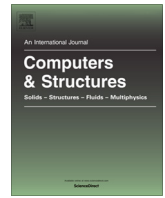




Contents lists available at ScienceDirect

Computers and Structures

journal homepage: www.elsevier.com/locate/compstruc

Probabilistic machine learning based predictive and interpretable digital twin for dynamical systems

Tapas Tripura^{a,*}, Arya Sheetal Desai^a, Sondipon Adhikari^c, Souvik Chakraborty^{a,b,*}

^a Department of Applied Mechanics, Indian Institute of Technology Delhi, Hauz Khas, Delhi, 110016, India

^b Yardi School of Artificial Intelligence (ScAI), Indian Institute of Technology Delhi, Hauz Khas, Delhi, 110016, India

^c James Watt School of Engineering, University of Glasgow, Glasgow G12 8QQ, United Kingdom

ARTICLE INFO

Article history:

Received 13 December 2022

Accepted 26 February 2023

Available online 9 March 2023

Keywords:

Predictive digital twin

Model update

Probabilistic machine learning

Stochastic differential equation

ABSTRACT

A framework for creating and updating digital twins for dynamical systems from a library of physics-based functions is proposed. The sparse Bayesian machine learning is used to update and derive an interpretable expression for the digital twin. Two approaches for updating the digital twin are proposed. The first approach makes use of both the input and output information from a dynamical system, whereas the second approach utilizes output-only observations to update the digital twin. Both methods use a library of candidate functions representing certain physics to infer new perturbation terms in the existing digital twin model. In both cases, the resulting expressions of updated digital twins are identical, and in addition, the epistemic uncertainties are quantified. In the first approach, the regression problem is derived from a state-space model, whereas in the latter case, the output-only information is treated as a stochastic process. The concepts of Itô calculus and Kramers-Moyal expansion are being utilized to derive the regression equation. The performance of the proposed approaches is demonstrated using highly nonlinear dynamical systems such as the crack-degradation problem. Numerical results demonstrated in this paper almost exactly identify the correct perturbation terms along with their associated parameters in the dynamical system. The probabilistic nature of the proposed approach also helps in quantifying the uncertainties associated with updated models. The proposed approaches provide an exact and explainable description of the perturbations in digital twin models, which can be directly used for better cyber-physical integration, long-term future predictions, degradation monitoring, and model-agnostic control.

© 2023 Elsevier Ltd. All rights reserved.

1. Introduction

A digital twin (DT) is described as a virtual counterpart of a physical entity. As a whole, it consists of a physical system denoting the physical space, a virtual model denoting mirror space, and a linking mechanism between them [1]. The origin of DT is mainly attributed to the concept of product life-cycle management [2,3]; however, the first practical definition was provided in [4]. Due to recent growth in the smart manufacturing industry, such as Industry 4.0 and Industrial Internet, the DT has received strategic priorities and is being increasingly explored to improve physical entities'

performance further [5,6]. In practice, the DT can exist in a local computer or the cloud platform [7,8], and recent innovations in computer-aided modeling, cloud computing, 5G networks, and wireless sensors have provided the opportunity to create digital twin models in such platforms [9,10]. A DT attempts continuous time synchronization of the physical and digital twins by means of modern machine learning/Internet-of-Things (IoT) and mechanical actuators [11]. While the machine learning models attempt to update the digital twin using real-time streaming data, the actuator can be used to invoke changes in the physical model. Using machine learning and modern artificial intelligence techniques [12–15], it is possible to seamlessly update the digital model based on real-time streaming data and perform required operations such as health monitoring, the discovery of failure causes, analysis of remaining useful life, regulating the supply chain, and optimizing product performance.

In academic publications, the “digital twin” terminology first appeared in [16]. Then the concept of DT was first put into use

* Corresponding authors at: Department of Applied Mechanics, Indian Institute of Technology Delhi, Hauz Khas, Delhi, 110016, India. (Tapas Tripura); Yardi School of Artificial Intelligence (ScAI), Indian Institute of Technology Delhi, Hauz Khas, Delhi, 110016, India (S. Chakraborty).

E-mail addresses: tapas.t@am.iitd.ac.in (T. Tripura), desai.aaryaaishwarya@gmail.com (A.S. Desai), Sondipon.Adhikari@glasgow.ac.uk (S. Adhikari), souvik@am.iitd.ac.in (S. Chakraborty).

by NASA's Apollo space program, whose aim was to build two identical space vehicles so that the space vehicle could be mirrored and monitored in real-time using the twin in the earth [3,17]. Since then, the presence of DTs in different industries has become more and more [5,18]. Applications of DTs in the engineering field for health monitoring, diagnosis, and prognostics are evident in the literature [19–24]. DT also enables intelligent manufacturing for efficient product planning, higher productivity, and low-cost manufacturing [25–28]. The presence of DT in the automotive and aerospace industry for automation of gearboxes, aircraft structural life prediction, and many more is evident in [29–31]. More details on the state-of-the-art DT research are available in [5,32]. In this work, we focus on predictive digital twins for nonlinear dynamical systems. In the purview of dynamical systems, the development of DT technology is a non-trivial task due to the presence of multiple time scales [33]. An example is the crack growth in the dynamical systems, where the dynamical behavior of the system dynamics is on a faster time scale, and the Paris-Erdogan law determines the propagation of a crack in the system on a slow time scale [34].

In the multi-time scale dynamics, the DT models can be broadly classified into pure physics-based, pure data-based, or a hybrid between them. In the literature [35], a physics-based DT is proposed for monitoring dynamical systems. When the physics is exact, and the real-time data are noise-free, a physics-based DT can outperform its data-driven counterparts. They can adapt and obtain accurate predictions in changing environments. However, the physics is often inexact and simplified, which may not represent the actual system. Another issue is the presence of environmental noise in the sensor measurements. In such cases, solely relying on the physics-based models may not cater to the intended purpose. Purely data-driven DTs provide an alternative solution to eliminating these issues [19]. Although the data-driven DTs can take into account the noise in the data, they fail to generalize in an unseen environment, such as a sharp change in the trajectory path. As a consequence, the synchronization between the physical and virtual space gets hindered, and the desired performance is not achieved. To alleviate these shortcomings, a hybrid alternative between the physics and data-driven DTs is proposed in [33,36]. This hybrid approach utilizes machine learning techniques to learn and compensates for incomplete physics from the real-time streaming data.

From the above discussion, it is clear that a DT is a virtual shadow of its physical counterpart, which is supposed to replicate the physical changes in the system during its entire service life. The DTs discussed above are based on the assumption that the underlying physics of the dynamical system does not evolve with time and evolution of the system can be perfectly captured through changes in its system parameters. This is merely an approximation and not practically true because the changes in the system may also arise due to the evolution of the underlying physics of the physical twin. This is evident in the degrading dynamical system [34], where the physics of the underlying system changes on a slow-time scale due to fatigue. If the physics of the DT is not updated properly, accounting for the effect of change in physics using changes in the parameter may not result in accurate predictions. Over the course of time, the physical and digital twins will be unsynchronized, and the actual purpose of devising the DT will not be achieved. Therefore, a robust DT should be able to track changes in both the system parameters as well as in the underlying physics. In this course, an important aspect of the development of DT is the identification of model-form errors. Such work is carried out in [37], which is again a hybrid framework that utilizes machine learning models to learn the missing physics. As a result, the updated model is difficult to interpret, and generalization is not

ensured. In a nutshell, currently, in the literature on DT, the concept of updating the physics of virtual models using physics-based interpretable functions is greatly missing.

In this paper, we aim to propose a framework for updating the DT model using interpretable physical functions from a predefined library. A predefined library is constructed from the various physical and mathematical expressions, each of which represents the physics of various dynamical characteristics of a dynamical system. Using this library, a regression problem is constructed. From the library, only the key terms representing the perturbations in the underlying physics of physical twins are identified for the purpose of updating the DT. The sparse Bayesian regression is used to identify the correct perturbation terms from the library accurately. Additionally, the epistemic uncertainty arising due to noisy and limited data is also captured using sparse Bayesian regression. Using the proposed approach, we discover any change in underlying physics through new physical terms from the predefined library. Since the perturbation in physics is learned using exact physical terms, the updated DTs are highly interpretable and possess the ability to generalize to unseen environments. This allows for accurate long-term predictions, estimates of remaining useful life, and failure probabilities.

The proposed predictive DT framework consists of two approaches. The first framework utilizes both the input–output sensor measurements, whereas the second framework uses only the output observation data from the physical twin. While the first approach is a simplified case, the second approach is more advanced and general. In the first approach, the library is constructed from the input–output observations, and then a regression problem is constructed using the derivatives of state measurements as labels. In the second approach, the output-only observations are treated as a stochastic process, and the governing physics of perturbed DT is expressed as stochastic differential equations (SDEs). The drift and diffusion components of the SDEs are identified separately using the Itô calculus [38] and Kramers-Moyal formula [39]. The labels of the regression are constructed from SDE generators, whereas the library is constructed for the drift and diffusion terms from the linear and quadratic variations of the output observations, respectively. The performance of the proposed approaches is exemplified using three nonlinear examples. Results show that both approaches provide exact identical expressions for the perturbations in the governing physics.

The remainder of the paper is organized as follows: firstly, in Section 2 the problem is stated where we have a readily available nominal model, and we aim to find the new perturbation terms in the nominal model. In Section 3, we review the concepts of sparse Bayesian regression briefly. In Section 4, two approaches for updating the digital twin from physics-based library functions are illustrated with sufficient mathematical descriptions. In Section 5, for demonstration purposes, three numerical examples are taken. Case studies for different levels of noise are also carried out here. Finally, in Section 6, a brief discussion on the salient features of the proposed framework and possible future extensions is given.

2. Problem statement

The main hypothesis behind the digital twin (DT) is that it begins with a nominal model, and in the future, the nominal model gets perturbed owing to the operational and environmental conditions. The nominal model can be imagined as either a laboratory-scale miniature or a computer model of the physical twin. During the evolution, it is assumed that the nominal model evolves at a considerably slower rate than the evolution of the system responses. This allows for the identification of perturbations in the system parameters as a function of slow time scale [33,40].

To put it into a mathematical framework, let us consider the following D -dimensional second-order partial differential equation,

$$\begin{aligned} & \mathbf{M}(t_s)\partial_{tt}\ddot{\mathbf{X}}(t, t_s) + \mathbf{C}(t_s)\partial_t\dot{\mathbf{X}}(t, t_s) + \mathbf{K}(t_s)\mathbf{X}(t, t_s) + \mathbf{H}(\dot{\mathbf{X}}, \mathbf{X}, t, t_s) \\ & + \mathbf{Q}(\dot{\mathbf{X}}, \mathbf{X}, t, t_s) \\ & = \Sigma\dot{\mathbf{B}}(t, t_s) \end{aligned} \quad (1)$$

where \mathbf{M} , \mathbf{C} and \mathbf{K} represent the $\mathbb{R}^{D \times D}$ mass, damping, and linear stiffness matrices of the system, respectively. The functions $\mathbf{H}(\dot{\mathbf{X}}, \mathbf{X}, t, t_s) : \mathbb{R}^D \mapsto \mathbb{R}^D$ and $\mathbf{Q}(\dot{\mathbf{X}}, \mathbf{X}, t, t_s) : \mathbb{R}^D \mapsto \mathbb{R}^D$, denote the linear and nonlinear perturbations in the system, respectively. The term $\mathbf{B}(t, t_s) \in \mathbb{R}^D$ on the right-hand side represents the white noise (the generalized derivative of Brownian motion) with noise intensity matrix $\Sigma \in \mathbb{R}^{D \times D}$. In the above equation, two-time scales, t and t_s are used, which represent the intrinsic time and the service time, respectively. The service time refers to the periods over which the underlying structure or a component is expected to be inspected. The time scale t_s is comparatively much slower than t and since $\mathbf{X}(t, t_s)$ is a function of both the time scale, therefore, the Eq. (1) is written in terms of the partial derivatives. It can be understood that the evolution in $\mathbf{M}(t_s)$, $\mathbf{C}(t_s)$ and $\mathbf{K}(t_s)$ occurs very slowly with respect to time scale t_s . The forcing term $\Sigma\dot{\mathbf{B}}(t, t_s)$ however can change with respect to both the times scales t and t_s . We call Eq. (1) as the model of the proposed digital twin. Since it is already mentioned that the system evolves with respect to the slower time scale, we rephrase the Eq. (1) when $t_s = 0$ as,

$$\mathbf{M}_0\ddot{\mathbf{X}}(t) + \mathbf{C}_0\dot{\mathbf{X}}(t) + \mathbf{K}_0\mathbf{X}(t) = \Sigma\dot{\mathbf{B}}(t). \quad (2)$$

The above equation denotes the beginning of the service life of the underlying system and is often called the nominal model in DT. Here, \mathbf{M}_0 , \mathbf{C}_0 and \mathbf{K}_0 are the parameters of the nominal model. Further, we assumed that as the time scale t_s shifts from the initial condition, the nominal system gets perturbed by new terms, expressed using the functions $\mathbf{H}(\dot{\mathbf{X}}, \mathbf{X}, t, t_s)$ and $\mathbf{Q}(\dot{\mathbf{X}}, \mathbf{X}, t, t_s)$ as,

$$\begin{aligned} & \mathbf{M}_0\ddot{\mathbf{X}}(t) + \mathbf{C}_0\dot{\mathbf{X}}(t) + \mathbf{K}_0\mathbf{X}(t, t_s) + \mathbf{H}(\dot{\mathbf{X}}, \mathbf{X}, t) + \mathbf{Q}(\dot{\mathbf{X}}, \mathbf{X}, t) \\ & = \Sigma\dot{\mathbf{B}}(t) \end{aligned} \quad (3)$$

It is straightforward to note that any changes in the physical model can be incorporated into the DT using the linear and nonlinear functions $\mathbf{H}(\cdot)$ and $\mathbf{Q}(\cdot)$. Therefore, in order to use the DT in practice, we need to characterize the functions $\mathbf{H}(\cdot)$ and $\mathbf{Q}(\cdot)$. In this work, it is assumed that a linking mechanism between the twins is established by using sensors and actuators. The sensors provide measurements of the system states and the force, whenever available, at the time instant t_s . At each time instant t_s , the measurements are obtained for $t_s + 1s$, which means that we have access to only one second of noisy data. We aim to discover the functions $\mathbf{H}(\cdot)$ and $\mathbf{Q}(\cdot)$ from these limited and noisy measurements. Once discovered, they are used to update the nominal model in Eq. (2). In the discovery of $\mathbf{H}(\cdot)$ and $\mathbf{Q}(\cdot)$, we aim to learn them in their interpretable forms instead of the black-box type surrogate models [36]. In order to assess the performance in an unseen scenario, we also aim to learn the uncertainties associated with the parameters of the functions $\mathbf{H}(\cdot)$ and $\mathbf{Q}(\cdot)$. For these, the sparse Bayesian inference is employed. The resulting framework thus is white in nature, and since the functions are learned in a probabilistic framework, the chances of overfitting are very low. Further, the physics of the underlying perturbations is learned using actual mathematical functions. Therefore it is conjectured that the proposed DT will be able to track the evolution of the physical twin accurately. In the coming chapters, we provide a brief introduction to sparse Bayesian regression and the proposed DT framework.

3. Background on sparse Bayesian regression

Let us consider that we have a set of the noisy measurements of system states $\mathbf{X} \in \mathbb{R}^{N \times m}$, where N is the number of measured points and m is the dimension of system states. Also, assume that we have the output of the system $\mathbf{Y} \in \mathbb{R}^{N \times 1}$. Then without loss of generality, any dynamical system can be expressed in the following form:

$$\mathbf{Y} = f(\mathbf{X}, \theta) + \epsilon, \quad (4)$$

where $f(\mathbf{X}, \theta)$ is the system that describes the governing physics of the system and $\epsilon \in \mathbb{R}^{N \times 1}$ is the measurement error. The system $f(\mathbf{X}, \theta)$ is a function of the inputs \mathbf{X} and a set of model parameters θ . Here we assume that the dependency of $f(\mathbf{X}, \theta)$ on \mathbf{X} and θ can be modeled as a linear superposition over a set of basis functions described by,

$$f(\mathbf{X}, \theta) = \mathbf{L}\theta \quad (5)$$

where $\mathbf{L} = [\ell_1, \ell_2, \dots, \ell_K]$ is the $\mathbb{R}^{N \times K}$ library with columns representing the candidate functions with respect to the state measurements \mathbf{X} and $\theta \in \mathbb{R}^{K \times 1}$ is the weight vector representing the system parameters. Instead of the deterministic values, we intend to identify the distribution of parameters. For estimating the distribution of the weight vector θ , we apply the Bayes formula in Eq. (4) yielding,

$$P(\theta|\mathbf{Y}) = \frac{P(\theta)P(\mathbf{Y}|\theta)}{P(\mathbf{Y})}. \quad (6)$$

where $P(\theta|\mathbf{Y})$ is the posterior distribution of θ , $P(\mathbf{Y}|\theta)$ is the likelihood function, $P(\mathbf{Y})$ is the normalizing factor and $P(\theta)$ is the prior distribution representing the prior knowledge about the model. The mismatch error ϵ are modeled as independent and identically distributed (i.i.d.) zero mean Gaussian variable with variance σ^2 . With this information, the conditional probability of \mathbf{Y} given the system parameters θ and noise variance σ^2 can be written as,

$$P(\mathbf{Y}|\theta) = \mathcal{N}(\mathbf{L}\theta, \sigma^2\mathbf{I}_{N \times N}), \quad (7)$$

where $\mathbf{I}_{N \times N}$ denotes the $N \times N$ identity matrix. In order to discover a parsimonious solution for the governing physics, i.e., to allow most of the components in the library to be removed, we introduce sparsity in the weight vector θ using the spike and slab (SS) prior [41–43]. The SS-prior has high shrinkage properties due to its sharp spike at zero and a diffused density spanned over a large range of possible parameter values. Although there are many variants of SS-prior available, we particularly model the spike using a Dirac delta function and the slab using Normal distribution. The Dirac-delta spike concentrates most of the probability mass at zero, thus allowing most of the samples to take a value of zero, and the diffused tail distributes a small amount of probability mass over a large range of possible values allowing only very few samples with very high probability to escape the shrinkage. More details on the selected SS-prior are available in Ref. [44].

In order to allow automated classification of weights into the spike and slab components of the SS-prior, we further introduce a latent indicator variable $\Psi = [\psi_1, \dots, \psi_K]$ for each of the component θ_k in θ . The latent indicator variables ψ_k behave as a Boolean function which takes a value of 1 if the weight corresponds to the slab component; otherwise, takes a value of 0. Since the latent variables ψ_k either takes a value 0 or 1, the automation is done by assigning the latent vector the Bernoulli prior with common hyperparameter p_0 as $p(\psi_k|p_0) = \text{Bern}(p_0)$ for $k = 1 \dots K$. The hyperparameter $p(p_0) = \text{Beta}(\alpha_p, \beta_p)$ is simulated from the Beta prior with the hyperparameters α_p and β_p . Since the effect of the components of the weight vector that belongs to the spike does not contribute

to the selection of the key functions, we construct a reduced weight vector $\theta_r \in \mathbb{R}^r : \{r \ll K\}$, composed from the elements of the weight vector θ for which $\psi_k = 1$. Denoting θ_r as the weight vector containing only those variables from θ for which $\psi_k = 1$, the SS-prior is defined as [41,44],

$$p(\theta|\Psi) = p_{slab}(\theta_r) \prod_{k, \psi_k=0} p_{spike}(\theta_k), \quad (8)$$

where the distributions are defined as, $p_{spike}(\theta_k) = \delta_0$ and $p_{slab}(\theta_r) = \mathcal{N}(\mathbf{0}, \sigma^2 \vartheta_s \mathbf{I}_{r \times r})$. The noise variance $p(\sigma^2) = IG(\alpha_\sigma, \beta_\sigma)$ is simulated from the Inverse-gamma distribution with the hyperparameters α_σ and β_σ , and the slab variance $p(\vartheta_s) = IG(\alpha_\vartheta, \beta_\vartheta)$ is assigned the Inverse-gamma prior with the hyperparameters α_ϑ and β_ϑ . The complete hierarchical prior distribution is summarized in Fig. 1. In the model the hyperparameters $\alpha_\vartheta, \beta_\vartheta, \alpha_p, \beta_p, \alpha_\sigma,$ and β_σ are provided as a deterministic constants.

Theoretically the random hyperparameters $\Psi, \vartheta_s, \sigma^2$ and p_0 should be inferred from the posterior distribution $p(\theta, \Psi, \vartheta_s, \sigma^2, p_0|\mathbf{Y})$, given as,

$$p(\theta, \Psi, \vartheta_s, \sigma^2, p_0|\mathbf{Y}) = \frac{p(\mathbf{Y}|\theta, \sigma^2)p(\theta|\Psi, \vartheta_s, \sigma^2)p(\Psi|p_0)p(\vartheta_s)p(\sigma^2)p(p_0)}{p(\mathbf{Y})} \quad (9a)$$

$$\propto p(\mathbf{Y}|\theta, \sigma^2)p(\theta|\Psi, \vartheta_s, \sigma^2)p(\Psi|p_0)p(\vartheta_s)p(\sigma^2)p(p_0), \quad (9b)$$

where $p(\theta, \Psi, \vartheta_s, \sigma^2, p_0|\mathbf{Y})$ denotes the joint distribution of the random variables, $p(\mathbf{Y}|\theta, \sigma^2)$ denotes the likelihood function, $p(\theta|\Psi, \vartheta_s, \sigma^2)$ is the prior distribution for the weight vector θ , $p(\Psi|p_0)$ is the prior distribution for the latent vector Ψ , $p(\vartheta_s)$ is the prior distribution for the slab variance ϑ_s , $p(\sigma^2)$ is the prior distribution for the noise variance, $p(p_0)$ is the prior distribution for the success probability p_0 and $p(\mathbf{Y})$ is the marginal likelihood. Due to the definition of the SS-prior direct sampling from the joint distribution in (9) is intractable. Thus, we adopt an MCMC using the Gibbs sampling technique to draw the random samples from the joint distribution [45]. The Gibbs sampler requires the conditional distributions of the random variables, which are derived in Ref. [44]. The pseudo-code to obtain the sequence of the random variables $\theta^{(i)}, \sigma^{2(i)}, \vartheta_s^{(i)}, p_0^{(i)}, \Psi^{(i)}$ in the i^{th} -iteration is given in Algorithm 1.

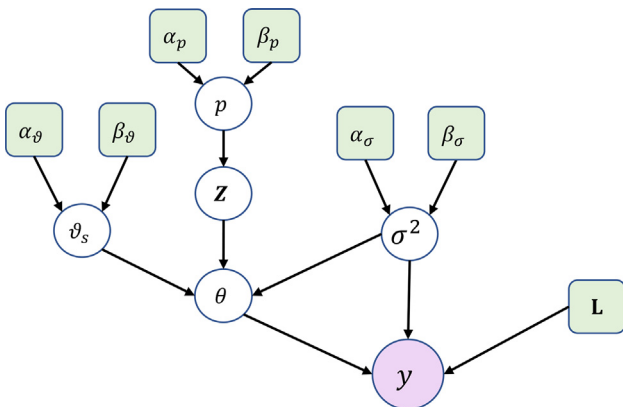


Fig. 1. Hierarchical Bayesian network of the spike and slab distribution. The green square boxes indicate the deterministic parameters, and the white circles represent the random variables. The parameters $\alpha_\vartheta, \beta_\vartheta, \alpha_p, \beta_p, \alpha_\sigma,$ and β_σ indicates the hyperparameters of the hierarchical prior distribution.

Algorithm 1. Pseudo code of the sparse Bayesian regression

Input: State measurements: $\mathbf{X}(t) \in \mathbb{R}^{N \times m}$ and hyperparameters: $\alpha_p, \beta_p, \alpha_\sigma, \beta_\sigma, \alpha_\vartheta, \beta_\vartheta, p_0^{(0)}, \vartheta_s^{(0)}$
 1: Obtain the library \mathbf{L} using the candidate basis functions.
 2: Estimate the initial variance of noise: $\sigma^{2, (0)} = \text{Var}(\mathbf{L}\theta - \mathbf{Y})$.
 3: Estimate the initial latent vector $\Psi^{(0)} = [\psi_1^{(0)}, \psi_2^{(0)}, \dots, \psi_K^{(0)}]$ subjected to $\arg \min_{\theta} \text{MSE}(\mathbf{L}\theta - \mathbf{Y})$.

4: Estimate

$$\mu_\theta^{(i)} = \sum_{\theta}^{(i)} \mathbf{L}_r^{(i)T} \mathbf{Y}, \Sigma_\theta^{(i)} = \sigma^{2(i)} \left(\mathbf{L}_r^{(i)T} \mathbf{L}_r^{(i)} + \vartheta_s^{(i-1)} \mathbf{R}_{0,r}^{(i-1)} \right)^{-1}, \text{ and}$$

find the initial weight vector $\theta_r^{(0)}$ from the Gaussian distribution with mean μ_θ and variance Σ_θ as,

$$p(\theta_r^{(i)}|\mathbf{Y}, \vartheta_s^{(i)}, \sigma^{2(i)}) = \mathcal{N}(\mu_\theta^{(i)}, \Sigma_\theta^{(i)}).$$

5: **for** $i = 1, \dots, \text{MCMC do}$

6: Estimate $u_k = \frac{p_0}{p_0 + \lambda(1-p_0)}$ and

$$\lambda = \frac{p(\mathbf{Y}|\psi_k^{(i)}=0, \Psi_k^{(i)}, \vartheta_s^{(i)})}{p(\mathbf{Y}|\psi_k^{(i)}=1, \Psi_k^{(i)}, \vartheta_s^{(i)})}. \quad \triangleright \text{Eq. (A.2)}$$

7: Then update the latent variable vector $\Psi^{(i+1)}$ from the Bernoulli distribution as,

$$p(\psi_k^{(i+1)}|\mathbf{Y}, \vartheta_s^{(i)}, p_0^{(i)}) = \text{Bern}(u_k).$$

8: Update the noise variance $\sigma^{2(i+1)}$ from the Inverse-gamma distribution as,

$$p(\sigma^{2(i+1)}|\mathbf{Y}, \Psi^{(i+1)}, \vartheta_s^{(i)}) = IG(\alpha_\sigma + 0.5N, \beta_\sigma + 0.5(\mathbf{Y}^T \mathbf{Y} - \mu_\theta^{(i)T} \Sigma_\theta^{(i-1)} \mu_\theta^{(i)})).$$

9: Update the slab variance $\vartheta_s^{(i+1)}$ from Inverse-gamma distribution as,

$$p(\vartheta_s^{(i+1)}|\theta^{(i)}, \Psi^{(i+1)}, \sigma^{2(i+1)}) = IG(\alpha_\vartheta + 0.5h_z, \beta_\vartheta + \frac{1}{2\sigma^2} \theta_r^{(i)T} \mathbf{R}_{0,r}^{(i-1)} \theta_r^{(i)}).$$

10: Estimate $h_z = \sum_{k=1}^K \psi_k^{(i+1)}$ and update the success rate $p_0^{(i)}$ from the Beta distribution as,

$$p(p_0^{(i+1)}|\Psi^{(i+1)}) = \text{Beta}(\alpha_p + h_z, \beta_p + K - h_z).$$

11: Update the weight vector $\theta_r^{(i+1)}$ from step 7.

12: Repeat steps 7–11

13: Discard the burn-in MCMC samples and calculate the marginal PIP values $p(\psi_k = 1|\mathbf{Y})$. $\triangleright \text{Eq. (10)}$

14: Select the basis function in the final model with desired PIP values.

Output: The mean μ_θ and covariance Σ_θ .

In the MCMC initial 500 samples are discarded as the burn-in samples for obtaining the posterior distributions of the parameters θ . In order to select the basis functions in the final model, the marginal posterior inclusion probabilities are calculated by taking the mean over the MCMC samples of latent variable vector Ψ . If n_{MC} denote the number of MCMC required to achieve the stationary distribution after the burn-in samples are discarded then for the k^{th} latent vector ψ_k the PIP is defined as [44],

$$p(\psi_k = 1|\mathbf{Y}) \approx \frac{1}{n_{MC}} \sum_{j=1}^{n_{MC}} \psi_k^j, \quad k = 1, \dots, K. \quad (10)$$

In this work, the basis functions that are observed more than fifty percent of the time in the MCMC simulations, i.e., whose corresponding PIP values are more than 0.5, are included in the final model. Once the final basis functions are selected after neglecting those with $PIP < 0.5$, the mean and covariance of the weight vector θ can be estimated using the empirical formulas. While the mean gives the expected value, the standard deviation provides the confidence interval of the predictions performed using the identified parameters in an unseen environment. The mean μ_θ and covariance Σ_θ will have non-zero entry at only those places where the $PIP > 0.5$. The future predictions using the obtained model can be performed using the following formulas,

$$\mu_{yp} = \mathbf{L}^p \mu_\theta \quad (11a)$$

$$\Sigma_{yp} = \mathbf{L}^p \Sigma_\theta \mathbf{L}^{pT} + \mu_{\sigma^2} \mathbf{I}_{N^p \times N^p}, \quad (11b)$$

where $\mathbf{L}^p \in \mathbb{R}^{N^p \times K}$ is the test dictionary obtained from the newly obtained unseen measurements, $\mu_{yp} \in \mathbb{R}^{N^p \times 1}$ is the predicted mean, $\Sigma_{yp} \in \mathbb{R}^{N^p \times N^p}$ is the predicted covariance, and $\mu_{\sigma^2} \in \mathbb{R}$ is the mean of the measurement noise variance.

4. Interpretable and predictive digital twin for model updating of dynamical systems

Digital Twins (DT) are one of the key components in the Mirror spaced model that was first tossed in the Ref. [2]. The mirror-spaced models consist of three primary components, namely the physical model, the nominal mirror model, and the linking mechanisms that connect the virtual model with its physical twin [1,46]. The idea behind creating the digital twin is to mimic the physical model using streaming data obtained from the physical model in real-time [47]. A few of the main applications of DT include remote access, real-time monitoring, future prediction of remaining useful life, predictive maintenance, prolonging the life-

cycle, and designing [17,48,49]. The available DT models are mostly grey in nature, whereas the physics-based nominal models are coupled with machine learning-based surrogate models. Although the machine learning models like Gaussian process [33,36,40] and neural networks [50,51] can learn almost everything, their ability to generalize beyond the training data in the long run is poor. On the other hand, learning the perturbations in the physical model using interpretable mathematical functions greatly enhances the predictive capability of the nominal model in case of unseen environmental conditions.

In this section, we illustrate the proposed predictive digital twin framework for dynamical systems. The schematic architecture of the framework is provided in Fig. 2. The network architecture has three primary components - (a) the physical model, (b) the digital twin, and (c) the linking mechanism. The linking mechanism further consists of three independent modules - (i) the data assimilation and processing module, (ii) the model updating module, and (iii) the prediction module. The data processing is performed by using a physics-based nominal model, and in the updating module, the sparse Bayesian regression is employed. Since the perturbations in the physical model are obtained in terms of interpretable functions, we consider the digital twin as white in nature. Although the proposed framework should theoretically work for higher-order dynamical systems, in this work, we assume only the second-order dynamical systems. Furthermore, we consider that the second-order dynamical systems can be completely represented using displacement and velocity measurements.

Due to the advances in the development of sensor technologies, it is possible to measure the displacement and velocity time histories of a dynamical system. However, often the measurement of input forces is not feasible. Towards this, we propose two frameworks - (i) when both the input-output measurements are available and (ii) when only the state measurements are measurable. In framework-1, we simply remove the information of the nominal model using the measured state measurements and then perform sparse regression to identify the perturbation terms. In framework-2, similar to the previous, we first remove the information of the nominal model and then employ the sparse Bayesian

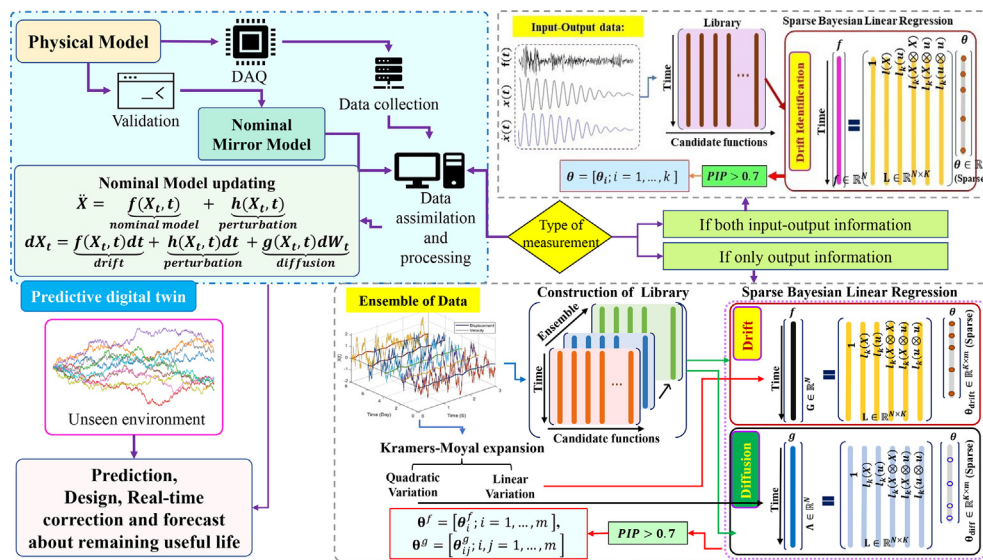


Fig. 2. Schematic architecture of the proposed predictive digital twin framework for model updating of dynamical systems. The network primarily consists of three components, namely the physical model, the mirror model, and the linking mechanism. The linking mechanism performs three simultaneous operations that are data assimilation and processing, the updating of the nominal mirror model, and making predictions in the presence of unseen environmental agents using the updated digital twin. For updating the virtual mirror model using explainable functions, the data assimilation and processing unit utilizes sparse Bayesian regression. The Bayesian regression makes use of two independent frameworks which are well equipped to obtain the parsimonious solution when both or either of the input and output measurements is available. The input refers to the source, and the output refers to the state measurements.

regression in the purview of the Kramers-Moyal expansion [39] to identify the perturbations in terms of stochastic differential equations [38,52].

4.1. Model updating using input–output measurement

The premise of identification of the perturbations from noisy input–output measurements and updating the original model is that the measurements can be expressed as a linear superposition of certain basis functions. The basis functions can be identity function, polynomial, trigonometric, exponential, logarithmic, signum, modulus, or combinations between them. These bases are evaluated on the input–output measurements, and a library, often called a design matrix is formed. However, due to the inclusion of a large number of candidate functions in the library, most of the candidate model components are likely to be incorrect. Further, there will always be some level of confusion as few of the basis functions will have high correlations. As a whole, the 'true' model components will not be identified, leading to the model discrepancy and bias in the identified system parameters. In order to allow the model components that do not provide a significant contribution in representing the data to be removed, the sparse Bayesian regression introduced in Section 3 is used. Further, the Bayesian nature of the model updating framework helps in removing the terms inside the library in a probabilistic manner, thereby requiring less human intervention.

Before moving into the mathematical description, we note that the higher-order dynamical systems are commonly described using a projected space, for example, the second-order systems are often expressed in terms of their displacement and velocity states. The benefit of the projected space is that all the states in this space are directly observable. Towards this, let us represent the projection by a map $T: \mathbb{R}^d \rightarrow \mathbb{R}^m$ where d and m are the dimension of the original and projected space. In the projected space, let us assume that there exists a dynamical system of the following form,

$$\dot{\mathbf{X}}_t = \mathbf{f}(\mathbf{X}_t, t); \quad \mathbf{X}(t = t_0) = \mathbf{X}_0; \quad t \in [0, T], \quad (12)$$

where $\mathbf{X}_t \in \mathbb{R}^m$ denotes the system states and $\mathbf{f}(\mathbf{X}_t, t): \mathbb{R}^m \mapsto \mathbb{R}^m$ represents the dynamics of the underlying system. Since we assume that due to operational and environmental conditions, the underlying nominal gets perturbed, we rephrase the above equation as,

$$\dot{\mathbf{X}}_t = \underbrace{\mathbf{f}(\mathbf{X}_t, t)}_{\text{Nominal model}} + \underbrace{\mathbf{h}(\mathbf{X}_t, t)}_{\text{Perturbation}}; \quad \mathbf{X}(t = t_0) = \mathbf{X}_0; \quad t \in [0, T], \quad (13)$$

where $\mathbf{h}(\mathbf{X}_t, t): \mathbb{R}^m \mapsto \mathbb{R}^m$ represents the perturbations terms. In the proposed digital twin framework, the nominal model $\mathbf{f}(\mathbf{X}_t, t)$ is known to us a priori and we aim correct the nominal model by learning the perturbation term $\mathbf{h}(\mathbf{X}_t, t)$ from freshly obtained input–output measurements. In order to only discover the perturbation terms, we first remove the information about the nominal model from the measured output as,

$$\begin{aligned} \dot{\mathbf{Z}}_t &= \dot{\mathbf{X}}_t - \mathbf{f}(\mathbf{X}_t, t) \\ &= \mathbf{h}(\mathbf{Z}_t, t) \end{aligned} \quad (14)$$

Let $\{\ell_k(\mathbf{Z}_t); k = 1, \dots, K\}$ be the set of candidate library functions, $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_K]^T$ be the associated parameters and K is the dimension of the library. In order to discover $\mathbf{f}(\mathbf{Z}_t, t)$ in terms of the analytical functions, we express $\mathbf{f}(\mathbf{Z}_t, t)$ as linear combination of the basis functions as,

$$\mathbf{f}_i(\mathbf{Z}_t, t) = \theta_{i_1} \ell_1(\mathbf{Z}_t) + \dots + \theta_{i_k} \ell_k(\mathbf{Z}_t) + \dots + \theta_{i_K} \ell_K(\mathbf{Z}_t). \quad (15)$$

In the above equation i represents the i^{th} state of m -dimensional state-space and θ_{i_k} denote the k^{th} basis function of i^{th} state. In the regression format, the above equation is expressed as,

$$\mathbf{Y}_i = \mathbf{L}\boldsymbol{\theta}_i + \boldsymbol{\epsilon}_i, \quad (16)$$

where $\mathbf{Y}_i = \dot{\mathbf{Z}}_i$ and $\mathbf{L} \in \mathbb{R}^{N \times K} := [\ell_1(\mathbf{Z}_t), \dots, \ell_K(\mathbf{Z}_t)]$ is the library. For constructing the target and library in the above equation, both the states \mathbf{Z}_t and $\dot{\mathbf{Z}}_t$ can be measured using the sensors. However, in case of restrictions one can choose to measure only \mathbf{Z}_t and obtain $\dot{\mathbf{Z}}_t$ from \mathbf{Z}_t using higher order numerical differentiation schemes. Once constructed, it can be understood that the above equation is identical to Eq. (4) and can be solved using the sparse Algorithm 1 in Section 3.

4.2. Model updating using output only measurements

In the previous section, we demonstrated the model updating framework using both the input–output information. However, often an accurate measurement of inputs is not feasible. In such cases, the library of candidate functions becomes ill-conditioned, and this leads to the selection of the wrong basis functions. Since the input information is assumed to be unavailable, we try to represent the underlying governing physics in terms of the stochastic differential equation (SDEs) [38,52]. To represent the systems in terms of SDEs we treat the output measurements as a stochastic process and perform sparse Bayesian learning in the purview of stochastic calculus. We again use the state-space to represent the higher-order dynamical systems in terms of the SDEs. Let the state-space be realized by a map $T: \mathbb{R}^d \rightarrow \mathbb{R}^m$ that maps the d -dimensional system to the m -dimensional SDEs with $d < m$. Then using T , any perturbed higher-order system can be reduced to the following SDEs:

$$\dot{\mathbf{X}} = \underbrace{\mathbf{f}(\mathbf{X}_t, t)}_{\text{Nominal model}} + \underbrace{\mathbf{h}(\mathbf{X}_t, t)}_{\text{Perturbation}} + \underbrace{\mathbf{g}(\mathbf{X}_t, t)\xi(t)}_{\text{Diffusion}} \quad (17)$$

where $\mathbf{f}(\mathbf{X}_t, t): \mathbb{R}^m \mapsto \mathbb{R}^m$ represents the dynamics of the nominal model, $\mathbf{h}(\mathbf{X}_t, t): \mathbb{R}^m \mapsto \mathbb{R}^m$ represents the nature of the perturbation, $\mathbf{g}(\mathbf{X}_t, t): \mathbb{R}^m \mapsto \mathbb{R}^{m \times n}$ represents the volatility associated with the dynamics, and $\xi(t)$ represents the stochastic input often represented as white noise [38]. To apply mathematical operations over the noise $\xi(t)$ it is often appropriate to represent the above equation through Itô SDEs which arises naturally in non-linear dynamical systems subjected to stochastic excitation such as earthquake, wind force, wave force, etc. [38]. Let (Ω, \mathcal{F}, P) be the complete probability space with $\{\mathcal{F}_t; 0 \leq t \leq T\}$ be the natural filtration constructed from sub σ -algebras of the filtration \mathcal{F} . Further note that the white noises are generalized derivative of Brownian motions, i.e. $\xi(t) = \dot{\mathbf{B}}(t)$. Therefore, under (Ω, \mathcal{F}, P) an m -dimensional n -factor SDE driven by n -dimensional Brownian motion $\{\mathbf{B}_j(t); j = 1, \dots, n\}$ can be written as,

$$d\mathbf{X}_t = (\mathbf{f}(\mathbf{X}_t, t) + \mathbf{h}(\mathbf{X}_t, t))dt + \mathbf{g}(\mathbf{X}_t, t)d\mathbf{B}(t); \quad \mathbf{X}(t = t_0) = \mathbf{X}_0; \quad t \in [0, T]. \quad (18)$$

Here $\mathbf{X}_t \in \mathbb{R}^m$ denotes the \mathcal{F}_t -measurable state measurements. In the Itô calculus, the first term within the bracket on the right-hand side is called the drift vector, the second term is termed as diffusion matrix, and $\mathbf{B}_j(t) \in \mathbb{R}^n$ is known as Brownian motion. In addition to classical methods, [52], for obtaining the solution of Eq. (18) many modern stochastic integration schemes are previously proposed by the authors [53,54]. In the digital twin framework, the nominal model $\mathbf{f}(\mathbf{X}_t, t)$ is known to us a priori with the help of which we aim to learn the perturbation term $\mathbf{h}(\mathbf{X}_t, t)$ using the freshly obtained measurements. For this, we first remove the information about the nominal model from the measured signal using the following operations,

$$\begin{aligned} d\mathbf{Z}_t &= (\mathbf{f}(\mathbf{X}_t, t) + \mathbf{h}(\mathbf{X}_t, t))dt + \mathbf{g}(\mathbf{X}_t, t)d\mathbf{B}(t) - \mathbf{f}(\mathbf{X}_t, t)dt \\ &= \mathbf{h}(\mathbf{Z}_t, t)dt + \mathbf{g}(\mathbf{Z}_t, t)d\mathbf{B}(t) \end{aligned} \quad (19)$$

The above SDE is now the function of the nominal model removed state measurements and contained the information of - (i) pertur-

bation in the drift and (ii) diffusion. At this point, it is straightforward to understand that the discovery of governing physics in terms of Eq. (18) requires the independent identifications of the drift and diffusion components. In contrast to the diffusion term, the deterministic drift functions behave as smooth functions, i.e., they are assumed to be at least twice differential. Thus there exists a finite variation of drifts. On the contrary, the stochastic Brownian motions are not differentiable everywhere with respect to the process $\mathbf{Z}(t)$. Due to such non-differentiability property, the Brownian motions are assumed to have only the quadratic variation. As a consequence, they are defined in the mean square sense only.

Mathematically, let us consider the interval $s \in [0, t]$ that is partitioned into n parts. If Z_t denotes arbitrary random process then according to the Itô calculus, as $n \rightarrow \infty$ the finite variation $\{V_n(\mathbf{Z}, t) : \sum_i^n |Z(s_i) - Z(s_{i-1})|\} \rightarrow V(\mathbf{Z}, t)$ and the quadratic variation $\{Q_n(\mathbf{Z}, t) : \sum_i^n |Z(s_i) - Z(s_{i-1})|^2\} \rightarrow Q(\mathbf{Z}, t)$ [38]. This suggests that if the sampling rate is sufficiently small, then the drift and diffusion components of an SDE in Eq. (18) can be learned using only the state measurements in terms of their linear and quadratic variations, respectively [45]. However, it should be noted that the diffusion components - (i) have zero finite variations and (ii) are bounded by the quadratic variations. Thus, the diffusion components are recoverable only through their covariation terms. Therefore, we express the drift and diffusion components of the SDE in Eq. (18) in terms of the state measurements as follows:

$$\mathbf{h}_i(\mathbf{Z}_t, t) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} E[Z_i(t + \Delta t) - Z_i(t)] \quad \forall k = 1, 2, \dots, N, \quad (20a)$$

$$\Gamma_{ij}(\mathbf{Z}_t, t) = \frac{1}{2} \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} E[|Z_i(t + \Delta t) - Z_i(t)| |Z_j(t + \Delta t) - Z_j(t)|] \quad \forall k = 1, 2, \dots, N, \quad (20b)$$

where $\mathbf{h}_i(\mathbf{Z}_t, t)$ is the i^{th} drift component and Γ_{ij} is the $(ij)^{\text{th}}$ component of the diffusion covariance matrix $\Gamma \in \mathbb{R}^{n \times n} := (\mathbf{g}\mathbf{g}^T)(\mathbf{Z}_t, t)$. In order to derive the analytical form of the drift and diffusion components from state measurements, we represent the drift and diffusion as a linear superposition of candidate basis functions.

Let $\{\ell_k(\mathbf{Z}_t); k = 1, \dots, K\}$ be the set of candidate library functions where $\ell_k(\mathbf{Z}_t)$ represents the various linear and non-linear mathematical functions defined with respect to the system states. We first construct the libraries $\mathbf{L}^f \in \mathbb{R}^{N \times K}$ and $\mathbf{L}^g \in \mathbb{R}^{N \times K}$ from the subsets $\{\ell_k^f(\mathbf{Z}_t)\} \subseteq \{\ell_k(\mathbf{Z}_t)\}$ and $\{\ell_k^g(\mathbf{Z}_t)\} \subseteq \{\ell_k(\mathbf{Z}_t)\}$ for drift and diffusion, respectively. Then, we express the i^{th} drift component and the ij^{th} term of diffusion covariance matrix as a linear superposition of the library functions as,

$$\mathbf{h}_i(\mathbf{Z}_t, t) = \theta_{i_1}^f \ell_{i_1}^f(\mathbf{Z}_t) + \dots + \theta_{i_k}^f \ell_{i_k}^f(\mathbf{Z}_t) + \dots + \theta_{i_K}^f \ell_{i_K}^f(\mathbf{Z}_t) \quad (21a)$$

$$\Gamma_{ij}(\mathbf{Z}_t, t) = \theta_{ij_1}^g \ell_{ij_1}^g(\mathbf{Z}_t) + \dots + \theta_{ij_k}^g \ell_{ij_k}^g(\mathbf{Z}_t) + \dots + \theta_{ij_K}^g \ell_{ij_K}^g(\mathbf{Z}_t), \quad (21b)$$

where $\theta_{i_k}^f$ and $\theta_{ij_k}^g$ are the weights associated with the k^{th} basis function of i^{th} drift and ij^{th} diffusion covariance components, respectively. In a compact form, the above equations can be represented as similar to Eq. (4),

$$\mathbf{Y}_i = \mathbf{L}^f \theta_i^f + \epsilon_i \quad (22a)$$

$$\mathbf{Y}_{ij} = \mathbf{L}^g \theta_{ij}^g + \eta_{ij}. \quad (22b)$$

In the above equations $\theta_i^f = [\theta_{i_1}, \theta_{i_2}, \dots, \theta_{i_K}]^T$, and $\theta_{ij}^g = [\theta_{ij_1}, \theta_{ij_2}, \dots, \theta_{ij_K}]^T$, which corresponds to i^{th} drift component and ij^{th} element of diffusion covariance matrix, respectively. Similarly, the target vectors \mathbf{Y}_i and \mathbf{Y}_{ij} corresponds to the i^{th} -drift component and $(ij)^{\text{th}}$ component of the diffusion covariance matrix, respectively. The terms ϵ_i and η_{ij} represent the corresponding mea-

surement error vectors. For the discovery, the target vectors are constructed using the Eq. (20) as,

$$\mathbf{Y}_i = \frac{1}{\Delta t} [(Z_{i,1} - \zeta_{i,1}), \dots, (Z_{i,N} - \zeta_{i,N})]^T \quad (23a)$$

$$\mathbf{Y}_{ij} = \frac{1}{\Delta t} [\{(Z_{i,1} - \zeta_{i,1})(Z_{j,1} - \zeta_{j,1})\}, \dots, \{(Z_{i,N} - \zeta_{i,N})(Z_{j,N} - \zeta_{j,N})\}]^T \quad (23b)$$

The straightforward application of the Algorithm 1 in Section 3 in the above directly yields - (i) the perturbation terms in drift and (ii) the diffusion, along with their parameters θ_i^f and θ_{ij}^g . For more information on the discovery of these drift and diffusion terms, one can refer [45,55].

5. Example problems

We demonstrate the efficacy, effectiveness, and robustness of the proposed digital twin framework using the following test beds, (a) an SDOF nonlinear Duffing oscillator, (b) a 2DOF nonlinear system, and (c) a degrading dynamical system. The system parameters are listed in Table 1. In the deterministic framework, it is assumed that we have access to the noisy measurements of both the system states and input. Therefore, we have included the force vector $f(t)$ as a basis function in the library. In this case, the input to the systems is modeled as zero mean Gaussian white noise with appropriate noise intensities, and the system responses are simulated using the Runge-Kutta 45 scheme. On the contrary, for the stochastic framework, it is assumed that we have access to the noisy measurements of system states only. The unknown input to the systems is modeled as Brownian motions with appropriate intensities. The Euler Maruyama (EM) scheme [52] with a sampling frequency of 1000 Hz is utilized to obtain the ensemble of stochastic system responses.

Reiterating the fact that we want to discover the interpretable form of any new function that might have introduced nonlinearity and/or change in the behavior in the original system, the previous information of the system is removed from the measured responses. Since the measurements in practice are always affected by the quality of sensors and the platform of operation, the measurements are corrupted with noise modeled as zero-mean Gaussian noises with a standard deviation equal to 5% of the standard deviation of the measurements. Once the corresponding responses are finally prepared, the target vectors and the libraries of candidate functions for the respective frameworks are obtained. With the target vectors and libraries ready, next, the sparse Bayesian regression is performed to identify the presence of new functions that accurately describes the change in the behavior of the underlying system. In this work, the dictionary $\mathbf{L} \in \mathbb{R}^{N \times K}$ is constructed from the following 8 types of mathematical functions, each function representing a mapping of the m -dimensional state vector $\mathbf{X} = \{X_1, X_2, \dots, X_m\}$:

$$\mathbf{L}(\mathbf{X}) = \begin{bmatrix} \mathbf{1} & p^1(\mathbf{X}) & p^2(\mathbf{X}) & \dots & p^6(\mathbf{X}) & \text{sgn}(\mathbf{X}) & e^{-\mathbf{X}} & e^{-\mathbf{X}\mathbf{X}} \end{bmatrix}. \quad (24)$$

Table 1
Parameters of the undertaken systems.

Simulated systems	Parameters
Example 1 ^(5.1)	$\{m, c, k, \alpha, \sigma\} = \{1, 2, 1000, 100000, 0.5\}$
Example 2 ^(5.2)	$\{m_1, m_2, c_1, c_2, k_1, k_2, \alpha, \sigma_1, \sigma_2\} = \{1, 1, 4, 4, 4000, 2000, 50000, 0.5, 0.5\}$
Example 3 ^(5.3)	$\{m, c, k, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \gamma, \beta, \sigma\} = \{1, 2, 2000, 0.5, 0.5, 1, 1, 0.001, 2, 1\}$

Here, for $i, j = 1, \dots, m$ $\mathbf{1} \in \mathbb{R}^N$ denotes the N -dimensional vector of 1, $\mathcal{P}^\varphi(\mathbf{X}) \in \mathbb{R}^{N \times m}$ denotes the set of terms present in the multinomial expansion $(X_1 + X_2 + \dots + X_m)^\varphi$, $\forall \varphi = 1, \dots, 6$, $\text{sgn}(\mathbf{X}) \in \mathbb{R}^{N \times m}$ represents the signum function of the states as, $\text{sgn}(X_i)$, $e^{-\mathbf{X}}$ and $e^{-\mathbf{X}\mathbf{X}}$ represents the functions e^{-X_i} and $e^{-X_i X_j}$, $|\mathbf{X}| \in \mathbb{R}^{N \times m}$ denotes the absolute mapping of the states as, $|X_i|$, $\mathbf{X}|\mathbf{X}|$ represents the set of functions: $X_i|X_j|$, and, $\sin(\mathbf{X})$ and $\cos(\mathbf{X})$ represents the sine and cosine functions of systems states.

To start the Bayesian model updating algorithm, the deterministic hyperparameters are initialized as $a_p=0.1$, $b_p=1$, $a_v = b_v=0.5$, $a_\sigma = b_\sigma=10^4$ and the following values are used as an initial guess for the random hyperparameters: $p_0^{(0)}=0.1$, $\vartheta^{(0)}=10$, and $\sigma^{2(0)}$ is set equal to the residual variance from ordinary least-squares regression [43,44]. The initial vector of binary latent variables $\mathbf{Z}^{(0)}$ is computed by initializing $\mathbf{Z}^{(0)}$ as zero vector and then activating the components $Z_k \in \mathbf{Z}^{(0)}$ that reduce the mean-squared error between the training data and the obtained model from ordinary least-squares. With the above parameters, the initial value of $\theta^{(0)}$ is then obtained from Algorithm 1. A Markov chain with 3000 Monte Carlo samples is utilized to obtain the posterior distributions of candidate library functions.

5.1. Example 1

As a first test bed, we consider an SDOF dynamical system, where we first assume that the nominal model is in the form of a linear mass-spring-dashpot system. However, due to the operational and environmental conditions, the physics is perturbed by a cubic dissipation term, in which case the system evolves as a Duffing oscillator. Towards the problem statement, the governing motion equations of the dynamical system in its nominal and perturbed form are described as,

$$\left. \begin{aligned} m\ddot{X} + c\dot{X} + kX &= \sigma f(t); && \text{Nominal model,} \\ m\ddot{X} + c\dot{X} + kX + \alpha X^3 &= \sigma f(t); && \text{Perturbed model,} \\ X(t = t_0) &= X_0; && t \in [0, T], \end{aligned} \right\} \quad (25)$$

where $m \in \mathbb{R}$, $c \in \mathbb{R}$, $k \in \mathbb{R}$ and $\alpha \in \mathbb{R}$ are the parameters of the oscillator. The parameter α represents the nonlinear spring constants, and based on the sign of α , the system exhibits the hardening ($\alpha > 0$) and softening ($\alpha < 0$) behavior. The term $f(t) \in \mathbb{R}^n$ denotes the n -dimensional external forcing function. For framework-1, where both the input–output measurements are available, the force term $f(t)$ is modeled as zero-mean standard Gaussian noise with intensity 1. In the case of framework-2, where only the output state measurements are available, the non-measurable force $f(t)$ is modeled as stochastic Brownian motion. The values of the parameters are provided in Table 1. Here our aim is to correct the nominal model using the freshly observed noisy sensor measurements. In addition to discovering the nonlinear dissipation term, we invoke the constraint that the new terms should be in explainable form. Further, when only the output information is available, we additionally aim to identify the diffusion term σ . For this purpose, we first simulated the perturbed system using the state-space $[X, \dot{X}] = [X_1, X_2]$ and then removed the information of the nominal model from the simulated data as, $\alpha X_1^3 = f(t) - m\ddot{X}_2 - kX_1 - cX_2$. The response simulation for framework-1 is done using fourth-order Runge–Kutta, and for framework-2, we used the EM scheme.

5.2. Example 2

In the second example, we consider a two-DOF dynamical system. Similar to the previous case, the nominal model is considered to be a two-DOF linear mass-spring-dashpot system which later

gets perturbed by the cubic dissipation terms. The governing motion equations of the nominal and perturbed model are as follows,

$$\underbrace{\begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \begin{bmatrix} \ddot{X}_1 \\ \ddot{X}_2 \end{bmatrix} + \begin{bmatrix} c_1 + c_2 & -c_2 \\ -c_2 & c_2 \end{bmatrix} \begin{bmatrix} \dot{X}_1 \\ \dot{X}_2 \end{bmatrix} + \begin{bmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_2 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}}_{\text{Nominal model}} + \underbrace{\begin{bmatrix} \alpha X_1^3 + \alpha(X_1 - X_2)^3 \\ \alpha(X_2 - X_1)^3 \end{bmatrix}}_{\text{Perturbation}} = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} f_1(t) \\ f_2(t) \end{bmatrix}, \quad (26)$$

where m_i, c_i and k_i for $i = 1, 2$ are the mass, damping and stiffness parameters of the i^{th} floor. The term α is the nonlinear spring constant, $f_i(t)$ is the external forcing functions, and σ_i is the intensity of the forcing function. Here, $\mathbf{X} = [X_1, X_2]$ is the solution vector of the system. In framework-1, the force vector is generated from the zero mean Gaussian white noise, whereas, in framework-2, the force vector is modeled from the Brownian motion. For solving the systems we adopted the same schemes as mentioned previously. In this example, we firstly aim to discover the nonlinear perturbation terms ($\alpha X_1^3 + \alpha(X_1 - X_2)^3$ in the first DOF and $\alpha(X_2 - X_1)^3$ in the second DOF) in their explainable forms. Secondly, when the input information is not available we also try to identify the diffusion (σ_1^2/m_1^2 in the first DOF and σ_2^2/m_2^2 in the second DOF). We have demonstrated the results in terms of the state-space $[X_1, \dot{X}_1, X_2, \dot{X}_2] = [X_1, X_2, X_3, X_4]$. Once the responses are simulated using the mentioned numerical schemes, the information about the nominal model is removed from the responses, as mentioned in the previous example.

5.3. Example 3

In this example, we considered a more sophisticated and near-realistic problem that involves the discovery of crack degradation in a linear dynamical system. The degradation of stiffness due to fatigue accumulation during the vibration process is a real phenomenon and has great importance in engineering practice. For simulating the degradation, we particularly adopted the model proposed in Ref. [34]. With this, the governing motion equations of the underlying problem are given as,

$$m\ddot{X}_1(t) + c\dot{X}_1(t) + k\lambda X_1(t) = f(t) \quad (27a)$$

$$\lambda = \alpha_1 + \alpha_2 \exp(-\alpha_3 q(t)^{\alpha_4}) \quad (27b)$$

$$\dot{q}(t) = \gamma \left(X_1^2(t) + \dot{X}_1^2(t) \right)^{\frac{\beta}{2}}. \quad (27c)$$

Here $m \in \mathbb{R}$, $c \in \mathbb{R}$, and $k \in \mathbb{R}$ are the mass, damping, and stiffness parameters of an actual linear system. The scalar $\lambda \in \mathbb{R}$ characterizes the dependency of stiffness on the degradation and $\alpha_1 \in \mathbb{R}^+$, $\alpha_2 \in \mathbb{R}^+$, $\alpha_3 \in \mathbb{R}^+$, $\alpha_4 \in \mathbb{R}^+$ defines the extent and rate of degradation. While the Eq. (27a) represents the actual system, the Eq. (27c) denotes the evolution of the degradation measure $q(t)$. For more details, the readers are referred to Ref. [34]. In a similar fashion to previous examples, here we aim to simultaneously discover the Eq. (27c), the nonlinear degradation terms $k(x_3) = k(\alpha_1 + \alpha_2 \exp(-\alpha_3 q_3^{\alpha_4}))$ in Eq. (27a) and the diffusion σ/m in Eq. (27a). Thus it can be noticed that although the initial system was linear, the identified system is highly nonlinear in nature. As it was explained in the previous examples, the forcing $f(t)$ is simulated as zero mean Gaussian white noise for framework-1 and as Brownian motion for framework-2.

5.4. Discovery results when both noisy input–output information are available

The discovery results for the DVP, two-DOF, and crack-degradation examples are illustrated both quantitatively and graphically in Table 2 and Fig. 3–6. In Fig. 3 the posterior inclusion probabilities of the candidate library functions are shown. In Fig. 3

(a), it is clearly evident that the proposed digital twin framework has correctly identified the basis function corresponding to the cubic nonlinearity in the perturbed model of Example 1. This similarity can be observed in Fig. 3(b) and Fig. 3(c), where all the coupled cubic nonlinear basis functions in the perturbed model of the two-DOF system are identified. In 3(d), the results are no exception, where we can observe that the proposed framework is

Table 2
Posterior mean and standard deviations of the selected basis functions.

Systems	Basis function	*Deterministic	†Stochastic	
Example 1 ^(5,1)	X^3	100000 ± 0.5587	100074 ± 124.48	
	u	$1.00 \pm 2.78 \times 10^{-7}$	–	
	$\dagger\sigma(X_t, t)$	–	0.5101 ± 0.0142	
Example 2 ^(5,2)	First DOF	X^3	-100000 ± 1.43	-100173 ± 1206.32
		$X^2\dot{X}$	150000 ± 2.38	150287 ± 1786.58
		$X\dot{X}^2$	-150000 ± 1.27	-150234 ± 841.33
		\dot{X}^3	50000 ± 0.2259	49967.9 ± 123.33
		u_1	$1.00 \pm 2.74 \times 10^{-7}$	–
	Second DOF	$\dagger\sigma_1(X_t, t)$	–	0.5346 ± 0.0042
		X^3	50000 ± 1.41	49939.9 ± 1322.00
		$X^2\dot{X}$	-150000 ± 2.35	-150098 ± 1934.90
		$X\dot{X}^2$	150000 ± 1.25	150051 ± 883.81
		\dot{X}^3	-50000 ± 0.2247	-49992.6 ± 128.34
Example 3 ^(5,3)	Crack Path	u_2	$1.00 \pm 2.89 \times 10^{-7}$	–
		$\dagger\sigma_2(X_t, t)$	–	0.5021 ± 0.0029
		X^2	$0.0099 \pm 5.29 \times 10^{-5}$	$0.0099 \pm 8.00 \times 10^{-4}$
	System	\dot{X}^2	$0.0100 \pm 4.02 \times 10^{-8}$	$0.0099 \pm 4.52 \times 10^{-7}$
		$\alpha_1 X$	$-1000 \pm 3.51 \times 10^{-5}$	-995.74 ± 22.62
		$\alpha_2 e^{-2s\psi^{24}} X$	$-1000 \pm 1.00 \times 10^{-4}$	-1003.93 ± 22.74
	u	$1.00 \pm 1.39 \times 10^{-5}$	–	
	$\dagger\sigma(X_t, t)$	–	1.00 ± 0.0549	

* The deterministic refers to the case when both the input–output information are available, and † stochastic refers to the case when the output-only response is available. ‡ Note that the diffusion terms are discovered in terms of their covariation, i.e., to discover the diffusion terms one needs to perform the square root operation on the covariance matrix Γ .

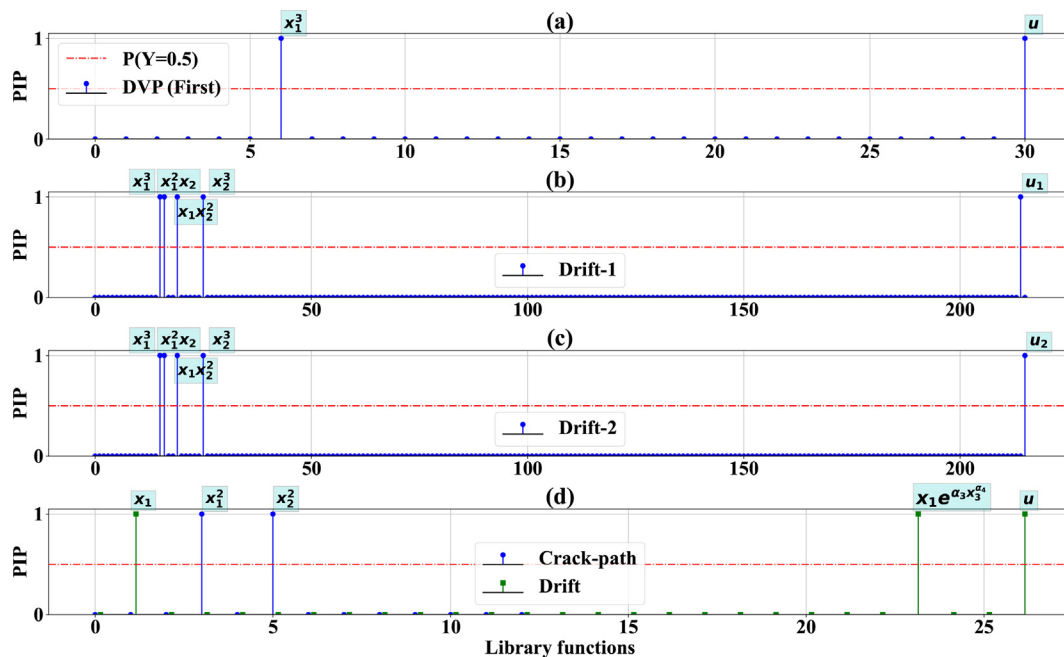


Fig. 3. Basis function selection for example problems when both input–output measurements are available. (a) Model selection in Example-1, (b) model selection in the first DOF of example-2, (c) model selection in the second DOF of example-2, and (d) model selection in the example-3. The library is constructed as $L \in \mathbb{R}^{31}$ for example-1, $L \in \mathbb{R}^{216}$ for example-2 and $L \in \mathbb{R}^{27}$ for example-3. Out of 31 only 2 in example-1, out of 216 only 5 in example-2, and out of 27 only 5 in example-3 have PIP > 0.5. This indicates the ability of the proposed framework to introduce sparsity and retain only the most important basis functions in the solution.

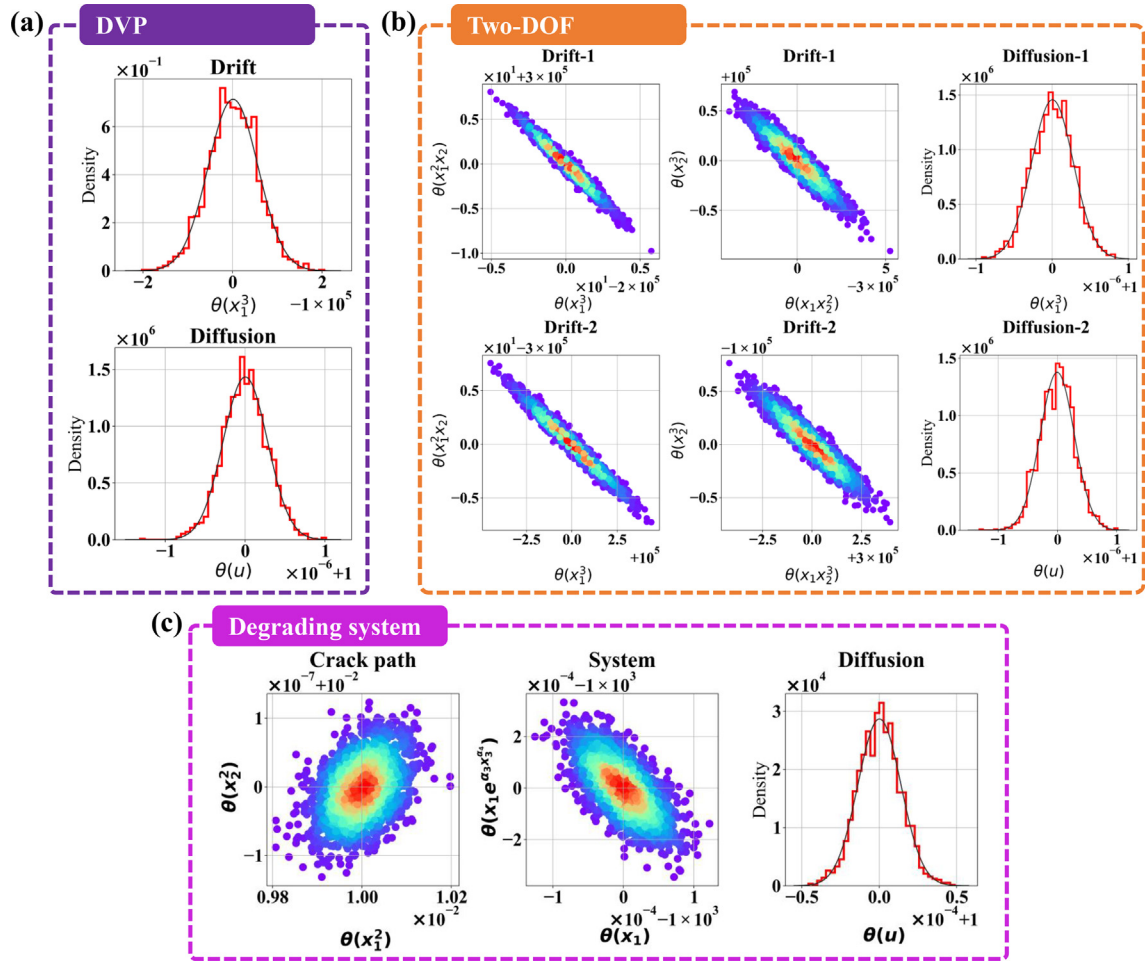


Fig. 4. Posterior probabilities of the selected basis functions in the final model obtained using framework-1. (a) Example-1: the density plots of the basis X^3 and u . (b) Example-2: posterior and joint posterior probabilities of the basis functions $X^2, X^3, X^2X, XX^2, u_1$ and u_2 . (c) Example-3: posterior densities of the basis functions $\theta(X^2), \theta(X^2)$ and $Xe^{(-\theta)}$. The red region in the joint posterior density indicates the mean of parameters, whose values are given in Table 2.

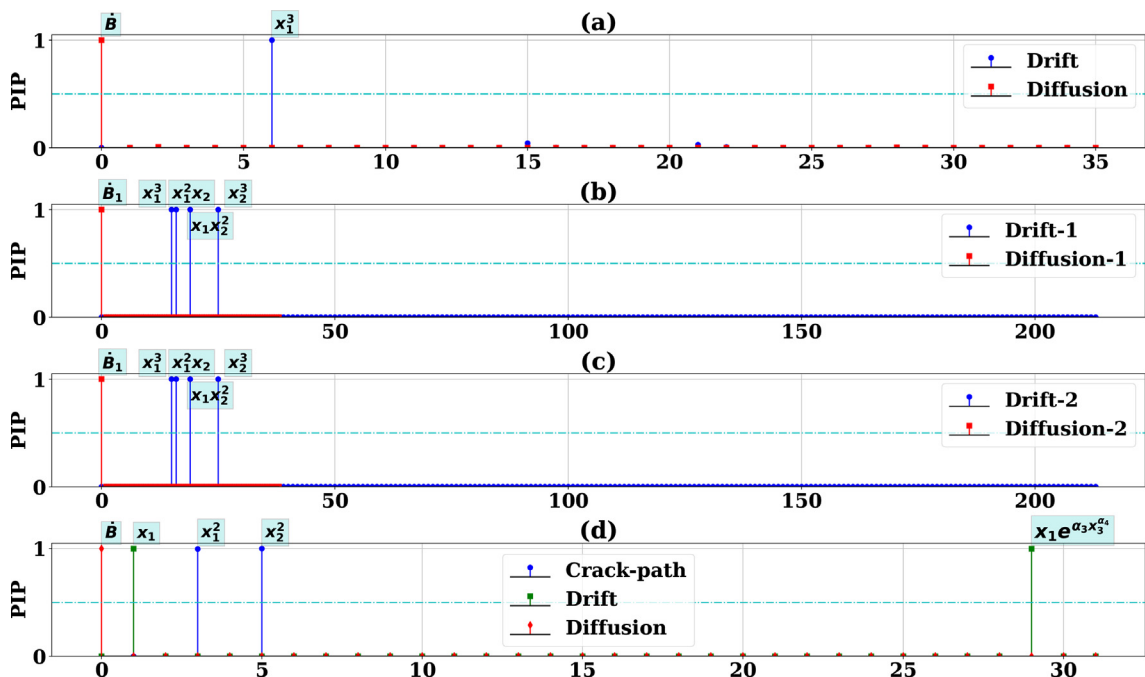


Fig. 5. Basis function selection for the example problems when only state measurements are available. (a) Example-1, (b) example-2: drift, (c) example-2: diffusion, (d) example-3. The blue bars correspond to perturbation terms in drift, and the red bars represent diffusion. A total of $\mathbb{R}^{36}, \mathbb{R}^{215}$, and \mathbb{R}^{32} basis functions are considered for the examples, respectively. The $PIP > 0.5$ criteria are invoked to discover the final model.

able to identify the basis functions corresponding to crack path and degradation exactly.

The posterior distributions of the parameters of the identified basis functions are further depicted in Fig. 4. For quantitative understanding, the mean and standard deviations of the parameters are provided in Table 2. For testing the fidelity of the proposed framework, we compare the parameter values in entries against the deterministic framework of Table 2 with those in the actual system (given in Table 1). From the comparison results, it can be stated that the mean values of the corresponding parameters (zero error) exactly match with actual values in Table 1. The standard deviation results demonstrate the uncertainties associated with the selection of basis functions in the final model. From the results, it is evident that the uncertainties associated with the corresponding basis functions that are identified using the proposed scheme are quite small. This indicates the efficacy of the proposed framework in discovering the perturbations in the system in their interpretable forms.

5.5. Discovery results when only the noisy output measurements are available

The results of the discovery of the drift and diffusion terms in the perturbed models of undertaken examples, when the input force information is unavailable are presented in Fig. 5 and 6. In

the first example, we aimed to discover the nonlinear drift term αx^3 and the diffusion constant σ from the output state measurements only. From the results for the basis function selection in example-1, displayed in Fig. 5(a), it is evident that the proposed digital twin framework is able to identify the perturbation terms in their interpretable forms correctly. In the second example, in addition to the nonlinear dissipating terms $(\alpha X_1^3 + \alpha(X_1 - X_2)^3$ in the first DOF and $\alpha(X_2 - X_1)^3$ in the second DOF), we also aim at discovering the diffusion terms σ_1/m_1 and σ_2/m_2 . Similar to example-1, from the basis function selection results presented in 5(b) and 5(c) it is easy to comprehend that the proposed framework is able to correctly identify the highly nonlinear basis functions without any greater difficulty. In the third example, similarly, we try to identify the diffusion term σ in addition to the crack growth and stiffness degradation without using the input force measurements explicitly. The results in Fig. 5(d) show the exact identification of the corresponding perturbation terms, which introduces the degradation in the underlying dynamical system.

For a quantitative understanding of the performance of the proposed framework, the posterior distributions of the parameters of the identified basis functions in the undertaken models are presented in Fig. 6. Further, the statistical properties of the parameters are given in Table 2 for easy reference. In the first observation on comparing the values provided in Table 2 against stochastic entry

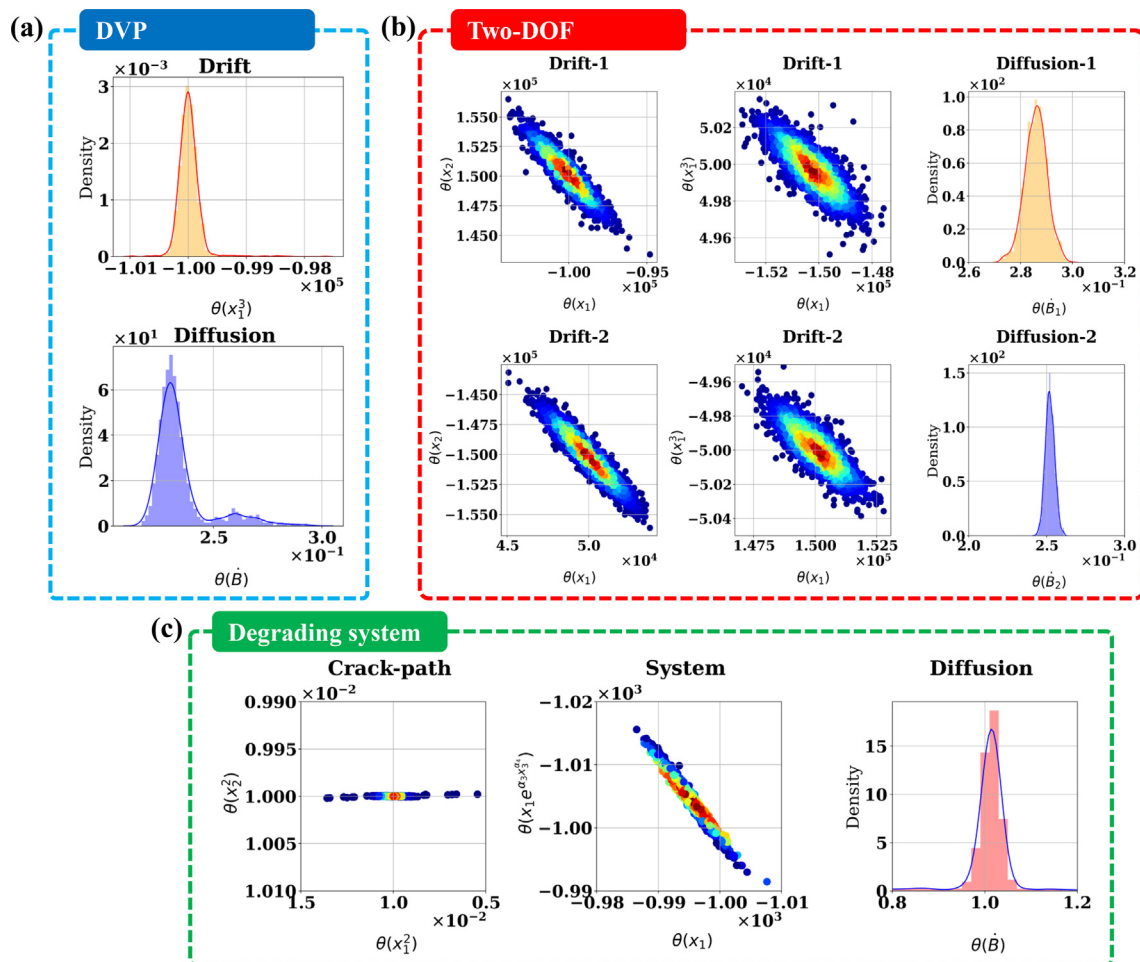


Fig. 6. Posterior distributions of the identified basis functions obtained using the framework-2. (a) Example-1: posterior distributions of the drift term X^3 and covariance σ^2 . (b) Example-2: joint posterior distributions of the basis functions $X^3, X^2 \dot{X}, X \dot{X}^2, \dot{X}^3, \sigma_1^2/m_1^2$ and σ_2^2/m_2^2 in first and second DOF, respectively. (c) Example-3: posterior distributions of the basis functions $X^2, \dot{X}^2, X e^{(-q)}$ and σ^2 . The values of the diffusion terms are obtained in terms of their covariance. The final values of diffusion provided in Table 1 are obtained by performing the square root operation on the covariance.

with those in Table 1, it can be stated that the error in the identified parameters is very small (around 0.06–3.93% for drift and 0.42–6.92% for diffusion). This means that the proposed digital twin framework can discover the correct perturbation terms with sufficient accuracy. On referring to the associated standard deviation values of the corresponding parameters, it is straightforward to understand that when the input force information is not available to the proposed algorithm, the uncertainty in the identified parameters increases. On the contrary, when both input–output information is available the uncertainty in the identified parameters decreases significantly.

5.6. Prediction using the proposed DT

The true potential of a DT is generally visualized by its ability to predict the future in the presence of unseen environmental disturbances. In order to judge the predictive performance of the proposed DT we have carried out predictions using the previously updated models on new random forcing values. The random forcing functions are simulated from a standard normal distribution with the intensities provided in Table 2. The predictions are performed using the identified models of both frameworks 1 and 2. The results on the prediction of system states of example-1, 2, and 3 are illustrated in Fig. 7–9, respectively. In Fig. 7, the prediction results for example-1 along with the 95% confidence interval are shown. It can be seen that the prediction results exactly emulate the original response. Additionally, it can be observed that the standard deviations are so small (given in Table 2) that the confidence interval completely overlaps the mean prediction results. This means that the predictions performed using the mean value of the parameters of the updated model have very less chance of diverging away from the actual results.

In Fig. 8, the prediction results along with their 95% confidence interval for system states of example-2 are demonstrated. The observations are similar to example-1, i.e. the predictions performed using the updated model are nearly the same as the original

predictions. On close observation, it can be seen that the updated model has also captured the sharp changes in the actual system. Similar to the previous case, the standard deviations in the identified system parameters are so small that the 95% confidence interval is very small, indicating negligible uncertainty in the identified model. Thus it can be envisaged that the proposed DT can learn the underlying physics of highly nonlinear perturbations and thereby can be used to update models in very complex environmental conditions. In Fig. 9, the results for example-2 are portrayed. The mean prediction results for the system states are very much similar to the actual system. The updated model has almost exactly predicted the actual system responses. However, in this case, it can be observed that when only the output measurements are used to update the DT, the uncertainties in the identified parameters increase. As a consequence, the uncertainties in the predictions performed using the updated model increase.

Further, to assess the long-term predictive ability of the proposed framework, we have performed predictions for the nonlinear oscillator in example-2 for a total duration of 500s at a sampling frequency of 1000 Hz using the EM scheme. During the predictions, the system parameters are kept the same as in Table 2. The results are shown for the second DOF in Fig. 10. We observe that the mean predictions emulate the actual solutions almost exactly in the initial phase. Consequently, the corresponding confidence interval is also very narrow. The long-term predictions after the 250s show very small deviations from the actual results, which is further depicted by the increased width of the confidence interval. Nevertheless, we observe that during the prediction period of 500s, the predicted responses oscillate very close to the actual solution indicating the robustness of the identified models and so as the robustness of the proposed frameworks.

5.7. Sensitivity to measurement noise

In the presence of environmental disturbances and operational faults, the measurements of system responses using sensors from

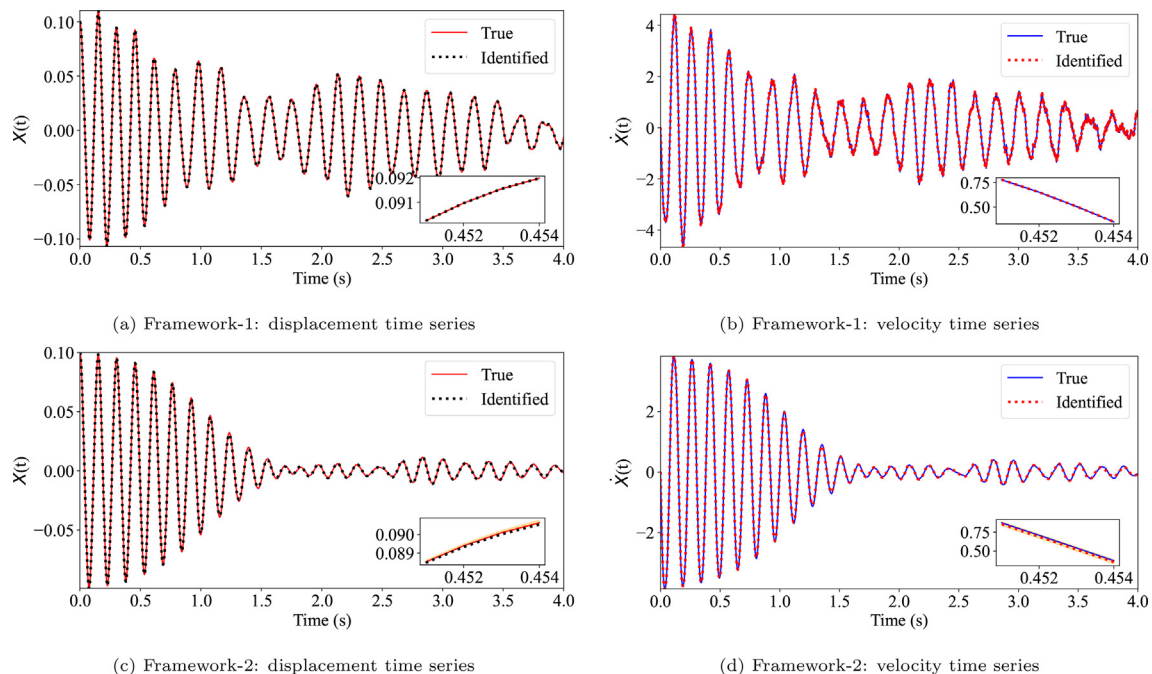


Fig. 7. Predictive performance of the proposed predictive digital twin. (a) and (b) Predictions result for the DT when both the input–output measurements are available. (c) and (d) Prediction results of the DT with output-only measurement. In both cases, the proposed DT showcases exceptional predictive ability. Additionally, the uncertainties associated with the identified parameters are so less that it does not get reflected in the predicted responses. This indicates that the proposed DT can be employed in implementations where high precision in the identifications is of primary importance.

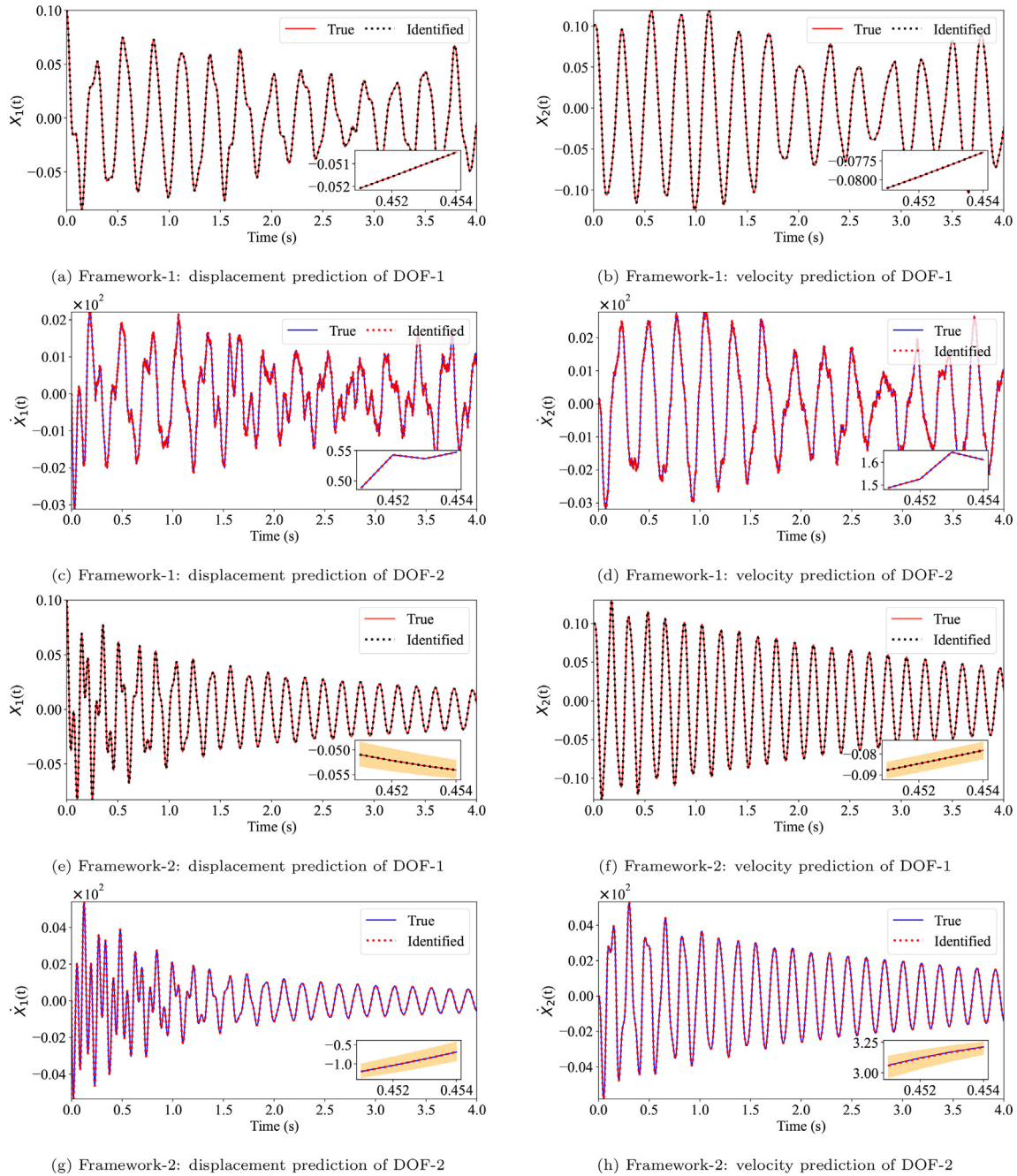


Fig. 8. Prediction performance of the proposed predictive digital twin for example-2. (a), (b), (c) and (d) Prediction results of the system states obtained using the model updated via input–output observations. (e), (f), (g), and (h) Prediction results of the system state when only the output measurements are used for updating the DT. The prediction results are highly accurate. The updated model is very effective in capturing the highly nonlinear behavior of the undertaken system, as it is able to capture the sharp changes in the system behavior with very small uncertainties.

field applications always get corrupted by surrounding noises. To be able to implement on an industrial scale, a novel framework should demonstrate the soundness of such noises and effectively correct the physics of nominal models using freshly observed noisy measurement data. While in the previous results, the noise of magnitude of 5% of the standard deviation of the input is considered, in this section, we carried out the same study using 13 different levels of noise ranging from 0 to 60%. The performance is compared using the absolute relative error between the actual parameter values and the mean of the posterior distributions. Similar to previous cases, the case studies are performed in search of new basis functions in drift and as well as for diffusion terms. The absolute rela-

tive error in the results of discovery using framework-1, along with their statistics, are provided in Fig. 11. In Fig. 11a, the sensitivity of the deterministic framework where both the input–output information is available is presented.

In these figures, the mean error is denoted by the marker star, and the median of the error is denoted by the black bars of the box-plot. In the identification of the basis function x^3 , it is observed that the mean values of the errors for different levels of noise are quite negligible (so as the confidence intervals) for noise levels of up to 45%, indicating the soundness of the proposed framework against noise. However, as the noise level reaches 50%, the proposed framework predicts the wrong parameters of the actual basis func-

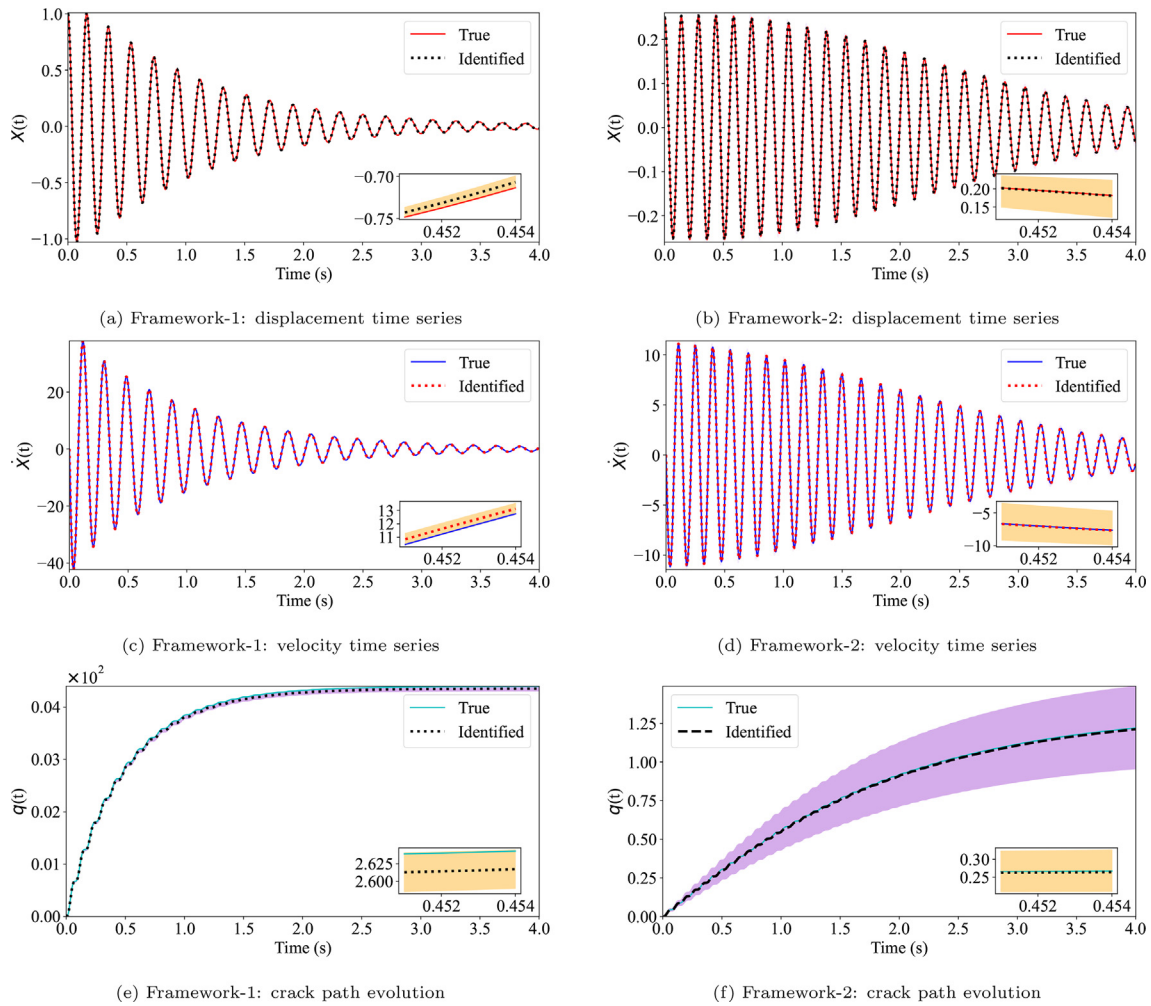


Fig. 9. Predictive performance of the proposed predictive digital twin for example-3. (a) and (b) Results for the DT using framework-1, where both the input–output observations are available. (c) and (d) Results of the DT when only output measurements are feasible. The DT perfectly identifies the terms of the perturbation along with their associated parameters. As a result, the prediction results match almost perfectly with the actual system responses. However, when the models are updated using only the output observations, the uncertainty in the predictions increases by some amount. This ability to learn the uncertainties in the identified system parameters helps us to perform reliability analysis on the systems designed using the proposed DT.

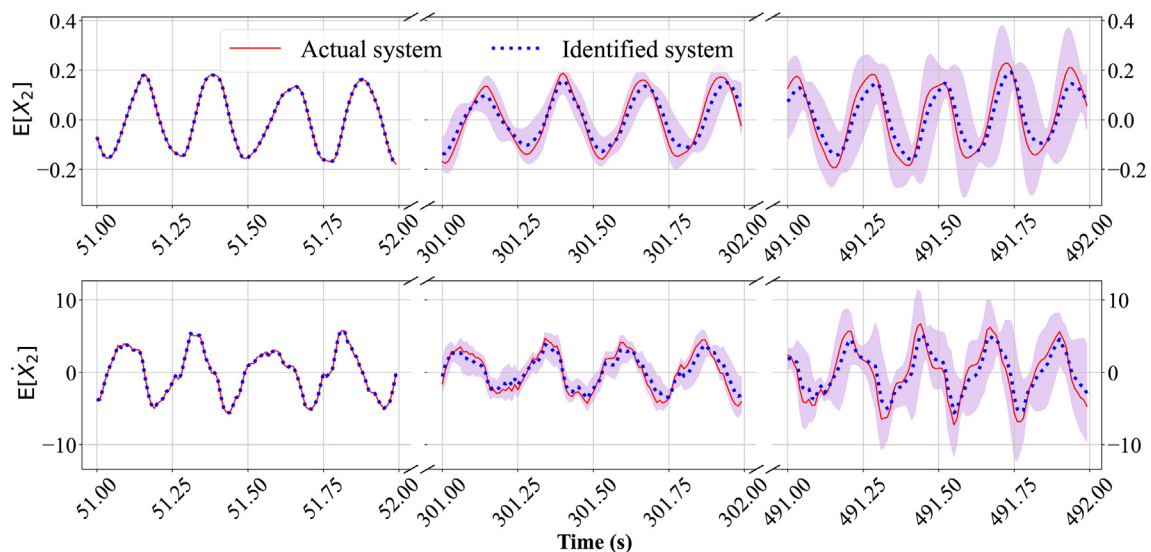


Fig. 10. Long-term predictive performance of the proposed predictive digital twin for the nonlinear oscillator in Example 2. The displacement and velocity responses are shown for the second DOF, where predictions are performed for unseen realizations of white noise. The mean prediction responses in the initial stage (< 250s) show an exact match to the true solution. However, the responses after the 250s show very small deviations. This phenomenon is also depicted by the evolution in the confidence interval, where the width of the interval in the initial phase is very narrow but later grows with time.

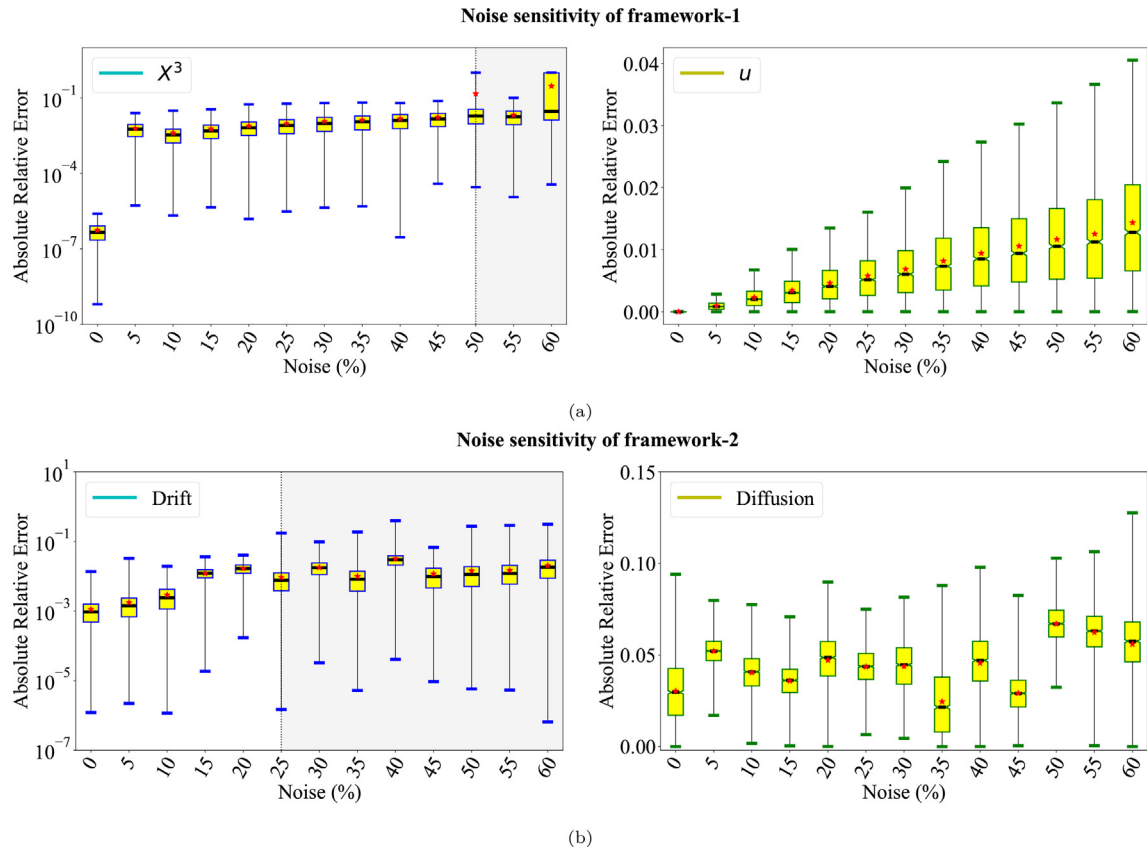


Fig. 11. Performance of the proposed predictive digital twin for example-1 at different noise levels. (a) The absolute relative error in the identified parameters of the basis functions X^3 and u , respectively. (b) The absolute relative error in the identification of the parameters of bases X^3 and \hat{B} , respectively. The marker star denotes the mean and the black bars denote the median of the error, and the shaded region indicates the breakpoint of the proposed frameworks.

tion. In cases of observations with 60% noise, the proposed framework predicts not only predicts wrong basis parameters but also populates the model with wrong basis functions. Consequently, the mean error also shoots up, and the deviation of the error becomes 100%. Similarly, in the identification of the forcing basis function u , we observe that as the noise level increases, the identification error of the forcing intensity increases, also indicated by the increasing confidence interval. Based on the analysis, it can be envisioned that when both the input–output information is measurable, the proposed framework can perform well for data that are corrupted with a noise of magnitude of up to 45%.

In Fig. 11b, the noise sensitivity of framework-2 is evaluated, where only the noisy measurements of systems responses are available. The absolute relative error in the identification results of drift shows that variations in the error mean with the increase in noise levels are very erratic. Although the associated confidence intervals are quite stable, the values of the identified parameters vary over a big range. A similar erratic trend can be seen in the results of diffusion identification. The proposed framework-2 is able to diffusion parameters with a mean error of $\leq 10\%$. Thus by considering the mean error levels and the associated bounds of the parameters we can envision that the proposed framework-2 can extract the exact basis functions with accurate parameters from output-only data that are corrupted with noises of the level of up to 20%.

6. Conclusions

A framework for real-time updating of the DTs using a library of physics-based functions is proposed. Two approaches for updating

the DT are proposed, where the first approach utilizes both input–output data, and the second approach uses output-only observations. Since, in noisy and limited data, dealing with the probability distribution of a random variable would be preferable to dealing with a predicted value, we utilize the sparse Bayesian regression to infer the perturbation terms from the library. As compared to the available grey-box DTs where the precise representation of the obtained governing physics is unresolved, the proposed framework has high predictive power since the actual governing physics of the perturbed model is learned using physical functions instead of surrogate models. Data availability is the foundation of any DT. However, sometimes obtaining the required data is not possible due to unavoidable circumstances. In such cases, the resulting framework can work in situations when either or both the input and output measurements are available.

Three numerical case studies are undertaken. In the first example, a simple SDOF linear dynamical system is considered where the perturbation is assumed to be a cubic dissipating force. In the second example, a two-DOF dynamical system is considered, which is assumed to be perturbed by coupled cubic nonlinear terms. In the third example, a near-realistic problem of stiffness degradation due to crack propagation is taken. In all the problems, the primary system is assumed to be linear, but as time progresses, the perturbed model becomes nonlinear. Further case studies on the performance of the proposed framework against different noise levels indicate that the framework can identify the exact perturbation from data that are corrupted with the noise of level - (i) 45% when both the input–output information are available and (ii) 20% when only the input information is available.

Overall, the salient features of the proposed work can be encapsulated in the following points:

- The actual physics of the perturbations in the physical twin is discovered using the proposed approach. The discovered terms are expressed in terms of interpretable functions; thus, the proposed approach is white in nature.
- The Bayesian approach reduces the chance of overfitting and avoids the requirement of human interventions for optimal tuning. Therefore, it can be directly applied for automation in the real-time monitoring, diagnosis, and prognosis of dynamical systems.
- The proposed approach presents a unique way of updating the governing physics of DT from either both input–output or output-only observations. The latter case is especially relevant when the measurement of input forces is intractable.
- Being probabilistic in nature, the proposed framework provides first and second-order statistics for quantifying the uncertainties arising due to noisy and low data.

A few potentials application of the proposed predictive framework includes (i) instances when the governing physics is misspecified or the physics of the underlying physical twin has changed, (ii) when the observations are noisy and limited, (iii) situations where only the output observations are available, and (iv) when long term predictions under a rapid change in environmental conditions are required for predictive maintenance. However, the proposed framework has certain issues, for e.g., the choice of correct basis functions. In the absence of correct basis functions and correct state measurements, the proposed framework will not yield a sparse solution, thereby limiting the generalization of the discovered model. In such cases, the knowledge of an expert can be combined with the associated uncertainties of the basis functions to gain some diagnostic information. In this work, we further modeled the input information as stationary Gaussian white noise, which may not hold true for most of the applications in DT. Therefore, a more practical framework is required which would generalize for realistic input such as nonstationary processes, jump adapted processes, narrow band processes, etc.

Additionally, in order to adopt the proposed framework for industrial and commercial implementations, a few more features are required to be added to the existing framework. For example, partial differential equations (PDEs) are commonly used to describe naturally arising problems like electrodynamics, fluid flow, heat and sound propagation, etc. Therefore, as an initial thought developing DT for model updating of systems involving PDEs can be carried out as an extension of this work. Secondly, if the noise level in the observations is very high, then performing the spares Bayesian inference using MCMC could be computationally time-consuming. Thus, the formulation of computationally tractable probabilistic frameworks for real-time deployment of DT could be another possible extension. Thirdly, the most significant improvement would be to integrate the proposed framework with control algorithms. This would, instead of physical intervention, enable the digital operation of the underlying process resulting increase in output, streamlining of services, and saving in post-investment cost.

Code availability

The source codes to reproduce the results reported in this study can be downloaded from <https://github.com/cscm-iitd/Interpretable-DT>.

Data availability

The data used for the research described in the article can be simulated from the source codes available at <https://github.com/cscm-iitd/Interpretable-DT>.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: [Souvik Chakraborty reports financial support was provided by Science and Engineering Research Board.]

Acknowledgements

T. Tripura acknowledges the financial support received from the Ministry of Education (MoE), India, in the form of the Prime Minister's Research Fellowship (PMRF). S. Chakraborty acknowledges the financial support received from Science and Engineering Research Board (SERB) via Grant No. SRG/2021/000467 and seed grant received from IIT Delhi.

Appendix A. The marginal likelihood function for estimating the latent indicator vector

Due to the Boolean nature of the latent variable ψ_k , it is sampled from the Bernoulli distribution using the hyperparameter p_0 as,

$$p\left(\psi_k^{(i+1)} | \mathbf{Y}, \vartheta_s^{(i)}, p_0^{(i)}\right) = \text{Bern}\left(\frac{p_0}{p_0 + \lambda(1 - p_0)}\right), \quad (\text{A.1})$$

where $\lambda = \frac{p(\mathbf{Y} | \psi_k^{(i)} = 0, \Psi_{-k}^{(i)}, \vartheta_s^{(i)})}{p(\mathbf{Y} | \psi_k^{(i)} = 1, \Psi_{-k}^{(i)}, \vartheta_s^{(i)})}$. Here, $\Psi_{-k}^{(i)} \in \mathbb{R}^{K-1}$ denotes the latent variable vector Ψ with k^{th} element removed. The k^{th} latent variable $\psi_k^{(i)}$ takes a value 0 or 1 with probabilities in Eq. (A.2a) and Eq. (A.2b), respectively,

$$p\left(\mathbf{Y} | \psi_k^{(i)} = 0, \Psi_{-k}^{(i)}, \vartheta_s^{(i)}\right) = \frac{\Gamma(\alpha_\sigma + 0.5N) \beta_\sigma^{2\alpha_\sigma}}{\Gamma(\alpha_\sigma) (2\pi)^{0.5N} \left(\beta_\sigma + \frac{1}{2} \mathbf{Y}^T \mathbf{Y}\right)^{(\alpha_\sigma + 0.5N)}}; \quad \text{when all } \left\{Z_k^{(i)} : k = 1, \dots, K\right\} = 0. \quad (\text{A.2a})$$

$$p\left(\mathbf{Y} | \psi_k^{(i)} = 1, \Psi_{-k}^{(i)}, \vartheta_s^{(i)}\right) = \frac{\Gamma(\alpha_\sigma + 0.5N) \beta_\sigma^{2\alpha_\sigma} \left(\left\|\mathbf{R}_{0,r}^{(i)}\right\| \left\|\Sigma_\theta^{(i)}\right\|\right)^{\frac{1}{2}}}{\Gamma(\alpha_\sigma) (2\pi)^{\frac{N}{2}} \vartheta_s^{\frac{N}{2}} \left(\beta_\sigma + \frac{1}{2} \mathbf{Y}^T \left(\mathbf{I}_{N \times N} - \mathbf{L}_r^{(i)} \Sigma_\theta^{(i)} \mathbf{L}_r^{(i)T}\right) \mathbf{Y}\right)^{(\alpha_\sigma + 0.5N)}}. \quad (\text{A.2b})$$

References

- [1] Främling K, Holmström J, Ala-Risku T, Kärkkäinen M. Product agents for handling information about physical objects. Report of Laboratory of information processing science series B, TKO-B 153 (03); 2003.
- [2] Grieves MW. Product lifecycle management: the new paradigm for enterprises. *Int J Product Develop* 2005;2(1–2):71–84.
- [3] Grieves M. Digital twin: manufacturing excellence through virtual factory replication. *White Paper* 2014;1(2014):1–7.
- [4] Tuegel EJ, Ingrassia AR, Eason TG, Spottswood SM. Reengineering aircraft structural life prediction using a digital twin. *Int J Aerospace Eng* 2011.
- [5] Tao F, Zhang H, Liu A, Nee AY. Digital twin in industry: State-of-the-art. *IEEE Trans Ind Informat* 2018;15(4):2405–15.
- [6] Jones D, Snider C, Nassehi A, Yon J, Hicks B. Characterising the digital twin: A systematic literature review. *CIRP J Manuf Sci Technol* 2020;29:36–52.
- [7] Coronado PDU, Lynn R, Louhichi W, Parto M, Wescoat E, Kurfess T. Part data integration in the shop floor digital twin: Mobile and cloud technologies to enable a manufacturing execution system. *J Manuf Syst* 2018;48:25–33.
- [8] Souza V, Cruz R, Silva W, Lins S, Lucena V. A digital twin architecture based on the industrial internet of things technologies. In: 2019 IEEE International Conference on Consumer Electronics (ICCE). IEEE; 2019. p. 1–2.
- [9] Wang J, Ma Y, Zhang L, Gao RX, Wu D. Deep learning for smart manufacturing: Methods and applications. *J Manuf Syst* 2018;48:144–56.
- [10] Ludwig S, Karrenbauer M, Fellan A, Schotten HD, Buhr H, Seetaraman S, et al. A5g architecture for the factory of the future. 2018 IEEE 23rd international conference on emerging technologies and factory automation (ETFA), Vol. 1. IEEE; 2018. p. 1409–16.
- [11] Arup, Digital twin: Towards a meaningful framework, Technical report, Arup, London, England, 2019.

- [12] Mosavi A, Salimi M, Faizollahzadeh Ardabili S, Rabczuk T, Shamshirband S, Varkonyi-Koczy AR. State of the art of machine learning models in energy systems, a systematic review. *Energies* 2019;12(7):1301.
- [13] Samaniego E, Anitescu C, Goswami S, Nguyen-Thanh VM, Guo H, Hamdia K, et al. An energy approach to the solution of partial differential equations in computational mechanics via machine learning: Concepts, implementation and applications. *Comput Methods Appl Mech Eng* 2020;362:112790.
- [14] Tripura T, Chakraborty S. Wavelet neural operator for solving parametric partial differential equations in computational mechanics problems. *Comput Methods Appl Mech Eng* 2023;404:115783.
- [15] Navaneeth N, Chakraborty S. Stochastic projection based approach for gradient free physics informed learning. *Comput Methods Appl Mech Eng* 2023;406:115842.
- [16] Hernández L, Hernández S. Application of digital 3d models on urban planning and highway design. *WIT Trans Built Environ* 1997;33.
- [17] Boschert S, Rosen R. Digital twin—the simulation aspect. In: *Mechatronic futures*. Springer; 2016. p. 59–74.
- [18] Tao F, Zhang M, Cheng J, Qi Q. Digital twin workshop: a new paradigm for future workshop. *Comput Integ Manuf Syst* 2017;23(1):1–9.
- [19] Booyse W, Wilke DN, Heyns S. Deep digital twins for detection, diagnostics and prognostics. *Mech Syst Signal Process* 2020;140:106612.
- [20] Wang J, Ye L, Gao RX, Li C, Zhang L. Digital twin for rotating machinery fault diagnosis in smart manufacturing. *Int J Prod Res* 2019;57(12):3920–34.
- [21] Millwater H, Ocampo J, Crosby N. Probabilistic methods for risk assessment of airframe digital twin structures. *Eng Fract Mech* 2019;221:106674.
- [22] Zhou M, Yan J, Feng D. Digital twin framework and its application to power grid online analysis. *CSEE J Power Energy Syst* 2019;5(3):391–8.
- [23] Alibrandi U. Risk-informed digital twin of buildings and infrastructures for sustainable and resilient urban communities. *ASCE-ASME J Risk Uncert Eng Syst, Part A: Civil Eng* 2022;8(3):04022032.
- [24] Angjeliu G, Coronelli D, Cardani G. Development of the simulation model for digital twin applications in historical masonry buildings: The integration between numerical and experimental reality. *Comput Struct* 2020;238:106282.
- [25] Haag S, Anderl R. Digital twin—proof of concept. *Manuf Lett* 2018;15:64–6.
- [26] Lu Y, Liu C, Kevin I, Wang K, Huang H, Xu X. Digital twin-driven smart manufacturing: Connotation, reference model, applications and research issues. *Robot Comput-Integ Manuf* 2020;61:101837.
- [27] Park KT, Lee D, Noh SD. Operation procedures of a work-center-level digital twin for sustainable and smart manufacturing. *Int J Precis Eng Manuf-Green Technol* 2020;7(3):791–814.
- [28] He B, Bai K-J. Digital twin-based sustainable intelligent manufacturing: A review. *Adv Manuf* 2021;9(1):1–21.
- [29] Li C, Mahadevan S, Ling Y, Choe S, Wang L. Dynamic bayesian network for aircraft wing health monitoring digital twin. *Aiaa J* 2017;55(3):930–41.
- [30] Kapteyn MG, Knezevic DJ, Willcox K. Toward predictive digital twins via component-based reduced-order models and interpretable machine learning. In: *AIAA Scitech 2020 Forum*; 2020. p. 0418.
- [31] Hoodorozhkov S, Krasilnikov A. Digital twin of wheel tractor with automatic gearbox. *E3S Web of Conferences*, vol. 164. EDP Sciences; 2020. p. 03032.
- [32] Liu Q, Liu B, Wang G, Zhang C. A comparative study on digital twin models. *AIP Conference Proceedings*, vol. 2073. AIP Publishing LLC; 2019. p. 020091.
- [33] Chakraborty S, Adhikari S, Ganguli R. The role of surrogate models in the development of digital twins of dynamic systems. *Appl Math Model* 2021;90:662–81.
- [34] Sobczyk K. Stochastic dynamics and reliability of degrading systems. *Bull Polish Acad Sci: Technical Sci* 2006;54(1).
- [35] Ganguli R, Adhikari S. The digital twin of discrete dynamic systems: Initial approaches and future challenges. *Appl Math Model* 2020;77:1110–28.
- [36] Garg S, Gogoi A, Chakraborty S, Hazra B. Machine learning based digital twin for stochastic nonlinear multi-degree of freedom dynamical system. *Probab Eng Mech* 2021;66:103173.
- [37] Garg S, Chakraborty S, Hazra B. Physics-integrated hybrid framework for model form error identification in nonlinear dynamical systems. *Mech Syst Signal Process* 2022;173:109039.
- [38] Oksendal B. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media; 2013.
- [39] Risken H. *Fokker-planck equation*. In: *The Fokker-Planck Equation*. Springer; 1996. p. 63–95.
- [40] Chakraborty S, Adhikari S. Machine learning based digital twin for dynamical systems with multiple time-scales. *Comput Struct* 2021;243:106410.
- [41] Mitchell TJ, Beauchamp JJ. Bayesian variable selection in linear regression. *J Am Stat Assoc* 1988;83(404):1023–32.
- [42] George EI, McCulloch RE. Approaches for bayesian variable selection. *Statistica Sinica* 1997:339–73.
- [43] O'Hara RB, Sillanpää MJ. A review of bayesian variable selection methods: what, how and which. *Bayesian Anal* 2009;4(1):85–117.
- [44] Nayek R, Fuentes R, Worden K, Cross EJ. On spike-and-slab priors for bayesian equation discovery of nonlinear dynamical systems via sparse linear regression. *Mech Syst Signal Process* 2021;161:107986.
- [45] Tripura T, Chakraborty S. A sparse bayesian framework for discovering interpretable nonlinear stochastic dynamical systems with gaussian white noise. *Mech Syst Signal Process* 2023;187:109939.
- [46] Singh M, Fuenmayor E, Hinchy EP, Qiao Y, Murray N, Devine D. Digital twin: Origin to future. *Appl Syst Innovat* 2021;4(2):36.
- [47] Gelernter T. *Mirror worlds: Or the day software puts the universe in a shoebox. How it will happen and what it will mean*. Oxford University Press; 1993.
- [48] Shafto M, Conroy M, Doyle R, Glaessgen E, Kemp C, LeMoigne J, et al. Modeling, simulation, information technology & processing roadmap. *National Aeronaut Space Administ* 2012;32(2012):1–38.
- [49] Rosen R, Von Wichert G, Lo G, Bettenhausen KD. About the importance of autonomy and digital twins for the future of manufacturing. *Ifac-papersonline* 2015;48(3):567–72.
- [50] Xu Y, Sun Y, Liu X, Zheng Y. A digital-twin-assisted fault diagnosis using deep transfer learning. *Ieee Access* 2019;7:19990–9.
- [51] Wang H, Wu Y, Min G, Miao W. A graph neural network-based digital twin for network slicing management. *Ieee Trans Industr Inf* 2020;18(2):1367–76.
- [52] Kloeden PE, Platen E. Higher-order implicit strong numerical schemes for stochastic differential equations. *J Stat Phys* 1992;66(1–2):283–314.
- [53] Tripura T, Gogoi A, Hazra B. An ito-taylor weak 3.0 method for stochastic dynamics of nonlinear systems. *Appl Mathe Modell* 2020.
- [54] Tripura T, Imran M, Hazra B, Chakraborty S. Change of measure enhanced near-exact euler-maruyama scheme for the solution to nonlinear stochastic dynamical systems. *J Eng Mech* 2022;148(6):04022025.
- [55] Tripura T, Chakraborty S. Model-agnostic stochastic model predictive control. *arXiv preprint arXiv:2211.13012*; 2022.