Parameter Estimation with Correlated Outputs using Fidelity Maps

Sylvain Lacaze[a], Samy Missoum[a,*]

[a]*Aerospace and Mechanical Engineering Department, University of Arizona, Tucson, Arizona, 85721*

Abstract

This paper introduces a new approach for parameter estimation and model update based on the notion of fidelity maps. Fidelity maps refer to the regions of the parameter space within which the discrepancy between computational and experimental data is below a user-defined threshold. It is shown that fidelity maps provide an efficient and rigorous approach to approximate likelihoods in the context of Bayesian update or maximum likelihood estimation. Fidelity maps are constructed explicitly in terms of the parameters and aleatory uncertainties using a Support Vector Machine (SVM) classifier. The approach has the advantage of handling numerous correlated responses, possibly discontinuous, without any assumption on the correlation structure. The construction of accurate fidelity map boundaries at a moderate computational cost is made possible through a dedicated adaptive sampling scheme. A simply supported plate with uncertainties in the boundary conditions is used to demonstrate the methodology. In this example, the construction of the fidelity map is based on several natural frequencies and mode shapes to be matched simultaneously. Various statistical estimators are derived from the map.

Keywords: Parameter estimation, Model Update, Likelihood, Bayesian Update, Support Vector Machines

1. Introduction

Computational models are used, for instance, to predict the static or dynamic behavior of structures. However, there might be marked discrepancies between the prediction of the model and experimental data. In order to reduce this difference, the model needs to be calibrated (or updated) by searching parameter values (e.g., material properties) that best "match" the data. For example, in modal analysis, the characteristics of the model (e.g., stiffness and mass distribution) will be modified so as to match experimental natural frequencies and mode shapes [1].

In engineering applications, the most widely used technique is the least square approach. However, uncertainties might have a pronounced effect on the responses of the system and this approach, often implemented in a deterministic way, is in general not suitable [2, 3]. For this reason, statistical approaches have been favored to extract distributions of update parameters and responses. The two most common statistical approaches are the maximum likelihood estimate and Bayesian update. While the maximum likelihood approach [4, 5] finds the most "probable" values of the parameters to be estimated, the Bayesian method [6, 7] focuses on refining the parameter distributions inferred from previous knowledge.

At the core of both approaches, lies the computation of the likelihood. In most engineering applications, the likelihood is difficult to compute and is approximated using assumptions on the correlation structure of the responses (e.g., independence). This difficulty is further exacerbated by computationally intense simulations, large num-

*Corresponding author

Email addresses: lacaze@email.arizona.edu (Sylvain Lacaze), smissoum@email.arizona.edu (Samy Missoum)

ber of responses [8, 9], and discontinuous responses.

The proposed update approach is designed to provide a flexible scheme which tackles the aforementioned difficulties. This is done through the identification of the regions of the parameter space where the discrepancy between model and experimental outputs is below a given threshold. These regions form a "fidelity map" and can be shown to provide a rigorous and efficient approximation of the likelihood without restrictive assumptions .

The boundaries of the fidelity maps are constructed using a Support Vector Machine (SVM) which is a classification technique. It is used to explicitly separate data belonging to two classes [10, 11, 12, 13]. In the context of the fidelity maps, this binary classification is performed based on the discrepancy between computational and experimental data which is either smaller or larger than a given user-defined threshold. In order to obtain an accurate fidelity map using a reasonable number of simulation calls, the SVM boundary is refined using an adaptive sampling scheme [14]. That is, most of the computational cost is concentrated in the construction of the fidelity maps. The likelihood can then be efficiently obtained as a sub-product of the fidelity map.

Many frameworks for model update have been developed. Of particular importance and impact is the work by Kennedy and O'Hagan [15] which has become a reference in the domain. For instance, the use of a Gaussian process was initially introduced in this work. This approach was subsequently used in other strategies [16]. There is also a large body of literature dedicated to Bayesian approaches [17, 18]. The proposed work sets itself apart from existing approaches by enabling model updating for problems with numerous responses (potentially discontinuous) without any a-priori on the correlation structure which is implicitly accounted for during the construction of the fidelity map. This flexibility stems from the use of a classification technique such as SVM for the construction of the fidelity map boundaries.

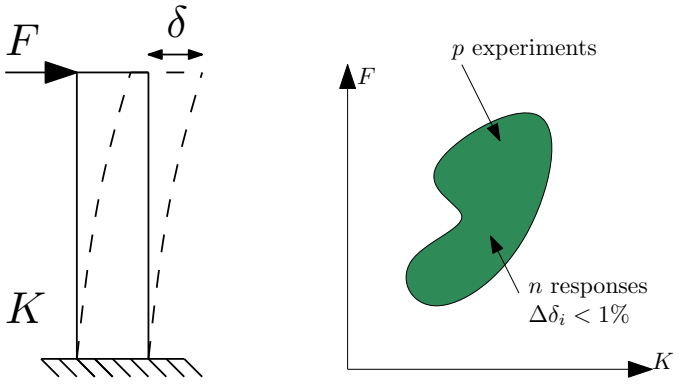
This article is constructed as follows. Section 2 provides the notation and the main concepts of the proposed approach. Section 3 presents the statistical estimators used in this article. Section 4 describes the fidelity map and its use to approximate the likelihood. Section 4.1 provides a background on SVM classifiers. Section 4.2 describes the adaptive sampling scheme used to accurately build the fidelity maps. Finally, Section 5 provides results on a demonstrative example consisting of a plate with uncertainty on the boundary conditions. In the example, several frequencies and mode shapes are to be matched simultaneously. For the sake of completeness, the results are compared to approaches where the responses are assumed independent or if a residual, which encompasses all the responses within one quantity, is used.

2. Illustrative example and notations

Consider the responses $\mathbf{y}$ of a model and the corresponding experimental measurements $\mathbf{y}^{exp}$. The responses of the system are governed by two types of parameters: the first set are the parameters to estimate $\mathbf{x}$ (e.g., material properties) while the second one, $\mathbf{A}$, are the "aleatory" parameters which are not to be estimated but introduce uncertainty (e.g., external load). The probability density function (PDF) of a random variable X is noted $\mathbf{f}_X$ and its cumulative distribution function (CDF) is noted $\mathbf{F}_X$.

As an illustrative example, consider a model in the form of a cantilever column (e.g., representative of a building) subjected to wind loading (Figure 1(a)). In this academic example, we wish to estimate the bending stiffness of the column $K \equiv x$ based on a set of experimental data $\boldsymbol{y}^{exp}$ (e.g., deflection δ) knowing that the column is subjected to a random load $F \equiv A$ with known probabilistic distribution.

Figure 1(b) depicts the construction of the fidelity map corresponding to p experiments and n responses (e.g., displacements, accelerations etc.)



(a) Model. (b) Fidelity map.

Figure 1: Illustrative example. Calibration of the stiffness K of a column subjected to a random (aleatory) load F based on experimental responses.

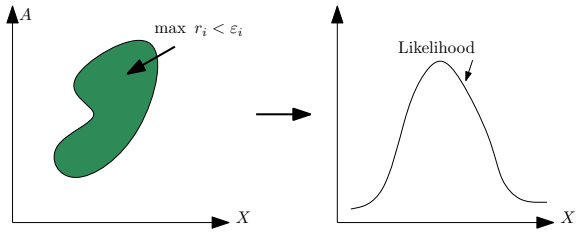


Figure 2: The fidelity map is then used to build an approximation of the likelihood (up to a constant).

per experiment. The fidelity map defines the region of the space where the relative discrepancy between model and experiments Δ_{fi} is lower than 1% for every response. It is accurately constructed with a Support Vector Machine classifier and an adaptive sampling scheme described in Section 4. The fidelity map is then used to approximate the likelihood (Figure 2) which allows one to update of the model through maximum likelihood or Bayesian update. For the reader who is not familiar with these statistical estimators, they are described in the following section which also underlines the advantages of using a fidelity map for their computation.

3. Background

3.1. Maximum Likelihood Estimate

Maximum likelihood estimates (MLE) were originally designed for the statistical inference of hyper-parameters of distributions:

$$\boldsymbol{\theta}^{MLE} = \underset{\boldsymbol{\theta}}{\mathbf{argmax}} \prod_{i=1}^{n_v} \mathbf{f}_X(x_i|\boldsymbol{\theta}) \tag{1}$$

where $\mathbf{x} = [x_1, \ldots, x_{n_v}]$ are n_v i.i.d observations of a random variable X following a PDF $\mathbf{f}_X(x|\boldsymbol{\theta})$ of hyper-parameters $\boldsymbol{\theta}$. This notion can be extended to engineering applications by considering that some output responses $\mathbf{Y}$ follow a joint PDF where uncertainties are due to $\mathbf{A}$ and parametrized by $\mathbf{x}$ (i.e., $\mathbf{f}_{\mathbf{Y}(\mathbf{x},\mathbf{A})}(\mathbf{y}|\mathbf{x})$). Therefore, the maximum likelihood estimate for parameter identification reads:

$$\mathbf{x}^{MLE} = \underset{\mathbf{x}}{\mathbf{argmax}} \prod_{i=1}^{p} \mathbf{f}_{\mathbf{Y}(\mathbf{x},\mathbf{A})}(\mathbf{y}^{exp,(i)}|\mathbf{x}) \tag{2}$$

where $\mathbf{y}^{exp,(i)}$ are the i^{th} experimental set of n responses. In the case of a single set of measurements ($p = 1$), as used in this work, Eq. 2 becomes:

$$\mathbf{x}^{MLE} = \underset{\mathbf{x}}{\mathbf{argmax}}\ \mathbf{f}_{\mathbf{Y}(\mathbf{x},\mathbf{A})}(\mathbf{y}^{exp}|\mathbf{x}) \tag{3}$$

3.2. Bayesian Estimate

While MLE follows a frequentist approach, and considers $\mathbf{x}$ as deterministic, Bayesian updating considers $\mathbf{X}$ as random variables. Bayesian estimators are derived from the Bayes formula:

$$\mathbf{f}_{A|B}\mathbf{f}_B = \mathbf{f}_{B|A}\mathbf{f}_A$$

Specializing it to engineering applications:

$$\mathbf{f}_{\mathbf{X}}(\mathbf{x}|\mathbf{y}^{\mathbf{exp}}) = \frac{\mathbf{f}_{\mathbf{Y}(\mathbf{x},\mathbf{A})}(\mathbf{y}^{exp}|\mathbf{x})\mathbf{f}_{\mathbf{X}}(\mathbf{x})}{\mathbf{f}_{\mathbf{Y}(\mathbf{X},\mathbf{A})}(\mathbf{y}^{exp})} \tag{4}$$

where:

i $\mathbf{f}_{\mathbf{X}}(\mathbf{x}|\mathbf{y}^{\mathbf{exp}})$ is the posterior distribution;
ii $\mathbf{f}_{\mathbf{Y}(\mathbf{x},\mathbf{A})}(\mathbf{y}^{exp}|\mathbf{x})$ is the likelihood;
iii $\mathbf{f}_{\mathbf{X}}$ is the prior distribution;
iv $\mathbf{f}_{\mathbf{Y}(\mathbf{X},\mathbf{A})}(\mathbf{y}^{exp})$ is a normalizing constant which represents the overall probability density to observe $\mathbf{y}^{exp}$.

One Bayes estimator is defined as the expectation of the posterior distribution:

$$x_i^{Bayes} = \mathbb{E}\left[X_i|\mathbf{y}^{\mathbf{exp}}\right] = \int x_i \mathbf{f}_{\mathbf{X}}(\mathbf{x}|\mathbf{y}^{\mathbf{exp}})\mathrm{d}\mathbf{x} \tag{5}$$

Such integral estimators are traditionally obtained through the use of Markov Chain Monte Carlo (MCMC) techniques [19, 20, 21].

3.3. Typical difficulties

If one knows the joint PDF of the responses with respect to $\mathbf{x}$, $\mathbf{f}_{\mathbf{Y}(\mathbf{x},\mathbf{A})}(\mathbf{y}|\mathbf{x})$, the previous estimators can be computed analytically. However in most engineering applications, such analytical expressions are not available as the functional relationship between $(\mathbf{x}, \mathbf{a})$ and $\mathbf{y}$ is only known through a black box model. In addition, these models are usually computationally expensive, limiting the number of model evaluations. Finally, the number of responses y_i can be very large (e.g., modal properties), making impractical the use of conventional surrogates to approximate each responses separately and $\mathbf{f}_{\mathbf{Y}(\mathbf{x},\mathbf{A})}(\mathbf{y}|\mathbf{x}^{(k)})$ for any given $\mathbf{x}^{(k)}$ (c.f., Appendix A). The proposed fidelity map approach provides a new level of flexibility by implicitly accounting for the correlations between multiple responses.

4. Fidelity maps and construction of the likelihood

A fidelity map is defined as the region of the parameter space corresponding to responses within a user-defined interval of the experimental data:

$$FM = \{(\mathbf{x}, \mathbf{a}) \mid r_i \leq \varepsilon_i,\ i = 1, \ldots, n\} \quad (6)$$

where

$$r_i = \left| \frac{y_i(\mathbf{x}, \mathbf{a}) - y_i^{exp}}{y_i^{exp}} \right|$$

It is shown (Section 4.3) that the likelihood of any given $\mathbf{x}^{(k)}$ can be efficiently approximated (up to a constant) as the probability of $\mathbf{x}^{(k)}$ to lie within the fidelity map, $\mathbb{P}\left[(\mathbf{x}^{(k)}, \mathbf{A}) \in FM\right]$.

4.1. SVM-based fidelity map

The fidelity maps are constructed using a Support Vector Machine (SVM) classifier. An SVM defines the boundaries between samples of two different classes (e.g., feasible and infeasible) [10, 11, 13, 12]. In the context of fidelity maps, SVM has the following advantages:

i Only one SVM decision function is needed irrespective of the number of responses $\mathbf{y}$.

ii Because it is a classifier, SVM implicitly accounts for the entire correlation structure of the responses, thus requiring no knowledge or assumptions for the construction of the likelihood.

iii It is insensitive to discontinuities and can handle binary responses [22];

iv The boundaries can be highly non-linear and represent disjoint non convex domains;

v The prediction of a class is very efficient, thus allowing the use of Monte-Carlo type sampling (Section 4.3).

Given N training sample, an SVM classifier is expressed as:

$$s(\mathbf{x}) = b + \sum_{k=1}^{N} \lambda^{(k)} l^{(k)} K(\mathbf{x}^{(k)}, \mathbf{x}) \quad (7)$$

where $\mathbf{x}^{(k)}$ is the k^{th} training sample, $\lambda^{(k)}$ is the corresponding Lagrange multiplier, $l^{(k)}$ is the label (class) that can take values +1 or -1, K is a kernel function (e.g., Gaussian in this work) and b is the bias. The boundary is defined as $s(\mathbf{x}) = 0$.

In order to build the fidelity map, an SVM is initially trained using a design of experiments (DOE). The class of each sample is defined based on the discrepancy between the model outputs and the experimental measurements. To be feasible (i.e., to belong to the fidelity map), a training sample must correspond to absolute relative differences r_i between the model outputs y_i and the measurements y_i^{exp} less than a given threshold ε_i. Therefore, the labels used to train the SVM are defined as:

$$l^{(k)} = \begin{cases} +1 & \text{if } r_i^{(k)} \leq \varepsilon_i,\ i = 1, \ldots, n \\ -1 & \text{otherwise} \end{cases}$$

where:

$$r_i^{(k)} = \left| \frac{y_i^{(k)} - y_i^{exp}}{y_i^{exp}} \right|$$

4.2. Map refinement and adaptive sampling

In order for the likelihood to be accurate, a small enough ε is needed as well as an accurate boundary. This section describes the adaptive sampling scheme as well as the case where the initial user-defined ε is too small.

4.2.1. Adaptive sampling

In order to build an accurate SVM at affordable cost, [23] introduced an adaptive sampling scheme that is used in this work. Assuming the existence of at least one point within the fidelity map (Section 4.2.2), the SVM can be refined using the following two types of samples ([23, 14] for details):

A primary sample, also referred as "max-min" sample. This sample is defined as the point in the space that maximizes the minimum distance to existing samples (i.e., sparse regions) under the constraint that it lies on the SVM boundary (i.e., $s(\mathbf{x})$=0).

A secondary sample is aimed at removing the locking of the SVM by positioning a sample in a region where data from one class is sparse in the vicinity of the boundary.

The sampling scheme used in this work adds two primary samples and one secondary sample [14] per iteration. The sampling scheme is summarized in Algorithm 1.

Algorithm 1 Fidelity map. Adaptive Sampling.

Require: Fidelity level by setting $\boldsymbol{\varepsilon}$;

1: Sample the space ($\{\mathbf{X}, \mathbf{A}\}$) according to a DOE of size n_p: $\mathbf{w}^{(k)} = [\mathbf{x}^{(k)}, \mathbf{a}^{(k)}]$;
2: Evaluate all samples: $\mathbf{y}^{(k)} = \mathbf{y}(\mathbf{w}^{(k)})$;
3: Define the relative difference for each measured response: $r_i^{(k)} = \left| \frac{y_i^{exp} - y_i^{(k)}}{y_i^{exp}} \right|$;
4: For each sample, compute their "feasibility": $r_{max}^{(k)} = \max_i \left(\frac{r_i^{(k)} - \varepsilon_i}{\varepsilon_i} \right)$;
5: Begin adaptive sampling:
6: **for** $j = n_p + 1 \rightarrow n_p + n_{adapt}$ **do**;
7: Set labels to -1: $l^{(k)} = -1,\ k = 1, \ldots, j-1$;
8: Define $\mathbb{K} = \left\{ k | r_{max}^{(k)} \leq 0 \right\}$;
9: **if** $\mathbb{K} = \varnothing$ **then**;
10: Find $k_c = \arg\min_l r_{max}^{(l)}$;
11: Define $\mathbb{K} = \{k_c\}$;
12: **end if**
13: Set $l^{(k)} = +1 \quad \forall k \in \mathbb{K}$;
14: Build the SVM as explained in Section 4.1;
15: Add an adaptive sample $\mathbf{w}^{(j)}$ as explain in Section 4.2.1;
16: Compute : $\mathbf{y}^{(j)}, \mathbf{r}^{(j)}, r_{max}^{(j)}$;
17: **end for**

4.2.2. Notion of most "feasible" sample

It might happen that following the initial DOE, there is not a single sample that satisfies the initial fidelity requirement imposed by the various ε_i. Therefore no feasible sample is available to construct an SVM. In order to solve this issue, the sample $\mathbf{x}^{(k_c)}$ with the minimum discrepancy over all the responses is searched and is defined as the most "feasible". The index k_c is:

$$k_c = \arg\min_k \quad r_{max}^{(k)}$$

where:

$$r_{max}^{(k)} = \max_i \left(\frac{r_i^{(k)} - \varepsilon_i}{\varepsilon_i} \right)$$

4.3. Likelihood approximation

This section shows how to relate the fidelity map to the likelihood. Without loss of generality and for the sake of simplicity, consider x and y^{exp} as two scalars. The relation between a PDF and a CDF is:

$$\mathbf{f}_{Y(x,\mathbf{A})}(y^{exp}|x) = \frac{\mathrm{d}\mathbf{F}_{Y(x,\mathbf{A})}}{\mathrm{d}y}(y^{exp}|x)$$

Therefore:

$$\mathbf{f}_{Y(x,\mathbf{A})}(y^{exp}|x) = \lim_{\varepsilon \to 0} \frac{\mathbb{P}[y^{exp} - \varepsilon \leq Y(x,\mathbf{A}) \leq y^{exp} + \varepsilon | x]}{2\varepsilon} \quad (8)$$

The probability $\mathbb{P}[y^{exp} - \varepsilon \leq Y \leq y^{exp} + \varepsilon | x]$ is the probability that the response lies within the user-defined interval around the experimental value knowing x, which is the definition of the fidelity map. From this, we obtain the following important result for one set of experiments (i.e., $p = 1$):

$$\mathbf{f}_{Y(x,\mathbf{A})}(y^{exp}|x) \propto \mathbb{P}\left[(x,\mathbf{A}) \in FM\right] \quad (9)$$

which can be extended to p sets of experiments:

$$\mathbf{f}_{Y(x,\mathbf{A})}(y^{exp}|x) \underset{\sim}{\propto} \prod_{l=1}^{p} \mathbb{P}\left[(x,\mathbf{A}) \in FM^{(l)}\right] \quad (10)$$

where $\underset{\sim}{\propto}$ stands for "approximately proportional to" and $FM^{(l)}$ is the fidelity map associated with the l^{th} set of experiments $\mathbf{y}^{exp,(l)}$.

Note that a proportional quantity leads to equivalent results both for maximization purposes and Markov Chain algorithms. This probability can be efficiently estimated using Monte Carlo Sampling over the aleatory variables. As ε tends to zero, this probability tends to the likelihood value.

4.4. Estimators based on the fidelity map

In order to quantify the changes in confidence in the estimates between the prior and the posterior distributions when Bayesian updating is used, a *fidelity index* is introduced:

$$F_I = \frac{\mathbb{P}[(\mathbf{X},\mathbf{A}) \in FM|\mathbf{y}^{exp}]}{\mathbb{P}[(\mathbf{X},\mathbf{A}) \in FM]} \quad (11)$$

where $\mathbb{P}[(\mathbf{X},\mathbf{A}) \in FM|\mathbf{y}^{exp}]$ (resp. $\mathbb{P}[(\mathbf{X},\mathbf{A}) \in FM]$) is computed using the posterior distribution $\mathbf{f}_{\mathbf{X}}(\mathbf{x}|\mathbf{y}^{exp})$ (resp. the prior distribution $\mathbf{f}_{\mathbf{X}}(\mathbf{x})$).

5. Results

The proposed methodology is applied to the update of a FE dynamic model based on modal properties. Model parameter estimation based on modal properties such as natural frequencies and mode shapes, is a typical example where the number of outputs $\mathbf{y}$ is rather large [24, 2, 1, 3].

Results are obtained for a plate with uncertainty in the boundary conditions. We wish to identify the Young's modulus given a set of natural frequencies and mode shapes. For comparison purposes, the results are performed along with an approach based on a residual which encompasses all the responses in one quantity. In addition, the product of the individual likelihoods corresponding to the various responses (See Appendix Appendix B) is used in order to show the influence of the correlation between the model outputs.

5.1. Finite element Model updating based on modal data

Traditional quantities used in model update using modal properties are:

i Differences in natural frequencies values (e.g., Euclidean norm of difference). This quantity is traditionally minimized in the form of a residual through optimization.

ii Differences between the mode shapes. This is typically measured using the Modal Assurance Criterion (MAC) matrix [24, 1].

iii Differences between the Frequency Response Function (FRFs) measured using the FRAC (Frequency Response Assurance Criterion).

iv Mode orthogonality.

The MAC criterion (Eq. 12) is by far the most widely used:

$$M_{ij} = \frac{(\Phi_i^{*T} A \Phi_{exp,j})^2}{(\Phi_i^{*T} A \Phi_i)(\Phi_{exp,j}^{*T} A \Phi_{exp,j})} \quad (12)$$

where Φ_i is the i^{th} computational mode shape and "exp" stands for experimental. Φ_i^{*T} is the conjugate transpose of the mode shape. A is often the identity matrix or the mass matrix. The MAC value is equal to unity for a perfect match of modes. It should be as close to zero as possible for cross terms.

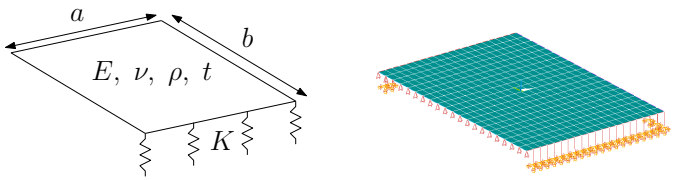
5.2. Test case description

Consider a simply supported rectangular plate modeled using finite elements. In order to model uncertainty in the displacement boundary conditions, one dimensional springs of stiffness K in the out-of-plane direction are used on three sides of the plate (Figure 3). The finite element model of the plate is constructed with 80 shell elements. We wish to identify the Young's modulus E of the plate based on $N_m = 4$ first modes for a total of $n = 14$ responses (4 natural frequencies and 10 MAC matrix terms). The parameters are summarized in Table 1.

In the absence of an actual experimental setup, "virtual experiments" are used by running the FE

Table 1: Parameters used for the plate example (S.I. units).

	Deterministic					To estimate	Aleatory
Parameter	a	b	ν	ρ	t	E	K
Value, Distribution	1	1.5	0.33	7800	0.01	N/A	$U(2\times10^5, 10^6)$



(a) Schematic representation.

(b) FEM representation.

Figure 3: Schematic and Finite Element representation of a simple plate. One side is simply supported while the others are connected to the ground through springs, to model uncertainties on the boundary conditions.

model with $E = E_{act}$ and $K = K_{act}$. The methodology is repeated for 6 combinations of E_{act} and K_{act} (see Table 2). All the configurations are run with $\varepsilon = 1\%$ for all the responses.

5.3. Fidelity map and likelihoods

The fidelity map is constructed in the (E, K) space with 15 Central Voronoi Tessellation (CVT) samples. CVT is chosen as a design of experiments as it fills the space uniformly. The boundary is then refined with 50 additional adaptive samples. For each value of E, the probability of being within the fidelity map (i.e., the approximated likelihood noted LH_{mcs}) is calculated with 10^5 Monte Carlo samples according to the distribution of K.

The proposed approach is compared to the results using the likelihood of the residual (LH_{res}) and the product of the individual likelihoods for the different responses LH_{prod} (c.f., Appendix B). These likelihoods are constructed using Kernel Smoothing [25] and Kriging models [26, 27, 28, 29] trained with 65 CVT samples for each of the $n = 14$ responses. LH_{res} uses a residual defined as follows:

$$R = \sum_{i=1}^{N_m} \left[\frac{(\lambda_i - \lambda_i^{exp})^2}{\lambda_i^{exp}} + (M_{ii} - 1)^2 + \sum_{j=1, j\neq i}^{N_m} M_{ij} \right]$$

The graphical results for the 6 configurations are depicted in Figure 4 and 5. Graphical inspection of the likelihoods show that LH_{mcs} exhibits a higher robustness than the two other methods for that example. The failure of the LH_{prod} is natural since the different natural frequencies are strongly correlated, therefore, the assumption of independence leads to incorrect results. On the other hand, the inaccuracy of LH_{res} is not straightforward. A loose explanation stems from the gathering of several responses that are correlated with different spreads within one quantity (similar to conclusion drawn in [3]).

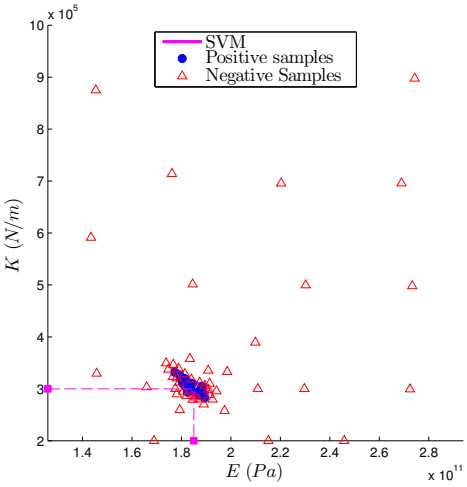
5.4. Maximum likelihood estimate

In the case where MLE is chosen for estimation, the results for the six cases are summarized in Table 2. As can be seen, the methodology is robust for this example.

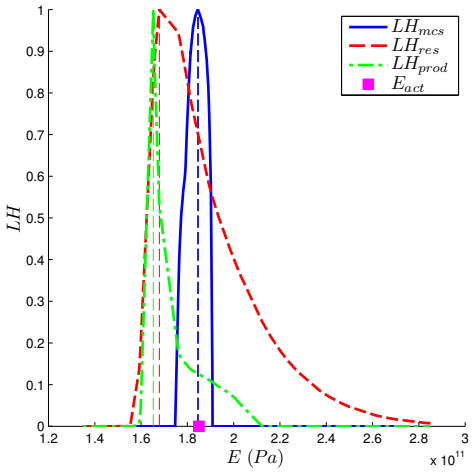
5.5. Bayesian update

In the case of Bayesian update, a wide prior, reflecting a significant lack of knowledge was chosen. The prior is set with a mean value of 210 GPa and a standard deviation of 21 GPa. Figure 6(a) depicts the likelihood function, the prior distribution, and the actual value, for the first case ($E_{act} = 185\times10^9\ Pa$ and $K_{act} = 3\times10^5\ N.m^{-1}$). Figure 6(b) shows the corresponding posterior distribution (Section 3.2). The Bayes estimators for the 6 cases are provided in Table 2.

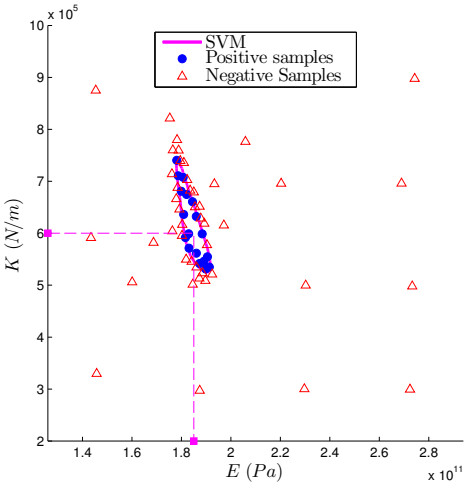
As an example of improvement brought by the Bayesian update in comparison to MLE, the fidelity index was computed for one case ($E_{act} =$



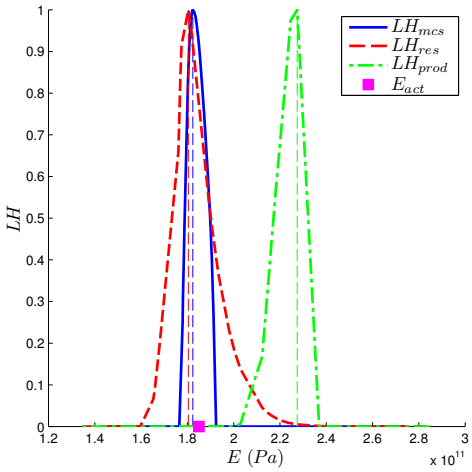
(a) Fidelity Map



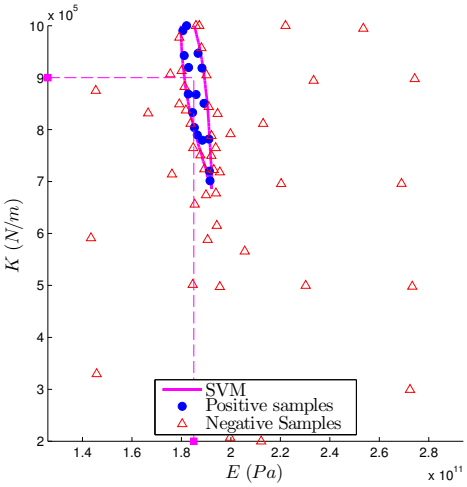
(b) Likelihood



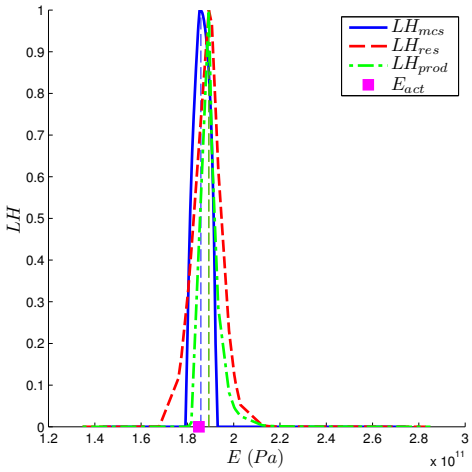
(c) Fidelity Map



(d) Likelihood

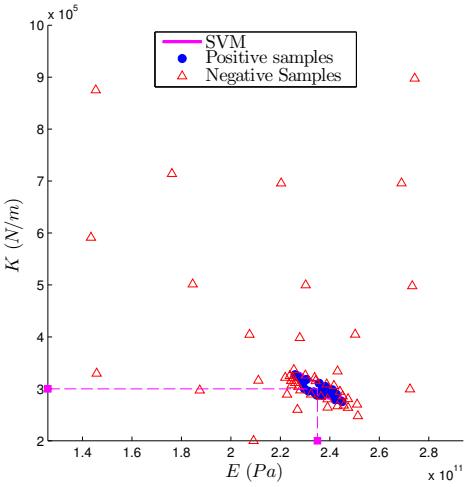


(e) Fidelity Map

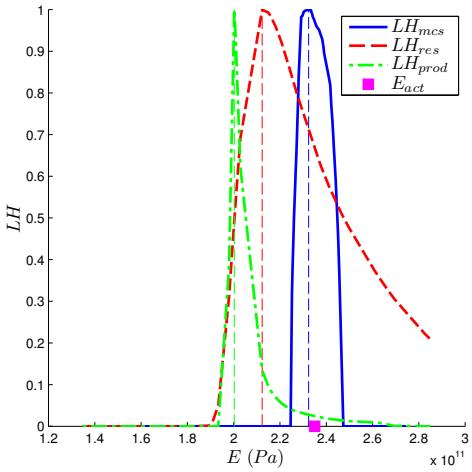


(f) Likelihood

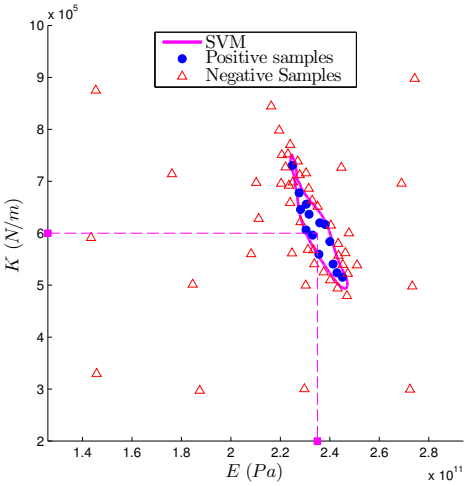
Figure 4: Graphical results of the plate example, showing the fidelity maps and the estimated likelihoods, for $E_{act} = 185 \times 10^9\ Pa$ and $K_{act} = 3 \times 10^5\ N.m^{-1}$ (a and b), $K_{act} = 6 \times 10^5\ N.m^{-1}$ (c and d), $K_{act} = 9 \times 10^5\ N.m^{-1}$ (e and f).



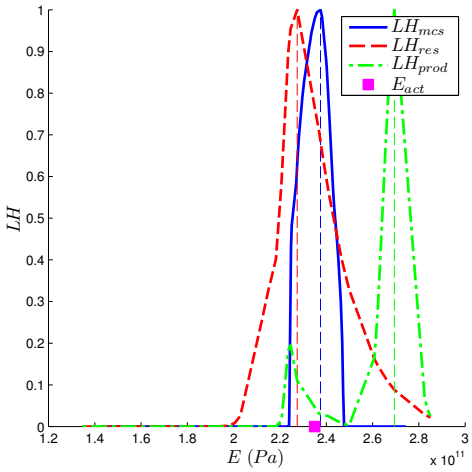
(a) Fidelity Map



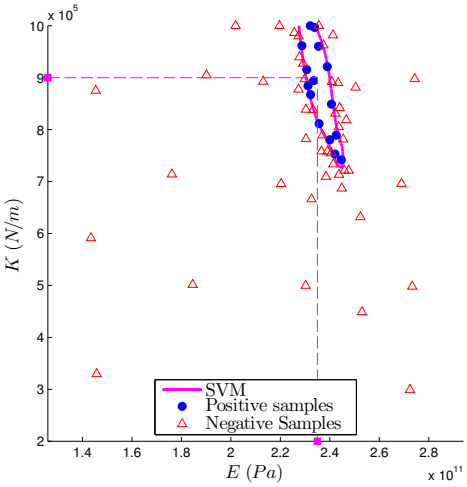
(b) Likelihood



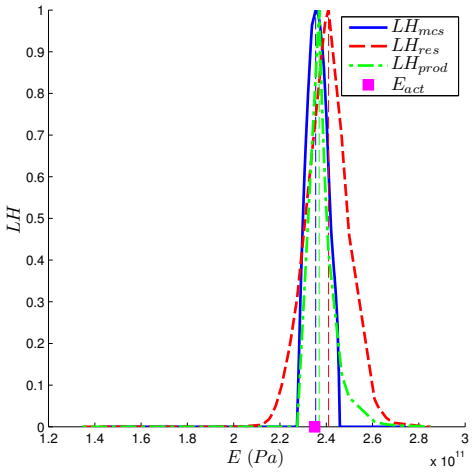
(c) Fidelity Map



(d) Likelihood



(e) Fidelity Map

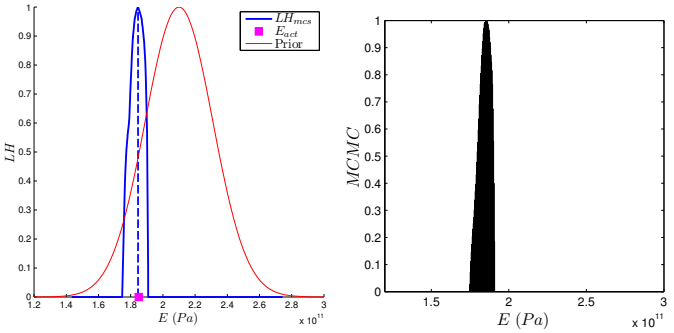


(f) Likelihood

Figure 5: Plate example. Fidelity maps and likelihoods for $E_{act} = 235 \times 10^9\ Pa$ and $K_{act} = 3 \times 10^5\ N.m^{-1}$ (a and b), $K_{act} = 6 \times 10^5\ N.m^{-1}$ (c and d), $K_{act} = 9 \times 10^5\ N.m^{-1}$ (e and f).

Table 2: Summary of the 6 experimental combinations and corresponding figures (S.I. units).

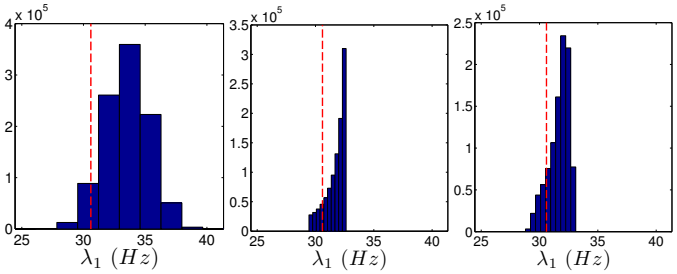
E_{act}	185×10^9			235×10^9		
K_{act}	3×10^5	6×10^5	9×10^5	3×10^5	6×10^5	9×10^5
Figures	4(a) & 4(b)	4(c) & 4(d)	4(e) & 4(f)	5(a) & 5(b)	5(c) & 5(d)	5(e) & 5(f)
E_{est}^{MLE} (Pa)	184.6×10^9	182.3×10^9	185.8×10^9	232.4×10^9	237.5×10^9	235.4×10^9
Error (%)	0.22	1.46	0.49	1.11	1.06	0.17
E_{est}^{Bayes} (Pa)	184.3×10^9	185.02×10^9	186.7×10^9	233.9×10^9	233.7×10^9	235.5×10^9
Error (%)	0.37	0.008	0.87	0.44	0.54	0.19



(a) Likelihood, prior knowledge, and actual parameter value E_{act}.

(b) Posterior through MCMC.

Figure 6: Bayesian update applied to the first case ($E_{act} = 185 \times 10^9\ Pa$ and $K_{act} = 3 \times 10^5\ N.m^{-1}$) of the plate example.



(a) Using prior.

(b) Ideal distribution.

(c) Using posterior.

Figure 7: First natural frequency distributions for first case ($E_{act} = 185 \times 10^9\ Pa$ and $K_{act} = 3 \times 10^5\ N.m^{-1}$) of the plate example. Uncertainty propagated using the prior, the posterior, and the ideal (unknown) distributions.

$185 \times 10^9\ Pa$ and $K_{act} = 3 \times 10^5\ N.m^{-1}$). A large value of 8.23 was obtained indicating that the posterior distribution gives roughly 8 times more chances to match the measurements.

In order to further gauge the benefits of the Bayesian update, the posterior was propagated to the first natural frequency (Figure 7(a)). For comparison, the prior was also propagated (Figure 7(b)). In addition the ideal, unknown, response distribution was computed (Figure 7(c)) using the actual value of the Young's modulus (E_{act}) along with the propagation of the aleatory variables (i.e., K). The figure clearly shows that the Bayesian update is closer to the actual response distribution.

6. Conclusion

An approach to perform model update using fidelity maps has been introduced. The construction of explicit fidelity maps using SVM in a space with parameter to estimate and aleatory uncertainties enables an efficient computation of the likelihood with Monte-Carlo simulations. In order to obtain an accurate boundary and reduce the number of model calls, an adaptive sampling scheme is used. Because SVM is a classification method, a large number of correlated model outputs can be used. This is done, and this is one of the strongest aspects of the fidelity maps approach, without the need for any assumptions on the correlation structure between the responses.

The next steps of this research will study the

scalability of the approach in higher dimensions. In addition, the approach will be tested on real world problems with actual experiments.

7. Acknowledgments

Support from the National Science Foundation (award CMMI-1029257) is gratefully acknowledged.

References

[1] Marwala T. Finite Element Model Updating Using Computational Intelligence Techniques: Applications to Structural Dynamics. Springer; 2010.

[2] Chen C, Duhamel D, Soize C. Probabilistic approach for model and data uncertainties and its experimental identification in structural dynamics: Case of composite sandwich panels. Journal of Sound and vibration 2006;294(1):64–81. doi: http://dx.doi.org/10.1016/j.jsv.2005.10.013.

[3] Gogu C, Haftka R, Le Riche R, Molimard J, Vautrin A, et al. Introduction to the bayesian approach applied to elastic constants identification. AIAA journal 2010;48(5):893–903. doi:10.2514/1.40922.

[4] Pratt J. Fy edgeworth and ra fisher on the efficiency of maximum likelihood estimation. The Annals of Statistics 1976;4(3):501–14. doi: 10.1214/aos/1176343457.

[5] Aldrich J. Ra fisher and the making of maximum likelihood 1912-1922. Statistical Science 1997;12(3):162–76. doi:10.1214/ss/1030037906.

[6] Press S, Press J. Bayesian statistics: principles, models, and applications. Wiley New York; 1989.

[7] Carlin B, Louis T. Bayes and empirical Bayes methods for data analysis; vol. 7. Springer; 1997.

[8] Jiang Z, Chen W, Apley DW. Preposterior analysis to select experimental responses for improving identifiability in model uncertainty quantification. In: Proceedings of the ASME 2013 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. 2013,.

[9] McFarland J, Mahadevan S, Romero V, Swiler L. Calibration and uncertainty analysis for computer simulations with multivariate output. AIAA journal 2008;46(5):1253–65.

[10] Gunn S. Support vector machines for classification and regression. ISIS technical report 1998;14.

[11] Vapnik V. The nature of statistical learning theory. Springer Verlag; 2000.

[12] Christianini N, Taylor S. An introduction to support vector machines (and othre kernel-based learning methods. Cambridge University Press; 2000.

[13] Schölkopf B, Smola A. Learning with kernels: Support vector machines, regularization, optimization, and beyond. the MIT Press; 2002.

[14] Basudhar A, Missoum S. An improved adaptive sampling scheme for the construction of explicit boundaries. Structural and Multidisciplinary Optimization 2010;42(4):517–29. doi: http://dx.doi.org/10.1007/s00158-010-0511-0.

[15] Kennedy M, O'Hagan A. Bayesian calibration of computer models. Journal of the Royal Statistical Society Series B, Statistical Methodology 2001;63:425–64. doi:http://www.jstor.org/stable/2680584.

[16] Xiong Y, Chen W, Tsui K, Apley D. A better understanding of model updating strategies in validating engineering models. Computer Methods in Applied Mechanics and Engineering 2009;198(15-16):1327–37. doi:http://dx.doi.org/10.1016/j.cma.2008.11.023.

[17] Zhang R, Mahadevan S. Model uncertainty and bayesian updating in reliability-based inspection. Structural Safety 2000;22(2):145–60.

[18] Mahadevan S, Rebba R. Validation of reliability computational models using bayes networks. Reliability Engineering & System Safety 2005;87(2):223–32.

[19] Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH. Equation of state calculations by fast computing machines. The Journal of Chemical Physics 1953;21(6):1087–92. doi: http://dx.doi.org/10.1063/1.1699114.

[20] Hastings W. Monte carlo sampling methods using markov chains and their applications. Biometrika 1970;57(1):97–. doi:10.1093/biomet/57.1.97.

[21] Nichols J, Moore E, Murphy K. Bayesian identification of a cracked plate using a population-based markov chain monte carlo method. Computers & Structures 2011;89:1323–32. doi: http://dx.doi.org/10.1016/j.compstruc.2011.03.013.

[22] Basudhar A, Missoum S, Harrison Sanchez A. Limit state function identification using support vector machines for discontinuous responses and disjoint failure domains. Probabilistic Engineering Mechanics 2008;23(1):1–11. doi: http://dx.doi.org/10.1016/j.probengmech.2007.08.004.

[23] Basudhar A, Missoum S. Adaptive explicit decision functions for probabilistic design and optimization using support vector machines. Computers & Structures 2008;86(19-20):1904–17. doi: http://dx.doi.org/10.1016/j.compstruc.2008.02.008.

[24] Allemang R. The modal assurance criterion (mac): twenty years of use and abuse. In: Proceedings, International Modal Analysis Conference. 2002,.

[25] Bowman A, Azzalini A. Applied smoothing techniques for data analysis: the kernel approach with S-Plus illustrations; vol. 18. Oxford University Press, USA; 1997.

[26] Sacks J, Welch W, Mitchell T, Wynn H. Design and analysis of computer experiments. Statistical science 1989;4(4):409–23. URL http://www.jstor.org/stable/2245858.

[27] Jones D. A taxonomy of global optimization

methods based on response surfaces. Journal of Global Optimization 2001;21(4):345–83. doi: http://dx.doi.org/10.1023/A:1012771025575.

[28] Forrester A, Keane A. Recent advances in surrogate-based optimization. Progress in Aerospace Sciences 2009;45(1-3):50–79. doi: http://dx.doi.org/10.1016/j.paerosci.2008.11.001.

[29] Basudhar A, Dribusch C, Lacaze S, Missoum S. Constrained efficient global optimization with support vector machines. Structural and Multidisciplinary Optimization 2012;46(2):201–21. doi:10.1007/s00158-011-0745-5. URL http://dx.doi.org/10.1007/s00158-011-0745-5.

Appendix A. Sequential uncertainty propagation

If one does not use the fidelity map approach, the approximation of a likelihood can be done as follows. For each given $\mathbf{x^{(k)}}$, the uncertainties due to $\mathbf{A}$ are propagated in order to fit $\mathbf{f_{Y(x,A)}}(\mathbf{y}|\mathbf{x}^{(k)})$. The probability density corresponding to the measurement $\mathbf{y}^{exp}$ is then read (Figure A.8 for x and y^{exp} scalar). For expensive black box models, this process requires two layers of surrogates (e.g., Kriging [26, 27, 28, 29]), one on the responses $\mathbf{y}$ to propagate the uncertainties, the other on the conditional joint PDFs $\mathbf{f_{Y(x,A)}}(\mathbf{y}|\mathbf{x}^{(k)})$ for any given $\mathbf{x^{(k)}}$. This approach is impractical in the case where the number of responses $\mathbf{y}$ is large.

For the sake of comparison to the proposed fidelity map approach, this process was applied to the plate problem (Section 5). The purpose was to estimate the parameters using the traditional residual-based approach or the product of likelihoods as described in the next section. Due to the fact that the conditional PDFs $\mathbf{f_{Y(x,A)}}(\mathbf{y}|\mathbf{x}^{(k)})$ are 1-dimensional in the plate problem, kernel smoothing is used [25].

Appendix B. Product of the likelihoods and Residual

For comparison with the proposed approach, this appendix introduces two other methods to compute a scalable likelihood approximation. These two approaches are used in the results Section 5.

Appendix B.1. A residual-based method

A seemingly intuitive approach is to use a residual that combines n responses into one quantity:

$$R(\mathbf{x}, \mathbf{a}) = \sum_{i=1}^{n} \left(\frac{y_i^{exp} - y_i(\mathbf{x}, \mathbf{a})}{y_i^{exp}} \right)^2 \tag{B.1}$$

This approach is implemented to provide a comparison with the method introduced in 4.3. A possible MLE then reads:

$$\mathbf{x^{MLE}} = \underset{\mathbf{x}}{\mathbf{argmax}}\ \mathbf{f}_{R(\mathbf{x},\mathbf{A})}(0|\mathbf{x}) \tag{B.2}$$

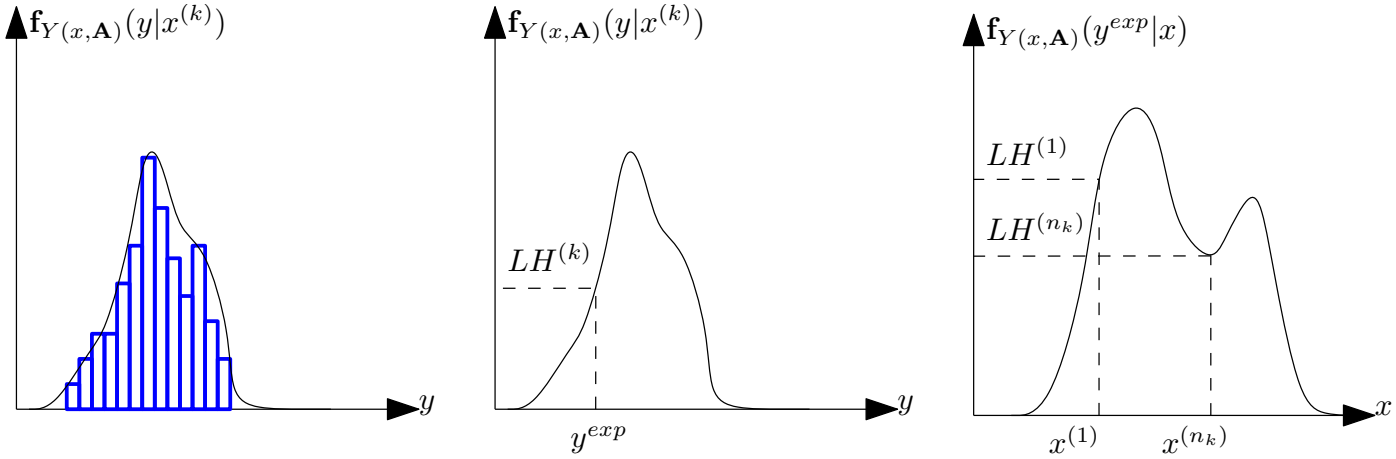
It is worth mentioning that this method does not compute exactly the likelihood as described in the results section.

Appendix B.2. Product of the likelihood formulation

A widely used approach is to consider the likelihood as a product of "marginal" likelihoods. However this approach relies on a strong assumption (usually wrong) that the measurements are independent. The formulation of the product of the likelihood is straightforward:

$$\mathbf{x^{MLE}} = \underset{\mathbf{x}}{\mathbf{argmax}} \prod_{i=1}^{n_m} \mathbf{f}_{Y_i(\mathbf{x},\mathbf{A})}(y_i^{exp}|\mathbf{x}) \tag{B.3}$$

The only appeal of this approach is that the "marginal" likelihoods are 1-dimensional and can be computed as explained in Appendix A.



(a) For $x^{(k)}$, propagate the uncertainty to y, due to $\mathbf{A}$ and fit the PDF on top of it.

(b) Read the value of the conditional PDF for $y = y^{exp}$ ($LH^{(k)}$).

(c) Repeat the process for $k = 1..n_k$ in order to construct the likelihood over the design space.

Figure A.8: Description of the likelihood construction using sequential uncertainty propagation.