, 1);
        F0 = F;
end
```

Python

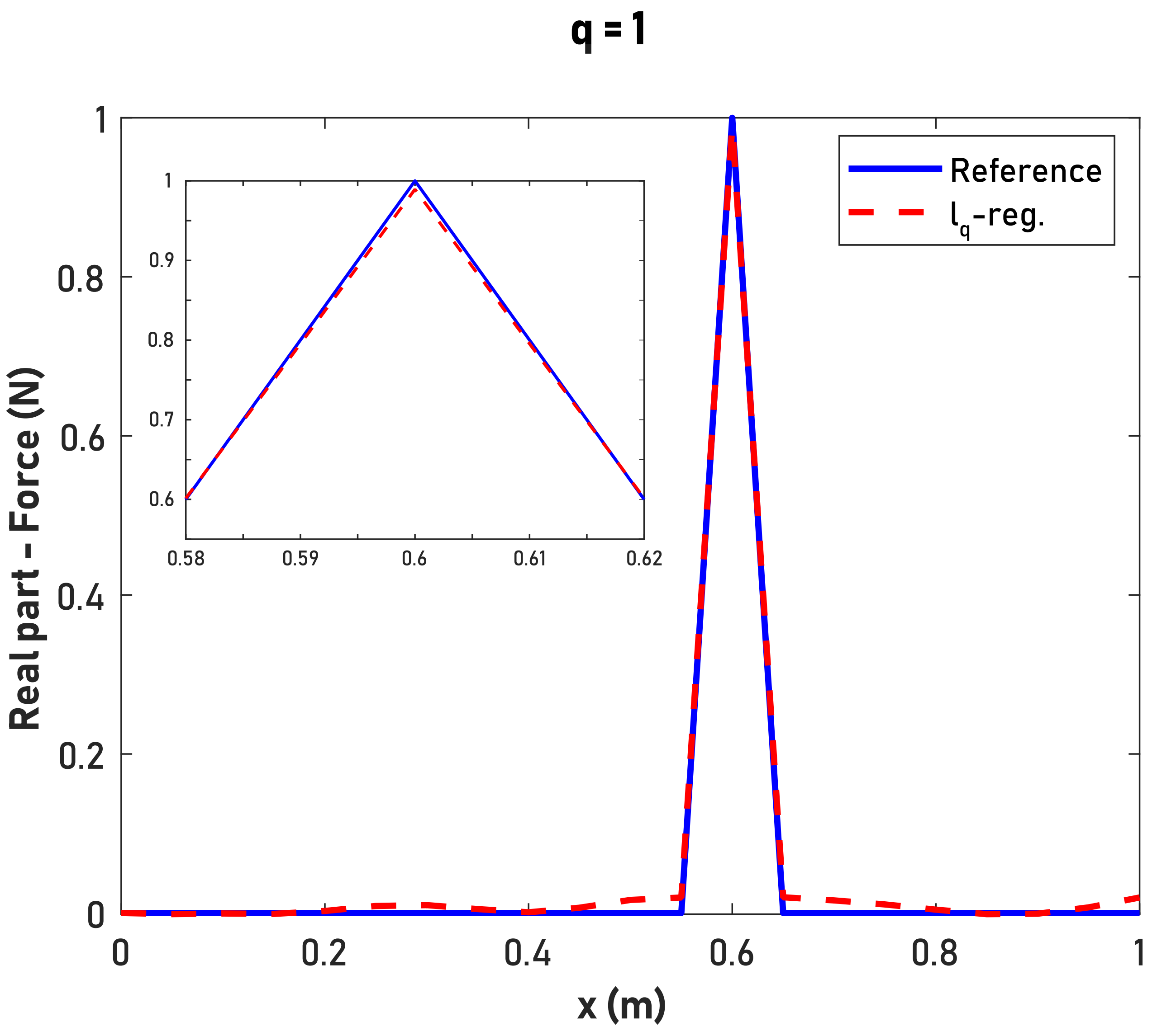
```python
def lp_reg(H, X, q, tol):

        # Initialization
        N = H.shape[1]
        Hh = H.T.conj() @ H # For speed
        Hx = H.T.conj() @ X

        L = np.eye(N)
        lamb = lcurve(H, L, X)
        F = spl.solve(Hh + lamb*L, Hx)
        F0 = F # For convergence monitoring

        # Iteration
        crit = 1 # Convergence criterion
        while crit > tol:
                W = weight(F, q)
                L = np.sqrt(W) # W = L.T.conj()*L;
                lamb = lcurve(H, L, X)
                F = spl.solve(Hh + lamb*W, Hx)

                # Convergence monitoring
                crit = spl.norm(F - F0, 1)/spl.norm(F0, 1)
                F0 = F

        return F, lamb
```

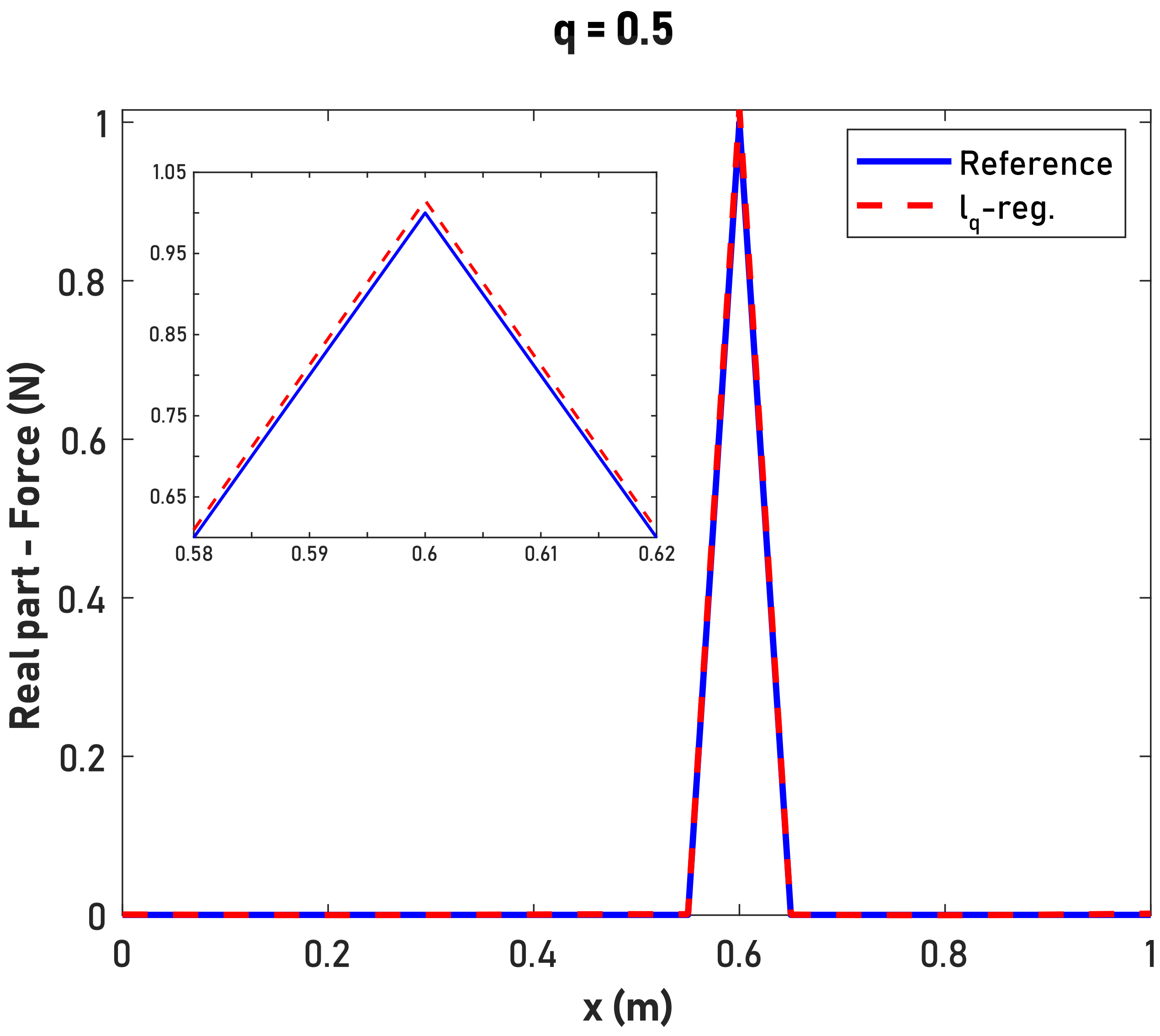
ℓ_q-regularization Sparse regularization





ℓ_q-regularization Sparse regularization





Filter factor analysis

Summary of regularization strategies

✓ Efficient approaches

✓ Easy implementation of resolution algorithms

But...

~ Require external procedures to determine the regularization parameter

~ Provide only point estimate ⇒ No uncertainty quantification of identified solutions

Possible solution ?

Summary of regularization strategies

- ✓ Efficient approaches
- ✓ Easy implementation of resolution algorithms

But...

- ∼ Require external procedures to determine the regularization parameter
- ∼ Provide only point estimate ⇒ No uncertainty quantification of identified solutions

Exploit the Bayesian paradigm !

Outline

1. Generalities
2. State of the art
3. **Bayesian Force regularization**
4. Extensions

Preliminaries Bayes' rule (1763 – posthumously)

For two events A and B

$$p(A|B) \propto p(B|A)\, p(A)$$

- $p(A|B)$ – **Posterior probability distribution**
  probability of A given a realization of B
- $p(B|A)$ – **Likelihood function**
  probability of B given a realization of A
- $p(A)$ – **Prior probability distribution**
  probability of A without any given conditions

The Bayes' rule updates our prior belief in A considering new information brought by an event B

Minimal formulation Basics

When choosing $A = \mathbf{F}$ and $B = \mathbf{X}$

$$p(\mathbf{F}|\mathbf{X}) \propto p(\mathbf{X}|\mathbf{F})\ p(\mathbf{F})$$

How to choose $p(\mathbf{X}|\mathbf{F})$ and $p(\mathbf{F})$?

Minimal formulation Likelihood function

The likelihood function describes the probability of the observed data as a function of the parameters of the chosen statistical model. Given our linear model $\mathbf{X} = \mathbf{H}\mathbf{F} + \mathbf{N}$, it reflects the uncertainty related to vibration measurements, i.e. related to measurement noise

Main assumption

The noise is due to multiple independent causes ⇒ **Gaussian white noise**

$$p(\mathbf{X}|\mathbf{F}, \tau_n) = \mathcal{N}_c(\mathbf{X}|\mathbf{H}\mathbf{F}, \tau_n^{-1}\,\mathbf{I})$$

Minimal formulation Likelihood function

The likelihood function describes the probability of the observed data as a function of the parameters of the chosen statistical model. Given our linear model $\mathbf{X} = \mathbf{HF} + \mathbf{N}$, it reflects the uncertainty related to vibration measurements, i.e. related to measurement noise

Main assumption

The noise is due to multiple independent causes ⇒ **Gaussian white noise**

$$p(\mathbf{X}|\mathbf{F}, \tau_n) = \left(\frac{\tau_n}{\pi}\right)^N \exp\left(-\tau_n \|\mathbf{X} - \mathbf{HF}\|_2^2\right)$$

- τ_n - Noise precision ($\tau_n > 0$)
- N - Number of measurement points

Minimal formulation Prior probability distribution

The prior probability distribution reflects the uncertainty related to $\mathbf{F}$ and can be seen as a measure of our prior knowledge on the sources to identify

Main assumption

$\mathbf{F}$ *is a real random vector, whose components are i.i.d. random variables following a Generalized Gaussian distribution*

$$p(\mathbf{F}|\tau_f, q) = \prod_{i=1}^{M} \mathcal{N}_g(F_i|\tau_f, q)$$

Minimal formulation Prior probability distribution

The prior probability distribution reflects the uncertainty related to $\mathbf{F}$ and can be seen as a measure of our prior knowledge on the sources to identify

Main assumption

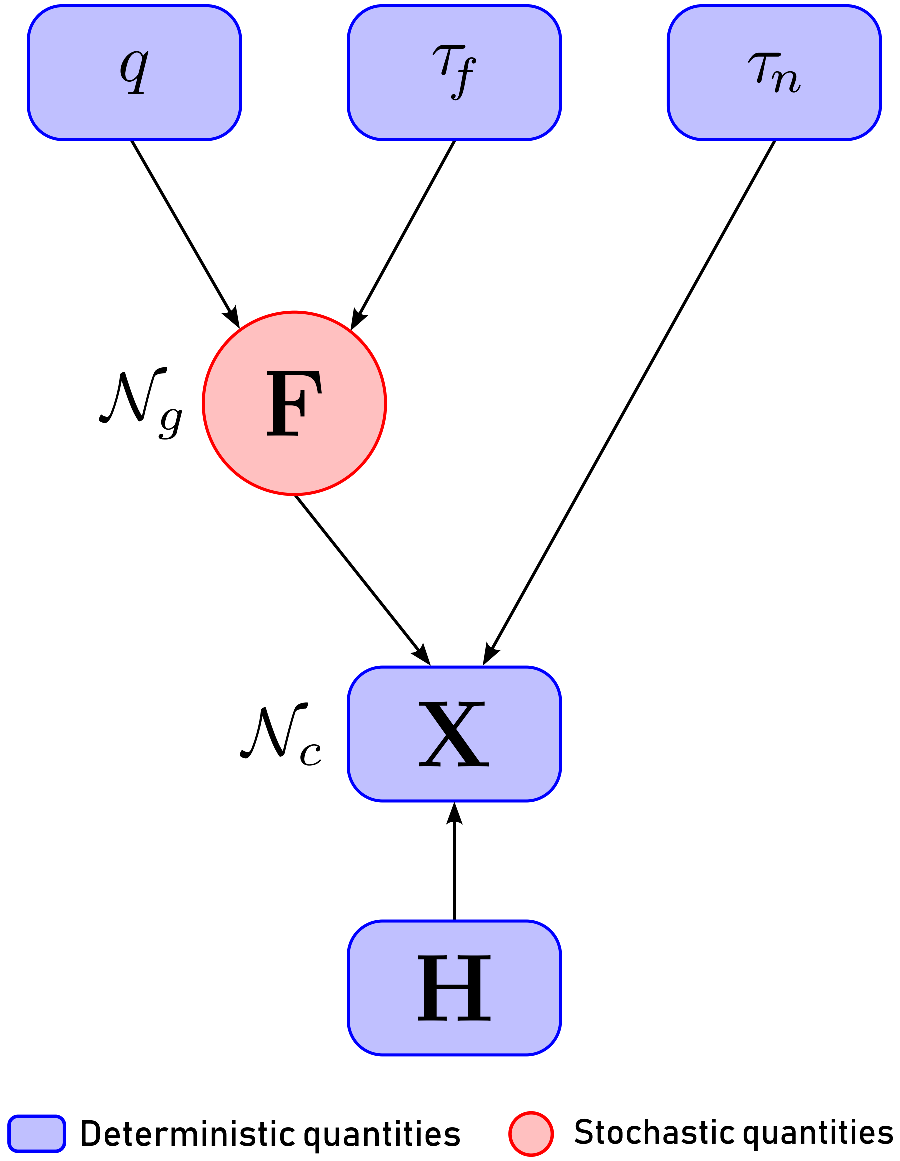
$\mathbf{F}$ *is a real random vector, whose components are i.i.d. random variables following a Generalized Gaussian distribution*

$$p(\mathbf{F}|\tau_f, q) = \left(\frac{q}{2\Gamma(1/q)}\right)^M \tau_f^{\frac{M}{q}} \exp\left(-\tau_f \|\mathbf{F}\|_q^q\right)$$

- q - Shape parameter of the distribution ($q > 0$)
- τ_f - Scale parameter of the distribution ($\tau_f > 0$)
- $\Gamma(x)$ - Gamma function
- M - Number of reconstruction points

Minimal formulation Overview

$$p(\mathbf{F}|\mathbf{X}, \tau_n, \tau_f, q) \propto p(\mathbf{X}|\mathbf{F}, \tau_n)\, p(\mathbf{F}|\tau_f, q)$$



Possible exploitations

- **MAP estimation** – Optimization
- **Uncertainty quantification** – Sampling

Minimal formulation MAP estimation

The MAP estimation consists in finding the most probable excitation field $\mathbf{F}$ given the available data $\mathbf{X}$ and for known precision parameters (τ_n, τ_f) and shape parameter q

$$\widehat{\mathbf{F}} = \underset{\mathbf{F}}{\operatorname{argmax}}\, p(\mathbf{F}|\mathbf{X}, \tau_n, \tau_f, q)$$

Minimal formulation MAP estimation

The MAP estimation consists in finding the most probable excitation field $\mathbf{F}$ given the available data $\mathbf{X}$ and for known precision parameters (τ_n, τ_f) and shape parameter q

$$\widehat{\mathbf{F}} = \underset{\mathbf{F}}{\operatorname{argmax}}\; p(\mathbf{X}|\mathbf{F}, \tau_n)\, p(\mathbf{F}|\tau_f, q)$$

Minimal formulation MAP estimation

The MAP estimation consists in finding the most probable excitation field $\mathbf{F}$ given the available data $\mathbf{X}$ and for known precision parameters (τ_n, τ_f) and shape parameter q

$$\widehat{\mathbf{F}} = \underset{\mathbf{F}}{\operatorname{argmin}} \quad -\log\left[p(\mathbf{X}|\mathbf{F}, \tau_n)\right] - \log\left[p(\mathbf{F}|\tau_f, q)\right]$$

Minimal formulation MAP estimation

The MAP estimation consists in finding the most probable excitation field $\mathbf{F}$ given the available data $\mathbf{X}$ and for known precision parameters (τ_n, τ_f) and shape parameter q

$$\widehat{\mathbf{F}} = \underset{\mathbf{F}}{\operatorname{argmin}} \; \|\mathbf{X} - \mathbf{HF}\|_2^2 + \lambda \|\mathbf{F}\|_q^q \;\; \text{with} \;\; \lambda = \frac{\tau_f}{\tau_n}$$

MAP estimation $\equiv \ell_q$-regularization !

Minimal formulation Uncertainty quantification

Idea for posterior sampling Transform the Generalized Gaussian into a multivariate Gaussian distribution

$$p(\mathbf{F}|\tau_f, q) \propto \exp\left(-\tau_f \|\mathbf{LF}\|_2^2\right)$$

where $\mathbf{L}^H\mathbf{L} = \mathbf{W}$ is a weigthing depending on $\mathbf{F}$ and q

In doing so, one has

$$\begin{aligned} p(\mathbf{F}|\mathbf{X}) &\propto \exp\left(-\tau_n \|\mathbf{X} - \mathbf{HF}\|_2^2 - \tau_f \|\mathbf{LF}\|_2^2\right) \\ &\propto \mathcal{N}_c(\mathbf{F}|\boldsymbol{\mu}_\mathbf{F}, \boldsymbol{\Sigma}_\mathbf{F}) \end{aligned}$$

where $\boldsymbol{\mu}_\mathbf{F} = \tau_n \boldsymbol{\Sigma}_\mathbf{F} \mathbf{H}^H \mathbf{X}$ and $\boldsymbol{\Sigma}_\mathbf{F} = \left(\tau_n \mathbf{H}^H\mathbf{H} + \tau_f \mathbf{W}\right)^{-1}$

Drawing samples

$$\mathbf{F}^{(k)} = \boldsymbol{\mu}_\mathbf{F} + \mathbf{S}\,\mathbf{z}^{(k)} \quad \text{with} \quad \mathbf{S}\mathbf{S}^H = \boldsymbol{\Sigma}_\mathbf{F} \quad \text{and} \quad \mathbf{z}^{(k)} \sim \mathcal{N}_c(\mathbf{z}^{(k)}|\mathbf{0}, \mathbf{I})$$

Properties of Gaussian distributions

Minimal formulation Uncertainty quantification

Estimation of $\boldsymbol{\tau_n}$ and $\boldsymbol{\tau_f}$

$\boldsymbol{\mu}_{\mathbf{F}}$ is the solution of the ℓ_q-regularization $\Rightarrow$ After convergence of the iterative process, one obtains $\boldsymbol{\mu}_{\mathbf{F}}$, $\mathbf{W}$ and λ

From these quantities, the most probable values of τ_n and τ_f given the data are computed such that

$$(\widehat{\tau}_n, \widehat{\tau}_f) = \underset{(\tau_n, \tau_f)}{\operatorname{argmax}}\, p(\tau_n, \tau_f | \mathbf{X})$$

Minimal formulation Uncertainty quantification

Estimation of τ_n and τ_f

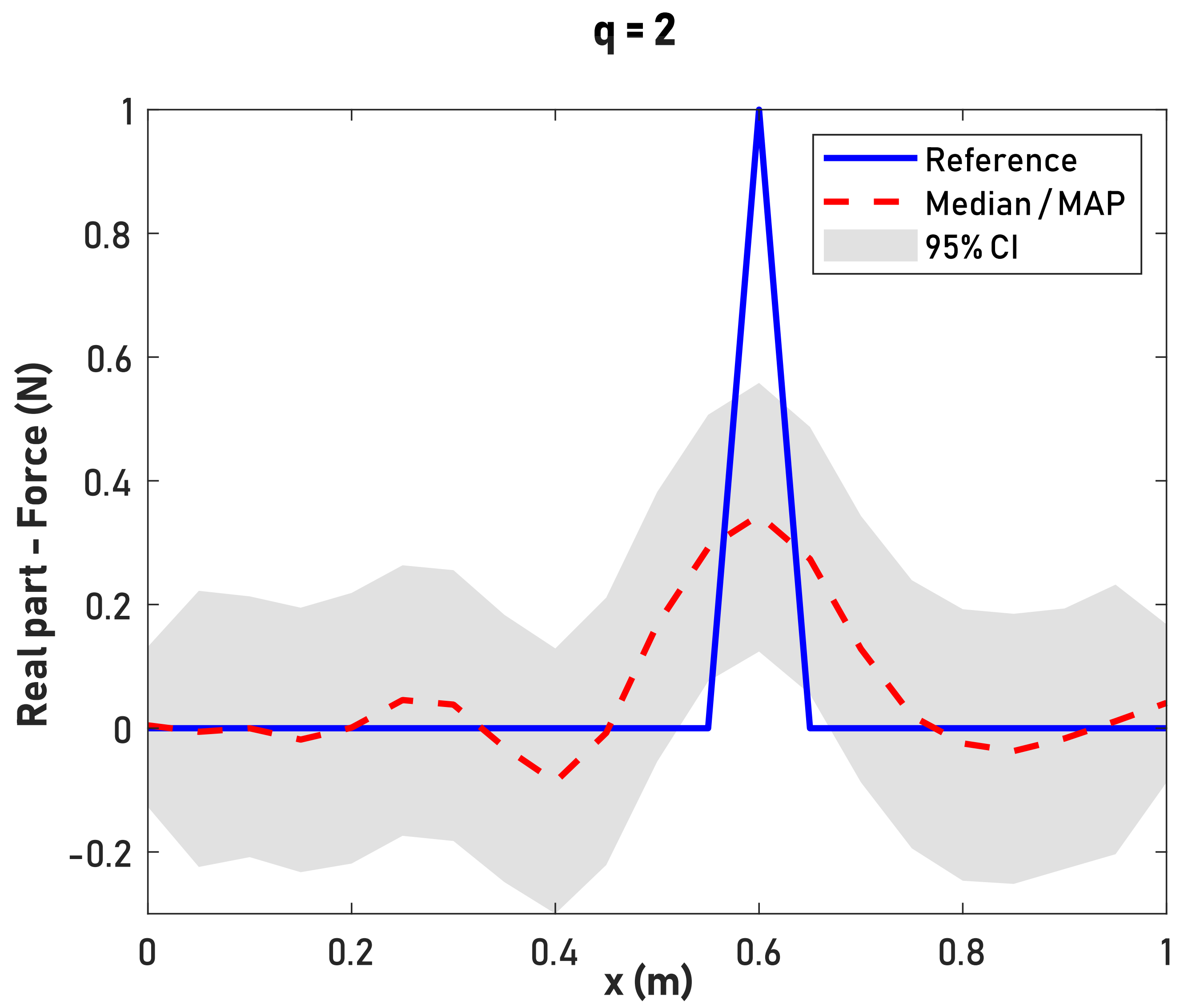
$\boldsymbol{\mu}_{\mathbf{F}}$ is the solution of the ℓ_q-regularization $\Rightarrow$ After convergence of the iterative process, one obtains $\boldsymbol{\mu}_{\mathbf{F}}$, $\mathbf{W}$ and λ

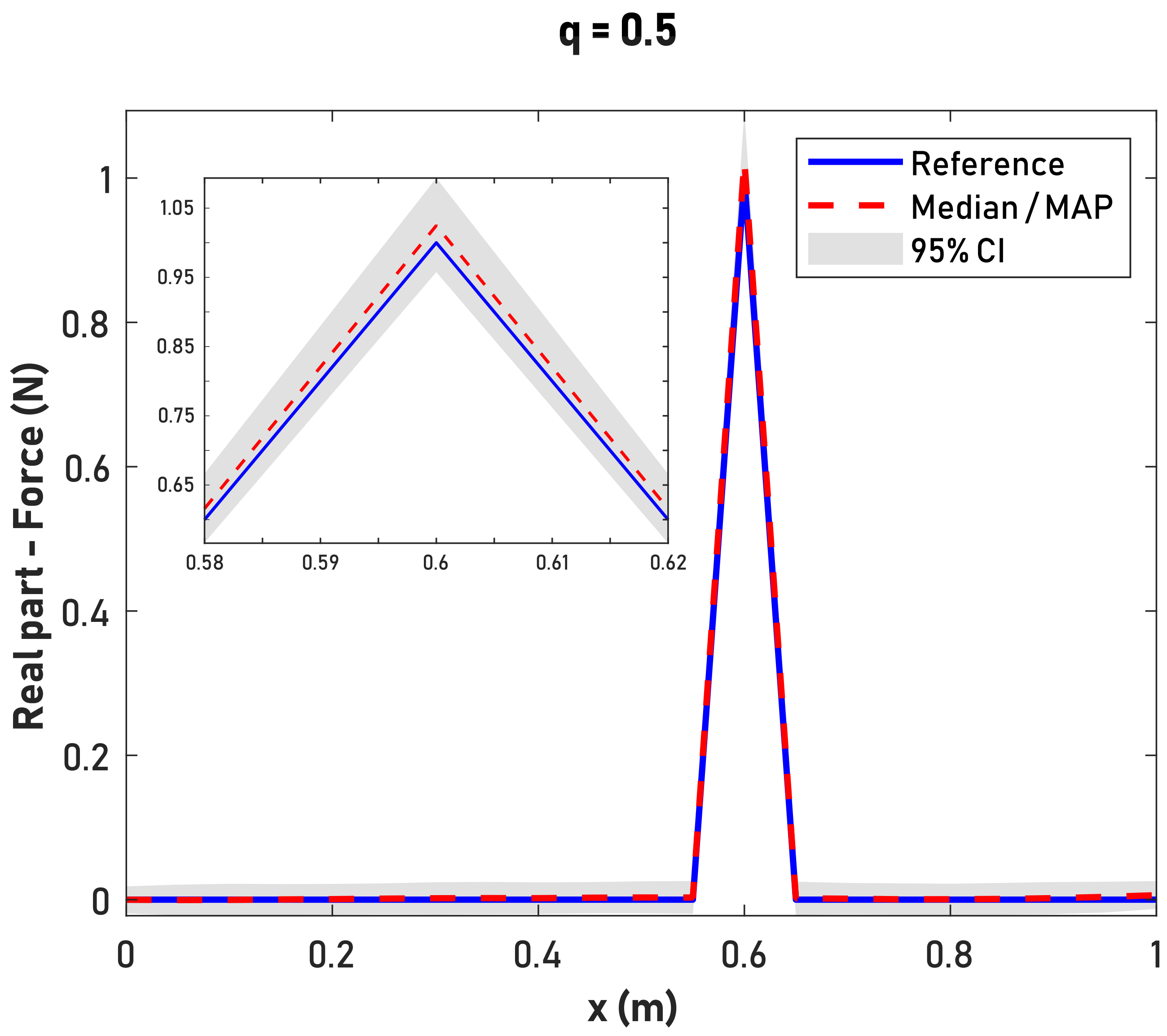
From these quantities, the most probable values of τ_n and τ_f given the data are computed such that

$$\widehat{\tau}_f = \frac{N}{\mathbf{X}^H \left(\mathbf{H}\mathbf{W}^{-1}\mathbf{H}^H + \lambda\mathbf{I}\right)^{-1}\mathbf{X}} \quad \text{and} \quad \widehat{\tau}_n = \frac{\widehat{\tau}_f}{\lambda}$$

Proof

Minimal formulation Application





Minimal formulation Summary

✓ MAP is equivalent to ℓ_q-regularization

✓ Easy implementation of uncertainty quantification

Provided that...

~ External procedures is implemented to estimate the precision parameters τ_n and τ_f

~ The shape parameter q is known a priori

Need for a more comprehensive formulation

Complete formulation Basics

Choosing a priori relevant values for τ_n, τ_f and q is far from an easy task for non-experienced users $\Rightarrow$ **Infer them !**

Main assumption $\boldsymbol{\tau_n, \tau_f}$ **and** $\boldsymbol{q}$ **are independent random variables**

$$p(\mathbf{F}, \tau_n, \tau_f, q|\mathbf{X}) \propto p(\mathbf{X}|\mathbf{F}, \tau_n)\, p(\mathbf{F}|\tau_f, q)\, p(\tau_n)\, p(\tau_f)\, p(q)$$

- $p(\tau)$ - Prior distrubution on the precision parameter τ
- $p(q)$ - Prior distribution on the shape parameter q

How to choose $p(\tau)$ and $p(q)$?

Complete formulation Prior distribution $p(\tau)$ - Gamma distribution

$$p(\tau|\alpha,\beta) = \mathcal{G}(\tau|\alpha,\beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)}\tau^{\alpha-1}\exp(-\beta\tau) \text{ with } \alpha > 0,\ \beta > 0$$

- α - Scale parameter
- β - Shape parameter

This choice is made for mathematical convenience (conjugate prior), but it does not reflect any real prior information on the precision parameters, except their positiveness

➡ Prior distribution on τ should be as minimally informative as possible (flat prior). For this reason, $\alpha = 1$ and $\beta = 10^{-18}$

Complete formulation Prior distribution $p(q)$ - Truncated Gamma distribution

$$p(q|\alpha_q, \beta_q, l_b, u_b) = \frac{\Gamma(\alpha_q)}{\gamma(\alpha_q, \beta_q u_b) - \gamma(\alpha_q, \beta_q l_b)} \mathcal{G}(q|\alpha_q, \beta_q) \mathbb{I}_{[l_b, u_b]}(q)$$

- $\mathbb{I}_{[l_b,u_b]}(q)$ - Truncation function, defined such that

$$\mathbb{I}_{[l_b,u_b]}(q) = \begin{cases} 1 & \text{if } q \in [l_b, u_b] \\ 0 & \text{otherwise} \end{cases}$$

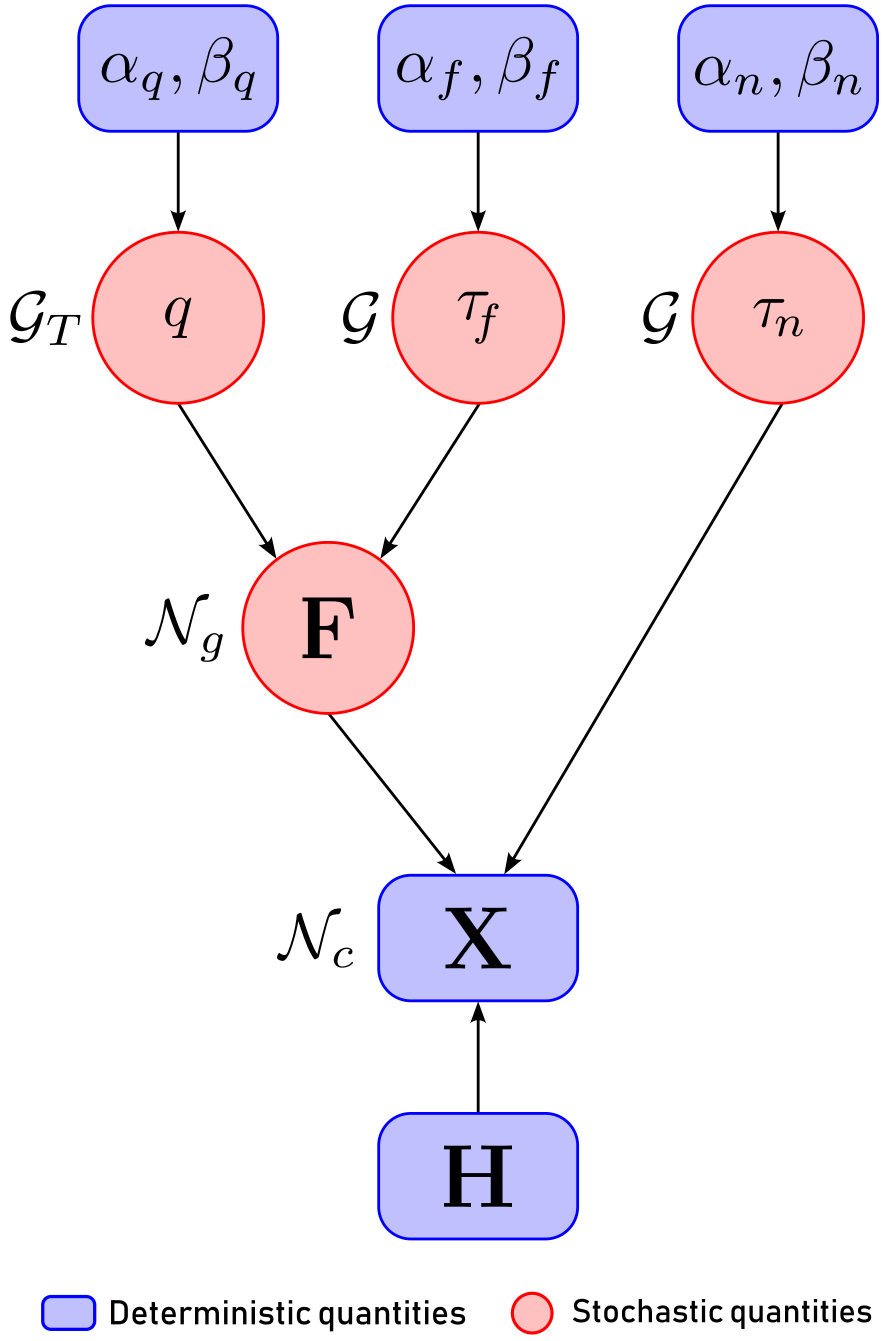
- $\gamma(s, x)$ - Lower incomplete Gamma function

Requirements

- Expert knowledge $\Longrightarrow l_b = 0.05$ and $u_b = 2.05$
- Weakly informative distribution $\Longrightarrow \alpha_q = 1$ and $\beta_q = 10^{-18}$

Complete formulation Overview

$$p(\mathbf{F}, \tau_n, \tau_f, q|\mathbf{X}) \propto$$
$$p(\mathbf{X}|\mathbf{F}, \tau_n)\, p(\mathbf{F}|\tau_f, q)\, p(\tau_n|\alpha_n, \beta_n)\, p(\tau_f|\alpha_f, \beta_f)\, p(q|\alpha_q, \beta_q)$$



Possible exploitations

- **MAP estimation** – Optimization
- **Uncertainty quantification** – Sampling

Complete formulation MAP estimation

The MAP estimate of the complete formulation is given by

$$\left(\widehat{\mathbf{F}}, \widehat{\tau}_n, \widehat{\tau}_n, \hat{q}\right) = \underset{\mathbf{F},\tau_n,\tau_f,q}{\operatorname{argmax}}\; p(\mathbf{F}, \tau_n, \tau_f, q|\mathbf{X})$$

The solution of the previous problem can be found by applying the first-order optimality condition to the dual minimization problem. An alternative, but equivalent, way of solving this problem consists in maximizing the full conditional probability distributions associated to each parameter

$$\begin{aligned}
\hat{q} &= \underset{q}{\operatorname{argmax}}\; p(q|\mathbf{X}, \mathbf{F}, \tau_n, \tau_f) \\
\widehat{\tau}_f &= \underset{\tau_f}{\operatorname{argmax}}\; p(\tau_f|\mathbf{X}, \mathbf{F}, \tau_n, q) \\
\widehat{\tau}_n &= \underset{\tau_n}{\operatorname{argmax}}\; p(\tau_n|\mathbf{X}, \mathbf{F}, \tau_f, q) \\
\widehat{\mathbf{F}} &= \underset{\mathbf{F}}{\operatorname{argmax}}\; p(\mathbf{F}|\mathbf{X}, \tau_n, \tau_f, q)
\end{aligned}$$

Complete formulation MAP estimation

The MAP estimate of the complete formulation is given by

$$\left(\widehat{\mathbf{F}}, \widehat{\tau}_n, \widehat{\tau}_n, \hat{q}\right) = \underset{\mathbf{F},\tau_n,\tau_f,q}{\operatorname{argmax}}\ p(\mathbf{F}, \tau_n, \tau_f, q|\mathbf{X})$$

The solution of the previous problem can be found by applying the first-order optimality condition to the dual minimization problem. An alternative, but equivalent, way of solving this problem consists in maximizing the full conditional probability distributions associated to each parameter

$$\begin{aligned}
\hat{q} &= \underset{q}{\operatorname{argmin}}\ f(q|\widehat{\mathbf{F}}, \widehat{\tau}_f) \\
\widehat{\tau}_f &= \frac{M + \hat{q}(\alpha_f - 1)}{\hat{q}(\beta_f + \|\widehat{\mathbf{F}}\|_{\hat{q}}^{\hat{q}})} \\
\widehat{\tau}_n &= \frac{N + \alpha_n - 1}{\beta_n + \|\mathbf{X} - \mathbf{H}\widehat{\mathbf{F}}\|_2^2} \\
\widehat{\mathbf{F}} &= \underset{\mathbf{F}}{\operatorname{argmin}}\ \|\mathbf{X} - \mathbf{H}\mathbf{F}\|_2^2 + \lambda\|\mathbf{F}\|_{\hat{q}}^{\hat{q}}
\end{aligned}$$

where $f(q|\mathbf{F}, \tau_f) = M \log\Gamma(1/q) - \frac{M}{q}\log\widehat{\tau}_f - [M + \alpha_q - 1]\log q + \beta_q\, q + \widehat{\tau}_f\|\widehat{\mathbf{F}}\|_q^q$ and $\lambda = \widehat{\tau}_f/\widehat{\tau}_n$

Complete formulation MAP estimation - Iterative resolution

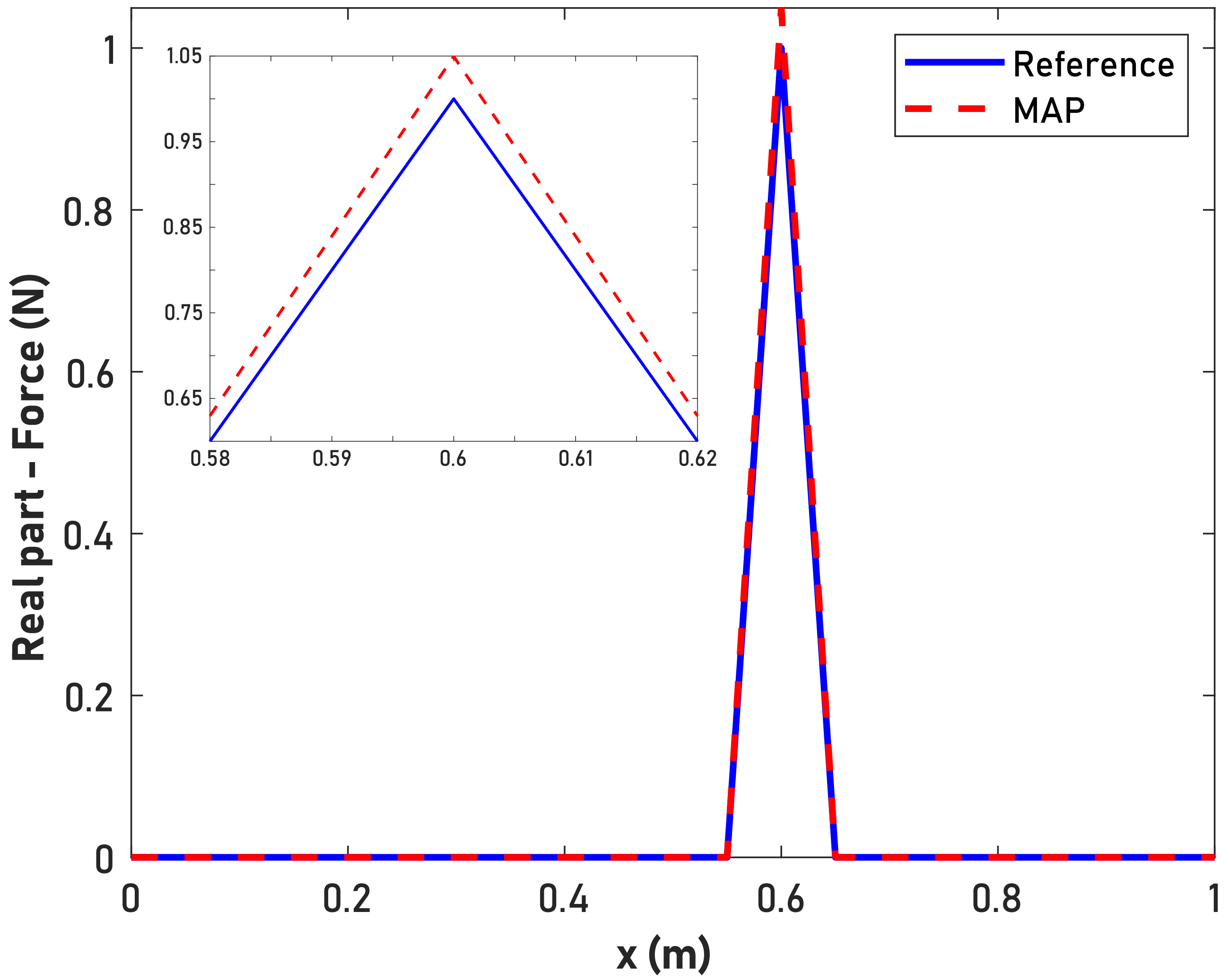
Initialization ℓ_2-regularization $\left(\widehat{\mathbf{F}}^{(0)}, \boldsymbol{\lambda}^{(0)}, \widehat{\mathbf{q}}^{(\mathbf{0})} = \mathbf{2}\right)$ + Estimation of $\boldsymbol{\tau}_{\boldsymbol{f}}^{(\mathbf{0})}$ from $\boldsymbol{\lambda}^{(\mathbf{0})}$

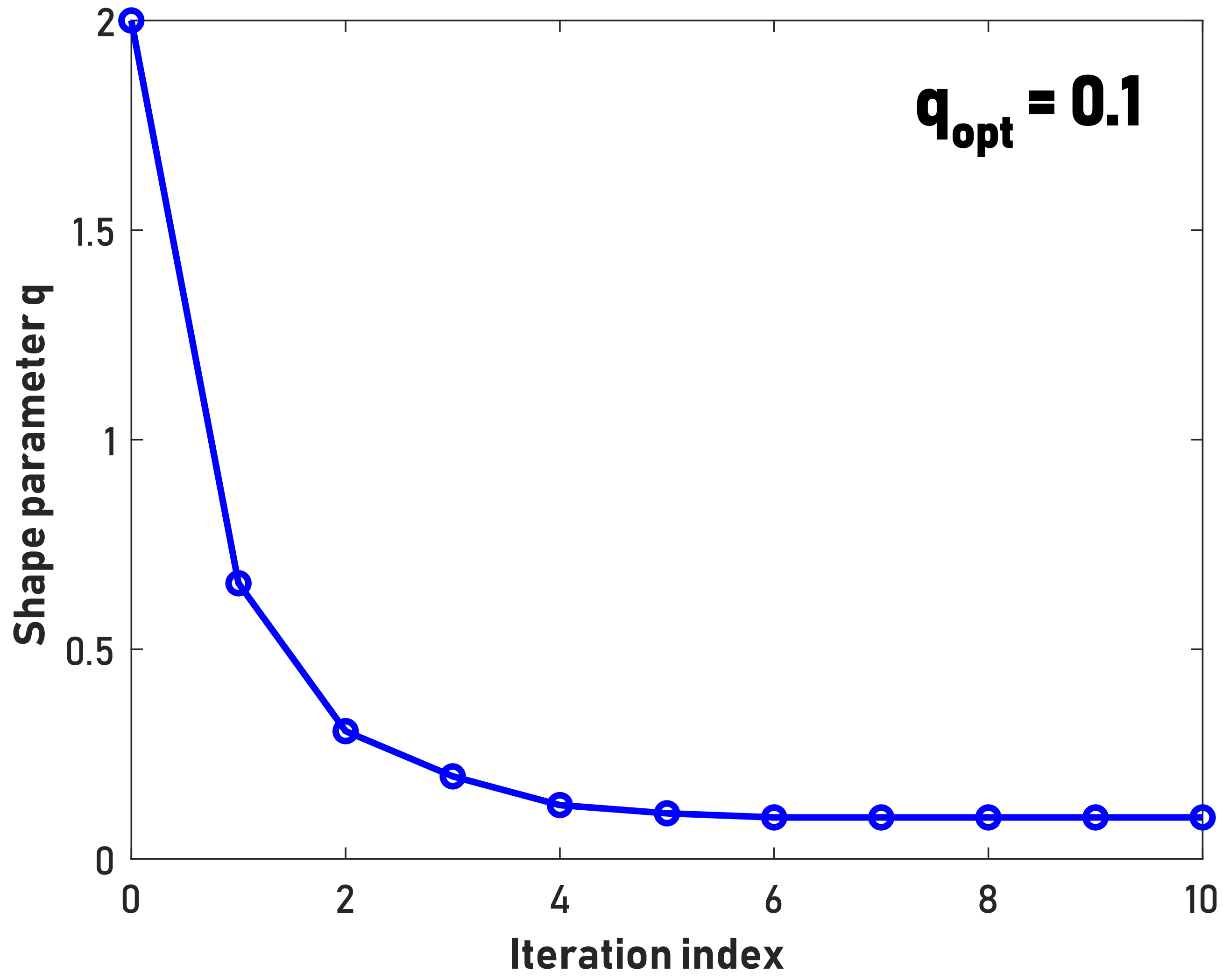
Iteration While convergence is not reached do

$$
\begin{aligned}
\hat{q}^{(k)} &= \underset{q}{\operatorname{argmin}}\ f\left(q|\widehat{\mathbf{F}}^{(k-1)}, \widehat{\tau}_f^{(k-1)}\right) \\
\widehat{\tau}_f^{(k)} &= \frac{M + \hat{q}^{(k)}(\alpha_f - 1)}{\hat{q}^{(k)}\left(\beta_f + \|\widehat{\mathbf{F}}^{(k-1)}\|_{\hat{q}^{(k)}}^{\hat{q}^{(k)}}\right)} \\
\widehat{\tau}_n^{(k)} &= \frac{N + \alpha_n - 1}{\beta_n + \|\mathbf{X} - \mathbf{H}\widehat{\mathbf{F}}^{(k-1)}\|_2^2} \\
\widehat{\mathbf{F}}^{(k)} &= \underset{\mathbf{F}}{\operatorname{argmin}}\ \|\mathbf{X} - \mathbf{H}\mathbf{F}\|_2^2 + \lambda^{(k)}\|\mathbf{F}\|_{\hat{q}^{(k)}}^{\hat{q}^{(k)}}
\end{aligned}
$$

Convergence monitoring $\boldsymbol{\delta} = \|\widehat{\mathbf{F}}^{(\mathbf{k})} - \widehat{\mathbf{F}}^{(\mathbf{k-1})}\|_{\mathbf{1}} / \|\widehat{\mathbf{F}}^{(\mathbf{k-1})}\|_{\mathbf{1}}$

Complete formulation MAP estimation – Application





Complete formulation Uncertainty quantification - MCMC

Markov Chain Monte Carlo (MCMC) is a class of algorithms that produce sequences of random samples converging to a target distribution for which direct sampling is difficult

Here, because the full conditional probability distributions are available, a Gibbs sampler (particular case of MH sampler) can be implemented

$$
\begin{aligned}
p(q|\mathbf{X},\mathbf{F},\tau_n,\tau_f) &\propto \frac{\tau_f^{M/q}}{\Gamma(1/q)} q^{M+\alpha_q-1} \exp\big(-\beta_q\, q - \tau_f \|\mathbf{F}\|_q^q\big)\mathbb{I}_{[l_b,u_b]} \\
p(\tau_f|\mathbf{X},\mathbf{F},\tau_n,q) &\propto \mathcal{G}(\tau_f|\alpha_f + M/q, \beta_f + \|\mathbf{F}\|_q^q) \\
p(\tau_n|\mathbf{X},\mathbf{F},\tau_f,q) &\propto \mathcal{G}(\tau_n|\alpha_n + N, \beta_n + \|\mathbf{X}-\mathbf{HF}\|_2^2) \\
p(\mathbf{F}|\mathbf{X},\tau_n,\tau_f,q) &\propto \exp\big(-\tau_n\|\mathbf{X}-\mathbf{HF}\|_2^2 - \tau_f\|\mathbf{F}\|_q^q\big)
\end{aligned}
$$

Build a markov chain

for each parameter to compute statistics

Complete formulation Uncertainty quantification - Gibbs sampling

General scheme

1. Set $k = 0$ and initialize $q^{(0)}$, $\tau_n^{(0)}$, $\tau_f^{(0)}$ and $\mathbf{F}^{(0)}$
2. Draw N_s samples from full conditional distributions

   for $k = 1 : N_s$
   - Draw $q^{(k)} \sim p\left(q|\mathbf{X}, \mathbf{F}^{(k-1)}, \tau_n^{(k-1)}, \tau_f^{(k-1)}\right)$
   - Draw $\tau_f^{(k)} \sim p\left(\tau_f|\mathbf{X}, \mathbf{F}^{(k-1)}, \tau_n^{(k-1)}, q^{(k)}\right)$
   - Draw $\tau_n^{(k)} \sim p\left(\tau_n|\mathbf{X}, \mathbf{F}^{(k-1)}, \tau_f^{(k)}, q^{(k)}\right)$
   - Draw $\mathbf{F}^{(k)} \sim p\left(\mathbf{F}|\mathbf{X}, \tau_n^{(k)}, \tau_f^{(k)}, q^{(k)}\right)$

   end for
3. Monitor the convergence (stationarity) of the Markov chains

Complete formulation Uncertainty quantification – Implementation

Initialization

- ℓ_2-regularization $\left(\mathbf{F}^{(0)}, \lambda^{(0)}, q^{(0)}\right)$ + Estimation of $\tau_n^{(0)}$ and $\tau_f^{(0)}$ from $\lambda^{(0)}$
- MAP estimate $\left(\mathbf{F}^{(0)}, \tau_f^{(0)}, \tau_n^{(0)}, q^{(0)}\right)$

Drawing samples

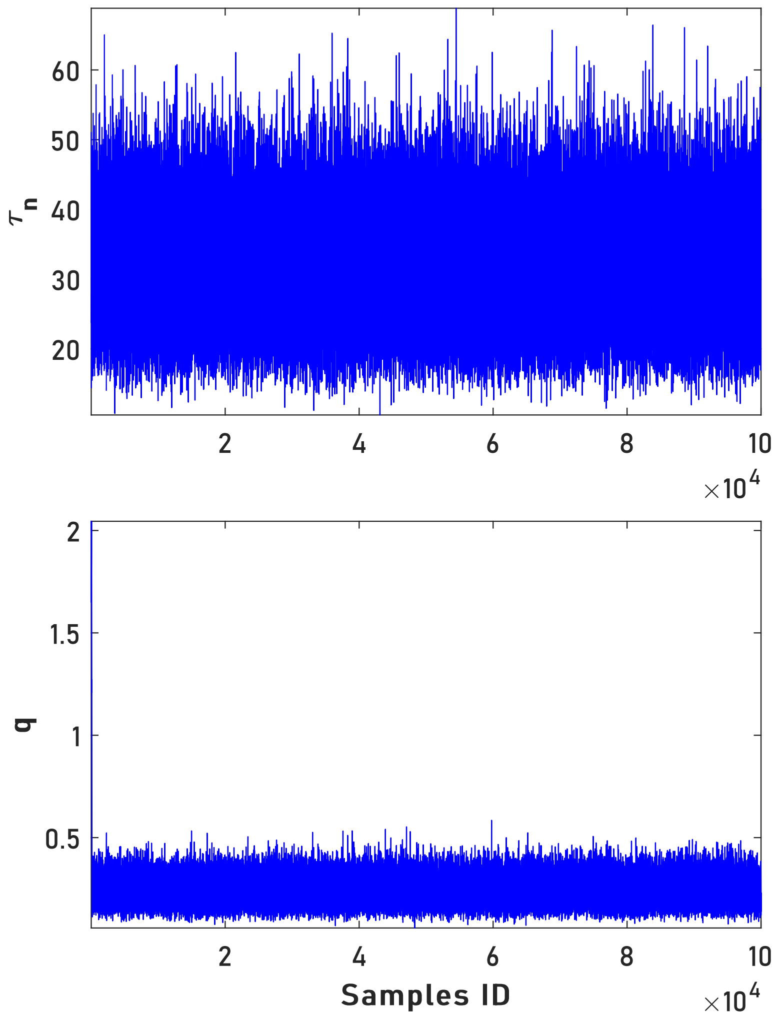
1. $p\left(q|\mathbf{X}, \mathbf{F}^{(k-1)}, \tau_n^{(k-1)}, \tau_f^{(k-1)}\right)$ – Non-standard probability distribution ⇒ Instance of MH sampler (or HMC, ...)
2. $p\left(\tau_i|\mathbf{X}, \mathbf{F}^{(k-1)}, \tau_j^{(k-1)}, q^{(k)}\right)$ – Gamma distribution ⇒ RNG implemented in standard programming languages
3. $p\left(\mathbf{F}|\mathbf{X}, \tau_n^{(k)}, \tau_f^{(k)}, q^{(k)}\right)$ – Multivariate Gaussian-like distribution ⇒ Procedure defined for the min. formulation

Convergence diagnostic

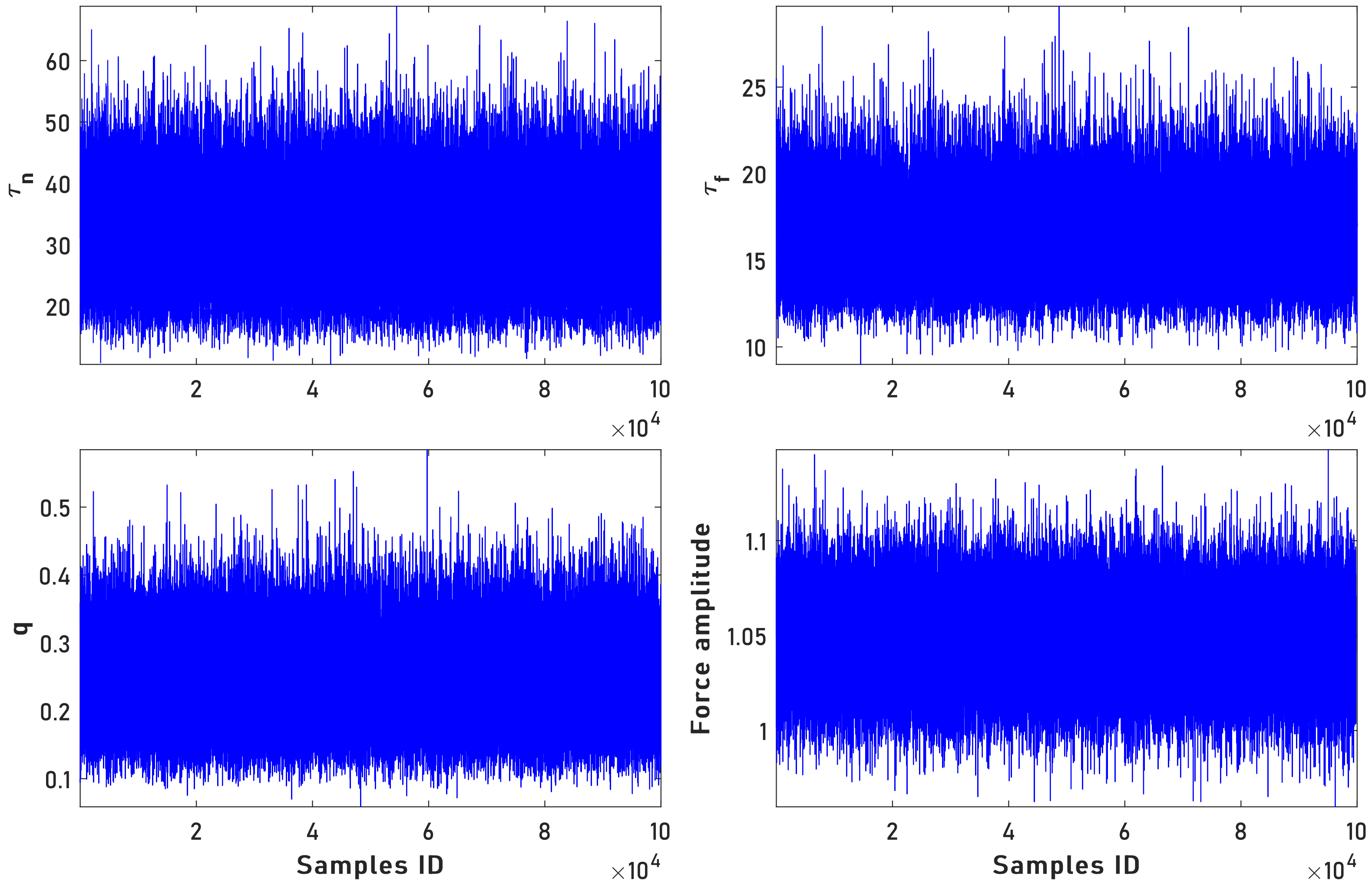
- **Burn-in period** – Number of samples to discard at the beginning of the chains (period before convergence)
- **Total length** – Number of samples required to compute statistics
- **Available diagnostics** – Raftery-Lewis, Geweke (one long chain), Gelman-Rubin (multiple chains) and more

Complete formulation Uncertainty quantification – Application

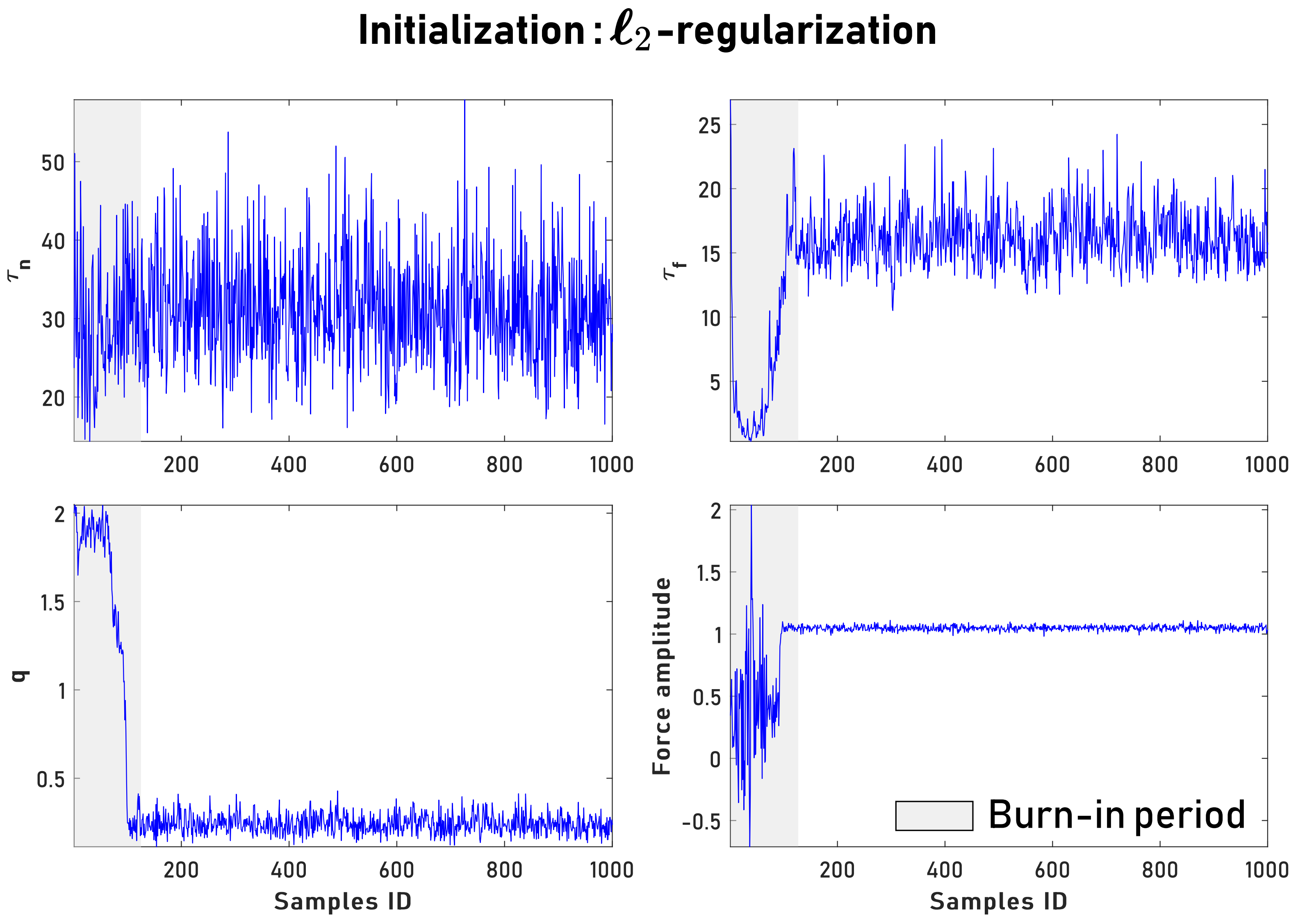
Initialization : ℓ_2-regularization

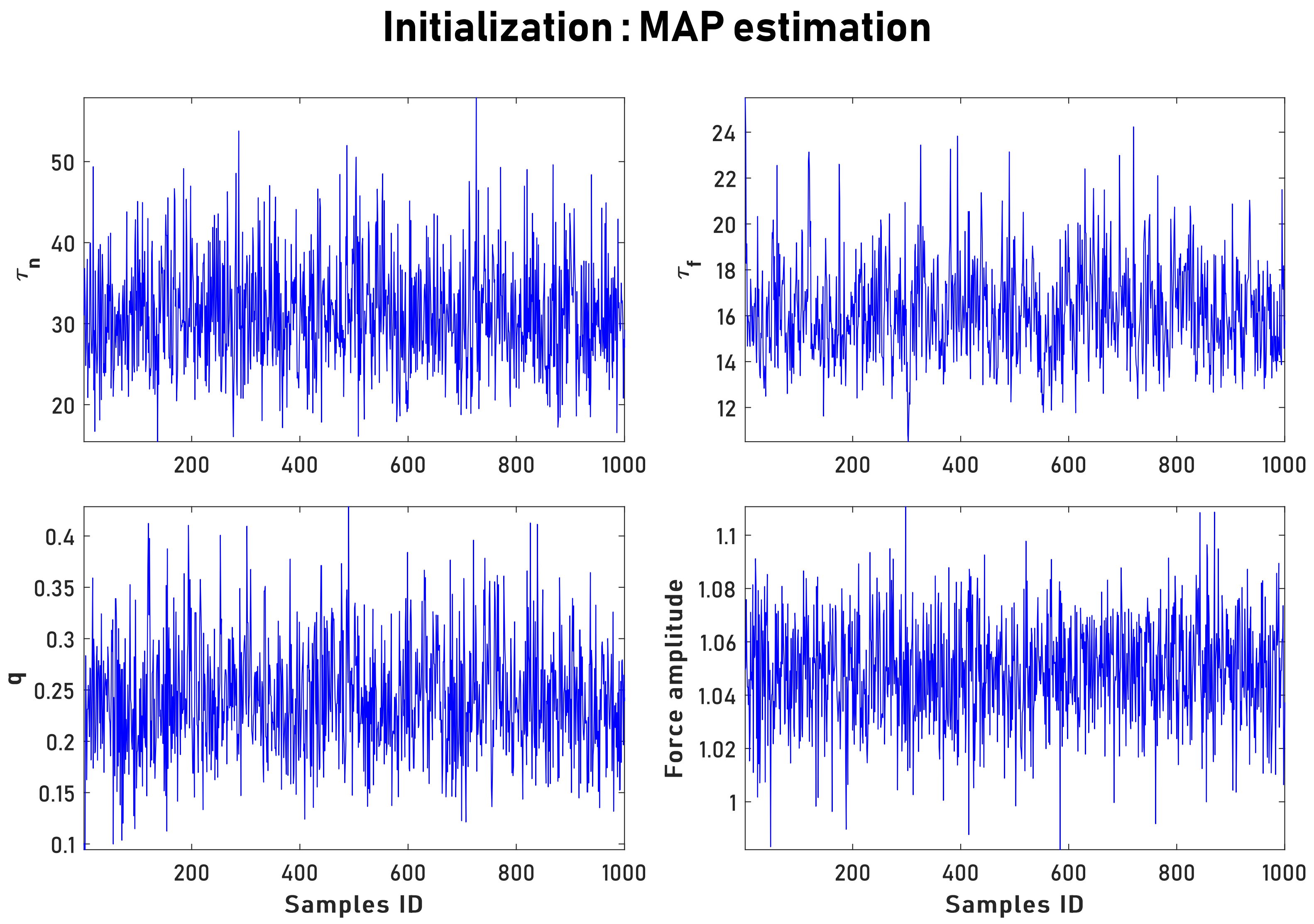


Initialization : MAP estimation

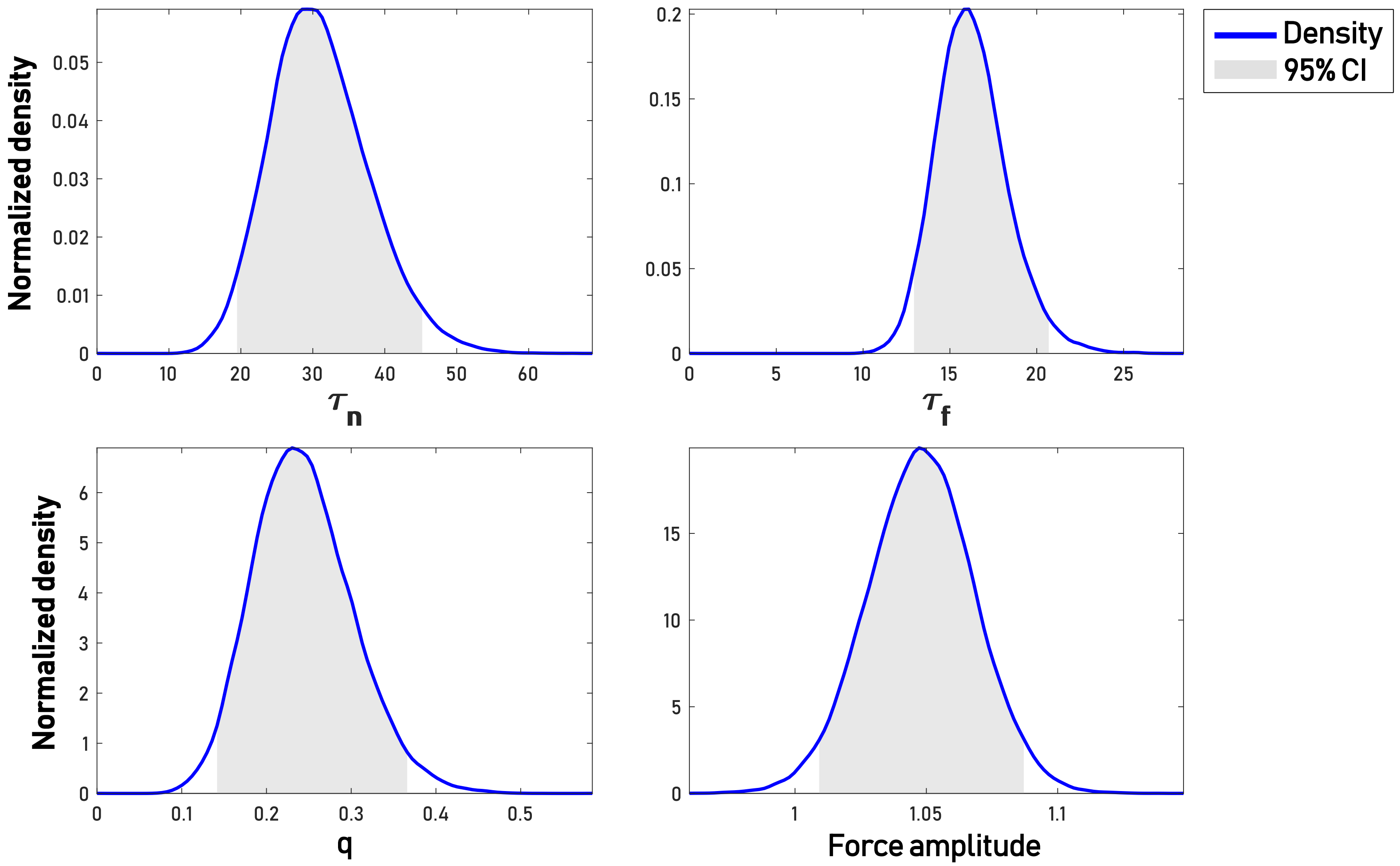


Complete formulation Uncertainty quantification – Application



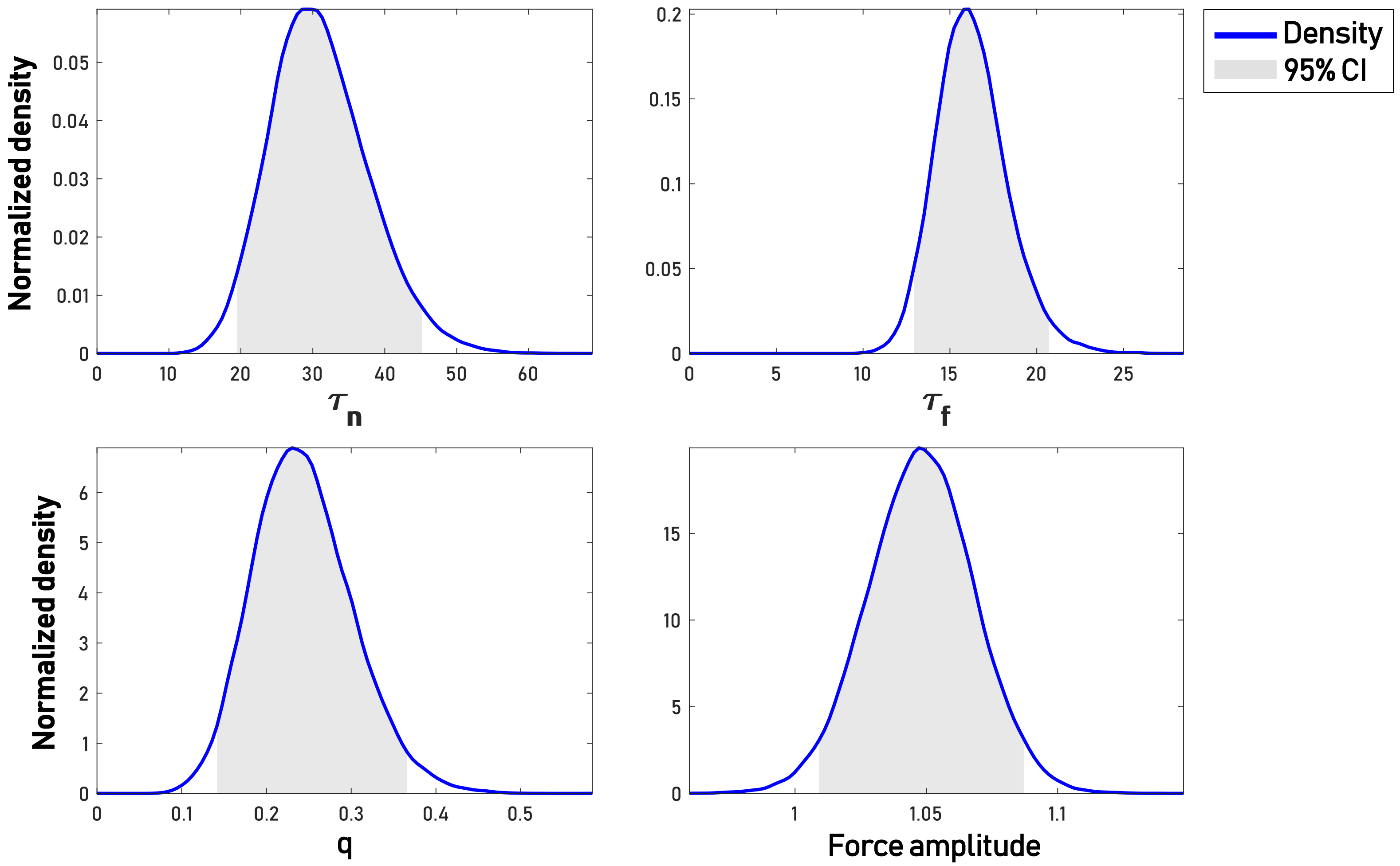


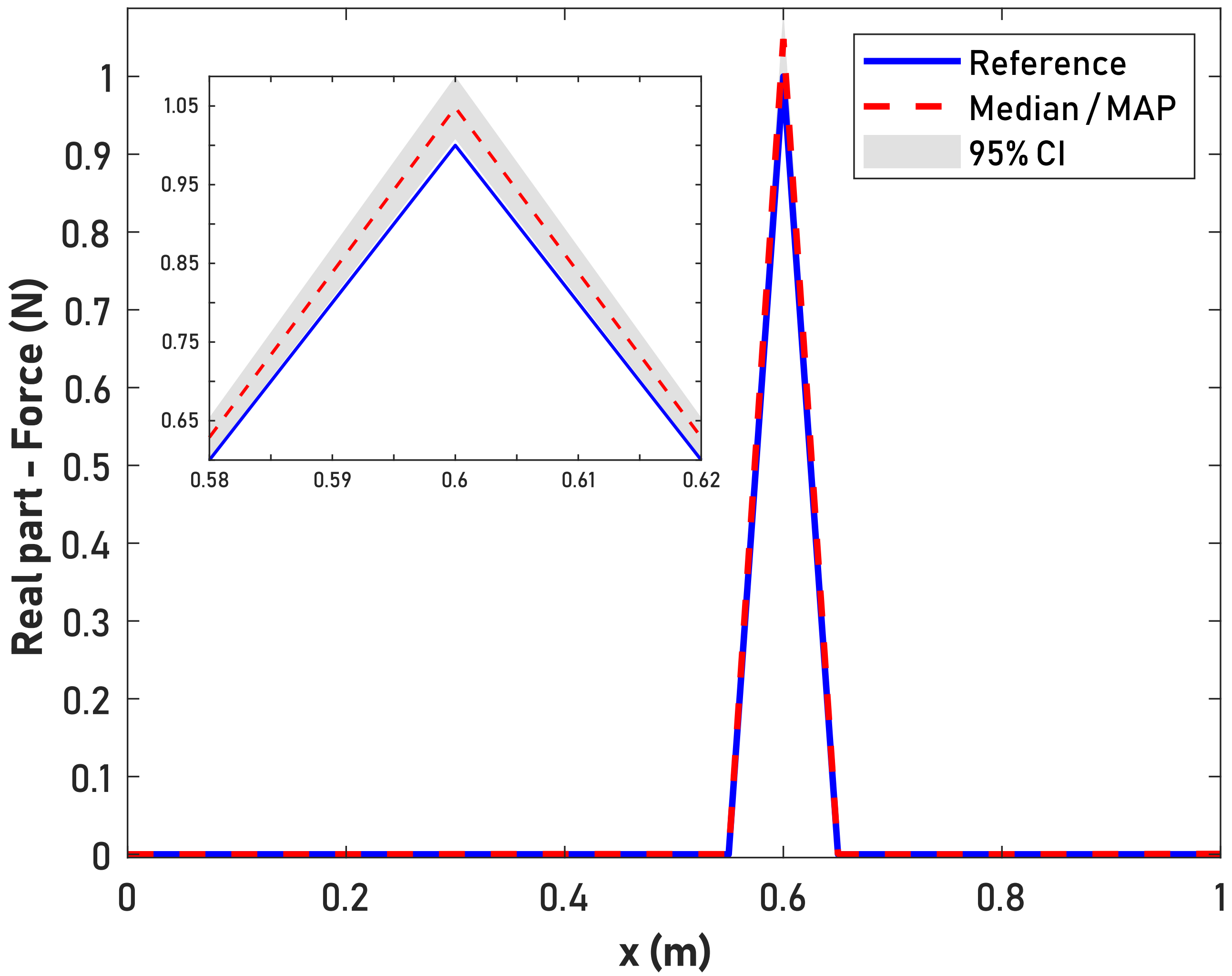
Complete formulation Uncertainty quantification – Application



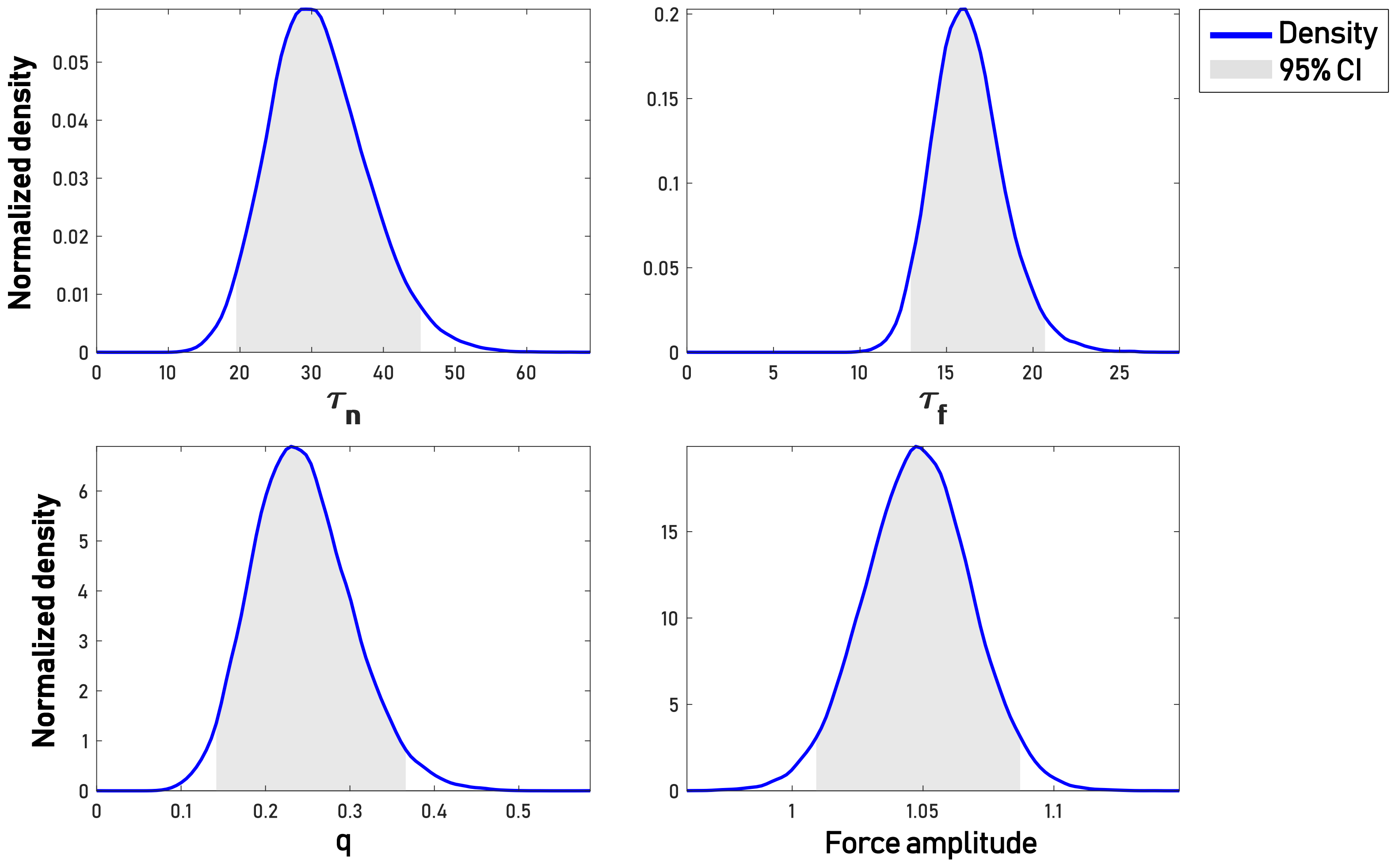
	F_0	τ_n	τ_f	q
Median	1.0481	30.50	16.12	0.240
Mean	1.0480	31.02	16.27	0.244
MAP	1.0472	29.21	16.09	0.230
95% CI	[1.0079, 1.0876]	[19.08, 45.77]	[12.66, 20.76]	[0.141, 0.368]

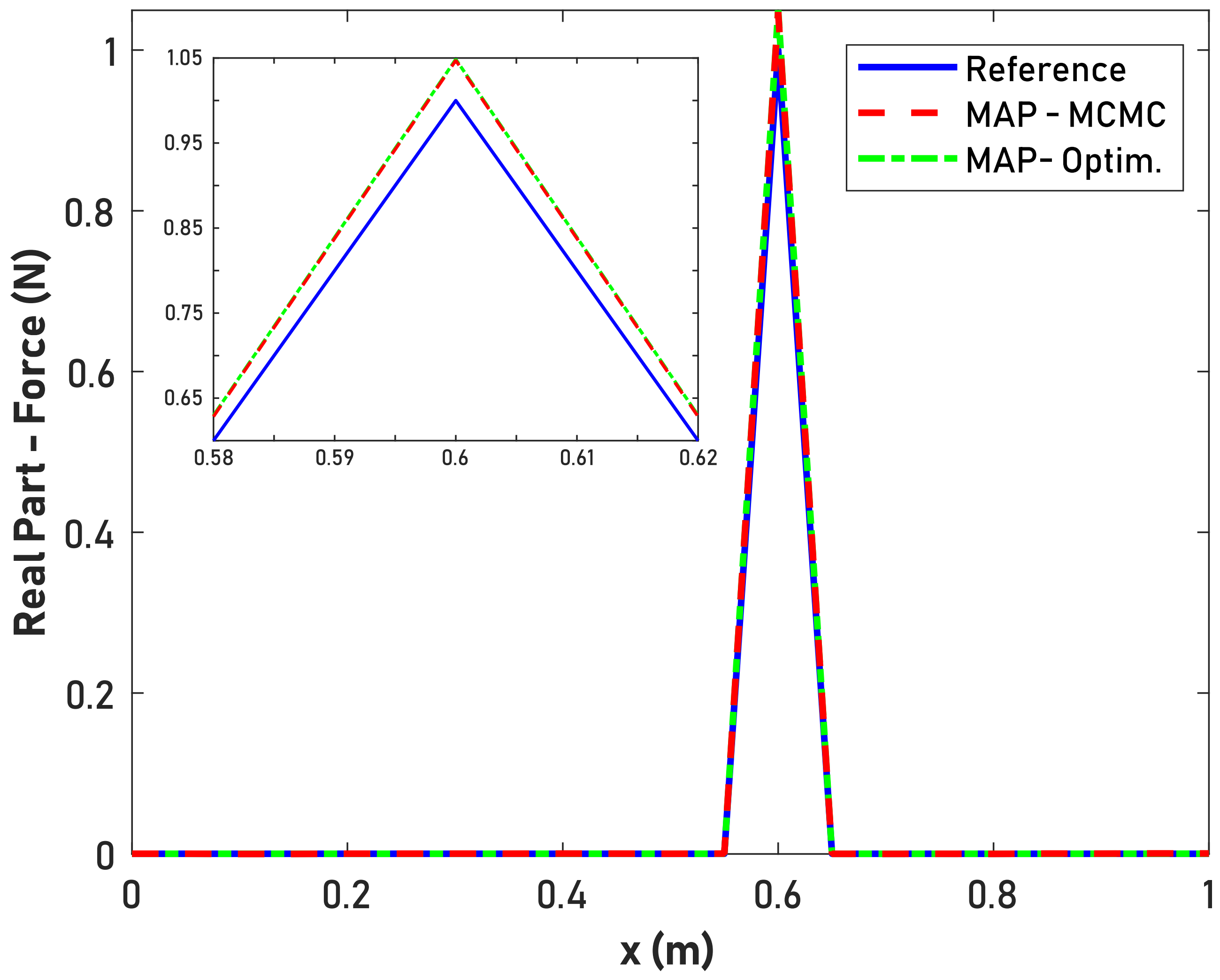
Complete formulation Uncertainty quantification – Application





Complete formulation Uncertainty quantification – Application





Complete formulation Summary

- ✓ Automatic identification of all the parameters
- ✓ Robust identification of the excitation field

Can we do better or at least different ?

Yes, of course !

Outline

Relevant Vector Regression Basics

RVR is a particular Bayesian approach for which the prior probability distribution over $\mathbf{F}$ is such that

$$p(\mathbf{F}) = \prod_{i=1}^{M} \mathcal{N}\left(F_i|0, \tau_{fi}^{-1}\right) \text{ with } \mathcal{N}(F_i|0, \tau_{fi}^{-1}) = \sqrt{\frac{\tau_{fi}}{2\pi}} \exp\left(-\frac{\tau_{fi}}{2}|F_i|^2\right)$$

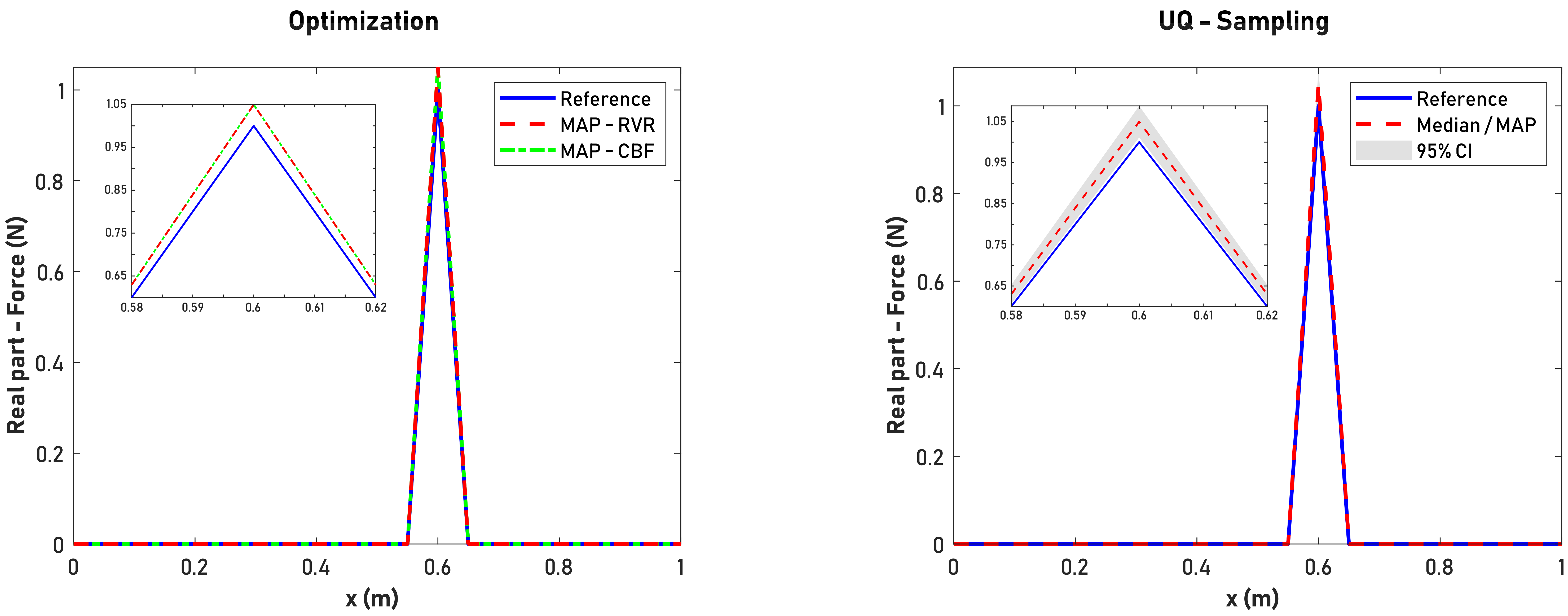
The corresponding Bayesian formulation is expressed as

$$p(\mathbf{F}, \tau_n, \tau_{f_i}|\mathbf{X}) \propto p(\mathbf{X}|\mathbf{F}, \tau_n) \prod_{i=1}^{M} p(F_i|\tau_{f_i})\, p(\tau_{fi}) \text{ with } p(\tau_{fi}) = \mathcal{G}(\tau_{f_i}|\alpha_{fi}, \beta_{fi})$$

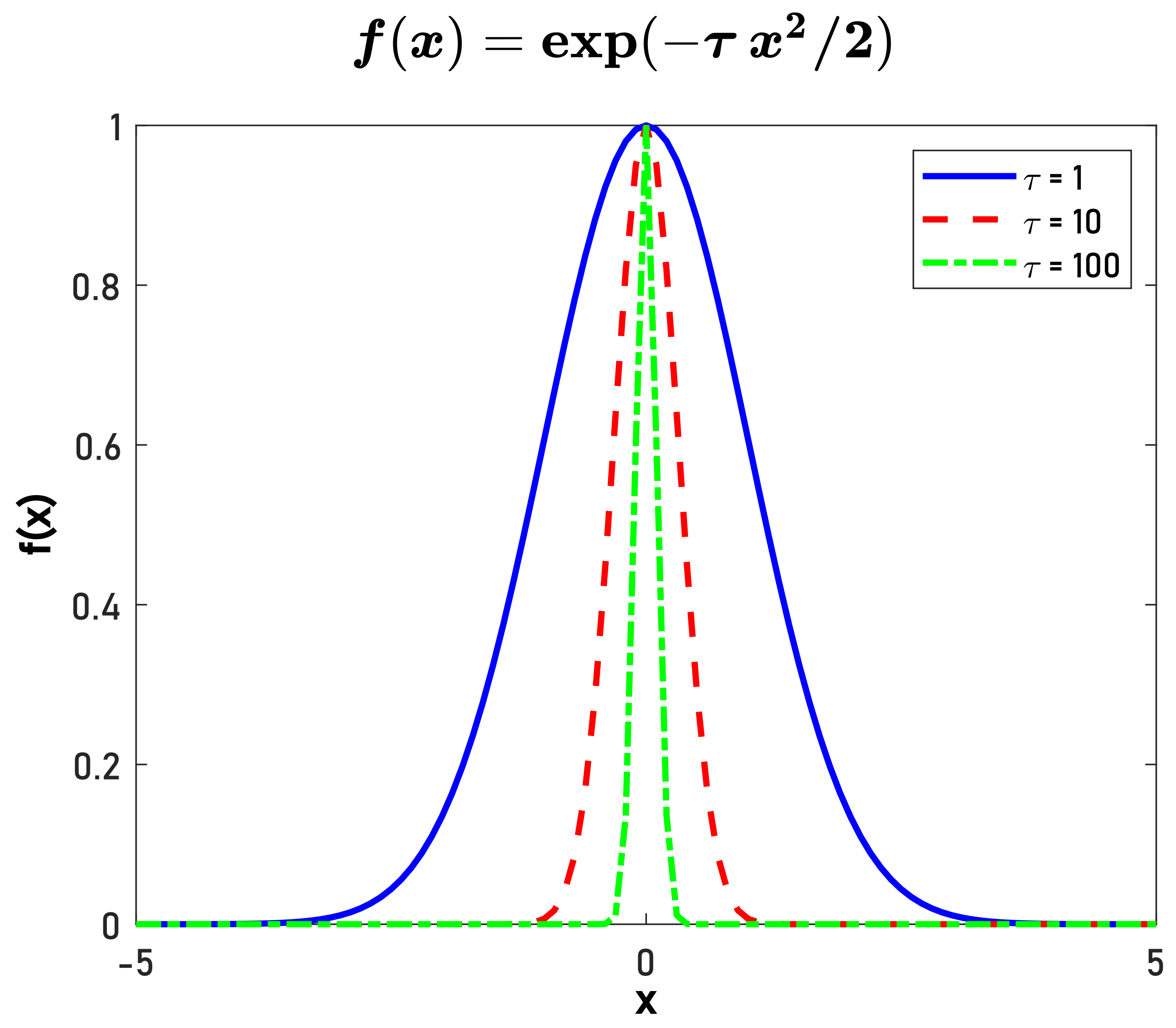
Main features

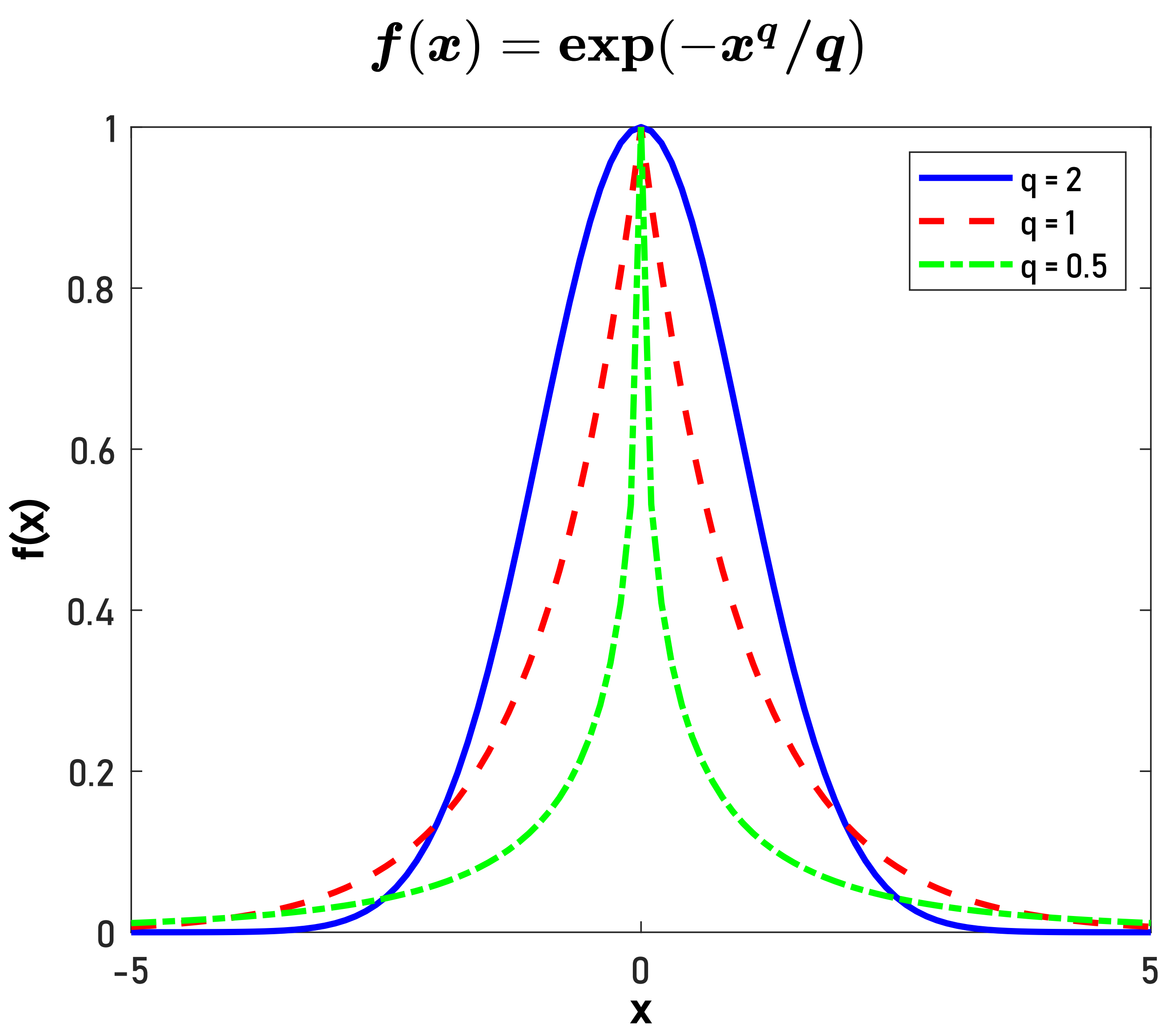
- Implementation of MAP estimation and UQ via Gibbs sampling require minor changes of the algorithms described previously
- More parameters needs to be infered ($M+3$ for CBF and $2M+1$ for RVR)
- Computationally more efficient than CBF

Relevant Vector Regression Application



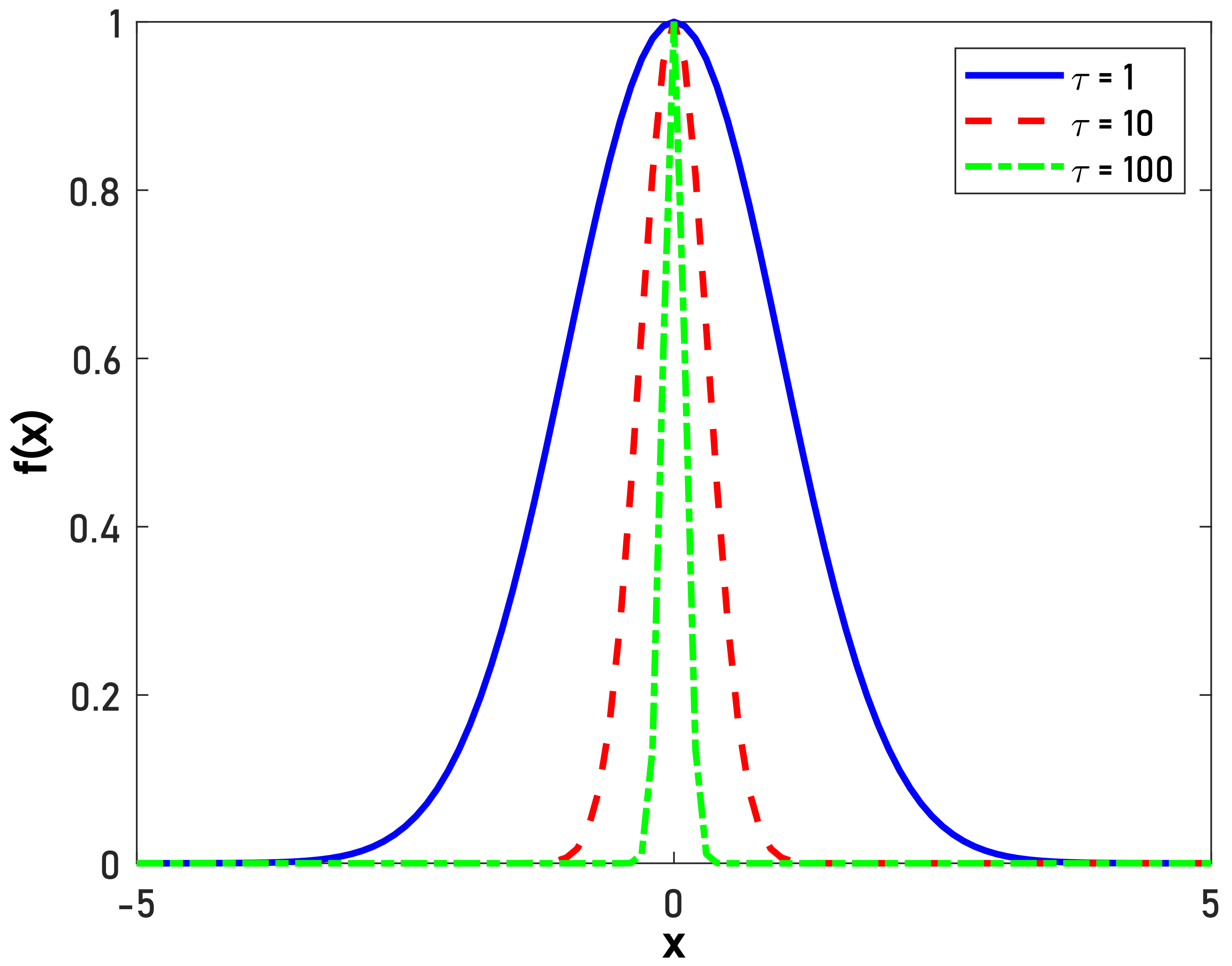
Relevant Vector Regression Why does it work so well ?





Relevant Vector Regression Why does it work so well ?

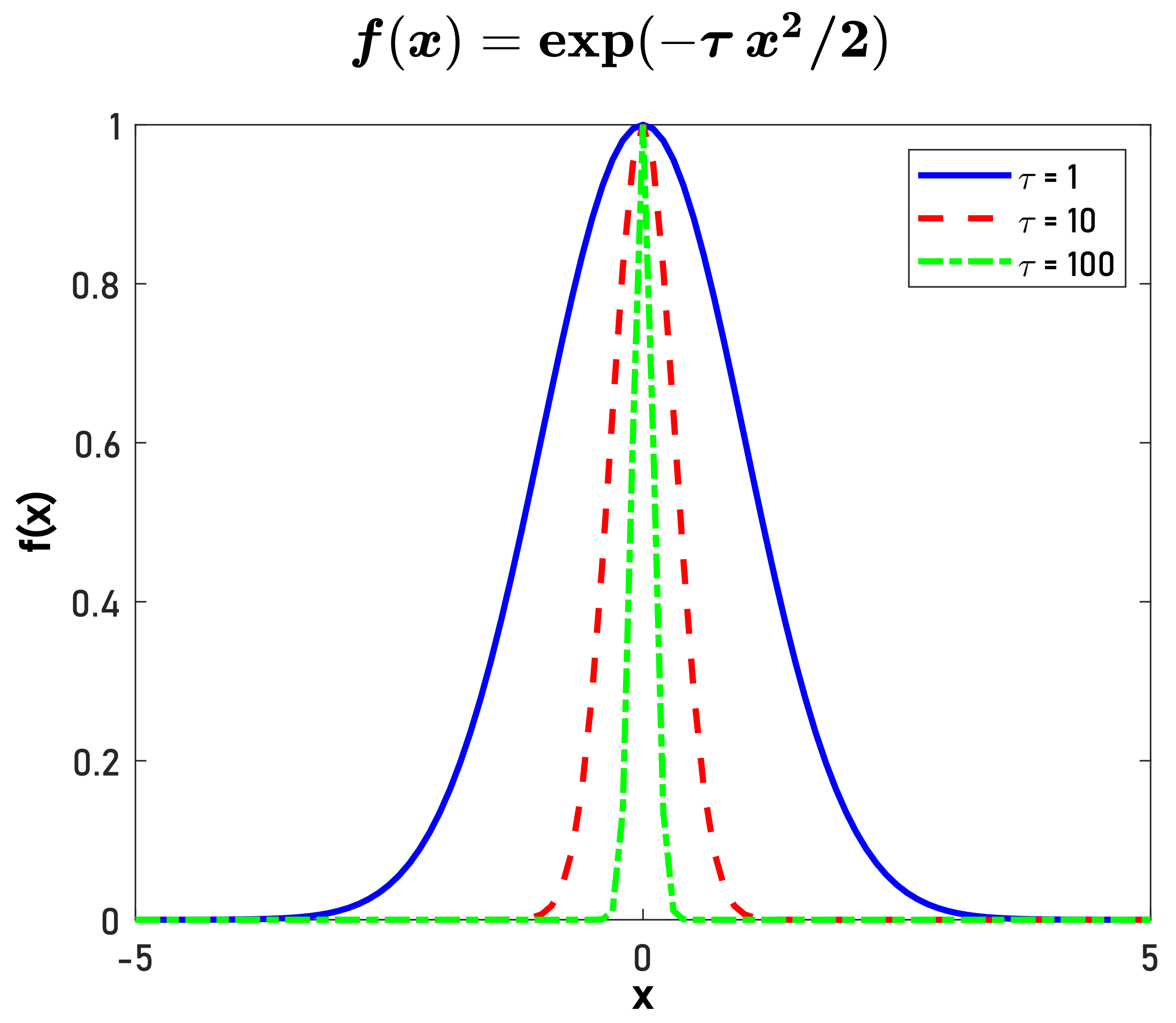
$$f(x) = \exp(-\tau\, x^2/2)$$

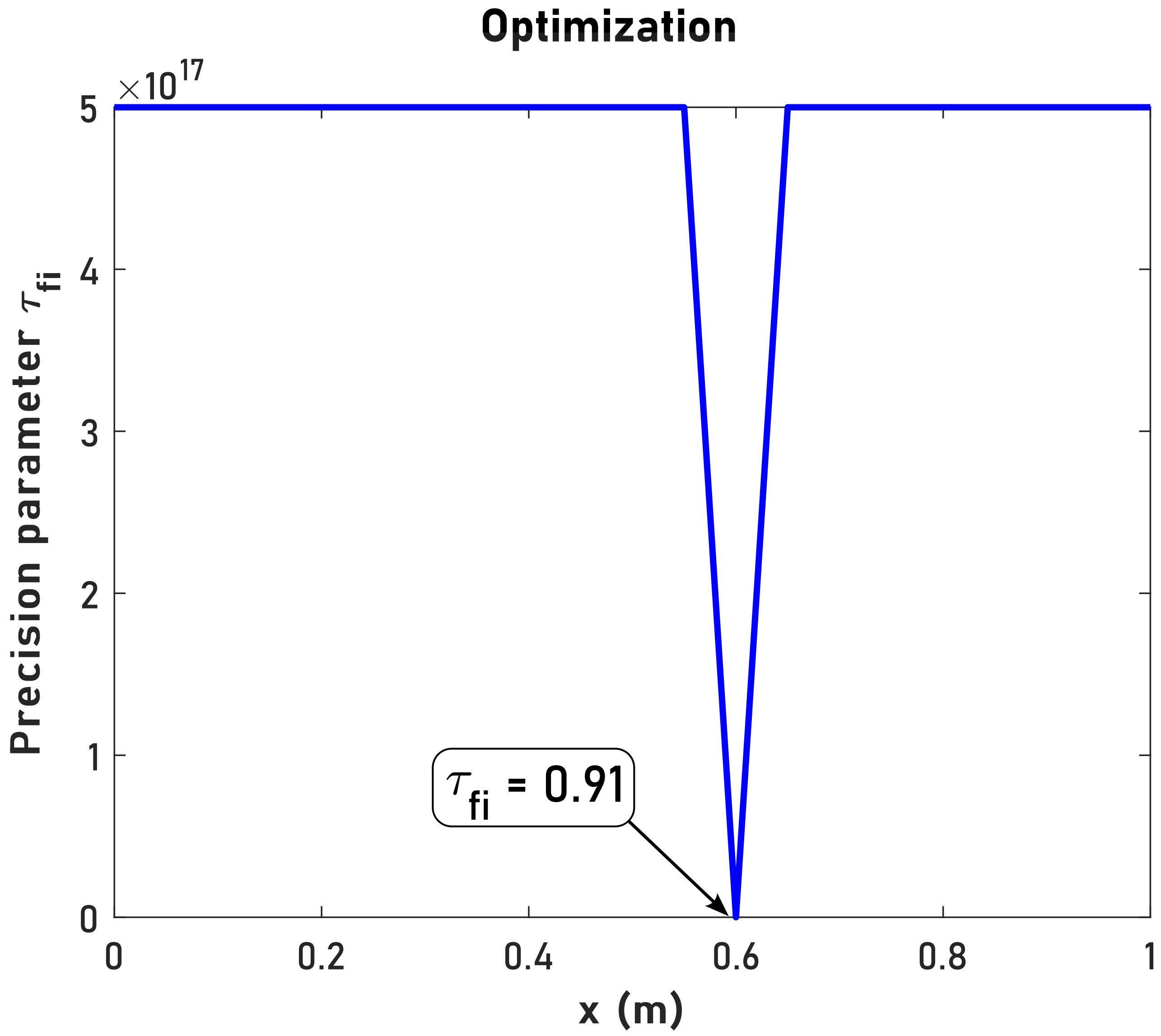


The parameters τ_{fi} and q play a similar role

➡ The larger the value of τ_{fi}, the closer the value of F_i is to 0

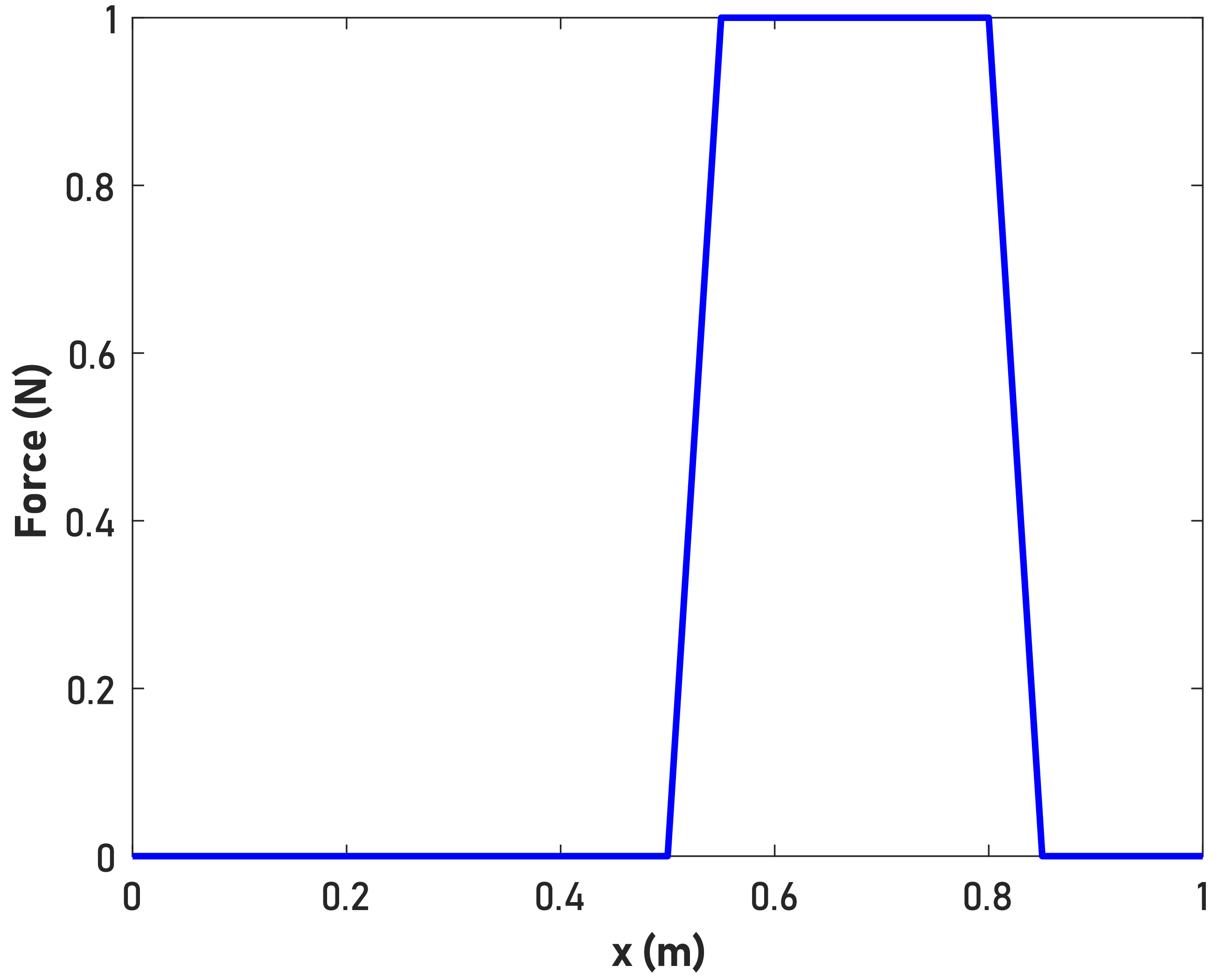
Relevant Vector Regression Why does it work so well ?



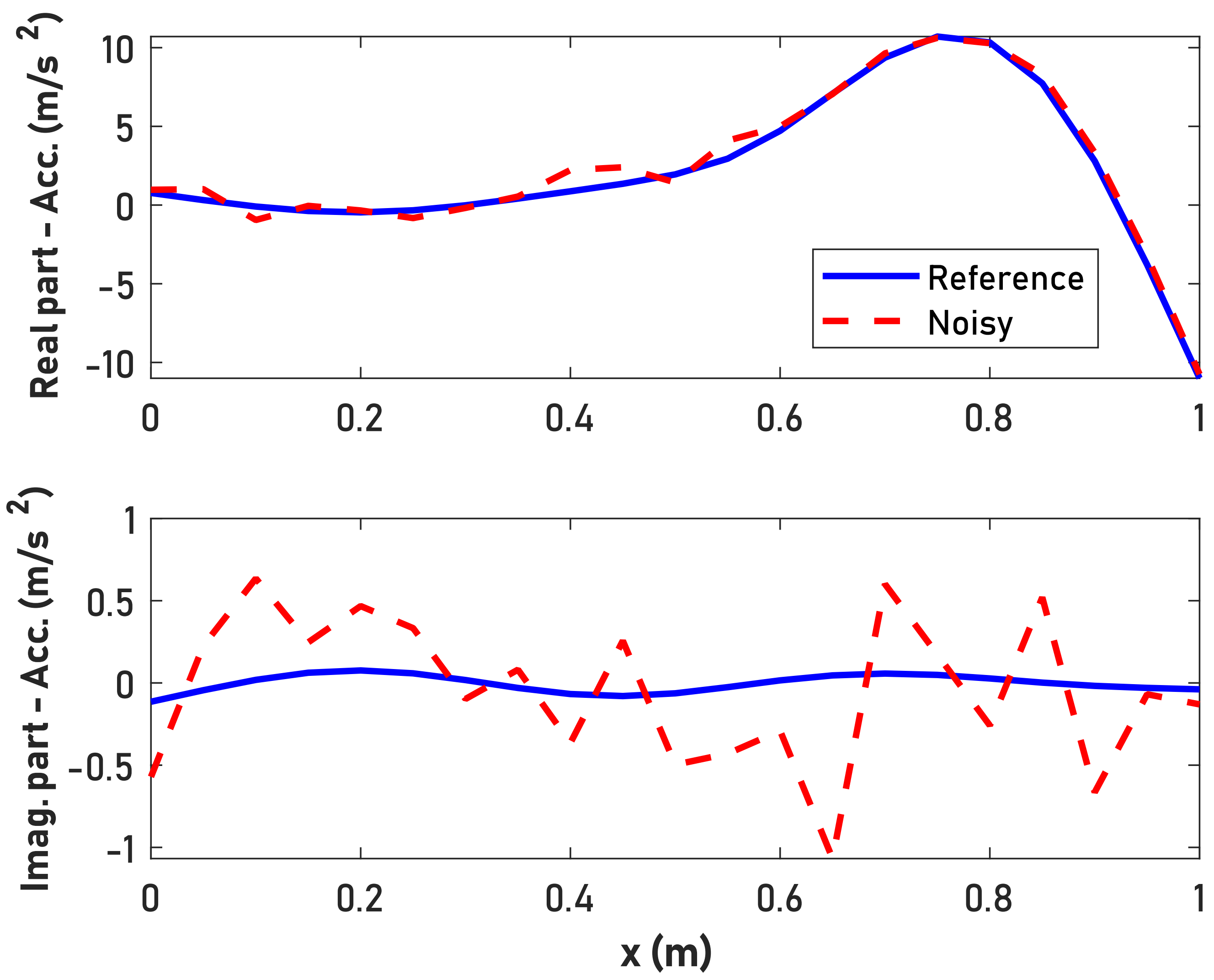


Piecewise constant excitation Objective

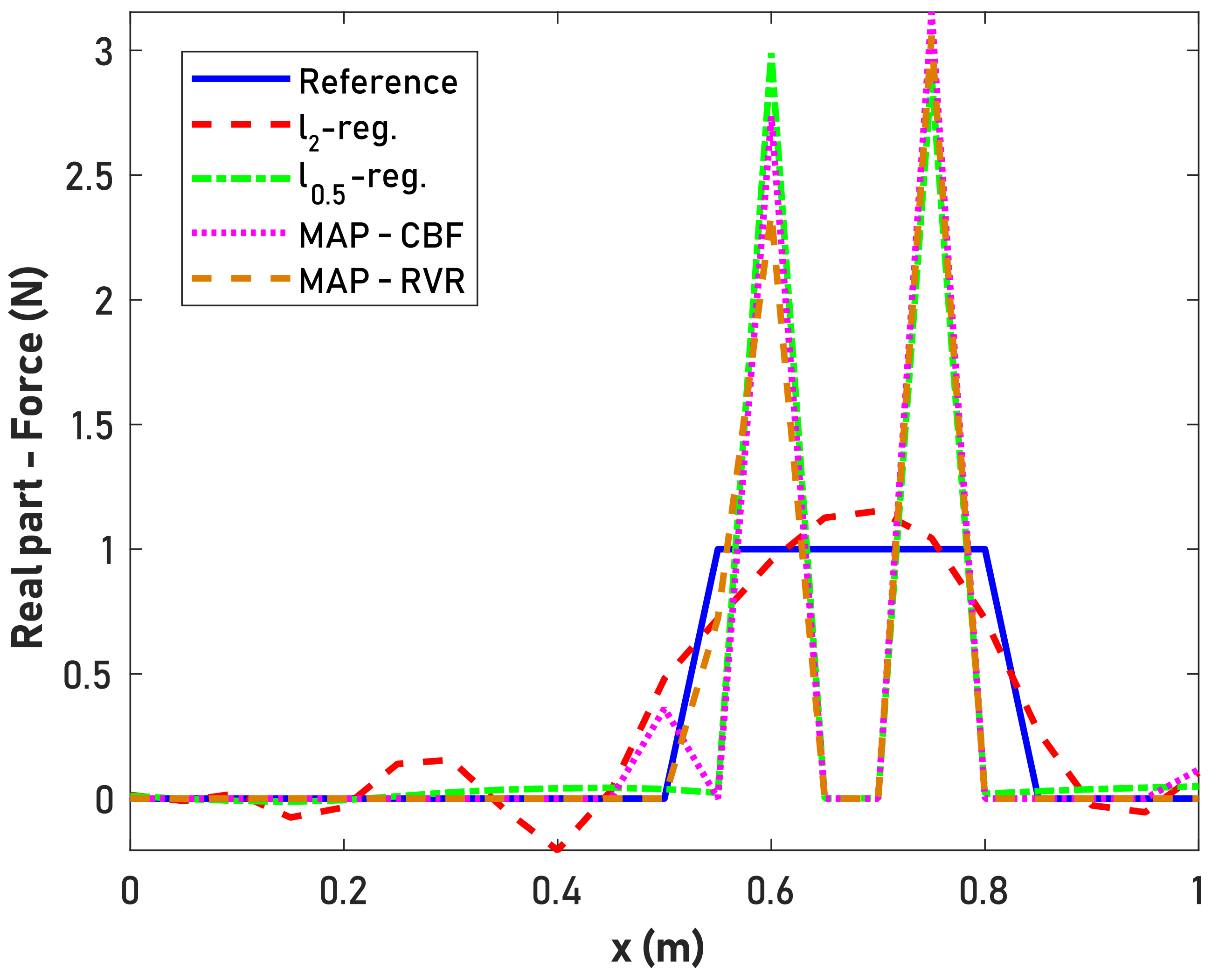
Reconstruct



From



Piecewise constant excitation Naive application

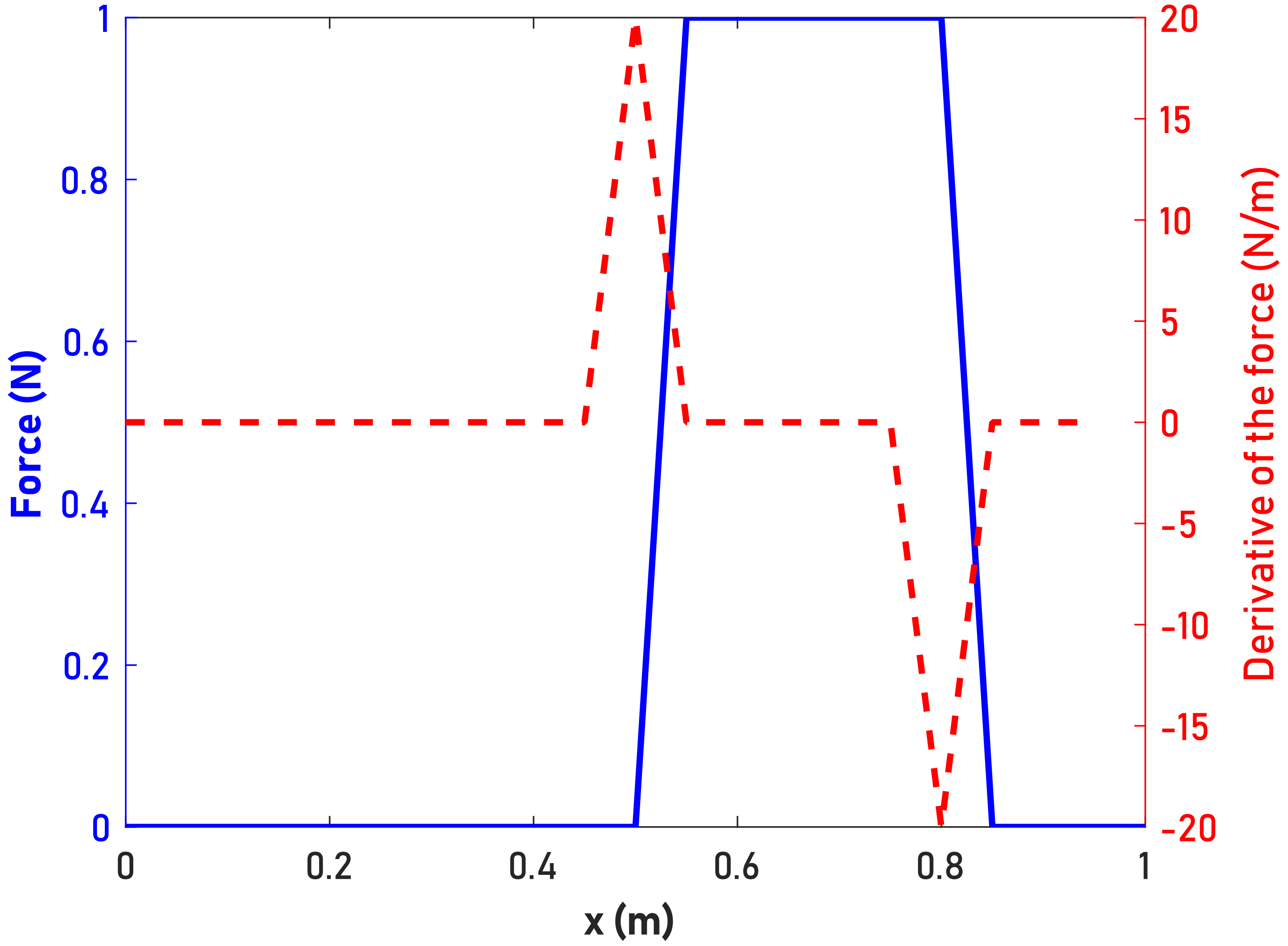


None of the strategies described previously is able to properly reconstruct the excitation field !

What to do ?

Promote piecewise constant solution !

Piecewise constant excitation Intuition



The first derivative of the excitation field is sparse

➡ Promote the sparsity of $\frac{\partial \mathbf{F}(x)}{\partial x}$

Piecewise constant excitation Implementation

Using the discretized first-order derivative operator $\mathbf{D}$

$$\mathbf{D} = \frac{1}{\Delta x} \begin{pmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \end{pmatrix}_{(M-1)\times M}$$

One has the following prior probability distributions

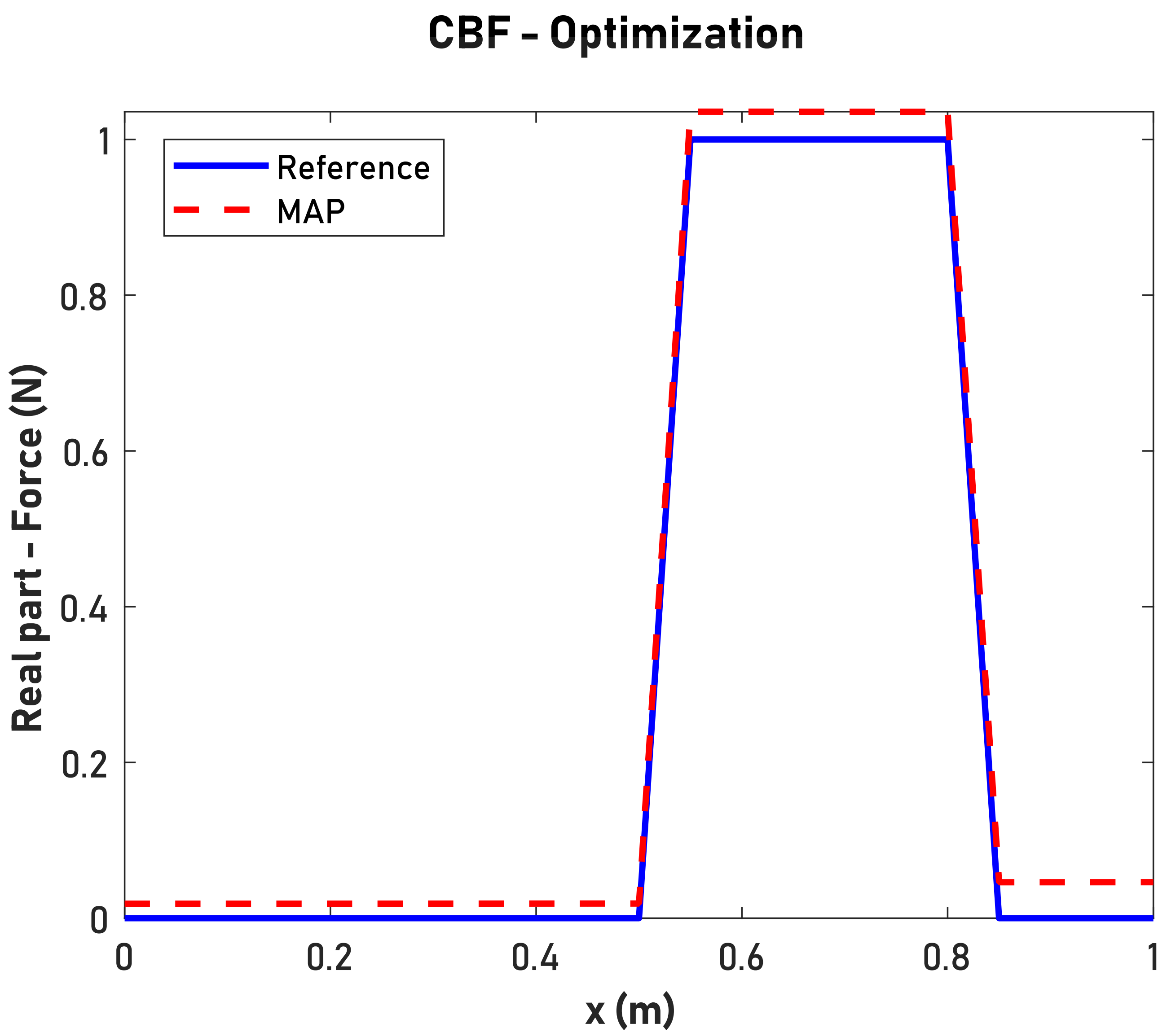
Complete Bayesian formulation

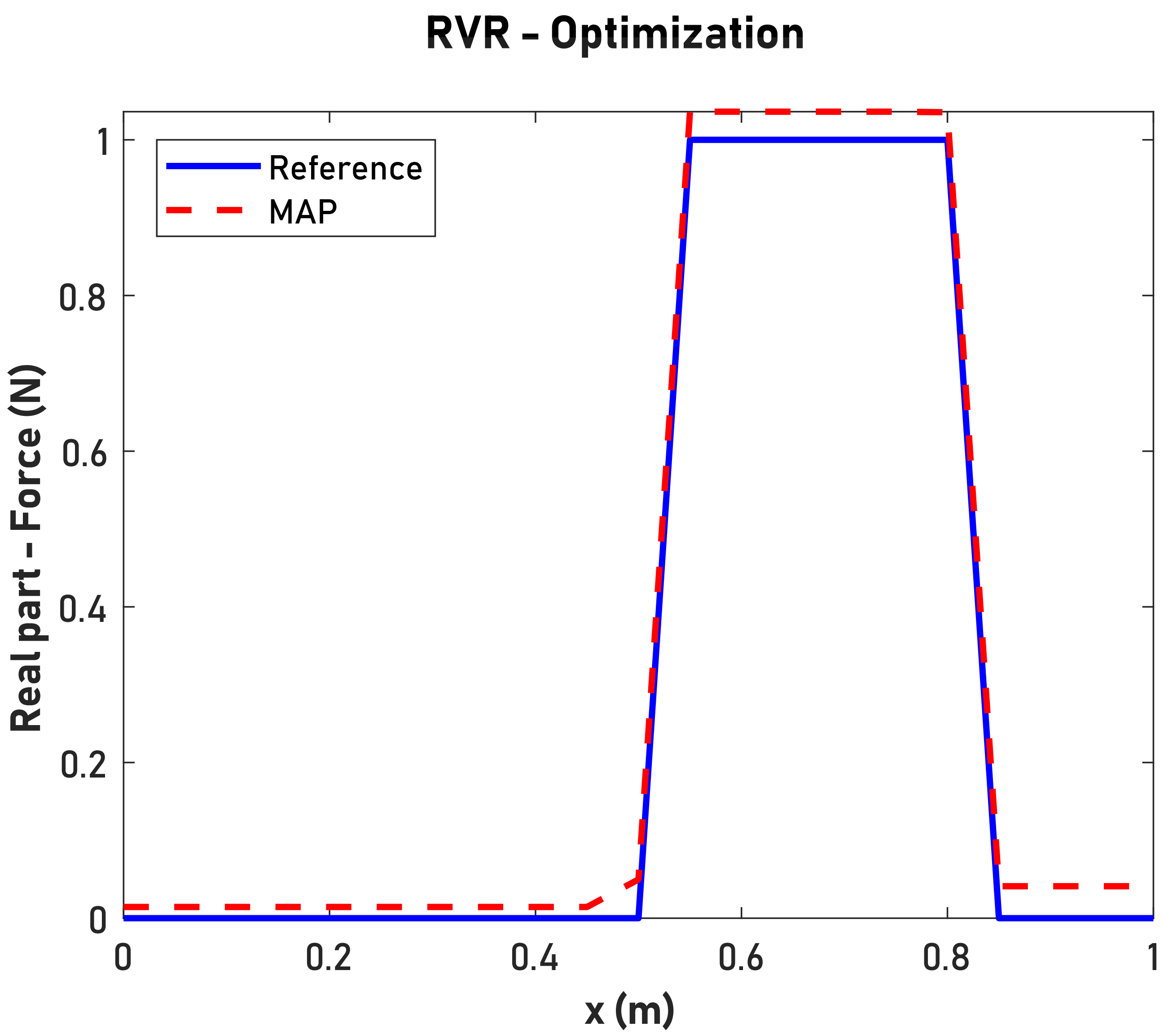
$$p(\mathbf{F}|\tau_f, q) \propto \exp\left(-\tau_f \|\mathbf{DF}\|_q^q\right)$$

Relevant vector regression

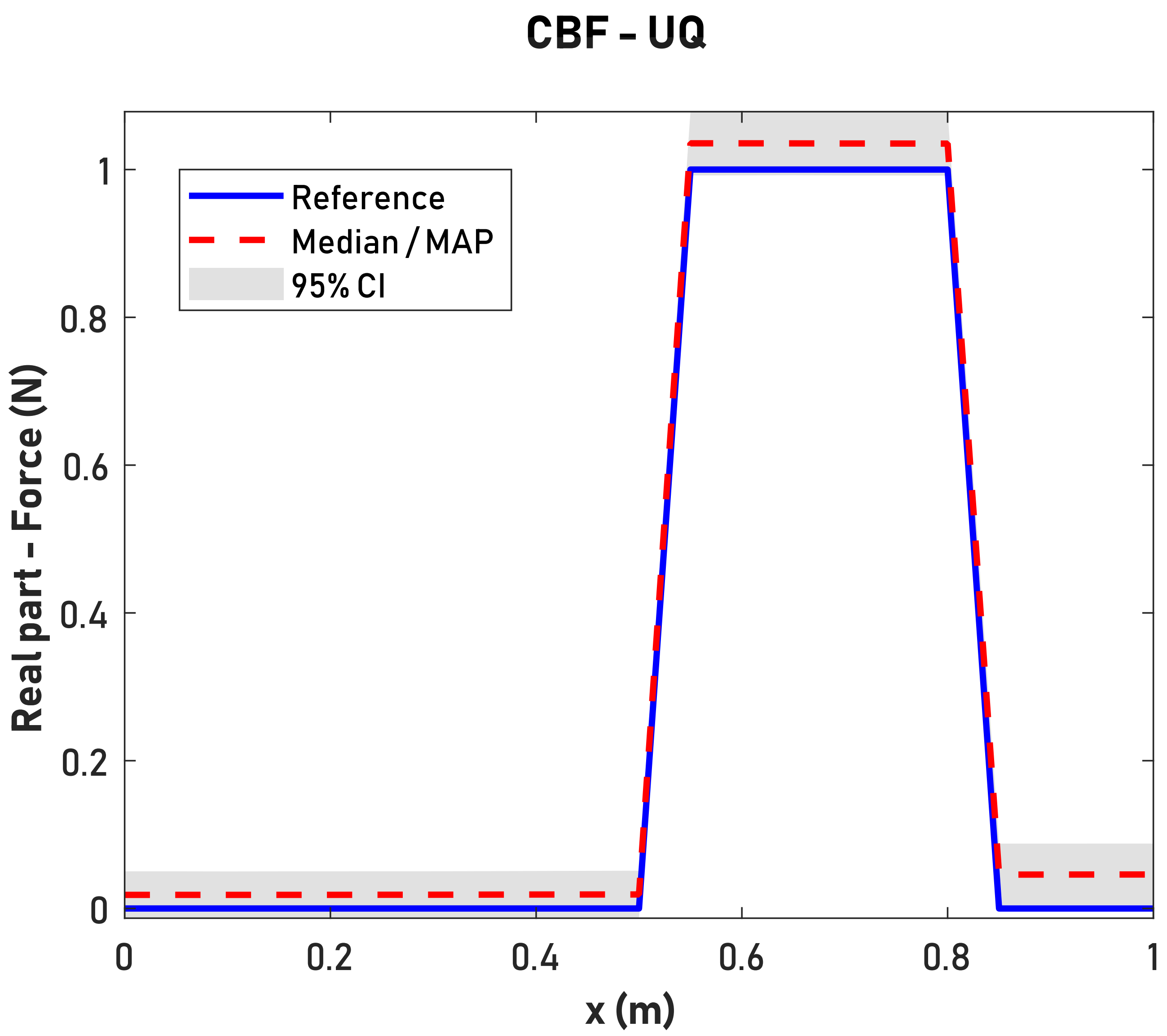
$$p(F_i|\tau_{fj}) \propto \exp\left(-\frac{\tau_{fj}}{2}|D_{ji}F_i|^2\right)$$

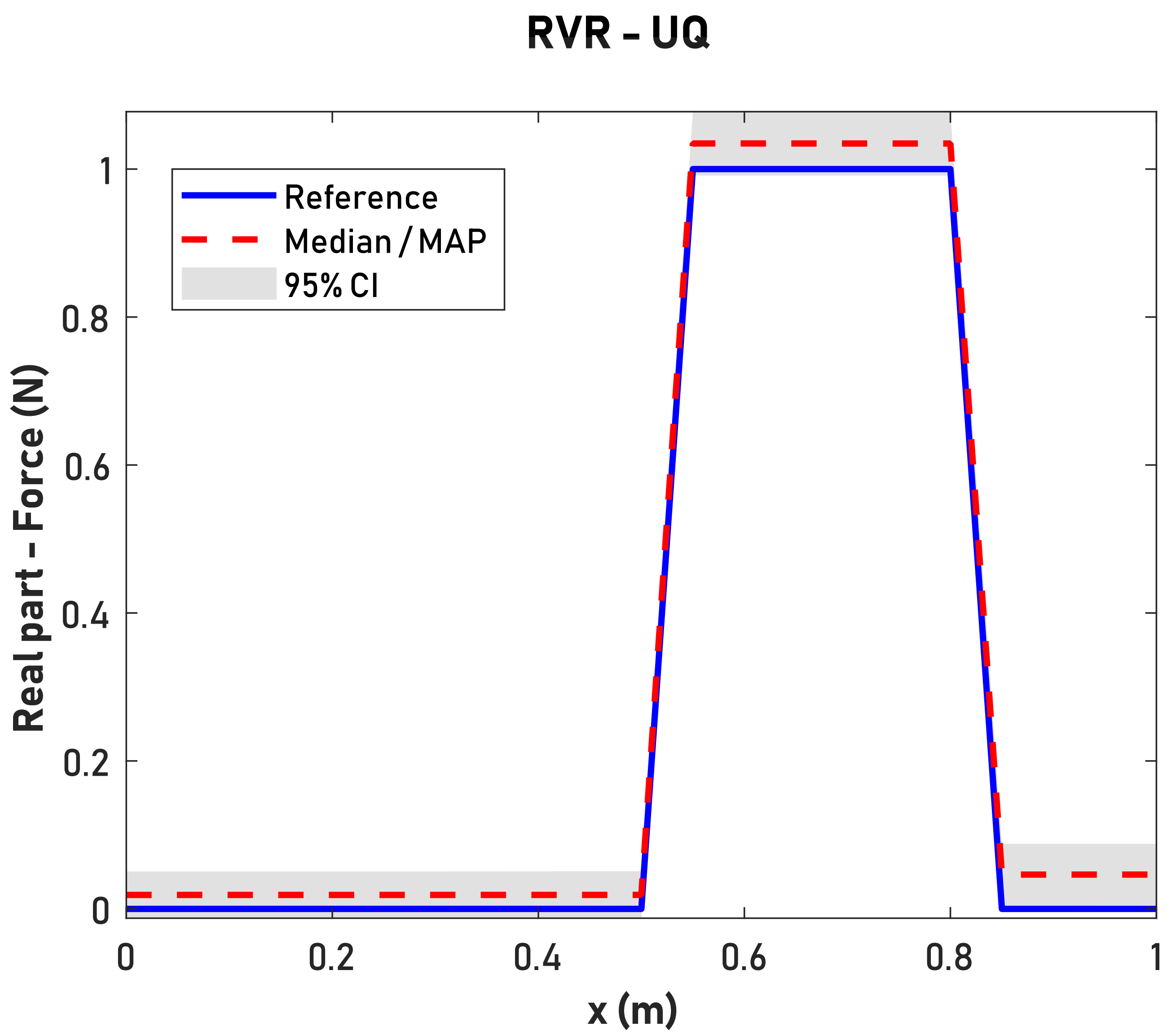
Piecewise constant excitation Application





Piecewise constant excitation Application





Conclusions

- The Bayesian framework provides an efficient and convenient way to combine probabilistic and mechanical data
- It allows exploiting one's prior knowledge of the sources to identify
- It includes an internal mechanism of regularization
- No external procedures are required to infer or optimize all the parameters of the model

Other applications in force reconstruction

- Group regularization – e.g. Identification of external forces and BC on plates
- Mixed-norm regularization – e.g. Identification of space-frequency/time features of excitation sources

Application in other fields

- Image/signal processing (e.g. denoising)
- Acoustics (e.g. fault diagnosis, source reconstruction)
- Material science, Structural mechanics (e.g. parameter estimation, OMA, cracks detection)
- Computer science (e.g. neural networks, bayesian programming)
- Thermal science, Econometrics, Epidemiology, ...

Only the sky is the limit !

Or, maybe, the quantity/quality of available data,
the complexity of the problem,
the computing power/resources, ...

Force reconstruction

A Bayesian perspective

https://github.com/maucejo/MOIRA_Workshop_BFR

Well-posed problem in the sense of Hadamard (1902)

- A solution exist
- The solution is unique
- The solution changes continuously with changes in the data

Back to presentation

Well-posed problem in the sense of Hadamard (1902)

✓A solution exist

✓The solution is unique

✗ The solution changes continuously with changes in the data

➥ The problem considered in this lecture is ill-posed

Back to presentation

ℓ_q-regularization Filter factor analysis at convergence

$$\widehat{\mathbf{F}} = \sum_{i=1}^{21} f_i \frac{\mathbf{v}_i \mathbf{u}_i^H \mathbf{X}}{\sigma_i} \quad \text{with} \quad f_i = \frac{\gamma_i^2}{\gamma_i^2 + \lambda}$$

where γ_i are the singular values of $(\mathbf{H}, \mathbf{L})$ and σ_i are the singular values of $\mathbf{H}$

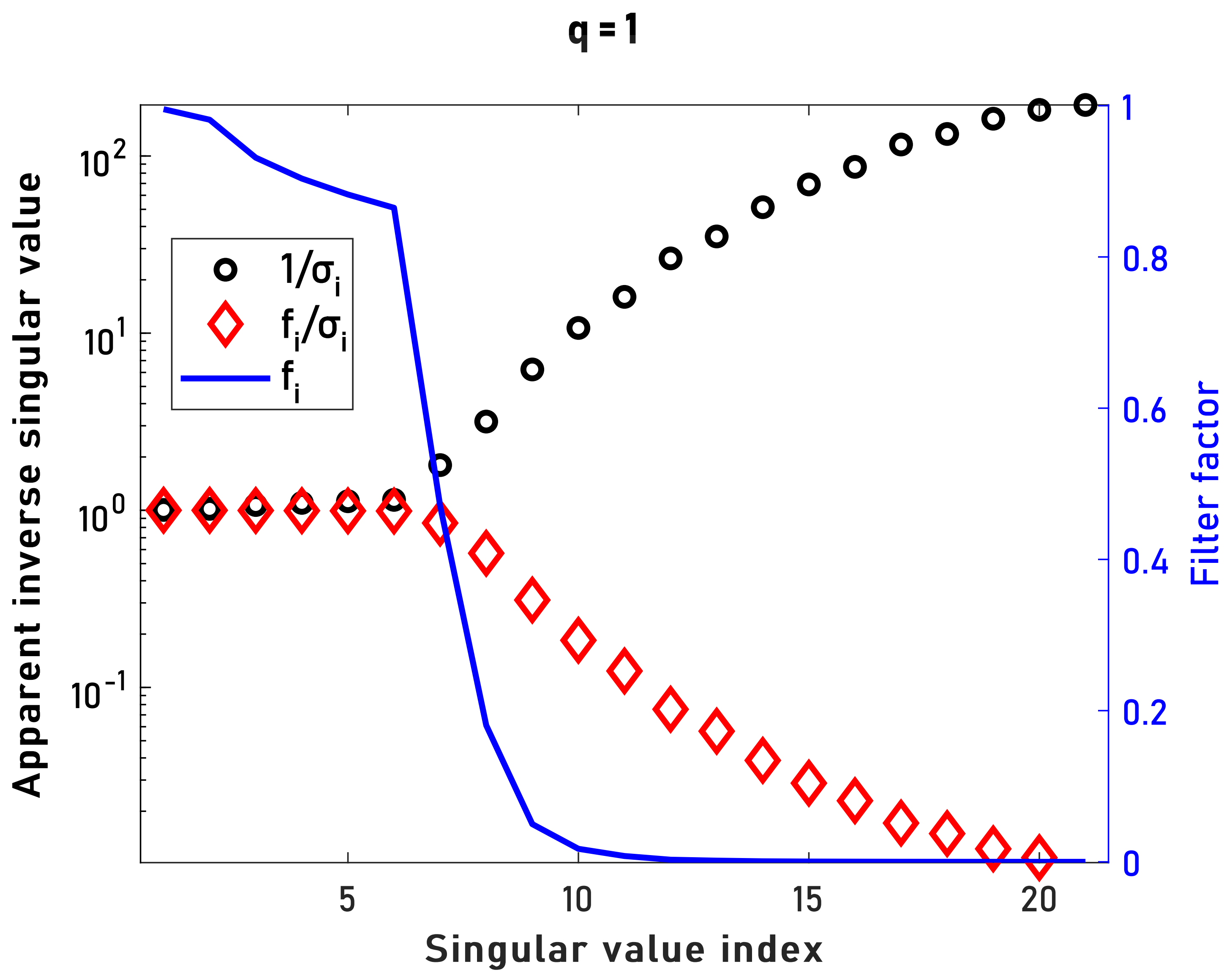
Generalized SVD

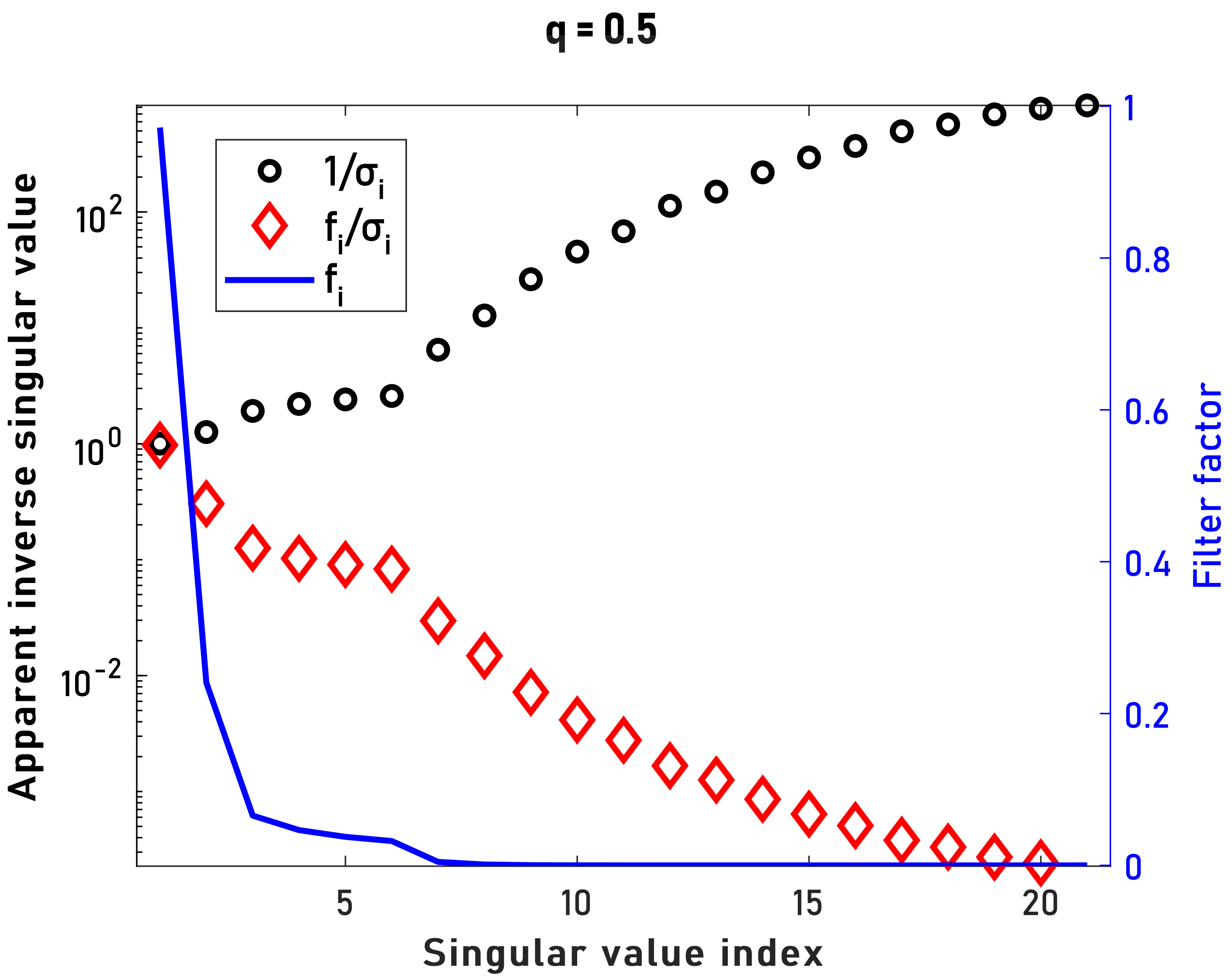
$$\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{Y}^H \quad \text{and} \quad \mathbf{L} = \mathbf{V}\boldsymbol{\Omega}\mathbf{Y}^H$$

Properties of GSVD

$$\boldsymbol{\Sigma}^H\boldsymbol{\Sigma} + \boldsymbol{\Omega}^H\boldsymbol{\Omega} = \mathbf{I} \quad \text{and} \quad \gamma_i = \frac{\sigma_i}{\omega_i}$$

ℓ_q-regularization Filter factor analysis at convergence





Back to presentation

Properties of Gaussian distributions Marginal and Conditional distributions

Let's consider two random vectors, $\mathbf{x}$ and $\mathbf{y}$, such that

$$p(\mathbf{x}) = \mathcal{N}_c(\mathbf{x}|\boldsymbol{\mu}_{\mathbf{x}}, \boldsymbol{\Sigma}_{\mathbf{x}}) \text{ and } p(\mathbf{y}|\mathbf{x}) = \mathcal{N}_c(\mathbf{y}|\mathbf{A}\mathbf{x} + \mathbf{b}, \boldsymbol{\Sigma}_{\mathbf{y}})$$

From these distributions, the marginal and conditional distributions, $p(\mathbf{y})$ and $p(\mathbf{x}|\mathbf{y})$ are given by

$$p(\mathbf{y}) = \mathcal{N}_c\left(\mathbf{y}|\mathbf{A}\boldsymbol{\mu}_{\mathbf{x}} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}_{\mathbf{x}}\mathbf{A}^H + \boldsymbol{\Sigma}_{\mathbf{y}}\right)$$
$$p(\mathbf{x}|\mathbf{y}) = \mathcal{N}_c\left(\mathbf{x}|\boldsymbol{\Sigma}\{\mathbf{A}^H\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}(\mathbf{y} - \mathbf{b}) + \boldsymbol{\Sigma}_{\mathbf{x}}^{-1}\boldsymbol{\mu}_{\mathbf{x}}\}, \boldsymbol{\Sigma}\right)$$

with $\boldsymbol{\Sigma} = \left(\mathbf{A}^H\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{A} + \boldsymbol{\Sigma}_{\mathbf{x}}^{-1}\right)^{-1}$

Drawing samples from multivariate Gaussian distribution

Let's consider a random Gaussian vector $\mathbf{x}$ such that

$$p(\mathbf{x}) = \mathcal{N}_c(\mathbf{x}|\boldsymbol{\mu}_{\mathbf{x}}, \boldsymbol{\Sigma}_{\mathbf{x}})$$

By assuming that $\boldsymbol{\Sigma}_{\mathbf{x}} = \mathbf{S}\mathbf{S}^H$, one has

$$\begin{aligned} \exp\left[-(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{x}})^H \boldsymbol{\Sigma}_{\mathbf{x}}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{x}})\right] &= \exp\left[-\{\mathbf{S}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{x}})\}^H\{\mathbf{S}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{x}})\}\right] \\ &= \exp\left[-\mathbf{z}^H\mathbf{z}\right] \end{aligned}$$

where $\mathbf{z} = \mathbf{S}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{x}})$ and $\mathbf{z} \sim \mathcal{N}_c(\mathbf{z}|\mathbf{0}, \mathbf{I})$

Consequently, to draw samples from a multivariate Gaussian distribution with mean $\boldsymbol{\mu}_{\mathbf{x}}$ and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{x}}$, it is enough to compute

$$\mathbf{x}^{(k)} = \boldsymbol{\mu}_{\mathbf{x}} + \mathbf{S}\,\mathbf{z}^{(k)} \quad \text{with} \quad \mathbf{S}\mathbf{S}^H = \boldsymbol{\Sigma}_{\mathbf{x}} \quad \text{and} \quad \mathbf{z}^{(k)} \sim \mathcal{N}_c(\mathbf{z}^{(k)}|\mathbf{0}, \mathbf{I})$$

Back to presentation

Calculation of τ_n and τ_f

By using the Bayes' rule, the conditional distribution $p(\tau_n, \tau_f | \mathbf{X})$ is expressed as

$$p(\tau_n, \tau_f | \mathbf{X}) \propto p(\mathbf{X} | \tau_n, \tau_f)\, p(\tau_n)\, p(\tau_f)$$

Assuming that $p(\tau_n) = p(\tau_f) \propto 1$, one has

$$p(\tau_n, \tau_f | \mathbf{X}) \propto p(\mathbf{X} | \tau_n, \tau_f) = \int_{\mathbf{F}} p(\mathbf{X} | \mathbf{F}, \tau_n)\, p(\mathbf{F} | \mathbf{W}, \tau_f) d\mathbf{F}$$

Using the fact that all the conditional distributions are Gaussian, one establishes that

$$p(\tau_n, \tau_f | \mathbf{X}) \propto \mathcal{N}_c(\mathbf{X} | \mathbf{0}, \mathbf{H}\mathbf{W}^{-1}\mathbf{H}^H / \tau_f + \mathbf{I}/\tau_n)$$

The MAP estimate is found by solving

$$(\widehat{\tau}_n, \widehat{\tau}_f) = \underset{(\tau_n, \tau_f)}{\operatorname{argmin}} \; -\log[p(\tau_n, \tau_f | \mathbf{X})]$$

By noting $\lambda = \tau_n / \tau_f$, it comes

$$(\widehat{\tau}_n, \widehat{\tau}_f) = \underset{(\tau_n, \tau_f)}{\operatorname{argmin}} \; \tau_f \, \mathbf{X}^H \left(\mathbf{H}\mathbf{W}^{-1}\mathbf{H}^H + \lambda \mathbf{I}\right)^{-1} \mathbf{X} - N \log \tau_f + \log \left|\mathbf{H}\mathbf{W}^{-1}\mathbf{H}^H + \lambda \mathbf{I}\right|$$

By applying the first-order optimality condition, one finds

$$\widehat{\tau}_f = \frac{N}{\mathbf{X}^H \left(\mathbf{H}\mathbf{W}^{-1}\mathbf{H}^H + \lambda \mathbf{I}\right)^{-1} \mathbf{X}}$$

Back to presentation