Simulation-based optimum Bayesian experiment design for high-voltage conductors

Carlos Frederico T. Matt  
Electric Energy Research Center (CEPEL), Department of Lines and Substations  
Rio de Janeiro, RJ, Brazil  
E-mail: cfmatt@cepel.br

Daniel Alves Castello  
Federal University of Rio de Janeiro, Department of Mechanical Engineering  
Poli/COPPE/UFRJ, Rio de Janeiro, RJ, Brazil  
E-mail: castello@mecanica.coppe.ufrj.br

William Pinto Hernandez  
Electric Energy Research Center (CEPEL), Department of Laboratories  
Rio de Janeiro, RJ, Brazil  
E-mail: hernandez-bolsista@cepel.br

Abstract. The aim of the present work is to apply a Bayesian framework coupled with information theoretic design criteria based on the expected gain in Shannon information from the prior to the posterior distribution of the model parameters in order to design “optimum” experiments. A computational model is proposed to predict the dynamic behavior of typical high-voltage conductors. The proposed optimum experiment design problem may be stated as follows: given a fixed number of sensors (accelerometers) to measure the conductor’s dynamic response on a laboratory span and a computational model to map the model parameters to observable quantities of interest, where should one install the sensors to extract the maximum available information for a more accurate model calibration and uncertainty quantification?

1. Introduction

Nowadays, computational models have been increasingly used for inference, prediction, control and optimization of engineering structures as, for instance, high-voltage transmission line conductors. These structures are stranded cables commonly composed by aluminum wires helically wrapped around a steel core with either seven or nineteen wires. They are commonly referred to as Aluminum Conductor Steel Reinforced (ACSR) conductors. In the field, they are strung to relatively high tensile loads (in the range from 10% to 20% of their rated tensile strength) and clamped at the suspension towers. Mechanical vibrations on ACSR conductors are induced by the wind due to the well-known vortex-shedding effect [1].

Many different computational models have been proposed to predict the dynamic behavior of ACSR conductors in the field. The energy balance principle has been the most widely used
theoretical model adopted in the literature to predict the maximum amplitude of vibration in steady-state regime as a function of the excitation frequency (defined in terms of the incoming wind velocity through the well-known Strouhal relationship) [2, 3]. Nevertheless, it requires knowledge about the energy dissipated by a stranded cable during its bending vibrations, which is quite difficult to measure even in a controlled experiment performed on a laboratory span. On the other hand, the much less used equivalent beam model has been demonstrated by some researchers [2, 3] to accurately reproduce the measured cable’s dynamic behavior both in frequency and time domains. Notwithstanding, it requires knowledge about cable’s bending stiffness and also on damping parameters. Owing to the complicated helicoidal geometry of a stranded cable, uncertainties in both its bending stiffness and damping parameters should be accounted for.

One essential step to build an accurate computational model for reliable prediction, control and optimization tasks is to calibrate its parameters. Calibration requires thus two essential components: (i) a computational model to predict an observable behavior and (ii) measured data for that observable behavior. Computational models for complex structures subject to complex physical phenomena, as high-voltage conductors excited by vortex shedding, are simplifications of the true physics and its predictive behavior contains thus an amount of uncertainty. Data quantifying an observable behavior are also uncertain due to unavoidable measurement errors, even in controlled experiments performed under essentially the same conditions. Hence, model calibration approaches that take into account the two aforementioned sources of uncertainty are highly desirable.

The design of an experimental set-up to extract the maximum information from the data set to accurately calibrate model parameters is also a critical issue. Several definitions for experimental design are available in the literature. Chaloner and Verdinelli [4] stated that experimental design involves the specification of all aspects of an experiment. According to those authors, available resources and knowledge of the goals for carrying out the experiment help in selecting important features that depend on the specific problem at hand. According to Huan and Marzouk [5], experimental design encompasses questions of where and what to measure, which variables to interrogate, and what experimental conditions to employ in order to extract the maximum available information from the data. Emery and Nenarokomov [6] defined optimum experiment design as the conditions under which an experiment is conducted (collectively named as experimental protocol by the authors) in order to maximize the accuracy of the estimated parameters. Here, it is implicitly assumed by those authors that the experimental goal is to infer the parameters of a computational model proposed to predict an observable behavior of the system under investigation. In designing an experiment, decisions must be made before data collection, and data collection is restricted by limited resources. As prior information is usually available to the experimenter and also motivates the experiment, Bayesian statistics can play an important role. The Bayesian statistical framework may also be particularly attractive for model calibration because it provides a rigorous mathematical tool for inference from noisy, indirect and incomplete data and a natural mechanism to incorporate uncertainties from heterogeneous sources of information.

Dynamic tests performed on a laboratory span with high-voltage conductors are particularly time-consuming and expensive to conduct. Hence, maximizing the value of experimental data by designing experiments to be “optimal” with respect to some appropriate measure may significantly increase the efficiency of the modeling and calibration processes and data analysis. The aim of the present work is thus to apply a Bayesian framework coupled with information theoretic design criteria based on the expected gain in Shannon information from the prior to the posterior distribution of the model parameters in order to design “optimum” experiments. The proposed optimum experiment design problem may be stated as follows: given a fixed number of sensors (accelerometers) to measure the conductor’s dynamic response on a laboratory span...
and a computational model to map the model parameters to observable quantities, where does one install the sensors to extract the maximum available information for a more accurate model calibration and uncertainty quantification?

2. General Experiment Design Formulation

2.1. Experimental goals

The first task to perform an optimal experiment design is to define the design criterion or the objective function which should reflect how valuable or relevant an experiment is expected to be. In order to specify the design criterion one must have in advance the answer to the question: what does one intend to do with results of the experiment? In other words, the experimental goals must be clearly defined prior to planning or designing the experiments. For instance, if one performs an experiment to estimate a physical constant, a good design criterion should reflect the uncertainty in its inferred values. On the other hand, if one is interested in making model predictions then a good criterion should reflect the uncertainty in model outputs. In the current work, one assumes that the experimental goal is to infer a finite number of model parameters of a computational model proposed to predict the dynamic behavior of a particular mechanical system, namely, a high-voltage conductor.

The design criterion developed in what follows uses a full information theoretic formalism [5] and can also be extended to other experimental goals. Penalties reflecting experimental effort or cost and/or restrictions associated with resource constraint, for instance, may also be included in the experimental design optimization problem. Nevertheless, for simplicity and since costs and constraints are problem-specific, they are not taken into account in this work.

2.2. Design criterion and expected utility

The experimental design criterion will be mathematically formulated in a Bayesian framework. Bayesian statistics offers a rigorous mathematical foundation for inference from noisy, indirect and incomplete data. It allows one to incorporate physical constraints and/or other sources of information (reflected in the prior knowledge) and also to quantify the uncertainty in parameters and model predictions. In Bayesian statistics, unknown parameters are treated as random variables and uncertainty is quantified by means of probability distributions. There is also a natural link between Bayesian statistics and statistical decision theory [7], which will be exploited to derive the experiment design criterion.

Let the triplet $(\Omega, \mathcal{F}, P)$ denote a probability space, where $\Omega$ denotes a sample space, $\mathcal{F}$ denotes a $\sigma$-algebra and $P$ is a probability measure on $(\Omega, \mathcal{F})$. Let the vector of real-valued random variables $\theta : \Omega \to \mathbb{R}^n_\theta$ denote the uncertain parameters of interest, i.e., the parameters to be conditioned on experimental data. The random vector $\theta$ is equipped with a continuous probability density function $p_{\theta}(\theta)$, which is assumed to always exist. The observable data $y$ is treated as an $\mathbb{R}^{n_y}$-valued random vector endowed with a continuous probability density function $p_Y(y)$. The data $y$ should be understood as one possible physical realization of the underlying continuous random process $\{Y(t)\}$ at discrete ‘time’ instants $t_k, k = 1, 2, \ldots, \infty$. The vector $d \in \mathbb{R}^{n_d}$ denotes the design variables or experimental conditions to be optimized. Unlike the vectors $\theta$ and $y$, the vector $d$ is deterministic. The symbols $n_\theta$, $n_y$ and $n_d$ denote, respectively, the number of uncertain model parameters, the number of observations and the number of design variables.

Assuming that one performs an experiment at conditions given by $d$ and observes a realization of the data given by the vector $y$, the change in the state of knowledge about the model parameters $\theta$ may be quantified by the posterior probability distribution $p_{\theta|Y}(\theta|y, d)$, given by Bayes’ rule as

$$p_{\theta|Y}(\theta|y, d) = \frac{p_Y(y|\theta, d) p_{\theta}(\theta|d)}{p_Y(y|d)} \quad (1)$$
where \( \pi_{\theta}(\theta|d) \) denotes the prior probability distribution of the model parameters; \( \pi_{\theta|Y}(\theta|y,d) \) denotes the likelihood function, i.e., the probability distribution of the data conditioned on the parameters, and \( \pi_{Y}(y|d) \) denotes the marginal probability distribution of the data (the evidence), for a given design \( d \). It is reasonable to assume that the prior distribution on the parameters does not depend upon the design variables; hence, \( \pi_{\theta}(\theta|d) = \pi_{\theta}(\theta) \).

Based on a decision theoretic approach, Lindley [8] proposed that an experiment design criterion should have the following general form

\[
U(d) = \int_{D_{\theta}} \int_{D_{Y}} u(d, y, \theta) \pi_{\theta|Y}(\theta|y,d) \, dy \, d\theta
\]

\[
= \int_{D_{\theta}} \int_{D_{Y}} u(d, y, \theta) \pi_{\theta} \pi_{Y}(y|d) \, d\theta
\]

\[
= \int_{D_{\theta}} \int_{D_{Y}} u(d, y, \theta) \pi_{\theta}(\theta) \pi_{Y}(y|d) \, d\theta
\]  
(2)

where \( u(d, y, \theta) \) is a utility function reflecting the purpose of the experiment performed at conditions \( d \), given a particular value of the parameters \( \theta \) and a particular realization \( y \); \( U(d) \) is the expected utility; and \( D_{\theta} \) and \( D_{Y} \) denote, respectively, the supports of the distributions \( \pi_{\theta}(\theta) \) and \( \pi_{Y}(y|d) \). As the precise values of the uncertain parameters \( \theta \) are not known and the outcomes \( y \) cannot be known before the experiment is performed, the utility function must be averaged over the joint distribution of \( \theta \) and \( y \). The last equality on Eq. (2) arises upon substitution of Eq. (1) for the posterior probability distribution \( \pi_{\theta|Y}(\theta|y,d) \).

Several choices for utility functions are available, depending on the goals of the experiment [4]. Our choice of utility function is identical to that used by Huan and Marzouk [5], which is appropriate when the goal of the experiment is to infer the values of uncertain parameters \( \theta \). Based on this choice, the utility function may be written as

\[
u(d, y, \theta) = \int_{D_{\theta}} \pi_{\theta}(\theta) \ln \left( \frac{\pi_{\theta}(\theta)}{\pi_{\theta}(\xi)} \right) d\xi = u(d, y). \]  
(3)

The utility function given by Eq. (3) involves an internal integration over the parameter space (the variable \( \xi \) is a dummy variable of integration representing the uncertain model parameters) and therefore it is not a function of the parameters \( \theta \). By substituting Eq. (3) into the second equality of Eq. (2) yields

\[
U(d) = \int_{D_{Y}} \int_{D_{\theta}} \pi_{\theta}(\theta|y,d) \, d\theta \, u(d, y) \pi_{Y}(y|d) \, dy
\]

\[
= \int_{D_{Y}} u(d, y) \pi_{Y}(y|d) \, dy
\]  
(4)

because by definition of a probability density function, the integral of \( \pi_{\theta|Y}(\theta|y,d) \) over the support \( D_{\theta} \) is identical to unity. Now, by inserting Eq. (3) into Eq. (4) yields (after changing

\[
D_{KL}(\pi_{A}(\theta)||\pi_{B}(\theta)) = \int_{D_{\theta}} \pi_{A}(\theta) \ln \left( \frac{\pi_{A}(\theta)}{\pi_{B}(\theta)} \right) d\theta
\]

\footnote{Given generic probability distributions \( \pi_{A}(\theta) \) and \( \pi_{B}(\theta) \) the Kullback-Leibler divergence from \( \pi_{A}(\theta) \) to \( \pi_{B}(\theta) \), \( D_{KL}(\pi_{A}(\theta)||\pi_{B}(\theta)) \) is defined as}
the dummy variable $\xi$ by $\theta$)

\[
U(\mathbf{d}) = \int_{D_Y} \int_{D_\theta} \pi_{\Theta|Y}(\theta|y, \mathbf{d}) \ln \left( \frac{\pi_{\Theta|Y}(\theta|y, \mathbf{d})}{\pi_{\Theta}(\theta)} \right) \pi_Y(y|\mathbf{d}) \, d\theta \, dy
\]

\[
= \int_{D_Y} \left[ \int_{D_\theta} \pi_{\Theta|Y}(\theta|y, \mathbf{d}) \ln \left( \frac{\pi_{\Theta|Y}(\theta|y, \mathbf{d})}{\pi_{\Theta}(\theta)} \right) \, d\theta \right] \pi_Y(y|\mathbf{d}) \, dy
\]

\[
= E_{y|\mathbf{d}[D_{KL}(\pi_{\Theta|Y}(\theta|y, \mathbf{d})||\pi_{\Theta}(\theta))].
\]

(5)

Hence, the *expected utility* is mathematically the expected value (ensemble average), under the probability distribution of the data, of the Kullback-Leibler divergence from the posterior to the prior distributions of the unknown model parameters. The *expected utility* is also referred to as the *expected information gain* in $\theta$. Larger KL divergences from posterior to the prior implies that the data $y$ are thus more informative for parameter inference as they decrease the entropy in $\theta$ by larger amounts [5]. As we generally have only one realization $y$ of the observable data, the information gain should be maximized on average. Huan and Marzouk [5] demonstrated that the expected utility is mathematically equivalent to the mutual information between the parameters $\theta$ and the data $y$. When applied to a linear Gaussian design problem, $U(\mathbf{d})$ reduces to the Bayesian D-optimality condition [5]. To sum up, in order to find the optimum experiment design conditions, denoted by $\mathbf{d}^*$, one must find the maximum of the expected utility $U(\mathbf{d})$ over the design space $D$; in mathematical terms,

\[
\mathbf{d}^* = \arg \max_{\mathbf{d} \in D} U(\mathbf{d}).
\]

(6)

The optimum experiment design problem becomes thus an optimization (maximization) problem. At a first glance, it is an unconstrained optimization problem; nevertheless, one should include penalization terms to account for experimental effort and/or cost, as previously mentioned. Here, no penalty terms have been added.

2.3. Numerical evaluation of expected utility

As discussed in the previous section, the optimum experiment design problem is mathematically formulated as the maximization of the expected utility. Hence, the main task to perform is to evaluate efficiently (i.e., without too much computational effort) the expected utility $U(\mathbf{d})$. The expected utility given by Eq. (5) generally involves integration in high-dimensional spaces and typically has no closed form specially when the map $G : (\theta, \mathbf{d}) \in \mathbb{R}^{n_\theta} \times \mathbb{R}^{n_d} \rightarrow y \in \mathbb{R}^{n_y}$ is nonlinear; therefore, $U(\mathbf{d})$ must be approximated numerically. The expression for the expected utility may be rewritten as

\[
U(\mathbf{d}) = \int_{D_Y} \int_{D_\theta} \pi_{\Theta|Y}(\theta|y, \mathbf{d}) \ln \left( \frac{\pi_{\Theta|Y}(\theta|y, \mathbf{d})}{\pi_{\Theta}(\theta)} \right) \pi_Y(y|\mathbf{d}) \, d\theta \, dy
\]

\[
= \int_{D_Y} \left[ \int_{D_\theta} \ln \left( \frac{\pi_{\Theta|Y}(\theta|y, \mathbf{d})}{\pi_Y(y|\theta, \mathbf{d})\pi_{\Theta}(\theta)} \right) \pi_Y(y|\theta, \mathbf{d}) \, d\theta \right] \pi_Y(y|\mathbf{d}) \, dy
\]

\[
= \int_{D_Y} \int_{D_\theta} \{ \ln[\pi_{\Theta|Y}(y|\theta, \mathbf{d})] - \ln[\pi_Y(y|\mathbf{d})] \} \pi_{\Theta|Y}(y|\theta, \mathbf{d}) \pi_{\Theta}(\theta) \, d\theta \, dy
\]

(7)

by inserting the Bayes’ expression given by Eq. (1) for the posterior probability distribution both inside and outside the logarithm. As noticed by Huan and Marzouk [5], for the special case in which the Shannon entropy of $\pi_{\Theta|Y}(y|\theta, \mathbf{d})$ is independent of the design variables $\mathbf{d}$, the first term in Eq. (7) becomes a constant and can be dropped from the objective function. Maximizing the remaining term, which is the entropy of $\pi_Y(y|\mathbf{d})$, is equivalent to the maximum entropy
sampling approach of Sebastiani and Wynn [9]. Several methods may be used to compute an estimate for the (general) high-dimensional integrals defined in Eq. (7). Here, Monte Carlo sampling is employed to estimate the expected utility; hence, 

$$U(d) \approx \frac{1}{N_{\text{out}}} \sum_{i=1}^{N_{\text{out}}} \left\{ \ln[\pi_Y(y^{(i)}|\theta^{(i)}, d)] - \ln[\pi_Y(y^{(i)}|d)] \right\}$$

(8)

where the samples $\theta^{(i)}$ and $y^{(i)}$, $i = 1, 2, \ldots, N_{\text{out}}$, are drawn, respectively, from the prior $\pi_\Theta(\theta)$ and the likelihood $\pi_{Y|\Theta}(y|\theta, d)$ distributions; and $N_{\text{out}}$ denotes the number of samples in this “outer” Monte Carlo estimate [5].

It is noteworthy to emphasize that models enter the formulation through the likelihood distribution. Henceforth, one assumes an additive error between the experimental observations and model predictions (which has been largely adopted in many scientific and engineering problems), i.e.,

$$y = G(\theta, d) + \epsilon$$

(9)

where the additive error $\epsilon$ is a $\mathbb{R}^n_{\text{y}}$ vector-valued random variable equipped with a (general) known probability density $\pi_\epsilon(\epsilon)$; the form of this probability density is left unspecified for now. A realization of the data $y^{(i)}$ may thus be obtained from Eq. (9), i.e.,

$$y^{(i)} = G(\theta^{(i)}, d) + \epsilon^{(i)},$$

where $\epsilon^{(i)}$ is a realization of the error drawn from $\pi_\epsilon(\epsilon)$. Hence, drawing a realization from the likelihood requires evaluating the forward model $G$ at a particular $(\theta, d)$. Note also that for the additive discrepancy model given by Eq. (9), $\pi_{Y|\Theta}(y|\theta, d) = \pi_\epsilon(y - G(\theta, d))$; therefore, evaluating the likelihood distribution at a particular realization $(y^{(i)}, \theta^{(i)})$ for a given $d$ also requires evaluating $G$.

With the additive error model, the Monte Carlo estimate for the expected utility may thus be rewritten as follows

$$U(d) \approx \frac{1}{N_{\text{out}}} \sum_{i=1}^{N_{\text{out}}} \left\{ \ln[\pi_E(y^{(i)} - G(\theta^{(i)}, d))] - \ln[\pi_Y(y^{(i)}|d)] \right\}$$

(10)

It remains to compute the evidence evaluated at $y^{(i)}$, $\pi_Y(y^{(i)}|d)$, which typically does not have an analytical closed-form. Nevertheless, it may also be approximately computed using an “inner” Monte Carlo estimate

$$\pi_Y(y^{(i)}|d) = \int_{\mathcal{D}_\Theta} \pi_{Y|\Theta}(y^{(i)}|\theta, d) \pi_\Theta(\theta) \, d\theta \approx \frac{1}{N_{\text{in}}} \sum_{j=1}^{N_{\text{in}}} \pi_{Y|\Theta}(y^{(i)}|\theta^{(i,j)}, d)$$

$$\approx \frac{1}{N_{\text{in}}} \sum_{j=1}^{N_{\text{in}}} \pi_\epsilon(y^{(i)} - G(\theta^{(i,j)}, d))$$

(11)

where $\theta^{(i,j)}$ are also drawn from the prior distribution $\pi_\Theta(\theta)$ and $N_{\text{in}}$ is the number of samples in this “inner” Monte Carlo sum. According to Huan and Marzouk [5], the combination of Eqs. (10) and (11) yields a biased estimator $\hat{U}(d)$ of $U(d)$, with the bias controlled by $N_{\text{in}}$ and the variance controlled by $N_{\text{out}}$ [5]. Evaluating and sampling from the likelihood for each new sample of $\theta$ constitutes the most significant computational cost. In order to reduce the computational cost of the nested Monte Carlo estimator, the same procedure employed by Huan and Marzouk [5] is adopted here. For every $d$, one draws a fresh batch of prior samples $\{\theta^{(k)}\}_{k=1}^{N_{\text{out}}}$ and uses this set for both the “outer” Monte Carlo sum (i.e., $\theta^{(i)} = \theta^{(k)}$) and all the “inner” Monte Carlo sums at that $d$ (i.e., $\theta^{(i,j)} = \theta^{(k)}$, and consequently $N_{\text{out}} = N_{\text{in}}$). Such a procedure reduces
the computational cost for a fixed \( d \) from \( O(N_{\text{in}}N_{\text{out}}) \) to \( O(N_{\text{out}}) \) [5]. The main advantage of sample reuse is that it avoids near-zero evidence estimates at small sample sizes [5]; on the other hand, the major shortcoming is that it contributes to the bias of the estimator but Huan [10] has shown that this effect is very small.

The results reported in Section 4 were computed with the general formulation described in this section but for the particular case in which \( \pi_{\mathcal{E}}(\epsilon) \) is a multivariate Gaussian distribution with zero mean value and a diagonal covariance matrix \( \Sigma = \sigma^2 I \), with \( I \) being the \( n_y \times n_y \) identity matrix and \( \sigma^2 \) the constant variance of the measurement error. Extension to other probability distributions for the additive error \( \epsilon \) is straightforward.

3. Application: Mechanical Vibrations on High-Voltage Conductors

3.1. Description of the experiment

On a laboratory test, both ends of a conductor are rigidly clamped at screwed fixtures mounted on rigid concrete blocks. One of the ends is coupled to a tighter and a load cell in order to apply and measure the tensile load. At the other end, an electrodynamic shaker is connected to the conductor in order to transmit the excitation. The distance between the two clamped ends defines the span length. Both a force transducer and an accelerometer are placed at the driving point for measurement and acquisition of the input excitation transmitted to the conductor and its response acceleration. One, two or more accelerometers are also placed along the conductor in order to measure and collect its response at different positions. By measuring the force transmitted to the conductor and its response one may compute, for instance, frequency response functions in terms of the excitation frequency and for a given position. Observable data on a laboratory test are either accelerations in time domain or frequency response functions. Here, time-domain accelerations comprise the observable data. Figure 1 illustrates a sketch of the experimental set-up for dynamic tests performed on a laboratory span with a typical high-voltage conductor and an overview of the laboratory span at Electric Energy Research Center (CEPEL).

![Figure 1. Sketch of the experimental set-up for dynamic tests on a laboratory span with a typical high-voltage conductor (left) and an overview of CEPEL’s laboratory span (right).](image)
of accelerometers to be placed along the conductor, where does one should place them in order to maximize the information gain on the unknown parameters?

3.2. Forward model and hybrid analytico-numerical solution

The mechanical vibrations of a typical high-voltage conductor on a laboratory span may be mathematically described by the following general partial differential equation

\[
\frac{\partial}{\partial x^2} \left( EI(x) \frac{\partial^2 w}{\partial x^2} \right) - \frac{\partial}{\partial x} \left( T(x,t) \frac{\partial w}{\partial x} \right) + \mu(x) \frac{\partial^2 w}{\partial t^2} = f(x,t) \tag{12}
\]

subjected to the boundary conditions \( w(0,t) = w'(0,t) = 0 \) and \( w(L,t) = w'(L,t) = 0 \) at \( x = 0 \) and \( x = L \), and to zero initial conditions \( w(x,0) = \frac{\partial w}{\partial t} |_{t=0} = 0 \); the prime denotes differentiation with respect to the \( x \)-variable. Hence, one assumes that both conductor’s ends are rigidly-clamped and that it is initially at rest. The independent variables are the position \( x \) and time \( t \) while the dependent variable is the conductor’s transverse displacement \( w(x,t) \). The symbols \( EI(x) \) and \( \mu(x) \) stand for, respectively, the conductor’s bending stiffness and mass per unit length (assumed to vary along the conductor’s length); \( T(x,t) \) denotes the tensile load applied to the conductor; \( L \) denotes the span length; and, finally, \( f(x,t) \) stands for the external excitation force.

The particular case in which the coefficients \( EI(x) \), \( T(x,t) \) and \( \mu(x) \) are constants has been deeply investigated by Castello and Matt [2], for an harmonic concentrated excitation \( f(x,t) = f_0 \delta(x-x_s) e^{i\omega t} \), with \( \delta(x-x_s) \) being the Dirac delta distribution of the first kind, \( x_s \) the excitation point \( (0 < x_s < L) \) and \( \omega \) the excitation frequency. Here, one retains the more general formulation with spatial- and/or temporal-dependence of the differential equation coefficients. One also assumes a concentrated sine chirp excitation defined by \( f(x,t) = f_0 \delta(x-x_s) \sin \left[ \omega(t) t \right] \) with a linear increase in the excitation frequency from \( \omega_l \) to \( \omega_u \) in \( t_f \) seconds: \( \omega(t) = \omega_l + (\omega_u - \omega_l) \frac{t}{t_f} \). The forward model just described will be rewritten in dimensionless form in what follows.

By defining the dimensionless independent and dependent variables \( X = x/L, \tau = \omega_0 t \) and \( W(X,\tau) = w(x,t)/D \), where \( \omega_0 \) and \( D \) denote, respectively, a characteristic frequency for the problem under investigation (the oscillation frequency of the tensile load, as described in the next paragraphs) and the conductor’s diameter, one may rewrite the initial- and boundary-value problem for the dimensionless transverse displacement \( W(X,\tau) \) as follows

\[
\frac{\partial}{\partial X^2} \left( A(X) \frac{\partial^2 W}{\partial X^2} \right) - \frac{\partial}{\partial X} \left( B(X,\tau) \frac{\partial W}{\partial X} \right) + C(X) \frac{\partial^2 W}{\partial \tau^2} = F_0 \delta(X-X_s) \sin \left[ \Omega(\tau) \tau \right] \tag{13}
\]

\[
W(0,\tau) = W'(0,\tau) = 0 \quad \text{at} \quad X = 0, \tau > 0 \tag{14}
\]

\[
W(1,\tau) = W'(1,\tau) = 0 \quad \text{at} \quad X = 1, \tau > 0 \tag{15}
\]

\[
W(X,0) = \frac{\partial W}{\partial \tau} \bigg|_{\tau=0} = 0 \quad \text{at} \quad \tau = 0, 0 \leq X \leq 1 \tag{16}
\]

where \( \Omega(\tau) \equiv \frac{\omega(\tau/\omega_0)}{\omega_0} \) and the dimensionless coefficients \( A(X) \), \( B(X,\tau) \), \( C(X) \) and \( F_0 \) are defined as

\[
A(X) \equiv \frac{EI(XL)}{\mu_0 \omega_0^2 L^2}, \quad B(X,\tau) \equiv \frac{T(XL,\tau/\omega_0)}{\mu_0 \omega_0^2 L^2}, \quad C(X) \equiv \frac{\mu(XL)}{\mu_0} \quad \text{and} \quad F_0 \equiv \frac{f_0}{\mu_0 \omega_0^2 D}. \tag{17}
\]

The conductor’s mass per unit length and the tensile load are henceforth assumed to be known and given as \( \mu(x) = \mu_0 \) and

\[
T(x,t) = T_0 \sqrt{1 + \left( \frac{\mu_0 g}{T_0} \right)^2 x^2 \cos \omega_0 t} \tag{18}
\]
such that \( C(X) = 1 \) and \( B(X, \tau) = B_0 \sqrt{1 + B_1 X^2} \cos \tau \). The coefficients \( B_0 \) and \( B_1 \) are defined as
\[
B_0 = \frac{T_0/\mu_0}{w_0^2 L^2} \quad \text{and} \quad B_1 = \left( \frac{\mu_0 g L}{T_0} \right)^2 .
\] (19)

On the other hand, the conductor’s bending stiffness \( EI(x) \) has been treated as an unknown function to be estimated from observable data and parameterized as
\[
EI(x) = EI_0 \left\{ \alpha + (1 - \alpha) \left[ 1 - e^{-\beta x (L-x)} \right] \right\}
\] (20)
such that \( A(X; \theta) = \theta_1 \left\{ \theta_2 + (1 - \theta_2) \left[ 1 - e^{-\theta_3 X (1-X)} \right] \right\} \) with a one-to-one correspondence among the \( \theta_i \)'s, \( i = 1, 2, 3 \), and the coefficients \( EI_0, \alpha \) and \( \beta^2 \). The modified notation \( A(X; \theta) \) will be used henceforth in order to emphasize the dependence of the coefficient \( A(X) \) upon the unknown parameter vector \( \theta \equiv [\theta_1 \ \theta_2 \ \theta_3]^T \). The reasoning behind the parameterization given by Eq. (20) relies on the fact that the conductor’s bending stiffness increases near the clamped ends as the relative movement among the wires becomes even more restricted.

The design variables are the dimensionless accelerometers’ positions \( d_j \in (0, 1), \ j = 1, 2, \ldots, n_d \). The observable data \( y \) are the conductor’s transverse displacement at the design positions and at prescribed time instants \( \tau_k, \ k = 1, 2, \ldots, N_\tau \). The components of the forward map \( G(\theta, d) \) (the predicted data) are arranged such that its \( l \)-th component \( G_l(\theta, d) = W(d_j, \tau_k; \theta) \), with \( l = j + (k - 1) n_d, \ l = 1, 2, \ldots, n_y \) and \( n_y = n_d \times N_\tau \). The modified notation \( W(d_j, \tau_k; \theta) \) again emphasizes the dependence of the predicted data upon the unknown parameter vector \( \theta \). The initial- and boundary-value problem given by Eqs. (13)-(16) has been solved by the Generalized Integral Transform Technique (GITT), a hybrid analytico-numerical technique based on eigenfunction expansions [11, 12]. The details are omitted here due to space limitation. Interested readers should consult Refs. [13, 14, 15, 16, 17] and references therein for mathematical and computational aspects about the GITT approach and its more recent developments to solve applied thermal and structural engineering problems governed by linear and/or nonlinear partial differential equations. The main idea is to eliminate all but one independent variable by choosing an auxiliary eigenvalue problem (which should carry as much information as possible about the original initial- and boundary-value problem to be solved) and also by performing an integral transformation. The solution of the eigenvalue problem comprises a discrete set of orthogonal eigenfunctions, which are the basis for the series expansion of the original potential. The GITT itself has also been employed to provide a hybrid analytico-numerical solution for the aforementioned auxiliary eigenvalue problem with arbitrarily general coefficients \( A(X; \theta) \). With the GITT approach, the auxiliary eigenvalue problem is reduced to a matrix eigenvalue problem [14, 16], which may be readily solved, for instance, with well-known Fortran routines for eigensystem analysis.

4. Numerical Results and Conclusions

In this section one reports the numerical results obtained for the optimum experiment design problem generally formulated in Section 2 using the forward model detailed in Section 3. Here, for simplicity, one assumes that there is only one accelerometer available to collect the conductor’s acceleration in time domain, i.e., \( n_y = 1 \). The optimum experiment design problem thus reduces to find the optimal location of that accelerometer \( d_1^* \), with \( d_1^* \in (0, 1) \). For typical

\[ \theta_1 \equiv \frac{EI_0}{\mu_0 w_0^2 L^4} \quad \theta_2 \equiv \alpha \quad \text{and} \quad \theta_3 \equiv \beta L^2 \]

The following relationships may be readily verified

\[ \theta_1 \equiv \frac{EI_0}{\mu_0 w_0^2 L^4} \quad \theta_2 \equiv \alpha \quad \text{and} \quad \theta_3 \equiv \beta L^2 \]
values of the known parameters $\mu_0$, $L$ and $T_0$, and by choosing $\omega_0 = 0.2 \pi \text{ rad/s}$, one assumes $B_0 = 15$ and $B_1 = 1.5 \times 10^{-3}$.

Prior information regarding the unknown parameters $\theta_1$, $\theta_2$ and $\theta_3$ suggests that they must be non-negative numbers and also that $\theta_2 > 1$, as $\theta_2 < 1$ implies that the conductor’s bending stiffness would increase as one moves away from the clamped ends, which has not been observed experimentally. One assumes that they are independent random variables with Gamma distribution, whose support comprises the set of non-negative real numbers $(0, \infty)$. The prior distribution $\pi_\Theta(\theta)$ thus simplifies to $\pi_\Theta(\theta) = \pi_{\Theta_1}(\theta_1) \pi_{\Theta_2}(\theta_2) \pi_{\Theta_3}(\theta_3)$ with each $\pi_{\Theta_i}(\theta_i)$ being a univariate Gamma distribution with shape and scale parameters $k_i$ and $s_i$, $i = 1, 2, 3$

$$\pi_{\Theta_i}(\theta_i) = \frac{1}{\Gamma(k_i) s_i^{k_i}} \theta_i^{k_i-1} \exp \left(-\frac{\theta_i}{s_i}\right)$$

(21)

The mean values $\bar{\theta}_i$ and variances $\sigma^2_{\theta_i}$ are, respectively, $\bar{\theta}_i = k_i s_i$ and $\sigma^2_{\theta_i} = k_i s_i^2$, $i = 1, 2, 3$. The shape and scale parameters are chosen such that the coefficient of variation for each parameter $\theta_i$, $\sigma_{\theta_i}/\bar{\theta}_i$, $i = 1, 2, 3$, is held fixed at 0.25. The following values are adopted: $k_1 = k_2 = k_3 = 16$; $s_1 = 0.1125$; $s_2 = 0.25$ and $s_3 = 0.25$; therefore, the prior mean, $\bar{\theta}$, and the prior covariance matrix, $\Sigma_\theta$, are given by $\bar{\theta} = [1.8 \quad 4 \quad 4]^T$ and $\Sigma_\theta = \text{diag}(0.2025, 1, 1)$ (a diagonal matrix with entries 0.2025, 1 and 1). With the chosen values for the prior distribution of the unknown parameter $\theta_2$, the probability of the event $0 < \theta_2 \leq 1$, $\Pr(0 < \Theta_2 \leq 1) = \int_0^{1} \pi_{\Theta_2}(\theta_2) d\theta_2$, is approximately $4.9 \times 10^{-6}$; in other words, it is a rare event. Regarding the external excitation, the following parameters are held fixed: $F_0 = 5 \times 10^3$; $X_s = 0.9$; $\omega_i = 2 \pi \text{ rad/s}$; $\omega_n = 10 \pi \text{ rad/s}$; and $t_f = 2 s$. One also assumes a sampling frequency $f_s = 400$ Hz (400 samples/s) for the measurement of the transverse displacements at every $\Delta t = 1/f_s = 0.0025$ s ($\Delta \tau = \omega_0 \Delta t \approx 0.0016$); hence, $N_T = 801$ (number of data points in time-domain).

**Convergence analysis.** In the first set of results, the convergence of the GITT solution for the transverse displacement has been checked by gradually increasing the truncation order $N_T$, with $N_T \in \{10, 15, 20, 25, 30\}$. For the convergence analysis, the unknown parameters $\theta_1$, $\theta_2$ and $\theta_3$ are held fixed at their prior means. Tables 1, 2 and 3 show the numerical results obtained for the transverse displacement $W(X, \tau)$ as a function of position $X$ and at three different dimensionless times $T_\tau$, $T_\tau \in \{T_\tau, 2T_\tau, 3T_\tau\}$, for each of the truncation orders $N_T$. From the analysis of Tabs. 1, 2 and 3 one may draw the following important remarks. Firstly, convergence with at least three significant digits may be achieved with $N_T = 25$. Secondly, faster convergence has been verified for the smaller times and also for those positions furthest from the end $X = 0$. Higher truncation orders are needed for higher times because for the chosen sine chirp excitation, higher excitation frequencies are excited as time elapses. Figure 2 graphically illustrates the convergence behavior of the transient displacement $W(X_a, \tau)$ at four different positions $X_a$, $X_a \in \{0.1, 0.3, 0.7, 0.9\}$, for the entire duration of the sine chirp excitation $\tau \in [0, \tau_f]$. Notice that the differences among the curves obtained with $N_T \geq 15$ are not visually distinguishable for the entire time interval of analysis. For times $\tau \lesssim 1$, quite accurate results were achieved with the lowest truncation order $N_T = 10$.

**Optimum experiment design.** In the second and final set of results, one computes estimates for the expected utility $U(d_1)$ with the aid of Eq. (10), for two different variances of the measurement error $\sigma^2$, $\sigma^2 \in \{0.01, 0.10\}$, and for increasing number of samples in the nested Monte-Carlo estimate ($N_{\text{out}} = N_{\text{in}}$), with $N_{\text{out}} \in \{10^3, 2 \times 10^3, 5 \times 10^3\}$. Based on the previous convergence analysis with the GITT solution of the associated forward model, all the computations of expected utility have been performed with a truncation order $N_T = 25$. Figure 3 shows the estimates obtained for $U(d_1)$ as a function of $d_1$, for $d_1 \in (0, 1)$. From
Figure 2. Plots of transient displacement $W(X_a, \tau)$ as a function of the dimensionless time $\tau$, $\tau \in [0, \tau_f]$, and at four different positions $X_a \in \{0.1, 0.3, 0.7, 0.9\}$, computed with the GITT approach with increasing truncation orders $N_T$, $N_T \in \{10, 15, 20, 25, 30\}$.

Figure 3. Expected utility $U(d_1)$ as a function of the accelerometer position $d_1$, for $d_1 \in (0, 1)$, for two different variances of the measurement error $\sigma^2$, $\sigma^2 \in \{0.01, 0.10\}$, and for three different number of samples used in its Monte-Carlo estimate, $N_{\text{out}} \in \{10^3, 2 \times 10^3, 5 \times 10^3\}$. 
Table 1. Convergence of the GITT solution for the transverse displacement at \( \tau = \frac{\tau_f}{4} \), for increasing truncation orders \( N_T \), \( N_T \in \{10, 15, 20, 25, 30\} \) and for \( \theta = [1.8 \ 4 \ 4]^T \).

<table>
<thead>
<tr>
<th>X</th>
<th>( N_T = 10 )</th>
<th>( N_T = 15 )</th>
<th>( N_T = 20 )</th>
<th>( N_T = 25 )</th>
<th>( N_T = 30 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.750</td>
<td>0.935</td>
<td>0.948</td>
<td>0.950</td>
<td>0.950</td>
</tr>
<tr>
<td>0.2</td>
<td>4.312</td>
<td>4.286</td>
<td>4.344</td>
<td>4.350</td>
<td>4.350</td>
</tr>
<tr>
<td>0.3</td>
<td>6.984</td>
<td>7.283</td>
<td>7.337</td>
<td>7.340</td>
<td>7.340</td>
</tr>
<tr>
<td>0.5</td>
<td>-1.403</td>
<td>-1.337</td>
<td>-1.346</td>
<td>-1.338</td>
<td>-1.335</td>
</tr>
<tr>
<td>0.7</td>
<td>-11.360</td>
<td>-11.552</td>
<td>-11.626</td>
<td>-11.619</td>
<td>-11.619</td>
</tr>
<tr>
<td>0.9</td>
<td>-3.165</td>
<td>-3.231</td>
<td>-3.283</td>
<td>-3.277</td>
<td>-3.274</td>
</tr>
</tbody>
</table>

Table 2. Convergence of the GITT solution for the transverse displacement at \( \tau = \frac{\tau_f}{2} \), for increasing truncation orders \( N_T \), \( N_T \in \{10, 15, 20, 25, 30\} \) and for \( \theta = [1.8 \ 4 \ 4]^T \).

<table>
<thead>
<tr>
<th>X</th>
<th>( N_T = 10 )</th>
<th>( N_T = 15 )</th>
<th>( N_T = 20 )</th>
<th>( N_T = 25 )</th>
<th>( N_T = 30 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>-3.480</td>
<td>-3.564</td>
<td>-3.514</td>
<td>-3.510</td>
<td>-3.506</td>
</tr>
<tr>
<td>0.2</td>
<td>-12.753</td>
<td>-12.852</td>
<td>-12.821</td>
<td>-12.812</td>
<td>-12.817</td>
</tr>
<tr>
<td>0.3</td>
<td>-16.687</td>
<td>-16.864</td>
<td>-16.966</td>
<td>-16.956</td>
<td>-16.951</td>
</tr>
<tr>
<td>0.5</td>
<td>19.404</td>
<td>19.479</td>
<td>19.513</td>
<td>19.527</td>
<td>19.525</td>
</tr>
<tr>
<td>0.9</td>
<td>-9.602</td>
<td>-10.469</td>
<td>-10.739</td>
<td>-10.730</td>
<td>-10.725</td>
</tr>
</tbody>
</table>

Table 3. Convergence of the GITT solution for the transverse displacement at \( \tau = \tau_f \), for increasing truncation orders \( N_T \), \( N_T \in \{10, 15, 20, 25, 30\} \) and for \( \theta = [1.8 \ 4 \ 4]^T \).

<table>
<thead>
<tr>
<th>X</th>
<th>( N_T = 10 )</th>
<th>( N_T = 15 )</th>
<th>( N_T = 20 )</th>
<th>( N_T = 25 )</th>
<th>( N_T = 30 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>-45.5</td>
<td>-62.7</td>
<td>-64.5</td>
<td>-64.4</td>
<td>-64.3</td>
</tr>
<tr>
<td>0.2</td>
<td>-20.4</td>
<td>-5.85</td>
<td>-5.21</td>
<td>-4.99</td>
<td>-5.02</td>
</tr>
<tr>
<td>0.3</td>
<td>8.91</td>
<td>0.311</td>
<td>-0.369</td>
<td>-0.129</td>
<td>-0.122</td>
</tr>
<tr>
<td>0.5</td>
<td>29.90</td>
<td>24.58</td>
<td>25.96</td>
<td>26.27</td>
<td>26.29</td>
</tr>
<tr>
<td>0.7</td>
<td>42.61</td>
<td>39.09</td>
<td>39.42</td>
<td>39.57</td>
<td>39.51</td>
</tr>
<tr>
<td>0.9</td>
<td>-11.63</td>
<td>-28.21</td>
<td>-29.06</td>
<td>-28.73</td>
<td>-28.78</td>
</tr>
</tbody>
</table>

the plots shown in Fig. 3 the following remarks should be highlighted. Firstly, the shape of the expected utility curve is quite symmetric about \( d_1 = 0.5 \). Notice that the expected utility falls down sharply as the accelerometer is placed next to both conductor’s clamped ends. Such behavior may be explained by the fact that time-domain acceleration becomes much more affected by the zero-displacement boundary conditions at the clamped ends, regardless of the conductor’s bending stiffness spatial variation (defined by the unknown parameter vector \( \theta \)). In other words, a time-domain acceleration collected at a position near to either conductor’s end would probably lead to inferring values of \( \theta \) with larger uncertainties. Secondly, for the larger variance \( \sigma^2 \), smaller differences have been verified among the expected utility curves computed with increasing number of samples. As the variance \( \sigma^2 \) increases, the contribution of the prior
distribution, \(\pi_\theta(\theta)\), to the expected utility becomes more relevant with regard to the likelihood distribution. The lowest number of samples probably captures quite well the high-probability region of the prior support, which has far fewer dimensions than the likelihood support. Finally, the optimum experiment design for the problem under investigation is thus \(d^*_1 = 0.5\), regardless of the magnitude of the variance of the measurement errors.

5. Final Remarks
In the current work, a Bayesian statistical framework developed by Huan and Marzouk [5] has been applied to solve an experiment design problem with high-voltage conductors in bending vibrations. The associated forward model has been mathematically formulated in dimensionless form and subsequently solved by the GITT approach. A convergence analysis for the transverse displacement has been reported in tabular and graphic formats by gradually increasing the truncation order of the series expansion. The main result reported for the optimum experiment design problem may be summarized as follows. Given only one accelerometer to measure the conductor’s time-domain acceleration and for a sine chirp excitation force with lower and upper frequencies \(\omega_l = 2\pi \text{ rad/s}\) and \(\omega_u = 10\pi \text{ rad/s}\), the optimum design, \(d^*_1\), is \(d^*_1 = 0.5\), i.e., the measurement accelerometer should be placed at the midpoint of the span length. Future extensions to be considered are (i) \(n_d > 1\) accelerometers for the measurement of the conductor’s time-domain acceleration; (ii) the use of design variables other than the position of the measurement accelerometers (as, for instance, the final time \(\tau_f\), the lower and upper excitation frequencies \(\omega_l\) and \(\omega_u\)); and (iii) the modification of the forward model to account for the interaction between the conductor and the electrodynamic shaker at the excitation point.

References