

Article

Recovery of differential equations from impulse response time series data for model identification and feature extraction

Merten Stender ^{1,2*} , Sebastian Oberst ²  and Norbert Hoffmann ^{1,3} 

¹ Hamburg University of Technology, Germany; m.stender@tuhh.de

² University of Technology Sydney, Australia; sebastian.oberst@uts.edu.au

³ Imperial College London, United Kingdom; n.hoffmann@imperial.ac.uk

* Correspondence: m.stender@tuhh.de; Tel.: +49-40-42878-3120

Abstract: Time recordings of impulse-type oscillation responses are short and highly transient. These characteristics may complicate the usage of classical spectral signal processing techniques for a) describing the dynamics and b) deriving discriminative features from the data. However, common model identification and validation techniques mostly rely on steady-state recordings, characteristic spectral properties and non-transient behavior. In this work, a recent method, which allows reconstructing differential equations from time series data, is extended for higher degrees of automation. With special focus on short and strongly damped oscillations, an optimization procedure is proposed that fine-tunes the reconstructed dynamical models with respect to model simplicity and error reduction. This framework is analyzed with particular focus on the amount of information available to the reconstruction, noise contamination and non-linearities contained in the time series input. Using the example of a mechanical oscillator, we illustrate how the optimized reconstruction method can be used to identify a suitable model and to extract features from uni-variate and multivariate time series recordings in an engineering-compliant environment. Moreover, the determined minimal models allow for identifying the qualitative nature of the underlying dynamical systems as well as testing for the degree and strength of non-linearity. The reconstructed differential equations would then be potentially available for classical numerical studies, such as bifurcation analysis. These results represent a physically interpretable enhancement of data-driven modeling approaches in structural dynamics.

Keywords: signal processing; sparse regression; system identification; impulse response; optimization; feature generation; structural dynamics; time series classification

1. Introduction

As measurement equipment and sensors have become less expensive and computationally more powerful in recent years, the amount and quality of data acquired in various engineering applications have steadily improved. While data have been collected and analyzed in classical laboratory testing ever since, on-line measurements of machines during operation are becoming increasingly popular. Revisiting the main reasons for data collection on mechanical structures, at least three fundamental motives can be identified:

1. Model identification, model validation and model updating for control and numerical simulation, mostly performed during the product development phase [1–3].
2. State observation and supervision, especially for quality assurance and human-machine interaction purposes during operation.
3. Future state and fault prediction, commonly known as structural health monitoring [4].

While each of these objectives has unique features, specialized methods and particularities, there are common themes throughout all purposes for data acquisition in a mechanical engineering

environment. Typically, time series data are collected, that is a number of values arranged in a sequence prescribed by time, and is then post-processed to yield quantities of interest. These quantities are the input values for the methods used and thus must be discriminative for the respective purpose. For example, a trend of changing feature values might indicate and predict a fault in a machine or its malfunction - a situation commonly studied in the field of structural health monitoring. While some of the post-processing methods do not rely on a sequential character, such as statistical moments, others, such as spectral analysis, are defined based on the temporal information of sequence data (frequencies). In both cases, the sample of values drawn from measurement are assumed to a) represent the underlying distribution well and b) remain stationary for the time of observation. However, real vibration data from complex operating structures are often non-stationary, leading to less descriptive classical signal processing features as the data become decreasingly stationary. Impact and impulse measurements represent a major challenge for feature extraction as the dynamics are highly transient and short, in some cases causing subsequent values to differ in orders of magnitudes [5]. Hence, discriminative features cannot always be derived from classical methods like the Fourier transform for those signals.

This work aims to derive new features for short and highly transient vibration time series data. Particularly, we propose to derive governing differential equations directly from the vibration data [6]. Along with increased data accessibility, a rise in data-driven methods can be observed. Recently, several methodologies have been proposed that enable the extraction of differential equations from time series data. Mostly driven by fluid dynamics applications and control, different core methodologies can be observed:

- neural network based methods [7,8],
- compressive sensing approaches [9–11],
- and sparse regression approaches [12,13].

In this work, we employ the Sparse Identification of non-linear Dynamics (SINDy) approach recently proposed by Brunton et al. [12]. Using SIND, a set of ordinary differential equations that models the input time series data can be obtained. Because not all states can be measured during experiments, state space reconstructions are utilized to find the missing active degrees of freedom that are encoded in a measurement. Then, the sparse regression allows for finding a set of coefficients that represents a set of ordinary differential equations (ODEs) in a prescribed space of admissible ansatz functions. The core contribution of this work is similar to that of Oberst et al. [14], but is more comprehensive and general in that it constitutes an extension of SINDy to a higher degree of automation. Here, the SINDy ODE approximation is subjected to a constrained non-linear optimization scheme to define equations that better replicate the input data and favor simplicity of the reconstructed model. The approximation quality is assessed by means of time series comparison measures. Hence, the proposed approach represents an integrated, data-driven model identification methodology that is specifically developed for short and transient signals [14]. The structure of the extracted set of ordinary differential equations (ODEs) can be used for several purposes, such as model identification, model updating, and as a feature engineering proxy for machine learning techniques. In the following, we illustrate the SINDy algorithm and related techniques in detail. Then, a single degree-of-freedom (DOF) damped oscillator is studied as a reference in various linear and non-linear configurations. The complexity of the ODE identification is gradually increased by removing information. Opportunities and limitations of the proposed method are discussed with regards to model identification and feature engineering purposes.

2. Methods

In the following, we introduce the data-driven reconstruction of differential equations as well as several techniques that are required to yield an optimal reconstruction given time-limited and noise-contaminated data.

2.1. Sparse Identification of non-linear Dynamics - SINDy

We apply the methodology proposed by Brunton et al. [12] to find differential equations that describe measured time signals. The sparse identification of non-linear dynamics (SINDy) makes use of a sparsity assumption on the observed dynamics: Most physical phenomena can be described by governing equations that are rather sparse in the huge space of possible non-linear functions. For instance, the Lorenz system consists of three ordinary differential equations with only two multiplicative terms of the system's states. Hence, this system can be considered very sparse in the possible function space of polynomials, irrational functions and others. However, the Lorenz system exhibits regular as well as irregular dynamics and is able to describe some features of atmospheric convection, which is inarguably complex, involving a plethora of qualitatively different dynamics.

Next, the SINDy framework, as described in detail by Brunton et al. [12], is briefly introduced here, with focus on its interface to non-linear optimization techniques. The procedure of SINDy is depicted in Figure 1. Here, only time-continuous systems physical flows of the type

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)) \quad (1)$$

are considered. The system states $\mathbf{x} \in \mathbb{R}^m$ as well as their temporal derivatives $\dot{\mathbf{x}}$ are assumed to be accessible, while the underlying system is unknown. For instance, the system states can be measured during vibration experiments and their derivatives can be approximated numerically. Then, a library of non-linear functions Θ is compiled. The central objective of the reconstruction is to approximate \mathbf{f} from (1) with as few terms as possible from the functions library, that is to find a sparse representation in the space of possible candidate functions. The state measurements and derivatives for time instances t_1, \dots, t_N are collected in matrices $\mathbf{X} = [\mathbf{x}_1^\top(t), \dots, \mathbf{x}_m^\top(t)]$ and $\dot{\mathbf{X}} = [\dot{\mathbf{x}}_1^\top(t), \dots, \dot{\mathbf{x}}_m^\top(t)]$, respectively. To reconstruct the underlying system of differential equations, the over-determined system of equations

$$\dot{\mathbf{X}} = \Theta(\mathbf{X}) \Xi \quad (2)$$

is evaluated using the function library Θ for each system state. The central idea of SINDy is then to solve the resulting system of equations using a sparse regression approach. In contrast to classical solution schemes, such as least-squares algorithms, this procedure promotes sparsity in the coefficient matrix Ξ , which provides the best solution to the over-determined system. As a result, \mathbf{f} , that is the underlying dynamical system, is approximated by only few functions from the library of ansatz functions. Hence, the solution to the reconstruction problem is sparse with respect to many possible candidate functions to approximate the non-linear relation between system states and derivatives. The reconstructed set of differential equations can be read directly from the non-zero entries in the solution Ξ

$$\dot{\mathbf{x}} = \Xi^\top \left(\Theta(\mathbf{x}^\top) \right)^\top \approx \mathbf{f}(\mathbf{x}), \quad (3)$$

where $\Theta(\mathbf{x})$ denotes the evaluation of the symbolic library of non-linear functions and $\tilde{\mathbf{x}}$ the reconstructed state. Essentially, each coefficient vector ξ defines the linear combination of non-linear functions for each state of the system:

$$\dot{x}_1 = \xi_{1,1} \cdot \theta_1(\mathbf{x}) + \xi_{1,2} \cdot \theta_2(\mathbf{x}) + \dots + \xi_{1,k} \cdot \theta_k(\mathbf{x}). \quad (4)$$

Here, θ denotes one of the k ansatz functions, such as a quadratic polynomial $(\cdot)^2$, in Θ , see also the example in Figure 1. The reconstruction result, i.e. the set of governing differential equations, is built from the sparse coefficient matrix and the library of non-linear functions. This system of equations can then be subjected to further analysis, such as forward time integration or bifurcation analysis. In this work, we use the time integration of the reconstructed equations to compare the reconstructed dynamics $\tilde{\mathbf{x}}(t)$ to the observed dynamics $\mathbf{x}(t)$, i.e. the input of SINDy, to define a measure for the reconstruction error estimate Γ .

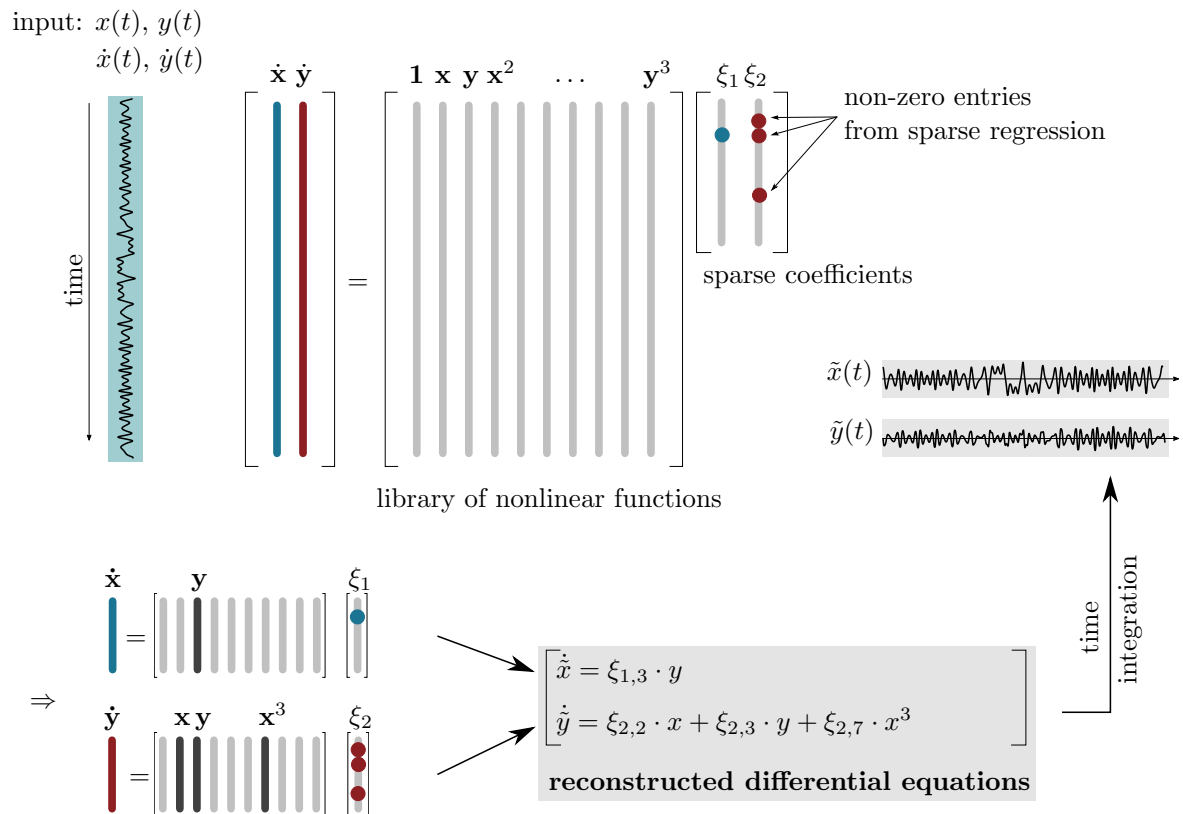


Figure 1. Schematic illustration of the SINDy procedure, closely related to the illustration given in [12]. System states and their derivatives are required as time series input. Then, a library of non-linear candidate functions is compiled. The resulting over-determined system of equations is solved by sparsity-promoting regression techniques to yield a sparse reconstruction of the underlying dynamical system in terms of ordinary differential equations

2.2. Time-delay embedding

In measurements, data is limited to a finite number of signals points. Hence, only a fraction of the relevant dynamical states of a dynamical system can be measured in laboratory experiments. Time delay embedding allows for the reconstruction of the phase space of a deterministic motion that is observed in an uni-variate fashion. Takens's theorem [15] guarantees that for strictly deterministic processes the trajectories in the reconstructed space have the same geometrical and dynamical properties as in the true phase space. Hence, time-delay embedding enables us to unfold the dynamics on an attractor, the invariant set, that is encoded in a single scalar time measurement. Then, the reconstructed attractor's properties, so-called invariant measures, can be studied, which include so-called *static* quantities like the correlation dimension D_2 or dynamic quantities like the Lyapunov spectrum Λ [16,17]. Time-delay embedding, given by

$$\mathbf{x}(t_n) = [s_{n-(m-1)\tau}, s_{n-(m-2)\tau}, \dots, s_{n-\tau}, s_n], \quad (5)$$

essentially re-arranges the scalar time series $s(t_n)$ via sampling into a m -variate time series $\mathbf{x}(t_n)$. Here, τ denotes the *delay*, which measures the distance of successive entries that are arranged to the instantaneous state vector of *embedding dimension* m , see Figure 2. The embedding parameters are obtained sequentially. First, the optimal delay value is derived from the temporal correlations present in the time series $s(t)$. The aim is to find temporally maximal de-correlated state space vectors. Hence, common approaches for deriving τ involve the first zero of the signal's auto-correlation function (ACF)

or the first minimum of the auto-mutual information (AMI) function which also takes non-linear correlations into account. The embedding dimension must be chosen to be larger than twice the attractor's dimension. For example, the motion on a limit cycle, i.e. a one-dimensional attractor, has to be embedded into two dimensions to fully unfold the attractor. While there exist multiple approaches to estimate the embedding dimension, the false nearest neighbor (FNN) algorithm is the standard tool for deriving the global dimension m [17–19]. The FNN algorithm iteratively increases the embedding dimension and at the same time tracks neighborhood relations in the m -dimensional phase space. False nearest neighbors are defined as states that are neighbors in dimension m , but lose their neighborhood relation once the embedding dimension is increased to $m + 1$. The fraction of false nearest neighbors measures the overall share of those points on all state space vectors. Once the attractor is fully unfolded, states remain neighbors even if the dimension is further increased. Hence, the required embedding dimension is characterized by a low fraction of false neighbors in the m -dimensional reconstructed phase space. In this work, we employ time-delay embedding to reconstruct multivariate trajectories from uni-variate time series measurements. Additionally, the SINDy-reconstructed dynamics can be compared to measured dynamics by employing embedding parameters as error metric, cf. [20].

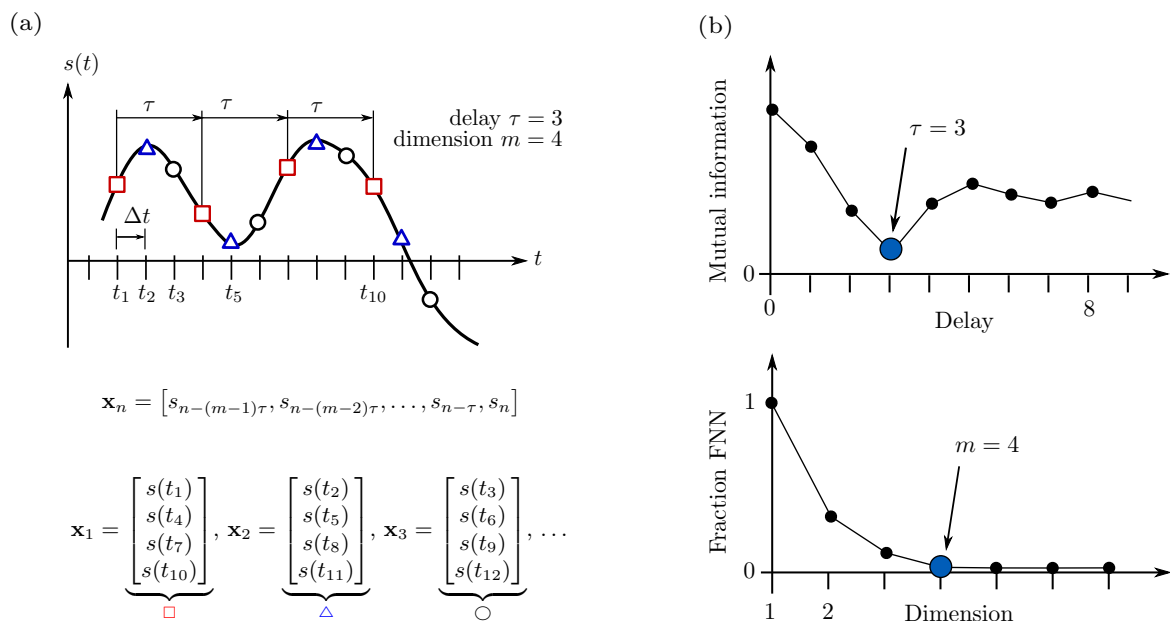


Figure 2. (a) schematic illustration of the time delay embedding procedure using the embedding parameters τ and m that were obtained from (b) the first minimum of the auto-mutual information function and the dimension with vanishing fraction of false nearest neighbors, respectively

2.3. Uncertainty suppressing numerical differentiation

Besides all states, SINDy also requires the states' time derivatives either measured or generated numerically, whereas the form is rarely conducted. Classical numerical differentiation schemes, such as finite differences, exhibit noise amplification - as a rule of thumb, accuracy decreases by one order of magnitude per differentiation [21]. As measurements are always contaminated with noise, a differentiation scheme is required that is particularly robust to account for this additional loss of accuracy. We employ Total Variation Regularized Numerical Differentiation, *TVRegDiff*, as introduced by [22] and proposed in [12], to be used within SINDy. The central idea of this differentiation scheme is to regularize the differentiation itself by balancing a) the irregularity introduced by the derivative and b) the error introduced by the smoothing property using a regularization parameter. Hence, this regularization parameter, denoted as α , plays a crucial role. In the following, we propose an integrated and completely data-driven approach for choosing an appropriate α for the specific data under study.

In fact, the regularized derivative can serve as a filtering method. The regularization, controlled here by the parameter α , can then be utilized to also dampen noise in the differentiation step. The smoothed derivative is then integrated in time to yield a qualitatively enhanced (filtered) version of the input signal, see illustrative examples in Appendix A. In this work, we use TVRegDiff to compute the numerical derivatives of states $\mathbf{x}(t)$ with respect to time. In order to arrive at a self-consistent procedure, we substitute the original states with the time-integrated derivatives for usage in the SINDy procedure as explained before. This approach ensures that states and derivatives are matching. Here, we propose to define a time series difference measure that compares the raw input signal to the one obtained through numerical derivation and time-integration similar to Oberst et al. [14]. The search of the difference minimum will yield the optimal regularization parameter, see Figure 3. In order to reduce boundary effects at the beginning and end of each time series, the five first and last samples are discarded after differentiation.

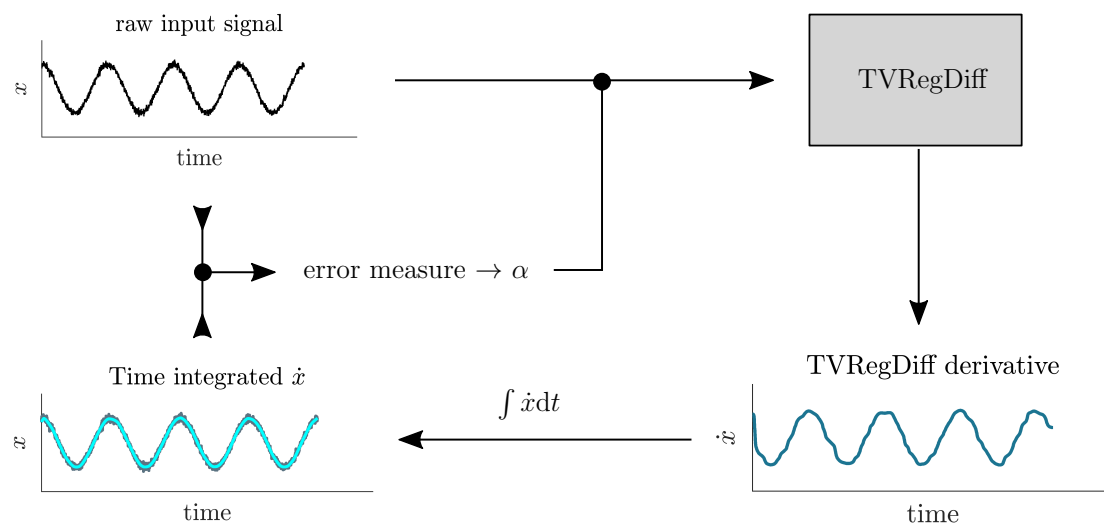


Figure 3. Integrated approach for finding the optimal regularization parameter α for the numerical differentiation using TVRegDiff: the time integration of the numerical derivative allows for deriving an error measure that can be used to adapt the regularization parameter for noise contaminated input time series

2.4. time series comparison measures

This work extracts differential equations from time series input data. Hence, appropriate error measures have to be chosen to compare the reconstructed dynamics with the input data and thus quantify the reconstruction quality. Particularly, the set of reconstructed ODEs can be integrated in time (setting the initial value to the first input data entry) to create the reconstructed time series. Depending on the qualitative character of the time series and the specific reconstruction purpose, the reconstruction error needs to be defined. This objective represents an analogon to research fields of time series classification and time series feature engineering, where finding discriminative features forms the central methodical challenge [23–25]. The types of features, and thus possible candidates for definition of the reconstruction error, can be clustered as follows:

- Instance-based schemes that compare contemporaneous pairs of time series instances. In the simplest case, sequences are subtracted from each other. Modifications and advanced approaches include warping methods that add more flexibility and the ability to also take into account phase shifts.
- Higher-level features based on transforms, sampling strategies or correlation measures. Examples include statistical moments and linear transforms, such as the Fourier transform. Time series

comparison is then performed based on features extracted from the transforms, such as the standard deviation or major periodicities.

- Quantifiers for qualitative behavior, mostly borrowed from non-linear time series analysis and complexity sciences [26]. Here, the actual shape of the sequence is rather irrelevant. Instead, the qualitative nature of the dynamics encoded in the time series is of interest: dynamical invariants quantify the degree of regularity, entropy, or fractal properties of the sequence when studied in a dynamical framework, cf. [14].

Depending on the specific signal, reconstruction purpose and availability of data, any of the aforementioned features may be chosen to quantify the reconstruction error between the input and SINDy-reconstructed time series. In this work, we use the cumulative difference between reconstructed and original signal as a measure for instance-based reconstruction error due to its higher robustness, computational efficiency, and simplicity of implementation compared to e.g. recurrence plot-based measures [14].

2.5. Constrained non-linear optimization

The sparse regression solution of the over-determined SINDy system of equations is not continuous in the sense of the coefficient matrix Ξ , which is a function of the sparsity parameter λ . This means that a fixed set of coefficients will result from a range of sparsification levels λ . If λ is increased, at some point, an additional coefficient entry will be set to zero. Until then, the coefficients remain constant. As a result, SINDy may be able to find the correct locations of non-zero coefficient entries, but not their optimal absolute value. We propose to use a non-linear optimization scheme to fine-tune the entries in the sparse coefficient matrix found by SINDy. In this work, the Matlab routine *fmincon* is employed to run the optimization using the sequential quadratic programming (sqp) method. For each non-zero coefficient, we prescribe a lower and upper bound. Then, the optimizer finds the optimal set of coefficients with respect to the cost function, which measures the difference between input signal and reconstruction signal. Hence, the SINDy results represent the first approximation of the reconstruction and the non-linear optimization fine-tunes the model for the best fit. Depending on the specific reconstruction objective, the cost function may be formulated such that it penalizes complexity and non-linearities, promotes matching spectral properties [27], or improves other characteristics of the reconstructed model.

2.6. Models used

We employ a linear, 1-DOF, viscously damped mechanical oscillator to illustrate the automation and optimization extensions of SINDy. Special focus is put on the short and highly transient nature of ring-down data, i.e. impulse-response-type time series data. As a reference, the oscillator is defined as shown in Figure 4. Owing to high damping levels, there are usually only a few significant oscillations. Hence, determining classical features of ring-down data, such as the decay rate, may be challenging and prone to errors. In the following, we discuss the reconstruction in the phase space, which is represented by a first-order description of the form

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\frac{c}{M}x_2 - \frac{k}{M}x_1.\end{aligned}\quad (6)$$

This dynamical system is integrated in time for 4 s using a sampling time of $\Delta t = 0.01$ s. If not denoted differently, a unit initial deflection is prescribed, i.e. $\mathbf{x}(t=0) = [1, 0]^T$. For the sake of clearer notation, we discard the physical units in the following discussions.

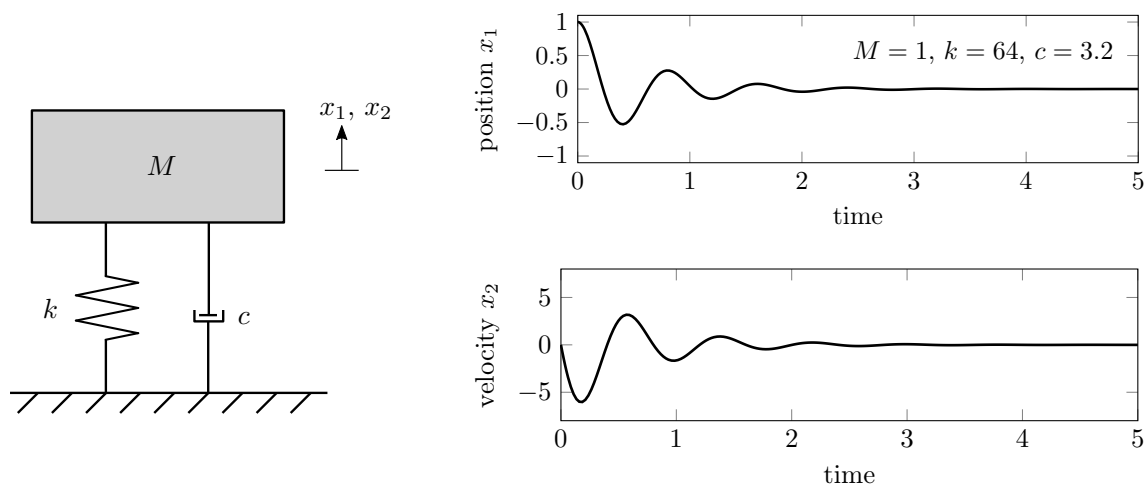


Figure 4. Linear 1-DOF viscously damped oscillator and time traces of position x_1 and velocity x_2 for an initial deflection of $x_1(t = 0) = 1$

3. Results

We elaborate on the proposed SINDy extensions using the example of the mechanical oscillator that oscillates freely. In the first step, the linear damped oscillator is subjected to the SINDy reconstruction for various degrees of input information level. The complexity of the reconstruction task is increased incrementally. First, the analytical derivatives are replaced by the regularized derivatives. Then, only the first state is recorded, which requires reconstructing the second state using time delay embedding. Secondly, we illustrate how the optimization procedure can improve both the reconstruction quality and the simplicity of the reconstructed models. Finally, a cubic stiffness term is introduced and identified by the proposed methods. These cascades of decreasing amounts of information available to SINDy are illustrated in Figure 5.

3.0.1. Instance-based error measure

As the error measure, we employ the cumulative absolute Euclidean distance, given by

$$\Gamma = \sum_{i=1}^{\text{DOF}} \sum_{j=1}^N \left| \frac{x_{\text{true},i}(t_j)}{\max(|\mathbf{x}_{\text{true},i}|)} - \frac{x_{\text{recon},i}(t_j)}{\max(|\mathbf{x}_{\text{true},i}|)} \right| \quad (7)$$

throughout the followings discussion. For each DOF, the difference is computed between the true \mathbf{x}_{true} and reconstructed $\mathbf{x}_{\text{recon}}$ states per time step t_j . Then, all differences are added up to form the scalar error measure Γ . To achieve a balanced error measure, each time series of length N is normalized by the respective maximum of the true states per DOF. Hence, differences in position and velocity are weighted equally in the error measure and do not depend on the actual value range per DOF.

3.0.2. Providing full information: all states and analytical derivatives

Initially, all information is provided for SINDy: position and velocity signals as well as their derivatives are provided from the time integration of the equations of motion. As we provide a library of polynomials up to order $p = 3$, the exact coefficients should be found by SINDy. However, as parameter remains the sparsification level, which has to be defined according to the length, the sampling time, the non-linear function library and the qualitative nature of the dynamics. Following the previous illustrations, the sparsification parameter λ is varied in the full range between non-zero entries (NZE) of 100% to NZE = 0%. Meanwhile, the error measure is computed from the time domain solutions of instantaneous SINDy results. Then, the optimal sparsification parameter is chosen

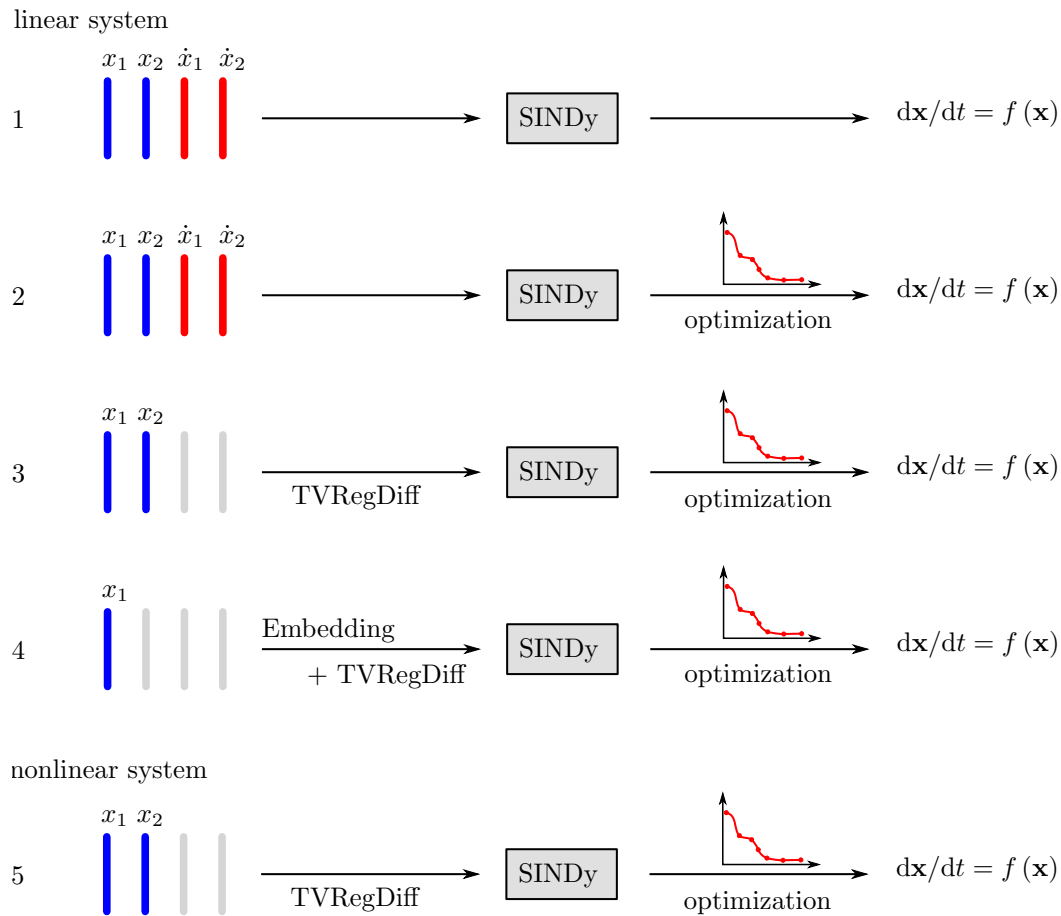


Figure 5. Overview of the studies performed using the 1-DOF oscillator. The reconstruction task becomes increasingly challenging due to decreasing amounts of information available and added non-linearity

such that it minimizes the error measure. The resulting SINDy reconstruction is integrated in time and compared to the input signal, see Figure 6. The result of SINDy matches perfectly with the analytical expression for the oscillator. Even though polynomials up to order $p_{\max} = 3$ were permitted, the optimal sparsification level yields the correct first-order differential equations $\dot{x}_1 = 1 \cdot x_2$ and $\dot{x}_2 = -64 \cdot x_1 - 3.2 \cdot x_2$ with coefficients equaling zero for all other entries, thus a NZE fraction of 3/20. As the error is close to machine precision, we do not need to further tune the non-zero coefficients, regardless of whether or not the true solution of the input signal is known.

3.0.3. Providing less information: all states and numerical derivatives

Analytical derivatives will hardly be available for an analysis as proposed here. In other words, if states and derivatives were available, there would be different solutions to find the best system of differential equations that links states with derivatives. Hence, in this scenario we record only the states x_1 and x_2 and compute their derivatives using TVRegDiff. Therefore, the regularization parameter α , has to be set prior to applying SINDy. Here, we consider noise-free signals, which makes the selection of α rather straight-forward. Following the description illustrated in Figure 3, we adapt the regularization such that it minimizes the difference between an input state and the state which is obtained after time integration of the numerical derivative, see Figure 7. When dealing with noise-contaminated signals, this is a semi-automated procedure that must be supervised by the user in order to make use of the filtering property of regularized derivatives A.

The filtered states, that is the time integrated numerical derivatives, are fed as input x_{in} to the SINDy procedure along with the numerical derivatives. Using a maximal polynomial order $p = 3$, the

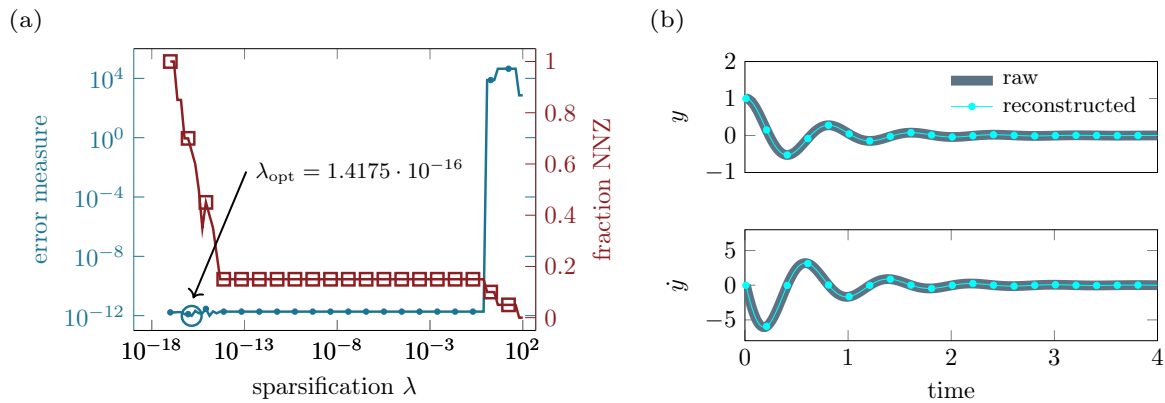


Figure 6. (a) finding the optimal sparsification level as minimum of the error measure. (b) comparison of input signal to SINDy reconstruction with maximal polynomial order $p_{\text{max}} = 3$

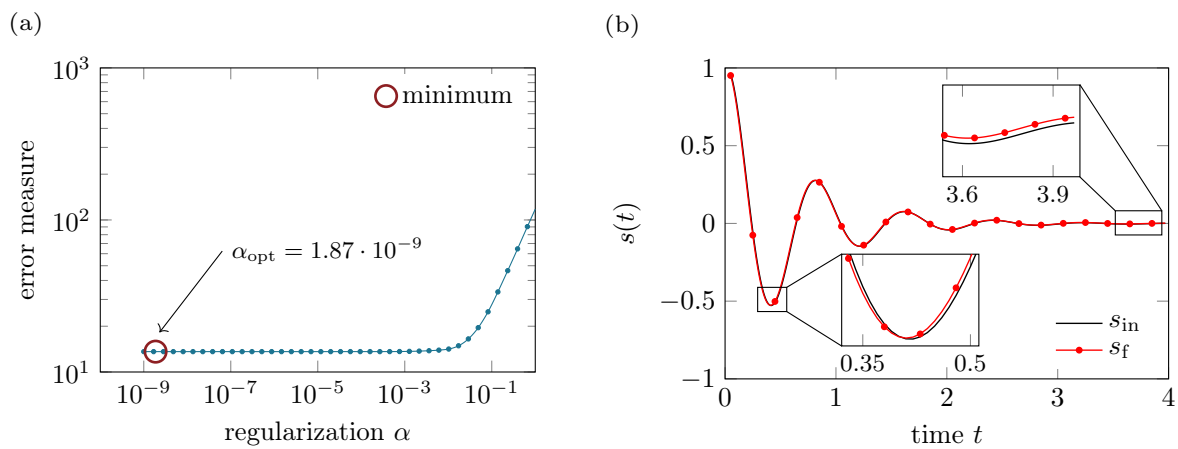


Figure 7. Comparison of the input time signal s_{in} and the filtered one s_{f} obtained through time integration of the regularized numerical derivative. (a) error measure as a function of the regularization parameter α and (b) the resulting signal for an optimal regularization α_{opt}

sparse reconstruction is obtained following the procedure elaborated before. The best reconstruction is obtained for a sparsification parameter $\lambda = 0.3583$. The reconstructed time signals for x_1 and x_2 match well with the input signals, see Figure 8(a). Out of the 20 possible coefficients, there are four non-zero entries: $\dot{x}_1 = 0.999 \cdot x_2$, $\dot{x}_2 = -64.8124 \cdot x_2 - 2.9204 \cdot x_1 - 0.5125 \cdot x_1^3$. The cubic term represents an artefact that was introduced by SINDy: Compared to the true system of equations, one can observe that the linear system's eigenfrequency is slightly too large ($\sqrt{64.81} > \sqrt{64.0}$), while the damping term is slightly smaller ($2.92 < 3.2$), which explains the additional cubic stiffness term. This cubic term creates larger restoring forces and thus compensates for the lack of damping in the system. As the reconstruction error is rather small, and the reconstructed dynamics are sparse in the admissible non-linear function library space, one may finish the reconstruction at this point.

However, we strive for an even smaller reconstruction error, fewer model terms and a test for non-linearity. Especially the latter objective may be of interest for various studies in system identification scenarios. In structural dynamics modeling, the degree of linearity is a point of crucial importance. Since linear systems obey the superposition principle, analysis of these systems can be performed with ease using a variety of established methods. Strongly non-linear systems, on the other hand, require more complex analytical methods and can exhibit a profusion of qualitatively different dynamics that typically have not been understood as thoroughly as those of linear systems. Hence, in the current case it might be interesting to see whether the cubic stiffness term is a decisive model parameter, or if the system can be modeled in linear fashion instead.

A central contribution of our work is the extension of the SINDy reconstruction with a non-linear optimization procedure. The optimization problem is formulated such that the non-zero coefficients are subjected to a bounded optimization with respect to a global cost function that measures the difference of input and reconstructed time signals. The non-zero entries (NZE) in Ξ found by SINDy represent the starting point for the optimization, which is permitted to change the NZE values within prescribed boundaries ξ_- and ξ_+ . The optimization boundaries are set in relative fashion, i.e. allowing for a certain fraction of change with respect to the initial value. However, to promote sparsification during the optimization, we set one of the boundaries to zero. This approach, which allows for further elimination of model parameters, corresponds to the *dropout* or *pruning* procedure, a well-known concept of complexity reduction in artificial neural networks [28]. Here, we allow a maximal relative change of 50% according to

$$\begin{aligned}\xi_i < 0 &\rightarrow \xi_-^+ = [0, \xi_i \cdot \gamma], & \gamma = 1.5 \\ \xi_i < 0 &\rightarrow \xi_-^+ = [\xi_i \cdot \gamma, 0] \quad .\end{aligned}\quad (8)$$

In our case, the non-zero entries in ξ obtained by SINDy are the coefficients of x_2 in the first differential equation and those of x_1 , x_2 and x_1^3 in the second differential equation. Figure 8 (b) depicts the optimization iterations of the bounded non-linear optimization. The error measure decays quickly within the first ten iterations. Only slight adaptations are made in the following iterations before the optimization algorithm finishes at 48 iterations. When observed in the time domain, the optimization iterations are not visible in the full-scale time signal, but can be identified through high magnification of the peak regions, see Figure 8 (c). The final minimisation result is then compared to the input signal, i.e. the reconstruction aim, and the SINDy result in Figure 8 (d). Through optimization, the reconstruction error Γ has been reduced by 81.95% to 2.345. While the reconstruction has been enhanced, the coefficient matrix Ξ has also become sparser thanks to this method: The auxiliary cubic stiffness term that was computed by SINDy is erased by the optimization, which makes use of the dropout option discussed before. The reconstructed set of differential equations after optimization is very similar to the analytical model which was used to create the input data, see Figure 8 (e).

3.0.4. Providing less information: state space reconstruction and numerical derivatives

In the next step, the amount of information provided by the signal input is further reduced. In real measurements, only a few measurement points are available for data acquisition. Therefore, only a fraction of the active degrees of freedom will be measured and available to the reconstruction procedure. We mimic this situation by restricting the input data to a single scalar time series, e.g. x_1 of the oscillator. The missing information about other relevant degrees of freedom in the true phase space has to be reconstructed using time delay embedding. Then, TVRegDiff allows for computing the corresponding derivatives that can be used for the SINDy reconstruction with subsequent coefficient optimization.

The embedding parameters $\tau = 18$, $m = 2$ are found for the x_1 time series by the first zero of the auto-correlation function and by the false nearest neighbor algorithm, see Figure 9. The resulting trajectories q_1 , q_2 in the reconstructed phase space and their respective numerical derivatives are used as inputs to SINDy.

The optimized SINDy reconstruction is computed for polynomials varying from degrees $p = 1$ to $p = 5$. For each configuration, the sparsity as well as the reconstruction error estimate Γ between trajectories $\mathbf{q}(t)$ and their reconstruction is computed. The corresponding sets of ODEs are listed in Table 1. Generally speaking, the reconstructed ODEs will not mimic the analytical oscillator as a result of the embedding step. The reconstruction of the trajectories $\mathbf{q}(t)$ maps the input $x_1(t)$ to another space and, therefore, the resulting states cannot be compared directly to the analytical model. In contrast, the general structure of the dynamical system as well as non-linearity serve as viable parameters for comparison.

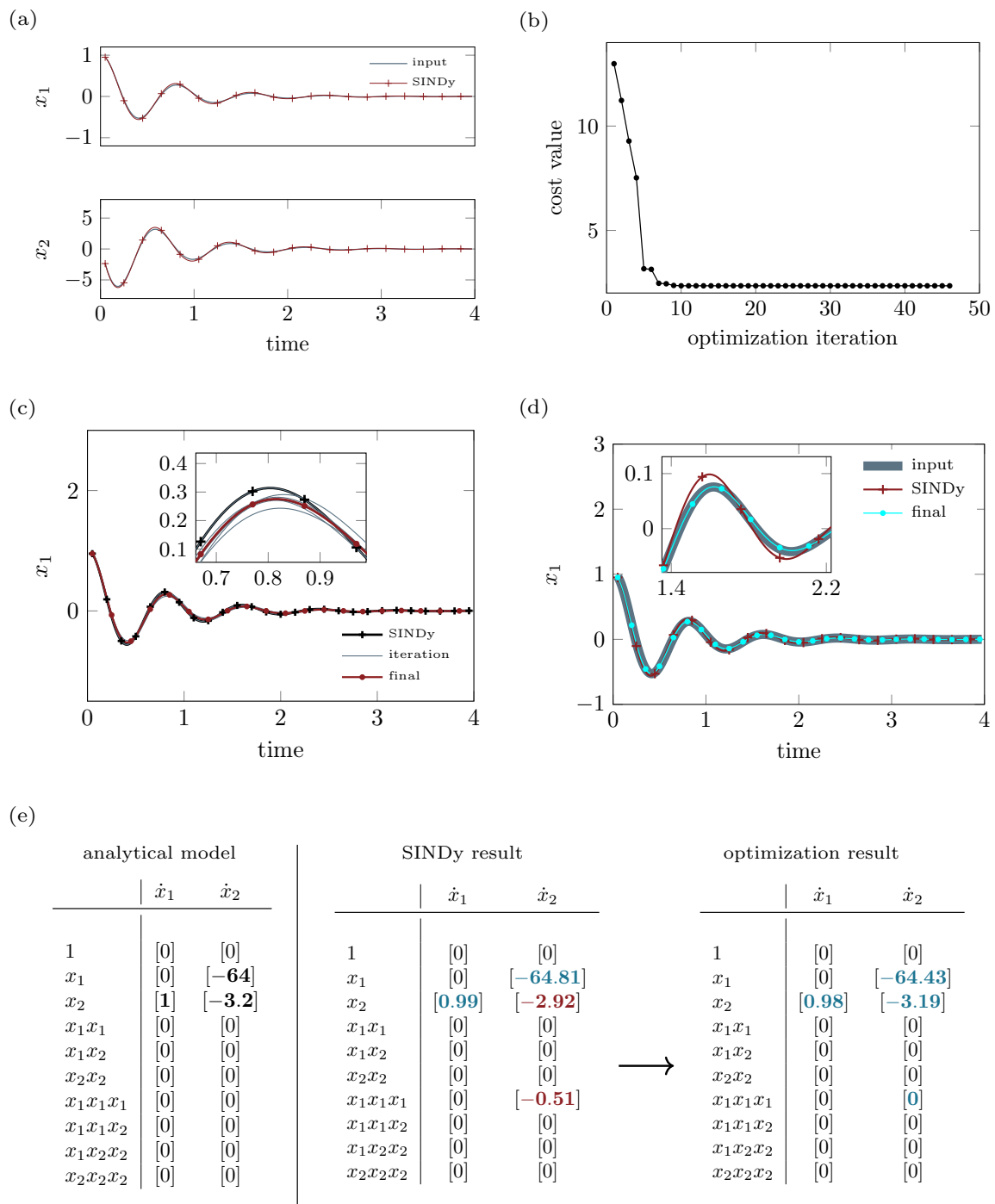


Figure 8. Two-step reconstruction procedure: (a) comparison of input data and SINDy reconstruction that used regularized numerical derivatives. (b, c) iterations of the bounded optimizer on the non-zero coefficients to reduce the difference between input and reconstruction signal. (d) result of the two-step reconstruction and (e) evolution of the coefficient matrices ζ obtained from SINDy and the optimization procedure in comparison to the analytical oscillator model that generated the initial input signals

The previous study involving both true model states \mathbf{x} exhibits a reconstruction error of 2.345. In the current study with less information, reconstruction errors of the same order of magnitude can be observed. As the polynomial degree is increased, the reconstruction error decreases slightly before it saturates for $p \geq 3$. The coefficient matrices exhibit predominantly constant entries beginning with the simplest model, where $p = 1$. As a result of this trend and the low reconstruction error, one can

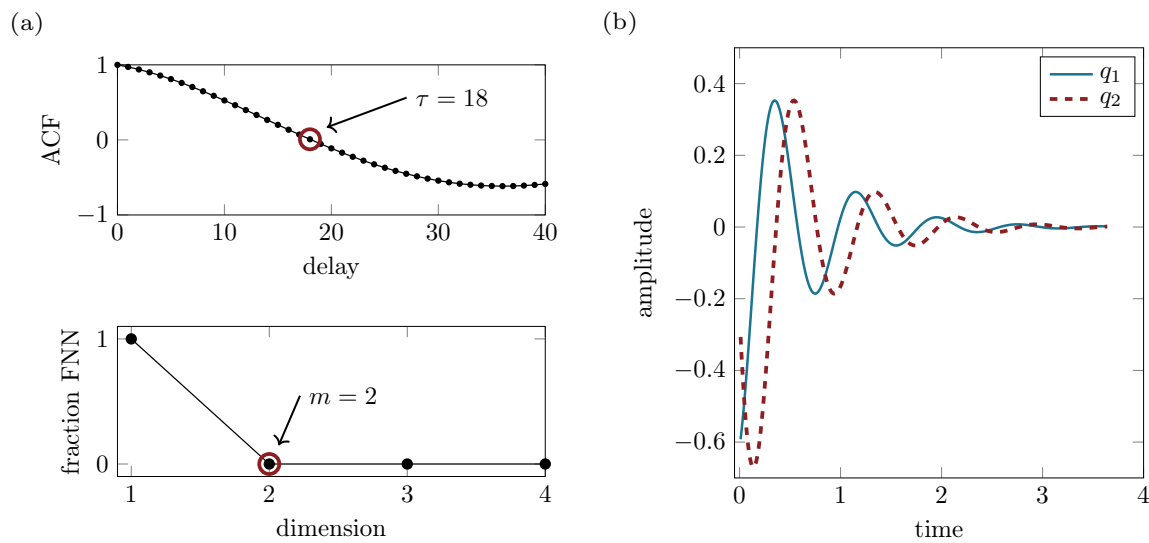


Figure 9. (a) determining the embedding parameters as first zero of the ACF and vanishing fraction of the FNN algorithm for x_1 . (b) resulting trajectories q_1 and q_2 in the reconstructed phase space

		1	q_1	q_2	$q_1 q_2$	$q_1^2 q_2$	q_2^3	other terms
$p = 1, \quad \Gamma = 1.88$	\dot{q}_1	0	-0.74	-5.84	-	-	-	-
	\dot{q}_2	-0.008	10.60	-2.45	-	-	-	-
$p = 2, \quad \Gamma = 1.82$	\dot{q}_1	0	-0.75	-5.82	-0.06	-	-	-
	\dot{q}_2	-0.007	10.62	-2.45	0	-	-	-
$p = 3, \quad \Gamma = 1.63$	\dot{q}_1	0	-0.81	-5.84	-0.13	0.51	0.16	-
	\dot{q}_2	0	10.60	-2.39	0	0	0	-
$p = 4, \quad \Gamma = 1.64$	\dot{q}_1	0	-0.85	-5.86	0	0.7080	0	$-0.96 q_1^3 q_2 - 1.44 q_1 q_2^3$
	\dot{q}_2	0	10.54	-2.36	0	0	0	0

Table 1. Coefficient matrices for optimized SINDy reconstruction for maximal polynomial degrees p and resulting reconstruction error Γ . The states q_1 and q_2 were taken from the time delay embedded signal x_1 and used as input signals, see Figure 9. Terms that are not displayed became zero in this analysis. The reconstruction for $p = 5$ failed with an error of $\Gamma = 159.66$ and 13 NZE after optimization

conclude that a reconstruction using only linear combinations of the monomials q_1 and q_2 is sufficient. However, since the original analytical model in this investigation was the simple linear oscillator, the resulting linear reconstruction is not very surprising. Still, the processes of embedding and numerical differentiation could have introduced artefacts that would have increased the model complexity after reconstruction. Depending on a specific objective, one may use the linear, or also the non-linear models that arise from this study. The latter models show a smaller reconstruction error, while the number of non-zero entries increases. This means that SINDy arrives at better reconstructions with an increasing number of terms and, as a result, greater model complexity. Thus, both the reconstruction error and the model complexity have to be balanced. For a specific application, either of those objectives may be favored.

In conclusion, SINDy is able to reconstruct dynamical models that reproduce the observed dynamics to a high degree, even if only a limited amount of input information is provided. In this case, a single state of the mechanical oscillator suffices to reconstruct a dynamical model that captures the

salient feature of the dynamics and exhibits low reconstruction errors for a linear model. This result is promising for the application of the proposed methods to real experimental impulse data, which offer limited access to the states of the system and thus require embedding approaches.

3.1. Balancing reconstruction error and model complexity

As illustrated in the previous study, reconstruction error and model complexity represent conflicting objectives when dealing with incomplete or noisy data. We classify those conflicting objectives in the framework of under- and over-fitting models, which is a classical framework in system identification and machine learning research. The number of non-zero coefficients in the reconstructed set of ODEs is a proxy for the model complexity. Hence, each SINDy model can be located in the plane spanned by model complexity and reconstruction error, see Figure 10 (a). Re-visiting the assumption of sparse dynamics, models of low complexity and low reconstruction error are the optimal solution to the time series based reconstruction challenge. A simple model with high reconstruction error does not capture all of the salient dynamics of the input and is thus under-fitted. Contrary, if high complexity is required to obtain a low reconstruction error, the model may over-fit the dynamics and is unlikely to yield a generally applicable result.

To display a sample case for selecting an optimal trade-off between model complexity and reconstruction error, we revisit the reconstruction from limited input data. In the previous study, the time delay embedding was computed from the first state x_1 of the analytical oscillator. In measurements, one will most likely measure a combination of the system's states rather than a single active degree of freedom. Thus, we study how the composition of the measured signal impacts the reconstruction error. The input $s(t)$ to the time delay embedding is now an aggregate of both states x_1 and x_2 according to $s(t) = \beta x_1 + (1 - \beta) x_2$, where the scaling parameter β prescribes the mixing ratio of both states. To study the effect of different input information, polynomial orders up to $p = 4$ are studied and β is varied in the range $[1, 0]$ with a step size of 0.05. The resulting reconstruction results are displayed in Figure 10 (b). The best reconstructions, consisting of low complexity models with low error, are found for $p = 3$. Lower polynomial orders result in simple but imprecise reconstructions, while $p = 4$ is prone to over-fitting. Depending on the specific purpose of the reconstruction, either configuration may be favorable. Still, the illustration of the resulting models in the complexity-error plane helps to find the optimal choice of parameters such as the number of candidate functions in the library Θ .

3.2. Model identification studies

For illustrative purposes, we simulate three qualitatively different model identification scenarios. First, data is generated for different system configurations driven by a control parameter. This control parameter is then identified in the reconstructed models. Second, reconstruction models are built for data that stems from the same system, but for different initial conditions. This case is designed to represent multiple measurements that contain different amounts of information about a given dynamical system. Finally, data are generated for a linear and weakly non-linear system configuration. Here, we illustrate the potential of the proposed methods to serve as a test for non-linearity.

3.2.1. Identification of parameter dependencies

In this study, the damping term δ is varied as a control parameter in the range of $\delta \in [0.1, 0.3]$. Optimized reconstruction models are created from each input time series and the non-zero coefficient matrix entries are recorded. As the reference system configuration, the standard value $\delta = 0.2$ is selected. Figure 11 depicts the relative change $\Delta\tilde{\zeta} = \tilde{\zeta}/\tilde{\zeta}_{\text{ref}}$ of each coefficient relative to the reference model. The two states x_1, x_2 are used with polynomial degree $p = 1$ such that six ODE coefficients result. It can be observed, that five out of those six coefficients remain constant throughout the control parameter variation, while only $\tilde{\zeta}_{2,3}$ changes significantly. Hence, the control parameter variation can be traced back to a single term in the reconstructed models. Consequently, this model parameter has to

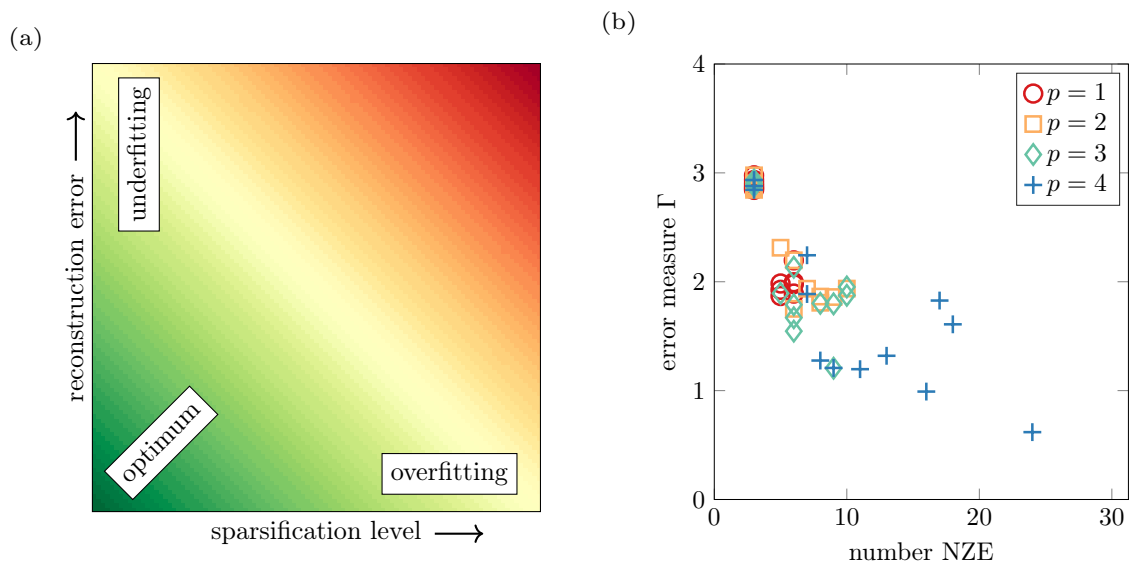


Figure 10. (a) balancing model complexity and reconstruction error: depending on a specific parameter selection, the reconstructed equations may represent under-, well-, and over-fitting models. (b) depicts the resulting model characteristics for various input configurations and varying polynomial orders of the non-linear functions library. In this case, most of the $p = 4$ configurations result in unsatisfactory reconstructions

be studied in order to investigate control parameter changes in simulations. Of course, this scenario is well-constructed to illustrate a simple parameter dependency. However, the concept of tracking model parameters along with control parameters can be transferred directly to more complicated and more realistic parameter identification challenges. Moreover, this procedure can be used in an inverse sense to identify the model parameters that are invariant under parameter changes.

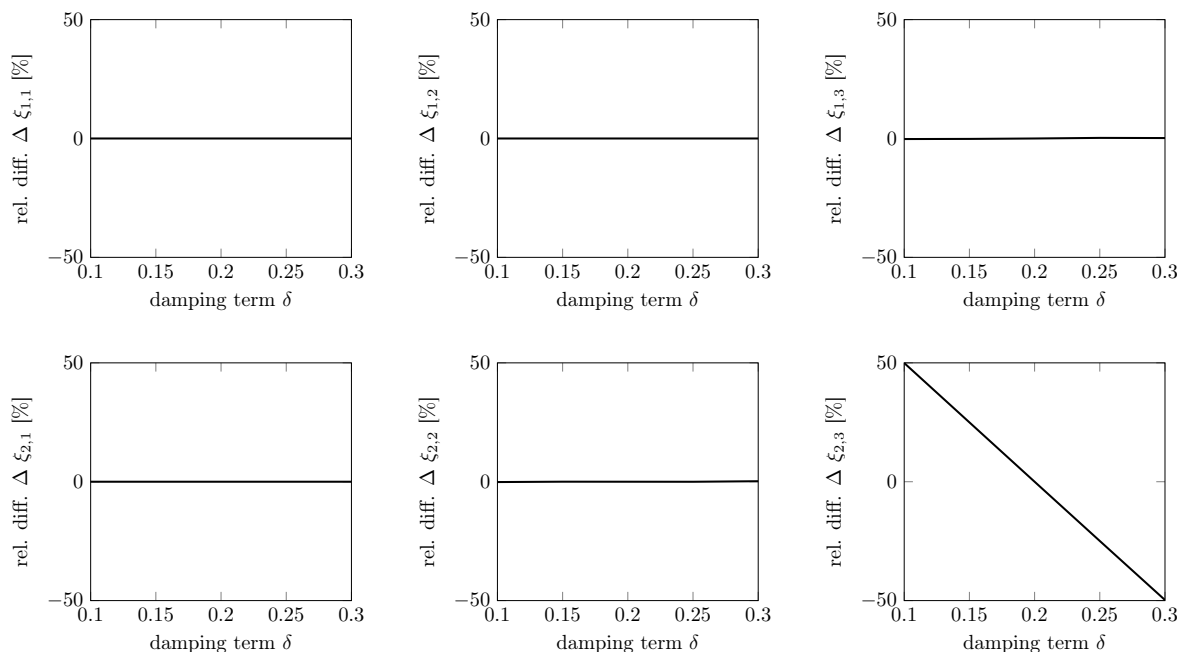


Figure 11. Relative deviation of the reconstructed ODE coefficients ξ_i from the reference configuration as function of the control parameter δ . As a reference, $\delta = 0.2$ is chosen to identify those terms that are invariant under the parameter change

3.2.2. Robust model parameter identification

Imagine testing a dynamical system using hammer impacts. Naturally, each excitation is different, but the system remains the same. We try to simulate this situation by varying the initial conditions for creating the input time series of the oscillator. Each realization stems from the same dynamical system, and thus we seek for one model reconstructed from all data. For each input data, an optimized SINDy model is computed and coefficient values are recorded. Figure 12 depicts histograms of the relevant coefficient matrix entries. In order to define the optimal model, one would select representative coefficient values, such as the medians, and form the final coefficient matrix. This procedure is different from the one proposed by Brunton et al. [12], where all data are stacked into a large data matrix and subjected to a single SINDy reconstruction. Due to the highly transient nature and resulting spectra of amplitudes of the data recorded in this work, we suggest deriving multiple models instead.

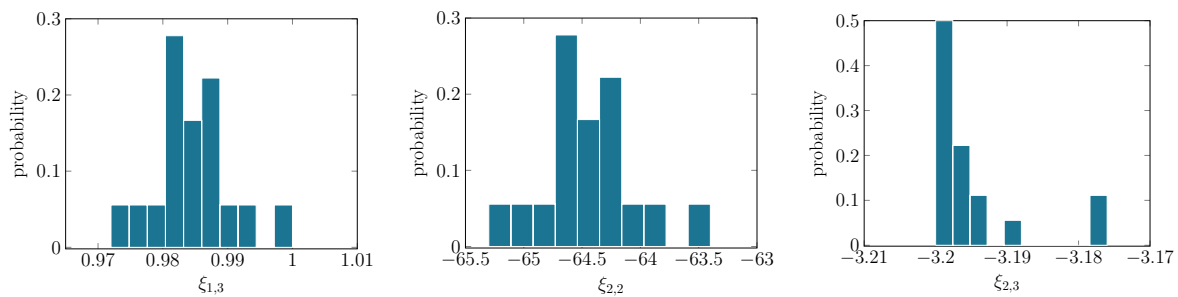


Figure 12. Histograms of the non-zero coefficient matrix entries resulting from 36 realizations of the linear oscillator for all possible combinations of the initial conditions $[x_1(0), x_2(0)] = [-1, -0.6, -0.2, 0.2, 0.6, 1]$ with highest polynomial order $p = 1$. The analytical model is given by $\xi_{1,3} = 1$, $\xi_{2,2} = -64$, $\xi_{2,3} = 3.2$. Thus, the reconstructions match well if the most frequent or median coefficient value is selected in the model identification process

3.2.3. Test for non-linearity

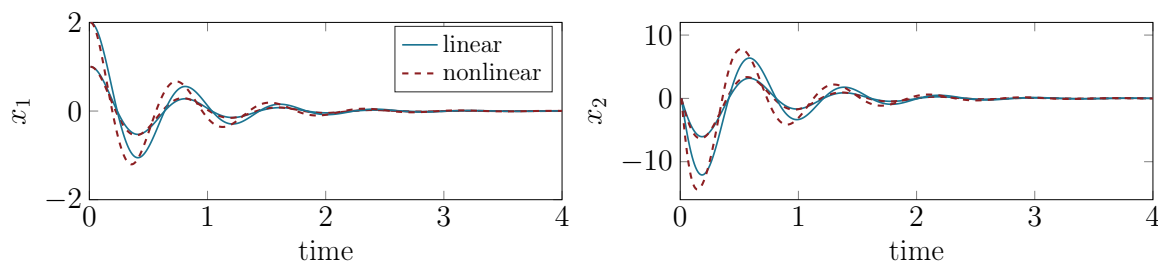
In order to extend the findings from the given mechanical oscillator to non-linear dynamics, we introduce an additional cubic stiffness term k_{nl} such that the governing equations read

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -cx_2 - kx_1 - k_{nl}x_1^3. \end{aligned} \quad (9)$$

The reference configuration is given by $k = 64$, $c = 3.2$ and, in the non-linear configuration, $k_{nl} = 10$, if not denoted differently. Figure 13 displays the analytical states x_1 , x_2 for both the linear as well as non-linear configuration of the 1-DOF oscillator. For small initial values, the non-linearity is practically inactive and only minimal differences to the linear configuration can be observed in time domain. However, the SINDy reconstruction clearly identifies a non-zero entry for the ansatz polynomial x_1^3 and thus allows for testing for non-linearity, even if the time traces are barely distinguishable in the time domain. Yet, while the cubic stiffness parameter is underestimated, the linear stiffness is overestimated. This result matches well with the understanding of linearization of the cubic stiffness for small displacement values. As the initial displacement is increased, the non-linearity becomes active and visible in time domain and the SINDy model exhibits a cubic stiffness term that is closer to that of the analytical model.

3.3. Feature generation for unsupervised time series classification tasks

Finally, we illustrate a sample case for using the SINDy reconstruction as feature generator. In various fields of research, the discrimination of time series data forms the central challenge for both



		1	x_1	x_2	...	x_1^3	...
linear config. $k_{nl} = 0, x_1(0) = 1$	\dot{x}_1	0 (0)	0 (0)	0.98 (1)	0	0 (0)	0
	\dot{x}_2	0 (0)	-64.60 (-64)	-3.19 (-3.2)	0	0 (0)	0
nonlinear config. $k_{nl} = 10, x_1(0) = 1$	\dot{x}_1	0 (0)	0 (0)	0.98 (1)	0	0 (0)	0
	\dot{x}_2	0 (0)	-65.32 (-64)	-3.19 (-3.2)	0	-5.41 (-10)	0
nonlinear config. $k_{nl} = 10, x_1(0) = 2$	\dot{x}_1	0 (0)	0 (0)	0.98 (1)	0	0 (0)	0
	\dot{x}_2	0 (0)	-65.31 (-64)	-3.19 (-3.2)	0	-8.86 (-10)	0

Figure 13. Comparison of linear and non-linear 1-DOF oscillators in time domain (top) and in terms of the optimized SINDy reconstruction models (bottom). The analytical model parameters are indicated by the coefficient values given in brackets in the table. Only non-vanishing terms are displayed here for a maximal polynomial order of $p = 3$

supervised as well as unsupervised classification tasks. To mimic the situation of an unsupervised classification task, we compute various ring-down time series data from varying model configurations. We then compute the reconstruction and employ the resulting ODE coefficients as features for describing the input data. These features are then used to solve the classification task in this setting. Particularly, the linear stiffness is varied over a wide range of values $k \in [36, 400]$, while the non-linear stiffness is kept constant for two different scenarios. First, the oscillator is considered in the linear configuration, i.e. $k_{nl} = 0$. Then, a constant cubic stiffness of $k_{nl} = 50$ is considered. For each of those configurations, 40 realizations with varying linear stiffness terms are created. Then, optimized SINDy reconstruction models are computed for each input set of time series. The highest polynomial order is set to $p = 3$ and both states x_1, x_2 from the analytical model are used as input to the reconstruction. All the resulting 80 time series are depicted in Figure 14. While it may seem easy to distinguish some of the linear system responses from the non-linear responses, most of the time traces overlap. Due to rather small amplitudes and strong damping, classical methods for describing and discriminating those dynamics may fail to cluster the data into two distinct groups representing the linear and non-linear configurations of the underlying model, as shown in the previous discussion. In particular, the non-linearity may be negligible for higher linear stiffness terms, which makes it nearly impossible to distinguish the system response from the raw time series data stemming from a purely linear system. We employ the proposed methods to derive optimized SINDy reconstructions and then collect the resulting ODE coefficients Ξ into a matrix. This matrix has 80 rows, each corresponding to one reconstruction, and 20 columns, each of which result from the coefficients ζ for the library of polynomials up to the third order for two states. For each observation, we treat the coefficients as discriminative features for describing the input time series data. The 20-dimensional feature space is reduced to two dimensions by principle component analysis (PCA). Hence, for each observation, two generalized features are given by the projection of the features onto the first and second principle components, see Figure 14. In this feature space, two distinct clusters form, which, since we know

the label for each observation, can be traced back to the underlying dynamical systems. The features created by our methods exhibit variance in mainly one direction for the linear oscillator, that is corresponding to the variation of the linear stiffness term. On the contrary, the features derived for the signals stemming from the non-linear systems show additional variance in the second principle component, which makes them distinguishable from the linear system features. We note that the linear stiffness has been varied over one order of magnitude for both cases in order to complicate the classification task. Still, our results illustrate that the unsupervised time series classification task can be solved using the SINDy reconstruction models as feature generator in a highly automated fashion. Naturally, this approach can be employed equally for supervised classification tasks, by assigning a label to an unknown input according to a-priori knowledge that was gathered from labeled training data.

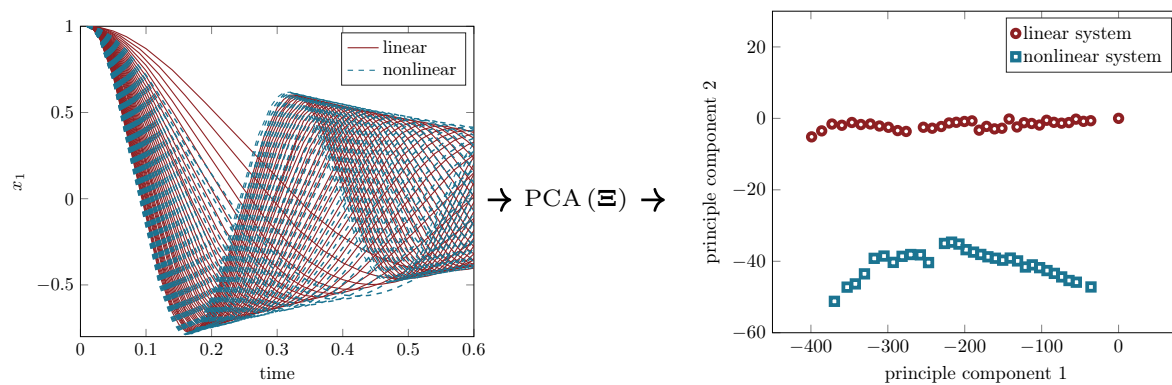


Figure 14. Left: 40 time series realizations for various linear stiffness terms $k \in [36, 400]$ for the a) linear configuration and b) the non-linear configuration with $k_{nl} = 50$. The resulting SINDy coefficient entries Ξ are used as descriptive features for classification of the input data. Using the principle component analysis, the feature space is reduced to two dimensions to find two clusters that can be traced back to the character of the underlying dynamical model that generated the input data

4. Conclusion

In this work, we propose several extensions to the SINDy method, which allows for reconstructing differential equations from time series input data. The reconstruction is sparse in the space of possible candidate functions and thus represents a minimal model for the observed dynamics. The main contribution of this work comprises of a highly automated parameter selection approach and of a sophisticated optimization procedure to fine-tune reconstructed models. Special focus is put on short and highly transient time series data as commonly obtained from impulse response and ring-down vibration measurements. Employing a 1-DOF oscillator, we illustrate the proposed extensions as well as common use-cases and challenges that may also arise and be observed in real-life laboratory data. These situations include limited data availability, noise contamination, and uncertain parameters of the dynamical system under study. The main findings and features of the proposed framework are summarized as follows.

- **Reconstruction of dynamical minimal models:** The sparse regression reconstructs systems of differential equations from time series data. Hence, these equations can be studied and analyzed by classical methods and provide detailed insight into the governing dynamics underlying an observation.
- **Model reconstruction for limited input data:** The proposed framework automates and optimizes the model reconstruction procedure while being suited well for accommodating limited data quality resulting from the amount of information, noise contamination, and unknown model dimensions.

- Test for non-linearity: The qualitative character of the underlying dynamical system can be estimated in terms of linearity and type and degree of non-linearity by inspecting the set of reconstructed differential equations.
- Model identification and model updating methods: The optimized reconstruction allows for identification of terms that depend explicitly on parameters that are prescribed or measured during experimentation. After identifying those terms in the reconstructed ODEs, uncertainty and bifurcation studies can be used in predictive modeling approaches to design safe and efficient structures without extensive testing.
- time series feature generation for classification and regression tasks: The reconstructed models represent features that are discriminative and possibly superior to classical time series features for uni-variate, short and highly transient input data.

Future research will use the proposed framework of methods to identify dynamically minimal models and changes the system underlying these minimal models taking real-life vibration measurements, such as hammer-impact testing during modal analysis of mechanical structures.

Author Contributions: conceptualization, M.S., S.O. and N.H.; software and investigation, M.S.; writing-original draft preparation, M.S.; writing- review and editing, S.O. and N.H.; supervision and project administration, S.O. and N.H.

Funding: This research was funded by Deutsche Forschungsgemeinschaft (DFG) within the priority program 'calm, smooth, smart' under grant HO 3852/21-1. MS acknowledges the financial support provided by SO over the UTS Centre for Audio, Acous- tics and Vibration (CAAV) international visitor funds.

Acknowledgments: The Matlab code provided by S. Brunton was used to generate the SINDy reconstruction. The authors would like to thank Edgar Hoover for his assistance in improving the clarity of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

DOF	Degree Of Freedom
NZE	Non-Zero Entries
ODE	Ordinary Differential Equation
SINDy	Sparse Identification of Non-linear Dynamics
TVRegDiff	Total Variation Regularized Numerical Differentiation

Appendix A. The filtering property of TVRegDiff numerical derivation schemes

Figure A1 illustrates the filtering property of the TVRegDiff scheme with appropriate choice of the regularization parameter for different signal-to-noise ratios.

References

1. Ondra, V.; Sever, I.A.; Schwingshackl, C.W. A method for non-parameteric identification of non-linear vibration systems with asymmetric restoring forces from a resonant decay response. *Mechanical Systems and Signal Processing* **2019**, *114*, 239-258, <https://doi.org/10.1016/j.ymssp.2018.05.010>.
2. Ondra, V.; Sever, I.A.; Schwingshackl, C.W. A method for detection and characterisation of structural non-linearities using the Hilbert transform. *Mechanical Systems and Signal Processing* **2017**, *83*, 210-227, <https://doi.org/10.1016/j.ymssp.2016.06.008>.
3. Pesaresi, L.; Stender, M.; Ruffini, V.; Schwingshackl, C. W. DIC Measurement of the Kinematics of a Friction Damper for Turbine Applications. *Dynamics of Coupled Structures, Volume 4, Conference Proceedings of the Society for Experimental Mechanics Series* **2017**, Springer, Cham, 93-101, https://doi.org/10.1007/978-3-319-54930-9_9.
4. Farrar, C.R.; Worden, K. An introduction to structural health monitoring. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* **2007**, *365*, 303-315, <https://doi.org/10.1098/rsta.2006.1928>.

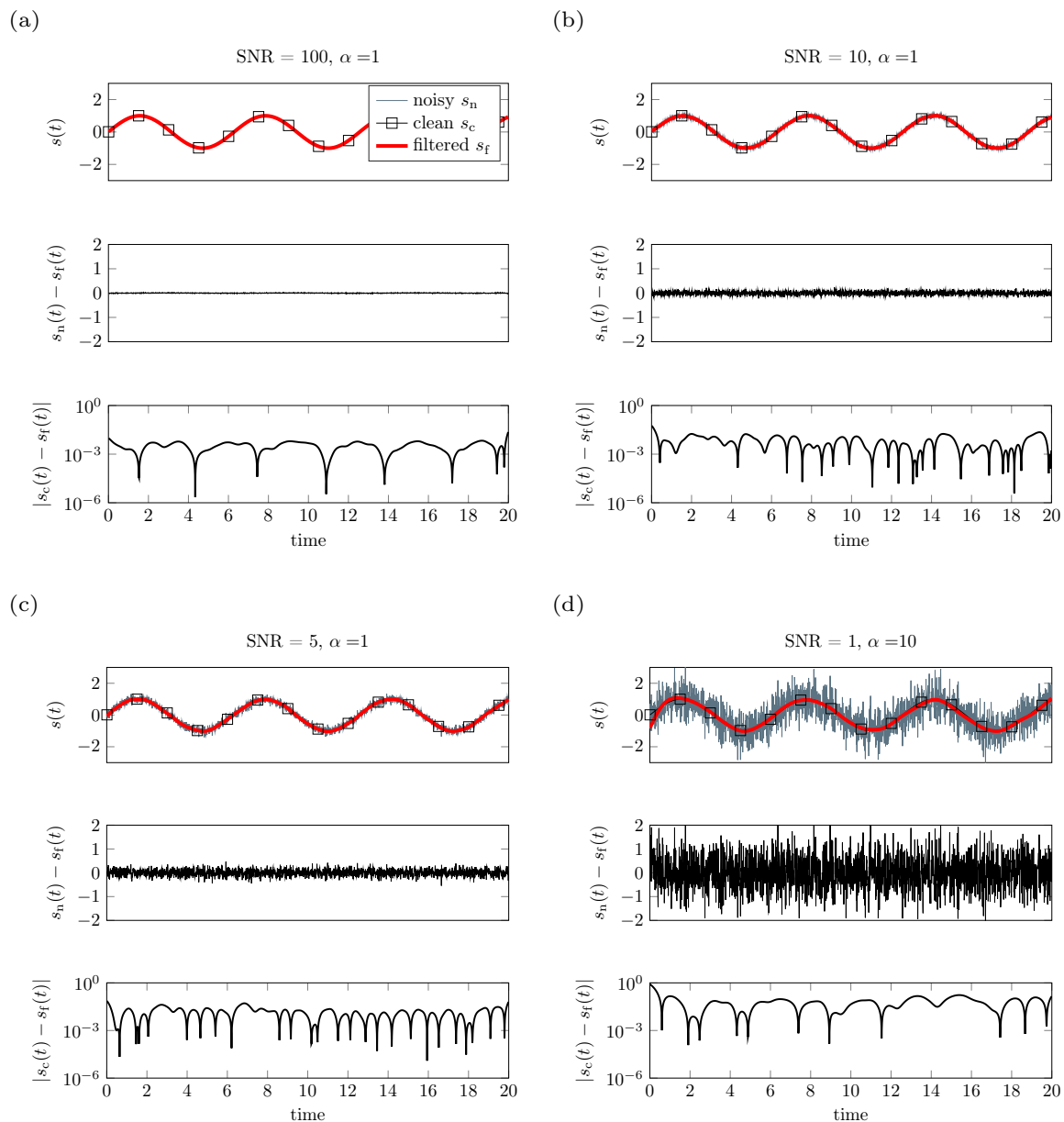


Figure A1. Filtering by integration of regularized derivatives: different noise levels applied to sinusoidal oscillation $s(t) = \sin(t)$. Clean signal is indicated by s_c , noisy (signal-to-noise ratio SNR) signal by s_n and filtered signal, that is obtained from time integration of the numerical derivative, s_f . Regularization parameters α employed for computing derivatives using TVRegDiff

5. Nuawi, M.Z.; Bahari, A.R.; Abdullah, S.; Ariffin, A.K.; Nopiah, Z.M. Time Domain Analysis Method of the Impulse Vibro-acoustic Signal for Fatigue Strength Characterisation of Metallic Material. *Procedia Engineering* **2013**, *66*, 539–548, <https://doi.org/10.1016/j.proeng.2013.12.106>.
6. Crutchfield, J.P.; McNamara, B.S. Equations of motion from a data series. *Complex systems* **1987**, *1*, 121.
7. González-García, R.; Rico-Martínez, R.; Kevrekidis, I.G. Identification of distributed parameter systems: A neural net based approach. *Computers & Chemical Engineering* **1998**, *22*, S965–S968, [https://doi.org/10.1016/S0098-1354\(98\)00191-4](https://doi.org/10.1016/S0098-1354(98)00191-4).
8. Raissi, M.; Karniadakis, G.E. Hidden physics models: Machine learning of nonlinear partial differential equations. *Journal of Computational Physics* **2018**, *357*, 125–141, <https://doi.org/doi:10.1016/j.jcp.2017.11.039>.
9. Gouesbet, G.; Letellier, C. Global vector-field reconstruction by using a multivariate polynomial L2 approximation on nets. *Physical Review E* **1994**, *49*, 4955–4972, <https://doi.org/10.1103/PhysRevE.49.4955>.

10. Wei-Dong, L.; Ren, K.F.; Meunier-Guttin-Cluzel, S.; Gouesbet, G. Global vector-field reconstruction of nonlinear dynamical systems from a time series with SVD method and validation with Lyapunov exponents. *Chinese Physics* **2003**, *12*, 1366.
11. Wang, W.X.; Yang, R.; Lai, Y.C.; Kovanis, V.; Grebogi, C. Predicting catastrophes in nonlinear dynamical systems by compressive sensing. *Physical review letters* **2011**, *106*, 154101, <https://doi.org/10.1103/PhysRevLett.06.154101>.
12. Brunton, S.L.; Proctor, J.L.; Kutz, J.N. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences of the United States of America* **2016**, *113*, 3932–3937, <https://doi.org/10.1073/pnas.1517384113>.
13. Rudy, S.H.; Brunton, S.L.; Proctor, J.L.; Kutz, J.N. Data-driven discovery of partial differential equations. *Science advances* **2017**, *3*, e1602614, <https://doi.org/10.1126/sciadv.1602614>.
14. Oberst, S.; Stender, M.; Baetz, J.; Campbell, G.; Lampe, F.; Morlock, M.; Lai, J.C.; Hoffmann, N. Extracting differential equations from measured vibro-acoustic impulse responses in cavity preparation of total hip arthroplasty. *15th Experimental Chaos and Complexity Conference*, **2018**, Madrid, <https://doi.org/10.13140/RG.2.2.28836.58240>.
15. Takens, F. Detecting strange attractors in turbulence. *Dynamical systems and turbulence* **1981**, Springer, pp. 366–381.
16. Kantz, H.; Schreiber, T. *Nonlinear Time Series Analysis*; Cambridge University Press: Cambridge, **2003**, <https://doi.org/10.1017/CBO9780511755798>.
17. Oberst, S.; Lai, J.C. A statistical approach to estimate the Lyapunov spectrum in disc brake squeal. *Journal of Sound and Vibration* **2015**, *334*, 120–135, <https://doi.org/10.1016/j.jsv.2014.06.025>.
18. Abarbanel, H.D.I.; Brown, R.; Sidorowich, J.J.; Tsimring, L.S. The analysis of observed chaotic data in physical systems. *Reviews of Modern Physics* **1993**, *65*, 1331–1392, <https://doi.org/10.1103/RevModPhys.65.1331>.
19. Kennel, M.B.; Brown, R.; Abarbanel, H.D.I. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical Review A* **1992**, *45*, 3403–3411, <https://doi.org/10.1103/PhysRevA.45.3403>.
20. Oberst, S.; Baetz, J.; Campbell, G.; Lampe, F.; Lai, J.C.; Hoffmann, N.; Morlock, M. Vibro-acoustic and nonlinear analysis of cadavric femoral bone impaction in cavity preparations. *International Journal of Mechanical Sciences* **2018**, *144*, 739–745, <https://doi.org/10.1016/j.ijmecsci.2018.05.051>.
21. Gilmore, R.; Lefranc, M.; Tuffiaro, N.B. The Topology of Chaos. *American Journal of Physics* **2003**, *71*, 508–510, <https://doi.org/10.1119/1.1564612>.
22. Chartrand, R. Numerical Differentiation of Noisy, Nonsmooth Data. *ISRN Applied Mathematics* **2011**, *2011*, 1–11, <https://doi.org/10.5402/2011/164564>.
23. Fulcher, B.D.; Jones, N.S. Highly Comparative Feature-Based Time-Series Classification. *IEEE Transactions on Knowledge and Data Engineering* **2014**, *26*, 3026–3037, <https://doi.org/10.1109/TKDE.2014.2316504>.
24. Bagnall, A.; Lines, J.; Hills, J.; Bostrom, A. Time-Series Classification with COTE: The Collective of Transformation-Based Ensembles. *IEEE Transactions on Knowledge and Data Engineering* **2015**, *27*, 2522–2535, <https://doi.org/10.1109/TKDE.2015.2416723>.
25. Fulcher, B.D.; Jones, N.S. hctsa: A Computational Framework for Automated Time-Series Phenotyping Using Massive Feature Extraction. *Cell systems* **2017**, *5*, 527–531.e3, <https://doi.org/10.1016/j.cels.2017.10.001>.
26. Marwan, N.; Romano, M.C.; Thiel, M.; Kurths, J. Recurrence plots for the analysis of complex systems. *Physics Reports* **2007**, *438*, 237–329, <https://doi.org/10.1016/j.physrep.2006.11.001>.
27. Stender, M.; Papangelo, A.; Allen, M.; Brake, M.; Schwingshackl, C.; Tiedemann, M. Structural Design with Joints for Maximum Dissipation. *Shock & Vibration, Aircraft/Aerospace, Energy Harvesting, Acoustics & Optics, Volume 9*, Springer, **2016**, pp. 179–187, https://doi.org/10.1007/978-3-319-30087-0_16.
28. Srivastava, N.; Hinto, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, **2014**, *15*, 1929–1958.