Computer Science Technical Report
TR-07-38
November 21, 2007

Emmanuel D. Blanchard, Adrian Sandu, and Corina Sandu

"A Polynomial Chaos Based Bayesian Approach for Estimating Uncertain Parameters of Mechanical Systems – Part I: Theoretical Approach"

Center for Vehicle Systems and Safety
Computer Science Department & Department of Mechanical Engineering
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061
Phone: (540)-231-2193
Fax: (540)-231-9218
Email: sandu@cs.vt.edu
Web: http://www.eprints.cs.vt.edu

A Polynomial Chaos Based Bayesian Approach for Estimating Uncertain Parameters of Mechanical Systems – Part I: Theoretical Approach

Emmanuel D. Blanchard (eblancha@vt.edu)

Advanced Vehicle Dynamics Lab

Center for Vehicle Systems and Safety, Virginia Tech, Blacksburg, VA 24061– 0238

Adrian Sandu (sandu@cs.vt.edu)

Computer Science Department and the Center for Vehicle Systems and Safety

Virginia Tech, Blacksburg, VA 24061

Corina Sandu (csandu@vt.edu)

Advanced Vehicle Dynamics Lab

Center for Vehicle Systems and Safety, Virginia Tech, Blacksburg, VA 24061– 0238

ABSTRACT

This is the first part of a two-part article. A new computational approach for parameter estimation is proposed based on the application of the polynomial chaos theory. The polynomial chaos method has been shown to be considerably more efficient than Monte Carlo in the simulation of systems with a small number of uncertain parameters. In the new approach presented in this paper, the maximum likelihood estimates are obtained by minimizing a cost function derived from the Bayesian theorem. Direct stochastic collocation is used as a less computationally expensive alternative to the traditional Galerkin approach to propagate the uncertainties through the system in the polynomial chaos framework. This approach is applied to very simple mechanical systems in order to illustrate how the cost function can be affected by undersampling, non-identifiablily of the system, non-observability, and by excitation signals that are not rich enough.. When the system is non-identifiable, regularization techniques can still yield most likely values among the possible combinations of uncertain parameters resulting in the same time responses than the ones observed. This is illustrated using a simple spring-mass system. Possible applications of this theory to the field of vehicle dynamics simulations include the estimation of mass, inertia properties, as well as other parameters of interest. In the second part of this article, this new parameter estimation method is illustrated on a nonlinear four-degree-of-freedom roll plane model of a vehicle in which an uncertain mass with an uncertain position is added on the roll bar.

Keywords: Parameter Estimation, Polynomial Chaos, Collocation, Bayesian Estimation, Hammersley Algorithm, Halton Algorithm, Vehicle Dynamics

1. INTRODUCTION AND BACKGROUND

The polynomial chaos theory has been shown to be consistently more efficient than Monte Carlo simulations in order to assess uncertainties in mechanical systems [13, 14]. This paper extends the polynomial chaos theory to the problem of parameter estimation, and applies it to a four degree of freedom roll plane model of a vehicle with a mass added on the roll bar. Parameter estimation is an important problem, because many parameters simply cannot be measured

physically with good accuracy, especially in real time applications. The method presented in this paper has the advantage of being able to deal with non-Gaussian parametric uncertainties.

Parameter estimation is a very difficult problem, especially for large systems, and a lot of effort devoted to it would be needed. Estimating a large number of parameters often proved to be computationally too expensive. This has led to the development of techniques determining which parameters affect the system's dynamics the most, in order to choose the parameters that are important to estimate [18]. Sohns, et al. [18] proposed the use of activity analysis as an alternative to sensitivity-based and principal component-based techniques. Their approach combines the advantages of the sensitivity-based techniques (i.e., being efficient for large models) and the component-based techniques (i.e., keeping parameters that can be physically interpreted). Zhang and Lu [23] combined the Karhunen–Loeve decomposition and perturbation methods with polynomial expansions in order to evaluate higher-order moments for saturated flow in randomly heterogeneous porous media.

The polynomial chaos method started to gain attraction after Ghanem and Spanos applied it successfully to the study of uncertainties in structural mechanics and vibration [5-8] using Wiener-Hermite polynomials. Xiu extended the approach to general formulations based on Wiener-Askey polynomials family [20], and applied it to fluid mechanics [19, 21, 22]. Authors applied for the first time the polynomial chaos method to multibody dynamic systems [13-16], terramechanics [12, 17], and parameter estimation [2, 3].

The fundamental idea of polynomial chaos approach is that random processes of interest can be approximated by sums of orthogonal polynomial chaoses of random independent variables. In this context, any uncertain parameter can be viewed as a second order random process (processes with finite variance; from a physical point of view they have finite energy). Thus, a second order random process $X(\theta)$, viewed as a function of the random event θ, can be expanded in terms of orthogonal polynomial chaos as [5]:

$$X(\theta) = \sum_{j=1}^{\infty} c^j \psi^j(\xi(\theta)) \tag{1}$$

Here $\psi^i(\xi_{i_1} \cdots \xi_{i_n})$ are generalized Askey-Wiener polynomial chaoses, in terms of the multi-dimensional random variable $\xi = (\xi_{i_1} \cdots \xi_{i_n})$.

The multi-dimensional basis functions are tensor products of 1-dimensional polynomial basis:

$$\psi^l(\xi_1 \cdots \xi_n) = \prod_{k=1}^{n} P_k^{j_k}(\xi_k),\ l = 1,2,\cdots,S;\ j_k =,2,\cdots,p \tag{2}$$

For Gaussian random variables the basis are Hermite polynomials, for uniformly distributed random variables the basis are Legendre polynomials, for beta distributed random variables the basis are Jacobi polynomials, and for gamma distributed random variables the basis are Laguerre polynomials [20, 21]. In practice, a truncated expansion is used,

$$X = \sum_{j=1}^{S} c^j \psi^j(\xi) \tag{3}$$

where $S = \dfrac{(n_u + p)!}{n_u!\, p!}$, n_u is the number of random variables ξ, and p is the maximum order of polynomial basis. The total number of terms increases rapidly with n_u and p.

In the deterministic case, a second order unconstraint multibody system can be described by the following Ordinary Differential Equation:

$$\begin{cases} \dot{x} = v \\ \dot{v} = f(x, v) \\ x(t_0) = x_0 \\ v(t_0) = v_0 \end{cases} \tag{4}$$

In this stochastic framework, for a second order unconstraint multibody system, the displacement x and the velocity v can be expanded as:

$$x_m(\xi) = \sum_{i=1}^{S} x_m{}^i \psi^i(\xi), \quad v_m(\xi) = \sum_{i=1}^{S} v_m{}^i \psi^i(\xi) \tag{5}$$

Propagating Eq. (5) through the deterministic system of equations of the multibody system, one obtains:

$$\begin{cases} \dot{x}_{m,k}^i = v_{m,k}^i \\ \sum_{i=1}^{S} \dot{v}_{m,k}^i \ \psi^i(\xi) = F_k(t, \sum_{j=1}^{S} x_m^j \ \psi^j(\xi), \sum_{j=1}^{S} v_m^j \ \psi^j(\xi); \sum_{j=1}^{S} \theta_m^j \ \psi^j(\xi)) \\ x_m(t_0) = x_{m,0}, \ t_0 \le t \le t_F \end{cases} \tag{6}$$

To derive evolution equations for the stochastic coefficients $x_m{}^i(t)$ we impose that Eq. (6) holds at a given set of collocation vectors $\mu^i = \left[\mu_1^i \ \cdots \ \mu_d^i\right]^T$ for all $1 \le i \le S$. This leads to:

$$\dot{x}_m{}^i = v_m{}^i, \ \sum_{i,j=1}^{S} A_{i,j} \ \dot{v}_m{}^j = F\left(t, \sum_{m=1}^{S} A_{i,m} \ x_m{}^m, \sum_{m=1}^{S} A_{i,m} \ v_m{}^m; \sum_{m=1}^{S} A_{i,m} \ \theta_m{}^m\right) \tag{7}$$

where $\mathbf{A}$ represents the matrix of basis function values at the collocation points:

$$A = \left(A_{i,j}\right), \ A_{i,j} = \psi^j(\mu^i), \ 1 \le i \le S, \ 1 \le j \le S \tag{8}$$

The collocation points have to be chosen such that $\mathbf{A}$ is nonsingular. The collocation system can be written as:

$$\dot{X}^i = V^i, \ \dot{V}^i = F\left(t, X^i, V^i, \Theta^i\right), \ 1 \le i \le S \tag{9}$$

After integration, the stochastic solution coefficients are recovered using:

$$x^i(t) = \sum_{j=1}^{S} \left(A^{-1}\right)_{i,j} \ X^j(t), \ v^i(t) = \sum_{j=1}^{S} \left(A^{-1}\right)_{i,j} \ V^j(t) \tag{10}$$

The mean values of $x(t)$ and $v(t)$ are $x^1(t)\psi^1(\xi)$ and $v^1(t)\psi^1(\xi)$, respectively.

The standard deviations of $x(t)$ and $v(t)$ are given by:

$$\sqrt{\int_\Omega \left(\sum_{i=2}^{S} x^i(t)\psi^i(\xi)\right)^2 d\xi}, \ \sqrt{\int_\Omega \left(\sum_{i=2}^{S} v^i(t)\psi^i(\xi)\right)^2 d\xi} \tag{11}$$

When the basis are orthogonal polynomials, the standard deviations of $x(t)$ and $v(t)$ are given by:

$$\sum_{i=2}^{S} \left(x^i(t)\right)^2 \sqrt{\int_\Omega <\psi^i(\xi), \psi^i(\xi)> d\xi}, \ \sum_{i=2}^{S} \left(v^i(t)\right)^2 \sqrt{\int_\Omega <\psi^i(\xi), \psi^i(\xi)> d\xi} \tag{12}$$

When the basis are orthonormal, the standard deviations of $x(t)$ and $v(t)$ are given by:

$$\sum_{i=2}^{S} \left(x^i(t)\right)^2, \ \sum_{i=2}^{S} \left(v^i(t)\right)^2 \tag{13}$$

The Probability Density Function (PDF) of $x(t)$ and $v(t)$ are obtained by drawing histograms of their values using a Monte Carlo simulation and normalizing the area under the curves that are obtained. It is not computationally expensive since the Monte Carlo simulation is run on the final result, and not for the while process. For instance, the ODE is run the same number of times than the number of collocation points, which is typically much lower than the number of runs used for the Monte Carlo simulation.

2. BAYESIAN APPROACH FOR PARAMETER ESTIMATION

Optimal parameter estimation combines information from three different sources: the physical laws of evolution (encapsulated in the model), the reality (as captured by the observations), and the current best estimate of the parameters. The information from each source is imperfect has associated errors. Consider the mechanical system model (6) which advances the state in time in a simpler notation:

$$y_k = \begin{bmatrix} x_k \\ v_k \\ \theta_k \end{bmatrix}, \quad y_k = M\left(t_{k-1}, y_{k-1}\right), \quad y_0 = y(t_0), \quad k = 1,2,\ldots,N \tag{14}$$

The state of the model $y_k \in \Re^n$ at time moment t_k depends implicitly on the set of parameters $\theta \in \Re^p$, possibly uncertain (the model has n states and p parameters). M is the model solution operator which integrates the model equations forward in time (starting from state y_{k-1} at time t_{k-1} to state y_k at time t_k).

For parameter estimation it is convenient to formally extend the model state to include the model parameters and extend the model with trivial equations for parameters (such that parameters do not change during the model evolution)

$$\theta_k = \theta_{k-1} \tag{15}$$

The optimal estimation of the uncertain parameters is thus reduced to the problem of optimal state estimation. Observations of quantities that depend on system state are available at discrete times t^k

$$z_k = h\left(y_k\right) + \varepsilon_{obs}^k \approx H_k y_k + \varepsilon_{obs}^k, \quad \left\langle \varepsilon_{obs}^k \right\rangle = 0, \quad \left\langle \left(\varepsilon_{obs}^k\right)\left(\varepsilon_{obs}^k\right)^T \right\rangle = R_k \tag{16}$$

where $z_k \in \Re^m$ is the observation vector at t_k, h is the (model equivalent) observation operator and H_k is the linearization of h about the solution y_k. Note that there are m observations for the n-dimensional state vector, and that typically m < n. Each observation is corrupted by observational (measurement and representativeness) errors [4]. We denote by $\langle \cdot \rangle$ the ensemble average over the uncertainty space. The observational error is the experimental uncertainty associated with the measurements and is usually considered to have a Gaussian distribution with zero mean and a known covariance matrix R_k.

Using polynomial chaoses the uncertain parameters are modeled explicitly as functions of a set of random variables $\xi \in \Omega \subset \Re^p$ with a joint probability density function $\rho(\xi)$. The explicit dependency of the system state on the random variables is obtained via a collocation approach:

$$\theta(\xi) = \sum_{i=1}^{S} \theta^i \phi^i(\xi), \quad y_k(\xi, t) = \sum_{i=1}^{S} \left(y_k(t)\right)^i \phi^i(\xi) \tag{17}$$

We adopt the point of view that the "state of knowledge" about the uncertain parameters can be described by probability densities. From Bayes' rule the probability density of the parameter distribution conditioned by all observations is

$$P\left[y_N \,\middle|\, z_N \ldots z_0\right] = \frac{P[z_N \,|\, y_N] \cdot P[y_N \,|\, z_{N-1} \ldots z_0]}{\int P[z_N \,|\, y] \cdot P[y \,|\, z_{N-1} \ldots z_0]\, dy} \tag{18}$$

where $P[z_N \,|\, y_N]$ is the Probability Density Function (PDF) of the latest observational error (taken at time t_N), $P[y_N \,|\, z_{N-1} \ldots z_0]$ is the "model forecast PDF" conditioned by all previous observations (taken at times t_0 to t_{N-1}), and $P[y_N \,|\, z_N \ldots z_0]$ is the "assimilated PDF". The assimilated PDF represents the aposteriori probability of the parameters after all the observations have been taken into account.

For simplicity denote by y the current state of the system (the best estimation obtained using all previous observations $z_{N-1} \ldots z_0$) and by $z = z_N$ the latest, yet-to-be-used set of observations. Moreover, consider that the

observational error has a Gaussian distribution with covariance R_k and that the observations at different times are independent. Then Bayes' formula becomes

$$P\left[y \middle| z\right] = \frac{P[z \middle| y] \cdot P[y]}{\int_{\Re^n} P[z \middle| y] \cdot P[y]\, dy} = \frac{e^{-\frac{1}{2}\sum_{k=0}^{N}(z_k - H y_k)^T R_k^{-1}(z_k - H y_k)} \cdot P[y]}{\int_{\Re^n} e^{-\frac{1}{2}\sum_{k=0}^{N}(z_k - H y_k)^T R_k^{-1}(z_k - H y_k)} \cdot P[y]\, dy} \tag{19}$$

The unconditional probability density $P[y]$ is the PDF of the current system state, and is implicitly represented by the polynomial chaos expansion of the state $y = y(\xi)$. Moreover, integration against this probability density can be evaluated by integration in the independent random variables

$$\int_{\Re^n} f(y) \cdot P[y] \cdot dy = \int_{\Omega} f\left(y(\xi)\right) \cdot \rho(\xi) \cdot d\xi \quad \text{for any } f(\cdot) \tag{20}$$

The denominator can be evaluated by a multidimensional integration. However, in our approach, there is no need to evaluate this scaling factor, since its omission does not change the estimation procedure. (The omission of this scaling factor is equivalent to adding a constant to the function we minimize, and this does not change the result of the minimization procedure).

The mean of the best state estimate that uses the new observations z is obtained from Bayes formula as

$$\langle y \rangle = \int_{\Re^n} y \cdot P[y \mid z] \cdot dy = \frac{1}{\text{den}} \int_{\Re^n} y \cdot e^{-\frac{1}{2}\sum_{k=0}^{N}(z_k - H y_k)^T R_k^{-1}(z_k - H y_k)} \cdot P[y] \cdot dy = \frac{1}{\text{den}} \int_{\Omega} y(\xi) \cdot e^{-\frac{1}{2}\sum_{k=0}^{N}(z_k - H y_k)^T R_k^{-1}(z_k - H y_k)} \cdot \rho(\xi) \cdot d\xi \tag{21}$$

For the parameter estimation the Bayes' formula specializes to:

$$P\left[\theta \middle| z\right] = \frac{P[z|\theta] \cdot P[\theta]}{\text{den}} = \frac{e^{-\frac{1}{2}\sum_{k=0}^{N}(z_k - H y_k(\theta))^T R_k^{-1}(z_k - H y_k(\theta))} \cdot P[\theta]}{\int_{\Re^m} e^{-\frac{1}{2}\sum_{k=0}^{N}(z_k - H y_k(\theta))^T R_k^{-1}(z_k - H y_k(\theta))} \cdot P[\theta]\, d\theta} \tag{22}$$

Note that the aposteriori probability defined by Bayes formula can be written (in principle) as a function of the independent random variables ξ

$$\hat{\rho}(\xi) = P\left[\xi \middle| z\right] = \frac{e^{-\frac{1}{2}\sum_{k=0}^{N}(z_k - H y_k(\xi))^T R_k^{-1}(z_k - H y_k(\xi))} \cdot \rho(\xi)}{\int_{\Omega} e^{-\frac{1}{2}\sum_{k=0}^{N}(z_k - H y_k(\xi))^T R_k^{-1}(z_k - H y_k(\xi))} \cdot \rho(\xi)\, d\xi} \tag{23}$$

In this setting polynomial chaos is used to model the *a priori* pdf of the parameters; the Bayes formula is employed to obtain the *a posteriori* pdf (i.e., the pdf conditioned by the observations).

The maximum likelihood estimate is given by that realization of the parameters (that value of ξ) which maximizes the *a posteriori* probability $P[\theta \mid z]$, or, equivalently, minimizes $-\log\left(P[\theta \mid z]\right)$:

$$\min_{\xi \in \Omega} J = \frac{1}{2}\sum_{k=0}^{N}(z_k - H y_k(\xi))^T R_k^{-1}(z_k - H y_k(\xi)) - \log\left(\rho(\xi)\right) \tag{24}$$

Note that for $\xi \notin \Omega$ we have $\rho(\xi) = 0$ and cost the function J becomes infinite. This cost function is composed of two parts:

$$J_{\text{total}}(\xi) = \underbrace{\frac{1}{2}\sum_{k=0}^{N}(z_k - H y_k(\xi))^T R_k^{-1}(z_k - H y_k(\xi))}_{J_{\text{mismatch}}} + \underbrace{\left(-\log\left(\rho(\xi)\right)\right)}_{J_{\text{apriori}}} \tag{25}$$

where J_{mismatch} comes only from the differences between the available measurements and the model response, while J_{apriori} encapsulates the apriori knowledge of the parameter uncertainty. The value $\hat{\xi} = \arg\min J$ minimizing the cost function (25) gives the most likely values of our uncertain parameters as $\hat{\theta} = \theta(\hat{\xi})$.

3. INSIGHT INTO THE BAYESIAN APPROACH USING SIMPLE MECHANICAL SYSTEMS

We now illustrate the proposed Bayesian approach for the estimation of parameters of several simple mechanical systems. We discuss how the cost function and the estimate can be affected by low sampling rates (i.e., below the Nyquist frequency), by measurement noise, and by non-identifiability issues.

The state of the model $y^k \in \Re^n$ at time moment t^k depends implicitly on the set of uncertain parameters $\theta \in \Re^p$, and therefore on the set of independent random variables ξ. This dependency is explicitly represented in the polynomial chaos framework, specifically, at each time moment t^k the state is given as a polynomial of the random variables $y^k = y^k(\xi)$. The probability density of the state can also be obtained from this relation. H is simply a matrix converting the states of the model $y(\xi)$ to the observable parameters of the system (i.e., the quantities which can be measured), which are contained in z. R_k is the covariance matrix of the uncertainty associated with the measurements, i.e. of the measurement noise.

$J_{apriori} = -\log(\rho(\xi))$ comes from the apriori knowledge of our uncertain parameters. Using polynomial chaoses the uncertain parameters can be modeled explicitly as functions of a set of random variables $\xi \in \Omega \subset \Re^p$ with a joint probability density function $\rho(\xi)$.

$J_{mismatch}$ is usually the most important component of the cost function, but $J_{apriori}$ is useful when $J_{mismatch}$ does not contain enough information in order to find a clear minimum value for our cost function. This is illustrated in the next section of this article.

3.1. Mass-Spring System with Uncertain Initial Velocity

This section applies the Bayesian approach to the simple Mass-Spring system shown in Figure 1.

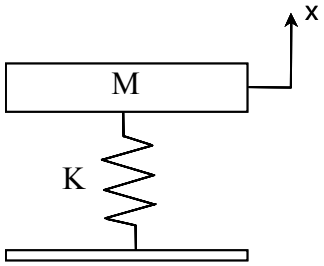


Figure 1. Mass –Spring System

The parameters K (the stiffness of the spring) and M (the mass of the body) are known. The system has a zero initial displacement $x_0 = 0$ but a nonzero initial velocity v_0 (e.g., created by hitting the mass from below with a hammer at $t = 0$, which will produce $v_0 > 0$). We want to estimate the uncertain initial condition v_0 based on measurements of the displacement $x(t)$ at later times.

The equation of motion of the system is

$$\mathrm{M}\ \ddot{x}(t) + K\ x(t) = 0 \tag{26}$$

and admits the general solution [11]:

$$x(t) = \sqrt{\frac{{v_0}^2 + (x_0\ \omega_n)^2}{{\omega_n}^2}}\ \sin\left(\omega_n\ t + \tan^{-1}\left(\frac{x_0\ \omega_n}{v_0}\right)\right)\ ,\ \omega_n = \sqrt{\frac{K}{M}} \tag{27}$$

The values chosen for the numerical experiments are $\mathrm{K} = 0.1\times(2\pi)^2 \approx 3.9478$ N/m, $\mathrm{M} = 0.1\,\mathrm{kg}$, and therefore $\omega_n = 2\pi$ rad/s $= 1$ Hz. For these values, and with $x_0 = 0$, the solution (27) becomes:

$$x(t) = \frac{v_0}{2\pi}\sin(2\pi t), \quad v(t) = v_0\cos(2\pi t). \tag{28}$$

The amplitude of $x(t)$ and the amplitude of $v(t)$ are both proportional to the uncertain parameter v_0, as illustrated in Figure 2. A *single* measurement of the displacement at a time $t_1 \neq m/2$ (with m integer) allows to estimate the initial velocity as $v_0 = 2\pi x(t_1)/\sin(2\pi t_1)$. Note that $\sin(2\pi t_1) \neq 0$. A single measurement of both the velocity and the displacement at any time t_2 is sufficient to retrieve the initial velocity, since for any t_2 at least one of the variables is nonzero, $\sin(2\pi t_2) \neq 0$ or $\cos(2\pi t_2) \neq 0$.

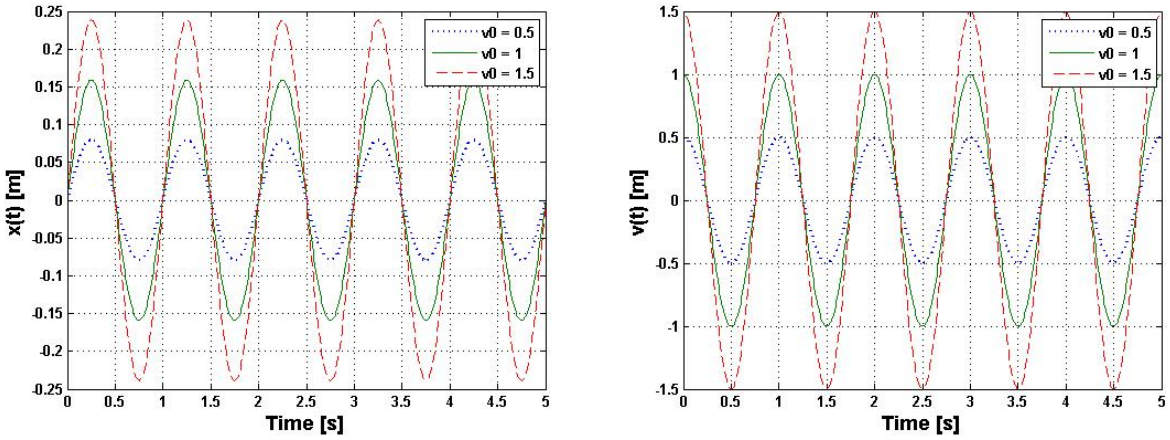


Figure 2. Displacements and Velocities of the Mass –Spring System.

We now consider the case where measurements of the displacement only are taken at multiple time moments $t_1, t_2, \cdots, t_N$ This will give insight on how the two parts of the Bayesian cost function can be affected by low sampling rates (i.e., below the Nyquist frequency) and by measurement noise.

We assume some prior knowledge of the initial velocity, which represents how hard different people can hit the mass with the hammer. The range of possible initial velocities is between 0.5 m/s and 1.5 m/s, with a most likely value of $v_0 = 1$ m/s. We model this prior knowledge as shown in Figure 3. Let ξ be a random variable with a Beta(1,1) probability distribution $\rho(\xi)$ in the range $\xi \in [-1,1]$. The random initial velocity is then

$$v(\xi,0) = v_{0\,nom}\left(1 + 0.5\,\xi\right)\ \ [m/s]. \tag{29}$$

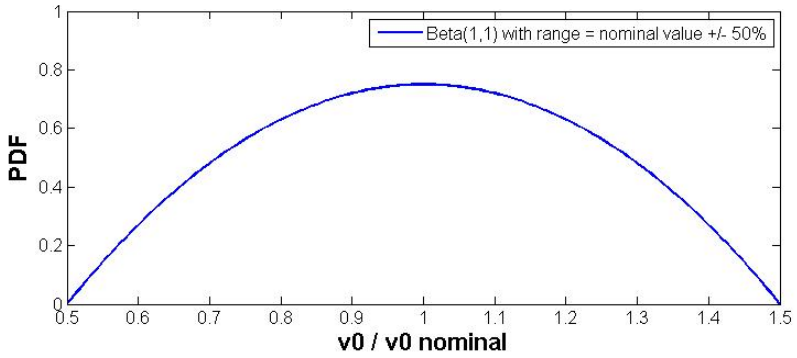


Figure 3. Beta (1, 1) Distribution for v0

The state of the system at future times depends on the random initial velocity and can be represented by $y(\xi,t)=\begin{bmatrix} x(\xi,t) \\ v(\xi,t) \\ \theta(\xi,t) \end{bmatrix}$. Synthetic measurements are obtained from a reference simulation with the reference value of the uncertain parameter $\xi^{ref}=0.23$. If we assume that only the displacement can be measured we have that $H=\begin{bmatrix}1 & 0 & 0\end{bmatrix}$ and the measurements yield

$$z_k = H\cdot y^{\text{ref}}(t_k)+\varepsilon_k = x^{\text{ref}}(t_k)+\varepsilon_k, \quad \varepsilon_k \in \mathrm{N}(0,R_k). \tag{30}$$

The measurement noise ε_k is assumed to be Gaussian with a zero mean and a variance 1% (or 0.01% or 10% when indicated) of the value of $x(t)$ plus the maximum measured value of $x(t)$ divided by 1000. Therefore, the covariance of the uncertainty associated with the measurements is $R_k=\max\left\{10^{-12},\left(0.01\ z_k+0.001\ \max_t(z_k)\right)^2\right\}[m^2]$. This value is always greater than zero and R_k^{-1} can always be computed. Measurement errors at different times are independent random variables.

The maximum likelihood estimate is obtained by minimizing the Bayesian cost function

$$J_{\text{total}}(\xi)=\underbrace{\tfrac{1}{2}\sum_{k=1}^{N}\left(z_k-Hy(\xi,t_k)\right)^T R_k^{-1}\left(z_k-Hy(\xi,t_k)\right)}_{J_{\text{mismatch}}}+\underbrace{\left(-\log\left(\rho(\xi)\right)\right)}_{J_{\text{apriori}}} \tag{31}$$

The random system output $y(\xi,t)$ is discretized using 6 terms in the polynomial chaos expansions, and 12 collocation points will be used to derive the polynomial chaos coefficients. The collocation points used in this study are obtained using an algorithm based on the Halton algorithm [9], which is similar to the Hammersley algorithm [10]. More details are provided about the Halton points and the concept of collocation in general in the second part of this article.

The frequency of the output signal $x(t,\xi)$ is 1 Hz for any value of ξ. If $x(t)$ is measured every 0.5 s from t = 0.5 to t = 5, then $x(t_k,\xi)=0$ for any value of ξ and $z_k=\varepsilon_k$ mismatch part gives no extra information, as shown in Figure 4, in which the plot for $x(t)$ was obtained with $\xi=0.23$.

$$J_{\text{mismatch}}(\xi)=\tfrac{1}{2}\sum_{k=1}^{N}\left(\varepsilon_k\right)^T R_k^{-1}\left(\varepsilon_k\right) \tag{32}$$

Let's illustrate this with detailing the step by step procedure using analytical formulas for this particular example.

The explicit dependency of $x(t,\xi)$ is obtained via a collocation approach. It can be represented as

$$x(t,\xi) = \sum_{i=1}^{S=6} x^i(t)\,\phi^i(\xi) \tag{33}$$

The equation of motion of the system is

$$M\,\ddot{x}(t) + K\,x(t) = 0 \tag{34}$$

As shown earlier, the solution for this equation for $\omega_n = \sqrt{\frac{K}{M}} = 2\pi$ rad/s $= 1$ Hz and for a zero initial displacement is

$$x(t) = \frac{v_0}{2\pi}\sin(2\pi t), \quad v_0 = \dot{x}(t) \;\; at \;\; t = 0. \tag{35}$$

In a polynomial chaos framework, the equation of motion of the system yields 6 equations for $S = 6$:

$$M\,\ddot{x}^i(t) + K\,x^i(t) = 0, \quad i = 1,\ 2,\ \ldots,\ S \tag{36}$$

Solving the 6 different equations of motions separately yields

$$x^i(t) = \frac{v_0^i}{2\pi}\sin(2\pi\, t_k), \quad i = 1,\ 2,\ \ldots,\ S \tag{37}$$

and the polynomial expression of the displacement is

$$x(t,\xi) = \sum_{i=1}^{S=6} \frac{v_0^i}{2\pi}\sin(2\pi\, t_k)\phi^i(\xi) \tag{38}$$

Let's note that in the general case, there is no need to know the closed form solution of the equations of motion. The approach presented in this paper still works when using the states variables obtained with numerical techniques to solve ODE's at the chosen collocation points.

The coefficients v_0^i are obtained using

$$v(\xi,0) = v_{0\,nom} + 0.5\,v_{0\,nom}\,\xi = v_0^1\,\phi^1(\xi) + v_0^2\,\phi^2(\xi) + v_0^3\,\phi^3(\xi) + v_0^4\,\phi^4(\xi) + .v_0^5\,\phi^5(\xi) + v_0^6\,\phi^6(\xi) \tag{39}$$

For beta(1,1) distributed random variables the basis are Jacobi (1,1) polynomials. With one random variable and for the range $\xi \in [-1,1]$, the normalized Jacobi (1,1) polynomials are:

$$\begin{cases} \phi^1(\xi) = 1/2 \\ \phi^2(\xi) = (3/28)\,(-1 + 2\,\xi) \\ \phi^3(\xi) = (1/106)\,(1 - 5\,\xi + 5\,\xi^2) \\ \phi^4(\xi) = (5/8344)\,(-1 + 9\,\xi - 21\,\xi^2 + 14\,\xi^3) \\ \phi^5(\xi) = (3/92822)\,(1 - 14\,\xi + 56\,\xi^2 - 84\,\xi^3 + 42\,\xi^4) \\ \phi^6(\xi) = (7/4455028)\,(-1 + 20\,\xi - 120\,\xi^2 + 300\,\xi^3 - 330\,\xi^4 + 132\,\xi^5) \end{cases} \tag{40}$$

Therefore, the coefficients v_0^i are

$$v_0^1 = \frac{5\,v_{0\,nom}}{2},\ v_0^2 = \frac{7\,v_{0\,nom}}{3},\ v_0^3 = 0, v_0^4 = 0, v_0^5 = 0,\ v_0^6 = 0 \tag{41}$$

The cost function can be written as

$$J(\xi) = \frac{1}{2}\sum_{k=1}^{N}\frac{1}{R_k}\left[z^k - \left(\sum_{i=1}^{S=6}\frac{v_0^i}{2\pi}\sin(2\pi\, t_k)\;\phi^i(\xi)\right)\right]^2 - \log(\rho(\xi)), \quad \rho(\xi) = \frac{3}{4}\left(1-\xi^2\right) \tag{42}$$

Using the fact that only v_0^1 and v_0^2 are nonzero and replacing v_0^1 by $(5/2)\,v_{0\,nom}$ and v_0^2 by $(7/3)\,v_{0\,nom}$, it yields

$$J(\xi)=\frac{1}{2}\sum_{k=1}^{N}\frac{1}{R_k}\left(z^k-\frac{(5/2)\,v_{0\,nom}}{2\pi}\sin(2\pi\,t_k)\left(\frac{1}{2}\right)-\frac{(7/3)\,v_{0\,nom}}{2\pi}\sin(2\pi\,t_k)\;(3/28)(-1+2\,\xi)\right)^2-\log\left(\frac{3}{4}\left(1-\xi^2\right)\right) \quad (43)$$

which can be simplified as

$$J(\xi)=\frac{1}{2}\sum_{k=1}^{N}\frac{1}{R_k}\left(z^k-\frac{v_{0\,nom}}{2\pi}\sin(2\pi\,t_k)\;-\frac{(v_{0\,nom}/4)}{2\pi}\sin(2\pi\,t_k)\;(2\,\xi)\right)^2-\log\left(\frac{3}{4}\left(1-\xi^2\right)\right) \quad (44)$$

which is also equal to

$$\begin{aligned}J(\xi)=&\frac{1}{2}\sum_{k=1}^{N}\frac{1}{R_k}\left(z^k-\frac{v_{0\,nom}}{2\pi}\sin(2\pi\,t_k)\right)^2-\sum_{k=1}^{N}\frac{1}{R_k}\left(z^k-\frac{v_{0\,nom}}{2\pi}\sin(2\pi\,t_k)\right)\frac{(v_{0\,nom}/4)}{2\pi}\sin(2\pi\,t_k)\;2\xi\\&+\frac{1}{2}\sum_{k=1}^{N}\frac{1}{R_k}\left(\frac{(v_{0\,nom}/4)}{2\pi}\sin(2\pi\,t_k)\right)^2(2\xi)^2-\log\left(\frac{3}{4}\left(1-\xi^2\right)\right)\end{aligned} \quad (45)$$

The closed form solution of the value minimizing the total cost function is a long expression that is not written here. The mismatch part of the cost function is

$$\begin{aligned}J_{mismatch}(\xi)=&\frac{1}{2}\sum_{k=1}^{N}\frac{1}{R_k}\left(z^k-\frac{v_0}{2\pi}\sin(2\pi\,t_k)\right)^2\\&-\sum_{k=1}^{N}\frac{1}{R_k}\left(z^k-\frac{v_0}{2\pi}\sin(2\pi\,t_k)\right)^2\frac{v_0^2}{2\pi}\sin(2\pi\,t_k)\;2\xi\\&+\frac{1}{2}\sum_{k=1}^{N}\frac{1}{R_k}\left(\frac{(v_0/4)}{2\pi}\sin(2\pi\,t_k)\right)^2(2\xi)^2\end{aligned} \quad (46)$$

The value $\hat{\xi}_{mismatch}=\arg\min J_{mismatch}$ minimizing mismatch part the cost function

$$\hat{\xi}_{mismatch}=\frac{\sum_{k=1}^{N}\frac{1}{R_k}\left(z^k-\frac{v_0}{2\pi}\sin(2\pi\,t_k)\right)^2\frac{v_0^2}{2\pi}\sin(2\pi\,t_k)}{2\sum_{k=1}^{N}\frac{1}{R_k}\left(\frac{(v_0/4)}{2\pi}\sin(2\pi\,t_k)\right)^2} \quad (47)$$

If $x(t)$ is measured every 0.5 s from t = 0.5 to t = 5, then $x(t_k,\xi)=0$ for any value of ξ and $z_k=\varepsilon_k$

$$J_{mismatch}(\xi)=\frac{1}{2}\sum_{k=1}^{N}\frac{(\varepsilon_k)^2}{R_k} \quad (48)$$

which is the formula that was already obtained in Eq. (32)

In this case, the denominator of Eq. (47) is not defined and $\hat{\xi}_{mismatch}$ is not defined, because the mismatch part does not depend on ξ and yields an estimation where all possible values of $v_0(\xi)=v_0\,(1+0.5\,\xi)$ are equally likely. Therefore, the value $\hat{\xi}=\arg\min J_{total}$ minimizing the total cost function is also the value minimizing the apriori part of the cost function, i.e., $\hat{\xi}=0$.

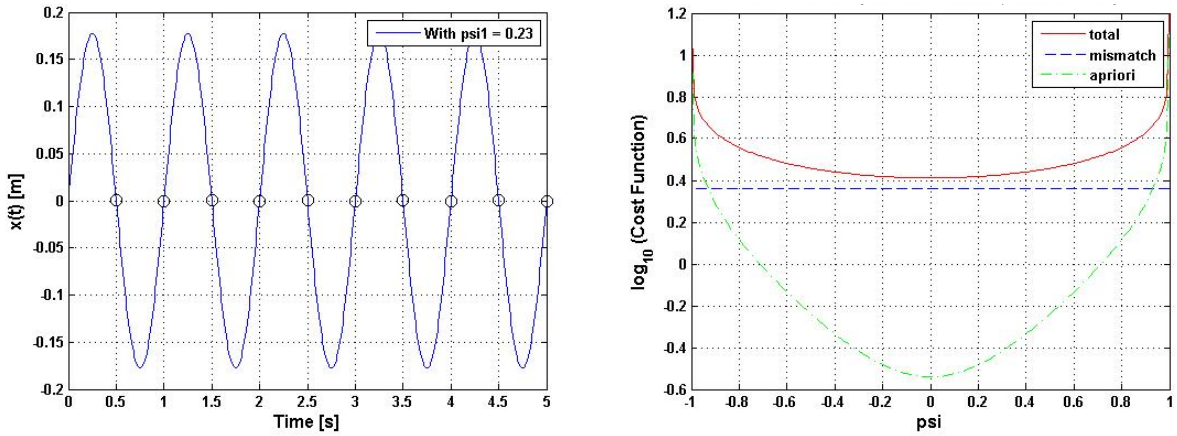
3.2. Possible Impact of Undersampling

Increasing the number of measurements generally yields a better estimation, and as a general rule, sampling above the Nyquist frequency rate should always be done when possible. This section studies the possible impact of

undersampling for two reasons. Measurements might not be available at a rate above the Nyquist frequency rate when this frequency is very high for instance. Another reason is that it will be shown that in some cases, it is possible to know that the estimation is already quite accurate when using only a very few measurements points instead of all of them. It can be useful when computational time is an issue and an answer is needed quickly, which does not prevent from continuing to process the extra information later on if needed, knowing that the extra measurements will generally yields more precision.

The mismatch part of the cost function is driven by the observational errors. The summed contribution of errors makes $J_{mismatch}$ a random variable with a χ^2 distribution with N degrees of freedom. For a relatively large number of measurements this distribution behaves like a normal one with mean N and variance N. The mismatch part does not depend on ξ and yields an estimation where all possible values of v_0 are equally likely. This is illustrated in Figure 4 where the mismatch part of the cost function is constant. The mismatch part does not depend on the noise level in this case since R_k^{-1} is inversely proportional to $(\varepsilon_k)^T(\varepsilon_k)$.

In this case the estimation relies entirely on the apriori part of the cost function, i.e. $-\log(\rho(\xi))$. The estimate coincides with the best initial guess, i.e., $\hat{\xi} = 0$. This is really the worst-case scenario: the frequency of sampling the output is below the Nyquist frequency rate, and the sampling points are exactly those time moments when the displacement is zero and the observations contain no information.



(a) (b) Estimate = 0

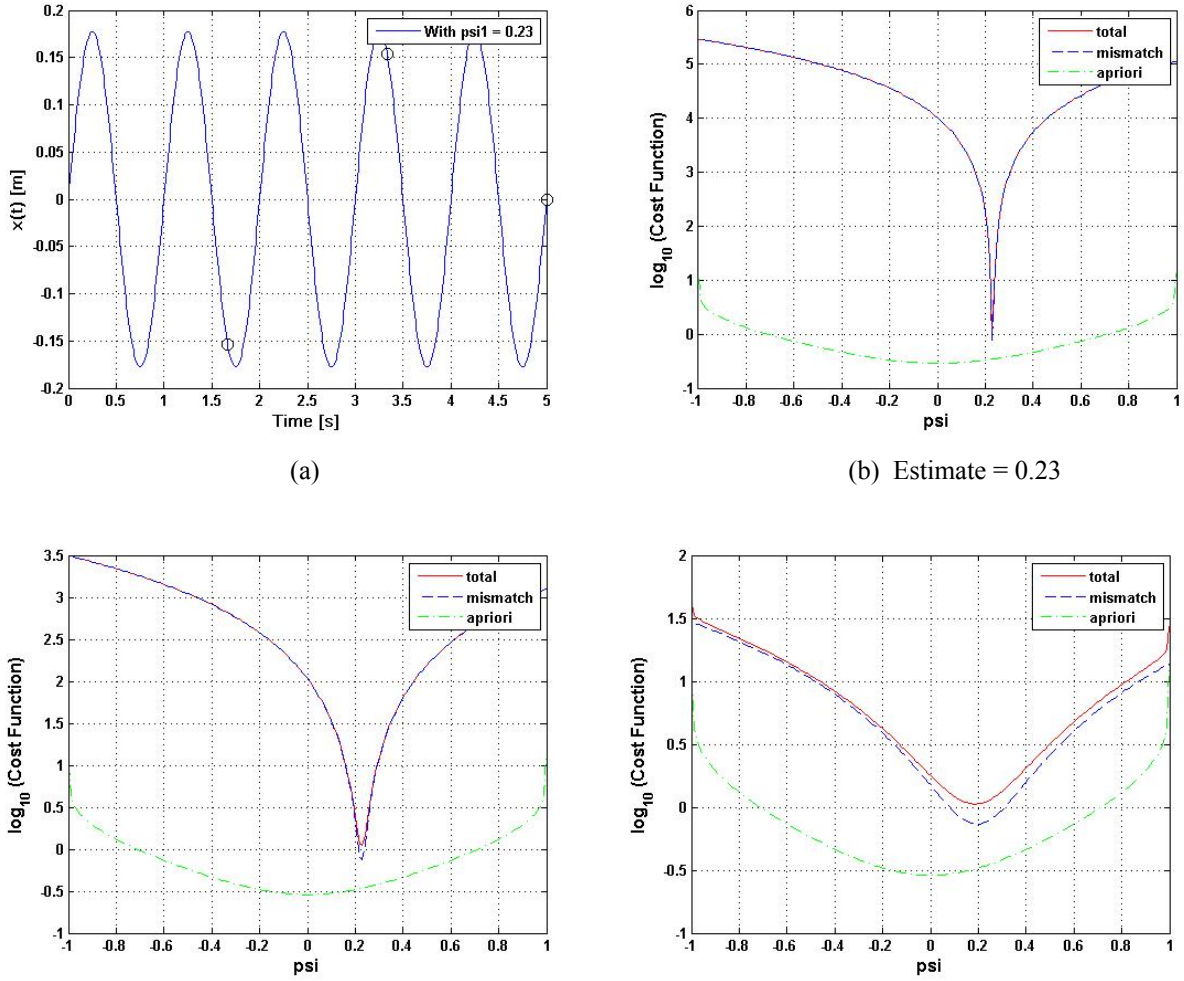
Figure 4. Bayesian Estimation with 10 Time Points

(a) Displacement when no noise added; (b) Estimation with Noise = 1%;

If we measure $x(t)$ at any other additional time point then the Bayesian approach yields an accurate estimation result (for any reasonable amount of measurement noise). We can interpret this fact as follows. If the output sampling is not done at least at the Nyquist frequency, one cannot guarantee that all the relevant information in the output signal is captured, i.e. one cannot guarantee that the mismatch part of the cost function will bring extra information. We still have a PDF of the possible values of the uncertain parameter, but in the worst case scenario, it will be no better than the apriori PDF.

In most practical situations, however, it is very likely that the Bayesian approach will provide an accurate estimate even when the output is sampled below the Nyquist frequency. In the example above a single measurement point is sufficient, provided that the measurement time is not one for which the displacement is zero. This is where the Parameter Estimation and Signal Reconstruction differ. In the above setting of parameter estimation one samples outputs of the system. If the outputs were arbitrary signals then their full reconstruction would require a sufficient sampling frequency. But the outputs are constrained by the input and by the system dynamics, and only a small set of all the possible reconstructed signals are consistent with both the known input and with the equations of motion. The

reconstruction of signals in this small family requires less information than the reconstruction of arbitrary signals. In our example all possible system outputs form a one-parameter family of signals (frequency of 1 Hz, phase equal to zero, and variable amplitude). Consequently a single measurement of the output is almost always sufficient to estimate the single uncertain parameter.



(a)

(b) Estimate = 0.23

(c) Estimate = 0.23

(d) Estimate = 0.19

Figure 5. Bayesian Estimation with 3 Time Points

(a) Displacement when no noise added; (b) Estimation with Noise = 0.01%;

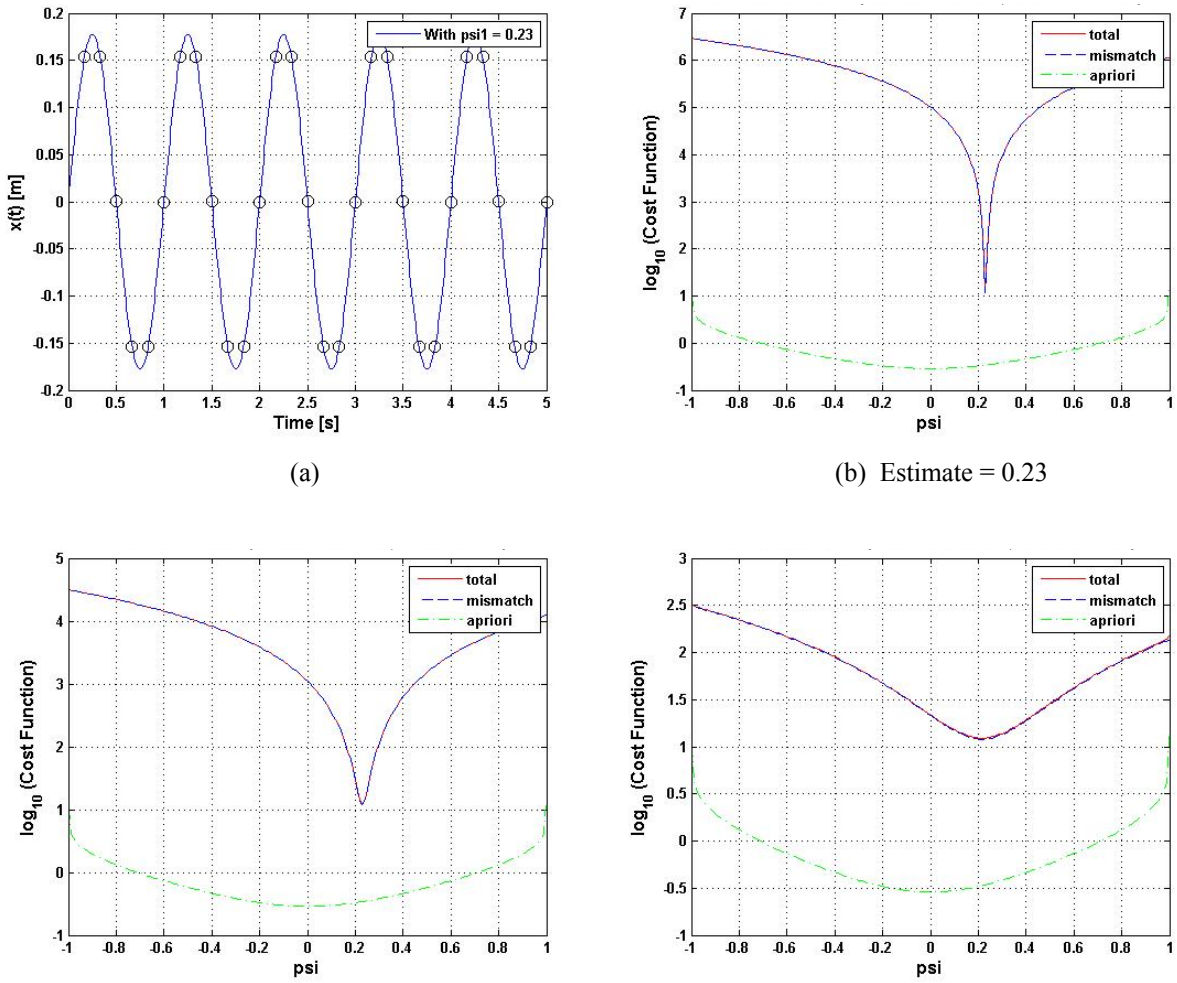
(c) Estimation with Noise = 1%; (d) Estimation with Noise = 10%

The practical question is now how to decide whether the sampling of the output is sufficient. The answer is given by the shape of the Bayesian cost function which indicates whether there is enough information to obtain a good estimate or not. The second derivative of the cost function at the minimum approximates the inverse of the covariance of the uncertainty in the estimate. Loosely speaking, the sharper the minimum of the cost function the more trustworthy the estimate is; and the wider the minimum the larger the estimation error can be.

The role of the shape of the cost function is illustrated in Figure 5, in which only three measurements points for $t > 0$ are used. Different levels of measurement errors lead to different shapes of the cost function, and to different estimation accuracies. For noise levels of 0.01% and 0.1% the total cost function is almost equal to its mismatch part for all values of ξ, it has a sharp minimum, and the Bayesian approach yields an accurate estimate. For very noisy

measurements (10%) the relative weight of the information coming from measurements is smaller, and the relative weight of the apriori information is higher. Consequently the apriori part of the cost function is more significant and the minimum of the total cost function is wider. In this scenario the output sampling is done below the Nyquist frequency, but we know that the estimates are accurate for noise levels of 0.01% and 0.1% because the cost functions have clear minima.

One very accurate output sample would be enough for a perfect estimation. Taking more sample points leads to a better estimation for noisy measurements because the effect of the noise averages out as we take more samples. Figure 6 illustrates the cost function when 30 measurements are used. The relative weight of the mismatch part increases and we get a better estimation when the noise of 10%.



(a) (b) Estimate = 0.23

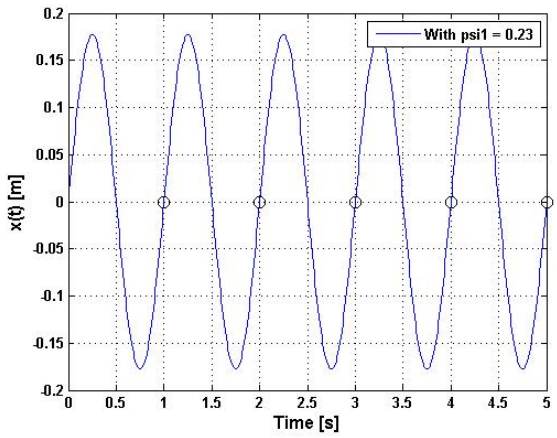
(c) Estimate = 0.23 (d) Estimate = 0.22

Figure 6. Bayesian Estimation with 30 Time Points

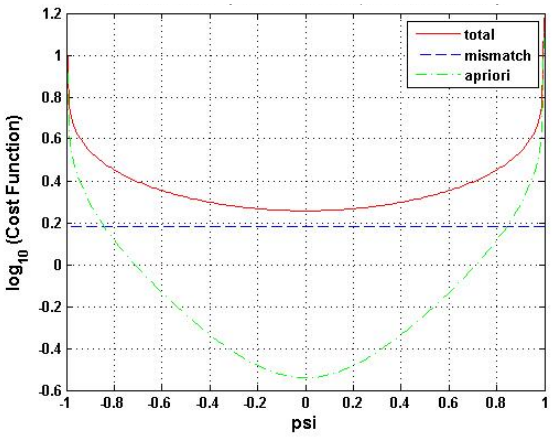
(a) Displacement when no noise added; (b) Estimation with Noise = 0.01%;

(c) Estimation with Noise = 1%; (d) Estimation with Noise = 10%

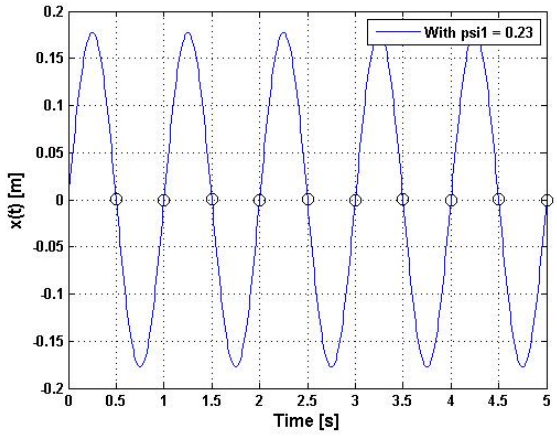
When the extra samples do not bring additional information to the estimation process the cost function changes as shown in Figure 7. The net results of the additional measurements is to add a constant to the mismatch part (which corresponds to the effect of measurement noise). The shape of the cost function does not change, in particular the minimum is not more pronounced, and the quality of the estimate is not improved.



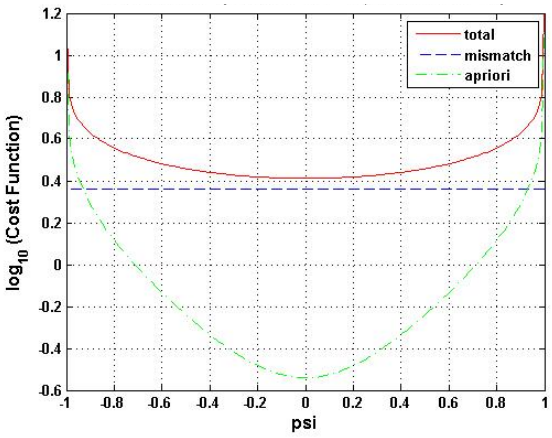
(a)



(b) Estimate = 0



(c)



(d) Estimate = 0

Figure 7. Effect of Adding Sample Points Containing No Useful Information:
(a) Displacement when no noise added with 5 Time Points; (b) Estimation with 5 Time Points and Noise = 0.01%;
(c) Displacement when no noise added with 10 Time Points; (d) Estimation with 10 Time Points and Noise = 0.01%

Next we consider the situation where both $x(t)$ and $v(t)$ are measured at times $t_k > 0$. The observation operator is $H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ and the measured values are $z_k = H\, y(\xi, t_k) + \varepsilon_k = \begin{bmatrix} x(\xi, t_k) \\ v(\xi, t_k) \end{bmatrix} + \begin{bmatrix} \varepsilon_k^x \\ \varepsilon_k^v \end{bmatrix}$. The measurement noise is assumed Gaussian with zero mean and covariance matrix

$$R_k = \begin{bmatrix} \max\left\{10^{-12}, \left(0.01\ \ x(t_k) + 0.001\ \ \max_t (x(t))\right)^2\right\} & 0 \\ 0 & \max\left\{10^{-12}, \left(0.01\ \ v(t_k) + 0.001\ \ \max_t (v(t))\right)^2\right\} \end{bmatrix}$$

The inverse R_k^{-1} can always be computed. One data point at any $t > 0$ is sufficient to estimate our unknown parameter v_0 for low noise levels, as shown in Figure 8. In the general case, however, measurements of the full state vector do not guarantee that they contain useful information when the sampling rate is below the Nyquist frequency.

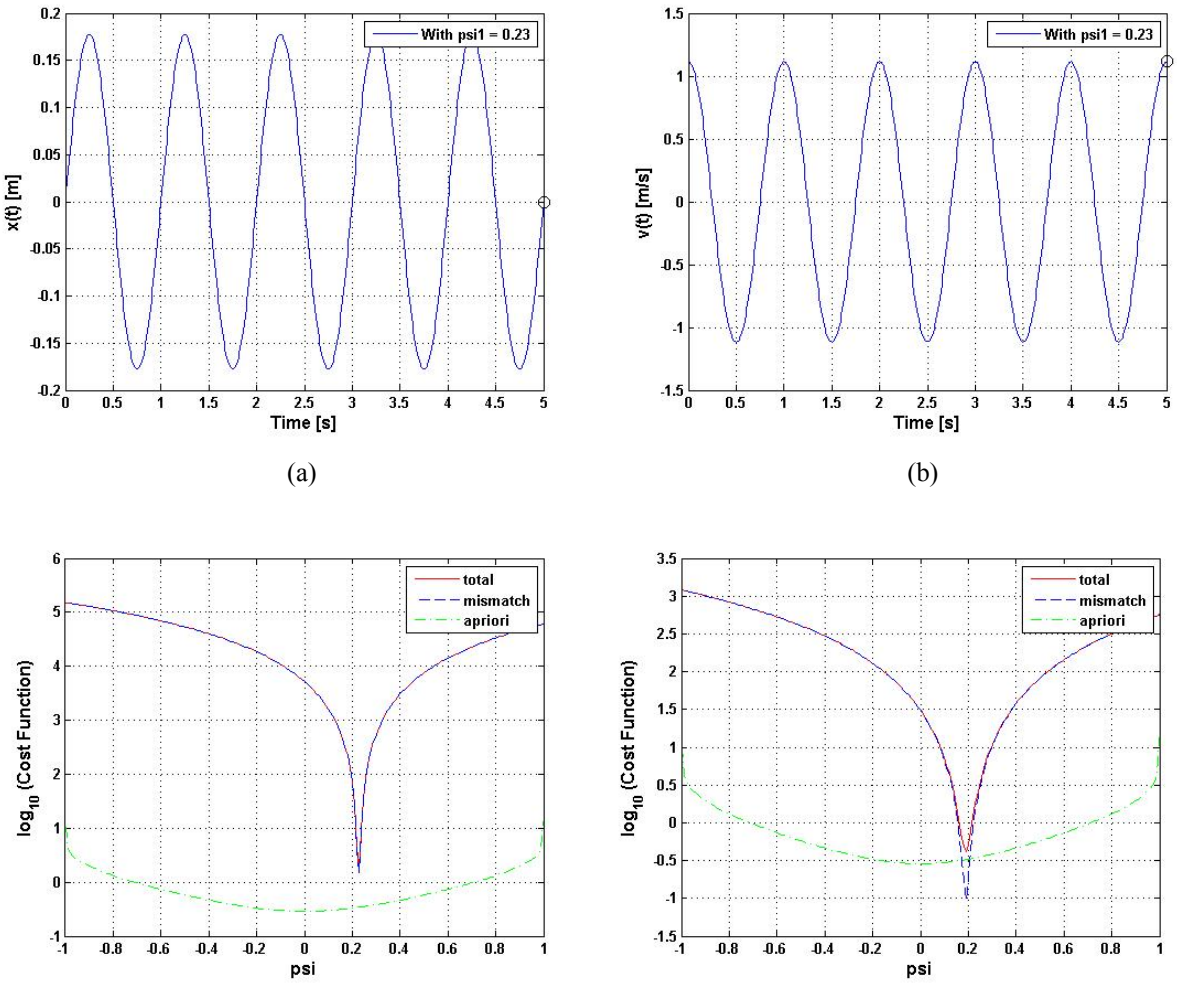


Figure 8. Bayesian Estimation with 1 Time Point when Velocity Measurement is Available

(a) Displacement when no noise added; (b) Velocity when no noise added;

(c) Estimation with Noise = 0.01%; (d) Estimation with Noise = 1%

3.3. Mass-Spring System with Sinusoidal Forcing Function

This section applies the Bayesian approach to the simple Mass-Spring system with sinusoidal forcing function shown in Figure 9.

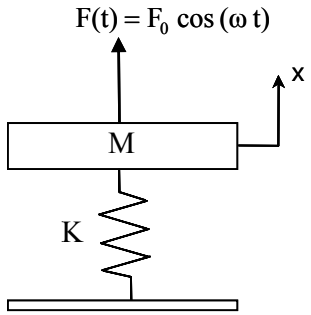


Figure 9. Mass –Spring System with Sinusoidal Forcing Function

The parameters K (the stiffness of the spring), M (the value of the mass) are known. The system is initially at equilibrium, i.e., it has zero initial displacement $x_0 = 0$ and velocity $v_0 = 0$. The problem is to estimate the uncertain the amplitude of the forcing function F_0.

We assume the following apriori information. F_0 has a Beta(1,1) distribution in the range $[500\,N,\ 1500\,N]$ with the most likely value $F_0 = 1{,}000$ N . With $\xi \in [-1,1]$ a Beta(1,1) distributed variable the apriori distribution of F_0 is

$$F_0 = 1{,}000N + \xi \cdot 500N \tag{49}$$

The reference value of the force amplitude is $F_0^{\text{ref}} = 1{,}115\ N$, or $\xi^{\text{ref}} = 0.23$. This reference value is used to generate artificial observations and is not available to the estimation procedure. The numerical values of the other parameters are as follows: $K = 200 \times (1.5 \times 2\pi)^2 \approx 17{,}765$ N/m, $\mathrm{M} = 200\,\mathrm{kg}$, $x_0 = 0$, $v_0 = 0$, and $\omega = \pi$ rad/s $= 0.5$ Hz. Note that $\omega_n = \sqrt{K/M} = 1.5 \times 2\pi$ rad/s $= 1.5$ Hz .

The equation of motion of the system is:

$$\mathrm{M}\ \ddot{x}(t) + K\ x(t) = F_0 \cos(\omega t) \tag{50}$$

The solution is sought in the time interval from t = 0 to t =5 . Since $x_0 = 0$ and $v_0 = 0$, the analytical solution of this equation of motion is [11]:

$$x(t) = \frac{2\,(F_0 / M)}{\omega_n{}^2 - \omega^2}\ \sin\left(\frac{\omega_n - \omega}{2} t\right)\ \sin\left(\frac{\omega_n + \omega}{2}\ t\right) \tag{51}$$

With our numerical values, the displacement of the mass can be written as:

$$x(t) \approx 1.2665 \times 10^{-4}\ F_0\ \ \sin(\pi t)\ \sin(2\pi t) \tag{52}$$

It can be seen that the amplitude of $x(t)$ and the amplitude of $v(t)$ are both proportional to the uncertain parameter F_0, as shown in Figure 10. Therefore, the estimation of F_0 can be in principle based on a single measurement of the output.

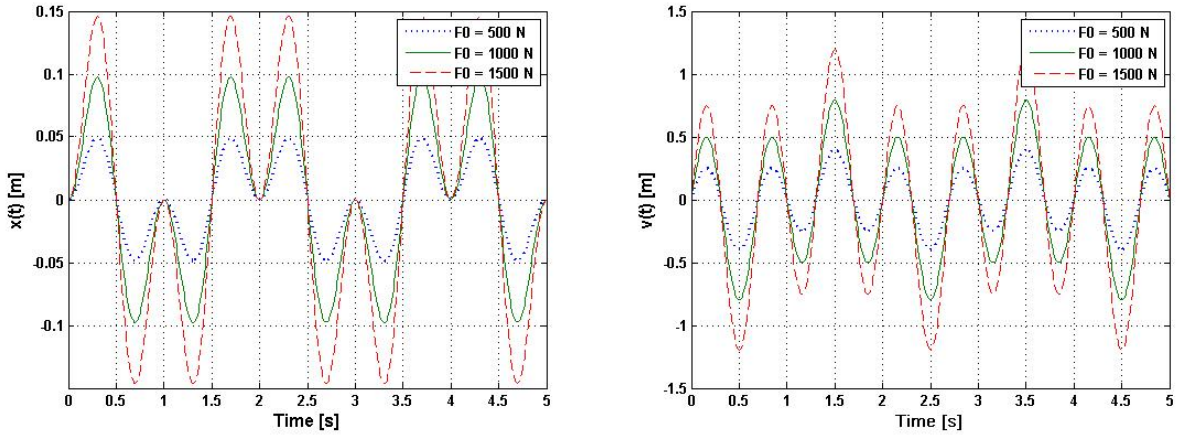
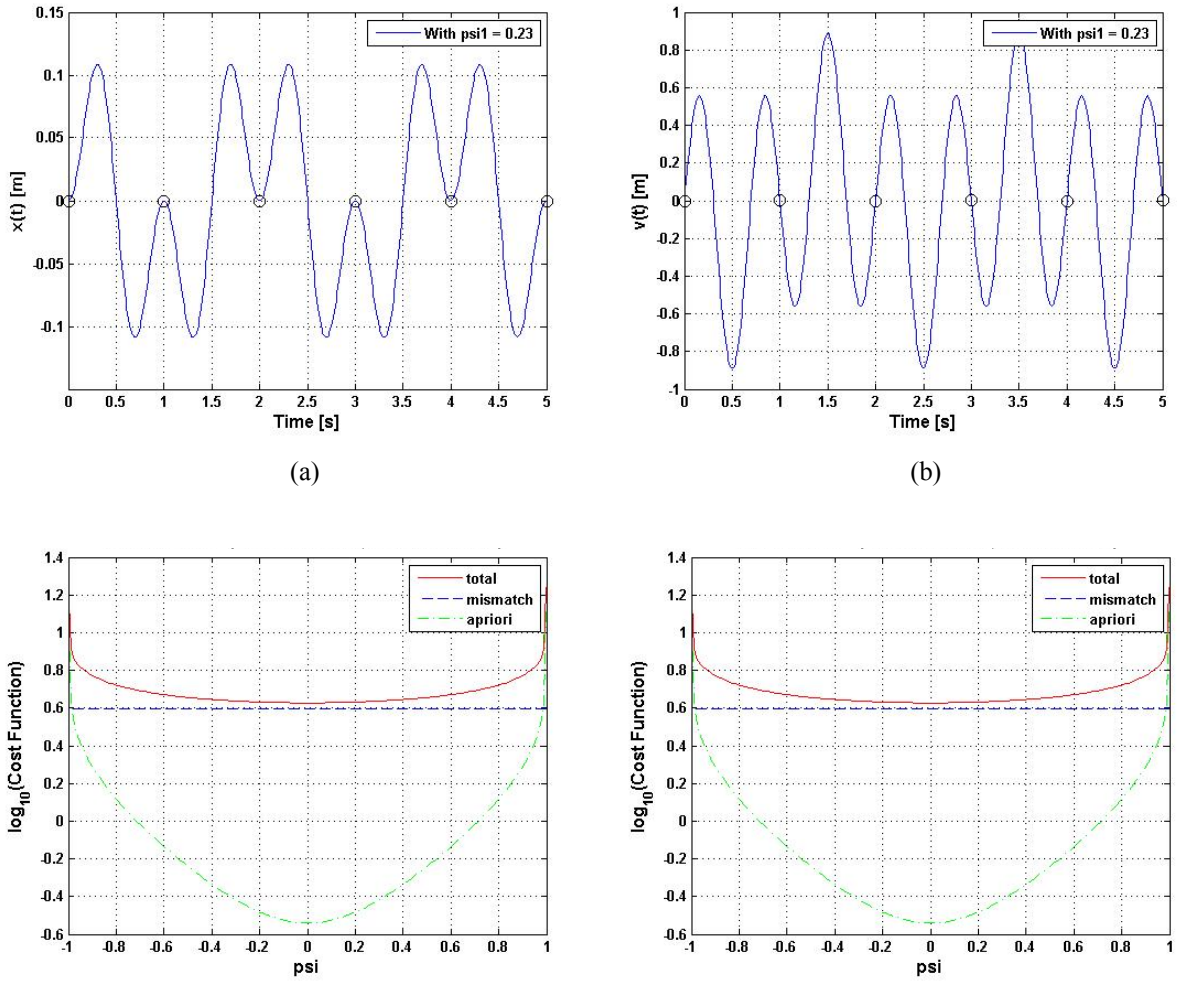


Figure 10. Displacement and Velocities of the Mass –Spring System with Sinusoidal Forcing Function

Figure 11 illustrates the effect of measurements of both $x(t)$ and $v(t)$ at five time points. This sampling provides no information on the uncertain parameter, and in this worst-case scenario the estimate is based solely on apriori information. A sampling of the output below the Nyquist frequency does not guarantee that we get sufficient

information from the output signal about the uncertain parameter. The mismatch part does not depend on the noise level in this case since R_k^{-1} is inversely proportional to $(\varepsilon_k)^T(\varepsilon_k)$, as shown in Eq. (32).



(a) (b)

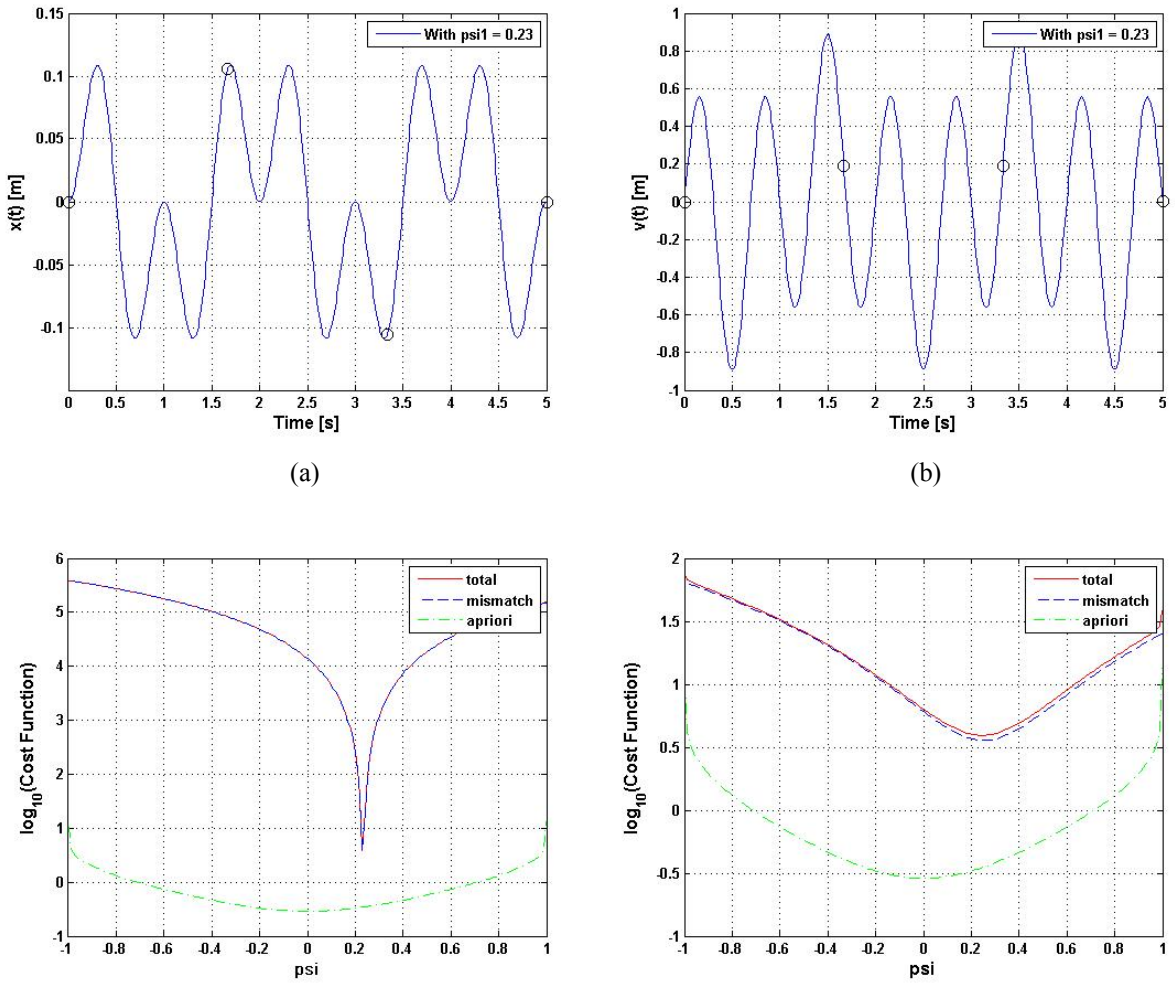
(c) Estimate = 0 (d) Estimate = 0.23

Figure 11. Bayesian Estimation with 5 Time Points when Velocity Measurement is Available

(a) Displacement when no noise added; (b) Velocity when no noise added

(c) Estimation with Noise = 0.01%; (d) Estimation with Noise = 10%

However, we can see in Figure 12 that 3 time measurements yield an accurate estimation for a low noise level, even though the sampling is well below the Nyquist frequency. Once again, the shape of the cost function indicates that for low noise levels we have enough information to accurately estimate our uncertain parameter. While there are many signals with a maximum frequency of 1.5 Hz which fit the observations at the chosen three measurement times, only one of them is consistent with the input signal and with equation of motion.



(a) (b)

(c) Estimate = 0.23 (d) Estimate = 0.25

Figure 12. Bayesian Estimation with 3 Time Points when Velocity Measurement is Available

(a) Displacement when no noise added; (b) Velocity when no noise added

(c) Estimation with Noise = 0.01%; (d) Estimation with Noise = 10%

3.4. Regularization Techniques Applied to a Mass-Spring System with Uncertain Stiffness and Uncertain Mass

This example addresses the issue of non-identifiability. The Bayesian approach is applied to the simple Mass-Spring system shown in Figure 1. The difference with section 3.1 is that the uncertain parameters are different: they are the stiffness of the spring (K) and the value of the mass (M). Our apriori information about the uncertain parameters is expressed in terms of probability densities as follows. The mass has a normal distribution with mean $M_0 = 200\,\text{kg}$ and a standard deviation $\mu = 6.667\,\text{kg}$. The stiffness has also a normal distribution with mean $K_0 = 200\times(2\pi)^2\ \text{N/m} \approx 7{,}895.68\ \text{N/m}$ and standard deviation $\sigma = 1{,}052.8\ \text{N/m}$. We represent the uncertain parameters as functions of a random vector of two independent normal random variables $\xi = (\xi_1, \xi_2)$ as follows

$$M(\xi) = M_0 + \mu\cdot\xi_1, \quad K(\xi) = K_0 + \sigma\cdot\xi_2, \quad \xi_1, \xi_2 \in \mathrm{N}(0,1). \tag{53}$$

We consider the "true" values of the parameters to be $M^{\text{ref}} = 201.533$ kg and $K^{\text{ref}} = 7495.62$ N/m, which correspond to the reference values of the random variables $\xi^{\text{ref}} = \left(\xi_1^{\text{ref}}, \xi_2^{\text{ref}}\right) = (0.23, -0.38)$. These values are not available to the estimation process, but are used in a reference simulation to generate synthetic observations.

We measure the values of the oscillation frequency ω_{obs} (along the reference solution) and use it to derive information about K and M. The measurement errors are assumed to have a normal distribution with zero mean (unbiased) and a standard deviation equal to $\rho = 0.01\sqrt{K_0/M_0} = 0.3938$ rad/s.

This example allows an analytical solution to the Bayesian approach and provides insight into the role of the mismatch and the a priori parts of the cost function in the estimation. The position of the mass is given by:

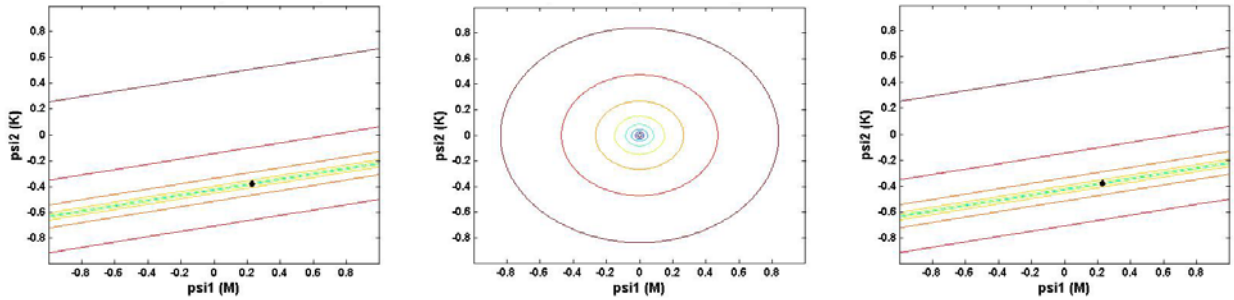
$$x(t) = \sqrt{\frac{v_0^{\,2} + (x_0\,\omega_n)^2}{\omega_n^{\,2}}}\,\sin\left(\omega_n\,t + \tan^{-1}\left(\frac{x_0\,\omega_n}{v_0}\right)\right), \tag{54}$$

and clearly it depends only on $\omega_n = \sqrt{K^{\text{ref}}/M^{\text{ref}}}$.

The Bayesian cost function is defined as:

$$\begin{aligned} J(\xi) &= \frac{1}{2}\,\frac{(M(\xi)-M_0)^2}{\mu^2} + \frac{1}{2}\,\frac{(K(\xi)-K_0)^2}{\sigma^2} + \frac{1}{2}\,\frac{\left(\sqrt{K(\xi)/M(\xi)} - \omega_{\text{obs}}\right)^2}{\rho^2} \\ &= \frac{1}{2}\xi_1^2 + \frac{1}{2}\xi_2^2 + + \frac{1}{2}\,\frac{\left(\sqrt{(K_0+\sigma\xi_2)/(M_0+\mu\xi_1)} - \omega_{\text{obs}}\right)^2}{\rho^2} \end{aligned} \tag{55}$$

The maximum likelihood value of the parameters is the argument that minimizes this cost function, $\hat{\xi} = \arg\min J(\xi)$. The contour plots shown in Figure 13 represent the mismatch part of the cost function, its apriori part, and the total value of the cost function in the space of random variables.



(a) (b) (c)

Figure 13. Contours of the Cost Function: (a) Mismatch part, (b) Apriori Part, (c) Total Cost Function

The magnitude of the apriori part is relatively small and the total cost function is roughly equal to its mismatch part. As expected the mismatch part yields a line of possible minima, because the measured ratio $\omega_n = \sqrt{K/M}$ does not contain information about the individual values of K and M. The point ξ^{ref} is plotted in Figure 13 and it lies on the line of minima. The *Bayesian interpretation is the following: all the pairs (K,M) along the line are equally likely to produce the value of the measured oscillation frequency we. We say that the (individual values of the) parameters K and M are non-identifiable.*

When multiple combinations of uncertain parameter values result in the same observed behavior of the system (same measurements) a regularization approach [1] can be used in estimation. In order to find the most likely parameter

values one increases the relative importance of the apriori knowledge of the system. This is done by multiplying the apriori part by a "regularization coefficient" large enough so that the new total cost function has a clear minimum value along the possible values.

$$J^{\text{regularized}}(\xi,\alpha)=\alpha^2\cdot\left(\frac{1}{2}\,\frac{(M(\xi)-M_0)^2}{\mu^2}+\frac{1}{2}\,\frac{(K(\xi)-K_0)^2}{\sigma^2}\right)+\frac{1}{2}\,\frac{\left(\sqrt{K(\xi)/M(\xi)}-\omega_{\text{obs}}\right)^2}{\rho^2} \tag{56}$$

The net effect of regularization in this example is to reduce the standard deviations in the apriori distributions (to μ/α and σ/α respectively), therefore to increase the trust in the apriori information. The contour plots of the regularized cost functions are shown in Figure 14 for different regularization coefficients.

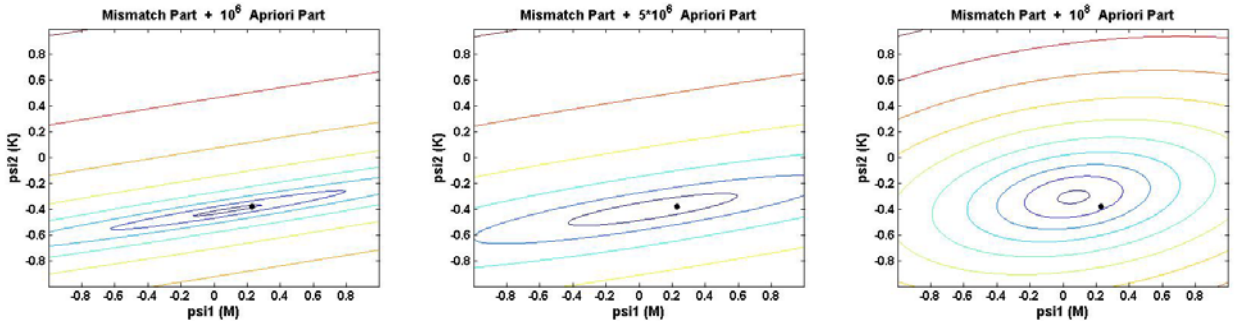


Figure 14. Contour Plots of the Cost Function after Regularization for Different Coefficients

The cost function looks like its mismatch part when the regularization coefficient is very low and looks like its apriori part when the regularization coefficient is very high. As the regularization coefficient gets larger, the line of possible minima becomes an ellipse, and starts moving away from the location of the original line of minima and toward (0, 0), the apriori most likely value. When the line becomes an ellipse with a well defined center, the regularization coefficient is large enough; it should not be further increased as this leads to an increase of the bias in the estimate. In our example $\alpha^2=5\times10^6$ seems to be a good value for the regularization coefficient and it results in the estimated values $(\xi_1,\xi_2)=(0.09,-0.40)$. *The choice of the regularization coefficient is problem-dependent and requires a careful analysis of the resulting estimates. Regularization leads to biased estimates, as stronger assumptions are being artificially imposed.*

3.5. Non-observability

We now discuss the effect of observability on parameter identifiability. Consider a linear system whose evolution depends linearly on a parameter θ. We add a trivial equation for the evolution of the parameter and represent the system as follows:

$$\begin{cases}\dot{x}_1=-2x_1+x_2+\theta+u_1\\ \dot{x}_2=x_1-2x_2+\theta+u_2\\ \dot{\theta}=0\end{cases}\quad\text{with the observed variable}\quad y=x_1-x_2\,. \tag{57}$$

The system is asymptotically stable in x_1,x_2, and neutrally stable in θ. The two states x_1,x_2 can be excited independently. Our goal is to estimate the uncertain parameter θ based on measurements of the output $y(t)$.

The observability matrix of the system is $\begin{bmatrix}1&-1&0\\-3&3&0\\9&-9&0\end{bmatrix}$, and has rank 1 and zeros in the θ column.

This system is non-observable and this leads to the non-identifiability of θ. Specifically, it can be seen that the output rate of change $\dot{y} = \dot{x}_1 - \dot{x}_2 = -3\,(x_1 - x_2) + u_1 - u_2$ does not depend on θ, and therefore the information provided by the measurements cannot distinguish between different values of the parameter. The effect on the cost function is that the mismatch part is constant since $y(t,\xi) - y_{\text{ref}}(t)$ does not depend on ξ. As far as the measurements are concerned all real values of θ are equally likely! Therefore, the result of the Bayesian estimation is based entirely on the apriori knowledge, and equals the most likely apriori value of the parameter. *In this case the non-identifiability problem can be addressed by including measurements of additional states in the estimation procedure.*

3.6. Choice of Excitation

Non-identifiability can also be the result of the choice of the inputs u_1 and u_2. The input signal may not be "rich enough" to excite all the relevant dynamics and the output values are similar for different possible parameter values.

As an example consider the two degree of freedom roll plane model in Figure 15. Let L be the length of the bar of mass M and inertia I. The two springs have equal stiffnesses $K_1 = K_2 = K$. We want to estimate the values of the uncertain parameters M and I from measurements of the left and right displacements $x_1(t)$ and $x_2(t)$.

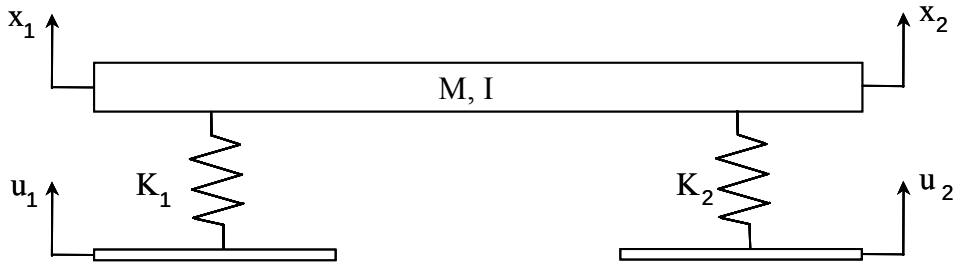


Figure 15. Two Degree of Freedom Roll Plane Model

For small angles (i.e., for $(\ddot{x}_2 - \ddot{x}_1)/L$ small) the equations of motion are:

$$\begin{cases} M\left(\dfrac{\ddot{x}_1 + \ddot{x}_2}{2}\right) + K_1\,(x_1 - u_1) + K_2\,(x_2 - u_2) = 0 \\ I\left(\dfrac{\ddot{x}_2 - \ddot{x}_1}{L}\right) + K_1\,(x_1 - u_1) - K_2\,(x_2 - u_2) = 0 \end{cases} \tag{58}$$

For the same excitations on the left and on the right, i.e. $u_1(t) = u_2(t) = u(t)$ the equations of motion become

$$\begin{cases} \dfrac{M}{2}\,(\ddot{x}_1 + \ddot{x}_2) + K\,(x_1 + x_2 - 2\,u) = 0 \\ \dfrac{I}{L}\,(\ddot{x}_2 - \ddot{x}_1) + K\,(x_1 - x_2) = 0 \end{cases} \tag{59}$$

If $u_1(t) = u_2(t) = u(t)$, $x_1(0) = x_2(0)$, and $\dot{x}_1(0) = \dot{x}_2(0)$, then $x_1(t) = x_2(t)$ for all future times. The second equation of motion is trivially satisfied and the system output does not depend on I (the system has the same evolution for any value of I). This means that the parameter I is non-identifiable. An excitation that is different on the left and on the right would easily lead to outputs that depend on the inertia, and would allow the estimation of this parameter. *In summary an input signal that is not rich enough can lead to non –identifiability. In this case the problem can be addressed by changing the kind of excitations applied to the system.*

3.7. Discussion of the Bayesian approach

The quality of the maximum likelihood estimate is related to the shape of the Bayesian cost function, with a sharp minimum indicating an accurate estimate. Inaccurate estimates can be caused by different factors, including a sampling rate below the Nyquist frequency, non-identifiability, non-observability, and an excitation signal that is not rich enough.

The parameters are non-identifiable when different parameter values lead to identical system outputs. In this case the Bayesian cost function has an entire region of minima (e.g., a valley), with each parameter value in the region being equally likely. A regularization approach based on increasing the weight of the apriori information can be used to select reasonable estimates.

For identifiable and observable systems accurate estimates can be obtained in most cases even if the output signal is sampled below the Nyquist rate. In the worst case, however, sampling below the Nyquist rate cannot guarantee that sufficient information is extracted from the output. In this worst case the apriori information becomes important and the estimate is biased toward the apriori most likely value.

4. SUMMARY AND CONCLUSIONS

This paper applies the polynomial chaos theory to the problem of parameter estimation, using direct stochastic collocation. The maximum likelihood estimates are obtained by minimizing a cost function derived from the Bayesian theorem. This approach is applied to very simple mechanical systems in order to illustrate how the cost function can be affected by undersampling, non-identifiablily of the system, non-observability, and by excitation signals that are not rich enough. Inaccurate estimates can be caused by those different factors. It has been shown that the quality of the maximum likelihood estimate is related to the shape of the Bayesian cost function, with a sharp minimum indicating an accurate estimate. The parameters are non-identifiable when different parameter values lead to identical system outputs. In this case the Bayesian cost function has an entire region of minima (e.g., a valley), with each parameter value in the region being equally likely. Regularization techniques can still yield most likely values among the possible combinations of uncertain parameters resulting in the same time responses than the ones observed. This was illustrated using a simple spring-mass system. For identifiable and observable systems accurate estimates can be obtained in most cases even if the output signal is sampled below the Nyquist frequency. In the worst case, however, sampling below the Nyquist rate cannot guarantee that sufficient information is extracted from the output. In this worst case the apriori information becomes important and the estimate is biased toward the apriori most likely value.

The proposed method has several advantages. Simulations using Polynomial Chaos methods are much faster than Monte Carlo simulations. Another advantage of this method is that it is optimal; it can treat non-Gaussian uncertainties since the Bayesian approach is not tailored to any specific distribution.

In the second part of this article, this new parameter estimation method is illustrated on a nonlinear four-degree-of-freedom roll plane model of a vehicle in which an uncertain mass with an uncertain position is added on the roll bar.

ACKNOWLEDGEMENTS

This research was supported in part by NASA Langley through the Virginia Institute for Performance Engineering and Research award. The authors are grateful to Dr. Mehdi Ahmadian, Dr. Steve Southward, Dr. John Ferris, and Mr. Carvel Holton for many fruitful discussions on this topic.

REFERENCES

[1] Aster , R.C., Borchers, B., and Thurber, C.H.: “Parameter Estimation and Inverse Problems”, Elsevier Academic Press, 2005. ISBN 0-12-065604-3.

[2] Blanchard, E., Sandu, C., and Sandu, A. – “A Polynomial-Chaos-based Bayesian Approach for Estimating Uncertain Parameters of Mechanical Systems”, *Proceedings of the ASME 2007 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference IDETC/CIE 2007, 9th International Conference on Advanced Vehicle and Tire Technologies (AVTT),* September 4-7, 2007, Las Vegas, Nevada, USA

[3] Blanchard, E., Sandu, A., and Sandu, C. – "Parameter Estimation Method Using an Extended Kalman Filter", *Proceedings of the Joint North America, Asia-Pacific ISTVS Conference and Annual Meeting of Japanese Society for Terramechanics*, June 23-26, 2007, Fairbanks, Alaska.

[4] Cohn, S. E. "An Introduction to Estimation Theory", *J. Meteor. Soc. Japan 75 (B)* (1997), 257-288.

[5] Ghanem, R.G., and Spanos, P.D. – "Stochastic Finite Elements", Dover Publications Inc, Mineola, New York, 2003.

[6] Ghanem, R.G., and Spanos, P.D. – "Polynomial Chaos in Stochastic Finite Element", *Journal of Applied Mechanics*, 1990, Vol. 57, 197-202.

[7] Ghanem, R.G., and Spanos, P.D. – "Spectral Stochastic Finite-Element Formulation for Reliability Analysis", *ASCE Journal of Engineering Mechanics*, 1991, Vol. 117, No. 10, 2351-2372.

[8] Ghanem, R.G., and Spanos, P.D. – "A Stochastic Galerkin Expansion for Nonlinear Random Vibration Analysis", Probabilistic Engineering Mechanics, 1993, Vol. 8, No. 3, 255-264.

[9] Halton, J. H., Smith, G. B. "Radical-inverse quasi-random point sequence". *Communications of the ACM*, 7(12):701–702, Dec. 1964.

[10] Hammersley, J. M. "Monte Carlo Methods for Solving Multivariables Problems", *Ann. New York Acad. Sci.*, 86:844–874, 1960.

[11] Inman, D. J., "Engineering Vibration", Second Edition, Prentice Hall, Inc., 2001. ISBN 0-13-726142-X.

[12] Li, L., Sandu, C., and Sandu, A. – "Modeling and Simulation of a Full Vehicle with Parametric and External Uncertainties", *Proc. of the 2005 ASME Int. Mechanical Engineering Congress and Exposition, 7th VDC Annual Symposium on "Advanced Vehicle Technologies", Session 4: Advances in Vehicle Systems Modeling and Simulation*, Paper number IMECE2005-82101, Nov. 6-11, 2005, Orlando, FL.

[13] Sandu, A., Sandu, C., and Ahmadian, M. – *Modeling Multibody Dynamic Systems With Uncertainties. Part I: Theoretical and Computational Aspects*, Multibody System Dynamics, Publisher: Springer Netherlands, ISSN: 1384-5640 (Paper) 1573-272X (Online), DOI 10.1007/s11044-006-9007-5, pp. 1-23 (23), June 29, 2006.

[14] Sandu, C., Sandu, A., and Ahmadian, M. – *Modeling Multibody Dynamic Systems With Uncertainties. Part II: Numerical Applications*, Multibody System Dynamics, Publisher: Springer Netherlands, ISSN: 1384-5640 (Paper) 1573-272X (Online), DOI: 10.1007/s11044-006-9008-4, Vol. 15, No. 3, pp. 241 - 262 (22), April 2006.

[15] Sandu, C., Sandu, A., Chan, B.J., and Ahmadian, M. – "Treating Uncertainties in Multibody Dynamic Systems using a Polynomial Chaos Spectral Decomposition", *Proc. of the ASME IMECE 2004, 6th Annual Symposium on "Advanced Vehicle Technology*", Paper number IMECE2004-60482, Nov. 14-19, 2004 Anaheim, CA.

[16] Sandu, C., Sandu, A., Chan, B.J., and Ahmadian, M. – "Treatment of Constrained Multibody Dynamic Systems with Uncertainties", *Proc. of the SAE Congress 2005*, Paper number 2005-01-0936, April 11-14, 2005, Detroit, MI.

[17] Sandu, C., Sandu, A., and Li, L. – "Stochastic Modeling of Terrain Profiles and Soil Parameters", *SAE 2005 Transactions Journal of Commercial Vehicles,* V114-2, 2005-01-3559, 211-220, Feb, 2006.

[18] Sohns, B., Allison, J., Fathy, H. K., Stein, J. L. "Efficient Parameterization of Large-Scale Dynamic Models Through the Use of Activity Analysis", *Proceedings of the ASME IMECE 2006,* IMECE2006, Nov 5-10, 2006, Chicago, Illinois.

[19] Xiu, D., Lucor, D., Su, C.-H., and Karniadakis, G.E. – "Stochastic Modeling of Flow-Structure Interactions using Generalized Polynomial Chaos", *J. Fluids Engineering*, Vol. 124, 51-59, 2002.

[20] Xiu, D., and Karniadakis, G. E. – "The Wiener-Askey Polynomial Chaos for Stochastic Differential Equations", *Journal of Sci Comput*, 2002: Vol. 24, No. 2: 619-644.

[21] Xiu, D., and Karniadakis, G.E. – "Modeling Uncertainty in Flow Simulations via Generalized Polynomial Chaos", *Journal of Computational Physics*, 2003: Vol. 187: 137-167.

[22] Xiu, D., and Karniadakis, G.E. – "Modeling Uncertainty in Steady-state Diffusion problems via Generalized Polynomial Chaos", Computer Methods in Applied Mechanics and Engineering, 2002: Vol. 191: 4927-4928.

[23] Zhang, D., Lu., Z. "An efficient, high-order perturbation approach for flow in random porous media via Karhunen–Loeve and polynomial expansions", *J Comp. Phys.* 54:265–291 (2006)

LIST OF FIGURES