Available online at www.eccomasproceedia.org
Eccomas Proceedia UNCECOMP (2021) 1-14

ECCOMAS Proceedia

UNCECOMP 2021
4th ECCOMAS Thematic Conference on
Uncertainty Quantification in Computational Sciences and Engineering
M. Papadrakakis, V. Papadopoulos, G. Stefanou (eds.)
Streamed from Athens, Greece, 28 -30 June 2021

NONLINEAR GAUSSIAN PROCESS LATENT FORCE MODELS FOR INPUT ESTIMATION IN HYSTERETIC SYSTEMS

T.J. Rogers[1], J. D. Longbottom[1], K. Worden[1] & E. J. Cross[1]

[1] Dynamics Research Group
Department of Mechanical Engineering, University of Sheffield
Mappin Street, Sheffield
S1 3JD, UK
e-mail: {tim.rogers, jdlongbottom1, k.worden, e.j.cross}@sheffield.ac.uk

Abstract. *One outstanding challenge in structural dynamics is lack of access to measurements while systems are in operation. For example, this could be wind or wave loads on offshore structures, contact forces between a vehicle's wheels and a roadway, the action of an earthquake on a tall building. In view of this, one specific and important problem is to infer both the unknown inputs to a dynamic system and its internal states. The difficulty of this task is linked to the inherent uncertainty in the system, which has two key sources. The first issue is the presence of noise on the measurements or uncertainty in the dynamics of the system (process noise); the second, is that it may not be possible to define* a priori *a functional form for the loading signal. The Gaussian process latent force model is a tool to address both of these tasks, it couples a known dynamic system with a flexible non-parametric representation of the unknown forcing. This representation is chosen to be a Gaussian process in time, which is used to estimate the distribution of possible functions which could have been the unmeasured inputs to the system. One key benefit of this approach is that it provides closed-form expressions for the process noise in the system, based on the characteristics of this function. This paper extends the range of systems for which this type of model can be applied to those exhibiting hysteresis. These systems represent a significant increase in difficulty; not only do they introduce a nonlinearity into the system parameters, but this nonlinearity requires an additional hidden state. It will be shown in this paper how a particle Gibbs approach allows Bayesian inference over the intractable state-space model that describes these systems. This approach allows joint input-state inference to be performed even in these highly nonlinear systems.*

Keywords: Bayesian methods, Particle filtering, Monte Carlo methods, Grey-box modelling.

ISSN:2623-3339
Peer-review under responsibility of the organizing committee of UNCECOMP 2021.
doi: 10.7712/120221.8017.18937

1 INTRODUCTION

This paper will address the problem where output measurements from a nonlinear dynamic system are available, from which the practitioner would like to infer the internal (hidden) states of that system and its unmeasured inputs. A motivating example may be the situation where a suspension system for a vehicle is identified in a laboratory environment, then in use, one wishes to investigate the displacement and velocity of that dynamic element and the unmeasured excitation from the road, based on measurements of the acceleration in the vehicle. One important class of systems for which input estimation is particularly difficult is *hysteretic* systems. Hysteresis, informally, introduces a memory effect into the system which can greatly complicate identification, including input-state estimation. The contribution of this work is to extend previous work on nonlinear input-state estimation [1] to the case of a hysteretic nonlinearity in order to demonstrate the general and powerful approach of nonlinear Gaussian process latent force modelling.

One sensible framework for attempting the challenge of input (load) estimation is that of a Bayesian state-space model (SSM) [2]. If measurements of the input to the system are available, recovering the hidden states of the system is achieved by inferring the filtering or smoothing distribution of the model. However, in the situation being considered here, these inputs are unknown. It would be possible to assume that the system was excited with a white-noise; however, this is often not true. It is necessary then to also model the unknown inputs to the system with a view to inferring them. The challenge becomes that of making an appropriate choice of tool to model these unknown inputs and how this may be incorporated into the model of the dynamical system.

It is desirable that the model chosen to represent the forcing should be flexible enough to accommodate the diversity of functions which may represent these unknown inputs. For this reason, a white-noise assumption on the input may be too restrictive. Instead, it may be appropriate to describe some of the characteristics of the function which may have generated the forcing, for instance how many times differentiable it is. Such a tool exists in the Gaussian process (GP), a flexible Bayesian nonparametric regression model (for an introduction see [3] or [4]). Given the existence of such an approach, a system can be imagined which has some defined dynamical structure where the input to that system is modelled as a GP in time which provides a richer class of possible forcing signals.

Alvarez et al. [5] proposed such an approach for linear systems which could be modelled as a dynamical system forced by a GP; they term this the *Latent Force Model*. By introducing this first or second-order dynamic component to the data-modelling process, they increased the range of scenarios which could be effectively modelled using the Gaussian process. However, the implementation and training of such models can be difficult and time consuming. A significant reduction in this complexity can be achieved given work by Hartikainen and Särkkä [6], which shows that a temporal GP can be written as an equivalent linear Gaussian state-space model (LGSSM) and an identical solution to the full GP is recovered by application of a Kalman filter and Rauch-Tung-Striebel (RTS) smoother. It is then possible to transform the GP latent force model of Alvarez et al. [5] into an LGSSM, where the states represent the dynamic system augmented by one or more additional states which are equivalent to the GP input to the system. This link was noted in [7]; however, one limitation remains — the dynamic model is restricted to linear systems. This paper will present a methodology for removing this limitation such that a nonlinear latent force model can be learnt, also within the state-space framework. In particular, the extension to more difficult hysteretic nonlinearities is attempted. The main question being investigated is: ”Does the introduction of additional dynamic states, because of hysteresis, impede joint input-state estimation with nonlinear Gaussian process latent force models?”.

2 NONLINEAR LATENT FORCE MODELS

Beginning with the case of a linear latent-force model, for a second-order dynamic system, the model being considered is,

$$M_s\ddot{\mathbf{q}} + C_s\dot{\mathbf{q}} + K_s\mathbf{q} = U, \qquad U \sim \mathcal{GP}\left(0, k(t,t')\right) \tag{1}$$

where M_s, C_s, and K_s are the mass, damping and stiffness matrices of a second order system which is forced by an unknown input U; the subscript s is used to indicate that these are the system matrices. The unknown input U is modelled as a Gaussian process with a zero mean and a covariance kernel (function) in time, $k(t,t')$. Defining the states of the model $\mathbf{x} = [\mathbf{q}, \dot{\mathbf{q}}, \mathbf{u}]^{\mathrm{T}}$, with $\mathbf{u}$ being the augmented states which correspond to the GP, this model has an equivalent continuous-time state-space representation,

$$\dot{\mathbf{x}}(t) = F\mathbf{x}(t) + v(t) \tag{2}$$

$$\mathbf{y}(t) = H\mathbf{x}(t) + w(t) \tag{3}$$

with H being the matrix which defines the observation of the states; for example, in the case of observing the acceleration of the system, $\ddot{\mathbf{q}}$, $H = [-M_s^{-1}K_s, -M_s^{-1}C_s, M_s^{-1}, 0]$. This observation is subject to white noise $w(t)$ on the measurements. The dynamics of the process are captured in the matrix F which can be considered to be made up of four block matrices,

$$F = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} = \left[\begin{array}{cc|c} 0 & \mathbb{I} & 0 \\ -M_s^{-1}K_s & -M_s^{-1}C_s & M^{-1} \\ \hline 0 & 0 & F_{GP} \end{array}\right] \tag{4}$$

In the top left, F_{11} is a block of the matrix which corresponds to the linear second-order dynamics of the system; in the top right, F_{12} relates the forcing states $\mathbf{u}$ to this dynamical system. The F_{21} block is zero, as the dynamics are assumed not to affect the forcing evolution in time, i.e. the applied forces are independent of the dynamics. Finally, the matrix F_{22} contains the state-space representation of the Gaussian process in time. How this matrix F_{22} is formed will be briefly reviewed now and it will be seen how it, along with the process noise $v(t)$ is fully defined by the GP over U.

Following [6], for a GP with a stationary covariance function $k(t,t')$, it is possible to consider the power spectral density of that covariance. Taking the popular Matérn 3/2 kernel as an example, with $r = |t - t'|$, one has,

$$k(r) = \sigma_f^2(1 + \frac{\sqrt{3}r}{\ell})\exp\left\{-\frac{\sqrt{3}r}{\ell}\right\} \tag{5}$$

which is governed by two *hyperparameters*; the length scale ℓ and the signal variance σ_f^2. Taking the Fourier transform gives the spectral density of the process as,

$$S(\omega) = 4\sigma_f^2\lambda^3(\lambda^2 + \omega^2)^{-2} \tag{6}$$

where $\lambda = \sqrt{3}/\ell$. Observing that this is a rational function with a denominator which is a polynomial in ω^2, Hartikainen and Särkkä [6] apply a spectral factorisation on this density to show that this is equivalent to a stochastic differential equation (SDE) of order p, if the denominator is a polynomial of order $(\omega^2)^p$. More intuitively, the mode can be seen as an ordinary differential equation (ODE) of order p forced by a white noise. Interestingly, this formulation links back to the interpretation of a Gaussian process as a linear filter on a continuous white noise sequence. For the Matérn 3/2 kernel being used as an example, this gives the SDE,

$$\frac{d^2u}{dt^2} - 2\lambda\frac{du}{dt} - \lambda^2 = \nu(t) \tag{7}$$

where $\nu(t)$ is a white noise process with spectral density q,

$$q = \frac{12\sqrt{3}\sigma_f^2}{\ell^3} \tag{8}$$

It should be noted at this point, that this formulation gives the state-space form for a single output GP; however, it is trivial to extend this to a number of independent GPs acting on the different degrees of freedom in equation (1).

Therefore, for a single-degree-of-freedom harmonic oscillator forced by a GP with a Matérn 3/2 covariance function, the equation of motion,

$$m\ddot{q} + c\dot{q} + kq = u, \quad u \sim \mathcal{GP}\left(0, k(t,t')\right) \tag{9}$$

has an equivalent state-space form,

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -k/m & -c/m & 1/m & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -2\lambda & -\lambda^2 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \nu(t) \tag{10}$$

$$\mathbf{y}(t) = \begin{bmatrix} -k/m & -c/m & 1/m & 0 \end{bmatrix} \mathbf{x}(t) + w(t) \tag{11}$$

when one is observing acceleration and if the augmented state vector is $\mathbf{x} = [q, \dot{q}, u, \dot{u}]$. The solution to this system can then be found by discretising and applying the Kalman filter and RTS smoother, since this is a LGSSM. For examples of the linear case, see [8], where inference over the model is extended to include the hidden states of the oscillator, the unknown input U and the parameters of the model M_s, C_s, K_s.

This paper however, is concerned with a more general case where the ODE which describes the dynamic system is nonlinear. Considering a single-degree-of-freedom system, this means inferring the displacement and velocity of an oscillator with an unknown input which can be expressed as,

$$m\ddot{q} + f(q, \dot{q}) = u, \quad u \sim \mathcal{GP}\left(0, k(t,t')\right) \tag{12}$$

under the same assumption that the unmeasured inputs can be modelled as a GP in time with zero mean and some stationary covariance $k(t,t')$. Here $f(q,\dot{q})$ is some function of the displacement and velocity of the oscillator, it is possible to recover the linear case as a subset of this model by setting $f(q,\dot{q}) = c\dot{q} + kq$.

It will now be seen how a similar procedure as for the linear case can be followed to develop a *nonlinear latent force model*. This approach is discussed in further detail in [8] where a similar methodology is applied to a Duffing oscillator. Inspecting the structure of the SSM in equations (10) and (11), it is clear that the state vector $\mathbf{x}$ can be viewed in two parts. The states related to the "physical" dynamical system (states related to q and $\dot{q}$) and the augmented states which are the GP modelling the forcing (in the case of the Matérn 3/2 kernel, u and $\dot{u}$). This interpretation gives some indication of how a nonlinear equivalent may be structured. The change in the system equation to include nonlinear dynamics will result in the system states related to q and its derivatives becoming nonlinear; however, the representation of the GP remains unchanged.

It is now possible to construct a nonlinear SSM with a similar structure to that seen in the linear case,

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} \dot{q} \\ \frac{u}{m} - \frac{1}{m} f(q, \dot{q}) \\ \dot{u} \\ -2\lambda u - \lambda^2 \dot{u} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \nu(t) \tag{13}$$

$$y(t) = \frac{u}{m} - \frac{1}{m} f(q, \dot{q}) + w(t) \tag{14}$$

This form covers many commonly-encountered nonlinearities; however, for certain types — notably hysteretic systems — additional states also need to be included in the model. For the purposes of this paper, the important point is that to apply the GP latent force approach, it is possible to augment the state vector of the system with a number of states which are a direct equivalent to assuming a GP in time as the input signal. The challenge then is to perform inference over this extended state-space model, to recover the smoothing distribution; in the vast majority of cases, this distribution will not be available in closed form and some numerical estimation must be employed. As a point of interest that will not be covered in any more depth, it would also be possible to accommodate inputs to the system which appear in a nonlinear manner as opposed to a simple additive forcing.

2.1 Inference

While the modification to the linear version of the GP latent force model may seem minor, it unfortunately severely complicates the inference procedure. Remembering that the quantities of interest are the hidden states of the model $\mathbf{x}$, which include the internal states of the oscillator and those related to the GP, the task to be solved (from a Bayesian filtering perspective) is one of inferring the smoothing distribution of this state-space model. That is to recover the distribution[1] $p\left(\mathbf{x}_{1:T} | \mathbf{y}_{1:T}\right)$. Since the model is no longer linear with Gaussian noise, it is not (generally) possible to recover this distribution exactly. Instead some approximation must be made.

The problem being addressed is one of estimating a high-dimensional intractable posterior distribution, which is the smoothing distribution of the nonlinear filter. This challenge will be addressed by the use of a particle filter. The particle filter can be used to form an efficient approximation to the filtering distribution of a nonlinear state-space model in a Monte Carlo manner. A number of weighted point masses (or particles) propagate through time and by repeated proposal, weighting and resampling, they form an importance sampling approximation to the filtering distribution at every time step; an introduction is given in Doucet and Johansen [9]. In general, algorithms of this type are referred to as *Sequential Monte Carlo* (SMC) methods and their applicability has been shown to extend beyond inference of the filtering distributions

[1] The notation $a : b$ is adopted to indicate the discrete index from time a to time b inclusively, e.g. $\mathbf{x}_{1:T}$ would denote the state vector $\mathbf{x}$ at all time points from step 1 to step T.

of nonlinear state-space models; for example, Andersson Naesseth et al. [10] show how this approach can be applied to probabilistic graphical models.

However, in this case, the object of interest is the smoothing distribution of the state-space model. A naïve approach to determining this distribution would be to apply a Markov Chain Monte Carlo (MCMC) approach; there are two major limitations to this, firstly, the likelihood of the system is not available in closed form. Secondly, the smoothing distribution can be very high-dimensional and it becomes far harder to define efficient proposals within standard MCMC schemes such as the Metropolis-Hastings algorithm. Andrieu et al. [11] show how the SMC approach can be combined with an MCMC framework, to allow more efficient inference over high dimensional distributions encountered in nonlinear state-space models. By showing that the estimate of the marginal likelihood of the filter provided by the SMC algorithm $\hat{p}_\theta(\mathbf{x}_{1:T}|\mathbf{y}_{1:T})$ can be used within a Metropolis-Hastings step, they create a valid MCMC algorithm. It is also shown that it is possible to construct an equivalent to the Gibbs sampling procedure, an approach which will be used in this paper.

In the case considered here, it will possible to employ a slightly more sophisticated approach, where multiple variables are sampled at once, specifically sampling all of the states together $\mathbf{x}_{1:T}$. A form of particle filter over the nonlinear state-space model will be able to provide an efficient proposal for the high-dimensional distribution being considered; in fact, it will be possible to ensure that the proposal generates a valid sample from the smoothing distribution on every iteration. In order to ensure the ergodicity of the Markov kernel which samples these states they are conditioned upon the previous sample, this gives rise to the *Conditional Particle Filter* (CPF) [12]. In a CPF, one state trajectory is fixed *a priori* which is included in the weighting and resampling steps. Conceptually, one could think of this as "anchoring" the particle filter to the previously-sampled state trajectory and ensuring it does not drift too far in a single iteration.

One challenge encountered when learning the smoothing distribution in an MCMC manner with a CPF is poor mixing in the chain because of path degeneracy, a phenomenon where re-sampling can mean particles all share a few common ancestors. This poor mixing more acutely affects samples of the states close to the beginning of the time series. To overcome this short-coming a simple yet powerful modification to the CPF approach was made by Lindsten et al. [13]. The contribution of that work is to include an *ancestor sampling* step in the CPF algorithm. At every point in time a new ancestor for the reference trajectory is sampled. Before presenting this algorithmically, it is worth considering what is being asked in the ancestor sampling step: "Given the location of the reference trajectory at time t, which of the particles at time $t-1$ could have generated this sample?". This possible ancestor for the reference trajectory at time t is then sampled based on those proposal probabilities.

The algorithm for inferring the smoothing distribution using an MCMC approach incorporating the CPF with ancestor sample will now be given, a more thorough review and comparison with an alternative method can be found in [12]. It should also be noted that, should the (hyper)parameters of the system need to be inferred this can be done as part of a full Particle Gibbs with Ancestor Sampling [13] scheme; an example of this approach in the context of a GP latent force model can be found in [1].

The algorithm for sampling P state trajectories using this approach can be found in Algorithm 1. Starting from some initial trajectory $X[0]$, each trajectory is drawn conditioned on the previous sample $X[p-1]$ by means of a CPF with ancestor sampling. The CPF runs as in the literature, a particle system is proposed from a prior $q(\mathbf{x}_1)$, except the final particle which is assigned the reference trajectory value; all particles are assigned equal weights. These particles are then weighted by some operation $W_{\theta,1}(\mathbf{x}_1^i)$; in the case of the bootstrap filter, this is the

Algorithm 1 MCMC Smoothing From a Conditional Particle Filter with Ancestor Sampling

Set $X[0]$ ▷ Initial reference trajectory.
for p=1,...,P **do**
    Sample $\mathbf{x}_1^i \sim q(\mathbf{x}_1)$ for $i = 1, \ldots, N-1$
    Set $\mathbf{x}_1^N = X[p-1]$
    Calculate $w_1^i = W_{\theta,1}(\mathbf{x}_1^i)$ for $i = 1, \ldots, N$
    for $t = 2, \ldots, T$ **do**
        Sample $\{a_t^i, \mathbf{x}_t^i\}_{i=1}^N \sim M_{\theta,t}(a_t, \mathbf{x}_t)$
        Calculate $\left\{\tilde{w}_{t-1|T}^i\right\}_{i=1}^N$
        Sample a_t^N with $\mathbb{P}(a_t^N = i) \propto \tilde{w}_{t-1|T}^i$
        Set $\mathbf{x}_t^N = \mathbf{x}_t'$
        Set $\mathbf{x}_{1:t}^i \leftarrow \left(\mathbf{x}_{1:t-1}^{a_t^i}, \mathbf{x}_t^i\right)$ for $i = 1, \ldots, N$
        Calculate $w_t^i = W_{\theta,t}(x_{1:t}^i)$ for $i = 1, \ldots, N$
    end for
    Draw k with $\mathbb{P}(k = 1) \propto w_T^i$
    return $X[p] = \mathbf{x}_{1:T}^{(k)}$
end for

observation likelihood. Then, sequentially for every time point, the whole particle system is moved from $t-1$ to t by means of a move kernel $M_{\theta,t}(a_t, \mathbf{x}_t)$ which describes the resampling and proposal steps. At this point the ancestor sampling operation takes place, the ancestral weights of the reference $\left\{\tilde{w}_{t-1|T}^i\right\}_{i=1}^N$, are given by $\tilde{w}_{t-1|T}^i = p_\theta\left(\mathbf{x}_t^N | \mathbf{x}_{t-1}^i\right)$ the probability of the value of the reference trajectory at time t given all the particles (including the reference) at time $t-1$. The ancestor for this point on the reference trajectory at time t can then be sampled with probability proportional to these ancestral weights. The N^{th} particle at this time step t is then replaced with the reference, and the ancestral paths of the particle system are updated. Finally, the weights of the particles at time t are calculated, $W_{\theta,t}(x_{1:t}^i)$. This pattern continues up to the end of the available time data, i.e. for $t = 1, \ldots, T$. Once the CPF has completed this forward pass, a path can be sampled with probability proportional to the weights at the final time step, $\mathbb{P}(k = 1) \propto w_T^i$; this is then assigned as the p^{th} sample $X[p]$. These P samples then provide a Monte Carlo approximation to the smoothing distribution of the state-space model.

Remembering that the aim is to recover the distribution over the unknown internal states and inputs to the nonlinear system, given the augmentation of the states with the state-space representation of the GP, this smoothing distribution is the object of interest. Therefore, employing this inference on the state-space model described in equations (13) and (14), will allow the input-state estimation task to be carried out.

3 INPUT ESTIMATION OF A BOUC-WEN SYSTEM

In this work, the Bouc-Wen hysteretic system [14] is considered as a typical and interesting benchmark. The model is formed in the same manner as in [15], with the equation of motion given as,

$$m\ddot{y} + c\dot{y} + ky + z(y, \dot{y}) = u(t) \tag{15a}$$

$$\dot{z}(y, \dot{y}) = \alpha\dot{y} - \beta\left(\gamma|\dot{y}||z|^{\nu-1}z + \delta\dot{y}|z|^{\nu}\right) \tag{15b}$$

The parameters of the system are chosen here to be $m = 2$, $c = 100$, $k = 5 \times 10^4$, $\alpha = 5 \times 10^4$, $\beta = 1 \times 10^3$, $\gamma = 0.8$, $\delta = -1.1$, $\nu = 1$. The response of this system is simulated subject to two different loading scenarios. First, the response of the system is simulated when the load is a random sample drawn from a Gaussian process in time, the exact form of the proposed model. Secondly, the response of the system to a sine wave excitation at 30 Hz, an input which is not explicitly drawn from the GP prior over the forcing, but for which the GP may be a suitable prior. It will now be shown that in both of these cases, very good estimates of the internal states of the Bouc-Wen system and estimates of the unmeasured forcing signals can be recovered. In both cases, the measured quantity is the acceleration of the oscillator, which is corrupted by an additive white Gaussian noise with a variance of approximately 2% of the variance of the measured acceleration. For both experiments a Matérn 5/2 kernel is used to model the unknown latent force which adds three additional states to the nonlinear SSM being identified.

3.1 Loaded with GP

Initially, loading the oscillator with a sample of a Gaussian process with a Matérn 5/2 kernel, the measured acceleration response to this load is simulated and shown in Figure 1.

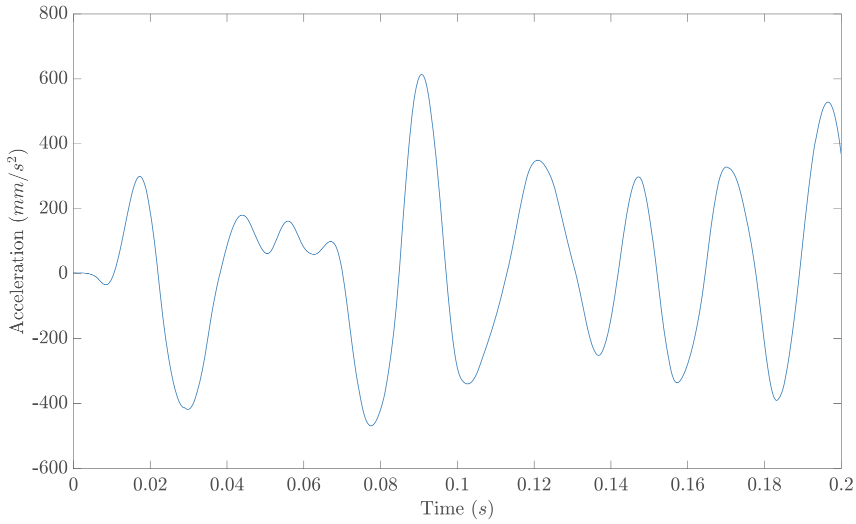

Figure 1: Measured acceleration from the Bouc-Wen oscillator when forced by a random sample from a Gaussian process with a Matérn 5/2 kernel.

This measured acceleration is used as the observed quantity in the model, as described in the previous section. The Bouc-Wen equations are used to describe a nonlinear state-space model of the dynamic system, which is coupled with the state-space representation of a Gaussian process in time with a zero mean and a Matérn 5/2 kernel (the same as used to generate the data). Given this known model form, it is expected that the model should perform well, provided the inference scheme used (PGAS) converges appropriately.

In Figure 2, the estimated states from the proposed inference scheme are shown. On the left hand side of the figure, the MCMC sampled paths are shown; the Markov chain was run for 5000 iterations with the first 1000 discarded as burn-in and the chain thinned by a factor of two. The particle filter was run as a bootstrap filter with 15 particles included; notably, even with this low number of particles the smoothing distribution can be seen to be well approximated. If the samples are used to construct a Gaussian approximation of the smoothing distribution, as shown in the left hand column of the figure, then it can be observed that all of the ground truth

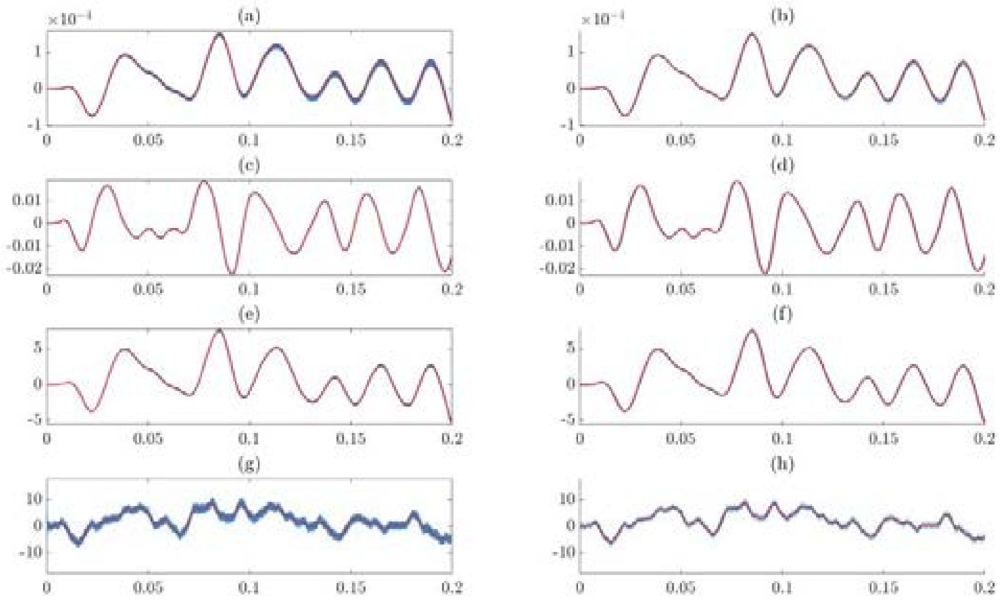

Figure 2: Estimated states inferred via the PGAS sampling of the smoothing distributions. In the left column, the samples obtained from the MCMC scheme and on the right Gaussian approximations to the distributions with one, two and three sigma intervals shaded. The four rows correspond to the first four states of the model: displacement, velocity, $z(y, \dot{y})$ and the latent force. In all plots the ground truth is indicated in red.

states lie within a three-sigma confidence interval.

To consider the quality of the fit in a purely deterministic sense, a normalised mean-squared error metric is used such that,

$$NMSE = \frac{100}{N\sigma^2} \sum_{i=1}^{N} \{(y_i - \hat{y}_i)\}$$

if y is the measured "true" signal of length N with variance σ^2 and the point estimates are given by $\hat{y}$. Intuitively, this leads to values bounded at the lower end by zero which corresponds to an exact fit and, as the quality of the fit degrades, the value increases. For some sense of this quality, a prediction where every point was equal to the sample mean of $\{y\}_{i=1}^{N}$ will give an NMSE of 100. Anecdotally, one can consider a value less than 5 to be a 'very good' fit and below 1 an 'excellent' fit.

For the estimations shown in Figure 2, if the mean of the samples is used as the expectation over the states to provide a point estimate, the NMSE values are calculated. For the displacement (frames (a) and (b)) the NMSE is 0.38; for the velocity (frames (c) and (d)), 0.33; for $z(y, \dot{y})$ (frames (e) and (f)) 0.25; and for the force estimate (frames (g) and (h)) 2.18. These metrics indicate that the method is performing well and confirms what is presented visually in Figure 2. Excellent recovery of the dynamic states of the oscillator is observed, including the additional hidden state related to the hysteresis in the system, and the estimate of the forcing is also very good. Some of the increase in the NMSE when considering the forcing can be attributed to the high-frequency components of this force which are smoothed out when the mean over the samples is used as the point estimate. Since the uncertainty in the forcing is also quantified and shown in the figure, it can be seen that the behaviour of the forcing has been captured remarkably well in this complex system.

3.2 Loaded with Sine Wave

In the previous experiment, it was known that the form of the latent force model exactly matched the true forcing signal being applied to the system, i.e. the system was loaded with a sample of a Gaussian process. A more challenging and realistic test case is one where the loading signal is not a draw from a Gaussian process, but instead some known deterministic function. A sine wave is chosen to be such a function with the forcing defined as $u(t) = 120 \sin(30 \cdot 2\pi \cdot t)$.

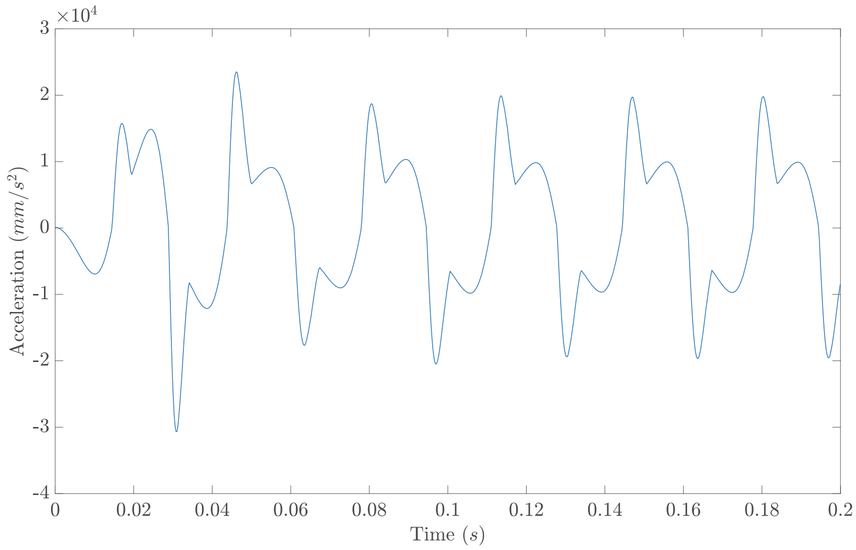

Figure 3: Measured acceleration from the Bouc-Wen oscillator when forced by a 30 Hz sine wave.

The response of the system to this sinusoidal forcing is shown in Figure 3. It is worth noting at this point that, if it were known that a sine wave had been used to load the system, then it is likely far less involved methods would work well to recover that signal. However, the power of the proposed approach is that it uses the GP as a functional prior over the loading function, which removes the need for prior knowledge with respect to the functional form of the forcing which has been applied. This flexibility is particularly important where the system may be subject to complicated and varying loads which are not easily expressed as a known function, e.g. wind or wave loading on offshore structures.

As previously, the nonlinear GP latent force model is used to draw samples from the smoothing distribution of a nonlinear SSM which is the Bouc-Wen model augmented by a GP with a Matérn 5/2 covariance function. 5000 samples of the states are sampled using PGAS, of which 1000 are discarded and 15 particles are used in the bootstrap particle filter. Doing so allows the results in Figure 4 to be constructed in the same way as Figure 2. In the left hand column, the samples obtained from MCMC are shown in blue, with the ground truth shown in red and on the right Gaussian assumed densities created by taking the expectation and variance of the samples at every point in time. The rows again correspond to the states: displacement, velocity, $z(y, \dot{y})$ and force. Reassuringly, the methodology continues to perform well, visually, even when the loading signal does not match the prior functional form of the GP very closely.

As with the previous case it is possible to consider the NMSE between the ground truth and

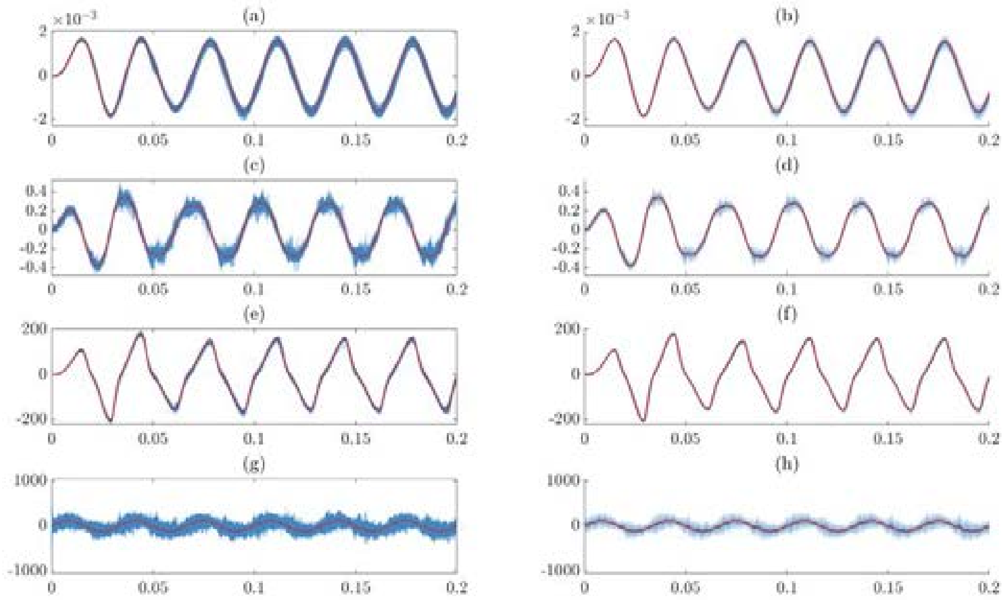

Figure 4: Estimated states inferred via the PGAS sampling of the smoothing distributions. In the left column, the samples obtained from the MCMC scheme and on the right Gaussian approximations to the distributions with one, two and three sigma intervals shaded. The four rows correspond to the first four states of the model: displacement, velocity, $z(y, \dot{y})$ and the latent force. In all plots the ground truth is indicated in red.

the expected values of the smoothing distributions. Considering the mean estimates gives values of 0.48, 0.36, 0.31 and 5.22 for the displacement, velocity, $z(y, \dot{y})$ and force states respectively. These values are larger than those observed in the previous experiment but only marginally so, which reinforces the visual quality of the fit seen in Figure 4. This slight increase in the error of the pointwise predictions is accompanied by an increase in the uncertainties estimated for all states. This uncertainty manifests as increased variance in the samples which have been generated from the Markov chain. These increases are most prominent in the forcing state, which sees a far larger degree of uncertainty, and in the velocity state where significant increases in uncertainty are seen close to the peaks of the response. It may be the case that the strong dependence of $z(y, \dot{y})$ on $\dot{y}$ causes uncertainty to be pushed onto the velocity state when the effect of the nonlinearity on the observed acceleration is greatest.

Considering more closely the estimate of the forcing, the estimates in Figure 4 are also shown in Figure 5 for the forcing state only. It can be seen that relative to the case study in Section 3.1 the uncertainty in the forcing estimate is far higher. However, clearly the main trend of the sine wave has been recovered, and the ground-truth signal lies close to the expected values of the smoothing distribution. In the lower frame of Figure 5, the Gaussian approximation to the smoothing distribution more clearly shows how well the method is performing, despite the slight increase in NMSE for this case. Remembering that this distribution has been approximated by a set of Monte Carlo samples, it can be seen that the main cause of error is high-frequency content in the mean estimate. These errors may well be a consequence of the limited number of samples used to form the estimate, which will be discussed with reference to the computational load of the proposed method. It is the opinion of the authors that the results shown here represent a good recovery of the unknown load on the system.

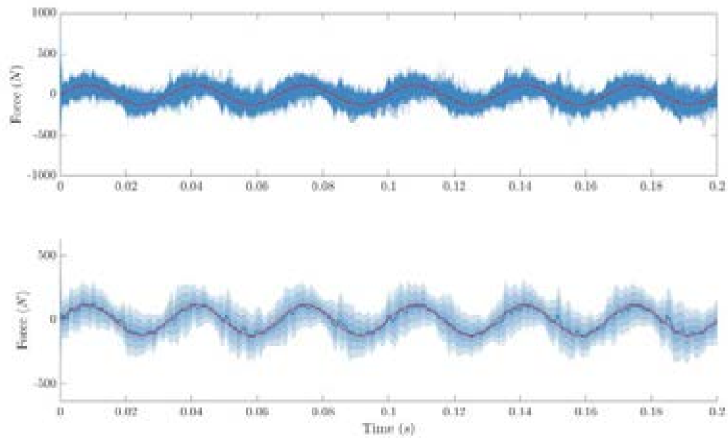

Figure 5: Estimated sine wave loading on the Bouc-Wen oscillator from the nonlinear latent force model. Ground truth shown in red. Above, samples of the possible loading signals. Below, mean estimate in blue and three sigma intervals shaded, of the approximated Gaussian distribution.

The results shown from these two case studies with the Bouc-Wen system provide reassuring evidence that the approach proposed for joint input-state estimation in nonlinear systems has the potential to work equally well when there are hysteretic nonlinearities present in the system. It would appear that the additional states introduced in the model do not cause significant observability issues, in that the estimates of the states and the unknown loads remain very good. However, it should be noted that there is still room for improvement. One shortcoming of the methodology presented in this work is the significant computational burden. This stems from the requirement to run multiple particle filters, which significantly increases the computation time relative to the linear case [8]. To overcome this practical limitation, it may be necessary to consider more efficient inference schemes in the future, or to resort to a different approximation of the nonlinear system, for example, a Gaussian filter which approximates the nonlinearity. This is expected to be of most benefit when the nonlinearity is weak in the system.

4 CONCLUSION

This paper has sought to understand the performance of a nonlinear input-state estimation methodology proposed in [1] when the nonlinear system contains a hysteretic nonlinearity. A Bouc-Wen system was chosen as a typical example of such a nonlinear dynamic model. It was shown that, adopting the nonlinear Gaussian process latent restoring force approach based on Particle Gibbs with Ancestor Sampling, it was possible to perform highly-accurate nonlinear input-state estimation on this challenging dynamical system. Errors in state estimates were seen to be consistently below 1%, including the additional internal state of the model, $z(y, \dot{y})$. The estimates of the forcing were also seen to result in low pointwise error: 2.18% in the case where a GP was used as the loading signal and 5.22% when a sine wave forcing was used. In conjunction with these low error values, reasonable estimation of the uncertainty in the states was also recovered which may be of more value when this type of identification is coupled with further analyses. Finally, it was discussed how the major drawback of the proposed method, i.e. its computational burden, is a promising area for further investigation.

5 ACKNOWLEDGEMENTS

The authors gratefully acknowledge funding from the Engineering and Physical Sciences Research Council (EPSRC), UK, under grants: EP/S001565/1 and EP/R003645/1. Thanks is also extended to Ramboll Energy for their support of J. D. Longbottom.

References

[1] Timothy J. Rogers, Keith Worden, and Elizabeth J. Cross. Bayesian joint input-state estimation for nonlinear systems. *Vibration*, 3(3):281–303, 2020.

[2] Simo Särkkä. *Bayesian Filtering and Smoothing*, volume 3. Cambridge University Press, 2013.

[3] Anthony O'Hagan. Curve fitting and optimal design for prediction. *Journal of the Royal Statistical Society: Series B (Methodological)*, 40(1):1–24, 1978.

[4] Christopher K.I. Williams and Carl Edward Rasmussen. *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006.

[5] Mauricio Alvarez, David Luengo, and Neil D. Lawrence. Latent force models. In *Artificial Intelligence and Statistics*, pages 9–16, 2009.

[6] Jouni Hartikainen and Simo Särkkä. Kalman filtering and smoothing solutions to temporal gaussian process regression models. In *2010 IEEE International Workshop on Machine Learning for Signal Processing*, pages 379–384. IEEE, 2010.

[7] Jouni Hartikainen and Simo Sarkka. Sequential inference for latent force models. *arXiv preprint arXiv:1202.3730*, 2012.

[8] Timothy J. Rogers, Keith Worden, and Elizabeth J. Cross. On the application of Gaussian process latent force models for joint input-state-parameter estimation: With a view to Bayesian operational identification. *Mechanical Systems and Signal Processing*, 140: 106580, 2020.

[9] Arnaud Doucet and Adam M. Johansen. A tutorial on particle filtering and smoothing: Fifteen years later. *Handbook of Nonlinear Filtering*, 12(656-704):3, 2009.

[10] Christian Andersson Naesseth, Fredrik Lindsten, and Thomas B. Schön. Sequential Monte Carlo for graphical models. *Advances in Neural Information Processing Systems*, 27: 1862–1870, 2014.

[11] Christophe Andrieu, Arnaud Doucet, and Roman Holenstein. Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(3):269–342, 2010.

[12] Andreas Svensson, Thomas B. Schön, and Manon Kok. Nonlinear state space smoothing using the conditional particle filter. *IFAC-PapersOnLine*, 48(28):975–980, 2015.

[13] Fredrik Lindsten, Michael I. Jordan, and Thomas B. Schön. Particle Gibbs with ancestor sampling. *The Journal of Machine Learning Research*, 15(1):2145–2184, 2014.

[14] Mohammed Ismail, Fayçal Ikhouane, and José Rodellar. The hysteresis Bouc-Wen model, a survey. *Archives of Computational Methods in Engineering*, 16(2):161–188, 2009.

[15] Maarten Schoukens and Jean Philippe Noël. Three benchmarks addressing open challenges in nonlinear system identification. *IFAC-PapersOnLine*, 50(1):446–451, 2017.