Sparse Bayesian Learning for Rational Polynomial Chaos Expansion and Application in Structural Dynamics

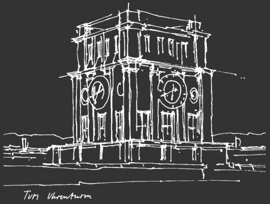
UQSay Seminar

Felix Schneider[1] Iason Papaioannou[2] Gerhard Müller[1]

[1]Chair of Structural Mechanics
Technical University of Munich

[2]Engineering Risk Analysis Group
Technical University of Munich

11. May 2023



Summary

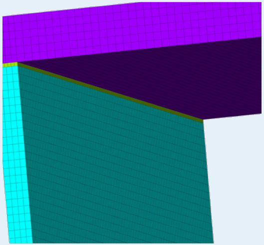
Motivation & Problem Statement

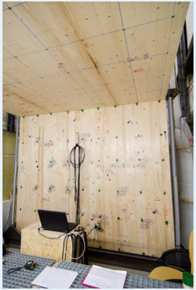
Motivating Example

Investigation of energy flow in cross-laminated timber structures

Motivating Example

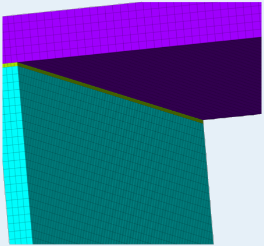
Investigation of energy flow in cross-laminated timber structures

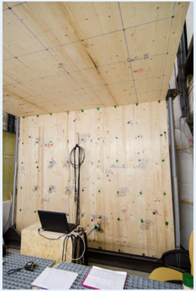




Motivating Example

Investigation of energy flow in cross-laminated timber structures





→ Goal: Update model parameters using available measurement data.

Structural Dynamics 101

We start with a space-discretized, linear, time invariant structural system with equation of motion

$$\mathbf{K}\mathbf{u}(t) + \mathbf{C}\dot{\mathbf{u}}(t) + \mathbf{M}\ddot{\mathbf{u}}(t) = \mathbf{f}(t) \tag{1}$$

where $\mathbf{K}$, $\mathbf{C}$ and $\mathbf{M}$ are the stiffness, mass and damping matrix and $\mathbf{u}(t)$ and $\mathbf{f}(t)$ denote the vector of degrees of freedom and the load vector. We apply the Fourier transform operator to the equation of motion and obtain

$$\mathbf{K}\tilde{\mathbf{u}}(\omega) + \mathrm{i}\omega\mathbf{C}\tilde{\mathbf{u}}(\omega) - \omega^2\mathbf{M}\tilde{\mathbf{u}}(\omega) = \tilde{\mathbf{f}}(\omega) \tag{2}$$

The solution to this equation is

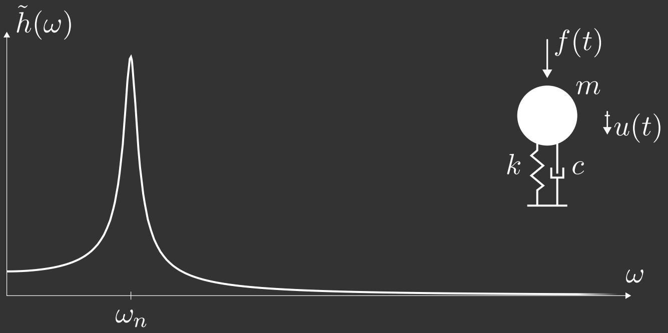
$$\tilde{\mathbf{u}}(\omega) = \underbrace{\left(\mathbf{K} + \mathrm{i}\omega\mathbf{C} - \omega^2\mathbf{M}\right)^{-1}}_{\tilde{\mathbf{H}}(\omega)}\tilde{\mathbf{f}}(\omega) \tag{3}$$

Structural Dynamics 101

The frequency response for a single degree of freedom system reads

$$\tilde{h}(\omega) = \frac{1}{k - \omega^2 m + \mathrm{i}\omega c} \tag{4}$$

The rational dependency on the frequency can be clearly observed:



Problem Statement

Given a set of system observations and corresponding forward models

$$\mathcal{D}_{\mathcal{O}} = \{y_{\mathcal{O},i} | i = 1, \ldots, n_{\mathcal{O}}\}, \tag{5}$$

$$\mathcal{M} = \{\mathcal{M}_i(\mathbf{x}) | i = 1, \ldots, n_{\mathcal{O}}\}, \tag{6}$$

find an updated set of model inputs $\mathbf{x} \in \mathbf{R}^d$ that best match the observations in some sense.

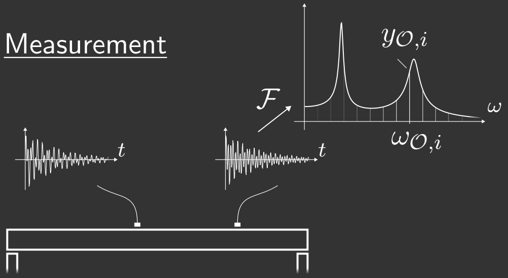
Problem Statement

Given a set of system observations and corresponding forward models

$$\mathcal{D}_{\mathcal{O}} = \{y_{\mathcal{O},i} | i = 1, \ldots, n_{\mathcal{O}}\}, \tag{5}$$

$$\mathcal{M} = \{\mathcal{M}_i(\mathbf{x}) | i = 1, \ldots, n_{\mathcal{O}}\}, \tag{6}$$

find an updated set of model inputs $\mathbf{x} \in \mathbf{R}^d$ that best match the observations in some sense.



Problem Statement

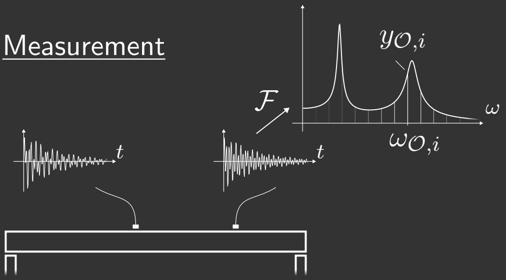
Given a set of system observations and corresponding forward models

$$\mathcal{D}_{\mathcal{O}} = \{y_{\mathcal{O},i} | i = 1, \ldots, n_{\mathcal{O}}\}, \tag{5}$$

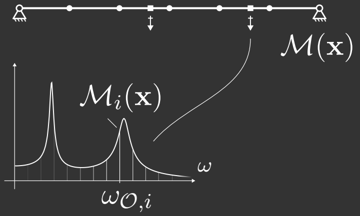
$$\mathcal{M} = \{\mathcal{M}_i(\mathbf{x}) \, | i = 1, \ldots, n_{\mathcal{O}}\}, \tag{6}$$

find an updated set of model inputs $\mathbf{x} \in \mathbf{R}^d$ that best match the observations in some sense.

Measurement



Simulation



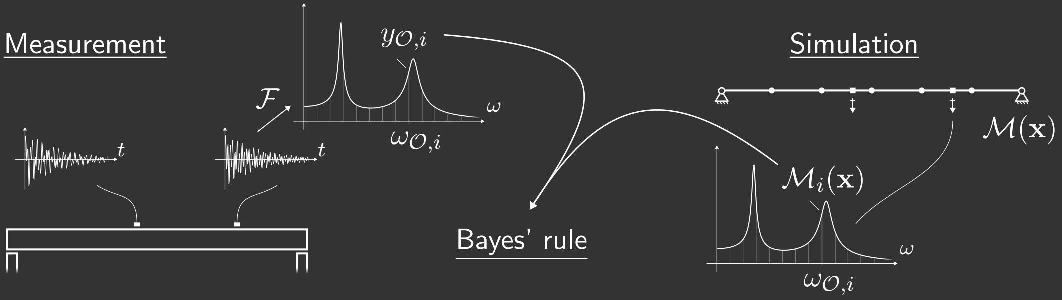
Problem Statement

Given a set of system observations and corresponding forward models

$$\mathcal{D}_{\mathcal{O}} = \{y_{\mathcal{O},i} | i = 1, \ldots, n_{\mathcal{O}}\}, \tag{5}$$

$$\mathcal{M} = \{\mathcal{M}_i(\mathbf{x}) | i = 1, \ldots, n_{\mathcal{O}}\}, \tag{6}$$

find an updated set of model inputs $\mathbf{x} \in \mathbf{R}^d$ that best match the observations in some sense.



Problem Statement

Bayes' rule states that

$$f(\mathbf{x}|\mathcal{D}_{\mathcal{O}}) = c_E^{-1} L(\mathbf{x}|\mathcal{D}_{\mathcal{O}}) f(\mathbf{x}) \,. \tag{7}$$

The posterior distribution $f(\mathbf{X}|\mathcal{D}_{\mathcal{O}})$ is seldom available in closed-form.

→ Infer posterior distribution through sampling or approximation

→ Resort to point estimates, such as the mode, i.e., the maximum-a-posteriori (MAP) estimate

Inference or estimation thus require repeated evaluation of the likelihood function, which require an evaluation of each of the forward models $\mathcal{M}_i(\mathbf{x})$.

Solution

→ Replace forward models by surrogates.

Accelerate Estimation and Sampling

To accelerate estimation we aim to surrogate the frequency response functions:

→ Application of polynomial chaos expansion (PCE) to frequency response function (FRF) models in [2, 3].

→ Due to slow convergence rates in PCE, PCE-based rational approximations in [4, 7, 8].

→ Stochastic frequency transformation and sparse PCE representation of the FRFs in [10].

→ Multi-output Gaussian process model for uncertainty quantification of FRF models in [5].

Accelerate Estimation and Sampling

To accelerate estimation we aim to surrogate the frequency response functions:

→ Application of polynomial chaos expansion (PCE) to frequency response function (FRF) models in [2, 3].

→ Due to slow convergence rates in PCE, PCE-based rational approximations in [4, 7, 8].

→ Stochastic frequency transformation and sparse PCE representation of the FRFs in [10].

→ Multi-output Gaussian process model for uncertainty quantification of FRF models in [5].

Polynomial Chaos Expansion based Rational Approximations

Rational Approximation

Consider a numerical model $\mathcal{M}(\mathbf{X})$ with outcome space $\mathbb{C}$. $\mathbf{X}$ is a random vector with outcome space $\mathbb{R}^d$ and given joint probability density function and models a set of uncertain parameters that represent the model input.

Definition

We define the rational approximation $\mathcal{R}(\mathbf{X};\mathbf{p},\mathbf{q})$ as

$$\mathcal{R}(\mathbf{X};\mathbf{p},\mathbf{q}) = \frac{P(\mathbf{X};\mathbf{p})}{Q(\mathbf{X};\mathbf{q})} = \frac{\sum_{i=0}^{n_p-1} p_i \psi_{p,i}(\mathbf{X})}{\sum_{i=0}^{n_q-1} q_i \psi_{q,i}(\mathbf{X})}$$

Rational Approximation

Definition

We define the rational approximation $\mathcal{R}(\mathbf{X};\mathbf{p},\mathbf{q})$ as

$$\mathcal{R}(\mathbf{X};\mathbf{p},\mathbf{q}) = \frac{P(\mathbf{X};\mathbf{p})}{Q(\mathbf{X};\mathbf{q})} = \frac{\sum_{i=0}^{n_p-1} p_i \psi_{p,i}(\mathbf{X})}{\sum_{i=0}^{n_q-1} q_i \psi_{q,i}(\mathbf{X})}$$

- $P(\mathbf{X};\mathbf{p})$ and $Q(\mathbf{X};\mathbf{p})$ are truncated polynomial chaos expansions with n_p and n_q terms, respectively.
- $p = \left[p_0, \ldots, p_{n_p-1}\right] \in \mathbb{C}^{n_p}$ and $q = \left[q_0, \ldots, q_{n_q-1}\right] \in \mathbb{C}^{n_q}$ denote the vectors of coefficients of the numerator and denominator polynomial, respectively.
- $\psi_{p,i}$ and $\psi_{q,i}$ denote some multivariate orthonormal polynomials built from the Wiener-Askey family.

Estimation and Fitting of the Rational Approximation

The task of fitting the surrogate model can be cast into a regression problem. Here, we apply

- Ordinary least squares regression
- Bayesian regression with sparsity inducing priors

The regression approach is based on measure of misfit. Two natural measures for the rational polynomial chaos expansion are

$$\varepsilon = \mathcal{M}(\mathbf{X}) - \frac{P(\mathbf{X}, \mathbf{p})}{Q(\mathbf{X}, \mathbf{q})} \tag{8}$$

$$\tilde{\varepsilon} = \varepsilon Q(\mathbf{X}, \mathbf{q}) = \mathcal{M}(\mathbf{X}) Q(\mathbf{X}, \mathbf{q}) - P(\mathbf{X}, \mathbf{p}) \tag{9}$$

Estimation and Fitting of the Rational Approximation

The task of fitting the surrogate model can be cast into a regression problem. Here, we apply

- Ordinary least squares regression
- Bayesian regression with sparsity inducing priors

The regression approach is based on measure of misfit. Two natural measures for the rational polynomial chaos expansion are

$$\varepsilon = \mathcal{M}(\mathbf{X}) - \frac{P(\mathbf{X}, \mathbf{p})}{Q(\mathbf{X}, \mathbf{q})} \tag{8}$$

$$\tilde{\varepsilon} = \varepsilon Q(\mathbf{X}, \mathbf{q}) = \mathcal{M}(\mathbf{X}) Q(\mathbf{X}, \mathbf{q}) - P(\mathbf{X}, \mathbf{p}) \tag{9}$$

Least-Squares Estimation for PCE-based Rational Approximation

Least-Squares Estimation

→ Determine the unknown coefficients in the rational approximation using a set of samples $\{\mathbf{x}_k, k = 1, \ldots, N\}$ of the input parameters $\mathbf{X}$ and corresponding model evaluations $\{\mathcal{M}(\mathbf{x}_k), k = 1, \ldots, N\}$.

→ Minimize a sample estimate of the modified mean-square error:

$$\widetilde{\mathrm{err}} = \mathbb{E}\left[\left|\mathcal{M}\left(\mathbf{X}\right) Q\left(\mathbf{X};\mathbf{q}\right) - P\left(\mathbf{X};\mathbf{p}\right)\right|^2\right]. \tag{10}$$

Least-Squares Estimation

→ Determine the unknown coefficients in the rational approximation using a set of samples $\{\mathbf{x}_k, k = 1, \dots, N\}$ of the input parameters $\mathbf{X}$ and corresponding model evaluations $\{\mathcal{M}(\mathbf{x}_k), k = 1, \dots, N\}$.

→ Minimize a sample estimate of the modified mean-square error:

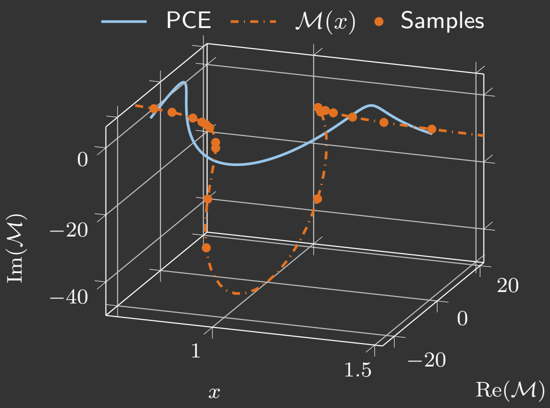
$$\widetilde{\mathrm{err}} = \mathbb{E}\left[\left|\mathcal{M}\left(\mathbf{X}\right) Q\left(\mathbf{X};\mathbf{q}\right) - P\left(\mathbf{X};\mathbf{p}\right)\right|^2\right]. \tag{10}$$

The minimizer is the solution of the following homogeneous linear system of equations [8]

$$\begin{bmatrix} \mathbf{\Psi}_P^{\mathrm{T}}\mathbf{\Psi}_P & -\mathbf{\Psi}_P^{\mathrm{T}}\,\mathrm{diag}\left(\mathbf{y}\right)\mathbf{\Psi}_Q \\ -\mathbf{\Psi}_Q^{\mathrm{T}}\,\mathrm{diag}\left(\overline{\mathbf{y}}\right)\mathbf{\Psi}_P & \mathbf{\Psi}_Q^{\mathrm{T}}\,\mathrm{diag}\left(\mathbf{y}\circ\overline{\mathbf{y}}\right)\mathbf{\Psi}_Q \end{bmatrix} \begin{bmatrix} \mathbf{p} \\ \mathbf{q} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \tag{11}$$

Matrices $\mathbf{\Psi}_P \in \mathbb{R}^{N\times n_p}$ and $\mathbf{\Psi}_Q \in \mathbb{R}^{N\times n_q}$ have as (i,j)-element $\psi_j(\mathbf{x}_i)$ and vector $\mathbf{y} \in \mathbb{C}^N$ has as i-element the model evaluation $\mathcal{M}\left(\mathbf{x}_i\right)$.

Simple Example



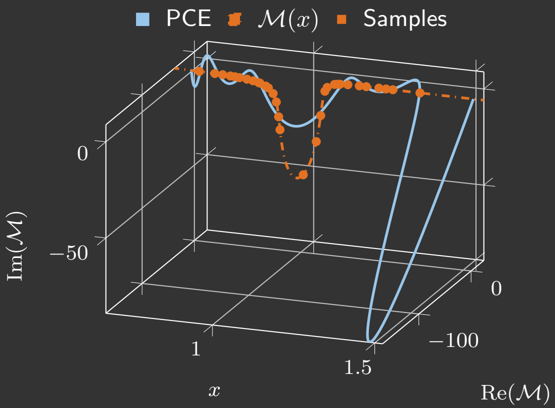
PCE with $m = 5$.

$$\mathcal{M}(x) = \frac{1}{x - 1 + \mathrm{i}0.02x}.$$

Problem settings:

- X follows a lognormal distribution with mean $\mu_X = 1$ and standard deviation $\sigma_X = 0.2$.
- Number of samples N chosen three times the number of polynomials, i.e., $N = 3(m+1)$, where m is the polynomial order.

Simple Example



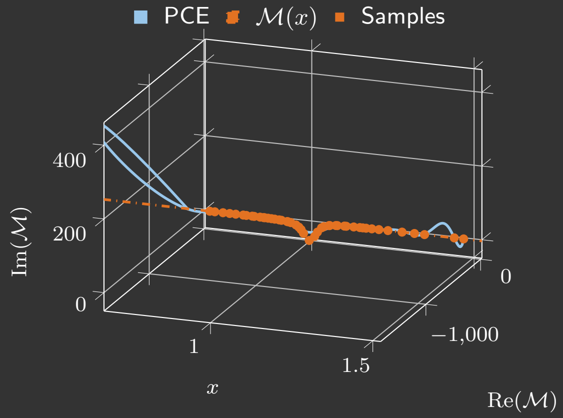
PCE with $m = 10$.

$$\mathcal{M}(x) = \frac{1}{x - 1 + \mathrm{i}0.02x}.$$

Problem settings:

- X follows a lognormal distribution with mean $\mu_X = 1$ and standard deviation $\sigma_X = 0.2$.
- Number of samples N chosen three times the number of polynomials, i.e., $N = 3(m+1)$, where m is the polynomial order.

Simple Example



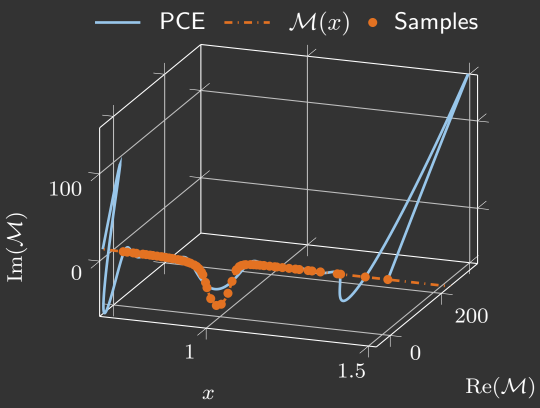
PCE with $m = 15$.

$$\mathcal{M}(x) = \frac{1}{x - 1 + \mathrm{i}0.02x}.$$

Problem settings:

- X follows a lognormal distribution with mean $\mu_X = 1$ and standard deviation $\sigma_X = 0.2$.
- Number of samples N chosen three times the number of polynomials, i.e., $N = 3(m+1)$, where m is the polynomial order.

Simple Example



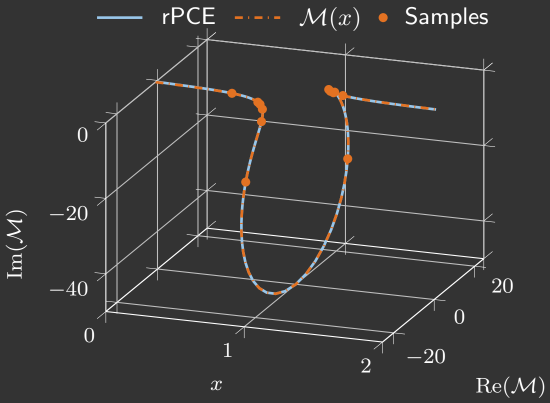
PCE with $m = 20$.

$$\mathcal{M}(x) = \frac{1}{x - 1 + \mathrm{i}0.02x}.$$

Problem settings:

- X follows a lognormal distribution with mean $\mu_X = 1$ and standard deviation $\sigma_X = 0.2$.
- Number of samples N chosen three times the number of polynomials, i.e., $N = 3(m+1)$, where m is the polynomial order.

Simple Example



rPCE with $m_p = m_q = 1$.

$$\mathcal{M}(x) = \frac{1}{x - 1 + \mathrm{i}0.02x}.$$

Problem settings:

- X follows a lognormal distribution with mean $\mu_X = 1$ and standard deviation $\sigma_X = 0.2$.
- Number of samples N chosen three times the number of polynomials, i.e., $N = 3(m+1)$, where m is the polynomial order.

Sparse Bayesian Learning of PCE-based Rational Approximations

Bayesian Formulation

Instead of casting the problem into a deterministic regression formulation, we develop a Bayesian learning strategy in [6].
$\rightarrow$ The coefficients in the rational approximation are treated as probabilistic and Bayes' theorem is applied:

$$f(\mathbf{p}, \mathbf{q}|\mathbf{y}) = c_E^{-1} L(\mathbf{p}, \mathbf{q}|\mathbf{y}) f(\mathbf{p}, \mathbf{q}) \,. \tag{12}$$

Bayesian Formulation

Instead of casting the problem into a deterministic regression formulation, we develop a Bayesian learning strategy in [6].
$\rightarrow$ The coefficients in the rational approximation are treated as probabilistic and Bayes' theorem is applied:

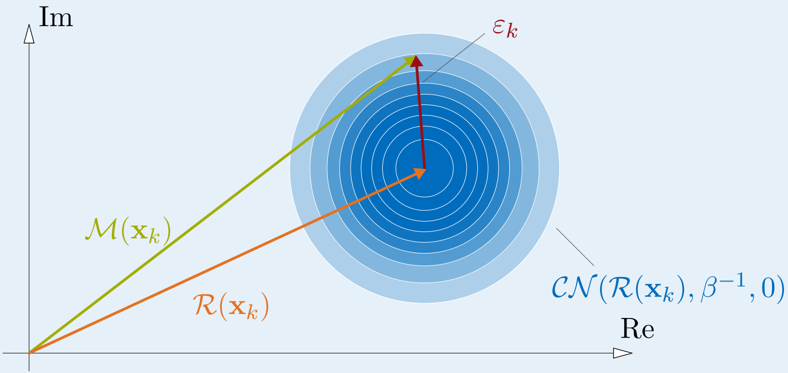
$$f(\mathbf{p},\mathbf{q}|\mathbf{y}) = c_E^{-1} L(\mathbf{p},\mathbf{q}|\mathbf{y}) f(\mathbf{p},\mathbf{q})\,. \tag{12}$$

$\rightarrow$ Likelihood function $L(\mathbf{p},\mathbf{q}|\mathbf{y}) \sim f_{\mathbf{Y}}(\mathbf{y}|\mathbf{p},\mathbf{q})$ in [6] is derived assuming the following additive error model

$$\mathcal{M}(\mathbf{x}_k) = \mathcal{R}(\mathbf{x}_k;\mathbf{p},\mathbf{q}) + \varepsilon_k\,. \tag{13}$$

Bayesian Formulation

Illustration of the error model



Bayesian Formulation

Instead of casting the problem into a deterministic regression formulation, we develop a Bayesian learning strategy in [6].
$\rightarrow$ The coefficients in the rational approximation are treated as probabilistic and Bayes' theorem is applied:

$$f(\mathbf{p},\mathbf{q}|\mathbf{y}) = c_E^{-1} L(\mathbf{p},\mathbf{q}|\mathbf{y}) f(\mathbf{p},\mathbf{q})\,. \tag{12}$$

$\rightarrow$ Likelihood function $L(\mathbf{p},\mathbf{q}|\mathbf{y}) \sim f_{\mathbf{Y}}(\mathbf{y}|\mathbf{p},\mathbf{q})$ in [6] is derived assuming the following additive error model

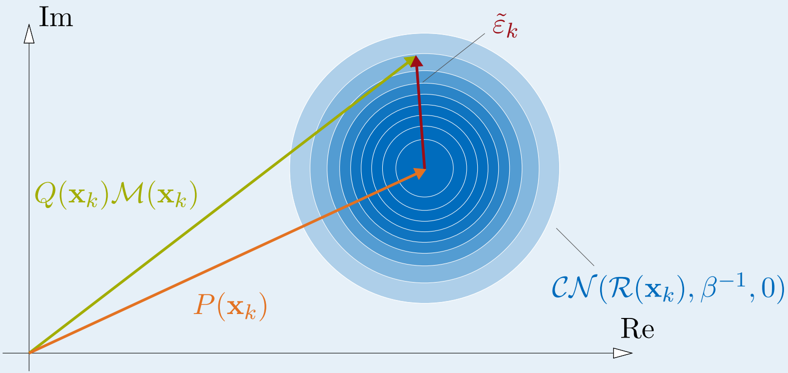
$$\mathcal{M}(\mathbf{x}_k) = \mathcal{R}(\mathbf{x}_k;\mathbf{p},\mathbf{q}) + \varepsilon_k\,. \tag{13}$$

$\rightarrow$ However, we can also consider the linearized residual formulation with

$$Q(\mathbf{x}_k;\mathbf{q})\mathcal{M}(\mathbf{x}_k) = P(\mathbf{x}_k;\mathbf{p}) + \tilde{\varepsilon}_k\,. \tag{14}$$

Bayesian Formulation

Illustration of the error model



Likelihood Formulation for Linearized Residual

Under the above error model, the expectation and covariance of the data are

$$\mathbb{E}\left[\mathbf{y}|\mathbf{p},\mathbf{q}\right] = \mathbf{Q}^{-1}\mathbf{\Psi}_p\mathbf{p}\,, \tag{12}$$

$$\mathbb{C}\mathrm{ov}\left[\mathbf{y}|\mathbf{p},\mathbf{q}\right] = \mathbf{Q}^{-1}\mathbf{\Sigma}_{\tilde{\varepsilon}\tilde{\varepsilon}}\mathbf{Q}^{-H}\,, \tag{13}$$

where $\mathbf{Q} = \mathrm{diag}(\mathbf{\Psi}_q\mathbf{q})$. Since $\mathbf{y}$ depends linearly on $\tilde{\varepsilon}_k$, the likelihood will have a Gaussian distribution with the moments as in Eqs. (12) and (13). With $\mathbf{\Sigma}_{\tilde{\varepsilon}\tilde{\varepsilon}} = \beta^{-1}\mathbf{I}_N$, we find

$$f(\mathbf{y}|\mathbf{p},\mathbf{q}) = \left(\frac{\beta}{\pi}\right)^N \det\left(\mathbf{Q}^{\mathrm{H}}\mathbf{Q}\right)\exp\left\{-\beta\left(\mathbf{Q}\mathbf{y}-\mathbf{\Psi}_p\mathbf{p}\right)^{\mathrm{H}}\left(\mathbf{Q}\mathbf{y}-\mathbf{\Psi}_p\mathbf{p}\right)\right\}. \tag{14}$$

Prior Assumptions

The prior distributions for both sets of coefficients are modeled as complex proper Gaussian distributions, i.e.,

$$f(\mathbf{p}|\boldsymbol{\alpha}_p) = \mathcal{CN}(\mathbf{p}|\mathbf{0}, \boldsymbol{\Lambda}_{pp}^{-1}, \mathbf{0}) \,,$$
$$f(\mathbf{q}|\boldsymbol{\alpha}_q) = \mathcal{CN}(\mathbf{q}|\mathbf{0}, \boldsymbol{\Lambda}_{qq}^{-1}, \mathbf{0}) \,,$$

with

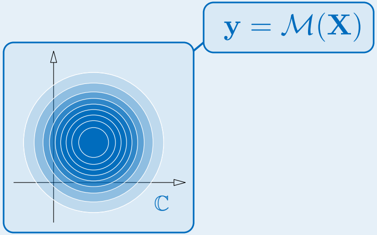
- $\boldsymbol{\Lambda}_{pp} = \operatorname{diag} \boldsymbol{\alpha}_p$ and $\boldsymbol{\Lambda}_{qq} = \operatorname{diag} \boldsymbol{\alpha}_q$
- $\boldsymbol{\alpha}_p = \left[\alpha_{p,0}; \ldots; \alpha_{p,n_p-1}\right]$ and $\boldsymbol{\alpha}_q = \left[\alpha_{q,0}; \ldots; \alpha_{q,n_q-1}\right]$

$\rightarrow$ Assume independence between the individual hyperparameters.

$\rightarrow$ Specify hyperpriors (Gamma) over $\boldsymbol{\alpha}_p$, $\boldsymbol{\alpha}_q$ and β according to [1, 9].

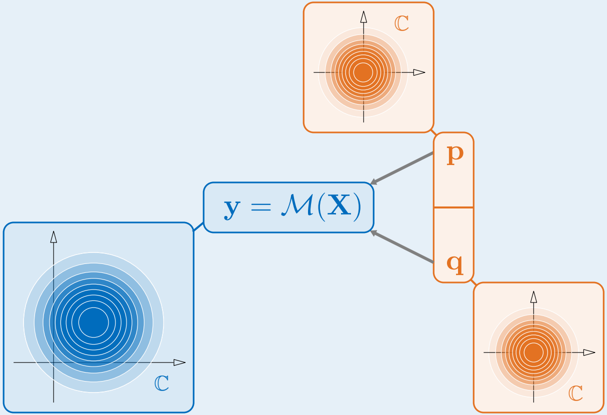
Bayesian Formulation

Illustration of the hierarchical Bayesian model



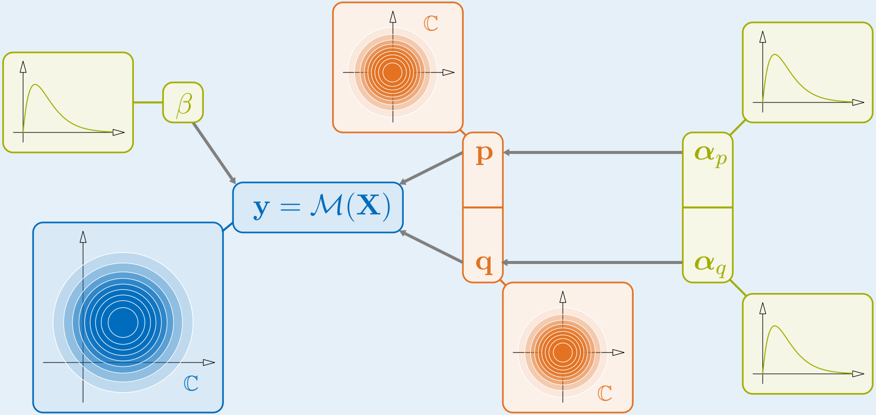
Bayesian Formulation

Illustration of the hierarchical Bayesian model



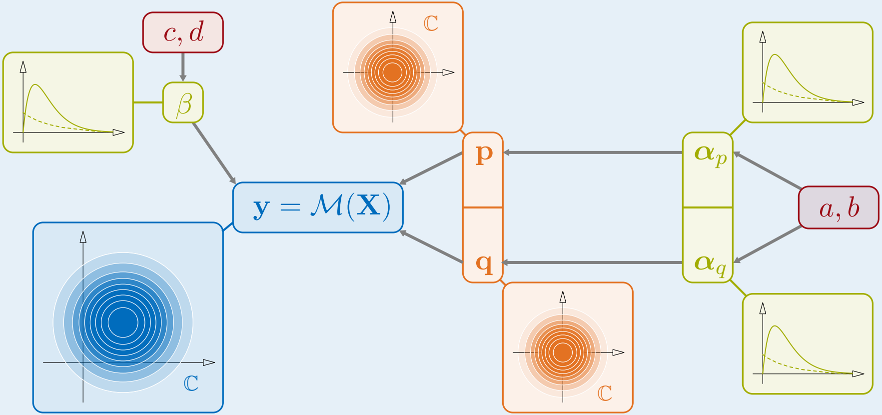
Bayesian Formulation

Illustration of the hierarchical Bayesian model



Bayesian Formulation

Illustration of the hierarchical Bayesian model



Posterior Distribution

The posterior distribution under these model assumptions cannot be computed in closed-form, since the required integration can not be solved directly.

$\rightarrow$ For linear models, an analytic solution is identifiable.

$\rightarrow$ Analytically determine the posterior distribution of $\mathbf{p}$ conditional on $\mathbf{q}$ and all hyperparameters:

$$f(\mathbf{p}|\mathbf{y},\mathbf{q},\boldsymbol{\alpha}_p,\beta) = \frac{1}{\pi^{n_p}\det\tilde{\boldsymbol{\Sigma}}_{\mathbf{pp}|\mathbf{y}}}\exp\Big\{\left(\mathbf{p}-\tilde{\boldsymbol{\mu}}_{\mathbf{p}|\mathbf{y}}\right)\tilde{\boldsymbol{\Sigma}}^{-1}_{\mathbf{pp}|\mathbf{y}}\left(\mathbf{p}-\tilde{\boldsymbol{\mu}}_{\mathbf{p}|\mathbf{y}}\right)\Big\}\,, \tag{15}$$

with posterior covariance matrix

$$\tilde{\boldsymbol{\Sigma}}_{\mathbf{pp}|\mathbf{y}} = \left(\boldsymbol{\Lambda}_{\mathbf{pp}} + \beta\boldsymbol{\Psi}_p^{\mathrm{T}}\boldsymbol{\Psi}_p\right)^{-1}\,, \tag{16}$$

and posterior mean

$$\tilde{\boldsymbol{\mu}}_{\mathbf{p}|\mathbf{y}} = \beta\tilde{\boldsymbol{\Sigma}}_{\mathbf{pp}|\mathbf{y}}\boldsymbol{\Psi}_p^{\mathrm{T}}\mathbf{Q}\mathbf{y}\,. \tag{17}$$

Posterior Distribution

The posterior distribution under these model assumptions cannot be computed in closed-form, since the required integration can not be solved directly.

$\rightarrow$ For linear models, an analytic solution is identifiable.

$\rightarrow$ Analytically determine the posterior distribution of $\mathbf{p}$ conditional on $\mathbf{q}$ and all hyperparameters:

$$f(\mathbf{p}|\mathbf{y},\mathbf{q},\boldsymbol{\alpha}_p,\beta) = \frac{1}{\pi^{n_p}\det\tilde{\boldsymbol{\Sigma}}_{\mathbf{pp}|\mathbf{y}}}\exp\left\{\left(\mathbf{p}-\tilde{\boldsymbol{\mu}}_{\mathbf{p}|\mathbf{y}}\right)\tilde{\boldsymbol{\Sigma}}_{\mathbf{pp}|\mathbf{y}}^{-1}\left(\mathbf{p}-\tilde{\boldsymbol{\mu}}_{\mathbf{p}|\mathbf{y}}\right)\right\}, \tag{15}$$

with posterior covariance matrix

$$\tilde{\boldsymbol{\Sigma}}_{\mathbf{pp}|\mathbf{y}} = \left(\boldsymbol{\Lambda}_{\mathbf{pp}}+\beta\boldsymbol{\Psi}_p^{\mathrm{T}}\boldsymbol{\Psi}_p\right)^{-1}, \tag{16}$$

and posterior mean

$$\tilde{\boldsymbol{\mu}}_{\mathbf{p}|\mathbf{y}} = \beta\tilde{\boldsymbol{\Sigma}}_{\mathbf{pp}|\mathbf{y}}\boldsymbol{\Psi}_p^{\mathrm{T}}\mathbf{Q}\mathbf{y}\,. \tag{17}$$

Posterior Distribution

The posterior distribution under these model assumptions cannot be computed in closed-form, since the required integration can not be solved directly.

$\rightarrow$ For linear models, an analytic solution is identifiable.

$\rightarrow$ Analytically determine the posterior distribution of $\mathbf{p}$ conditional on $\mathbf{q}$ and all hyperparameters:

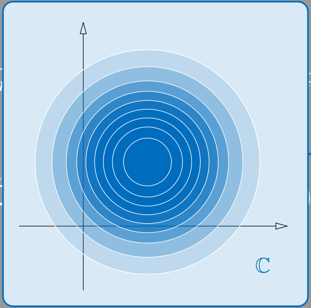
$$f(\mathbf{p}|\mathbf{y},\mathbf{q},\boldsymbol{\alpha}_p,\beta) = \frac{\quad}{\pi^{n}} \quad \ldots \quad {}_{\mathbf{p}|\mathbf{y}}\big)\,\tilde{\boldsymbol{\Sigma}}^{-1}_{\mathbf{pp}|\mathbf{y}}\,\big(\mathbf{p}-\tilde{\boldsymbol{\mu}}_{\mathbf{p}|\mathbf{y}}\big)\Big\}\,, \tag{15}$$

with posterior covariance matrix

$$\tilde{\boldsymbol{\Sigma}} \quad \ldots \quad)^{-1}\,, \tag{16}$$

and posterior mean

$$\ldots\,. \tag{17}$$



Posterior Distribution

The posterior distribution of the denominator coefficients can only be expressed up to proportionality as

$$f(\mathbf{q}|\mathbf{y}, \boldsymbol{\alpha}_p, \boldsymbol{\alpha}_q, \beta) \propto f(\mathbf{y}|\mathbf{q}, \boldsymbol{\alpha}_q) f(\mathbf{q}|\boldsymbol{\alpha}_q)\,. \tag{18}$$

Posterior Distribution

The posterior distribution of the denominator coefficients can only be expressed up to proportionality as

$$f(\mathbf{q}|\mathbf{y},\boldsymbol{\alpha}_p,\boldsymbol{\alpha}_q,\beta) \propto f(\mathbf{y}|\mathbf{q},\boldsymbol{\alpha}_q) f(\mathbf{q}|\boldsymbol{\alpha}_q)\,. \tag{18}$$

$f(\mathbf{y}|\mathbf{q},\boldsymbol{\alpha}_q)$ can be found analytically as

$$f(\mathbf{y}|\mathbf{q},\boldsymbol{\alpha}_q) = \frac{\det\left(\mathbf{Q}\overline{\mathbf{Q}}\right)}{\pi^N \det\left(\tilde{\mathbf{C}}\right)} \exp\left\{-\mathbf{y}^{\mathrm{H}}\overline{\mathbf{Q}}\tilde{\mathbf{C}}^{-1}\mathbf{Q}\mathbf{y}\right\}, \tag{19}$$

where $\tilde{\mathbf{C}} = \beta^{-1}\mathbf{I}_N + \boldsymbol{\Psi}_p\boldsymbol{\Lambda}_{pp}^{-1}\boldsymbol{\Psi}_p^{\mathrm{H}}$ and $f(\mathbf{q}|\boldsymbol{\alpha}_q)$ denotes the prior distribution of the denominator coefficients.

Posterior Distribution

The posterior distribution of the denominator coefficients can only be expressed up to proportionality as

$$f(\mathbf{q}|\mathbf{y}, \boldsymbol{\alpha}_p, \boldsymbol{\alpha}_q, \beta) \propto f(\mathbf{y}|\mathbf{q}, \boldsymbol{\alpha}_q) f(\mathbf{q}|\boldsymbol{\alpha}_q)\,. \tag{18}$$

$f(\mathbf{y}|\mathbf{q}, \boldsymbol{\alpha}_q)$ can be found analytically as

$$f(\mathbf{y}|\mathbf{q}, \boldsymbol{\alpha}_q) = \frac{\det\left(\mathbf{Q}\overline{\mathbf{Q}}\right)}{\pi^N \det\left(\tilde{\mathbf{C}}\right)} \exp\left\{-\mathbf{y}^{\mathrm{H}}\overline{\mathbf{Q}}\tilde{\mathbf{C}}^{-1}\mathbf{Q}\mathbf{y}\right\}, \tag{19}$$

where $\tilde{\mathbf{C}} = \beta^{-1}\mathbf{I}_N + \boldsymbol{\Psi}_p\boldsymbol{\Lambda}_{pp}^{-1}\boldsymbol{\Psi}_p^{\mathrm{H}}$ and $f(\mathbf{q}|\boldsymbol{\alpha}_q)$ denotes the prior distribution of the denominator coefficients.

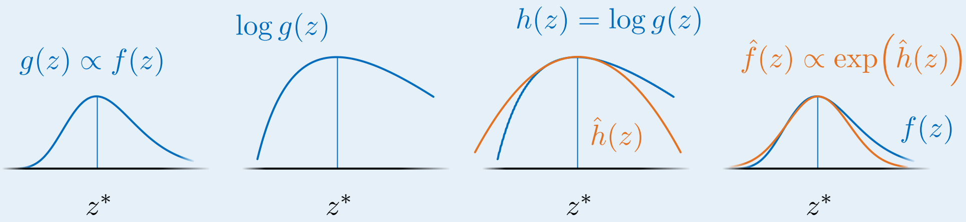
→ Maximum a posteriori approximation: $f(\mathbf{q}|\mathbf{y}, \boldsymbol{\alpha}_p, \boldsymbol{\alpha}_q, \beta) \approx \delta\left(\mathbf{q} - \mathbf{q}^*\right)$.

→ Laplace approximation: $f(\mathbf{q}|\mathbf{y}, \boldsymbol{\alpha}_p, \boldsymbol{\alpha}_q, \beta) \approx \mathcal{CN}\left(\mathbf{q}\,\middle|\,\mathbf{q}^*, \left(-\mathbf{H}_{\mathbf{qq}}\right)^{-1}\right)$.

Approximating the Posterior Distribution of $\mathbf{q}$

Instead of concentrating all probability in the posterior distribution at the MAP value, we can consider Laplace approximation. Under this approximation the logarithm of the posterior distribution is expressed through a second-order Taylor expansion.

Laplace Approximation



Laplace Approximation in Complex Variables

Denote by $g(\mathbf{x})$ some PDF that is known up to its normalization constant. The PDF of interest $f_{\mathbf{X}}(\mathbf{x})$ is then

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{g(\mathbf{x})}{\int_{\mathbb{C}^n} g(\mathbf{x})\,\mathrm{d}\mathbf{x}} \,. \tag{20}$$

A second-order expansion in complex-variables is applied using generalized calculus leading to

$$\ln g(\mathbf{x}) \approx \ln g(\mathbf{x}_0) + \left.\frac{\partial \ln g(\mathbf{x})}{\partial \underline{\mathbf{x}}}\right|_{\mathbf{x}=\mathbf{x}_0} (\underline{\mathbf{x}} - \underline{\mathbf{x}}_0) + \frac{1}{2}\,(\underline{\mathbf{x}} - \underline{\mathbf{x}}_0)^{\mathrm{H}}\,\underline{\mathbf{H}}_{\mathbf{xx}}\,(\underline{\mathbf{x}} - \underline{\mathbf{x}}_0)\;, \tag{21}$$

where $\underline{\mathbf{x}} = [\mathbf{x}; \overline{\mathbf{x}}]$ and $\underline{\mathbf{H}}_{\mathbf{xx}}$ is the augmented Hessian matrix, which can be written in block form as

$$\underline{\mathbf{H}}_{\mathbf{xx}} = \begin{bmatrix} \frac{\partial}{\partial \mathbf{x}}\left(\frac{\partial \ln g(\mathbf{x}_0)}{\partial \mathbf{x}}\right)^{\mathrm{H}} & \frac{\partial}{\partial \overline{\mathbf{x}}}\left(\frac{\partial \ln g(\mathbf{x}_0)}{\partial \mathbf{x}}\right)^{\mathrm{H}} \\ \frac{\partial}{\partial \mathbf{x}}\left(\frac{\partial \ln g(\mathbf{x}_0)}{\partial \overline{\mathbf{x}}}\right)^{\mathrm{H}} & \frac{\partial}{\partial \overline{\mathbf{x}}}\left(\frac{\partial \ln g(\mathbf{x}_0)}{\partial \overline{\mathbf{x}}}\right)^{\mathrm{H}} \end{bmatrix} \tag{22}$$

Laplace Approximation in Complex Variables

Denote by $g(\mathbf{x})$ some PDF that is known up to its normalization constant. The PDF of interest $f_{\mathbf{X}}(\mathbf{x})$ is then

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{g(\mathbf{x})}{\int_{\mathbb{C}^n} g(\mathbf{x})\,\mathrm{d}\mathbf{x}} \,. \tag{20}$$

A second-order expansion in complex-variables is applied using generalized calculus leading to

$$\ln g(\mathbf{x}) \approx \ln g(\mathbf{x}_0) + \left.\frac{\partial \ln g(\mathbf{x})}{\partial \underline{\mathbf{x}}}\right|_{\mathbf{x}=\mathbf{x}_0} (\underline{\mathbf{x}} - \underline{\mathbf{x}}_0) + \frac{1}{2} (\underline{\mathbf{x}} - \underline{\mathbf{x}}_0)^{\mathrm{H}} \underline{\mathbf{H}}_{\mathbf{xx}} (\underline{\mathbf{x}} - \underline{\mathbf{x}}_0) \,, \tag{21}$$

where $\underline{\mathbf{x}} = [\mathbf{x}; \overline{\mathbf{x}}]$ and $\underline{\mathbf{H}}_{\mathbf{xx}}$ is the augmented Hessian matrix, which can be written in block form as

$$\underline{\mathbf{H}}_{\mathbf{xx}} \approx \begin{bmatrix} \frac{\partial}{\partial \mathbf{x}} \left(\frac{\partial \ln g(\mathbf{x}_0)}{\partial \mathbf{x}} \right)^{\mathrm{H}} & \mathbf{0} \\ \mathbf{0} & \frac{\partial}{\partial \overline{\mathbf{x}}} \left(\frac{\partial \ln g(\mathbf{x}_0)}{\partial \overline{\mathbf{x}}} \right)^{\mathrm{H}} \end{bmatrix} \tag{22}$$

Laplace Approximation in Complex Variables

Denote by $g(\mathbf{x})$ some PDF that is known up to its normalization constant. The PDF of interest $f_{\mathbf{X}}(\mathbf{x})$ is then

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{g(\mathbf{x})}{\int_{\mathbb{C}^n} g(\mathbf{x})\,\mathrm{d}\mathbf{x}}\,. \tag{20}$$

A second-order expansion in complex-variables is applied using generalized calculus leading to

$$\ln g(\mathbf{x}) \approx \ln g(\mathbf{x}_0) + \left.\frac{\partial \ln g(\mathbf{x})}{\partial \underline{\mathbf{x}}}\right|_{\mathbf{x}=\mathbf{x}_0} (\underline{\mathbf{x}} - \underline{\mathbf{x}}_0) + \frac{1}{2}\,(\underline{\mathbf{x}} - \underline{\mathbf{x}}_0)^{\mathrm{H}}\, \underline{\mathbf{H}}_{\mathbf{xx}}\,(\underline{\mathbf{x}} - \underline{\mathbf{x}}_0)\;, \tag{21}$$

where $\underline{\mathbf{x}} = [\mathbf{x}; \overline{\mathbf{x}}]$ and $\underline{\mathbf{H}}_{\mathbf{xx}}$ is the augmented Hessian matrix, which can be written in block form as

$$\underline{\mathbf{H}}_{\mathbf{xx}} = \begin{bmatrix} \mathbf{H}_{\mathbf{xx}} & \mathbf{0} \\ \mathbf{0} & \overline{\mathbf{H}}_{\mathbf{xx}} \end{bmatrix} \tag{22}$$

Laplace Approximation in Complex Variables

A second-order expansion in complex-variables is applied using generalized calculus leading to

$$\ln g(\mathbf{x}) \approx \ln g(\mathbf{x}_0) + \left.\frac{\partial \ln g(\mathbf{x})}{\partial \underline{\mathbf{x}}}\right|_{\mathbf{x}=\mathbf{x}_0} (\underline{\mathbf{x}} - \underline{\mathbf{x}}_0) + \frac{1}{2} (\underline{\mathbf{x}} - \underline{\mathbf{x}}_0)^{\mathrm{H}} \underline{\mathbf{H}}_{\mathbf{xx}} (\underline{\mathbf{x}} - \underline{\mathbf{x}}_0) \ , \tag{20}$$

where $\underline{\mathbf{x}} = [\mathbf{x}; \overline{\mathbf{x}}]$ and $\underline{\mathbf{H}}_{\mathbf{xx}}$ is the augmented Hessian matrix, which can be written in block form as

$$\underline{\mathbf{H}}_{\mathbf{xx}} = \begin{bmatrix} \mathbf{H}_{\mathbf{xx}} & \mathbf{0} \\ \mathbf{0} & \overline{\mathbf{H}}_{\mathbf{xx}} \end{bmatrix} \tag{21}$$

For $\mathbf{x}_0 = \mathbf{x}^*$ (MAP), the PDF of $\mathbf{X}$ is approximately proper complex normal with mean $\mathbf{x}^*$ and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{xx}} = -\left(\mathbf{H}_{\mathbf{xx}}\right)^{-1}$.

Sequential Solution Strategy

In order to find suitable choices for the remaining parameters, our strategy is as follows:

1. Find the maximum a-posteriori (MAP) estimate for the denominator coefficients, $\mathbf{q}^*$,

$$\mathbf{q}^* = \underset{\mathbf{q} \in \mathbb{C}^{n_q}}{\arg\max} \, f(\mathbf{y}|\mathbf{q}, \boldsymbol{\alpha}_p, \beta) f(\mathbf{q}|\boldsymbol{\alpha}_q) \,. \tag{22}$$

   and approximate the posterior distribution of $\mathbf{q}$ as a proper complex Gaussian distribution with mean $\mathbf{q}^*$ and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{qq}|\mathbf{y}} = -\left(\mathbf{H}_{\mathbf{qq}}\right)^{-1}$.

Sequential Solution Strategy

In order to find suitable choices for the remaining parameters, our strategy is as follows:

1. Find the maximum a-posteriori (MAP) estimate for the denominator coefficients, $\mathbf{q}^*$,

$$\mathbf{q}^* = \underset{\mathbf{q}\in\mathbb{C}^{n_q}}{\arg\max}\, f(\mathbf{y}|\mathbf{q}, \boldsymbol{\alpha}_p, \beta) f(\mathbf{q}|\boldsymbol{\alpha}_q)\,. \tag{22}$$

   and approximate the posterior distribution of $\mathbf{q}$ as a proper complex Gaussian distribution with mean $\mathbf{q}^*$ and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{q}\mathbf{q}|\mathbf{y}} = -\left(\mathbf{H}_{\mathbf{q}\mathbf{q}}\right)^{-1}$.

2. Under the approximation compute the evidence conditional on the hyperparameters as

$$f(\mathbf{y}|\boldsymbol{\alpha}_p, \boldsymbol{\alpha}_q, \beta) \approx f(\mathbf{y}|\mathbf{q}^*, \boldsymbol{\alpha}_p, \beta) f(\mathbf{q}^*|\boldsymbol{\alpha}_q)\pi^n \det\left(-\mathbf{H}_{\mathbf{q}\mathbf{q}}\right)^{-1} \tag{23}$$

Sequential Solution Strategy

In order to find suitable choices for the remaining parameters, our strategy is as follows:

1. Find the maximum a-posteriori (MAP) estimate for the denominator coefficients, $\mathbf{q}^*$,

$$\mathbf{q}^* = \underset{\mathbf{q}\in\mathbb{C}^{n_q}}{\arg\max}\, f(\mathbf{y}|\mathbf{q}, \boldsymbol{\alpha}_p, \beta) f(\mathbf{q}|\boldsymbol{\alpha}_q)\,. \tag{22}$$

   and approximate the posterior distribution of $\mathbf{q}$ as a proper complex Gaussian distribution with mean $\mathbf{q}^*$ and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{qq}|\mathbf{y}} = -\left(\mathbf{H}_{\mathbf{qq}}\right)^{-1}$.

2. Under the approximation compute the evidence conditional on the hyperparameters as

$$f(\mathbf{y}|\boldsymbol{\alpha}_p, \boldsymbol{\alpha}_q, \beta) \approx f(\mathbf{y}|\mathbf{q}^*, \boldsymbol{\alpha}_p, \beta) f(\mathbf{q}^*|\boldsymbol{\alpha}_q)\pi^n \det\left(-\mathbf{H}_{\mathbf{qq}}\right)^{-1} \tag{23}$$

3. Maximize the evidence under the Laplace approximation in Eq. (23) at the given MAP estimate $\mathbf{q}^*$ (type-II-maximum likelihood).

Sequential Solution Strategy

In order to find suitable choices for the remaining parameters, our strategy is as follows:

1. Find the maximum a-posteriori (MAP) estimate for the denominator coefficients, $\mathbf{q}^*$,

$$\mathbf{q}^* = \underset{\mathbf{q}\in\mathbb{C}^{n_q}}{\arg\max}\, f(\mathbf{y}|\mathbf{q}, \boldsymbol{\alpha}_p, \beta) f(\mathbf{q}|\boldsymbol{\alpha}_q)\,. \tag{22}$$

   and approximate the posterior distribution of $\mathbf{q}$ as a proper complex Gaussian distribution with mean $\mathbf{q}^*$ and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{q}\mathbf{q}|\mathbf{y}} = -\left(\mathbf{H}_{\mathbf{q}\mathbf{q}}\right)^{-1}$.
2. Under the approximation compute the evidence conditional on the hyperparameters as

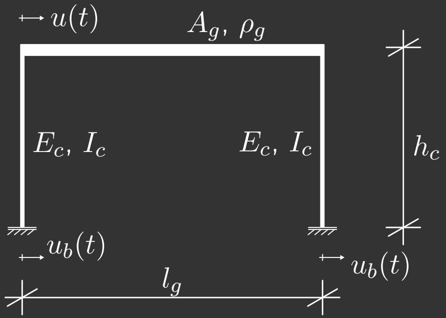
$$f(\mathbf{y}|\boldsymbol{\alpha}_p, \boldsymbol{\alpha}_q, \beta) \approx f(\mathbf{y}|\mathbf{q}^*, \boldsymbol{\alpha}_p, \beta) f(\mathbf{q}^*|\boldsymbol{\alpha}_q)\pi^n \det\left(-\mathbf{H}_{\mathbf{q}\mathbf{q}}\right)^{-1} \tag{23}$$

3. Maximize the evidence under the Laplace approximation in Eq. (23) at the given MAP estimate $\mathbf{q}^*$ (type-II-maximum likelihood).
4. Repeat 1. - 3. and prune non-significant terms until convergence.

Numerical Example
Transmissibility of Frame Structure

Mechanical Model

We consider a single storey frame structure and aim at approximating the frequency response function at $f = 5,1\,\mathrm{Hz}$.



$$\mathcal{M}\left(\mathbf{X};\omega\right) = \frac{\tilde{u}\left(\mathbf{X};\omega\right)}{\tilde{u}_b\left(\mathbf{X};\omega\right)}$$

Mechanical Model

The input random variables $\mathbf{X} = [E_c, I_c, h_c, \rho_g, A_g, l_g, \zeta]$ are assumed to be independent and marginally distributed as follows.

Table 1: Distribution Parameters

Parameter		Mean value	Coefficient of variation
Columns' Young's modulus	E_c	$3 \cdot 10^{10}\,\frac{\mathrm{N}}{\mathrm{m}^2}$	0.1
Columns' moment of inertia	I_c	$\frac{\pi(0.3\mathrm{m})^4}{64}$	0.1
Columns' height	h_c	$4\,\mathrm{m}$	0.1
Girder's density	ρ_g	$2.5 \cdot 10^3\,\mathrm{kgm}^{-3}$	0.05
Girder's cross-sectional area	A_g	$0.3\,\mathrm{m} \cdot 0.5\,\mathrm{m}$	0.1
Girder's length	l_g	$10\,\mathrm{m}$	0.1
Damping Ratio	ζ	0.02	0.3

The nominal eigenfrequency of the system is $\mathbf{f}_n = 5.5\,\mathrm{Hz}$.

Surrogate Model

The surrogate model parameters are chosen as follows.

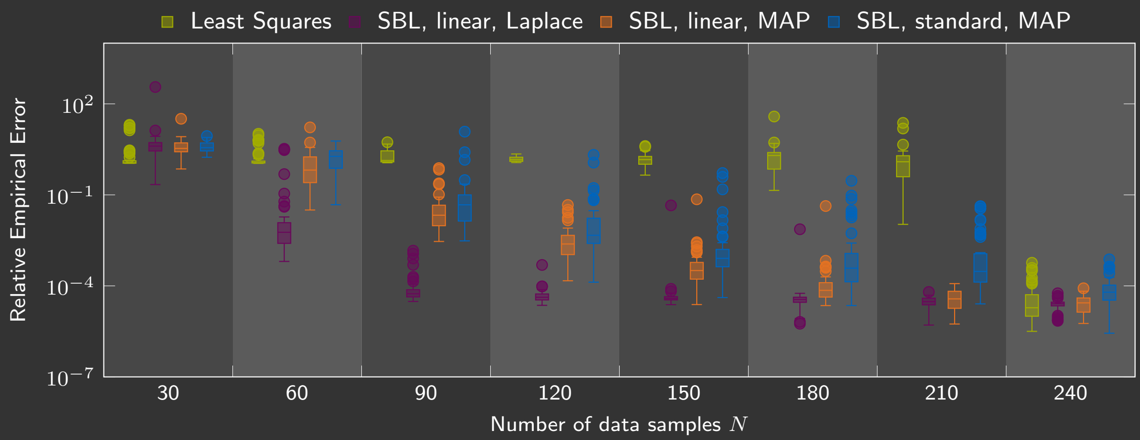
Table 2: Surrogate Parameters

Parameter		Value
Maximum polynomial degrees	$m_p = m_q$	3
Hyperbolic truncation	$q_p = q_q$	1
Number of polynomial terms	N_{pol}	240
Size of training set	N_{train}	$30, 60, \ldots, 240$
Size of test set	N_{test}	10^5
Number of repetitions	N_{rep}	50

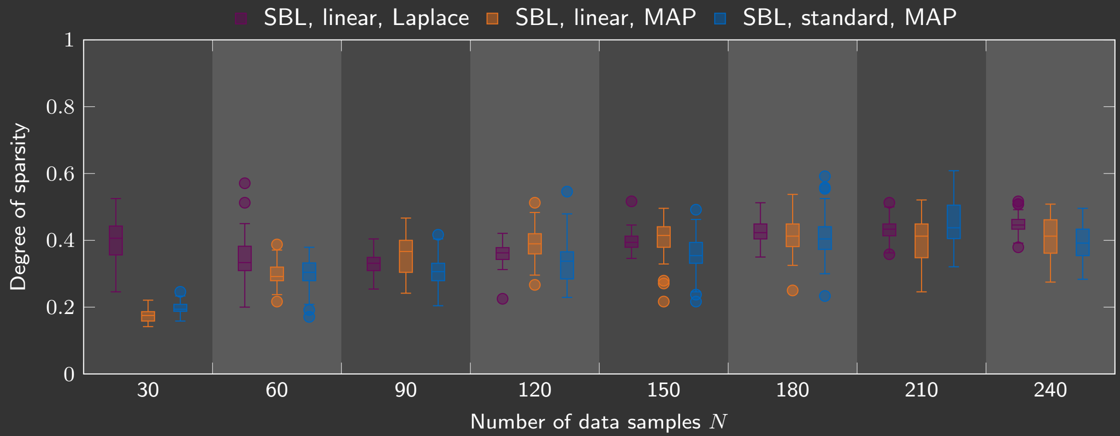
Model accuracy is assessed through the relative empirical error

$$\mathrm{err}_{\text{emp}} = \frac{\sum_{i=1}^{N_{\text{test}}} \left|\mathcal{M}(\mathbf{x}_i;\omega) - \mathcal{R}(\mathbf{x}_i;\omega)\right|^2}{\widehat{\mathrm{Var}}\left[\mathcal{M}(\mathbf{x}_i;\omega)\right]}. \tag{24}$$

Relative Empirical Error Comparison



Degree of Sparsity



Conclusion & Outlook

Conclusion

- Uncertainty Quantifcation and a Bayesian framework for model updating in structural dynamics based on frequency response data is presented.
- To reduce the computational burden, the model responses are emulated using a rational polynomial chaos expansion surrogate model.
- The coefficients are found through an iterative algorithm that maximizes the model evidence. We exploit the linearity in the numerator polynomials and describe the posterior distribution of the numerator polynomials conditional on all other parameters analytically.
- A Laplace approximation is incorporated to find a more accurate representation of the posterior distribution of the denominator coefficients. Through this, a measure of the uncertainty in each coefficient is reflected in the posterior distribution.
- The Bayesian learning framework allows to obtain a measure of the uncertainty in the surrogate prediction that can be utilized in an Bayesian Bayesian optimization framework.

Outlook

- We are currently working on an active-learning approach to estimating the rational PCE model in an Bayesian inference problem.
- Investigate alternative sparsity inducing hierarchical prior structures.
- Extend the applicability of rational PCE to real-valued data.

References I

[1] James O. Berger. *Statistical Decision Theory and Bayesian Analysis*. Springer New York, 1985. DOI: 10.1007/978-1-4757-4286-2. URL: https://doi.org/10.1007/978-1-4757-4286-2.

[2] E. Jacquelin et al. "Polynomial Chaos Expansion and Steady-State Response of a Class of Random Dynamical Systems". In: *Journal of Engineering Mechanics* 141.4 (2015), p. 04014145. ISSN: 0733-9399. DOI: 10.1061/(ASCE)EM.1943-7889.0000856.

[3] E. Jacquelin et al. "Polynomial chaos expansion in structural dynamics: Accelerating the convergence of the first two statistical moment sequences". In: *Journal of Sound and Vibration* 356 (2015), pp. 144–154. ISSN: 0022460X. DOI: 10.1016/j.jsv.2015.06.039.

References II

[4] Jacquelin E. et al. "Polynomial chaos–based extended Padé expansion in structural dynamics". In: *International Journal for Numerical Methods in Engineering* 111.12 (2016), pp. 1170–1191. ISSN: 0029-5981. DOI: 10.1002/nme.5497.

[5] Jun Lu et al. "Uncertainty propagation of frequency response functions using a multi-output Gaussian Process model". In: *Computers & Structures* 217 (2019), pp. 1–17. ISSN: 0045-7949. DOI: https://doi.org/10.1016/j.compstruc.2019.03.009. URL: https://www.sciencedirect.com/science/article/pii/S004579491831681X.

[6] Felix Schneider, Iason Papaioannou, and Gerhard Müller. *Sparse Bayesian Learning for Complex-Valued Rational Approximations*. 2022. DOI: 10.48550/ARXIV.2206.02523. URL: https://arxiv.org/abs/2206.02523.

References III

[7] Felix Schneider et al. "Bayesian parameter updating in linear structural dynamics with frequency transformed data using rational surrogate models". In: *Mechanical Systems and Signal Processing* 166 (2022), p. 108407. ISSN: 08883270. DOI: 10.1016/j.ymssp.2021.108407. URL: https://www.sciencedirect.com/science/article/pii/S0888327021007573.

[8] Felix Schneider et al. "Polynomial chaos based rational approximation in linear structural dynamics with parameter uncertainties". In: *Computers & Structures* 233 (2020), p. 106223. ISSN: 0045-7949. DOI: https://doi.org/10.1016/j.compstruc.2020.106223. URL: http://www.sciencedirect.com/science/article/pii/S0045794920300262.

[9] Michael E Tipping. "Sparse Bayesian learning and the relevance vector machine". In: *Journal of machine learning research* 1.Jun (2001), pp. 211–244.

References IV

[10] Vahid Yaghoubi et al. "Sparse polynomial chaos expansions of frequency response functions using stochastic frequency transformation". In: *Probabilistic Engineering Mechanics* 48 (2017), pp. 39–58.

Thank You