

Geometry-Informed Optimal Parameter Selection for Phase Space Warping with Application to Early Fatigue Damage Detection in 3D-Printed Beams

Hewenxuan Li

University of Rhode Island College of Engineering

David Chelidze (✉ chelidze@uri.edu)

University of Rhode Island <https://orcid.org/0000-0003-0655-1439>

Research Article

Keywords: Phase space warping, hierarchical dynamical systems, chaotic dynamical systems, sensitivity analysis, fatigue damage identification, experimental nonlinear dynamics

Posted Date: October 18th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-978063/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Geometry-Informed Optimal Parameter Selection for Phase Space Warping with Application to Early Fatigue Damage Detection in 3D-Printed Beams

He-Wen-Xuan Li and David Chelidze*

Department of Mechanical, Industrial and Systems Engineering,
University of Rhode Island, Kingston, RI 02881, USA

ABSTRACT

Phase space warping (PSW) methodology reconstructs a non-stationary hidden process from quasi-stationary observable dynamics, where these two coupled dynamical processes have disparate time scales. PSW has been applied to multivariate damage identification and tracking, bio-mechanics, and manifold characterization in nonlinear dynamical systems. However, its theory is not clearly connected to its practice. Furthermore, there is no associated sampling theory or guidelines for optimal parameter selection to estimate the hidden dynamics reliably. This paper focuses on a geometrical interpretation of PSW that coherently bridges its theory and practice by providing the needed theoretical insights and explaining practical constraints. The corresponding algorithm's parameter space is explored to provide reliable and accurate estimates of the PSW function guided by the obtained geometrical properties and insights. Numerical examples of a nonlinear hierarchical dynamical system with various hidden processes and observable dynamics are used to guide the parameter selection for the PSW algorithm. Parameter selection guidelines are obtained through global sensitivity analysis to the estimation accuracy of the simulation results. The established guidelines are used to extract fatigue damage evolution in 3D-printed beams from experimentally obtained vibration data. The obtained results show how the PSW-based fatigue tracking can be used for early fatigue damage detection.

Keywords— Phase space warping, hierarchical dynamical systems, chaotic dynamical systems, sensitivity analysis, fatigue damage identification, experimental nonlinear dynamics

1 Introduction

Phase space warping (PSW) concept describes and quantifies the manifold deformation in the phase space of a *hierarchical dynamical system* (HDS) where the fast- and slow-time subsystems are temporally coupled through a set of system parameters. The PSW-based algorithms are able to extract or reconstruct the *hidden* slow-time phase space dynamics from the *observable* fast-time dynamic measurements [1]. Applications of the PSW ranges from tracking structural degradation [2–4], estimating slow-time multivariate parameter drifts in electromechanical systems [5], identifying physiological fatigue and failure detection in biomechanics [6–8] to identifying invariant manifolds [9]. The PSW applies to systems that share two common features: (1) each system has a pair of state variables that evolve over different time-scales (e.g., the fatigue crack propagates slowly compared to the structural dynamics or oscillations); and (2) only the slowly-varying states are partially observable or non-observable (e.g., the fatigue crack is difficult to be tracked in real-life applications whereas the displacement can be readily measured).

The PSW algorithm is entirely data-driven. It uses non-parametric estimation to reconstruct slow-time dynamics from fast-time measurements. However, two main factors limit the use of PSW in broader applications. First, users unfamiliar with nonlinear time series analysis may find it difficult to understand its underlying theory and algorithmic composition. Second, a set of predetermined inputs to the PSW algorithm considerably affect the accuracy and robustness of the estimation. Nevertheless, no corresponding guidelines for the parameter selection were provided in

*Corresponding author.

E-mail address: chelidze@uri.edu (D.Chelidze).

URL:<http://web.uri.edu/engineering/meet/chelidze> (D.Chelidze).

the previous works [1, 3–5, 10]. Recently, several attempts have been made to improve the damage tracking quality of the PSW method by incorporating more sophisticated algorithms and metrics. In one, the auto-regressive model was used in place of the local linear models to improve the accuracy of the prediction of useful remaining life in rolling bearings [11]. In another, multivariate mutual information entropy and correlation-based multi-modal trend fusion were used to improve the fault tracking quality in rotating machinery when the observation was subjected to noise [12]. However, there are no general guidelines for parameter selection in the available literature. Thus, we still need to obtain insights and strategies to improve the tracking accuracy and robustness by optimizing the set of parameters used in the PSW algorithm with a broader range of applicability. This paper addresses these underlying issues through a geometrical (or graphical) interpretation of the PSW concept and mechanism. This interpretation provides a new perspective on the PSW concept and its application enabling a clear connection between the theory and practical implementation.

We start by examining the parameter space of the PSW algorithm focusing on application to tracking or estimating damage (i.e., the slow-time dynamics) from the displacement measurement of the nonlinear mechanical structure (i.e., the fast-time dynamics). The quality of the PSW-based tracking depends on both the fast- and slow-time dynamic behaviors in the HDS. Thus, the effects of both of these dynamics on parameter selection are examined. Based on the heuristics obtained from the graphical interpretation of the PSW, the parameter selection strategy is investigated by focusing the scope of this study on fatigue damage tracking and early damage detection. Specifically, we study a parameterized nonlinear oscillator with the prescribed slow-time parameter drift, which simulates a nonlinear beam with *known* dynamics of structural degradation. These priorly known slow-time dynamics are used as ground truths to examine the quality of the PSW estimation. The simulated fast-time dynamics models the response of a nonlinear beam to prescribed base excitation. The chosen system exhibits various complexity of the system response, ranging from simple periodic to a more complex chaotic motion. The estimation of the slow-time dynamics is examined by studying two types of damage trajectories: (1) transition from the flat to the ramp damage evolution (piecewise-constant-and-linear function) and (2) the linear-exponential damage accumulation function. The effect of both the fast-time dynamics and the slow-time dynamics on optimal parameter selection are examined under both damage functions. The contribution of the selected parameters to the quality of damage identification are statistically analyzed using the *global sensitivity analysis* (GSA) of a series of *quasi-Monte-Carlo* (QMC) simulations [13, 14]. Aside from applying the GSA in model evaluation, it was recently applied to experimental design for nonlinear system identification purposes [15]. Here, the dimension of the parameter space is reduced using the *factor fixing setting* [13, 16] of the GSA results to help apportion the importance of parameters in providing reliable PSW. A statistical description of the reduced parameter space that yields a reliable damage estimation is formalized into a strategy that guides the parameter selection. Finally, this strategy is applied to a series of field measurements obtained to study damage detection and prognosis in a nonlinear system whose structural integrity is affected by material anisotropy.

The rest of this paper is arranged as follows: Section 2 provides the geometrical interpretation of the basic ideas, assumptions, and techniques used to characterize the PSW and its algorithm; Section 3 describes the PSW algorithm’s parameter space reduction and parameter selection based on GSA; Section 4 discusses the parameter selection strategy and its application to field measurement based on the quasi-Monte-Carlo simulations; and, finally, Section 5 summarizes the identified heuristics and strategy for reliable tracking and detecting of fatigue damage using PSW.

2 Geometry of Phase Space Warping

The HDS models nonlinear systems with multiple temporally-coupled dynamics with separable time scales. PSW was proposed based on an algorithmic procedure [2], and it can be summarized as: PSW characterizes the deformation of the fast-time manifold induced by the slow-time subsystem’s dynamics using the fast-time dynamics sampled from the flow on that manifold. This deformation, or PSW, is estimated through an algorithmic framework that involves:

- the reconstruction of the fast-time *phase space* from the direct measurements of the fast-time dynamics,
- the *inference* of the *reference* fast-time flow as if the slow-time subsystem is static, and
- estimation of the *difference* between the observed fast-time flow and the inferred reference flow using the *phase space warping function* (PSWF).

In the following section, we emphasize the continuous-time formulation of the PSW, the assumptions made, and the approximations required to obtain reliable PSW estimation. Later, practical considerations of the PSW theory and its algorithmic development through sampling are discussed. Aside from the theoretical developments, a comprehensive graphical illustration of the PSW theory’s geometry and its algorithmic considerations is provided.

2.1 PSW Concept for a Continuous-Time Hierarchical Dynamical System

PSW is applicable to the following HDS

$$\begin{aligned}\dot{x} &= f(x, t; \mu(\phi)), \\ \dot{\phi} &= \epsilon g(\phi; \eta(x)), \\ s &= h(x),\end{aligned}\tag{1}$$

where $x \in \mathcal{X} \subset \mathbb{R}^m$ is the state variables that evolves in fast-time and it is directly observable (e.g., the structural dynamics); $t \in \mathcal{T} \subset \mathbb{R}$ is the time variable; $\phi \in \mathcal{D} \subset \mathbb{R}^n$ is the slow-time dynamic variable which is unobservable or partially observable (e.g., the damage evolution); $\mu(\phi) \in \mathbb{R}^k$ is the fast-time parameter vector affected by the slow-time dynamics; $\eta(x) \in \mathbb{R}^q$ is the slow-time loading function; $0 < \epsilon \ll 1$ constant reflects the difference between the fast and slow dynamics (for the rest of the paper, slow/fast system, slow-time/fast-time system, and damage/system response are used interchangeably); and s is the observation of the fast dynamic variable (e.g., measurement function of structural dynamics).

PSW quantifies the deformation (i.e., warping) of the fast-time phase space $\mathcal{S} \triangleq \mathcal{X} \times \mathcal{T}$ described by the vector field $f(\cdot)$ (please, refer to Fig. 1 for the following discussion in this section). This deformation is assumed to be caused by the change in the fast-time system parameters as a function of the slow-time variable, $\mu(\phi)$. More explicitly, we assume that for all $\phi \geq 0$, there exists an invariant, attractive, and finite-dimensional manifold $\mathcal{M}_\phi \subset \mathcal{S}$, that smoothly depends on the change in ϕ . To track this deformation, *the key operation in PSW is to estimate how the fast system would have evolved for a healthy reference system* (i.e., when $\dot{\phi} = 0$ and $\mu = \mu(\phi_R) \triangleq \mu_R$) and we call this the *reference dynamics*. If one knows exactly how the fast dynamics evolves for some fixed reference damage state ($\phi = \phi_R$), there is no need to estimate this reference dynamics. However, in practice, we do not expect to have enough knowledge of the reference dynamics. As a result, one has to model this reference dynamics from a designated *reference portion* of the acquired field measurements s .

To quantify the PSW, we first denote a fast-time flow in phase space \mathcal{S} starting from initial conditions $x(t_0) = x_{t_0}$ and $\phi(t_0) = \phi_{t_0}$, with t_0 an arbitrary initial time, as

$$x(t_0 + t, x_{t_0}; \phi_{t_0}) = F^t(x_{t_0}; \mu(\phi_{t_0})) .\tag{2}$$

Now, we can define the *phase space warping function* (PSWF) as

$$e^t(x_{t_0}, \phi_{t_0}) = F^t(x_{t_0}; \mu(\phi_{t_0})) - F^t(x_{t_0}; \mu_R) .\tag{3}$$

To see the utility of this function in tracking the changes in slow-time variable $\Delta\phi \triangleq \phi_{t_0} - \phi_R$, we expand the flow about the reference value of damage variable

$$F^t(x_{t_0}; \mu(\phi_{t_0})) = F^t(x_{t_0}; \mu_R) + \left. \frac{\partial F^t}{\partial \mu} \frac{\partial \mu}{\partial \phi} \right|_{\phi=\phi_R} \Delta\phi + \mathcal{O}(\|\Delta\phi\|^2) .\tag{4}$$

Provided that the matrix

$$P_R(x_{t_0}) = \left. \frac{\partial F^t}{\partial \mu} \frac{\partial \mu}{\partial \phi} \right|_{\phi=\phi_R}\tag{5}$$

exists and $\mathcal{O}(\|\Delta\phi\|^2)$ is negligible, the PSWF is an *affine transformation* of the change in slow-time variable

$$e^t(x_{t_0}, \phi_{t_0}) \approx P_R(x_{t_0}) \Delta\phi .\tag{6}$$

Therefore, given the reference damage variable ϕ_R and if $m \geq n$, we can write

$$\phi_{t_0} \approx \phi_R + P_R^\dagger(x_{t_0}) e^t(x_{t_0}, \phi_{t_0}) ,\tag{7}$$

where $P_R^\dagger(x_{t_0})$ is the pseudo-inverse of $P_R(x_{t_0})$. If $m < n$, we cannot obtain $P_R^\dagger(x_{t_0})$ and thus we do not have full observability of ϕ_{t_0} if we use only one initial condition x_{t_0} . However, we may recover this observability by considering a set of initial points in the fast-time phase space and estimating P_R^\dagger in the least-squares sense over all those points. Even if matrix in Eq. (5) does not exist due to the nonlinear nature of the fast-time flow, we may still obtain the higher-order observability of the damage variable by including higher-order terms in the flow expansion Eq. (4).

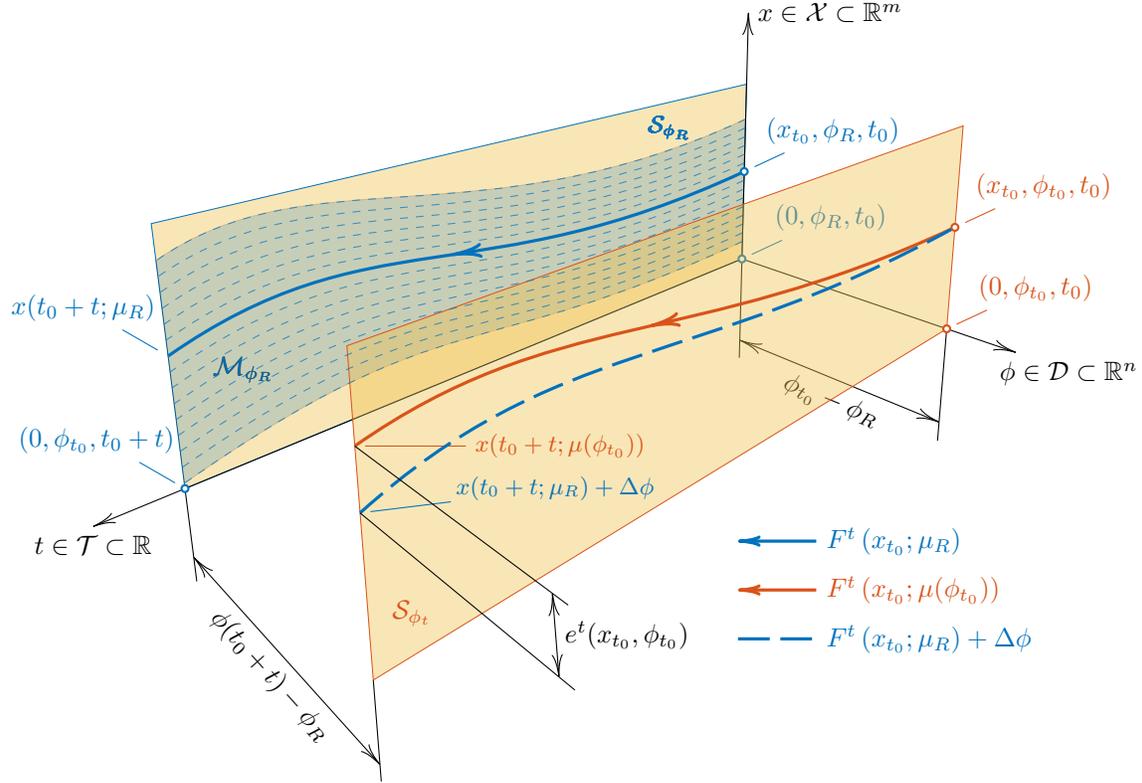


Fig. 1. The figure depicts three-dimensional illustration of the HDS's $(m+n+1)$ -dimensional phase space $\mathcal{X} \times \mathcal{D} \times \mathcal{T}$, where x and ϕ axes are meant to represent m - and n -dimensional coordinate axes, respectively. The two instants of fast-time phase space \mathcal{S}_{ϕ_R} and \mathcal{S}_{ϕ_t} are shown as orange sections through the reference and current values of slow-time variable ϕ_R and ϕ_t , respectively. Two different flows are originated from their initial conditions, where the blue curve depicts the flow in the *reference fast phase space* starting at (x_{t_0}, ϕ_R, t_0) and the red curve shows the flow in the *current fast phase space* starting at $(x_{t_0}, \phi_{t_0}, t_0)$, where t_0 is an arbitrary initial point of time. The image of the reference flow embedded into the current phase space \mathcal{S}_{ϕ_t} is shown by the dashed blue curve. The phase space warping function $e^t(x_{t_0}, \phi_{t_0})$ is defined as difference between the fast flows starting from the same point but for the reference and current values of slow variable ϕ_R and ϕ_t , respectively. In practice, we will not have access to the reference flow starting exactly from (x_{t_0}, ϕ_R, t_0) , and we have to use other available reference fast trajectories (shown by dashed thin blue lines on the light blue reference manifold \mathcal{M}_{ϕ_R}) to estimate it. Please note that \mathcal{M}_{ϕ_R} from the actual system will have some small thickness along the ϕ axis, as damage evolves *ad interim*, which is less than $\mathcal{O}(\epsilon T_R)$.

2.2 Practical Considerations for PSW Implementation

The measurement s in Eq. (1) almost always has a lower dimensionality compared to the fast state vector x . Assuming that $s \in \mathbb{R}$ and using ideas from the embedology [17], we can obtain a reconstructed state variable y topologically equivalent to x using the *delay coordinate embedding* $y = y(s) = y(h(x)) = y(x)$ or

$$y(t) = [s(t); s(t - \tau); s(t - 2\tau); \dots; s(t - (d - 1)\tau)] \in \mathcal{Y} \subset \mathbb{R}^{d \times 1}, \quad (8)$$

where $d \leq 2(m + 1)$ is the dimensionality of the embedded state vector and τ is the appropriate delay time usually a multiple of a sampling period t_s of s .

2.2.1 Autonomous Hierarchical Dynamical System for Reconstructed Fast Phase Space

Now, we interpret the PSW in this reconstructed fast-time phase space. Therefore, we recast our HDS, Eq. (1), as an autonomous dynamical system

$$\begin{aligned} \dot{y} &= \tilde{f}(y; \mu(\phi)) \\ \dot{\phi} &= \epsilon \tilde{g}(\phi; \eta(y)) \end{aligned} \quad (9)$$

131 Since the phase spaces of HDSs described by Eq. (1) and Eq. (9) are diffeomorphic, the dynamics between the two
 132 systems is topologically equivalent. Now, we denote the corresponding flows of the dynamical subsystems as

$$\begin{aligned} y(t_0 + t, y_{t_0}; \mu(\phi_{t_0})) &= F^t(y_{t_0}; \mu(\phi_{t_0})), \\ \phi(t_0 + t, \phi_{t_0}; \eta(y_{t_0})) &= G^t(\phi_{t_0}; \eta(y_{t_0})) \end{aligned} \quad (10)$$

133 where $y_{t_0} = y(t_0)$ and $\phi_{t_0} = \phi(t_0)$. By definition,

$$\dot{\phi}(t_0) = \lim_{\Delta t \rightarrow 0} \frac{\phi(t_0 + \Delta t) - \phi(t_0)}{\Delta t}. \quad (11)$$

134 Thus, the corresponding slow-time dynamics over a short time interval Δt can be approximated as

$$G^{\Delta t}(\phi_{t_0}; \eta(y_{t_0})) \approx \phi(t_0, \phi_{t_0}; \eta(y_{t_0})) + \Delta t \dot{\phi}(t_0, \phi_{t_0}; \eta(y_{t_0})) = \phi_{t_0} + \epsilon \Delta t \tilde{g}(\phi_{t_0}; \eta(y_{t_0})). \quad (12)$$

135 If the same map is applied $k - 1$ more times to the above slow-time state variable, the following is deduced (please
 136 refer to Appendix A for detailed derivations),

$$\phi(t_0 + k\Delta t; \phi_{t_0}) = G^{k\Delta t}(\phi_{t_0}; \eta(y_{t_0})) = \phi_{t_0} + \epsilon k \Delta t \tilde{g}(\phi_{t_0}; \eta(y_{t_0})) + \mathcal{O}((\epsilon \Delta t)^2). \quad (13)$$

137 Therefore, the change in the damage state variable $\|\phi(t_0 + t, \phi_{t_0}) - \phi_{t_0}\| \sim \mathcal{O}(\epsilon t)$.

138 Due to the inherent time-scale separation, PSW focuses on reconstructing *averaged* slow-time damage variable

$$\phi_t \triangleq \frac{1}{T} \int_t^{t+T} \phi(\nu) d\nu \quad (14)$$

139 where T represents the window size chosen for averaging over which fast-time dynamics can be treated as stationary.
 140 Using the result from above, we can see that for any $\nu \in [t, t + T]$, we have

$$\phi(\nu) - \phi_t = \mathcal{O}(\epsilon(\nu - t)) \leq \mathcal{O}(\epsilon T). \quad (15)$$

141 In this framework, we can specify the *reference damage state* of the slow dynamics as

$$\phi_R \triangleq \frac{1}{T_R} \int_0^{T_R} \phi(\nu) d\nu \quad (16)$$

142 where $T_R \geq T$ as we expect the rate of damage evolution to be the slowest initially, we may need more data to build
 143 a suitable reference model. In general, as illustrated in Fig. 2, the total change in the averaged slow-time variable is
 144 of the following order

$$\|\phi_t - \phi_R\| \leq \mathcal{O}\left(\epsilon \left[t + \frac{1}{2}(T_R + T)\right]\right), \quad (17)$$

145 and it indicates that the change in the slow variable is detectable only if the observation time in the fast variable is
 above that order.

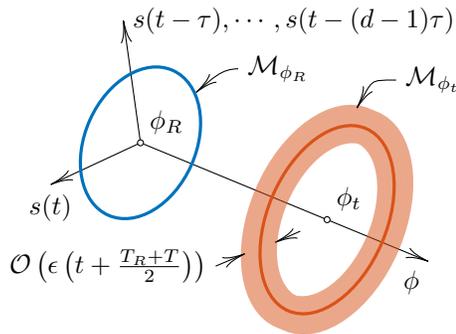


Fig. 2. The close blue curve represents the original manifold \mathcal{M}_{ϕ_R} for the reference slow dynamics. The corresponding observable change after some time is indicated by the red band \mathcal{M}_{ϕ_t} about the original manifold \mathcal{M}_{ϕ_R} .

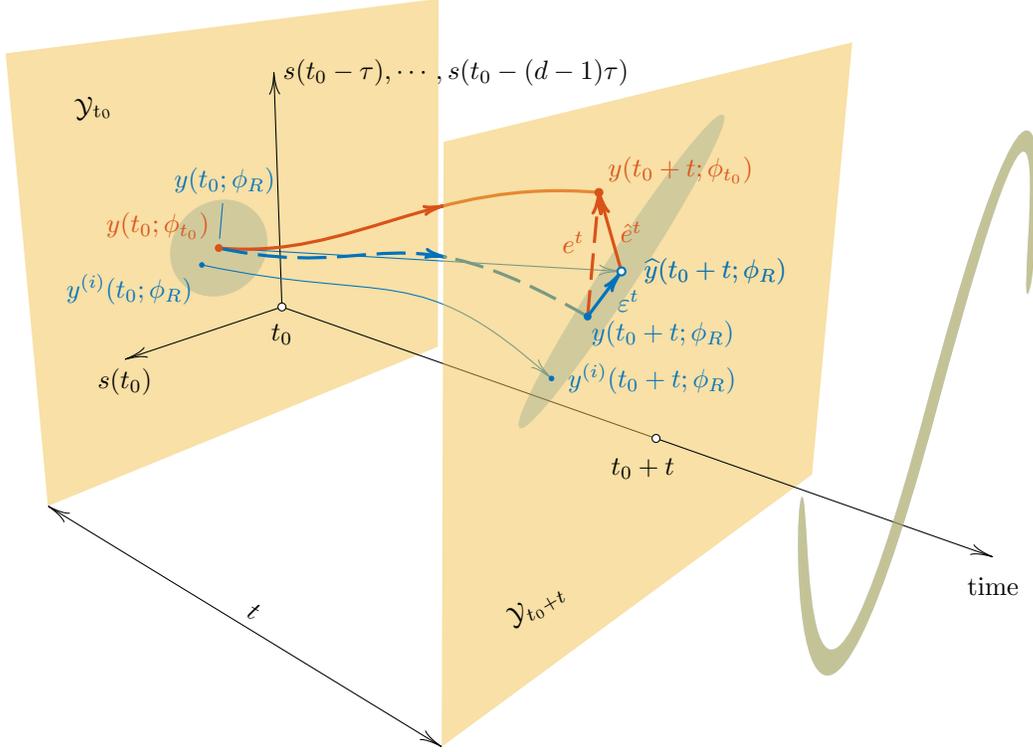


Fig. 3. PSWF e^t (dashed red vector) is estimated by approximating $F^t(y_Q; \mu(\phi_R))$ (dashed blue trajectory) using the tangent space (straight blue line) of the reference trajectories starting from the nearest neighboring reference points $\{y^{(i)}(t_0; \phi_R)\}_{i=1}^r$ of $y_Q = y(t_0; \phi_{t_0})$ (only one of those trajectories is shown as solid thin blue curve). The rest of the nearest neighbor trajectories (NNTs) are encircled by the blue-shaded areas which is mapped into an ellipse over a short period of time, t . When the short-time assumption breaks, the NNTs are affected by the globally nonlinear dynamics (please refer to the distorted s-shaped ellipse), making the prediction unreliable. Base on those reference trajectories we build a LLM to estimate $\hat{y}(t_0+t; \phi_R)$, this introduces the prediction error in the estimate $\varepsilon^k = \hat{y}(t_0+t; \phi_R) - y(t_0+t; \phi_R)$. If this error is small, then the estimated PSWF \hat{e}^k is close to the true value $e^k = \hat{e}^k + \varepsilon^k$.

2.2.2 Estimating PSWF by Approximating Reference Fast-Time Dynamics

If a model of the fast-time dynamics for the initial damage state $\phi(0)$ is not available, we can use the measured reference fast-time dynamics to approximate it over short time interval $t \ll T$ and for any $t_0 \in [0, T_R]$

$$F^t(y; \mu(\phi(0))) = F^t(y; \mu(\phi_R)) + \mathcal{O}\left(\epsilon\left(t + \frac{1}{2}T_R\right)\right) \approx F^t(y; \mu_R). \quad (18)$$

$F^t(y; \mu_R)$ can be viewed as a model for a fast-time subsystem for the reference value of $\phi_R \approx \phi(0)$.

To infer how the fast-time flow would evolve for the reference slow variable, we work in the reconstructed phase space as shown in Fig. 3. In practical applications, we use *local linear models* (LLMs) to estimate PSW-based

153 metrics as long as the following assumptions are satisfied: (1) the prediction time t is relatively short and (2) the
 154 dynamic behaviors of the flow under observation and the reference flow are correlated. The first assumption ensures
 155 that the tangent $\frac{\partial F^t}{\partial y}$ is a sufficient approximation to the actual reference flow, and the second assumption ensures
 156 slow dynamic variable does not significantly affect the fast-time flow during the prediction. Specifically, in the
 157 reconstructed phase space, we consider some arbitrary state vector on the *current* flow, $y_Q = y(t_0; \phi_{t_0})$, and look for
 158 a set of nn state vectors on the *reference* flow, which are similar to y_Q , we find them using $y^{(i)}(t_0; \phi_R) = y^{(i)}(\sigma(y_Q))$
 159 some similarity metric $\sigma(\cdot)$. Then, we approximate the globally nonlinear autonomous map with an LLM-based on
 160 a set of known state vectors $\{y^{(i)}\}_{i=1}^{nn} = \{y^{(i)}(t_0; \phi_R)\}_{i=1}^{nn}$,

$$F^t(y^{(i)}; \mu_R) = F^t(y_Q; \mu_R) + \frac{\partial F^t}{\partial y} \Big|_{y=y_Q} (y^{(i)} - y_Q) + \mathcal{O}(\|y^{(i)} - y_Q\|^2) \quad (19)$$

$$\approx A^t(y_Q) y^{(i)} + b^t(y_Q),$$

161 where $A^t(y_Q) = \frac{\partial F^t}{\partial y} \Big|_{y=y_Q}$ and $b^t(y_Q) = F^t(y^{(i)}; \mu_R) - A^t(y_Q) y^{(i)} = y^{(i)}(t_0 + t; \phi_R) - A^t(y_Q) y^{(i)}(t_0; \phi_R)$ form a
 162 tuple that describes the linearized map. These model parameters can be estimated in the least-squares sense using
 163 all points in the neighborhood $\{y^{(i)}\}_{i=1}^{nn}$. Therefore, the inferred flow can be expressed as an estimate based on this
 164 LLM,

$$\widehat{F}^t(y_Q; \mu_R) = A^t(y_Q) y_Q + b^t(y_Q). \quad (20)$$

165 The PSWF for the current fast-time phase space point $y_Q = y(t_0; \phi_{t_0})$ and the current slow-time state ϕ_{t_0} with
 166 respect to the reference value ϕ_R is given by

$$e^t(\phi_{t_0}, y_Q) = F^t(y_Q; \mu(\phi_{t_0})) - F^t(y_Q; \mu_R). \quad (21)$$

167 where $e^t(\cdot)$ is the *short-time reference model prediction error* or the *damage tracking function* in damage estimation
 168 applications; $F^t(y_Q; \mu(\phi_{t_0}))$ is a reconstructed fast-time phase space point that is t -time ahead of y_Q point, which is
 169 directly available from measurements, $F^t(y_Q; \mu(\phi_{t_0})) = y(t_0 + t; \phi_{t_0})$. However, the corresponding fast-time evolution
 170 that would have occurred for the reference/healthy state of the system $F^t(y_Q; \mu_R)$ needs to be estimated. Using
 171 LLMs or Eq. (20) we estimate the PSW tracking function as

$$\begin{aligned} \widehat{e}^t(\phi_{t_0}, y_Q) &= y(t_0 + t; \phi_{t_0}) - \widehat{y}(t_0 + t; \phi_R) \\ &= y(t_0 + t; \phi_{t_0}) - A^t(y_Q) y_Q - b^t(y_Q), \end{aligned} \quad (22)$$

172 where $A^t(y_Q)$ and $b^t(y_Q)$ are estimated from the nearest neighbors $\{y^{(i)}(t_0; \phi_R)\}_{i=1}^{nn}$ found in the reference fast-time
 173 data record and their images t -times later $\{y^{(i)}(t_0 + t; \phi_R)\}_{i=1}^{nn}$. In particular, we find the best fit parameters in the
 174 least squares sense to the following map

$$y^{(i)}(t_0 + t; \phi_R) \approx A^t(y_Q) y^{(i)}(t_0; \phi_R) + b^t(y_Q), \quad (23)$$

175 for all points in the neighborhood of y_Q . The solution to this in the least squares sense is:

$$[A^t, b^t] = Y_{t_0+t} \tilde{Y}_{t_0}^\dagger, \quad (24)$$

176 where $(\cdot)^\dagger$ is the pseudo-inverse of (\cdot) ,

$$Y_{t_0}^\dagger = Y_{t_0}^T (\tilde{Y}_{t_0} \tilde{Y}_{t_0}^T)^{-1}, \quad Y_t = [y^{(1)}(t; \phi_R), y^{(2)}(t; \phi_R), \dots, y^{(nr)}(t; \phi_R)], \quad \text{and} \quad \tilde{Y}_t = [Y_t; \text{ones}(1, nn)]. \quad (25)$$

177 2.3 Practical Implementation of PSW Algorithm

178 In practice, the estimation quality of the PSW is affected by the sampling and the subsequent data analysis method.
 179 The acquired fast-time series are first preprocessed, including data visualization, data truncation, and filtering.
 180 What is followed hereby is a framework that includes *reference dynamics modeling*, *PSWF estimation*, and *slow-time*
 181 *damage extraction*, which is referred to as the PSW algorithm as shown in Fig. 4.

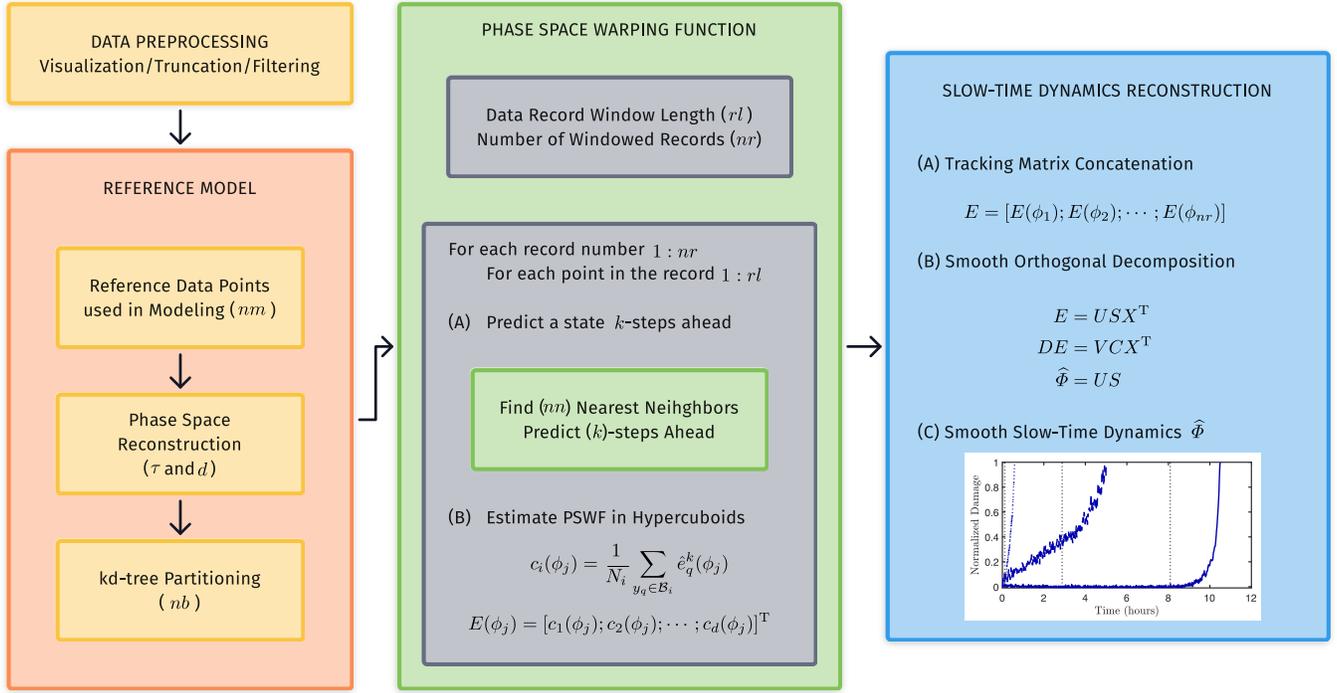


Fig. 4. Phase space warping algorithm in a flowchart. The PSW algorithm comprises of three main parts after the data is preprocessed: (1) building the reference model, (2) estimating the PSWF, and (3) extract the slow-time dynamics using the smooth orthogonal decomposition.

2.3.1 Reconstruction of the Reference Dynamics

We assume the reference fast-time dynamics is measured by uniformly-sampled-scalar-fast-time series $\{s_i\}$ with appropriate sampling period t_s , i.e., $s_i = s(it_s)$. Using this data and Eq. (8), a fast-time phase space trajectory points $\{y_i\} \in \mathbb{R}^d$ ($(m+1) \leq d \leq 2(m+1)$) are obtained. The corresponding nonlinear flows in Eqs. (10) and (12) become a sampled set of two coupled autonomous maps in Eq. (9).

There are two important issues in the *reference phase space* reconstruction: (1) how many phase space points, nm , should be used in the reconstruction? and (2) how to reconstruct this phase space? The answer to the first issue depends on how slow the damage dynamics is compared to the fast-time oscillations, i.e., how small the ϵ is. We usually use some finite time samples nm over which we assume that the inference of the reference fast-time behavior is independent of changes in ϕ . From Eqs. (13) and (17), we know that $\|\phi_{(i+nm)t_s} - \phi_{it_s}\| \sim \mathcal{O}(\epsilon nm t_s)$ and the reference model should not violate the quasi-stationary condition, as long as $nm \ll 1/\epsilon$. Based on the reference data, phase space reconstruction through delay coordinate embedding is accomplished using the proper selection of the *delay time*, τ (now in samples), and the *embedding dimension*, d , with conventional methods [18, 19]; and the reconstructed state vector are expressed as,

$$y_i = [s_i; s_{i-\tau}; \dots; s_{i-(d-1)\tau}] \in \mathbb{R}^{d \times 1} \quad (26)$$

where $\{s_i, s_{i-\tau}, \dots, s_{i-(d-1)\tau}\}_{i=1}^{nm}$ is a set of delayed coordinates of the reconstructed phase space. No sophistication is introduced herein regarding embedding the observed data; this provides insights to parameter selection with the simplest possible algorithmic construct of PSW. Readers are referred to [12] for the implementation of an improved embedding strategy.

2.3.2 Phase Space Warping Function

The fundamental difficulty for a reliable estimation of the PSW is the *inference* of the reference dynamics over time as if slow dynamics stays unaltered in its healthy state. To this end, the reference dynamics is resolved by identifying the reference state vectors that have similar dynamic behavior to the current state of the system (i.e., similar state vectors or phase space points), *in the reference phase space* for $\phi = \phi_R$. PSW algorithm infers $F^k(y_i, \mu_R) \triangleq F^{kt_s}(y_i, \mu_R)$ by characterizing the secant space (approximation to the tangent space) of the reference flow. Following Eqs. (19)

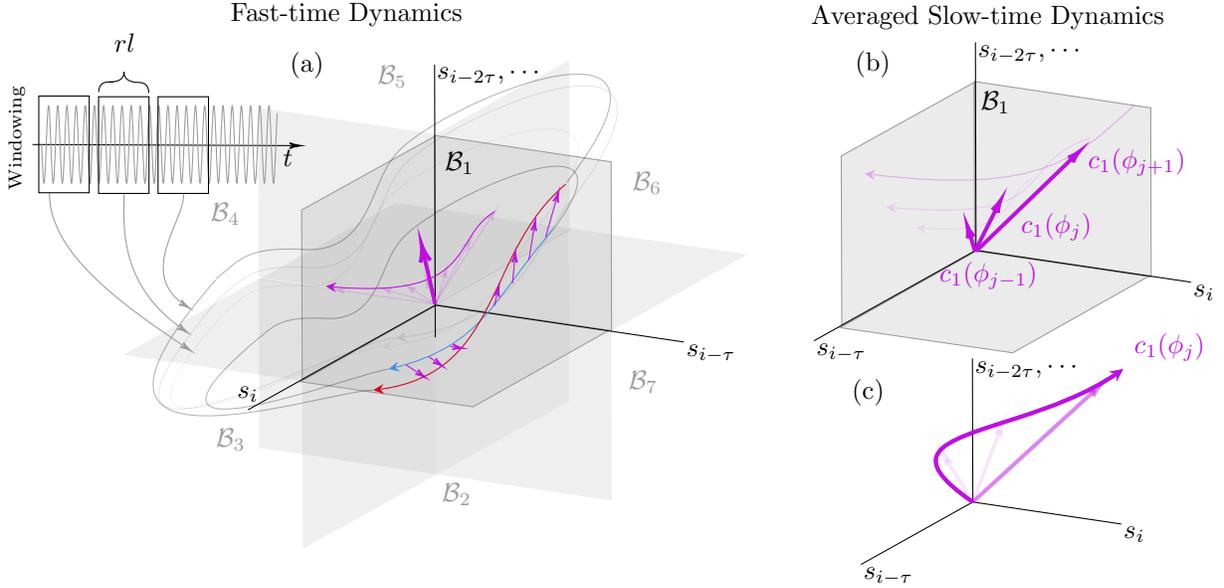


Fig. 6. Geometrical interpretation of the phase space warping feature vector estimation. A k - d tree sectioning of the whole phase space yields subspaces or hypercuboids \mathcal{B}_i . In subplot (a), the red curve inside the \mathcal{B}_1 hypercuboid stands for the corresponding portion of the current or warped flow, and the blue curve depicts the corresponding estimated reference flow. The small magenta vectors show how the PSWF changes over the sampled data inside the \mathcal{B}_1 . For each data record number j , the temporal average to the PSWF (the large magenta vector) is then obtained within each of the hypercuboids using the corresponding portions of rl samples shown as the thin magenta curve reconstructed from the PSWF magenta vectors. The flows in other hypercuboids are shown using grey curves. A close-up view of the first hypercuboid is shown in subplot (b) with only the corresponding PSWFs evolution for each data record number or slow-time stamp $(j - 1)$, j , and $(j + 1)$ and their corresponding averages, showing how the local PSWF evolves in the averaged (slow) time. The locally averaged PSWF in each of hypercuboids \mathcal{B}_i varies over the averaged “slower” time scale represented by the corresponding rl -long data record number j . As the slow-time or record number j varies, we can observe the corresponding time evolution of the time-averaged local PSWF $c_1(\phi_j)$ as shown in subplot (c). For ease of illustration, hypercuboids are shown as equispaced. However, in reality, the volume is partitioned according to the probability of density of the data points.

225 $(NN_i$ and $NN_{i+k})$. Thus, a shorter prediction time (i.e., keeping k small) would provide a more accurate inference
 226 of the reference flow. Longer prediction steps k may push the predictions into a normal space of the reference flow,
 227 resulting in a significant prediction error ε^k .

228 To track multi-dimensional slow-time variables and increase the robustness of the tracking metric, higher-dimensional
 229 vector-valued PSWF was proposed by sectioning the reconstructed phase space into nb subspaces or hypercuboids
 230 $\{\mathcal{B}_i\}_{i=1}^{nb}$ with equal probability of the presence of the reference state vectors [10], see subplot (a) in Fig. 6. This
 231 provides an ensemble of local estimates (or observers) of phase space distortion (i.e., PSWFs) in each of these
 232 hypercuboids,

$$c_i(\phi_j) = \frac{1}{N_i} \sum_{y_q \in \mathcal{B}_i} \hat{e}^k(\phi_j, y_q) = \frac{1}{N_i} \sum_{y_q \in \mathcal{B}_i} \hat{e}_q^k(\phi_j) \quad (29)$$

233 where $c_i(\phi_j) \in \mathbb{R}^{d \times 1}$ and N_i is the number of current points in the hypercuboid \mathcal{B}_i . One should realize that Eq. (29)
 234 involves averaging over the sampled time in each of the hypercuboids.

235 Figure 6 shows the illustration of assembling of the vector-valued PSWF. The whole time series is segmented with
 236 a predetermined window size called the *record length*, rl . For each of the windowed time records numbered j , the
 237 PSWF at each reconstructed sample point is estimated as visualized by the magenta vectors in subplot (a) of Fig. 6.
 238 For large enough rl values, there will be more than one trajectory passes through each of the hypercuboids and
 239 less probability that one of them is empty. For each j , the temporal average of the PSWF is calculated in each of
 240 the hypercuboids, giving one PSWF vector $c_i(\phi_j) \in \mathbb{R}^d$ in the averaged slow time for each of the hypercuboids, see
 241 the large magenta vector located at the center of the phase space in Fig. 6 (a). Keeping rl large will also smooth

242 out the local undulations in the local estimate of the PSWF. To increase the algorithmic robustness, in this paper,
 243 the averaged PSWF values in the empty boxes for the current step will be replaced by their predecessors (i.e.,
 244 last observation carried forward method in [20]). To help extract the slow-time dynamics from the obtained local
 245 averaged PSWFs, they are reshaped into a $N_d = (d \times nb)$ -dimensional damage feature vector $E(\phi_j) \in \mathbb{R}^{1 \times N_d}$ for
 246 each windowed data records $j = 1, \dots, nr$:

$$E(\phi_j) = [c_1(\phi_j); c_2(\phi_j); \dots; c_d(\phi_j)]^T \quad (30)$$

and then concatenated into a *tracking matrix*, $E \in \mathbb{R}^{nr \times N_d}$

$$E = [E(\phi_1); E(\phi_2); \dots; E(\phi_{nr})], \quad (31)$$

which can also be written as

$$E = \begin{bmatrix} | & | & \cdots & | \\ E_1(\phi) & E_2(\phi) & \cdots & E_{N_d}(\phi) \\ | & | & \cdots & | \end{bmatrix}, \quad (32)$$

247 where each column $E_i, i = 1, 2, \dots, N_d$ represents the projection of the variation of the locally averaged PSWF onto
 248 each independent local embedding dimension in Fig. 6 (c).

249 Reference [10] claims this phase space sectioning suppresses spurious fluctuations that are not related to the slow-time
 250 dynamics without giving any explanations. Here, we address this by providing a supporting explanation as to why
 251 the tracking accuracy favors partitioning. Consider a system whose reconstructed phase space is two-dimensional,
 252 and the slow dynamics either expands the flow globally or warps the flow locally, see Fig. 7. When no partitions
 253 exist, a global observer tends to average out the effect of the uniform expansion and predicts no change in the mean,
 254 see case (a1) in Fig. 7. Moreover, the global observer tends to obscure the global expansion and the local warping,
 255 ending up with a global observation vector whose direction is primarily affected by the local warping, see case (b1) in
 256 Fig. 7. Finally, the global observer may be insensitive to the nature of local warping, as shown in case (c1) in Fig. 7.
 257 In contrast to a global observer, k - d partitioning provides nb number of locally averaged PSWFs or an *ensemble of*
 258 *local observations* which enrich the spatial description of the PSW. As shown in cases (a2), (b2), and (c2), splitting
 259 the corresponding two-dimensional space into four subspaces (quadrants) allows for clear differentiation of local and
 260 global expansion or warping. In particular, partitioning helps describe and delineate spatially hierarchical warping
 261 in the phase space through a “voting mechanism” (the local observers infer the big picture by discussing what they
 262 saw through a conversation) provided by the slow-time dynamics reconstruction process. On the other hand, when
 263 the PSW results in large deformation of the flow or when a translation of the flow is dominant, larger nb may result
 264 in more empty subspaces, which will cause rank-deficiency in the error matrix E .

265 After the error matrix is obtained, multivariate data analysis is applied to obtain a desirable representative of the
 266 estimated damage-time history. In this paper, the *smooth orthogonal decomposition* (SOD) is used to extract the
 267 smoothest temporal variations from the feature vectors $E(\phi_j)$ [2]. For the details on SOD and its properties please
 268 see Refs. [2, 3, 21]. Specifically, SOD uses a generalized singular value decomposition of the matrices E and its time
 269 derivative DE , usually estimated using a finite difference operator to get:

$$\begin{aligned} E &= USX^T, \\ DE &= VCX^T, \end{aligned} \quad (33)$$

270 where U and V are unitary matrices, D is some finite difference operator in time, S and C are diagonal, and X is
 271 a full rank matrix composed of *smooth modes*. Then, the corresponding *smooth orthogonal coordinates* (SOCs) are
 272 obtained using:

$$\hat{\Phi} = US, \quad (34)$$

273 where $\hat{\Phi} \in \mathbb{R}^{nr \times N_d}$ is used to estimate the hidden slow dynamics trajectory by considering only the smoothest
 274 columns of $\hat{\Phi}$ indicated by the magnitudes of the *smooth orthogonal values* (SOVs) given by term by term division
 275 of $\text{diag}(S^T S)$ and $\text{diag}(C^T C)$.

276 3 Practical Parameter Selection based on Global Sensitivity Analysis

277 The PSW algorithm is a non-parametric estimator for the slow-time dynamics in the HDS that requires numerous
 278 input parameters that determine the algorithm’s performance. However, the parameter selection was not explicitly

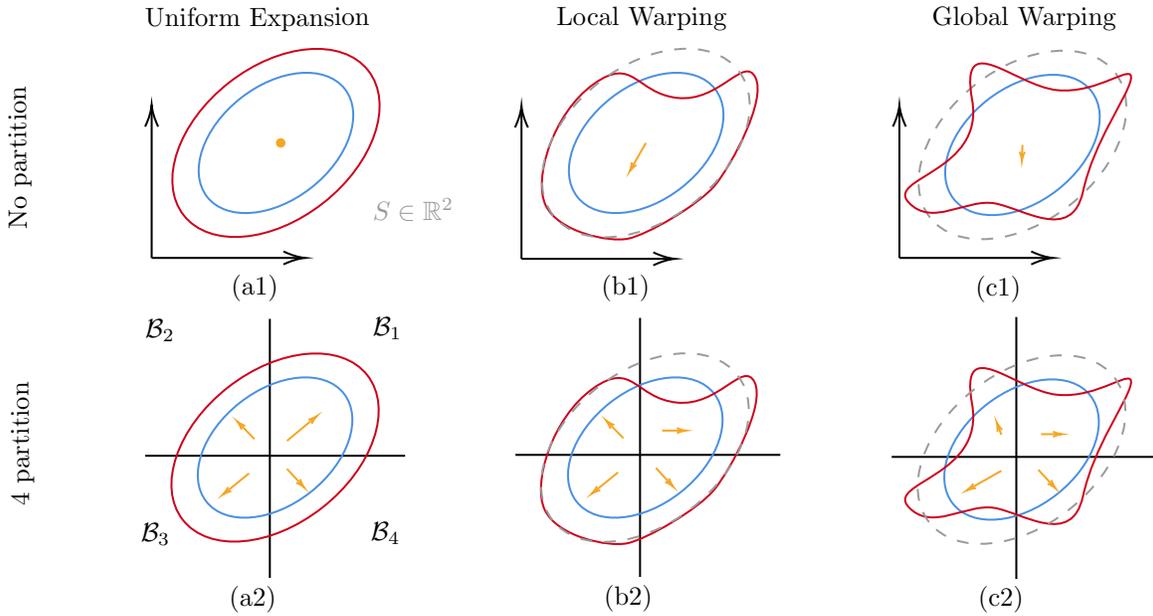


Fig. 7. Graphical Illustration of the effect of partitioning the reconstructed phase space into subspaces on the accuracy of the PSW estimation. For clarity, the phase space is confined to two dimensions. Blue curves indicate the reference manifold, and the red curve indicates the warped manifold.

279 explained or discussed in the available literature apart from providing some basic ranges for the selection [1, 3, 4, 10].
 280 In this section, we use geometrical heuristics and variance-based global sensitivity analysis to guide the parameter
 281 selection strategy focusing on the application to fatigue damage identification from vibration data of a nonlinear
 282 system.

283 3.1 Variance-Based Global Sensitivity Analysis

284 For an accurate estimation, *sensitivity analysis* is applied to the PSW procedure to determine a practical scheme of
 285 parameter selection. Given an existing model, sensitivity analysis determines how the uncertainty in the output of
 286 the model is apportioned to the uncertainties in the model inputs [13]. In contrast to *local sensitivity analysis* (LSA),
 287 the *global sensitivity analysis* (GSA) is used here as it can sort out the parameters in terms of their contribution
 288 to the variance of the output. In particular, the GSA helps one to prioritize (or skip) the selection of the parameters
 289 (i.e., the *factor prioritization*). Given a square integrable function $f(\cdot)$ —in this context, the PSW algorithm—over
 290 a m -dimensional parameter space (usually represented by a m -dimensional hypercuboid),

$$\Omega^k = (X | \min(X_i) \leq X_i \leq \max(X_i); i = 1, \dots, m), \quad (35)$$

291 where the random variable X represents the input variables to the PSW algorithm, and for each of the components
 292 X_i , the domain is designated by some type of probability distributions. Sobol proposed the *higher-order model*
 293 *representation* which decompose the function $f(\cdot)$ into,

$$f = f_0 + \sum_i f_i + \sum_i \sum_{j>i} f_{ij} + \dots + f_{12\dots k}, \quad (36)$$

294 in which each individual term is also square integrable over the domain of the functions. Based on the higher-order
 295 model representation, the model can be treated as the function described in Eq. (36). To characterize the sensitivity,
 296 one first square integrates every term in Eq. (36) and by implementing the notation of the variances, one can obtain
 297 the ANOVA-HDMR (*analysis of variance-higher dimensional model representation*) decomposition,

$$V(Y) = \sum_i V_i + \sum_i \sum_{j>1} V_{ij} + \dots + V_{12\dots k} \quad (37)$$

298 where $V(\cdot)$ now is the variance estimator, $V_i \triangleq V(f_i) = V[E(Y|X_i)]$ is the variance to the univariate function f_i ,
 299 $V_{ij} \triangleq V(f_{ij}) = V[E(Y|X_i, X_j)]$ is the variance to the bivariate function f_{ij} , and so do the rest of the components.

300 The decomposition of the sensitivity can be obtained by normalizing the equation by the total variance of the output,

$$\sum_i S_i + \sum_i \sum_{j>i} S_{ij} + \sum_i \sum_{j>i} \sum_{k>j} S_{ijk} + \dots + S_{123\dots m} = 1, \quad (38)$$

301 where $S_i \triangleq \frac{V[E(Y|X_i)]}{V(Y)}$ represents the sensitivity of the first order (or the univariate part of the function decom-
 302 position), which indicates how much variance is contributed by fixing only one variable in the input space; the
 303 second-order sensitivity indices indicate how much output variance will change if two variables are fixed (not in-
 304 cluding the *marginal effect* contributed from the univariate cases); sensitivity indices with higher order are defined
 305 in the same sense, and the total sensitivity should add up to unity. Therefore, if the summation of the first-order
 306 sensitivity indices does not equal to one, i.e., $\sum_i S_i \neq 1$, *interaction* between various inputs exists. However, a valid
 307 estimation of the higher-order interaction sensitivity terms requires a large amount of samples of the output; this
 308 is simply because the number of samples in the parameter subspace shrinks exponentially after constraining more
 309 than one input variable. It is suggested to use *first order sensitivity index* and *total sensitivity index* to examine the
 310 contribution of each variable and their interaction with all other possible variables, respectively [16].

311 3.2 Heuristics-Based Hypotheses on Parameter Selection

312 The conduct of a GSA includes (1) identifying the goal of the GSA, (2) deciding which input variables to be considered,
 313 (3) choosing proper distribution functions for each input variable, (4) executing the GSA, and (5) interpreting the
 314 results. To make the formal sensitivity analysis more tangible, the *expert’s opinion* or the appropriate *heuristics*
 315 is a good start for steps (2) and (3); this helps minimize the computational costs of the GSA [13]. Based on the
 316 previous studies using PSW, the parameter selection during the phase space reconstruction is standardized by existing
 317 methodologies in *nonlinear time series analysis* [18, 19]. In this paper, the rest of the parameters are categorized
 318 into two groups, namely, the ones tracking the *short-time phase-space dynamics* (nm , nb , nn , and k) and the ones
 319 related to *long-term temporal averaging* (rl and nr). The first group is intimately related to the *reference model*
 320 in the reconstructed phase space during the construction of the LLMs. In contrast, the second group determines
 321 how many data points are used for averaging the *fast-time dynamics* to track *slow-time dynamics*. The following
 322 *heuristics-based hypotheses* on these parameters are made:

- 323 1. nm , the number of data points used in the reference model establishes the baseline for the rest of the procedure.
 324 Therefore, we treat it as the most critical parameter to obtain an accurate estimation of the slow-time dynamics;
 325 it determines the accuracy of the reference model in terms of its *quasi-stationarity*. In particular, as discussed
 326 in section 2.2.1, nm should provide enough data for the reference model to capture the local dynamics in the
 327 reconstructed phase space but not be so large as to include dynamics related to the slow-time dynamics (thus,
 328 breaking the quasi-stationarity assumption).
- 329 2. nb , the number of hypercuboids used in the reference model’s phase space, is equally important to nm . This
 330 partitioning of the phase spaces helps delineate multi-spatial-scale PSW and eliminates spurious fluctuations
 331 predominated by the complexity of the fast-time dynamics.
- 332 3. nn , the number of nearest neighbors used to construct the reference LLMs has its effects on the model accuracy
 333 due to two main reasons: (1) when nn asymptotically approaches zero, the linear regression in Eq. (28)
 334 is unstable during the matrix inversion; and (2) very large nn causes LLM to lose its “locality” and when
 335 $nn \rightarrow nm$, the *local reference dynamics* becomes *global*, which a linear model cannot represent.
- 336 4. k , the number of steps (sampled time intervals) used in Eq. (28) determines the temporal distance between the
 337 two clouds of points, namely, the blue lines NN_i and NN_{i+k} on the reference manifold in Fig. 5 (or, equivalently,
 338 the circles and the ellipse in Fig. 3). When k is too large, the local affinity of the transformation is violated
 339 due to the nonlinear nature of the flow. This parameter also dictates the “locality” of the LLMs, e.g., if the k
 340 is too large, the mapping may lose its temporal correlation, causing larger error between the PSWF and the
 341 real PSW.
- 342 5. rl and nr are the record length and number of records used for temporal averaging, respectively. We consider
 343 these factors less critical due to their irrelevance to the reconstruction of the reference dynamics. We propose
 344 to treat rl as the window size and nr as the control of overlapping during the temporal averaging; that is
 345 when $rl \times nr > \text{length}(y)$, there exists overlapping between two windows, see Fig. 6. In general, nr should be
 346 larger than $d \times nb$ to allow for the full-rank SOD, and rl should not be too small as to yield too many empty
 347 hypercuboids in Eq. (29), which may cause the rank-deficiency of E .

348 In this paper, only the first four parameters are included in the GSA as rationalized by the above heuristics. Since no
 349 prior knowledge is available on the distribution of the four factors (as random input variables), uniform distribution
 350 for each parameter is assumed. However, since the previous studies proposed valuable ranges for the parameters to
 351 consider [1, 3, 4, 10], the distributions are designated as follows

$$\begin{aligned}
 nm &\sim \mathcal{U}(2^5, 2^{13}) \\
 nb &\sim \mathcal{U}(2^0, 2^5) \\
 nn &\sim \mathcal{U}(2^2, 2^5) \\
 k &\sim \mathcal{U}(1, 12).
 \end{aligned}
 \tag{39}$$

352 where $\mathcal{U}(\cdot, \cdot)$ represents the uniform distribution whose lower bound and upper bound are separated by a comma.
 353 All the samples drawn from these distributions are of integer values.

3.3 Practical Parameter Selection through Quasi-Monte-Carlo Simulations

355 Statistical examination of the GSA and visual inspection of the results guides the practical parameter selection
 356 strategy for an accurate estimation of slow-time dynamics (i.e., the damage evolution). The estimation of the PSWF
 357 is highly dependent on the nature of the fast and slow dynamics. Here, we use simulations of a nonlinear mechanical
 358 oscillator with prescribed damage to gain insight into parameter selection through GSA. These simulations are simple
 359 to implement and easy to understand but reflect the complexity of the nonlinear system response. In particular, the
 360 following issues are examined:

- 361 1. The effect of reference data quasi-stationarity: damage dynamics of the ramp type with various lengths of
 362 damage initiation are investigated;
- 363 2. The effect of underlying damage dynamics on the parameter selection: damage of the linear-exponential growth
 364 type with various growth rates are considered;
- 365 3. The effect of the nature of the fast-time response on the parameter selection: a comparative illustration among
 366 (i) period-1 responses, (ii) period-3 responses, and (iii) chaotic responses is conducted.

367 The ramp and linear-exponential damage function can be expressed explicitly as,

$$\phi = \phi_e \Theta \left(t - \frac{T_e}{r_\phi} \right) \circ \left[\frac{a_1}{T_e} \left(t - \frac{T_e}{r_\phi} \right) + a_2 \exp \left(a_3 \frac{t}{T_e} \right) \right],
 \tag{40}$$

368 where ϕ_e is the critical damage value when the simulations stop; $\Theta(\cdot)$ is the Heaviside function; \circ indicates a
 369 Hadamard product; T_e is the total time to failure; r_ϕ represents the damage initiation time as a ratio to the total
 370 time to failure; t is the time variable; and a_i , $i = 1, \dots, 3$ are the parameters that control the proportionality of the
 371 linear and exponential growth as well as the intensity of the exponential growth. This damage model mimics the
 372 Paris'-Law-type crack propagation in linear elastic fracture mechanics of solids [22]. The actual damage evolutions
 373 used in the simulations are shown in Fig. 8. When damage of the ramp type is desired, a_2 is set to zero, leaving only
 374 the linear part of the damage dynamics,

$$\phi_R = \phi_e \Theta \left(t - \frac{T_e}{r_\phi} \right) \circ \left[\frac{a_1}{T_e} \left(t - \frac{T_e}{r_\phi} \right) \right];
 \tag{41}$$

375 and in the case of linear-exponential damage accumulation, one can set the damage initiation ratio to infinity (i.e.,
 376 $r_\phi \sim \infty$) and obtain,

$$\phi_{LE} = \phi_e \left[\frac{a_1}{T_e} t + a_2 \exp \left(a_3 \frac{t}{T_e} \right) \right].
 \tag{42}$$

377 To examine the parameter selection under various measurable fast-time dynamics, a Duffing oscillator with deteriorating
 378 stiffness is considered,

$$\ddot{x} + c\dot{x} + [\kappa(1 - \phi) + \alpha]x + \beta x^3 = f(t),
 \tag{43}$$

379 which is *mass normalized*. In the Eq. (43): c is the damping coefficient; κ is the stiffness when there is no damage;
 380 $\kappa(1 - \phi)$ is the effective stiffness and is a function of the damage variable $\phi \in [0, 1]$; α and β are the linear portion

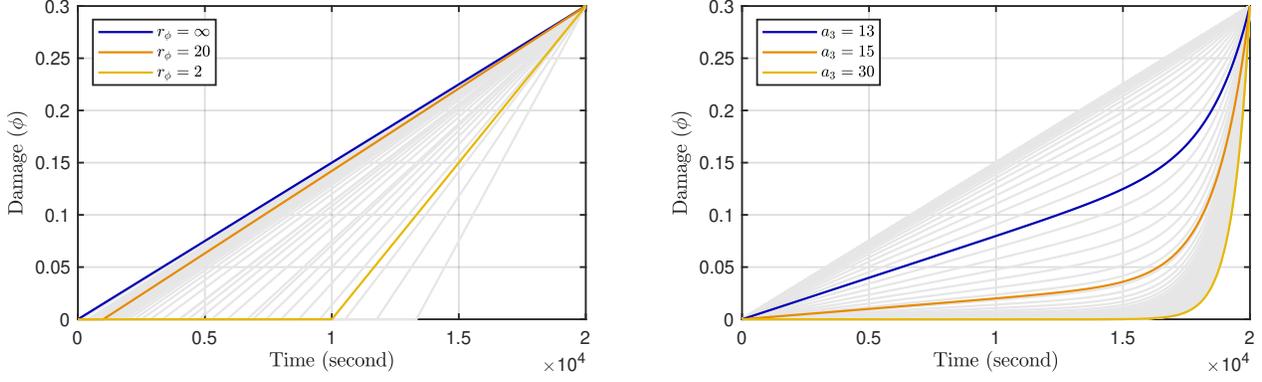


Fig. 8. Prescribed damage-time histories used in the quasi-Monte-Carlo simulations. Left: the damage of the ramp type (Eq. (41)) where three samples are designated which correspond to damage initiation at (1) the beginning of the life span ($r_\phi = \infty$), (2) 5% of the life span ($r_\phi = 20$), and (3) 50% of the life span ($r_\phi = 2$). Right: the damage of the linear-exponential type (Eq. (42)) where samples with three distinct damage growth rates are designated (i.e., $a_3 = 13$, $a_3 = 15$, and $a_3 = 30$, respectively). Grey lines in the plots indicate other possible damage evolutions of the same type and the colored functions are selected from the family of damage curves.

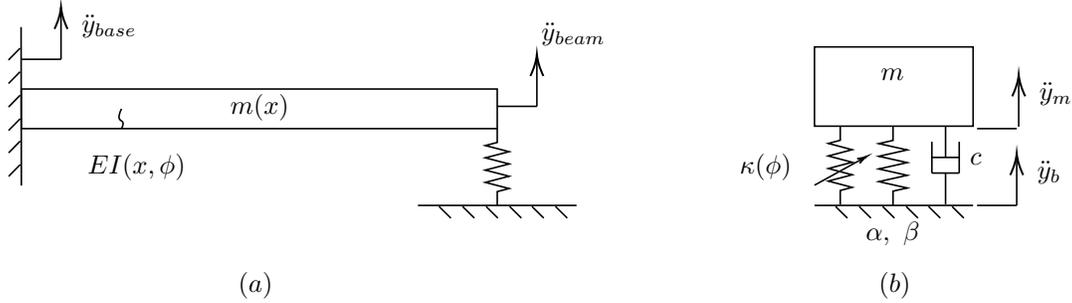


Fig. 9. Sketches of the mechanical system considered in the quasi-Monte-Carlo simulations. (a) a distributed parameter system (a cantilevered beam) with initial damage and nonlinear stiffness on one of the boundaries; (b) a simplified version of (a) where the mass is lumped to m , the damage is modeled by the variable stiffness $\kappa(\phi)$.

381 and the nonlinear portion of the nonlinear stiffness, respectively; and $f(t)$ describes the external excitation. One
 382 possible configuration of the assumed model is given in Fig. 9, if the first mode of vibration dominates the response.
 383

384 Since the GSA is computationally intensive, we select the numerical examples using the above *heuristics-based*
 385 *hypotheses*. Therefore, to examine the effect of the reference data quasi-stationarity on parameter selection, we
 386 use the ramp function with different periods of damage initiation (i.e., different time spans when damage does not
 387 propagate), see left subplot in Fig. 8. To further examine the effect of various rates of damage propagation on
 388 the parameter selection, we use the linear-exponential model cases shown in the right subplot of Fig. 8. Three
 389 types of responses from the nonlinear oscillator are considered through variations of the final damage level, system
 390 parameters, and the excitation magnitude to address the effects on fast-time dynamics. The parameters that control
 391 the system response are summarized in Table 1. These parameters ensure that no bifurcations happen during the
 392 damage evolution. The sample responses for the ramp damage function where damage initiates from 50% of the
 393 total life span are illustrated in Fig. 10, where the fast-time responses indicate that it is difficult to infer what the
 394 underlying damage dynamics is, especially when the response is chaotic.

395 One of the challenges of the GSA is establishing a reasonable *number of samples* needed from the prescribed sample
 396 space. Here, we used a trial-and-error approach to obtain the final sample size of 10^4 , where we observed no significant
 397 variations from the obtained sensitivity indices. This selection also ensures that no excessive samples are drawn from
 398 the predefined sample space Ω , where three of four parameters are sparse. The samples of the input parameters are
 399 drawn through a *Sobol quasi-random set* and evaluated through the PSW algorithm with a fixed set of temporal

Response Type	Parameter Values							
	System				Loading		Damage	
	κ	c	α	β	A	ω	ϕ_{t_0}	ϕ_e
Period-1	0.5	0.3	0	0	0.37	1.3	0	0.3
Period-3	0.5	0.3	-1.5	1	0.37	1.3	0	0.3
Chaotic	0.5	0.3	-1.5	1	0.32	1.3	0	0.3

Table 1: Parameter used to produce system responses with various complexity

400 averaging parameters $rl = 2^6$ and nr ensuring 0% overlapping. The first dominant SOC from each PSW estimation,
 401 obtained from Eq. (34) is treated as the estimate of the damage-time history and is scaled to its original range.
 402 Eventually, the prediction quality is quantified through the *mean absolute error* between the estimated damage
 403 and the actual damage, $\sum_{i=1}^{N_\phi} |\hat{\phi}_i - \phi_i| N_\phi^{-1}$. Some of the estimated damage time series obtained from the quasi-
 404 Monte-Carlo simulations are shown in Fig. 11. These plots indicate that the PSW damage estimates can track the
 405 underlying ground truth given a set of optimal parameters, regardless of the underlying type of responses and the
 406 damage dynamics.

3.4 Parameter Space Reduction based on the GSA Results

408 Here, we focused on the first-order sensitivity indices to illustrate the effects of parameters used in estimating the
 409 PSWF on the mean absolute error in damage tracking. The number of data used for constructing the reference phase
 410 space nm has a higher weight when the fast-time dynamics is periodic (see the nm bar cluster in the second row of
 411 Fig. 12) in contrast to the chaotic response. The significance of its effect decreases as the system response gets more
 412 complex, i.e., as fast-time response varies from period-1 to chaotic (see the nm bars in the first subplot of Fig. 12).
 413 However, nm loses its significance when the underlying damage is of linear-exponential type (bottom row plots in
 414 Fig. 12). The number of hypercuboids used for the PSWF (nb) shows increasing significance as the complexity of
 415 the system response increases. This feature is shared by both slow-time dynamics considered in this paper, which
 416 can be observed from the increment of sensitivity of the nb bars illustrated in the second row of Figs. 12–13. The
 417 contribution in the prediction error variance by the number of nearest neighbors (nn) is similar to the effect of nb
 418 for the ramp damage function. It can be seen from the first row of Fig. 12 that nn 's effect on prediction error
 419 increases as the quasi-stationarity breaks down. Moreover, its weight gets more prominent when the complexity
 420 of the fast-time response increases, see Fig. 12. However, the sensitivity-complexity correlation breaks down when
 421 the quasi-stationarity assumption is violated (in the linear-exponential damage case); the prediction error variation

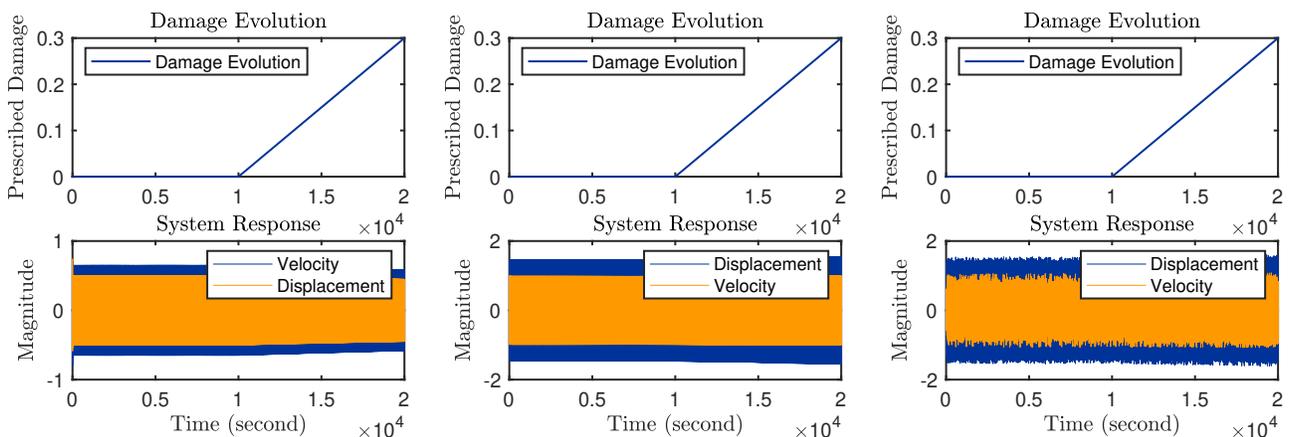


Fig. 10. Plots depicting the responses due to prescribed damage evolution. Upper: damage-time histories of the ramp type with the damage initiation at 50% of the total life span; lower: system responses with the prescribed damage. Left: system response is of period-1 type; middle: system response is of period-3 type; right: system response is of chaotic type.

422 significantly drops when the initial crack growth rate is low, see the decrements in the sensitivity of nn in the first
 423 row of Fig. 13. Among all the parameters considered in the sensitivity study, k or $step$ has the least contribution to
 424 the changes in the prediction error variance, under the considered range of k . Since the prediction error is insensitive
 425 to k , this parameter may be ruled out from further consideration. In addition, we are free to choose any value of k ,
 426 within the considered range, without significantly affecting the prediction quality under the considered distribution.
 427 Therefore, we reduce the final size of the parameter space to three, spanned by nm , nb , and nn in the sense of *factor*
 428 *fixing setting* [13]. One example of the visual inspection of the reduced parameter space from which the sampled
 429 parameters yield under five-percentile error is shown in Fig. 14. The figure shows the clustering of the parameter
 430 triplets in the reduced parameter space as the complexity of the response varies. Moreover, the two-dimensional
 431 marginal statistics show very distinct distributions for each response. For clarity, the rest of the analysis will focus
 432 on the moment description of the one-dimensional marginal probability density functions from all the cases.

433 3.5 Parameter Normalization and Parameter Selection Guideline

434 The probability mass functions show the effect of both the fast- and slow-time dynamics on the parameter selection as
 435 observed in the reduced parameter space. Table 4 summarizes the statistics of the marginal probability distribution
 436 of each parameter in the reduced three-dimensional parameter space that yields within 1-decile prediction error
 437 among all the considered cases with the ramp and LE damage function.

438 This prediction accuracy region is selected based on the visual inspection of the damage prediction quality from
 439 the simulation samples, see Fig. 11. To eliminate the effects from sampling and embedding, the statistics of the
 440 first three parameters in Table 4 are *embedology-parameter-normalized*; temporally, nm is normalized by the delayed
 441 sample time, nm/τ ; spatially, nb and nn are normalized by the embedding dimension d , these normalized statistics

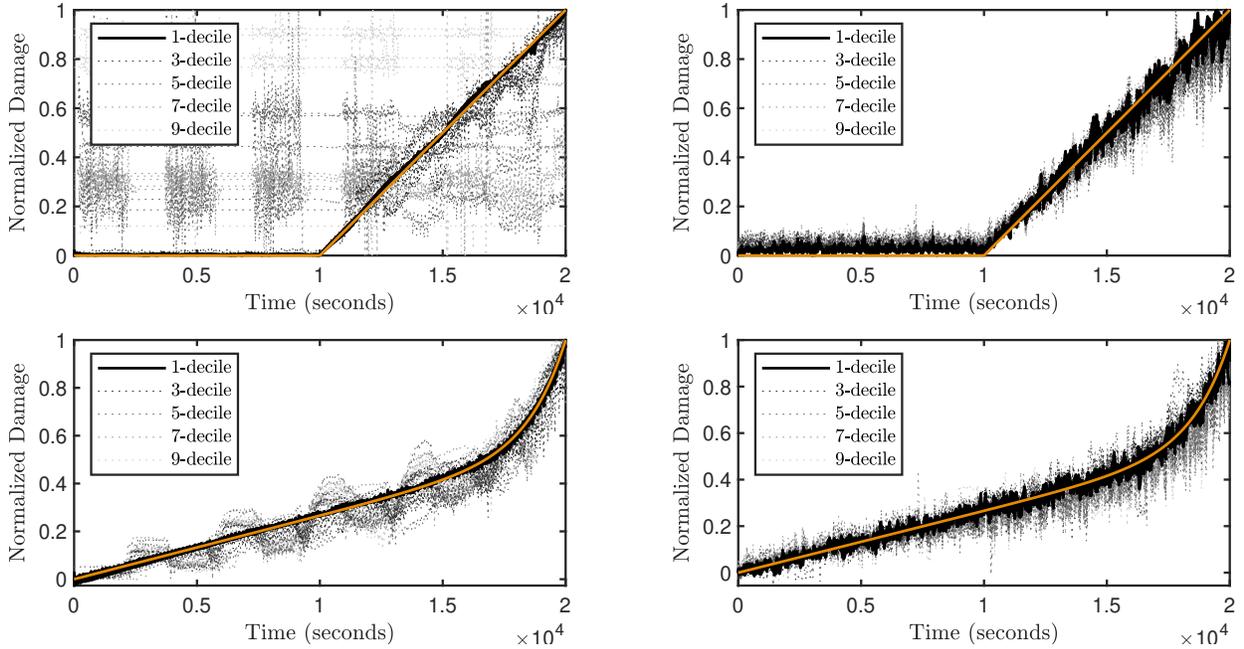


Fig. 11. Sampled damage estimation results obtained from the quasi-Monte-Carlo simulations. The results indicate that the PSW-estimated damage dynamics match the true damage dynamics as long as the optimal parameters are used, regardless of the type of the fast- and slow-time dynamics. The results are presented in terms of their decile statistics of the mean absolute error between the estimation and the ground truth and only the first ten estimates of each decile group are plotted. Left: the simulations that correspond to the harmonic fast-time responses; right: the simulations that correspond to the chaotic fast-time responses; upper: the slow-time dynamics of the ramp type (Eq. (41)) with 50% of the life span being static ($r_\phi = 2$); lower: the damage of the linear-exponential type (Eq. (42)) with $a_3 = 13$. Black-solid lines correspond to the damage estimates that are within 1-decile prediction error. Grey-dashed lines indicate damage estimates with higher estimation errors. The yellow curves are the ground truths among the illustrated cases.

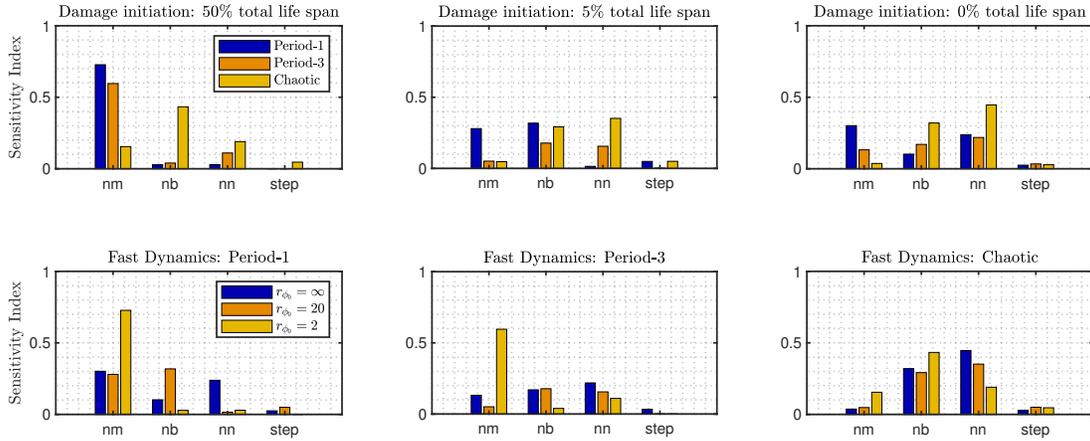


Fig. 12. First-order global sensitivity indices obtained from the quasi-Monte-Carlo simulation using the ramp damage function illustrated in the left subplot of Fig. 8. First row: the sensitivity indices compared across various quasi-stationarity during the damage propagation (slow-time dynamics), from left to right, the damage-free portion of the damage propagation decrease from 50% to 0; second row: the sensitivity indices compared across different types of fast-time responses, from left to right, the complexity of the fast-time responses increase.

442 are rounded to their ceiling integer and concluded in Table 2. The normalization is conducted to eliminate the
 443 temporal irregularity and the spatial complexity of the data. For example, it helps reduce the effects induced by
 444 various sampling frequencies across various trials of experiments; it also helps mitigate the significant discrepancies in
 445 *nb* induced by the response complexity. For a more detailed graphical representation of the PDFs that are discussed
 446 in this section, please refer to Figs. 17–18 in the Appendix B. From the Table 2,

447 1. for the parameter selection based on the *fast-time* dynamics:

448 (a) As the complexity of the system response increases, high-quality predictions require more sample vectors

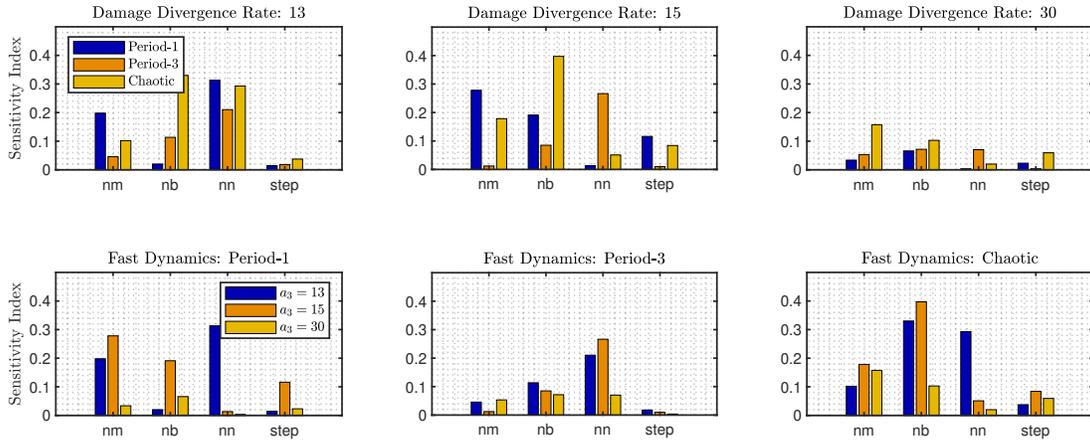


Fig. 13. First-order global sensitivity indices obtained from the quasi-Monte-Carlo simulation using the linear-exponential damage function illustrated in the right subplot of Fig. 8. First row: the sensitivity indices compared across various exponential damage growth rate during the damage propagation (slow-time dynamics), from left to right, the damage-free portion of the damage propagation increase from 13 to 30; second row: the sensitivity indices compared across different types of fast-time responses, from left to right, the complexity of the fast-time responses increase.

449 to be included in the reference model, i.e., a larger number of nm is preferred. Some cases may seem
 450 like outliers, e.g., the P-1 cases in the LE damage group; however, they have a large standard deviation,
 451 indicating the tracking quality is less sensitive to the selection of nm , which does not contradict the general
 452 conclusion made herein.

453 (b) For the normalized statistics, the dependence of the optimal nb on the complexity of the response is
 454 nullified when compared to the original statistics in Table 4. Generally, the normalized nb does not
 455 correlate with the complexity of the response, and one should use less nb to reduce computation time, if
 456 possible. However, when considering non-normalized nb , a significant separation in parameter sensitivity
 457 between the chaotic case and the period-1 response is expected; this supports the hypothesis made in
 458 section 3.2, see the difference between P-3 cases and chaotic cases in Table 4 in the Appendix.

459 (c) The nm is less sensitive to the responses' complexity than to the other two parameters. However, it favors
 460 a higher number of nearest neighbors as the complexity of the response increases for accurate reference
 461 modeling. Moreover, if the normalized statistics are used, the nm has a negative correlation to the response
 462 complexity but spans a limited range of selection, see the rows of nm in Table 2.

463 2. As for the parameter selection based on the *slow-time* dynamics:

464 (a) As the quasi-stationarity condition breaks down (e.g., from $a_3 = 30$ to $a_3 = 13$ in Table. 4), the selection
 465 of nm favors less data used for the reference model.

466 (b) The selection of nb is less affected by the quasi-stationarity condition, but it is more sensitive to the damage
 467 growth rate. The period-3 response data concentrates on an intermediate number of hypercuboids; the
 468 increment in its lower bound is due to the decrements in the linear part of the damage growth at the
 469 very beginning of the fatigue life. Moreover, the period-1 response favors a lower value for nb as the

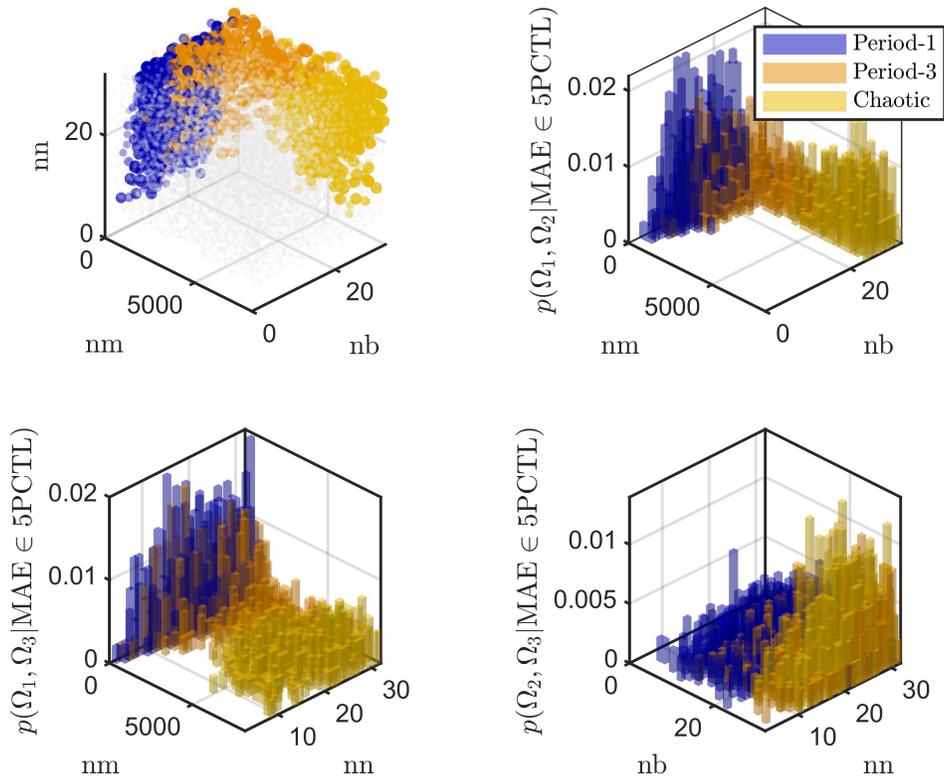


Fig. 14. An example of visual inspection of the reduced parameter space from which the sampled parameters result in high prediction accuracy with the prediction errors within 5-percentile (5PCTL) of the total variance of the prediction error among 10^4 trials from the quasi-Monte-Carlo simulations. The upper-left figure is the parameter triplets' scatter plot, leading to very low prediction error for the prescribed damage. Larger and darker dots indicate the triplets that yield a lower error. The rest of the subplots show the 2-dimensional marginal PDF of the parameters, which gives the 5-percentile error.

exponential growth rate increases. This observation is also consistent with the other two cases; for the period-3 response, its higher bound decreases, whereas for the chaotic one, its lower bound fades into lower values of nb . As the growth rate increases, the damaged manifold diverges from the reference manifold drastically, which causes empty local hypercuboids in the warped phase space if large nb is used. Therefore, a conservative number of nb is always preferred.

- (c) From the non-normalized statistics for both ramp and LE damage dynamics in Table 4, one can conclude that a higher number of nn is preferred for the chaotic responses as the quasi-stationarity breaks down. For example, as the exponential growth rate decreases from $a_3 = 30$ to $a_3 = 13$, the fast dynamics is less stationary, more nn is preferred by the chaotic cases. And the opposite is true for the periodic cases.

Based on the preceding statistical inference, the general guidelines for input parameter selection for the PSW algorithm can be summarized as follows:

1. The selection of nm depends sensitively on both the fast-time dynamics (i.e., the complexity of the response) and the slow-time dynamics (i.e., the damage dynamics). An increase in the complexity of the response necessitates a larger nm . In addition, as the quasi-stationarity breaks down, the lower complexity response favors smaller nm due to the damage contaminating the reference fast-time dynamics.
2. The nb used for the construction of the k-d tree and the high-dimensional representation of the tracking function is more sensitive to the complexity of the fast-time dynamics. Complex responses prefer larger nb before the parameters are normalized for its ability to capture more complex embedded dynamics. Periodic responses have a wider range of valid nb values, whereas the chaotic case needs significantly larger nb to improve damage estimation accuracy. Also, when the slow-time dynamics aggressively deforms or translates the flow, smaller nb is preferred due to the violation of the first assumption, $\Delta t \rightarrow \epsilon t$. Therefore, it is recommended to take a conservative estimate to reduce prediction error and save computation time, e.g., a normalized $nb = 6$ may be a good start.
3. The selection of nn has lower sensitivity compared to the last two parameters. While a larger number of nn is preferred for LLM building, the situation may require fewer nn to keep the neighborhood local. For the

Parameters	Statistics	$r_\phi = 2$			$r_\phi = 20$			$r_\phi = \infty$			
		P-1	P-3	C	P-1	P-3	C	P-1	P-3	C	
Ramp	nm	Mean	179	541	903	310	680	688	207	212	726
		Std	109	425	178	338	418	347	203	157	342
	nb	Mean	7	8	7	8	8	7	5	5	6
		Std	4	2	0.75	3	2	0.5	3	2	0.4
	nn	Mean	11	8	6	11	7	6	7	5	5
		Std	4	3	2	4	2	2	3	2	1
Embedology tuple (τ, d) :		(4, 2)	(3, 3)	(6, 4)	(4, 2)	(3, 3)	(6, 4)	(4, 3)	(5, 4)	(6, 5)	
Parameters	Statistics	$a_3 = 13$			$a_3 = 15$			$a_3 = 30$			
		P-1	P-3	C	P-1	P-3	C	P-1	P-3	C	
LE	nm	Mean	191	208	726	858	455	993	1033	871	976
		Std	120	208	325	558	473	246	635	490	263
	nb	Mean	7	5	6	8	5	6	7	5	5
		Std	4	2	1	3.5	1.5	0.6	4	1	1
	nn	Mean	10	5	5	7	5	4	9	5	4
		Std	4	2	1	3	2	1	4	2	1
Embedology tuple (τ, d) :		(4, 2)	(5, 4)	(6, 5)	(4, 3)	(5, 4)	(6, 5)	(4, 2)	(5, 4)	(6, 5)	

Table 2: A summary of the embedology-parameter-normalized statistics of the marginal probability density function of the reduced parameter space from which the prediction errors stay below 1-decile of the total mean square error. This group of simulations considers both the ramp damage and the linear-exponential damage dynamics with various damage initiation times.

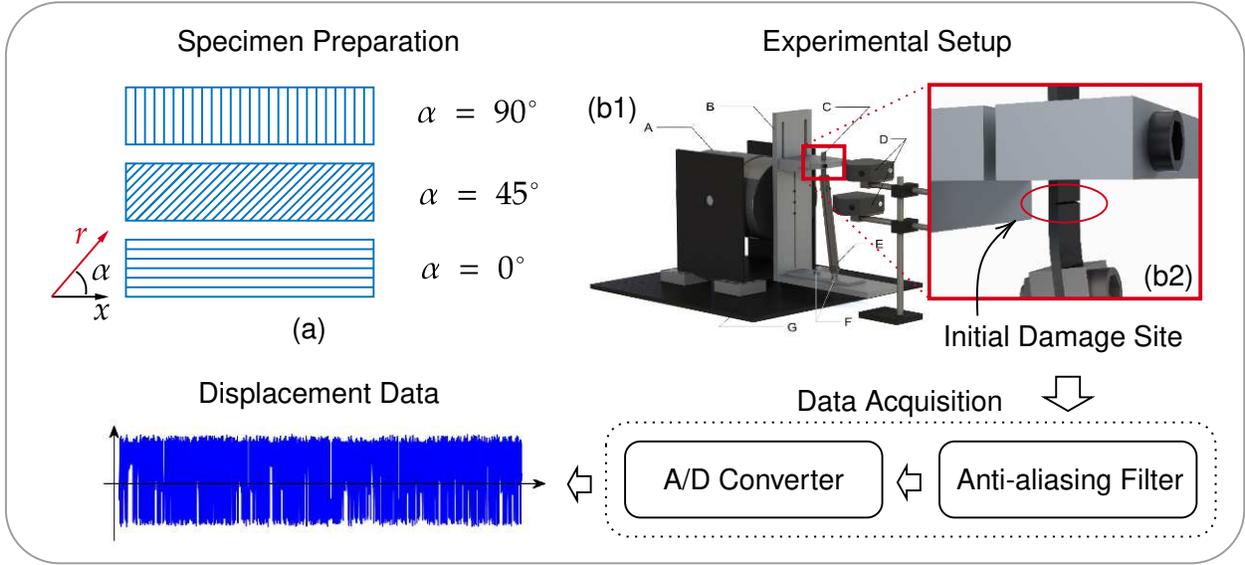


Fig. 15. A graphical illustration of the experimental setup from which the data is obtained and used in this section. (a) illustrates the three distinct raster angles used during the specimen preparation where x is the reference direction, and r is the direction along which the raster is deposited; therefore, the three raster angles are named according to the angle α , the angle between the raster direction and the reference direction. (b1) illustrates the setup of the structural health monitoring system based on a Moon beam setup, where the specimen, C , is cantilevered to the frame, B , which provides harmonic base motion to the beam; two magnets, F , provides nonlinear restoring force; two laser displacement sensors are used to record the beam motion and the base motion, respectively. (b2) illustrates the boundary condition and the initial damage created about 1 inch below the fixture. The effective load at the damage site (initial notch in the red circle) is tension-compression of variable amplitude. The analog signal is low-pass filtered and properly sampled for its use for the PSW algorithm.

- 495 normalized parameters, $nm = 5$ to 6 can be used for the first trial.
- 496 4. The careful selection of LLM's prediction time k is not as significant when compared to all the other parameters
497 within the ranges considered. This result is not consistent with our intuition, and it may be due to the limited
498 range considered in the assumed distribution. Nevertheless, this range of selection is of practical consideration,
499 and we should stick to the lower boundary if practical.
- 500 5. The rl should be large enough so that the temporal averaging is effective and not too many empty boxes exist
501 in the phase space, a power of two closest to nm is suggested. In practical field applications, rl also depends
502 on the rate of updating of the tracking results, e.g., when damage propagation gets faster, a shorter rl may
503 facilitate more frequent updates of the damage state.
- 504 6. As for selecting nr , it should at least satisfy the requirement of a valid matrix decomposition, $nr \geq d \times nb$. In
505 addition, it directly determines how frequently is the state of damage accumulation updated.

506 To reliably obtain the PSW estimate, the authors advocate using the normalized values of nm , nb , and nn , listed
507 in Table 2 during the parameter selection. Since the spatially related parameters are less affected by the type of
508 slow-time dynamics (except for the case where there is less stationarity in the case of $r_\phi = \infty$), the decision of nm
509 is of one's priority. The rest of the windowing-related parameters rl and nr are less important than the others.
510 They should be selected based on the requirement of the resolution of monitoring and the processing power of the
511 monitoring systems.

512 4 Application to Field Measurement Data

513 The PSW is applied to three experimentally obtained displacement-time histories to validate the parameter selection
514 guidelines based on the empirically determined "trust-region" for the PSW algorithm. These displacement-time

515 histories are recorded from a magneto-elastic beam system with polymeric beam specimens whose cross-sections
516 undergo area reduction due to an initial notch and fatigue loading. The beams are 3D-printed with three distinct
517 raster orientations, and their responses to the harmonic forcing are all chaotic, see the graphical abstract of the
518 experiment setup in Fig. 15. This system can be perceived as nonlinear with (1) nonlinear stiffness and (2) time-
519 varying linear stiffness whose sub-structure uses anisotropic materials. Since obtaining the damage evolution in such
520 non-conductive material is difficult, PSW using non-destructive health monitoring signals is more convenient for such
521 applications. The selection of PSW parameters is guided by the delay-coordinate embedding parameters (τ and d).
522 The number of points used in the reference phase space reconstruction is determined using $nm = 1000 \times \tau$. The
523 number of hypercuboids is chosen using $nb = 6 \times d$ and converting it into nearest power of two. The number of
524 nearest neighbors is set to the closest power of two of $nn = 4 \times d$. The number of steps used for local linear model
525 prediction is set as $k = \tau$. The record length for temporal averaging is selected such that the $nr \geq nb \times d$ when there
526 is no overlapping between the windows. If the window size could not satisfy the minimum $nr = nb \times d$, overlapping
527 is applied, e.g., since the 90-degree case has very fast time-to-failure, it uses 500 records with 80% overlapping for
528 all the windowed fast-time histories. All the selected parameters are summarized in Table 3.

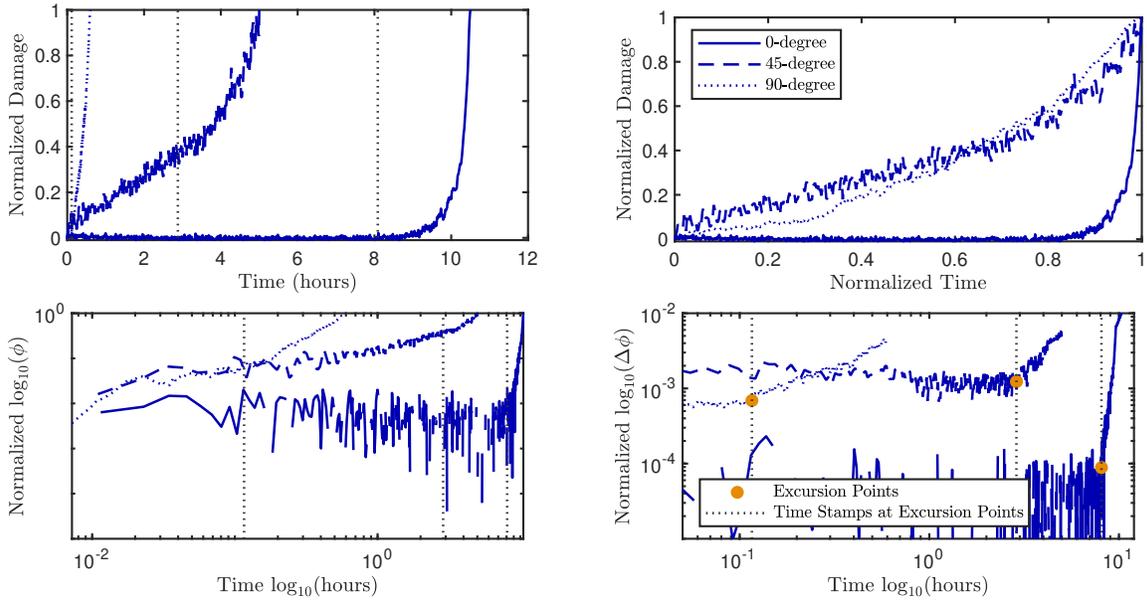


Fig. 16. Estimated damage dynamics (slow-time dynamics) from experimentally obtained displacement laser vibrometry data (fast-time observables). Upper left: estimated damage-time histories for 3D-printed polymeric beam specimens which undergoes chaotic motion; upper right: normalized damage-time histories which resembles the linear-exponential damage simulations; lower left: the estimated damage-time histories in log-log scale, indicating higher rate of damage evolution in the early fatigue lives of the 45-degree and the 90-degree cases; lower right: estimated damage-growth-time histories under log-log scale, which indicating exponential excursions of damage evolution, which can be used as indicators for incoming failures at an early stage.

529 The results of PSW-based damage tracking and identification are shown in Fig. 16. The upper-left subplot shows the
530 extracted damage-time history with a linear scale. One can see that the beam with the 0-degree raster orientation
531 is the most damage resistant, and as the raster angle increases from 0-degree to 90-degree, the damage propagation
532 gets significantly faster. By scaling the three damage-time histories within a range from zero to one, one can clearly

Raster Orientation	f_{ds}	Embedding	Record Time	nm	nb	nn	$step$	rl	nr (Overlap)
0-degree	50Hz	$\tau = 4, d = 6$	10.3 hours	2^{12}	2^5	2^4	4	2^{11}	946 (0%)
45-degree	50Hz	$\tau = 4, d = 6$	5 hours	2^{12}	2^5	2^4	4	2^{10}	457 (0%)
90-degree	50Hz	$\tau = 3, d = 6$	0.6 hours	2^{11}	2^4	2^4	3	2^{10}	500 (80%)

Table 3: Parameters used during the application of the PSW algorithm to the field measurement data

533 see that, under the normalized fatigue life, the 45-degree case has the largest rate of damage evolution, followed by
534 the 90-degree case, and then the 0-degree one. However, the rate of propagation increases drastically in the 90-degree
535 case after 35% of its entire life. As for the 45-degree case, even though its initial rate of damage accumulation is
536 higher, it remains constant for a significant amount of its life (approximately 80 percent of its total fatigue life before
537 a qualitative change in its response is observed). By plotting these damage-time histories using the log-log scale, one
538 can see more closely the damage accumulation rate during the early stage, refer to the lower-left subplot in Fig. 16.
539 The 90-degree case shares a similar damage propagation rate with the 45-degree case up until the first dashed line
540 and diverges with a higher rate than the 45-degree case, which deteriorates with the same rate until the very late
541 life. To better illustrate the amount of damage evolution in terms of their damage accumulation $\Delta\phi$, the $\Delta\phi - t$ in
542 the log-log scale is shown in the lower-right subplot of the Fig. 16. The amount of damage accumulation over time is
543 almost constant for all cases. However, the 0-degree case is one order of magnitude lower compared to the other two
544 cases. Aside from the absolute amount of damage accumulation, prominent inflection points were detected across all
545 three cases at different times. The post-inflection damage-accumulation-time histories show linear trends under the
546 log-log scale. This piece-wise constant to linear dynamics in the $\Delta\phi - t$ plot strongly indicates a linear-exponential
547 damage-time dynamics of the damage evolution. Hence, these inflection points are indicators where the exponential
548 dynamics overtakes the linear growth, after which the damage diverges significantly. Therefore, the inflection points
549 may be used for early damage detection purposes, and the extracted damage-time histories can be used for damage
550 modeling.

551 5 Conclusions

552 This paper establishes guidelines for the reliable application of phase space warping (PSW) for reconstructing slow-
553 time dynamics from fast-time measurements. The focus is on the application to damage tracking and early damage
554 detection in a nonlinear mechanical system. As an aid in understanding the guidelines, a new geometrical inter-
555 pretation of PSW that bridges its theory and algorithm is presented. Based on this novel interpretation, intuitive
556 selection criteria for the input parameters of the PSW algorithm are obtained. Heuristics-based hypotheses in param-
557 eter selection that ensure reliable estimations of PSW are proposed and qualitatively tested through global sensitivity
558 analyses of the input parameters, which are needed to infer the reference *fast-time flow* in the reconstructed phase
559 space. The GSA is conducted without loss of generality based on a set of quasi-Monte-Carlo simulations with various
560 distinct slow-time and fast-time dynamics. The first-order global sensitivity indices lead to a decision of reducing
561 the effective parameter space to three dimensions, spanned by the number of points used in the reference model
562 (nm), the number of hypercuboids created during k-d tree sectioning (nb), and the number of nearest neighbors
563 used in building the local linear models (nn), by ruling out the number of time steps used in prediction. Based
564 on a visual inspection between the marginal probability density of the parameter space and the prediction error,
565 trust regions in which the input parameter quadruple (i.e., nm , nb , nn , and k) yield 1-decile error are characterized
566 statistically. To facilitate the applicability of these statistics in real-world problems, mean values from the PDF of
567 the trust region, normalized by the parameters used in the embedding, are used as metrics to guide the parameter
568 selection. The effectiveness of these metrics is validated through the application to a set of field measurement data
569 that have distinct damage dynamics induced by material anisotropy. The PSW-estimated damage increment versus
570 time histories, under the log-log scale, clearly show inflection points in the damage evolution. These inflection points
571 locate where the excursion of the exponential damage growth overtakes the slow-linear one, which can be further
572 explored and used as indicators for early fatigue damage detection before catastrophic failures.

573 6 Acknowledgement

574 This study is supported by the National Science Foundation Grant No. 1561960.

575 7 Declaration

576 The authors declare that they have no known competing financial interests or personal relationships that could have
577 appeared to influence the work reported in this paper.

578 8 Data Availability

579 The data sets generated during and/or analysed during the current study are not publicly available but can be made
580 available from the corresponding author on reasonable request.

581 A Detailed Derivations

582 A.1 Derivation to Equation (13)

583 Given the expanded flow in Eq. (12), applied the iteration in the same sense, one obtains,

$$\begin{aligned}
G^{2\Delta t}(\phi_{t_0}; \eta(y_{t_0})) &\approx \phi_{t_0} + \epsilon\Delta t \tilde{g}(\phi_{t_0}; \eta(y_{t_0})) + \epsilon\Delta t \tilde{g}(\phi_{t_0} + \epsilon\Delta t \tilde{g}(\phi_{t_0}; \eta(y_{t_0})); \eta(y_{t_0})) \\
&= \phi_{t_0} + \epsilon\Delta t \tilde{g}(\phi_{t_0}; \eta(y_{t_0})) + \epsilon\Delta t \tilde{g}(\phi_{t_0}; \eta(y_{t_0})) + \epsilon\Delta t \left. \frac{\partial \tilde{g}}{\partial \phi} \right|_{\phi=\phi_{t_0}} \underbrace{(\phi(t) - \phi_{t_0})}_{\mathcal{O}(\epsilon t)} + \underbrace{\epsilon\Delta t \mathcal{O}(\|\phi(t) - \phi_{t_0}\|^2)}_{\mathcal{O}((\epsilon\Delta t)^3)} \\
&\quad + \epsilon\Delta t \left. \frac{\partial g}{\partial \eta} \frac{\partial \eta}{\partial y} \right|_{y=y_{t_0}} (y(t) - y_{t_0}) + \epsilon\Delta t \mathcal{O}(\|y(t) - y_{t_0}\|^2) \\
&= \phi_{t_0} + 2\epsilon\Delta t \tilde{g}(\phi_{t_0}; \eta(y_{t_0})) + \epsilon\Delta t \underbrace{\frac{\partial g}{\partial \eta} \frac{\partial \eta}{\partial \phi}}_{\mathcal{O}(1)} \underbrace{\frac{\partial \phi}{\partial y}}_{\mathcal{O}(\epsilon\Delta t)} \bigg|_{y=y_{t_0}} (y(t) - y_{t_0}) + \mathcal{O}((\epsilon\Delta t)^2) \\
&= \phi_{t_0} + 2\epsilon\Delta t \tilde{g}(\phi_{t_0}; \eta(y_{t_0})) + \mathcal{O}((\epsilon\Delta t)^2).
\end{aligned} \tag{44}$$

584 Similarly, the next iteration results in,

$$\begin{aligned}
G^{3\Delta t}(\phi_{t_0}; \eta(y_{t_0})) &= G^{2\Delta t}(\phi_{t_0}; \eta(y_{t_0})) + \epsilon\Delta t \tilde{g}(\phi_{t_0+2\Delta t}; \eta(y_{t_0})) \\
&\approx \phi_{t_0} + \epsilon\Delta t (2\tilde{g}(\phi_{t_0}; \eta(y_{t_0})) + \tilde{g}(\phi_{t_0} + 2\epsilon\Delta t (\tilde{g}(\phi_{t_0}; \eta(y_{t_0})); \eta(y_{t_0}))) \\
&= \phi_{t_0} + \epsilon\Delta t (2\tilde{g}(\phi_{t_0}; \eta(y_{t_0})) + \tilde{g}(\phi_{t_0}; \eta(y_{t_0}))) + 2 \underbrace{\frac{\partial g}{\partial \eta} \frac{\partial \eta}{\partial y}}_{\mathcal{O}(\epsilon\Delta t)} \bigg|_{y=y_{t_0}} (y(t) - y_{t_0}) + \mathcal{O}((\epsilon\Delta t)^2) \\
&= \phi_{t_0} + 3\epsilon\Delta t \tilde{g}(\phi_{t_0}; \eta(y_{t_0})) + \mathcal{O}((\epsilon\Delta t)^2).
\end{aligned} \tag{45}$$

585 As a result, one can obtain the iteration of any integer multiples $k \in \mathbb{Z}^+$ and its subsequent iteration $k+1$; by
586 mathematical induction, the iteration of the flow in the continuous time with a constant time frame Δt can be
587 obtained as,

$$G^{k\Delta t}(\phi_{t_0}; \eta(y_{t_0})) = \phi_{t_0} + \epsilon k \Delta t \tilde{g}(\phi_{t_0}; \eta(y_{t_0})) + \mathcal{O}((\epsilon\Delta t)^2). \tag{46}$$

588 B Marginal Probability Density Function Estimates of the Reduced 589 Parameter Space

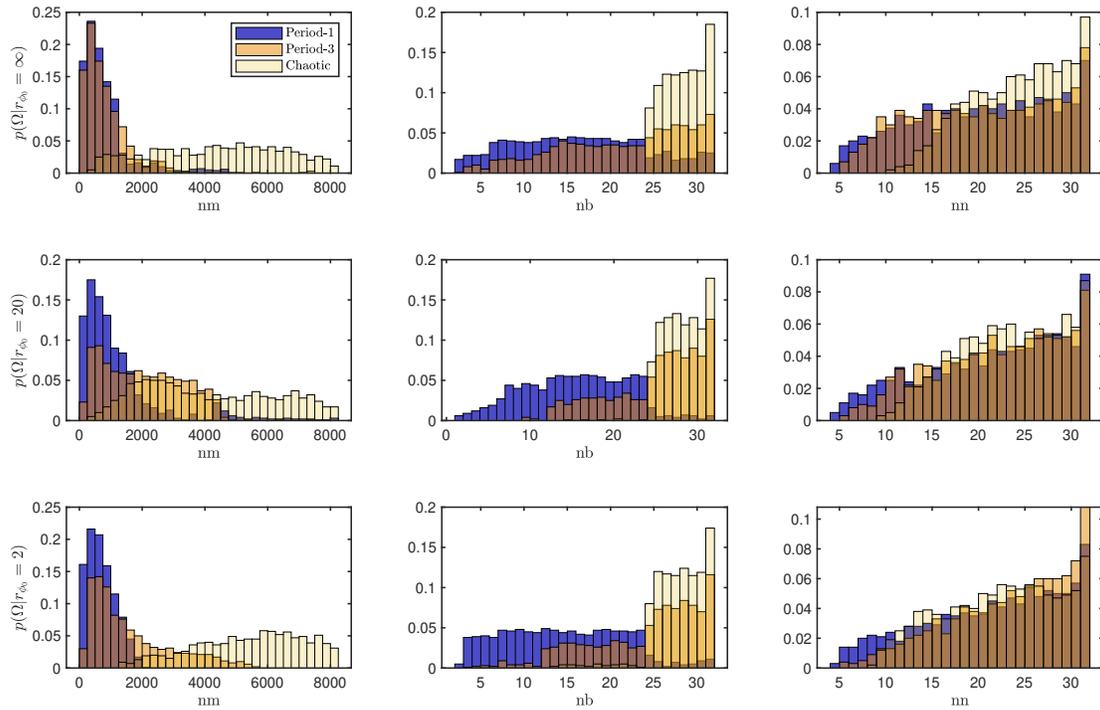


Fig. 17. Histograms to the parameters which yields first-dencile error, the slow dynamics is of the ramp type. From top to bottom, each row specifies ramp damage dynamics with various length of damage-free responses; and from left to right, each column represents one parameter considered in the reduced parameter space.

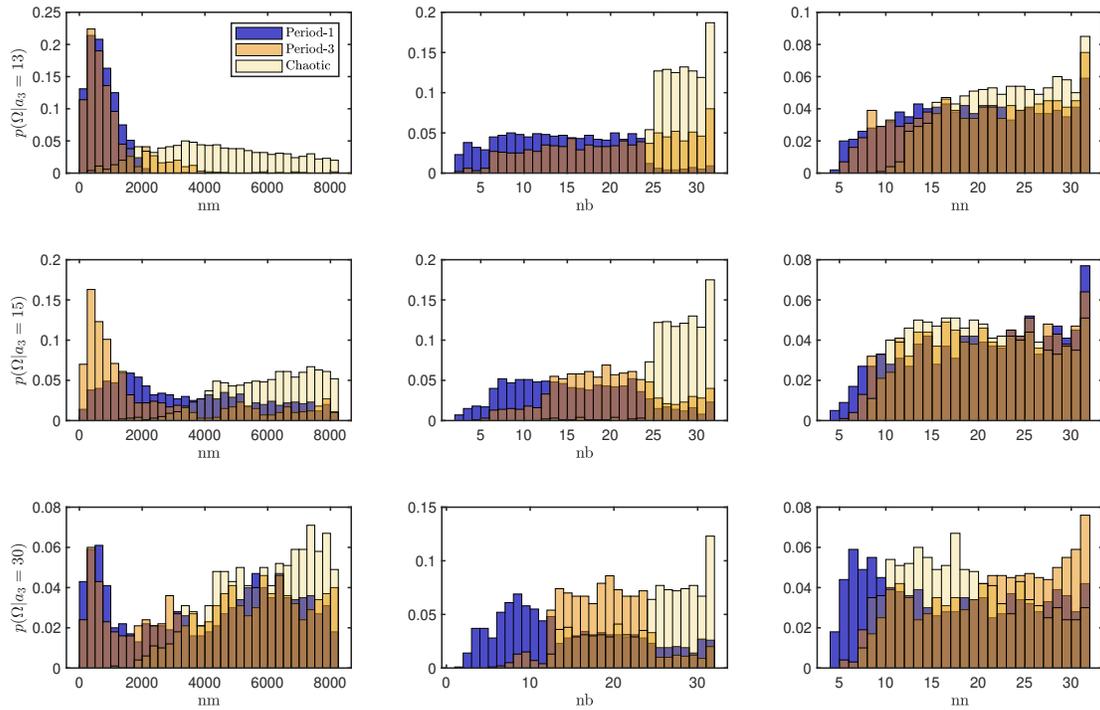


Fig. 18. Histograms to the parameters which yields first-decide error, the slow dynamics is of the linear-exponential type. From top to bottom, each row specifies ramp damage dynamics with various length of damage-free responses; and from left to right, each column represents one parameter considered in the reduced parameter space.

C Table of Statistical Description of the Optimal Parameter Selection

Parameters	Statistics	$r_\phi = 2$			$r_\phi = 20$			$r_\phi = \infty$			
		P-1	P-3	C	P-1	P-3	C	P-1	P-3	C	
Ramp	nm	Mean	716	1624	5416	1238	2040	4125	826	851	4361
		Median	644	1177	5623	819	1925	3821	609	648	4479
		Std	434	1277	1702	1353	1254	2084	814	789	2053
		Skewness	0.60	1.09	-0.38	2.47	0.36	0.24	2.54	3.34	-0.11
	nb	Mean	14	24	28	15	25	28	16	21	28
		Median	14	26	28	15	26	28	16	22	28
		Std	7	6	3	6	6	2	8	8	2
		Skewness	0.22	-0.80	-1.23	0	-0.77	-0.31	0.16	-0.46	-0.03
	nn	Mean	21	23	22	21	22	23	20	20	24
		Median	22	24	22	22	22	23	20	20	24
		Std	7	7	6	8	7	6	8	7	5
		Skewness	-0.42	-0.57	-0.23	-0.41	-0.36	-0.24	-0.22	-0.15	-0.33
LE	nm	$a_3 = 13$			$a_3 = 15$			$a_3 = 30$			
		P-1	P-3	C	P-1	P-3	C	P-1	P-3	C	
		Mean	765	1038	4353	3431	2273	5957	4123	4356	5854
		Median	680	711	4204	2990	1144	6139	4640	4702	6033
	nb	Std	480	1014	1948	2231	2364	1477	2541	2449	1576
		Skewness	1.12	2.55	0.14	0.43	1.18	-0.45	-0.18	-0.21	-0.42
		Mean	14	20	28	15	19	28	14	18	24
		Median	14	21	28	15	19	28	12	18	25
	nn	Std	7	8	2.32	7	6	3	8	5	6
		Skewness	0.21	-0.27	0.02	0.28	0.07	-1.25	0.45	0.30	-0.54
		Mean	19	20	23	20	20	20	17	21	18
		Median	20	20	23	20	20	20	16	22	17
nn	Std	8	8	6	8	7	7	8	7	7	
	Skewness	-0.05	-0.15	-0.15	-0.20	-0.09	-0.09	0.18	-0.28	0.33	

Table 4: A summary of the statistics of the marginal probability density functions of the reduced parameter space from which the prediction errors stay below 1-decile of the total mean square error. This group of simulations consider the ramp and LE damage dynamics with various damage initiation time. The PDFs corresponding to this table can be found in Figs. 17–18 in the Appendix.

REFERENCES

- [1] D. Chelidze and M. Liu, “Reconstructing slow-time dynamics from fast-time measurements,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 366, no. 1866, pp. 729–745, 2008.
- [2] D. Chelidze and J. P. Cusumano, “Phase space warping: nonlinear time-series analysis for slowly drifting systems,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 364, no. 1846, pp. 2495–2513, 2006.
- [3] D. Chelidze and M. Liu, “Multidimensional damage identification based on phase space warping: an experimental study,” *Nonlinear Dynamics*, vol. 46, no. 1, pp. 61–72, 2006.
- [4] M. L. Cler, J. J. Kuehl, C. Skurla, and D. Chelidze, “A novel method for bone fatigue monitoring and prediction,” *Bone reports*, vol. 11, p. 100221, 2019.
- [5] D. Chelidze, “Multimode damage tracking and failure prognosis in electromechanical systems,” in *Component and Systems Diagnostics, Prognostics, and Health Management II*, vol. 4733. International Society for Optics and Photonics, 2002, pp. 1–12.
- [6] M. Song, D. B. Segala, J. B. Dingwell, and D. Chelidze, “Slow-time changes in human emg muscle fatigue states are fully represented in movement kinematics,” *Journal of biomechanical engineering*, vol. 131, no. 2, 2009.
- [7] A. Alwasel, M. Yung, E. M. Abdel-Rahman, R. P. Wells, and C. T. Haas, “Fatigue detection using phase-space warping,” *Journal of biomechanical engineering*, vol. 139, no. 3, 2017.
- [8] J. B. Dingwell, D. F. Napolitano, and D. Chelidze, “A nonlinear approach to tracking slow-time-scale changes in movement kinematics,” *Journal of biomechanics*, vol. 40, no. 7, pp. 1629–1634, 2007.
- [9] J. Kuehl and D. Chelidze, “Identifying invariant manifold using phase space warping and stochastic interrogation,” *International Journal of Non-Linear Mechanics*, vol. 45, no. 1, pp. 42–55, 2010.
- [10] D. Chelidze, “Identifying multidimensional damage in a hierarchical dynamical system,” *Nonlinear Dynamics*, vol. 37, no. 4, pp. 307–322, 2004.
- [11] Y. Qian, R. Yan, and R. X. Gao, “A multi-time scale approach to remaining useful life prediction in rolling bearing,” *Mechanical Systems and Signal Processing*, vol. 83, pp. 549–567, 2017.
- [12] P. Luo, N. Hu, L. Zhang, J. Shen, and Z. Cheng, “Improved phase space warping method for degradation tracking of rotating machinery under variable working conditions,” *Mechanical Systems and Signal Processing*, vol. 157, p. 107696, 2021.
- [13] A. Saltelli, S. Tarantola, F. Campolongo, and M. Ratto, *Sensitivity analysis in practice: a guide to assessing scientific models*. Wiley Online Library, 2004, vol. 1.
- [14] B. Iooss and P. Lemaître, “A review on global sensitivity analysis methods,” in *Uncertainty management in simulation-optimization of complex systems*. Springer, 2015, pp. 101–122.
- [15] A. Lund, S. J. Dyke, W. Song, and I. Bilonis, “Global sensitivity analysis for the design of nonlinear identification experiments,” *Nonlinear Dynamics*, vol. 98, no. 1, pp. 375–394, 2019.
- [16] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola, *Global sensitivity analysis: the primer*. John Wiley & Sons, 2008.
- [17] F. Takens, “Detecting strange attractors in turbulence,” in *Dynamical systems and turbulence, Warwick 1980*. Springer, 1981, pp. 366–381.
- [18] J. C. Sprott and J. C. Sprott, *Chaos and time-series analysis*. Citeseer, 2003, vol. 69.
- [19] H. Kantz and T. Schreiber, *Nonlinear time series analysis*. Cambridge university press, 2004, vol. 7.
- [20] H. Kang, “The prevention and handling of the missing data,” *Korean journal of anesthesiology*, vol. 64, no. 5, p. 402, 2013.

- 634 [21] D. Chelidze and W. Zhou, “Smooth orthogonal decomposition-based vibration mode identification,” *Journal of*
635 *Sound and Vibration*, vol. 292, no. 3-5, pp. 461–473, 2006.
- 636 [22] R. J. Sanford, *Principles of fracture mechanics*. Pearson College Division, 2003.