Data-Driven Techniques for the Nonlinear Dynamics of Mechanical Structures

Vom Promotionsausschuss der
Technischen Universität Hamburg

zur Erlangung des akademischen Grades

Doktor-Ingenieur (Dr.-Ing.)

genehmigte Dissertation

von
Merten Stender

aus
Eutin

2020

Mündliche Prüfung am: 26.10.2020

Gutachter: Prof. Dr. rer. nat. habil. Norbert Hoffmann
Prof. Dr.-Ing. Utz von Wagner

Preface

This dissertation was conducted under the supervision of Professor Norbert Hoffmann at Hamburg University of Technology between October 2016 and May 2020. I would like to thank my dissertation committee members, and especially Professor von Wagner, for realizing the last chapter of this project in such challenging times.

I am extremely grateful to a large number of people who supported me on this journey. Amongst all of them, I would like to thank some of them in particular.

My colleagues and friends at AUDI in Ingolstadt, thank you for providing data, advice, and for always making sure that chaos-theoretic thinking needs to be grounded to everyday-engineering business in order to let research truly make an impact. Thanks to my flatmates for great times off the computer keyboard.

Merten Tiedemann, I am particularly grateful for sending me on this road and for being a role model throughout. You taught me to finish things off and always keep the 80/20 rule in mind. Sebastian Oberst, thanks for challenging my results, spending much time on reviewing my articles, starting inspiring discussions, and for hosting me in Sydney for a remarkable research stay.

Norbert, thank you for sharing your immense knowledge, providing guidance and - most importantly - giving me the freedom to take two steps into the foggy dark of new research domains. I truly enjoyed your outstanding ability to think more than one step ahead, our enthusiastic discussions and your view for the bigger picture.

My friends Leon, Thilo and the Lerncamp crew, thank you for making my studies an easy ride, for always being there for almost ten years now, and for countless memories of good times in Hamburg.

To my parents: you laid the foundations for this dissertation through teaching me discipline and persistence, by encouraging action over talking, and for always believing in me. Thank you, and also to my brothers and family, for your unconditional support and wise counsel.

And finally Anne: thank you for your unfailing encouragement, for constantly pushing my own standards, for ensuring a healthy work-life balance, and for your unwavering support! I could not have done it without you.

Kurzfassung

Trotz vielfältiger Fortschritte im Bereich der System-Identifikation stellen komplexe Schwingungen von multi-physikalischen, multi-skaligen und gefügten mechanischen Systemen auch heute noch Herausforderungen für die numerische Simulation dar. Insbesondere Dämpfungs- und Reibungseffekte sind aufgrund ihrer Unzugänglichkeit für Messungen und ihrer Variabilität der Grund für unzureichende Modellierungsansätze. Gleichzeitig stehen heutzutage die datengetriebene Modellierung und die Datenwissenschaften immer mehr im Fokus von Wissenschaft und Gesellschaft. Diese Thesis präsentiert neue Ansätze für die Anwendung von datengetriebenen Verfahren in der mechanischen Strukturdynamik. Die Potentiale und Limitierungen solcher Methoden werden anhand von Anwendungsstudien diskutiert. Diese Studien erstrecken sich von Ein-Freiheitsgrad-Schwingern bis hin zu komplexen Bremssystemen von Automobilen. Insbesondere stehen die physikalische Konformität der erlernten Modelle sowie deren schwingungstechnische Interpretation im Fokus dieser Arbeit.

Abstract

In structural vibrations, several modeling approaches have helped to develop better, i.e. safer, less-vibrating and more controllable designs for machines and mechanical structures. However, complex nonlinear vibrations of multi-physics, multi-component, and multi-scale systems still represent challenges to today's identification and vibration prediction approaches. Particularly, damping and friction can play a crucial role for vibration mitigation while being inherently difficult to characterize, quantify or even approximate. At the same time, the data sciences have become omnipresent not only in different fields of science, but also in society. This thesis introduces a rigorous framework and discusses chances and limitations for using machine learning in complex structural dynamics. Special focus is put on the physical peculiarities of the vibration signals and physics-informed learning. In the context of case studies, new methodologies are presented for several systems ranging from single-degree-of-freedom oscillators to complete automotive disk brake systems.

Contents

List of Figures

List of Tables

List of Symbols

Preliminary remarks

$\boldsymbol{x}$	small symbols printed in bold font denote vectors
$\boldsymbol{X}$	capital symbols printed in bold font denote matrices
$[]^\top$	transpose
$\dot{[]}$	an overdot denotes the first derivative with respect to time
$\ddot{[]}$	a double overdot denotes the second derivative with respect to time
$\bar{[]}$	a bar placed over a symbol denotes an averaged quantity
$\tilde{[]}$	a tilde over a symbol denotes a normalized quantity
$\hat{[]}$	a hat over a symbol denotes effective values and machine learning predictions

If different notation than the one given hereafter is used, this is explicitly indicated in the respective sections of the work.

Latin notation

A	amplitude
b	bias value
c	viscous damping coefficient
C	confidence value
D	dimension
e	Euler's number (2.71828 ...)
$\boldsymbol{f}$	general dynamical system
f	frequency
F	force
h	humidity
i, j	iteration indices
J	loss function, error
$\boldsymbol{J}$	Jacobian matrix
k	stiffness

L	recurrence plot line length
m	state space dimension
M	mass
n	iteration index
N	normal load
P	probability distribution
q	general coordinate
r	radius
$\boldsymbol{R}$	recurrence plot
$\boldsymbol{s}$	univariate time series
$\boldsymbol{S}$	multivariate time series
t	time
T	temperature
u	displacement
v	velocity
w	window
$\boldsymbol{W}$	weight matrix
$\boldsymbol{x}$	general state vector
x, y, z	general states or coordinates

Greek notation

α	general parameter
δ	real part of an eigenvalue
Δ	difference
ϵ	small parameter ($\ll 1$)
γ	scaling parameter
Γ	envelope function
λ	Lyapunov exponent, eigenvalue
Λ	Lyapunov spectrum
μ	friction coefficient
Ω	rotational velocity
ω	(natural) frequency
ϕ	feature, feature extraction function
Φ	moment of inertia
ψ	mother wavelet
ρ	similarity index
σ	activation function

τ	embedding delay value
$\boldsymbol{\Xi}$	sparse regression coefficient matrix
ξ	SINDy coefficient, loading quantity
ζ	damping value

Other symbols

$\Im$	imaginary part
$\Re$	real part
$\mathscr{M}$	machine learning model
$\mathscr{D}$	data set

Indices

symbol	**meaning**
$[]_0$	initial
$[]^*$	fixed point
$[]_{\text{amb}}$	ambient
$[]_{\text{end}}$	end
$[]_{\text{ext}}$	external
$[]_{\text{fluc}}$	fluctuation
$[]_{\text{fluid}}$	fluid
$[]_{\text{irreg}}$	irregular
$[]_{\text{lin}}$	linear
$[]_{\text{L}}$	Lorenz system
$[]_{\text{max}}$	maximum
$[]_{\text{min}}$	minimum
$[]_{\text{nl}}$	nonlinear
$[]_{\text{reg}}$	regular
$[]_{\text{rel}}$	relative
$[]_{\text{rot}}$	rotational, rotor
$[]_{\text{s}}$	sampling
$[]_{\text{sl}}$	slip
$[]_{\text{start}}$	start
$[]_{\text{st}}$	stick

$[]_{\mathrm{t}}$	tangential
$[]_{\mathrm{test}}$	test
$[]_{\mathrm{train}}$	training
$[]_{\mathrm{val}}$	validation
$[]_{\mathrm{x,y,z}}$	in x, y, z direction

List of Abbreviations

abbreviation	meaning
ACC	acceleration
AI	artificial intelligence
AMI	auto-mutual information function
ANN	artificial neural network
AP	average precision
ARMA	auto-regressive moving average model
AUC	area under curve
CAE	computer aided engineering
CEA	complex eigenvalue analysis
CNN	convolutional neural network
COCO	common objects in context
CV	computer vision
CWT	continuous wavelet transform
DET	determinism
DL	deep learning
DT	decision tree
EMD	empirical mode decomposition
ENTR	recurrence plot entropy
EP	equilibrium solution
E-PMC	extended periodic motion concept
FAN	fixed amount of neighbors norm
FC	fully connected network
FEM	finite element method
FFT	fast Fourier transform
FIV	friction-induced vibrations
FN	false negative
FNN	false-nearest neighbors
FP	false positive
GPU	graphics processing unit
GRU	gated recurrent unit
HBM	harmonic balance method
HHT	Hilbert-Huang transform

IMF	intrinsic mode function
IoU	intersection over union
KS	Kolmogorow-Sinai (entropy)
LAM	laminarity
LC	limit cycle solution
LDV	laser doppler vibrometry
LSTM	long-short-term memory network
mAP	mean average precision
MCC	Matthews correlation coefficient
ML	machine learning
MLP	multi-layer perceptron
NNM	nonlinear normal modes
NTSA	nonlinear time series analysis
NVH	noise, vibration, harshness
NZE	non-zero elements
ODE	ordinary differential equation
OPTME	obtain, pre-process, transform, model, explain
OSEME	obtain, scrub, explore, model, explain
PCA	principle component analysis
PRC	precision recall curve
R-CNN	region based convolutional neural network
ReLu	rectified linear unit
R-FCN	region based fully convolutional network
RNN	recurrent neural network
ROC	receiver operating characteristic
ROM	reduced order modeling
RP	recurrence plot
RPN	region proposal network
RPQA	recurrence plot quantification analysis
RR	recurrence rate
SINDy	sparse identification of nonlinear dynamics
SP(L)	sound pressure (level)
SSD	single-shot detection
STFT	short-time Fourier transform
SVM	support vector machine
TN	true negative
TP	true positive
TVRegDiff	total variation regularized numerical differentiation
UPO	unstable periodic orbit

1 Introduction

Engineering dynamics and vibrations of mechanical structures have a long history of modeling approaches that matured from single-body equations in the ages of Newton, to multi-body considerations, and nowadays to complex numerical models, all based on first principles of physics. However, there are mechanical systems that have not yet benefited as much from the simulations, since the simulated behavior significantly differs from the one observed in experimental testing and during operation. Examples for this class of systems can be found in friction-affected systems, such as brakes, clutches and drill-strings, systems affected by multi-physics and transient loads, such as renewable energy converters, systems composed of a large number of components, such as turbines, and many more. For those systems, the predictive power of numerical simulations, mostly assuming linearity and time invariance, must be considered weak even today. Consequently, system identification strategies have focused on learning about unknown and uncertain system elements from experimental measurements for updating the corresponding numerical models. These approaches work well for weakly nonlinear systems with regular vibration behavior. Nonetheless, there is a significant number of systems and phenomena for which this type of physics-based bottom-up modeling is lacking acceptable quality levels.

For centuries, the natural and life sciences have relied on evidence-based techniques to draw conclusions about natural systems from observations, i.e. pursuing top-down approaches. Here, the physical complexity, sparse observations and the number of unknown subsystems with hidden links encourage a data- or measurement-driven research approach rather than building bottom-up analytical models by hand. The related scientific disciplines have developed a vast corpus of signal processing, data assimilation and evidence-based methods. Recently, computing power has enabled data-driven methods framed by terms of Machine Learning (ML) and Artificial Intelligence (AI) to allow for processing of massive amounts of data and for the extraction of hidden patterns. Owing to their versatility, those novel data-driven methods have enabled breakthroughs in various disciplines such as medicine, astro-physics, climate research, and many more. Interestingly, data-driven approaches, being 'bread-and-butter' techniques in the natural sciences, have only had limited impact in structural dynamics until today.

This thesis aims at answering three major questions that arise from the situation described above. First, is the present disambiguation of complex natural systems and complex mechanical structures in terms of methodical approaches, dynamic behavior and system understanding justified? Second, what is required to employ modern evidence-based and data-driven approaches in nonlinear structural dynamics? And third, what are possible applications to exemplary evaluate chances and limitations of those novel techniques in structural dynamics of mechanical structures? These research motives are addressed by various interdisciplinary techniques for sequential data, which is the most common form of measurement recordings in mechanical systems. Focus is put on deterministic dynamics of friction-affected structures. The analysis incorporates regular as well as irregular dynamics and related transitions. Furthermore, multi-scale and non-stationary behavior is taken into account. In the field of machine learning, methods for explicit consideration of sequences and time dependency are evaluated in respect to their compatibility and potential benefit. After all, eligibility for the usage in an engineering-compliant environment is regarded as a key factor for potential benefits of the proposed methods.

This thesis is structured as follows: In Chapter 2, today's challenges in mechanical vibrations are outlined by revising some of the core characteristics of nonlinear dynamics and related phenomena. The observed dynamic behavior is compared to nonlinear dynamics in the natural sciences to present commonalities and differences, especially with respect to the notion of complexity. Furthermore, a generalized framework for a consistent system description is developed that sets the base for a data-driven treatment of vibrating mechanical structures. Chapter 3 illustrates how a classical data science process can be adapted for the use in nonlinear structural dynamics in an engineering-compliant environment. A systematical data science process for structural vibrations is presented. A high-level overview on dynamics-specific methods is given that builds on the domain knowledge developed in Chapter 2. Potential pitfalls and approaches to explaining black-box data-driven models are indicated. Chapter 4 depicts the use of the novel data science process for structural dynamics. Specifically tailored data-driven methods are applied to six application cases from nonlinear structural vibrations. The data-driven methodology is applied to system identification, vibration detection, vibration characterization and vibration prediction in various real mechanical systems. Most of these systems are affected by friction that gives rise to various aspects of dynamical complexity. Chances and limitations of the novel data science process and the related techniques are evaluated. Finally, Chapter 5 concludes the results of the thesis and gives outlines for future work.

2 Complex mechanical vibrations

In this chapter, a selection of characteristics of mechanical vibrations is presented that render systems and their dynamics complex. Namely, nonlinearity, irregularity, dynamic transitions, and dynamics at multiple scales are re-visited. Dynamical and structural complexity is discussed and identified in mechanical systems. In the first part, special focus is put on the class of mechanical systems that are affected by friction. In the second part, aspects of complex dynamic behavior are shown to exist in small and well-known analytical models. The third part closes this chapter with a generalized system framework for describing and modeling complex machine dynamics using data-driven methods. The rather extensive review of nonlinear vibrations is considered to be of fundamental importance: only the knowledge about the wide range of dynamical phenomena and their characteristics will allow to select appropriate methods for a data-driven system identification that is at the heart of this thesis.

2.1 Nonlinear and transient vibrations

'Nonlinearity is generic in nature, and linear behavior is an exception' as stated by Kerschen et al. [1][1]. Linear behavior can be assumed for some mechanical structures that undergo only small deflections. However, the ever-increasing demand for lighter structures results in more flexibility, and thus nonlinear behavior. Other systems, such as those affected by friction, exhibit strongly nonlinear characteristics especially for small-amplitude motions. As a result, nonlinearity gives rise to a plethora of different phenomena in dynamical systems. This section begins with a brief historical introduction to nonlinear dynamics before the core aspects of regular and irregular motion, self-excitation and spatially distributed dynamics are discussed in more detail.

Initiated by Kepler and Newton [2] for the two-body problem, Poincaré [3] studied nonlinear dynamics using topographical quantities and, nowadays most famously, introduced the idea of the Poincaré section. Also, he discovered dynamics with sensi-

[1] Alternatively, a humorous phrase says: 'to divide all systems into linear and nonlinear is like dividing the world into bananas and non-bananas'

tive dependence on initial conditions - today known as *chaos*. However, his findings remained without much impact during his life time. It is Lorenz in 1963 [4], who can be considered the founder of the modern research discipline of chaotic dynamics. For the first time making use of an early computer, he showed the sensitivity on initial conditions of the Rayleight-Bénard convection model, now referred to as the *Lorenz system*, and thus established the well-known *butterfly effect*. Feigenbaum [5] studied the bifurcation patterns of iterative maps and identified the period-doubling bifurcations as a route to chaos in the simple map. Next, Mandelbrot [6] played a crucial role in the discovery of fractals, i.e. geometric object that exhibit self-similarity. Interestingly, solving the equations for the famous three-body problem has remained practically unsolved until today due to its chaotic character. This thesis is concerned with phenomena and analysis methods for *deterministic dynamics*. Here, determinism denotes the evolution of the dynamical system's states to follow explicit rules that do not contain any stochasticity. Given exact knowledge about all present states, the mathematical description of the system dynamics would allow to compute the next state in time exactly. Since knowledge is always limited and afflicted with uncertainty, and our world is not fully deterministic, predictions are only valid for short times. Furthermore, the dynamics of completely deterministic systems are not necessarily simple. The following sections illustrate how strictly deterministic systems can exhibit complicated dynamics with seemingly random characteristics. The reader is referred to the rich corpus of literature, such as the seminal textbooks [7, 8, 9, 10, 11], for more detailed discussions on nonlinear systems and their dynamics.

Generally, mechanical systems can be categorized into *forced*, *parametrically-excited*, and *self-excited* systems from a vibrations point of view. Forced systems experience an external source of energy that prescribes central aspects of the system dynamics. Typical examples include systems with periodic forcing from rotational motors and engines. In nonlinear systems of this class, periodic motion with phenomena of subharmonic and superharmonic responses and internal resonances are well-known [1]. Parametrically-excited systems exhibit fast varying, and mostly periodic, coefficients that can produce large vibrations even far from the system's natural frequencies. On the contrary, self-excited systems develop their individual dynamics in the absence of an explicit external forcing. Energy is transferred into self-excited motion through positive feedback mechanisms between the structural motion and a source of energy, such as a friction interface or aero-elastic forces from an airstream. Naturally, also combinations of those different system classes exist. In forced self-excited systems, a variation in forcing amplitude can cause amplitude reduction, so-called *quenching*, while a forcing frequency variation can cause synchronization, i.e. entrainment of the self-excited response.

In this thesis, deterministic dynamical systems $\boldsymbol{f}$ of the form

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \dot{\boldsymbol{x}}, t) \quad \boldsymbol{x} \in \mathbb{R}^m \tag{2.1}$$

are considered, where $\boldsymbol{x}(t)$ denotes the m-dimensional state and $\dot{\boldsymbol{x}}(t)$ denotes its temporal derivative. The states are represented in the state space[2]. The temporal evolution of a dynamical system along a path in the state space is denoted as a *trajectory*. In the range of non-equilibrium responses, period-1 motion $\boldsymbol{x}(t) = \boldsymbol{x}\left(t + \bar{t}\right)$ with period $\bar{t}$ is at the lower end of dynamic complexity. Period-n motion $\boldsymbol{x}(t) = f(\omega_1 t, \ldots, \omega_n t)$ is observed if the harmonic components ω_i of the response are rationally commensurable, i.e. if $k_1\omega_1 + k_2\omega_2 + \cdots + k_n\omega_n = 0$ for non-zero integer coefficients $k_i \in \mathbb{Z}^*$, $i = 1, \ldots, n$. On the contrary, quasi-periodic motion on a n-torus is observed for rationally incommensurable frequencies ω_n, i.e. $k_1\omega_1 + k_2\omega_2 + \cdots + k_n\omega_n \neq 0$ for any non-zero integer k_i. A limit cycle denotes periodic motion on a closed trajectory that is isolated, i.e. neighboring states either converge to, or diverge from, this solution [10]. Forced systems without self-excitation and linear systems cannot exhibit limit cycles, since closed trajectories are not isolated in these systems. *Regular motion* is a summary term for the motion on a periodic or quasi-periodic attracting set. Regular motion is characterized by negative semi-definite Lyapunov exponents. Regular dynamics have been studied extensively until Ruelle and Takens [12] reported an object in the state space that is globally attracting, but shows locally repelling trajectories with erratic dynamic behavior. They introduced the term of a *strange attractor* for this object, even though Lorenz [4] had already discovered such an object before. While irregular motion is also possible in conservative systems, strange attractors can only exist in dissipative systems that exhibit contracting phase space volumes. The following sections cover the definition of the term of *chaotic motion* and various transitions from regular dynamics to chaotic dynamics.

2.1.1 Irregular motion, friction and self-excitation

Deterministic systems can exhibit irregular motion that corresponds to the observation of *chaos*. Chaotic motion is characterized by *sensitive dependence on initial conditions* and *aperiodic* trajectories. Throughout this thesis, the terms of *chaotic* and *irregular* motion will be used in an interchangeable way for indicating deterministic dynamics that exhibit sensitive dependency on initial conditions. If a set is approached by the trajectories, it is termed the *strange attractor* [12]. A strange attractor exists if the system dynamics are bounded by non-smooth manifolds with complex,

[2] Typically, the term *phase space* refers to continuous systems (also called 'flows'), while the term *state space* refers to discrete or discretized systems [10].

i.e. strange, topology. The sensitive dependence on initial conditions causes two trajectories starting in very close vicinity to remain close only for a short time before they diverge at an exponential rate, which is measured by the maximal *Lyapunov exponent*. The Lorenz system [4]

$$\begin{aligned} \dot{x} &= \sigma\left(y - x\right) \\ \dot{y} &= x\left(R - z\right) - y \\ \dot{z} &= xy - \beta z \end{aligned} \tag{2.2}$$

is one of the most-studied dynamical systems that exhibit chaotic motion. The simplistic mathematical model features only two nonlinear terms, yet it shows a wide range of regular and irregular dynamical responses. In the standard configuration for $\sigma = 10$, $R = 28$, and $\beta = {}^{8}/_{3}$, the system exhibits a chaotic response that creates the typical butterfly-alike attractor in the state space, see Figure 2.1. As the attractor is

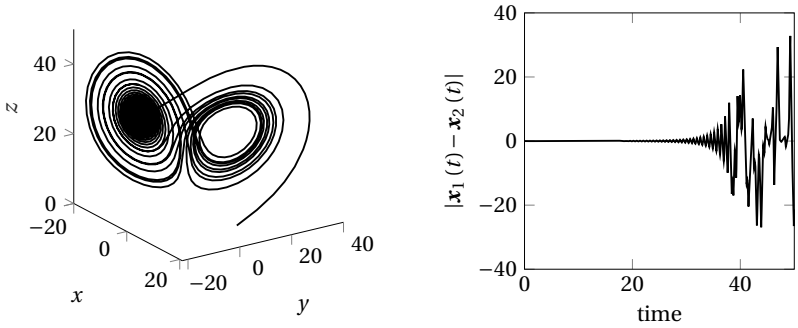


Figure (2.1) (a) trajectory of the standard Lorenz system in Equation (2.2) starting from $\boldsymbol{x}_1 = \left[x, y, z\right]^\top = [10, 10, 0]^\top$ and converging to a strange attractor. (b) Euclidean difference between two trajectories starting nearby at $\boldsymbol{x}_1$ and $\boldsymbol{x}_2 = [10, 10.1, 0]^\top$

bounded, also the difference between two nearby starting trajectories is bounded after the initial phase of exponential divergence. Practically, chaotic behavior means that long-term predictions are not possible, and that the prediction of the next state in time becomes exponentially harder as time passes [13]. As a result, and following the paradigm of quantum mechanics and sensitive dependency on initial conditions, one may have to get familiar with a world in which even more precise and massive measurements will not allow for a precise prediction of future states in chaotic dynamical systems, even in the complete absence of random elements. Research on nonlinear dynamical systems has revealed a plethora of different routes to chaos, i.e.

the transition from regular to irregular dynamics through parametric changes to the dynamical system [14]. The period doubling route to chaos proposed by Feigenbaum [15, 16] involves a cascade of pitchfork bifurcations from period-2 orbits to period-4 orbits and so forth until chaos emerges [17]. The bifurcation rate, i.e. the shrinking distance between period doubling bifurcation points, is given by the universal Feigenbaum constant [18]. On the contrary, the intermittency route to chaos introduced by Pomeau and Manneville [19] is characterized by a sudden transition of a periodic orbit to a chaotic attractor, also called *hard chaos* [20]. Here, the trajectories exhibit regular behavior, i.e. laminar motion, for a long time, before a sudden burst initializes irregular motion. Examples for this kind of transitions can be found in the Lorenz system. The *Ruelle-Takens-Newhouse route to chaos* [21] considers a cascade of Hopf bifurcations as the root cause for turbulent flow at increasing Reynolds numbers. A periodic orbit transits into a quasi-periodic orbit, i.e. a 2-torus, which in turn transits to a 3-torus and finally chaos arises through arbitrarily small perturbations. In the case of the *crisis route to chaos* [22], a chaotic transient turns into a chaotic attractor by the duration of the chaotic transient approaching infinity. Even more routes to chaos are likely to exist in the wide range of nonlinear dynamic systems.

Attractors, whether strange or regular, are typically characterized by dynamical and topological properties that are rooted in ergodic theory [23]. Ergodic theory assigns statistical quantities, so-called *invariant measures*, to the deterministic long-term dynamics on the attractor. Exemplary, the Lyapunov spectrum measures the average divergence of nearby starting trajectories in each state space dimension for dissipative systems. Thus, chaotic systems exhibit at least one positive exponent. Several entropy metrics, such as the Shannon entropy [24] and the more general Kolmogorow-Sinai (KS) entropy [25], are measures for the degree of order [26] and for the amount of information that is required for predicting the future of a dynamical system. As a result, regular dynamics have a vanishing KS entropy. Most dynamical invariant measures are temporal averages, not spatial averages, and hence are based on a single trajectory. Ergodic theory investigates if the invariant measures are independent from the spatial dimension, i.e. the choice of initial conditions and the choice of the coordinate system. More practically, invariant measures represent a direct way to compare (strange) attractors and measure their dynamical and geometrical properties. The invariant measures are preserved under transformations, which is the central idea of Takens' theorem [27], which will be discussed in more detail in Chapter 3.

In addition to the dynamical invariant measures, geometric invariant measures complement the description of a strange attractor that can exhibit fractal properties. *Fractals* are deeply connected to dynamics through iterations, i.e. via an iterative process, such as the Koch curve [28] or the Cantor set. Fractal geometry allows to measure the extension of fractal objects and their dimension, i.e. the amount of cascaded details of

an object. Well-known dimension metrics are the Hausdorff dimension and the box-counting method[3]. The *correlation dimension* introduced by Grassberger and Procaccia [29] quantifies the density of states in sections of the state space to measure the dimensionality of strange attractors. For example, the correlation dimension of the Lorenz attractor shown in Figure 2.1 is $D_{\text{corr}} = 2.05 \pm 0.01$ [29]. Other dimensionality estimators, such as the Renyi [30] dimension exist for the quantification of strange attractors. Fractality and self-similarity can also be found in time series, where it is expressed by different self-similar temporal scales, as exemplary found in stock price time series [31].

Chaotic dynamics have been observed through a wide range of scientific disciplines, such as electrical engineering [32], chemistry [33], the life sciences [34], mechanical engineering [35, 36, 37, 38], geology [39] and many more. Generally, since conservative systems cannot be asymptotically stable, there is no possibility of fixed points, limit cycles or strange attractors in those systems owing to the lack of volume contraction in phase space. Still, irregular motion is possible, which is then deeply interlaced with regular regimes [40]. Furthermore, *strange non-chaotic attractors* can exist in non-autonomous and quasi-periodically driven dynamical systems [41, 42, 43]. Here, the notation of *strange* relates the not piecewise differentiable, i.e. fractal, geometry, while the motion itself does not exhibit sensitive dependence on initial conditions as indicated by non-positive Lyapunov exponents.

This thesis focuses on friction-affected and self-excited mechanical systems and their dynamics. Almost every mechanical structure must be considered to be affected by friction owing to the mechanical joints that are used for assembling the structure from multiple components [44]. These mechanical joints and contacts have found increasing interest in the research on nonlinear structural dynamics [45, 46]. The uncertainty and nonlinearity introduced by the frictional interfaces cause severe difficulties for accurate modeling [47]. In fact, it was shown that the nonlinear response of two simple beams connected by a bolted joint exhibits such complex dynamics that even highly sophisticated identification strategies and numerical simulations can only approximate the dynamics in certain situations [48]. Since the contact interface, its wear status, and its current stick or slip state are generally inaccessible to measurements, the parametrization of contact models can still be considered an art today [49]. Friction results in nonlinearity, and particularly for small-amplitude motions. This effect can be counter-intuitive regarding the dogma of nonlinearity arising from large displacements. Hence, it must seem natural that the numerical simulation has only limited predictive capabilities for complex structures that are assembled from a large number of component.

[3] The box-counting dimension D is computed via $\log(\text{number of boxes}) = D\log\left(1/\text{length-of-side}\right)$ for a segmentation of space into small boxes and calculating the number of boxes that contain parts of the object.

Self-excited oscillations can be observed in electronic circuits, aero-elastic flutter and flow-induced vibrations [50], supersonic flow past a liquid film, hip endoprosthesis systems [51], violin strings [52], multimode operation of lasers, and beam-plasma systems amongst others [8]. In mechanical systems, the source of self-excitation is often found to be friction. Friction-induced vibrations (FIV) have been observed for a wide range of systems and machines, such as the so-called 'chatter' in machine tools [53, 54, 55] and high-frequency vibrations of brake systems [56]. Mechanisms for FIV in mechanical structures have been studied extensively and among all the works, four main mechanisms can be identified: A coefficient of friction that decreases with relative sliding velocity, binary flutter type instability [57, 58] with neighboring modes coupling, intermittent stick-slip behavior [59] with sliding and sticking phases that can feed energy into the system, and sprag-slip [60]. Tribology has proposed a wide range of friction models [61, 62, 63, 64] to approximate the velocity-dependent Stribeck friction behavior [65]. For high-frequency vibrations in brake systems, the flutter instability has been mostly agreed on to be the dominant driving mechanism [58, 66]. However, multiple mechanisms may be present simultaneously or excite each other, which has not been studied until today [67]. Several types of regular and irregular FIV have been observed in brake systems. Wernitz, Hoffmann and Vitanov [68, 69] studied brake systems in the steady sliding condition, i.e. in the non-vibrating configuration. They revealed signs of low-dimensional deterministic chaos with significant energy shares in frequency ranges way above the audible range that stem from fast-scale frictional processes. Stochastic processes were found to play only a secondary role. Oberst and Lai [37, 70, 71] studied the linearly unstable configuration that exhibits high-intensity FIV. These authors found evidence for periodic, quasi-periodic, and chaotic motion through nonlinear time series analysis. In general, friction-affected structures and friction-induced vibrations are not sufficiently understood today [72, 73, 74]. Hence, friction is one of the major sources of nonlinearity, uncertainty, and unpredictable behavior in most mechanical systems [72, 75, 46], which can be one restricting element for predictive numerical modeling.

2.1.2 Multistability, localization and synchronization

Nonlinearity enables the phenomenon of *multistability*, a situation in which the system has multiple co-existing stable solutions. Therefore, only the choice of initial conditions and instantaneous perturbations will dictate the state to which the system converges. A wide range of complex dynamical systems across the sciences, such as the human brain [34], lasers, power grids [76] and the Amazon rainforest [77, 78], are well-known for multistability [79]. In mechanical vibrations, the observation of multistability seems to be well accepted [1], but has, until today, remained a phenomenon

that is largely not reflected in actual engineering designs, except for bi-stable energy harvesting devices. Figure 2.2 depicts a simplistic oscillator with a velocity-dependent weakening friction characteristic that gives rise to a subcritical bifurcation of the equilibrium position [80]. In a specific velocity range, the steady sliding state and a high-amplitude stick-slip cycle co-exist. The related basins of attraction are separated by the unstable periodic orbit. As a result, small perturbations of the initial conditions can cause jumping phenomena and hysteretic behavior. Given the simplicity of this oscillator, multistability in complex mechanical systems cannot be surprising, but must rather considered to be natural. Therefore, it can seem rather paradoxical that such systems are often studied by means of local and linearizing approaches that cannot resolve multistability.

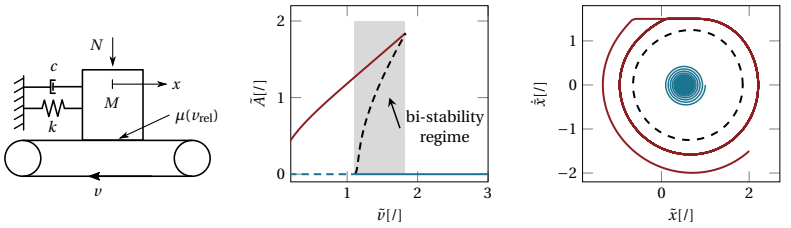


Figure (2.2) Bi-stability in a single-degree-of-freedom oscillator (a) with a velocity-dependent friction characteristic [80]. A stick-slip limit cycle (solid red line) co-exists with the steady sliding solution (blue line) for the belt velocity $\tilde{v} < 1.84$. Owing to the subcritical Hopf bifurcation of the equilibrium solution, bi-stability is observed for $1.1 \leq \tilde{v} \leq 1.84$ in (b). (c) depicts the state space at $\tilde{v} = 1.5$ with the unstable periodic orbit (dashed black line) and the two stable solutions. Parameters and the equations of motion are given in Appendix A.1

Explicit multistability studies are often limited to systems that feature few degrees-of-freedom and, hence, allow for compute-intensive solution path continuation procedures. Examples can be found in self-excited frictional oscillators [45, 80, 81], forced oscillators [1, 82, 83, 84], and rotor dynamics [85] amongst others. Examples for multistability considerations in large numerical models are rather limited to the study of bifurcation paths in turbo-machinery [86, 87]. One possible reason for the lack of consideration of multistability in structural dynamics can be found in the difficulty of obtaining a complete picture of all possible solutions. Particularly hard to find are those solutions which are not connected to other solutions, the so-called *isolated* branches [88]. Once the number of degrees-of-freedom grows, the branching behavior of (even weakly) nonlinear systems can quickly become complicated [82, 88, 89]. Furthermore, most techniques are limited to low-dimensional regular vibrations. Even if extensive studies can reveal the bifurcation diagrams and all possi-

ble solutions, it still remains unknown to which of the competing solutions the system will converge during operation: Brute-force search for the different basins of attraction and their boundaries [90], for example for evaluating the basin stability [79], can become computationally intractable even for low-dimensional systems.

Up to now, uniform motion of dynamical systems has been discussed without a reference to spatially distributed dynamics. Real systems, however, feature multiple components that are geometrically coupled and can interact. Coupled oscillators are known for exhibiting various complicated dynamics that involve salient features of localization and synchronization. The collective dynamics of coupled oscillators, as found in biology, chemistry, physics, brain dynamics, and many more [91], crucially depend on the type, whether local, non-local or global, and strength of coupling [92]. In linear systems, *localized dynamics* were found to arise from inhomogeneities as initiated by the Nobel prize winning '*Anderson localization*' [93] in solid state physics. Here, localization patterns emerge from the distribution of natural frequencies or other symmetry-breaking elements of the structure. These localizations are mostly understood today. On the contrary, understanding of localization in nonlinear systems has not yet been achieved to a satisfactory level. In nonlinear conservative systems, families of breathers and solitary states can be observed [94]. Understanding spatial localization in dissipative systems was enhanced by studying branching behavior stemming from subcritical bifurcations in pattern forming systems [81]. If a single element in an oscillator chain exhibits bi-stability (see Figure 2.2), then the chain of oscillators can exhibit various vibration patterns in which some oscillators coherently follow one solution, while other families exhibit motion similar to the other stable state of the single bi-stable element.

Other phenomena of spatially sharply separated, and concurrently existing coherent and incoherent dynamics exist: Discovered by Kuramoto and Battogtokh in 2002 [92], and termed *chimera states* by Abrams and Strogatz [95] in 2004, those patterns represent another feature of localization in nonlinear dynamical systems. Chimera states denote the remarkable co-existence of phase-locked and drifting dynamics in chains of identical oscillators with non-local coupling for particular initial perturbations. Coherent dynamics with oscillation synchrony exist in one spatial regime, while incoherent dynamics with complete asynchrony exist in the remaining spatial regime [84]. Hence, depending on the spatial viewpoint, one can come to different conclusions about the qualitative character of the dynamics. Localization can play a crucial role in practical vibration research: depending on the sensor placement, qualitatively different dynamics may be observed at the same time. Sensitivity to initial conditions adds additional challenges for experimental analysis. *Snaking* [96, 97] is inherently connected to the spatial localization of vibration states and has been observed across mechanics, fluid dynamics and nonlinear optics [97, 98, 99]. Particularly, snaking bi-

furcations relate to a special structure of the bifurcation diagram where two intertwined solutions branches cross various times in a helix-alike picture for increasing energy levels. Localization and snaking bifurcations are intrinsically related to local resonances and nonlinear normal modes [100] owing to the nonlinear dependence of frequency and energy [101] in the system.

Synchronization is most often denoted as the collective behavior of various components or subsystems of a dynamical system owing to interactions between two processes [91, 102], such as the locking of phases. Synchronization phenomena in structural dynamics are often related to multi-component systems that experience some externally pre-scribed motion in the presence of intrinsic energy sources. For example, a self-excited oscillator may synchronize with a periodic external forcing in the so-called *entrainment* scenario [8]. Also, several self-excited oscillators can synchronize to either regular or irregular motion [13]. Synchronization can be especially relevant when energy transfer happens across time scales, i.e. when slow-scale and fast-scale processes interact in such a way that vibrations grow. Such a scenario is well-known in geology and climate research [103], but has attracted little attention in structural dynamics. The following section discusses several aspects of multi-scale dynamics, which can be assumed to play a role in complex structural dynamics.

2.1.3 Transient, multi-scale and multi-physics dynamics

A further increase of dynamical complexity can be observed in transient, multi-scale and multi-physics dynamics. Transient dynamics include the response to perturbations and chaotic dynamics, which themselves are inherently transient owing to the non-negative Lyapunov spectrum. Considering realistic dynamical systems, such as engines or other machines, one may even wonder if something like a steady-state[4] even exists at all in those systems? Realistic multi-component structures consistently interact with the environment, loads, and operation strategies, and hence would need to synchronize to a uniform motion that is robust against perturbations to remain in a regular steady state. Taking such scenarios into account, transient behavior may seem to be the natural vibration mode in complex systems. Therefore, the current modeling and identification approaches seem to be limited by the inherent assumption of mostly periodic and steady-state behavior. For example, the structural dynamics of an offshore wind turbine blade are intrinsically transient due to cyclo-stationary blade rotation, inertial properties, and variations in the instantaneous wind characteristics or ocean motion. These operational conditions cover a wide range of temporal scales, such as sudden fluctuations, daily variations, and seasonal effects. In response to the instantaneous power demand and the environmental conditions, the

[4] static equilibrium states and those that are periodically recurring are denoted as *steady-state* in this thesis

blade rotation is constantly subjected to a control strategy. Built-up models would have to incorporate multiple physics and multiple temporal scales to adequately represent and predict the aero-elastic structural vibrations, which is still out of scope today [104]. Aspects of such multi-scale and multi-physics dynamics are discussed in the following paragraphs. Particularly, those aspects are closely related to the selection of system boundaries when trying to build mathematical models for complex dynamical systems.

Multi-scale dynamics are well-known in a range of scientific disciplines, such as the life sciences [105], economics [106], geophysics [107], and climatology [103, 108]. Common to all of these scientific disciplines is the finding that complex system dynamics can be governed by latent coupling mechanisms [109] that link different spatial or temporal scales. When dominant processes act on different time scales, the individual analysis of a single time scale might not reveal the relevant mechanisms for the observed dynamics. Thus, a multi-scale analysis may be required to capture the complete mechanisms, event synchronization, and driver-response relationships [103]. Several techniques have been proposed to make certain features of the signal more amenable to study. Exemplary, the wavelet transform is a well-accepted technique for analyzing abnormal ECG recordings, respiratory patterns, blood pressure trends, and other biosignals that carry multi-scale features [105]. In structural dynamics, most research efforts on multi-scale analysis have been devoted to geometrically multi-scale and multi-component materials [110], and multi-scale friction interfaces [111]. Thus, it seems surprising that structural dynamics are sometimes assumed to be governed by a single time scale, while a broad range of other scientific disciplines has already integrated multi-scale behavior into the general knowledge. In this context, the *inherent* assumption of single-scale dynamics can be read from predominant signal processing methods based on the Fourier transform and from computational methods that include only a limited number of harmonic functions or modes. In simulations, whether linear or fully nonlinear, almost all approaches are concerned with the concepts of low-dimensional regular dynamics and, hence, with single-scale dynamics. Temporal and spatial multi-scale behavior is part of many considerations in tribology [112], but once incorporated into structural models of complex systems, rather massive simplifications and homogenization approaches with simplistic friction formulations dominate. Vibrating structures that experience fluid-structure-interaction are, most probably, one of the few structures that are considered in a multi-scale framework owing to the multi-scale nature of turbulence or seasonal effects that arise in in geology. In machine dynamics, only few works have explicitly mentioned vibrations across scales, such as Sever et al. [113] who report on nonlinear dynamics at all spatial and temporal scales in aero engines. For automotive brake systems, Wernitz et al. [68] recorded vibrations in a steady sliding state. Using sampling rates up to $f_s = 200\,\mathrm{kHz}$, they showed significant energy shares in tempo-

ral scales that are magnitudes above the dominant structural modes or dynamics from the rotating disk. Therefore, one may state that temporal multi-scale dynamics are under-represented in the current research on structural vibrations of mechanical systems.

Most of the aforementioned multi-scale dynamics in fluid-structure-interaction (FSI) and friction-excited systems are strongly affected by multi-physics. Multi-physical interactions are the cause for fluid-induced vibrations, such as multi-mode vortex-induced vibrations of flexible offshore marine structures (pipelines, cables) owing to marine currents and ice-induced vibrations of offshore wind turbines. Other multi-physics interactions exist, such as thermo-acoustic instabilities in aircraft turbines. Furthermore, friction is a highly complex phenomenon that involves multiple physics, such as mechanical, thermal, and chemical processes [61]. The energy dissipation in the contact can give rise to temperature effects that in turn affect material properties. Often, the temperature effects act on different temporal scales [114] than the structural dynamics, which highlights the intrinsic coupling of multi-physics and multi-scale effects in some dynamical systems. These examples may imply how those dynamical systems interact with their environment, which consequently needs to be taken into account during the analysis. As a result, the selection of a proper system boundary can be a decisive step during the data acquisition and the modeling step.

2.2 Complexity in structural dynamics

After having illustrated various aspects of nonlinear dynamics, this section aims at a description of *complexity*. However, precisely defining the term of structurally or dynamically *complex systems* appears to be a challenging task. Hence, this discussion starts with a collection of examples given in other scientific disciplines, main properties of complex systems, and common goals for the analysis of such systems.

2.2.1 A review across the sciences

Mitchell [115] gives examples for a range of different complex systems: Ants, which are primitive as an individual, work collaboratively in a colony to accomplish difficult tasks without any central control, i.e. through self-organization. The brain is made of interconnected neurons, which themselves are rather simple structures that fire an output given a sufficient input activation. The huge and self-organized assemblage of neurons gives rise to cognition, intelligence, and creativity. Other examples for complex systems are social networks, food networks, and financial systems. Weaver [116] gave a categorization of problems with different complexity levels in 1948. He stated that the first two classes could be dealt with using the methods available at that time, but that the third class would represent the most important challenge to science in the second half of the 21^{st} century:

(i) Problems of simplicity (involving only few variables), such as pressure and temperature in thermodynamics, dealt with before the early 20^{th} century.

(ii) Problems of disorganized complexity (involving billions or trillions of variables), such as understanding the laws of temperature and pressure as emerging from trillions of air molecules. Here, the assumption of little interaction among variables allows to describe the collective behavior by averages.

(iii) Problems of organized complexity (involving moderate numbers of variables), where components interact in a strong and nonlinear way, such that averages cannot be taken over the complete population of variables, the so-called 'organic whole' [116].

Since the definition of the term *complexity* can be very specific to the research discipline at hand, Seth [117] proposed to measure the complexity quantitatively. In 2001 he introduced a list of metrics and the three classes of *difficulty of description*, *difficulty of creation*, and *degree of organization*. These classes are proposed to be measured by means of information-theoretic metrics, such as the Shannon entropy, fractal dimension, degree of hierarchy, thermodynamic depth, and others.

2.2.2 Properties of complex systems

From this small selection of works on the term *complexity*, a minimal set of properties can be deduced that seem to be common to most complex systems from a wide range of scientific disciplines:

(i) Individually simple (i.e. with respect to the whole system) components or agents are coupled in a strongly nonlinear manner and can produce complicated overall system behavior. Complex systems encode long time history and involve aspects of feedback behavior.

(ii) There is no need for central control, but the system organizes itself through nonlinear information processing, interactions, synchronization, and time-variant adaptations to the environment and the other agents.

(iii) *Emergent behavior* can be observed, i.e. behavior that cannot be explained by the individual component, but must be understood from the system level perspective. Despite the high number of states, only a subset of the phase space is covered, i.e. dimensionality reduction or pattern formation takes place. Such behavior can be captured by measures of entropy.

Additionally, most complex systems, such as the brain, social networks, or chaotic weather dynamics, do not follow a compact or closed-form description. The complex systems and their dynamical behavior cannot be encapsulated into few simple equations owing to the interdependent action of several agents, and adaptation of the agents to environmental impacts, and the limited knowledge of system properties. These characteristics of complex systems prevent direct modeling using first principles, such as Newton's laws or Maxwell's equations. Hence, observations of the dynamical response are studied using evidence-based techniques. Typically, the goals of research on complex systems are the development of mathematical tools for cross-disciplinary insights, and, ultimately, the development of a general theory. Methodically, these objectives are approached by a combination of experimental studies, theoretical work, and computer simulation. Often, the objective is to find causal relations between time series observations of different system states, such as for understanding the El Niño weather phenomenon [118]. Overall, one may say that in complex systems *the whole is more than the sum of the parts*. However, systems can have all the features of complex systems, but yet they can end up exhibiting non-complex behavior, such as equilibrium states.

2.3 Physics-based and data-driven modeling approaches

Mechanical engineering structures have a long history of analytical and numerical modeling. In contrast to complex systems from other scientific disciplines, mechanical structures are assembled from individually well-known parts. Therefore, the component-based superposition principle seems to be a natural approach, even for analyzing complete machines. As a result, an imbalance can be observed between highly specialized and sophisticated computational methods and rather rudimentary signal processing and data analysis, as indicated by various authors [1, 37, 72] for the field of friction-self-excited vibrations of mechanical structures. Numerical modeling can be understood as a *bottom-up* approach, while system identification from observations is a *top-down* approach. The following paragraphs briefly introduce core aspects of modeling and identification approaches employed for the analysis of vibrating mechanical structures today.

2.3.1 System identification and numerical modeling

System identification is the process of developing or improving a mathematical model of a dynamical system from observations [119]. System identification aims at finding a *validated* model, such that the chosen model terms can describe the physical phenomenon as accurately as possible. Linear system identification approaches have matured to a well-established methodology with sound theoretical foundation [1, 120]. *Modal analysis* [121, 122] is probably the most popular linear identification procedure that determines natural frequencies, modal damping values and mode shapes from vibration testing. However, in systems with (even weakly) nonlinear behavior, the basic principle of linear system theory, the *superposition principle*, generally does not hold anymore. Therefore, nonlinear system identification approaches [11, 123] are required for the multitude of nonlinear dynamical phenomena discussed in the previous sections. In contrast to linear identification approaches, the general form of the dynamical system $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \dot{\boldsymbol{x}})$ from Equation (2.1) is unknown owing to the unknown type and form of the relevant nonlinearities. Instead of updating parametrized models (as it is generally done in linear system identification), conventional nonlinear identification approaches [1] propose to iteratively detect the nonlinearity, characterize the location, type and form of the nonlinearity, and finally estimate the related parameters.

In their seminal work, Kerschen et al. [1][5] classify different approaches to nonlinear system identification into the techniques of linearization, time and frequency domain

[5] Interestingly, system identification for systems with chaotic responses is not discussed in that work

methods, modal methods, time-frequency analysis, black-box modeling and structural updating. Frequency response functions, although only valid in linear systems in a strict sense, are often used to study changes of quasi-modal properties under increased excitation amplitudes, or to identify linearized models [124]. Time domain methods include the restoring-force-surface method [125], various variants of time series analysis including auto-regressive moving average models (ARMA) [126] and nonlinear ARMA with exogenous input models (NARMAX) [127], as well as several time-domain transformations, cf. Hilbert and Hilbert-Huang transformation [128, 129]. Frequency domain methods cover adaptations of the frequency response function to nonlinear systems, higher-order spectra [130], and inverse harmonic balance approaches. When it comes to nonlinear behavior, nonlinear normal modes (NNM) [100] are the classical analogues to modal analysis in linear systems. Most applications of NNMs for system identification can be found in forced nonlinear systems with well-separated modes. In time-frequency analysis [131], the Hilbert transform can be used to compute backbone and damping curves from time series, but is limited to single-component signals. Other time-frequency techniques aim at separating different (nonlinear) modes in free oscillations for parameter identification of the individual modes [1]. Naturally, a multitude of novel system identification approaches have been and still are developed for nonlinear systems. For example, sparse regression techniques [132], compressive sensing [133], and Koopman operators [134] have attracted attention recently.

Most of the aforementioned approaches rely on white-box or gray-box models, i.e. underlying assumptions about the mathematical structure of the dynamical system. On the contrary, black-box approaches [135, 136] make no a priori assumption about the dynamical model. First black-box system identification approaches using artificial neural networks (ANN) go back to 1990 [127, 137, 138]. ANN can approximate arbitrary nonlinear functions, and can hence be used to represent unknown input-output behavior. As a downside, most of the black-box and machine learning models have only limited practical use owing to the lack of interpretability. Even though a group of researchers denotes this type of modeling as system identification, this thesis understands system identification as an approach to the inverse problem, i.e. the task of extracting physically meaningful parameters from measured data. Today, system identification of weakly nonlinear single- and multi-degree-of-freedom lumped models with localized nonlinearities are in reach. Complex engineering systems with a high modal density and high modal damping values have been, and still can be, a challenge for system identification.

Numerical models, whether derived from bottom-up or top-down approaches, are useful for predicting the response to some input, for inferring the effect of structural design changes, or for developing a control strategy via virtual prototyping. In struc-

tural dynamics, bottom-up models are most-often derived from discretizations of geometric representations of the structure, such as the finite element method (FEM). More abstract or reduced representations can result in simpler mathematical structures, such as low-dimensional sets of differential equations. The analysis of the dynamical systems aims at the core aspects of stability analysis, bifurcation analysis and the prediction of the dynamics for a given load. Tremendous success of numerical modeling has been achieved, but in some areas there is a striking lack of predictive models. Examples for the latter case have been given in the context of complex, multiscale and multi-physics dynamical systems. Some of the typical approaches to the numerical analysis of nonlinear dynamical systems are outlined in the following.

The complex eigenvalue analysis (CEA) is an established method with a long history in structural dynamics [56, 139, 140]. The system is linearized around a point of operation $\boldsymbol{x}^*$ and the eigenvalues λ of the Jacobian indicate the linear stability of $\boldsymbol{x}^*$ against small perturbations. The eigenvalues $\lambda = \delta + \mathrm{i}\omega$ are associated with the vibration frequency ω and the stability measured by the sign of δ: vibrations grow for a positive real part according to $\left|\mathrm{e}^{\delta t}\right|$. The eigenvectors represent the related mode shapes. However, the relevance of this local linear stability measure is rather limited for the analysis of nonlinear systems. Once a system exhibits stable non-equilibrium attractors, nonlinear and global stability concepts are required, as displayed in the Appendix A.6 for the van der Pol oscillator. The eigenvalues denote only the convergence rate for small perturbations, but not the likelihood of the system to converge to the equilibrium. Bifurcation analysis investigates qualitative changes of the dynamics when a system parameter is varied. Nonlinear normal modes (NNM) can be used to study the system dynamics under increased energy levels, i.e. when nonlinear effects become more pronounced. Complex nonlinear modes, being the analog to the linear complex modes for systems with non-modal damping [141], provide solutions for nonlinear non-conservative systems by means of generalized Fourier series. In contrast to linear modes, NNMs do not obey the principle of linear superposition and hence a system response cannot be composed by addition of NNMs. When it comes to geometrically complex systems, such as those composed of a multitude of components, reduced order modeling (ROM) techniques come to play. ROM aims at reducing the size of the mathematical model by omitting irrelevant states or modes. However, most numerical analysis is limited to weakly nonlinear regimes, because here the systems are amenable to perturbation analysis. Fully nonlinear analysis through time marching integration of the equations of motion is cost-intensive [139, 140]. Depending on the highest mode considered in the analysis, very short time steps must be chosen. Furthermore, the appropriate choice of initial conditions, consideration of time-variant loadings and contact formulations and their parametrization are challenging [139] in complex systems. Some progress has been achieved through uncer-

tainty quantification and uncertainty propagation techniques.

Damping can be considered one of the greatest challenges for numerical modeling. Among the various system parameters that are uncertain or unknown, damping can have the most crucial impact on the system dynamics. It has been shown that even a minor change of a local damping value can de-stabilize systems [142, 143, 144] and qualitatively change the overall motion [145]. The damping parametrization in a simulation model plays a crucial role as it is a sensitive tuning knob for the linear stability margin, thus resulting in common under- and over-prediction issues [139, 140, 145, 146, 147] in practice. This circumstance has already been reported by Bathe and Wilson [148] more than 40 years ago: '*in practice, it is difficult, if not impossible, to determine for general finite element assemblages the element damping parameters, in particular the damping properties are frequency dependent*' and Ouyang et al. [146] add '*The notion of viscous damping is a convenient one rather than an accurate one*'. Along these lines, Tiedemann et al. [75] recently illustrated that material damping has a neglecting effect when compared to damping originating from mechanical joints in brake systems. Since joints and their properties are mostly inaccessible, their characterization is a major challenge [46, 149]. Thus, it may not seem surprising that the numerical simulation of some complex dynamical structures has not arrived at a satisfactory level yet.

2.3.2 Chances and limitations of data-driven approaches

This paragraph attempts to put together most phenomena, identification and modeling approaches, and challenges in structural dynamics of complex dynamical systems that have been discussed in the previous sections. First, mechanical systems are assembled from a multitude of components. In an optimal case, the individual properties of the components are known. However, the mechanical joints in the assembled structure remain mostly unknown, while contributing severe amounts of nonlinearity and damping. Second, the resulting structural nonlinearity must not stem from large-amplitude motions, but can also be initiated by frictional contacts during small-amplitude motion. Third, nonlinearity gives rise to a wide range of dynamical phenomena, such as (quasi-)periodic motion, multistability, localization, chaotic and emergent behavior. Aspects of self-excitation can add autonomous vibrations in the absence of an external forcing. Fourth, besides the internal nonlinearity, energy sinks and sources, the dynamical system interacts with the environment, external loads and operational conditions, which can be multi-scale, multi-physics, and non-stationary. As a result, multi-component dynamical structures must be considered as complex systems since they share salient features of complexity with other systems, such as the climate, the human body or social networks. In this situation, it

seems counter-intuitive to assume the resulting structural dynamics to be stationary or strictly periodic under real-world circumstances. Moreover, given the plethora of unknown and uncertain system parameters, and the possibility of various competing stable solutions, bottom-up numerical simulations cannot be assumed to be as predictive as they are for linear systems. Nonlinear system identification approaches have matured, but still remain a loose collection of specialized techniques for specific use cases. Overall, an imbalance can be observed between highly sophisticated numerical modeling approaches, even though conceptually limited when considering irregularity and complexity, and evidence-based approaches in structural dynamics. Under these circumstances, steady incremental progress can be expected from conventional approaches for several classes of systems that pose severe challenges to understanding and modeling today. Among these systems, highly nonlinear, multi-physics and multi-component structures can be found.

In the age of exponentially growing data, this thesis aims at shedding light on data-driven and evidence-based approaches to understand mechanisms, sensitivities, and representative mathematical models in complex mechanical systems. In sum, there are two aspects of structural dynamics that make novel data-driven system analysis and identification approaches appealing. First, some classes of mechanical systems exhibit complex dynamical behavior that does not allow to isolate specific aspects for conventional identification techniques. Various phenomena and driver-response relationships can be buried in the overall non-stationary dynamics. Secondly, the amount of data acquired from those systems grows rapidly. Hence, data-driven techniques from the quickly growing field of machine learning are promising approaches to discover patterns that are deeply buried in the diverse measurements, so-called *dark data*. Dark data denotes data that was recorded, but due to a lack of appropriate methods, remains unused and does not contribute to new scientific discoveries. Some classes of systems, such as friction-excited brake systems, are only partly understood today, even though large amounts of testing data have already been acquired. The fact that system identification is inherently concerned with the input-output behavior of dynamical systems makes the application of machine learning techniques, also concerned with approximating input-output relations, plausible.

2.4 Analytical models with complex dynamics

Various features of complex mechanical vibrations discussed before may be attributed to the fact that realistic structures are composed of many and partially unknown components. In the following discussions, two analytical minimal models with frictional self-excitation are presented that can produce complex dynamics without the need for complicated structures or severe nonlinearities. A single-mass slider [45, 58, 88, 142, 150] connected through a cubic stiffness joint to a linear subsystem is revealing aspects of subcritical bifurcation behavior, multistability and the emergence of irregular solutions through minor system parameter changes. In a second model [151, 152] featuring gyroscopic terms, a multi-physics friction law is shown to drive the nonlinear system dynamics into an irregular and multi-scale domain. The common aspects of both studies are the simplicity of the weakly nonlinear models and the possibility of complex dynamic behavior - a scenario that is well-known from other scientific disciplines such as the Lorenz system [4]. All governing differential equations, parameter specifications and initial conditions are given in Appendix A. For reasons of readability, physical units are omitted in this section. Bifurcation diagrams were computed using orthogonal collocation methods. Parts of the content and figures presented in this section have already been published in

[88] M. Jahn, M. Stender, S. Tatzko, N. Hoffmann, A. Grolet and J. Wallaschek: *The Extended Periodic Motion Concept for Fast Limit Cycle Detection of Self-Excited Systems.* Computers and Structures, 106-139, 2019

[150] M. Stender, M. Tiedemann and N. Hoffmann: *Characterization of complex states for friction-excited systems.* PAMM, 17(1): 45-46, 2017

[152] M. Stender, M. Tiedemann, N. Hoffmann and S. Oberst: *Impact of an irregular friction formulation on dynamics of a minimal model for brake squeal.* Mechanical Systems and Signal Processing, 107:439-451, 2018

[153] M. Stender, M. Jahn, N. Hoffmann and J. Wallaschek: *Hyperchaos co-existing with periodic orbits in a frictional oscillator.* Journal of Sound and Vibration, 115-203, 2020.

2.4.1 Multi-stable and irregular self-excited vibrations

A minimal model for self-excited instabilities through mode-coupling [58] is extended by a linear substructure via a nonlinear joint element [45] and is shown in Figure 2.3. The Coulomb friction element and the diagonal attachment couple the horizontal and vertical movement of the sliding mass and thus allow for the flutter-type instabil-

ity. The linear subsystem is composed of a single oscillator, denoted as *3DOF model* in the following, or two oscillators arranged in parallel, denoted as *4DOF model* in the following. The subsystem is linked to the oscillator via the joint comprised of a linear spring-damper element and a weak cubic stiffness k_{nl}.

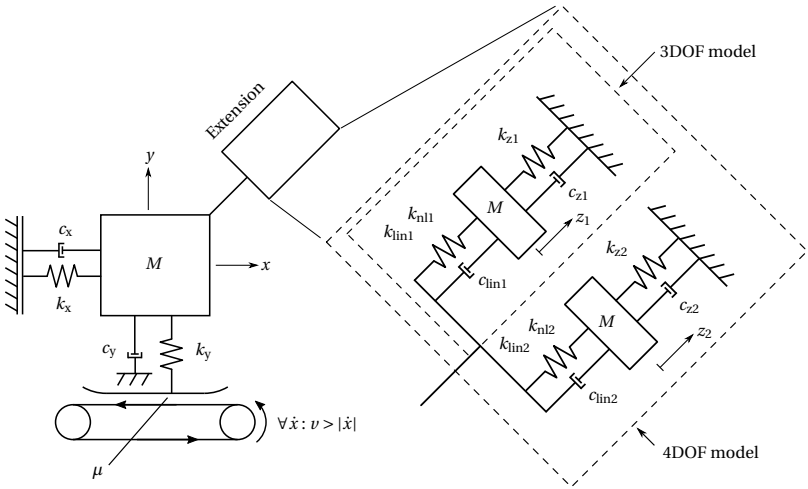


Figure (2.3) Frictional oscillator as proposed by Hoffmann and Gaul [58] and two extensions by linear subsystems that are connected to the slider via a nonlinear stiffness joint k_{nl} as studied in [45, 88, 153]. Adapted from Figure 3 from [88]

The instability mechanism, destabilizing effects of damping and the bifurcation behavior for the linear frictional oscillator and the nonlinear 3DOF model are described in [58, 142] and [45], respectively. In this thesis, focus is put on the qualitatively different types of dynamics that emerge in this system. The system definition, parameters and initial conditions are given in Appendix A.2. First, the 3DOF model exhibits Hopf-Andronov bifurcations [154] of the steady sliding state from which a limit cycle solution grows and links to a second Hopf point as shown in the bifurcation diagram for the horizontal stiffness k_{x} in Figure 2.4 (a). Owing to the subcritical character of the Hopf points, a stable equilibrium solution and two limit cycle solutions, an unstable and a stable one, co-exist locally. A reduced damping level in the structure enables the birth of a detached limit cycle solution, i.e. an *isola*. This solution co-exists with the previously described limit cycle, such that for a large parameter regime $8.0 \leq k_{\mathrm{x}} \leq 29.8$, the steady-state system response depends crucially on the initial conditions. As damping is generally hard to quantify and control in mechanical

structures, a scenario of varying damping levels that create new solutions seems to be natural in mechanical systems, posing massive challenges for numerical simulations and *a-priori* prediction of vibrations. Depending on the shape and size of the corresponding basins of attraction, perturbations of some initial condition can force the trajectories onto a different attractor which will result in significantly different steady-state response amplitudes. Furthermore, when converged to the isola solution, the system dynamics cannot easily be damped by control parameter variations due to the lack of links to the equilibrium solution and the other co-existing solutions in the system at hand. For example, the system can jump from small periodic vibrations at $k_x = 20$ to large amplitudes on the isola in Figure 2.4 (b). To arrive at the stable equilibrium, one would have to increase the stiffness to values larger than 30, jump to the lower solution branch and further increase the stiffness to finally arrive at the stable equilibrium.

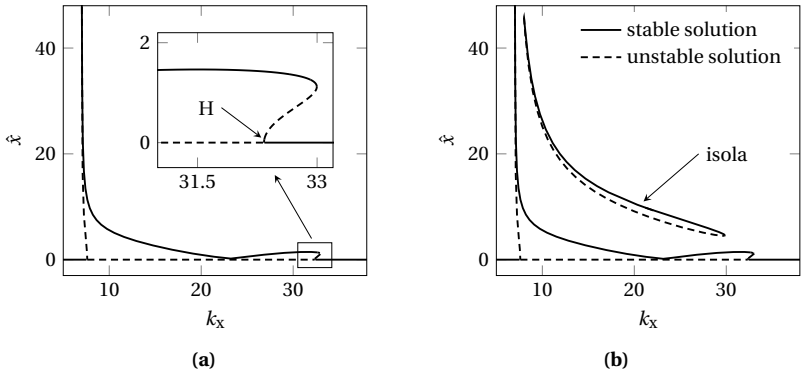


Figure (2.4) Multiplicity of solutions enabled by a local cubic stiffness nonlinearity in the 3DOF frictional oscillator 2.3. (a) limit cycle solution connected to subcritical Hopf bifurcation points H of the equilibrium solution along varying horizontal stiffness values k_x for damping level $c = 0.02$. (b) birth of an isola through reduced damping $c = 0.002$, creating two co-existing stable limit cycle solutions. Adapted from Figure 4 and Figure 9 from [88]

Already in the 3DOF model with its six states, all possible steady-state solutions are difficult to find for a given parameter range. The nonlinear modes of the system and the extended periodic motion concept (E-PMC) [155] can ease this search as shown in [88][6]. The E-PMC creates artificial periodic solutions by adding mass-proportional energy terms that balance out the system's energy intake and dissipation along its nonlinear modes. Whenever the artificial energy term vanishes, a true periodic solu-

[6] These findings are part of joint work with Martin Jahn and Sebastian Tatzko from Leibnitz Universität Hannover and Aurélien Grolet from Arts et Métiers Lille

tion of the autonomous system is found. Therefore, the E-PMC represents a robust and fast method for finding the spectrum of periodic solutions for a given system configuration. Otherwise, brute-force search for the basins of attraction remains as a computationally intractable approach even for small systems.

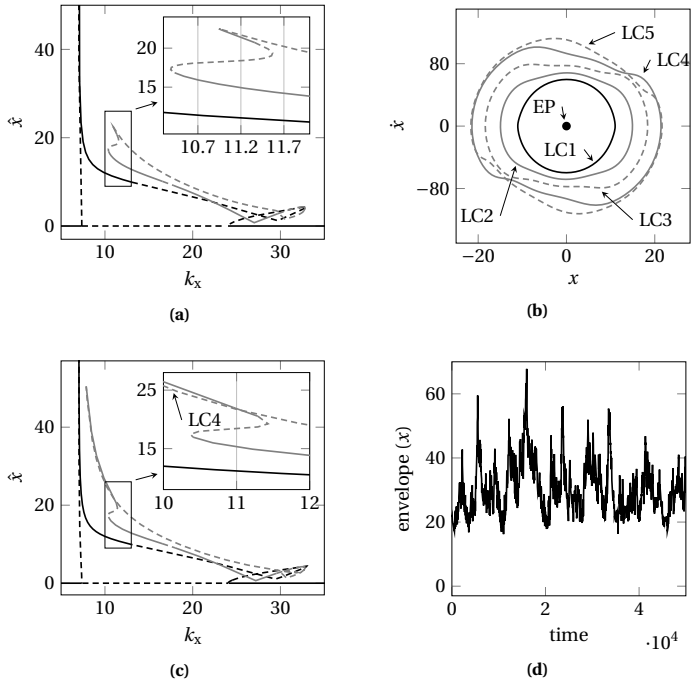


Figure (2.5) Dynamics of the 4DOF frictional oscillator, adapted from [153]. The bifurcation diagram and the state space ($k_x = 11.2$) of the horizontal motion are displayed in (a) and (b), respectively, for the damping $c = 0.02$. Limit cycle solutions are enumerated by their amplitudes in x direction, stable (unstable) solutions are indicated by solid (dashed) lines, and isola solutions are displayed in gray color. Reduced damping $c = 0.002$ gives rise to a more complicated bifurcation behavior in (c) and hyperchaotic motion of trajectories starting from the unstable limit cycle LC4. The upper envelope of a trajectory starting from LC4 for $k_x = 11$ is displayed in (d). Parameters and initial conditions are given in Appendix A.3

A second oscillator in the symmetric linear subsystem (except for the linear stiffness terms $k_{z1,2}$) creates additional phenomena. In the reference damping configuration, five limit cycle solutions (three stable and two unstable ones) are found to co-exist for $10.3 \leq k_x \leq 11.6$ as displayed in Figure 2.5 (a). The bifurcation map shows a compli-

cated branching behavior, which is similar to snaking patterns. Depending on the initial condition or instantaneous perturbations, the system converges to significantly different periodic attractors that differ amplitude-wise by up to a factor of three, see the state space of the horizontal motion in Figure 2.5 (b). Multiple periodic solutions co-exist with the stable equilibrium for $24 \leq k_x \leq 33$. Hence, a mechanical system in this region can exhibit potentially dangerous jumping phenomena, i.e. the transition from a stable equilibrium to a high-intensity periodic vibration through instantaneous perturbations or external forcing.

In a next step, the overall damping level is decreased to $c = 0.002$ to replicate situations in which the damping is highly variable, frequency-dependent or only roughly known. This setting renders the numerical stability and bifurcation analysis even more challenging, see Figure 2.5 (c). For the chosen configuration, there still exist five limit cycle solutions in parallel to the unstable equilibrium. The time evolution of trajectories starting nearby each of the periodic orbits at $k_x = 11$ were studied to confirm the bifurcation diagram. The trajectories settle on one of the three stable periodic solutions for most of the initial conditions. Interestingly, the time evolution of trajectories starting from the unstable periodic orbit LC4 do not settle on a periodic attractor but exhibit irregular long-term behavior displayed in Figure 2.5 (d). It turns out that the irregular solutions arising from that unstable limit cycle are in fact hyper-chaotic [156], i.e. there exists more than one positive Lyapunov exponent [7] such that exponential divergence is observed into multiple dimensions of the phase space. Even though transient chaos [157] can never be ruled out completely, extensive numerical studies involving long integration times and various time marching solver schemes confirm the irregular nature of those particular solutions. Hence, there seems to exist a set of initial conditions that with time evolve 'bouncing between unstable orbits' and create the chaotic motion. Such behavior is reported here for the first time in FIV.

The classical understanding (mainly driven by intuition) of a system exhibiting a single dynamic response for a fixed parameter configuration is contradicted by these findings. Not only different periodic steady-states, but also fundamentally different dynamics are possible, which is a chimera-alike behavior. This observation indicates that, if possible in such simple and weakly nonlinear systems, it may be very likely for such scenarios to occur in realistic complex mechanical structures. Then, the often mentioned non-repeatable and elusive character of friction-induced vibrations [146, 145] can be seen as a natural observation for multi-stable systems.

[7] The Lyapunov spectrum computed via the variational approach B.1 at $k_x = 11$, $c = 0.002$ initiated from LC4 is $\Lambda \approx [0.381, 0.213, 0.075, -0.001, -0.001, -0.079, 0.218, -0.386]^T$

2.4.2 Multi-physics and multi-scale self-excited vibrations

As a second model, a friction oscillator with gyroscopic terms and friction-induced damping is studied along with a multi-physics and multi-scale friction law. The wobbling disk model by von Wagner et al. [151], displayed in Figure 2.6, features a central spherical support of the spinning rigid disk that is in point-contact with an upper and a lower friction element at the outer radius. The contact is modeled by a nonlinear stiffness k_{nl} and a linear viscous damping element c in vertical direction and a preload N_0 that ensures constant contact. Particularly, the nonlinear stiffness is defined as a quadratic function $k_{\mathrm{nl}}(u) = k_0 + k_2 u^2$ of the vertical contact displacement $u = q_1 r$ as proposed by Gräbner [158]. The rotational stiffness k_{t} and the viscous damper c_{t} in the support allow for a wobbling disk motion about the generalized cardan angles q_1 and q_2. This systems exhibits mode-coupling instability through the a-symmetric displacement proportional coupling terms and gyroscopic terms from the spinning disk and the frictional contact. The dynamics of this model and the nonlinear ex-

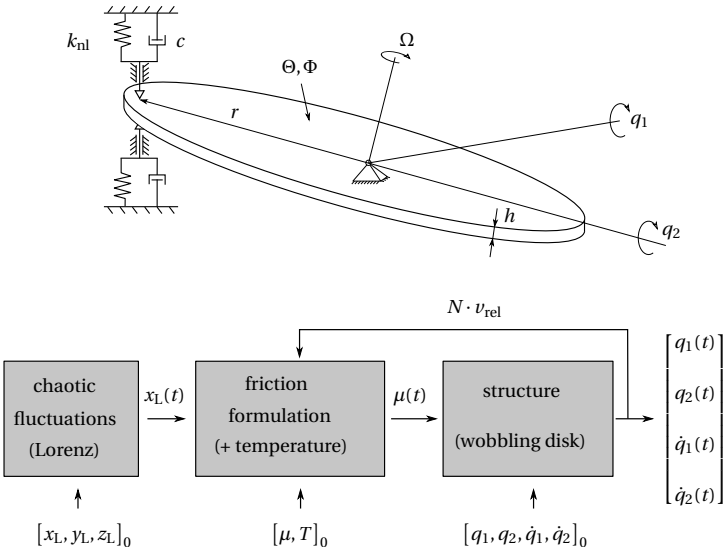


Figure (2.6) Wobbling disk model by von Wagner et al. [151] that is extended by a quadratic contact stiffness nonlinearity k_{nl}. The lower panel illustrates the multi-physics and multi-scale friction formulation in Equation (2.4) coupled to the structure as a feedback loop with fast irregular fluctuations stemming from a co-simulation of the Lorenz system. Reference parameters are given in Appendix A.4. Adapted from Figure 3 from [152]

tension have been studied extensively before [151, 158] using the common Coulomb friction law in the contacts.

The objective of this study is to evaluate the effect of a multi-scale and irregular friction law on the overall structural dynamics. A temperature dependent friction term is introduced following Ostermeyer and Bode [159]. Here, the temperature T can be seen as a slow internal state that is controlled by the friction energy accumulating in the contact. The friction coefficient μ is no longer constant but governed by the differential relationship

$$\begin{aligned} \dot{\mu} &= -A_\mu \left(\mu - \mu_0 - C_\mu T \right) \\ \dot{T} &= -B_\mathrm{T} \left(T - T_0 - D_\mathrm{T} \left| N v_\mathrm{rel} \right| \right) \end{aligned} \tag{2.3}$$

that was proposed by Graf and Ostermeyer [160]. Given the initial temperature and friction values T_0, μ_0, respectively, the inner state T evolves as a function of the absolute energy input, and consequently the friction coefficient changes. The changes of both quantities are controlled by the gain constants C_μ and D_T, while the instantaneous time behavior is controlled by A_μ and B_T. Conceptually, the mechanical vibrations of the disk couple into this formulation via the frictional energy, which in turn changes the inner variable T that causes changes in the friction value, which finally couples back into the nonlinear structure. The time constants A_μ and B_T introduce a first-order lag behavior as a low-pass filter with respect to the driving signals $\left| N v_\mathrm{rel} \right|$ (for temperature $T(t)$) and T (for the friction $\mu(t)$). As a result, the friction coefficient shows rather slow dynamics compared to the structural dynamics and depends mostly on the vibration and force levels in the contact. Inspired by the results on very fast and irregular interface dynamics reported by Wernitz and Hoffmann [68], fast fluctuations of the friction value are introduced through a co-simulation of the Lorenz system. These fluctuations are configured such that their dominant dynamics act on a time scale that is one order of magnitude faster than that of the wobbling disk in the reference configuration. Particularly, the dominant frequencies arising from the main branch switching periodicity of the chaotic Lorenz system are found in the range of 1.7 Hz. The Lorenz system in Equation (2.2) with the states $\left[x_\mathrm{L}, y_\mathrm{L}, z_\mathrm{L} \right]^\top$ is set up to create irregular dynamics using $\sigma = 10$, $R = 28$ and $\beta = 8/3$. The selection of this particular chaotic element is not motivated by physical reasoning or experimental observations. Instead, it serves as a simple, well-studied and prototypical system to create irregular motion that can be coupled into the time-dependent friction value. The irregular friction value formulation including the external chaotic fluctuations

reads

$$\dot{\mu}(t) = -A_\mu \left(\underbrace{\mu(t) - \mu_0 - C_\mu T(t)}_{\mu_{\text{reg}}} - \underbrace{\mu_{\text{fluc}} \cdot \frac{x_{\text{L}}(t)}{18.75}}_{\mu_{\text{irreg}}} \right) \tag{2.4}$$

where μ_{reg} denotes the contribution from the temperature-dependent law in Equation (2.3) and μ_{irreg} denotes the fast irregular contribution from the Lorenz system. Note that the first Lorenz state x_{L} is normalized by its maximum value such that μ_{fluc} allows to scale the amplitude of the irregular friction component. As a result, the irregular term will fluctuate about μ_{reg} by a maximum amplitude of $\pm\mu_{\text{fluc}}$. The complete equations of motions and model parameters are given in Appendix A.4. The dynamical system is comprised of four states related to the mechanical structure and two states for the dynamic and temperature-dependent friction value. The Lorenz system can be considered as an external input. Thus, the complete system has the form $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \dot{\boldsymbol{x}}) + f_{\text{ext}}(t)$, where the states $\left[q_1, q_2, \dot{q}_1, \dot{q}_2, \mu, T\right]^\top$ form the state vector $\boldsymbol{x}$. As there is no feedback connection from the structure to the Lorenz system, $f_{\text{ext}}(t)$ represents a co-simulation that solves the Lorenz equations simultaneously to the system $\boldsymbol{f}$. The resulting feedback loop is depicted in the lower panel of Figure 2.6. The effect of three different frictional systems on the structural dynamics is studied in the following: first, the original system with Coulomb friction, then the system with temperature-dependent friction and finally the system with superimposed irregular fluctuations in the dynamic friction formulation.

Figure 2.7 depicts three representative results of the study reported in [152]. Here, only the state q_2 is displayed as qualitatively similar behavior was observed for q_1. First, the bifurcation behavior with respect to a change of the disk rotation Ω is studied. For the constant friction value, the system exhibits a supercritical Hopf bifurcation at $\Omega = 0.0065$. The equilibrium position is rendered unstable for larger velocities and a stable limit cycle solution exists. The corresponding Lyapunov spectrum Λ was computed using the semi-analytical approach illustrated in Appendix B.1. As the bifurcation point is approached, two of the four exponents grow and approach the zero value. At the bifurcation, one exponent turns zero, therefore indicating regular motion on a limit cycle. The time integration for a system configuration in the linearly unstable regime indicates period-1 motion on a limit cycle. Adding the dynamic temperature-dependent friction formulation to the system is the next step in the analysis. As the nonlinear dynamic interaction of the friction formulation and the structure is unclear and chaos is possible in systems featuring at least three states, the system with dynamic friction is studied by means of the orbit diagram, see Appendix B.2, that is capable of displaying both regular and irregular motion. The dots in Figure 2.7 (b) and (c) indicate the upper envelope maxima of the steady-state time

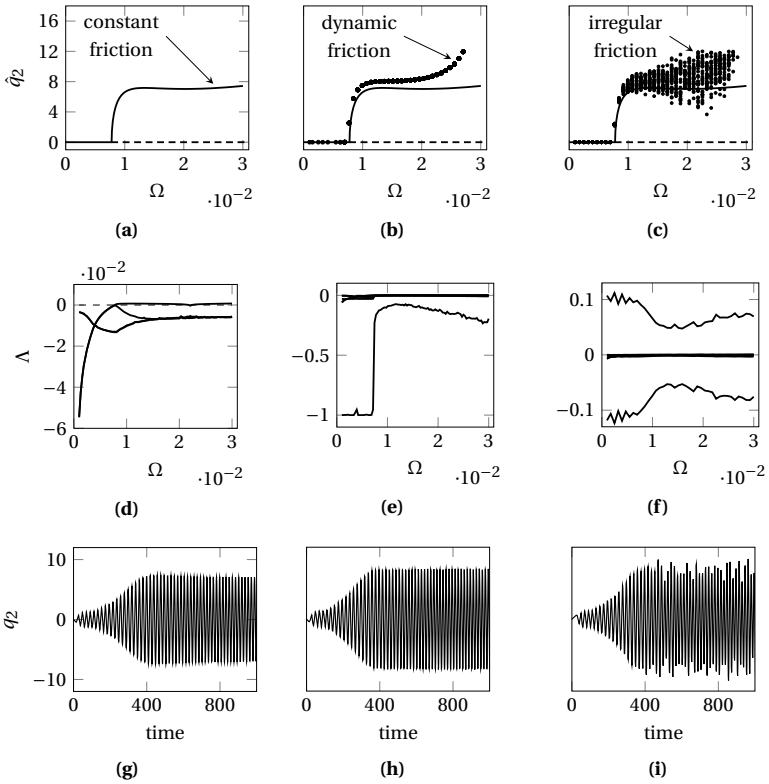


Figure (2.7) Impact of different friction formulations on the dynamics of the wobbling disk model in Figure 2.6: Coulomb friction ($C_\mu = 0$, $\mu_{\text{fluc}} = 0$) in the left column, dynamic friction with temperature effects ($\mu_{\text{fluc}} = 0$) in the center column and dynamic friction with chaotic fluctuations in the right column. The top panels (a)-(c) depict the bifurcation and orbit diagrams with respect to the disk rotational velocity Ω, the middle panels (d)-(f) display the corresponding Lyapunov spectra Λ and the lower panels (g)-(i) show time marching solutions of the second state q_2 for the parameter value Ω = 0.002. Initial conditions and model parameters are given in A.4

traces of q_2. For the temperature-dependent friction, the qualitative stability behavior remains unchanged compared to the static friction case. For all velocity values, the orbit diagram features a single point, and hence indicates periodic-1 motion. This observation is supported by the Lyapunov spectrum with a single zero exponent. Note, that in panel (e) the smallest exponent is constantly (−1) and is not displayed for rea-

sons of clarity. Overall, and for this parameter range, the dynamic friction formulation does not change the regular character of the self-excited vibrations. However, when compared to the constant friction case, the resulting amplitudes of the structural vibrations are larger in the linearly unstable parameter regime. The dynamics resulting from the irregular friction formulation involving fast-scale chaotic fluctuations are depicted in panels (c), (f) and (g). While the overall stability behavior remains unchanged also for this system, the orbit diagram exhibits a scattering range of steady-state maximum values, i.e. aperiodic motion. The Lyapunov spectrum features at least one positive exponent that provides evidence for the overall chaotic structural vibrations. The rather non-smooth evolution of exponents in panel (f) can be traced back to the numerically stiff character of the dynamical system resulting from the multi-scale dynamics and the resulting numerical challenges. Overall, the irregularly fluctuating friction value enforces irregular motion of the structure. Damping out the irregularity, synchronization onto periodic cycles, or homogenization cannot be observed here. Such a picture may seem intuitive in the environment of analytical models - inserting some chaotic parameter fluctuations will most likely result in irregular system dynamics, even if the parameter fluctuations act on a different time scale. Interestingly, the picture seems to be different in experimental testing of friction-excited systems: Despite the knowledge from tribology about ever-changing and multi-scale surface roughness [112], asperity dynamics, wear, adhesion, temperature, and other effects, the periodic character of the resulting self-excited dynamics is somewhat ad-hoc assumed. Aperiodic behavior, non-repeatable results and other effects are most often assumed to be caused by sub-optimal measurements, uncontrolled laboratory conditions or environmental noise. This discrepancy between the modeling and the testing world is interesting to note here.

In fact, the irregular friction fluctuations are not transferred to the structural dynamics in a *one-to-one* fashion: the irregularity introduced at a significantly faster time scale excites irregularity in the slower time scale of the structure. This observation is given in Figure 2.8 (a) where the fast friction dynamics and slow structural dynamics are displayed. From this temporally limited viewpoint, the structural dynamics of q_1 and q_2 even seem periodic at first glance. In fact, the increased vibration amplitudes do not solely stem from instantaneously higher friction values compared to the Coulomb configuration: the minimal and maximum values of the irregular friction coefficient are collected along the bifurcation parameter. Then, the corresponding limit cycle amplitudes of the Coulomb configuration are computed for the minimum and maximum friction value. Figure 2.8 (b) indicates the corridor spanned by the range of limit cycle amplitudes when considering a constant friction value. It can be observed that the vibrations of the irregular configuration, even if the friction coefficient is only fluctuating between μ_{min} and μ_{max}, leave the corridor. Hence, the amplification cannot solely stem from a parametric excitation, but has to be rooted

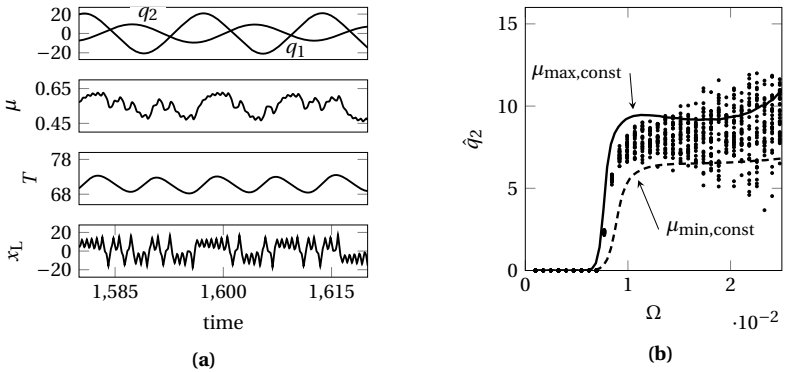


Figure (2.8) (a) steady-state evolution of the system states for the irregular friction formulation at $\Omega = 0.02$. (b) depicts the corridor of limit cycle amplitudes that would result from the constant friction formulation when parametrized with the minimum and maximum friction value observed in the irregular formulation. Dots indicate the system response of the irregular model configuration in terms of the orbit diagram. Adapted from Figures 8 and 11 from [152]

in intrinsic dynamic interactions of the structure with the chaotic friction fluctuations. A deeper analysis [152] of the characteristics and dynamical invariants of both the Lorenz-based fluctuations and the chaotic structural motion exhibits clear differences: even if the fast-scale irregularity excites structural aperiodicity, their qualitative properties are not similar. Hence, observing only the structural vibrations of the system at hand would not allow to draw conclusions about the dynamical properties of the actual friction dynamics in the interface.

Concluding, the two aforementioned studies reveal that complex dynamics do not necessarily require complicated structures or severe nonlinearities. Moreover, the transition from regular to irregular dynamics can be gentle, and different physical effects can interact across multiple time scales. In the first system, damping was found to play a crucial role in the formation of multistability scenarios and the emergence of chaotic motions. Therefore, one may speculate that aspects of multi-scale dynamics, multistability, and chaotic motion are likely to occur in real structures that are affected by friction. Hence, these aspects need to be taken into account when identifying and modeling those systems.

2.5 A framework for describing complex dynamical systems

Following the previous discussions on the nonlinear dynamics of complex systems, this section proposes a generic framework for a data-driven description of complex mechanical systems and their vibrational behavior. In this context, the *data-driven description* denotes an observation-based description that can be built from experimental testing or numerical simulations. A schematic of the framework is presented in Figure 2.9 for illustrative purposes.

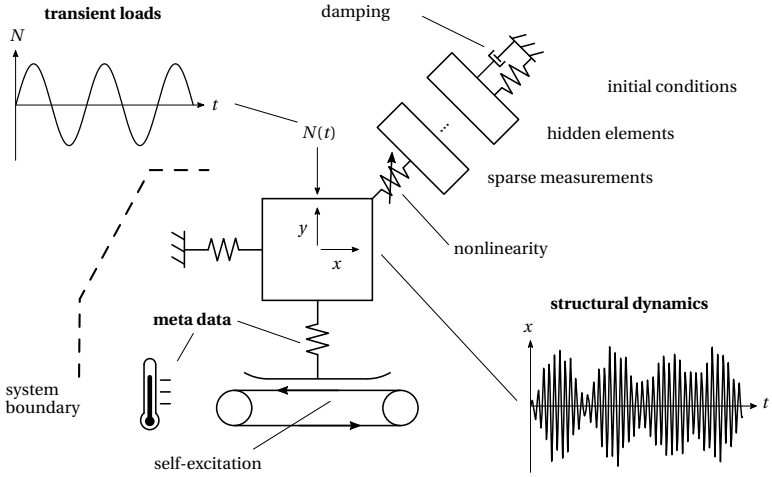


Figure (2.9) Schematic representation of the framework for describing complex dynamical systems by means of **structural dynamics**, **transient loads** and **meta-data** to be compatible with a data science process. Additionally, major keywords capturing the salient features of complex dynamics are given

Conceptually, the framework consists of three main classes of observations or types of measurements:

1. the **structural system dynamics** observed through motion sensors,
2. the potentially multi-physics external **loads** and operational conditions, and
3. static system configurations, initial conditions, and component properties described by **meta-data**.

The first two classes, i.e. vibration measurements and loads, typically take the form of

time series data. Owing to the transient and potentially multi-scale characteristics of these measurements, adequate sampling rates and recording durations are required. The third class, i.e. meta-data, is typically realized by a collection of scalar values, such as components' eigenfrequencies, constant environmental factors, set-up specifications, or others. The meta-data are static values which are not affected by the current loadings and the vibration of the structure. For each of the measurements, limited observability is caused by unknown system properties, inaccessible or hidden elements that allow only for sparse measurements. These circumstances render direct numerical modeling and model updating challenging, even for comparatively simple structures in well-controlled environments. Hence, the objective of this thesis is to explore the potential of evidence- and data-driven techniques to facilitate the understanding of vibrating mechanical systems.

Given a specific system, the selection of (i) a spatial and temporal scale to analyze, (ii) the selection of a system boundary, and (iii) limitations in observability need to be considered. These factors dictate the kind and amount of data that is required for capturing most of the physics involved to reveal hidden mechanisms buried in sparse measurements. Then, a data science process can be formulated before adequate methods and models are selected. This framework enables a consistent and compatible treatment of the various classes of data gathered from vibrating structures, which will be discussed the upcoming Chapter 3.

3 A data science process for structural dynamics

Dramatically decreasing sensor costs and de-facto unlimited data storage capabilities have enabled the collection of large amounts of data in the fields of structural dynamics and mechanical engineering. Big data sets can enable the discovery of patterns that would not be observable in small numbers of observations. However, novel approaches are required to facilitate a highly automatic and consistent vibration analysis of those data. Additionally, a somewhat extrinsic motivation for data-driven approaches is omnipresent today: given the success of the data sciences across all fields of science, why should data-driven approaches *not* enable significant progress also in structural dynamics? This thesis aims to shed light on the possibilities and limitations of data-driven analysis and modeling approaches for research in nonlinear mechanical vibrations. Pathways are proposed for handling complex time series measurements in a data science process that obeys the underlying laws of physics and nonlinear dynamics.

Possible motives for the use of data-driven analysis techniques for mechanical vibrations and structural dynamics are depicted in Figure 3.1. Starting from time series data acquired during vibration measurements, the underlying dynamical system can be studied through **identification** strategies. **Classification** of sequential data can be useful for recognizing particular states of operation and early signs of a change in the vibrational responses. **Control** of complex systems is possible through data-based approaches [161], but is not covered explicitly in this thesis. Additionally, future state **prediction** is one motive for developing digital twins of vibrating structures using data assimilation. Several strategies are presented hereafter to approach those tasks following a novel data science process tailored to the peculiarities of mechanical vibrations.

A classical data science process is introduced first. Conceptual differences to engineering vibration research are identified that motivate adaptations in the context of structural dynamics. Finally, a summary is given on methodologies that can be employed for physics-informed data assimilation and data analytics, with a focus on nonlinear dynamics to take into account the multitude of vibration characteristics discussed in Chapter 2. The concepts, methods and thoughts introduced hereafter are not meant to give a complete picture on data-driven methods, but to provide a high-level overview. In this thesis, most of the methods and approaches given in the

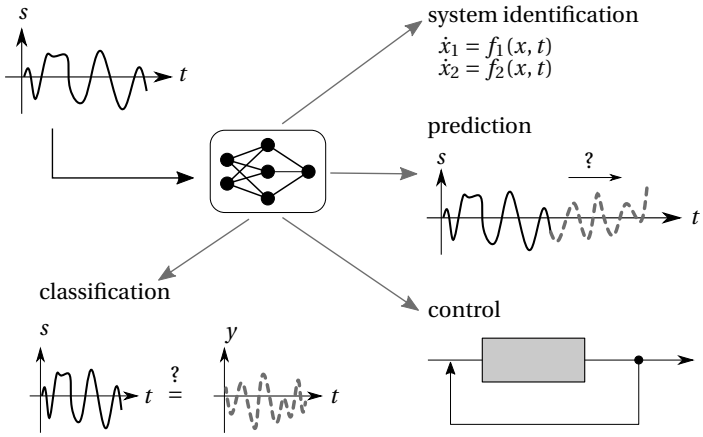


Figure (3.1) Major motives for the use of data science in the fields of mechanical vibrations and structural dynamics

following sections are tailored to time series data that stem from deterministic systems with only little or vanishing aspects of stochasticity. Although this may seem to be a strong restriction, several recent studies have revealed strong signs of determinism in complex and seemingly random mechanical vibrations [68, 70, 162].

3.1 The classical data science process

The process of obtaining insights from observations through data-driven techniques can be described by a *data science process* or *data science lifecycle* [163]. The individual phases of this process are not following strict definitions, but are rather best-practice guidelines to data science projects. Generally, single instances of data are denoted as *observations* or *samples*, which are typically composed of multiple *features* or *dimensions*. In the simplest case, an observation takes the form of a vector carrying multiple features. Analogously, model outputs can take different forms, such as vectors of scalars and sequences. Machine learning problems can be cast into the general classes of unsupervised or supervised learning, as well as *regression* or *classification* tasks. Supervised learning denotes the task of learning patterns from historic data for which a ground truth label is available. Hence, the model can learn an approximate mapping from the input features x to the requested outputs y from the labeled data.

The model error can be directly accessed by comparing the prediction $\hat{y}$ to the ground truth y. The validated model $\mathcal{M}$ can then be used to make predictions on new, that is unseen and unlabeled, data. On the contrary, unsupervised learning processes consider unlabeled data, which in most cases play a role in clustering tasks. Generally, feeding some inputs to a learning model and obtaining an output is referred to as 'making predictions' in the following. Hence, there is no restriction to a specific task or problem setting whenever this term is used.

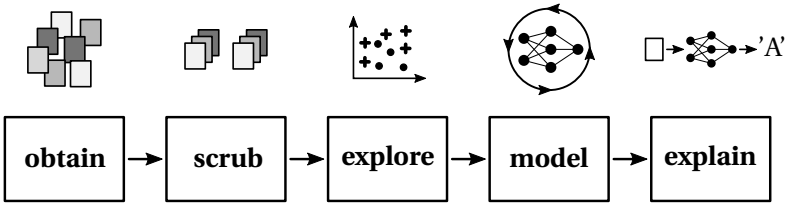


Figure (3.2) A possible description and schematic illustration of a classical data science process following the 'OSEME' work flow

One possibility to describe a data science process is by the phases of **obtain**, **scrub**, **explore**, **model**, and **explain**, i.e. the **OSEME** process [163] displayed in Figure 3.2. Another way to describe a data science lifecycle is the Cross Industry Standard Process for Data Mining (CRISP-DM) [163] that is more business-focused, helping decision-making for commercial interests. The most important aspects of the OSEME data science process and its phases are:

obtain: Following the problem description, potentially relevant raw data are acquired, collected from existing historic recordings, or aggregated from multiple sources. The data types and formats involved may vary.

scrub: The data are filtered and transformed into consistent formats. The data are cleaned, i.e. outliers and missing values are detected, and treated in a way that is coherent with the problem description. The succeeding steps of the project strongly depend on the data preparation, since, data-driven models are governed by the principle that colloquially reads 'garbage in, garbage out'.

explore: To get a general impression of the data distributions, the cleaned data set undergoes an exploratory data analysis step. Dominant patterns and relations between features may become visible through simple graphical representation or the calculation of correlations. Specific domain knowledge may allow to calculate features from the raw data that collect characteristics in an abstract form.

Dimensionality reduction techniques can be used to remove redundant or linearly dependent feature dimensions.

model: Depending on the problem description, one or multiple models are selected to discover patterns, make predictions, or derive forecasts in an unsupervised or supervised manner. The model architecture depends on the type of data, the complexity of the underlying patterns, and the amount of data that is available. Quality metrics must be defined that allow the validation of the model once it has been built. Hyperparameter optimization may improve the quality metrics through changes to the model configuration.

explain: The predictive capabilities of data-driven model lie in their ability to learn from historic data and make predictions on new data. Additionally, the study of the learned approximation can help to understand the possibly highly nonlinear relationship between input and output quantities.

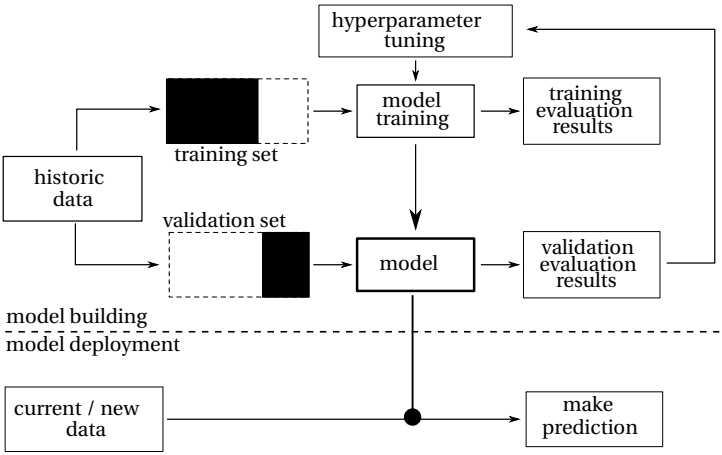


Figure (3.3) Building data-driven models through training on a subset of all available data and evaluating the performance on the validation set. Variants of the models are compared in hyperparameter studies to find the optimal model which can be deployed to make predictions on new observations

A fundamental concept in most data science processes is the evaluation of the model generalization by means of a validation data set. The complete set of available historic data $\mathscr{D}$ is initially split into the *training* set $\mathscr{D}_{\text{train}}$ and the *validation* set $\mathscr{D}_{\text{val}}$, where the training set is often chosen as 60 to 80% of the overall data set. *Training*

of a model denotes the process of finding the model parameters that optimize a loss function for the observations in the training set [1]. After training, the model parameters are frozen and the model is validated against observations from the validation set. A schematic of the typical work flow is illustrated in Figure 3.3. High quality models perform with small error on the validation set, and are denoted as *generalizing well*. However, there is a possibility of selecting a model that cannot capture the salient patterns in the data due to insufficient model complexity. For example, approximating a nonlinear process with a linear model can lead to *underfitting*. Analogously, if the model complexity allows to *overfit* the training observations, the model over-adapts to the training data and shows weak generalization on the validation set. Therefore, finding a good balance between model complexity and model error is one of the key challenges while building and selecting effective data-driven models. Ultimately, validated models can be used to make predictions on new observations. This final stage in the data science lifecycle is commonly referred to as *deployment* phase or *sending a model to production*.

3.2 Peculiarities of data-driven structural dynamics

In this section, some aspects of dynamics and dynamics-related data are re-visited that motivate the adaptation of the classical OSEME data science process for the use in mechanical dynamics. Two main themes are identified that make structural dynamics different from classical data science applications: the **problem description** and the **data characteristics** in the field of mechanical vibrations. Although it is impossible to generalize the vast applications of machine and deep learning, it is probably safe to say that, in most cases, the model prediction quality is more important to the user than the learning process and model itself. For example, in the highly complex task of natural language processing, the most important goal is to decrease the errors regrading syntax, grammar, and semantics. The way the model achieves this goal can be considered to be of significantly minor importance compared to the output quality for customer applications. However, achieving high prediction scores in some modeling task is only the first step for making use of the data sciences in structural dynamics. The ultimate goal is knowledge discovery for understanding complex dynamical phenomena and developing design countermeasures. Hence, the predictions given by a data-driven approach must be related to the actual mechanical structure, component properties, loading conditions, or other system properties. Interpretability [164, 165, 166] of the models and their predictions is a core objective to make practical use of data-driven approaches in structural dynamics. As a con-

[1] For example, in a simple regression task, the best model is typically found by minimizing the root-mean-squared error between the regressor and the actual data.

sequence, dimensionality reduction techniques that do not preserve physically explainable features, such as the famous principle component analysis (PCA), may not be suitable for the analysis of structural vibrations. Furthermore, interpretable models that assign importance values to input dimensions may be favored over complex models that lack the possibility to understand the model decision making process. Thus, not only the data treatment, but also the modeling phase must be compatible with the objective to discover causal relations in the mechanical vibrations data.

Dynamics, being the science of motion, rely on the evolution of states and, hence, on temporal relations and change. Data acquired for structural dynamics is **sequential** as discussed in Section 2.5. The sequence of measurement values carries the crucial temporal information. Stationarity cannot be assumed a-priori, such that behavior may change over time. Considering continuum mechanics, mechanical systems feature an infinite number of degrees of freedom or states. In experiments, only a very small number of those states can be measured at finite precision over some time span. Hence, the available data is **sparse** in a spatial (considering all possible degrees of freedom), temporal (limited by the temporal resolution of the acquisition system), and physical (only some physical quantities are measured) sense. In some cases, it is unclear if a given measurement carries essential pieces of information to understand a dynamical phenomenon, or if relevant dimensions are not available to the current measurement setup. Another aspect of sparsity is given by meta data, according to the system description framework developed in Section 2.5. Environmental and loading conditions, individual system component properties, and measurement circumstances may be unknown or inaccessible in practical engineering reality. While these aspect seem solvable through extensive measurement instrumentalization, key aspects of nonlinear dynamics pose more crucial implications on meta data: As described in Section 2.1, nonlinearity promotes the sensitivity to initial conditions that is causing, amongst others, load history effects, irregularity, and multi-stable solutions. Initial conditions are challenging to measure, especially in structures assembled of many mechanical elements. Furthermore, the identification and precise description of omnipresent frictional contacts must be considered impossible. As a result, those data must be considered to be very sparse, even though having a notoriously high impact on the nonlinear dynamics of a system. Lastly, mechanical vibrations are often related to a malfunction, inefficiency, and undesired dynamics of a mechanical structure. As such, the observation of different vibrational patterns is commonly heavily imbalanced, meaning that the desired operation mode dominates the recorded data, and only few observations of abnormal behavior exist. However, these observations are of crucial interest to the researcher or engineer.

3.3 A novel data science process for vibrations

Considering the characteristics of the data and problems typically encountered in structural dynamics, a novel data science process for mechanical vibrations is proposed and discussed in the following. Adapting the notion of OSEME, the new process follows **O-P-T-M-E**, indicating the phases of **obtain**, **pre-process**, **transform**, **model**, and **explain**. The details of this process are discussed hereafter. One of the key differences to the classical data science process is the explicit consideration of physics in each step, the stronger link between individual phases and, most importantly, the stronger relation of the learning model to the mechanical structure itself. To be compliant with the type of problems addressed in structural dynamics, the process is more of a cyclic instead of a sequential process. A schematic, illustrating the OPTME process, is given in Figure 3.4.

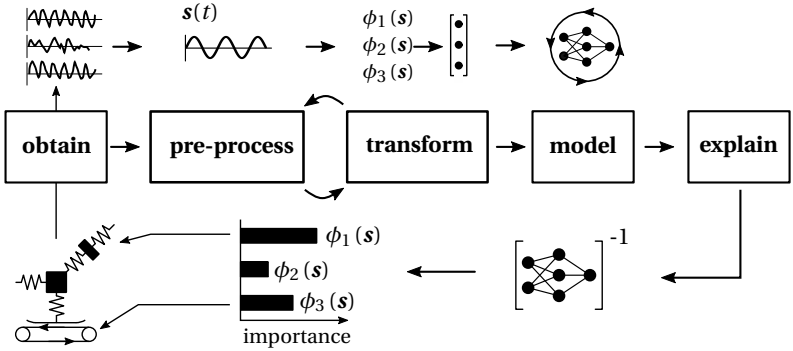


Figure (3.4) The 'OPTME' data science process that requires the phases to be compliant with the underlying physics. The interpretation of the learned model can help to understand the role of system components and their impact on the observed vibrational behavior

3.3.1 The *obtain* phase

The first phase is concerned with data acquisition. Contrary to modal and operational deflection shape analysis [121], it is proposed to store the complete time series data from experiments, including transients at the very beginning of a measurement. One of the reasons for this recommendation is the observation presented by the author in [167] and in Section 4.1, where the identification of the underlying system crucially

depends on the consideration of transient dynamics. Transient behavior is very characteristic for a particular system at hand and forms a distinct signature in contrast to steady-state behavior. Hence, and in analogy to concepts of entropy, the amount of information contained in transients, i.e. including chaotic dynamics, is larger than the amount of information contained in regular steady-state motion. As most realistic dynamics can be assumed instationary, sparsity in the temporal sense should be addressed by long measurements and selecting the highest sampling rates [2]. Spatial and physical sparsity is in practice dictated by the available sensor instrumentation and the accessibility of the mechanical structure, which is related to optimal sensor placement [113]. Again, it can only be advised to measure at as many locations of the structure and as many physical quantities as possible. Identification of irrelevant dimensions and data reduction are part of the later stages of the data science process. Furthermore, seemingly irrelevant quantities can represent proxy measures for essential states that allow for a better system description. Choosing appropriate observables for the measurements through data-driven methods is at the core of the proposed methodology, and represents one of the main contributions of this thesis to the current state-of-the-art research on mechanical vibrations.

3.3.2 The *pre-processing* phase

The pre-processing phase focuses on data exploration and basic signal processing to carve out relevant characteristics for the following *transform* phase. It is closely related to the transformation phase and may be part of an iterative search for the optimal signal pre-processing using the insights from the transformation phase and vice versa. Signal pre-processing covers aspects of filtering, data slicing into shorter sequences, time scale selection, as well as annotating the data. Filtering noise from the measurements can be achieved through a wide range of techniques, such as spectral bandwidth filtering, rolling mean averaging, or other methods from the signal processing community [168, 169, 170]. However, noise filtering should be conducted carefully with a first test for determinism [171, 172] in the signals. Seemingly noisy signals may in fact stem from irregular but deterministic processes. As filtering is very specific to the application case, no further discussion of appropriate methods is given in this thesis. The reader is referred to [168] for a thorough overview. Depending on the successive transformations applied to the measurement data for extracting meaningful features, a test for stationarity [173, 172] can be essential. If strong non-stationarity is observed that may cause conflicts or inconsistencies in the feature engineering phase, time-frequency analysis [131] and **data slicing** can create shorter se-

[2] Down-sampling, sequence truncation, and other typical signal processing steps will be part of the pre-processing step within the OPTME process, and not of the measurement phase.

quences that better fulfill stationarity assumptions. For sequential data $\boldsymbol{s}(t)$, *moving window processing* following

$$\boldsymbol{s} \in \mathbb{R}^{n_\mathrm{t}} \mapsto \boldsymbol{s}_k = [\boldsymbol{s}(t_0 + (k-1)\,\Delta w), \ldots, \boldsymbol{s}(t_0 + w + (k-1)\,\Delta w)], \quad k = 1, \ldots, \frac{n_\mathrm{t} - w}{\Delta w} \quad , \tag{3.1}$$

is a viable tool. A window of length w is shifted by increments of Δw along the signal $\boldsymbol{s}$ that has an original length of n_t samples to obtain the sequences $\boldsymbol{s}_k$.

Time scale selection is relevant for dynamical systems that are simultaneously governed by slow changes and fast fluctuations. Knowledge of the scales that are involved in the dynamics is required for a successful modeling and interpretation. Spectral analysis by the Fourier transform [174] can be a good starting point, but is only useful for stationary signals. The short-time Fourier transform (STFT) or *Gabor* transform [175] displays the frequency information along time. Likewise, the continuous wavelet transform (CWT) [176, 177] localizes the frequency information along time using wavelets as orthogonal basis. Hence, the Gabor and the wavelet transform are suitable candidates for non-stationary data with time-variant dynamics. Appendix B.4 gives the mathematical definitions of these transforms. The uncertainty principle, denoted as the *Gabor limit*, is a fundamental constraint of time-frequency analysis [178]. Since the STFT has a fixed resolution in time for a given frequency sampling, it is only expressive for signals with one dominant spectral scale. The wavelet transform has a high time resolution at low scale and a low time resolution at high scale, see Figure 3.5. Hence, the wavelet transform and its inverse are especially useful for the analysis of multi-scale signals. When only individual coefficients of the wavelet transform are used to compute its inverse, time signals with a certain frequency bandwidth content are recovered.

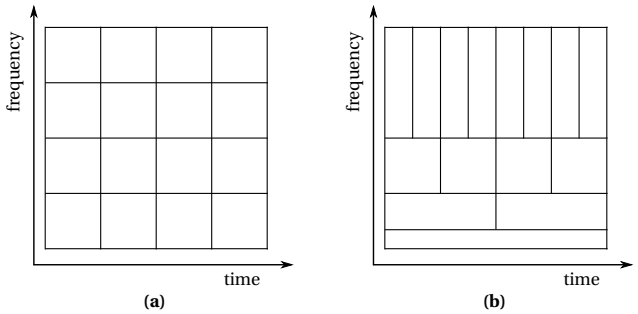


Figure (3.5) The uncertainty principle illustrated canonically by *Heisenberg boxes*, having all the same area, for the time-frequency analysis of signals [179]: (a) short time Fourier transform with a fixed resolution in time and frequency and (b) multi-resolution analysis with a hierarchical structure stemming from the wavelet transform

Figure 3.6 (a) depicts the discrete wavelet decomposition of a signal using the coefficients at five levels of scale. For increasing levels, the spectral resolution decreases, such that the fast components of the signal can be found in the first decomposition levels.

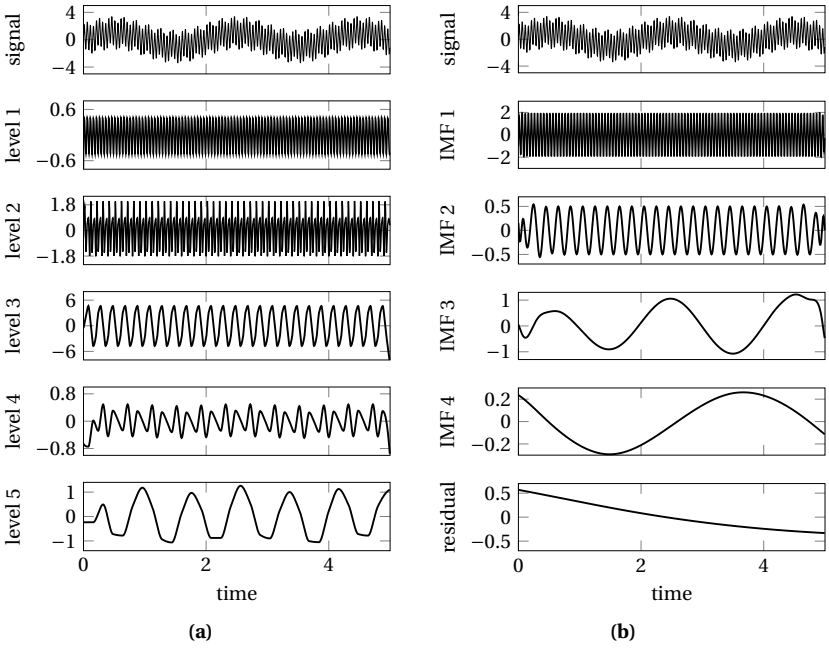


Figure (3.6) Multi-resolution analysis of the sample signal $\boldsymbol{s}(t) = \sin(\omega_1 t) + \frac{1}{2}\sin(\omega_2 t) + 2\sin(\omega_3 t)$ with $\omega_1 = 0.5 \cdot 2\pi$, $\omega_2 = 5 \cdot 2\pi$, $\omega_3 = 20 \cdot 2\pi$ sampled at $f_s = 200\,\text{Hz}$. (a) one-dimensional wavelet decomposition using Daubechies-2 wavelets at five detail coefficient levels after temporal upsampling and (b) Hilbert-Huang decomposition of the signal into four intrinsic mode functions (IMF)

As a second approach to time scale separation, the Hilbert-Huang Transform (HHT) [128, 129], presented in Appendix B.3, can be utilized. The HHT decomposes the original multi-component signal $\boldsymbol{s}(t)$ into so-called mono-component *intrinsic mode functions* (IMF) that encode the time-frequency localization on different time scales. The HHT lacks a sound theoretical description and is given by an empirical algorithm that subtracts the IMFs from the given signal, following the so-called *sifting process*. Empirical mode decomposition (EMD) is used to calculate the IMF and Hilbert spectral analysis and obtains the instantaneous frequency information from the differ-

ent IMFs. IMFs represent oscillatory modes with variable amplitudes and frequency along time, which renders the HHT especially useful for nonlinear and non-stationary time series data. For comparison to the DWT scale separation, the IMFs for the sample signal are depicted in Figure 3.6 (b). As a major difference to the wavelet transform, the scales computed by the HHT are defined by characteristics of the input signal, to which the HHT adapts. Additionally, adding up the IMFs in time domain will approximate the original signal. Other approaches to multi-resolution analysis have mostly been proposed in the life sciences [109].

Overall, and depending on the application case, the time scale analysis helps to identify the relevant time scales and their extraction from the measurement data without prior knowledge, such that sequential data for a specific spectral range can be transferred to the next phase in the process.

3.3.3 The *transformation* phase

After signal pre-processing, the data need to be transformed into quantities that are the most descriptive for the researched phenomenon and into a format that can be handled by a machine learning model. In Figure 3.4, these transformations are indicated by $\phi_i(x)$. Commonly, the activities in this phase are denoted as *feature engineering*. Instead of feeding the pre-processed time series measurement to a machine learning model, few quantities that are abstract representations of the main characteristics of the data are opted for. These representations are denoted as *features* in the data sciences. Exemplary, the amplitude distribution of a time series can be quantified by statistical moments, such that the mean represents one possible feature. Generally, choosing an appropriate transformation for creating features can be considered more important than the model selection or data pre-processing step. The representation of the data in a different domain can simplify the processing, emphasize dominant properties, or decouple complex relations. Interestingly, the life sciences [180], climate research [181, 182], and other disciplines are well-aware of this importance and have developed a large corpus of transformations and feature generators for time series analysis, while classical engineering dynamics seem to have been largely limited to the time and frequency domain representation of vibration measurements. For the usage in structural dynamics, the data transformation should be interpretable with respect to the underlying physics, such that the features are meaningful for nonlinear dynamics. Otherwise, even if the model allows for explaining its decision making process, the overall data science process would lack explainability owing to features without a physical meaning. For example, spectral analysis can extract the essential features from the Fourier transform for a signal dominated by few periodicities. However, when the spectral content is broader, such as in the case of

chaotic dynamics, spectral analysis may not be suitable anymore to extract quantities that express the discriminative features of the time series [183]. Instead, dynamical invariant measures, such as the Lyapunov spectrum [70, 184, 185], Rényi entropy [181, 186], and the fractal dimension [187], represent quantities that are essential for the description of irregular motion in an abstract form [173, 188]. Among the various transforms for time series data, the following discussion highlights concepts from nonlinear time series analysis (NTSA) and recurrence plot (RP) analysis since they are especially viable tools for non-stationary and irregular processes.

Nonlinear time series analysis [172, 173, 189, 190], particularly famous in the 1980s, is the scientific discipline that is concerned with *estimating* [191] nonlinear invariant measures from time series data. Computing invariant measures [192] typically requires knowledge of the evolution of trajectories in the full, i.e. potentially high-dimensional, phase space of the dynamical system. In experimental dynamics, there may be only an univariate measurement of a single quantity available to describe the dynamics of a mechanical structure or machine. Many degrees of freedom constitute the actual state phase of this machine, such that the acquired data are sparse, see Section 3.2. To resolve the issue of missing dimensions in the measurement, Takens' theorem [27] allows to unfold an univariate time series into a reconstruction state space in which the attractor is topologically [193] and dynamically equivalent to the one in the actual state space, see Figure 3.7. Determinism in the dynamics is the strong assumption underlying this type of *attractor reconstruction* and can be tested for in various ways [171, 194]. Given a time series $\boldsymbol{s}(t_i)$, sampled at equidistant times t_i, *delay embedding* reconstructs the evolution of a trajectory following

$$\boldsymbol{x}(t_i) = \left[\boldsymbol{s}(t_i), \boldsymbol{s}(t_{i+\tau}), \boldsymbol{s}(t_{i+2\tau}), \ldots, \boldsymbol{s}(t_{i+(m-1)\tau})\right] \tag{3.2}$$

in the m-dimensional state space. $n_\mathrm{t} - (m-1)\tau$ vectors in the reconstructed state space result from the embedding. Here, τ denotes the delay that needs to be chosen such that it most effectively de-correlates the temporal relations in $\boldsymbol{s}(t_i)$. Typically, τ is chosen as the first zero-crossing of the signal's auto-correlation function or as the first minimum of the auto-mutual information function [195]. The embedding dimension m ensures a proper unfolding of the trajectory in the reconstructed phase space. m can be chosen based on the false-nearest neighbor (FNN) algorithm [173, 196] or based on the number of dominant principle components. Essentially, these dimension estimators increase the embedding dimension until the geometric and dynamic properties of the object, formed by the reconstruction, do not change anymore. Apart from delay embedding, there exist other approaches such as differential embedding or singular value decomposition [197] of the Hankel matrix for the signal $\boldsymbol{s}(t_i)$. Most of the techniques for embeddings and dynamical invariants involve the selection of several problem-specific parameters, which renders a physically-meaningful signal processing in a highly-automated fashion challenging [152].

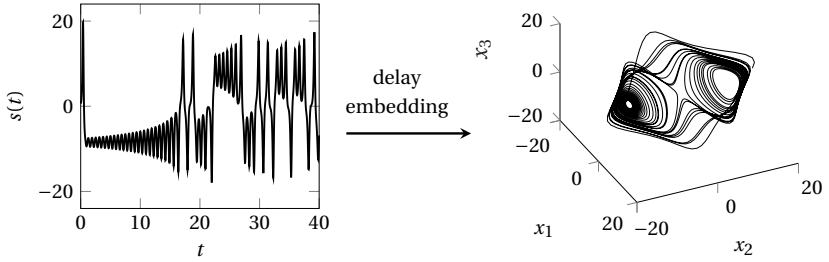


Figure (3.7) Reconstruction of a state space $[x_1, x_2, x_3]$ representation of an univariate time series $\boldsymbol{s}(t_i)$ using time delay embedding. The time series $\boldsymbol{s}(t)$, sampled at $\Delta t = 0.02$, is the first state of the Lorenz system in Equation (2.2), which was embedded into $m = 3$ dimensions using a time delay of $\tau = 10$ according to Equation (3.2)

Another branch of physically meaningful, i.e. interpretable, transformations for sequential data beyond time-frequency analysis have been introduced by N. Marwan and co-authors [198, 181] through the quantification of recurrence plots (RP). Introduced by Eckmann et al. [199] in chaos theory, the recurrence plot illustrates the short-term and long-term recurrent behavior of a trajectory in a m-dimensional space by pairwise comparison of the current state to all n states over time. Using a distance measure $|\cdot|$ and a small threshold value ϵ, the recurrence of a trajectory $\boldsymbol{x}(t_i)$ to the point $\boldsymbol{x}(t_j)$ in state space is indicated by $\boldsymbol{R}$

$$R_{i,j} = \theta\left(\left|\boldsymbol{x}_i - \boldsymbol{x}_j\right| - \epsilon\right), \quad i,j = 1,\ldots,n, \quad \boldsymbol{x} \in \mathbb{R}^m, \quad \boldsymbol{R} \in \mathbb{N}_{\{0,1\}}^{n \times n} \tag{3.3}$$

that uses the heaviside function θ to encode distances into a binary value. Hence, the RP comprises ones for the locations in phase space where x_i and x_j are closer than ϵ, and zeros for all other locations. The RP captures salient features of the underlying dynamics, such that it allows for a visual distinction between periodic, quasi-periodic, chaotic, and stochastic motion. $\bar{t}$-periodic signals create diagonal line structures as the trajectory runs through the same location after every period $\bar{t}$. Driven by the exponential divergence of nearby starting trajectories, irregular motion results in shorter and interrupted diagonal lines in the RP. Vertical and horizontal lines arise from intermittent dynamics, while stochastic signals generally create no distinct line structures in a RP. The recurrence plot quantification analysis (RPQA) defines metrics to quantitatively capture this visual representation. Exemplary, the determinism metric DET denotes the percentage of recurrence points that form diagonals, and the longest diagonal line $L_{\max}$ is related to the largest Lyapunov exponent. Several metrics have been proposed to measure the distribution of line lengths in the RP. Generally, metrics for diagonal line structures can indicate transitions from regularity to irregularity, while metrics for vertical line structures can reveal chaos-chaos transitions. Appendix

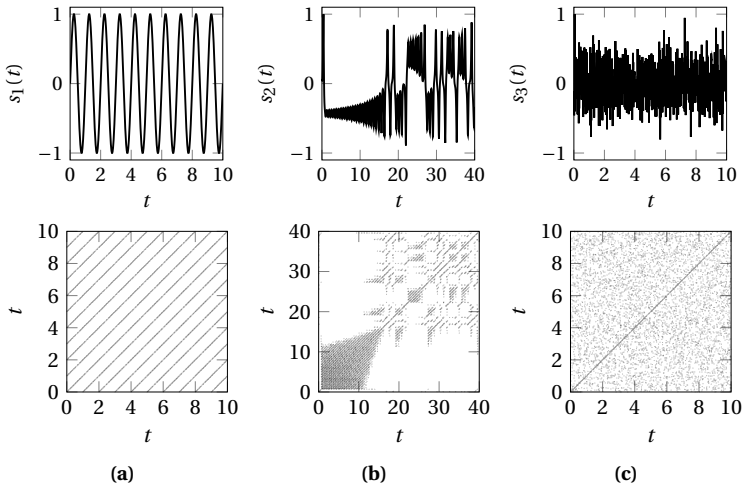


Figure (3.8) Recurrence plots for time series of canonical dynamical processes: (a) periodic signal $s_1 = \sin(2\pi t)$, (b) irregular motion in the first state of the Lorenz system (2.2) for $(\sigma = 10, \beta = 8/3, \rho = 28)$ and (c) Gaussian noise. The time sampling Δt, embedding dimension m, embedding delay τ and ϵ neighborhood are $[0.01, 2, 3, 0.01]_{(a)}$, $[0.05, 3, 4, 0.3]_{(b)}$, and $[0.01, 1, 1, 0.01]_{(c)}$, respectively. Adapted from Figure 4 from [152]

B.5 lists all common RPQA metrics and Figure 3.8 illustrates the recurrence plots for three representative signals from the classes of regular, irregular, and stochastic motion. RPQA and its variations can be considered established [200] in climate research, and the reader is referred to [181] for a comprehensive review.

Concluding, the selection of observables is the most crucial aspect for studying larger amounts of vibration measurements using data-driven techniques. Table 3.1 gives an overview on feature extractors and transformations that are mostly derived from physical principles and hence can be considered useful for the proposed physics-informed data science process.

3.3.4 The *modeling* phase

A detailed and comprehensive review of machine learning models is clearly out of scope of the thesis at hand. In the quickly emerging fields of the data sciences, new models are proposed on a weekly basis. However, there is a number of well-established methods that constitute the core building blocks of many more involved

Table (3.1) Overview of transformations and related extractors for an univariate time series in $\mathbb{R}^1$ to a different domain that is physically interpretable

domain	transformation	dimensionality
classical statistics	• 1^{st} order moments	$\mathbb{R}^1 \mapsto \mathbb{R}$
	• 2^{nd} order moments	$\mathbb{R}^1 \mapsto \mathbb{R}$
spectral analysis	• Fourier transformation (B.4)	$\mathbb{R}^1 \mapsto \mathbb{R}$
	• auto-correlation	$\mathbb{R}^1 \mapsto \mathbb{R}$
	• auto-mutual information	$\mathbb{R}^1 \mapsto \mathbb{R}$
time-frequency analysis	• Gabor transformation (B.5)	$\mathbb{R}^1 \mapsto \mathbb{R}^2$
	• wavelet transformation (B.6)	$\mathbb{R}^1 \mapsto \mathbb{R}^2$
nonlinear time series analysis [172]	• delay embedding (see Eq. (3.2))	$\mathbb{R}^1 \mapsto \mathbb{R}^m$
	• nonlinear invariant measures	$\mathbb{R}^1 \mapsto \mathbb{R}$
recurrence plot analysis [181]	• quantification measures (B.1)	$\mathbb{R}^1 \mapsto \mathbb{R}$
	• recurrence plots (Eq. (3.3))	$\mathbb{R}^1 \mapsto \mathbb{N}^2_{\{0,1\}}$

models [201]. A classical regression is often referred to as the simplest machine learning model. Support vector machines (SVM) map the input to a higher-dimensional representation through kernels, and then find hyperplanes that separate classes. Decision trees (DT) formalize a successive data splitting based on simple decision rules at the nodes of the tree. The outer leaves of the DT exhibit a high purity in the class distribution. Boosting and growing many DTs in parallel [202] has built a wide range of DT-inspired classifiers [203]. A perceptron, introduced by Rosenblatt in 1958 [204], multiplies the inputs ϕ_i with the weights w_i

$$\boldsymbol{y} = \sigma\left(\boldsymbol{W}\boldsymbol{\Phi} + \boldsymbol{b}\right) \tag{3.4}$$

and generates the output y_i through the activation function σ. Typically, a bias value b_i is added. Common activation functions are given in Table B.4. The arrangement of multiple perceptrons follows the general inspiration of neural science and the brain: An artificial neural network (ANN), built up from multi-layered perceptrons (MLP), is mostly inspired by the visual system of humans and animals, where several levels of abstractions are generated from a complex input. The models presented so far take a vector of features ϕ_i as input. However, spatially distributed data, such as an image, form multi-dimensional inputs. Furthermore, the location of a single feature, e.g. a pixel, in the overall array is relevant for forming distinct structures through neighborhood relations. Convolutional neural networks (CNN) [205] have proven to be highly

successful for such kinds of input data. Stacked CNN layers can learn increasingly abstract representations of the input, which allows to recognize lines, edges and, finally, geometric objects in a picture when used for computer vision. The order of input features matters also in the case of sequential data, such as time signals of speech or text. Here, recurrent neural networks (RNN) are used to learn about patterns that are encoded through correlations in the sequence. Other well-known learning strategies and model architectures are given by reinforcement learning, reservoir computers, auto-encoders, neural ordinary differential equations and others.

A fundamental approach to training models is the so-called error *backpropagation* [206] which is used for the majority of the models presented before. The models are formulated in terms of a large number of parameters w_i, which can be initialized randomly. Predictions for the training observations are generated by feeding the data through the network, the so-called *forward pass*. A loss function J measures the deviation of the model output from the desired output. The partial derivative of the loss with respect to every model parameter $\frac{\partial J}{\partial w_i}$ is computed in the *backward pass*. An optimization scheme, such as stochastic gradient descend [207] or adam [208], is utilized to update the parameters according to their individual contribution to the overall prediction error. This process is repeated until the model has converged to a local minimum with a satisfactory prediction quality in the error functional in the high-dimensional optimization space. The dimension of this space can easily grow into regions of 10^9 parameters for deep learning models. The reader is referred to the standard literature [201, 209, 210, 211, 212, 213] for in-depth discussions of machine and deep learning from various perspectives.

3.3.5 The *explain* phase

The explain phase may be considered the second most important aspect of the data science process in structural dynamics besides the feature engineering. In most generic data science projects, obtaining a high-quality prediction is the ultimate objective. On the contrary, data science for structural dynamics requires an extra effort to make actual use of data-driven techniques for engineering decisions through knowledge discovery. Essentially, the model decisions have to be traced back to the input quantities and, hence, the mechanical system at hand. Only then machine learning models can help to understand the role of a physical quantity, such as some component's eigenfrequencies or external loads, for the emergence of structural vibrations. In Figure 3.4 this objective is indicated by the inverse of the model and a feature relevance scoring. Linking the model output to the actual physical system can be achieved through explainability approaches. These techniques make the nonlinear relations and the decision making process accessible to humans. As a result, the

interpretation of data-driven models can indicate especially relevant input features, important relations between features and sensitivities of the output with respect to perturbations of the model input. For example, the analysis of a data-driven model may point out that the output, e.g. the vibrational behavior, is strongly driven by few inputs, e.g. certain modes of a sub-component and the phase of some forcing. Now, the data-driven modeling process has created actual value for understanding vibration phenomena that can guide engineers and designers.

3.4 Further remarks and potential pitfalls

The particular characteristics of vibration measurement data and engineering dynamics problems need to be taken into account in the transformation and the modeling phase. The discussions given hereafter highlight additional key aspects that the author finds to be crucial for the success of a data science research project in structural dynamics.

3.4.1 Explainability

In machine learning, there is a general trade-off to be made between model complexity and model interpretability. The more complex the input data and the underlying patterns are, the larger the required models are. However, such complex models lack direct ways to explain their decision making process [164]. Explainability denotes the ability of a human to a) understand the importance of individual features and b) understand the decision rules and relations between relevant features that lead to a model prediction. Even though deep neural networks can be written down as highly nonlinear equations, those representations are not intuitively interpretable. On the contrary, simpler models like decision trees formulate their decision rules by simple expressions in the fashion of *'if x_i is larger than a, then the observation belongs to class 1'*. Figure 3.9 depicts a small selection of models in the plane spanned by model accuracy and model interpretability. The region of interest is located at high model accuracy and high levels of explainability. Typically, this region can only be approached by explainable deep learning models.

Some models allow direct access to feature importance values, such as linear models and decision trees. For more complex models, there exist various approaches to obtain interpretable decisions. One of which are additive feature attribution methods that try to approximate a complex model by a simpler and explainable model. Other approaches rely on perturbations of the inputs, such as *LIME* [214]. These model-agnostic techniques are applicable to a broad range of machine learning mod-

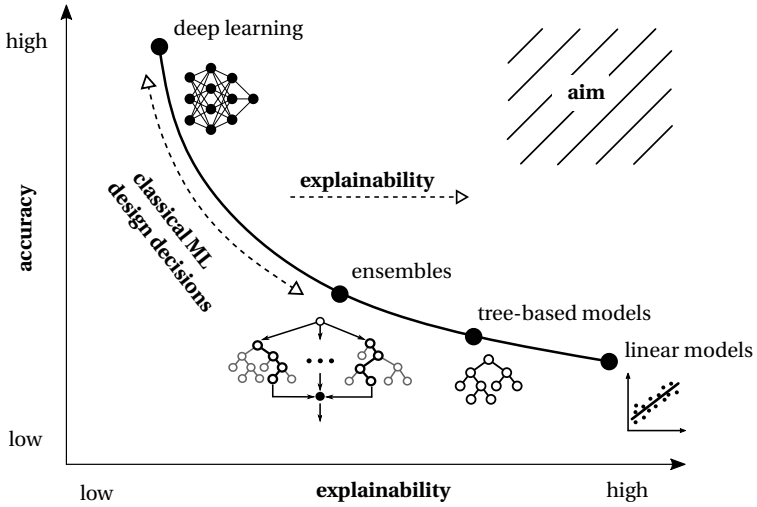


Figure (3.9) The spectrum of interpretability in the decision making process depends on the complexity of the chosen (machine learning) modeling approach. Explainability approaches help to make decision-making processes transparent while keeping the model complexity level

els. Other global approaches rely on the concept of Shapley values from game theory [215]. Additionally, particular model architectures enable the derivation of class activation maps [216] that indicate the most relevant regions in the inputs.

3.4.2 A-priori knowledge

Especially in the transformation and feature engineering phase, it is of crucial importance to have some a-priori knowledge of the essential properties of the data, dominant relations between dimensions and the type of observables required to capture the most discriminative characteristics of the physical process giving rise to the observations. The more expressive and discriminative the set of model input features is, the simpler the models can be, and consequently, the higher the degree of explainability is throughout the complete process. In the optimal case, the set of relevant features is clear from theoretical considerations or practical domain expertise. Then, only the modeling part needs to be researched to find the simplest model that

achieves a desired prediction quality. On the other end of the spectrum, it can be completely unclear which set of features will be required to model a physical process by means of data-driven techniques. In this case, two general approaches can be followed. First, an end-to-end learning path can be taken. The available data is provided in its raw form to a model that is able to process such kind of input. For example, recurrent neural networks can be used to process time series data. As the data is provided in the raw form without abstraction into a more condensed form, the required model may be rather complex. The second way for approaching a situation with very limited a-priori knowledge incorporates *massive feature extraction*. Here, a large number of transformations is applied to the data in order to extract as many features as possible. The underlying assumption is that some of the various features carry essential pieces of information that encode the input-output relation. If predictive modeling is possible using the full set of features, feature selection approaches are required in a second step. There is a wide range of selection approaches through filter, wrapper, and embedded techniques [217, 218] making use of correlations [219] in the input dimensions through hypothesis testing [220] and directly evaluating feature significance scores from models. Feature elimination schemes represent a very generic but computationally demanding approach [221] to find the most relevant features. For example, greedy backward elimination [180, 222, 223] recursively removes input features to find the one that affects the model performance the least. This process is continued until a highly discriminative set of features is found obtaining high prediction quality. To the knowledge of the author, today there are four publicly available software packages that provide functionalities for extracting massive amounts of features from time series data: *hctsa* [180, 224], *FATS* [225], *COTE* [226], and *tsfresh* [220]. Transformations from nonlinear time series analysis and recurrence plots are available in *hctsa*, which returns more than 7700 features for a single input time series. However, most of the features rooted in nonlinear dynamics and physics require manual inspection of the numerical results and careful selection of hyperparameters, such that those highly automated approaches may not extract the true values for the Lyapunov spectrum or others. Manual feature engineering may be required to derive physically meaningful features that allow for a deeper understanding of the underlying dynamics.

3.4.3 Measuring performance for small and imbalanced data

The confusion matrix, see Figure B.2 in Appendix B.8, is a standard tool to evaluate a classifier's prediction quality. In a binary classification task, a single prediction can be a *true positive* (TP), *false positive* (FP), *false negative* (FN), or *true negative* (TN). Several classification performance measures can be derived from the confusion ma-

trix, see Table B.2 in Appendix B.8. The selection of a quality metric strongly depends on the available data set size, the class distributions and the specific objective of the modeling problem. Mechanical vibrations data carry signatures of sparsity in multiple dimensions as discussed in Section 3.2. As a result, *small data* may be a frequent observation. There is no generic definitions of the terms *small* and *big* data, but the ratio of the number of input dimensions to the number of observations can be a valid indicator [227]. Small data sets render the training process sensitive to overfitting and the data split. In such a scenario, k-fold cross-validation makes use of the complete data set during training validates the model robust fashion. The data are split into k subsets, and during each fold the model is trained on $k-1$ subsets and validated against the remaining subset.

Secondly, the class distribution must be taken into account when splitting data for classification tasks. Mechanical structures and machine design have matured for centuries, such that vibration issues are exceptional cases and, optimally, occur only rarely. Hence, data from mechanical vibrations can be assumed to be inherently imbalanced [228, 229], such that one class (mostly the desired state of operation) is strongly over-represented in the recorded data. Class imbalance can be accounted for on the data level and on the model level [230]. On the data level, *resampling* denotes the process of artificially adding members of the minority class, removing members of the majority class or even creating new samples from the existing ones through sampling from the feature space [228]. However, if the data set is small, such techniques typically are not desired owing to weak statistical consistency. *Stratified splitting* is crucial when splitting imbalanced data sets for training and validation. Contrarily to random splitting, stratified splitting assures the same class distribution in each subset[3]. On the model level, one seeks to use cost functions and models that penalize misclassifications on the minority class stronger than misclassifications on the majority class [232]. For example, if the model is trained for the most common *accuracy* metric, it will tend to over-predict the majority class. Well-known scenarios for critically imbalanced classification tasks is the detection of tumor cells or credit card frauds [233]. Over-sensitivity to the positive class (malignancy, fraud) may not be desired, but has significantly less serious consequences than insensitivity, i.e. false negative predictions. Here, insensitivity would be fatal and potentially deadly, such that the number of false negatives must be kept very low. Even if machine dynamics research may not face as drastic consequences, appropriate selection of the quality metric is crucial. Especially with regards to strongly imbalanced data sets, *Matthews*

[3] It has been shown that increased numbers of minority class samples in the training set can increase the classifier performance [231]. However, in practice it remains unclear how to select the class distribution for the training set.

correlation coefficient (MCC) [234]

$$\mathrm{MCC} = \frac{\mathrm{TP}\cdot\mathrm{TN} - \mathrm{FP}\cdot\mathrm{FN}}{\sqrt{(\mathrm{TP}+\mathrm{FP})\,(\mathrm{TP}+\mathrm{FN})\,(\mathrm{TN}+\mathrm{FP})\,(\mathrm{TN}+\mathrm{FN})}} \tag{3.5}$$

is a classification metric that accounts for the actual distribution of positive and negative samples in the data set. The MCC ranges from (-1) (complete disagreement) over 0 (prediction no better than random) to (1) (complete agreement). The MCC is commonly used in bioinformatics [235] and other medical-related research domains.

3.4.4 Dynamics-compliant error measures

The beginning of this chapter claimed the importance to define a metric for measuring success in a data science project. In this thesis, it is proposed to select a quality metric that respects the underlying physics and the observed dynamics themselves. State prediction can be taken as an example, where the future evolution of some trajectory is achieved by learning from a sequence of historic time series data. Hence, given sequence data up to time t_0, the task is to predict the dynamics for future times $t_1, \ldots, t_n$. The prediction quality of some model can be measured by a distance metric between the ground truth and the predicted time series data per time step. However, depending on the qualitative character of the dynamics, different metrics need to be chosen. While the Euclidean is valid for periodic response data, chaotic dynamics require fundamentally different metrics. Owing the inherent property of exponentially diverging trajectories in chaotic systems, nonlinear invariant measures may represent appropriate distance metrics here. Hence, when evaluating the quality of a regression fit, or the generalization property of a neural network, specifically tailored metrics are favorable to arrive at a solution that is compliant with the underlying physics. This example may explain why sound domain knowledge in nonlinear dynamics is a profound prerequisite for the application of data-driven tools to structural dynamics.

3.5 Data science in structural dynamics today

The application of data-driven, data-assisted, or evidence-based approaches is not a new motive in dynamics. In fact, first system identification approaches using neural networks date back to 1990 [127, 137, 138]. Since then, various data-driven system identification, state observation and system diagnosis applications have been presented in the field of structural dynamics. Most of these contributions stem from the disciplines of structural health monitoring (SHM), fault diagnosis, and fault prediction. Particularly, the status and life-time prediction of bearings has been treated by

numerous authors owing to their versatile scope of application in any kind of machine and the critical failure thread to highly optimized production plants. Based on sparse vibration measurements (often, central components of machinery are inaccessible), a wide range of data treatment strategies and transformations have been proposed to find early warning signs in the observations [104, 236]. Still, spectral considerations dominate. Other applications involve wind turbines fault detection [104], milling machine diagnosis [237], and gear box dynamics [238, 239]. Linear correlation between several states of an aero engine, such as temperatures, with the vibrational response have been presented [240], but the transient character of the loads is not considered as such and explainability is prevented through PCA data reduction. Phase space reconstruction was used by Nichols et al. [241] to assess the status of a bolted joint. The geometric shape of the system's attractor was used to detect parametric changes [242] in the field of structural health monitoring.

Despite the various examples for data-driven approaches to structural dynamics, this thesis aims to make a new, and more methodical, scientific contribution. The proposed data science process captures most of the procedures that can be observed in individual studies and can therefore provide a guideline for future research. Furthermore, the proposed system description framework including structural dynamics measurements, loads, and meta data is a general consideration that can help practitioners as well as researchers to develop test campaigns that allow for data-driven treatment. To the knowledge of the author, the explicit consideration of peculiarities of structural vibration data and the underlying physics in the data flow is novel and has not been reported elsewhere. Furthermore, data-driven research and machine learning is rare in friction-affected and friction-excited systems today. The following Chapter 4 illustrates several application cases that follow the proposed data science process for system identification, state observation and future state prediction in various systems. New data-driven techniques are presented that help to overcome current limitations in the analysis of complex structural dynamics.

4 Case studies and novel methods

This chapter illustrates several novel methods in the framework of the proposed OPTME data science process for nonlinear mechanical vibrations. First, regression-based approaches with a high degree of interpretability are discussed. Second, vibration measurements are analyzed for identifying linear and nonlinear properties of the underlying dynamical system by means of signal pre-processing, several data transformations and exploratory data analysis. Neural network-based approaches are demonstrated in the third part for several challenges in friction-induced vibrations, such as vibration detection, characterization and prediction. Conceptually, the cases discussed hereafter span a wide range in the spectrum spanned by model complexity and model explainability that is depicted in Figure 3.9. Model complexity increases and the level of model interpretability decays along the course of this chapter. The specific methods and data dimensions used in every phase of the data science process are given in an overview table prior to each case study.

4.1 System identification through sparse regression

Parts of the work and figures presented hereafter have been published in

[167] M. Stender, S. Oberst and N. Hoffmann: *Recovery of Differential Equations from Impulse Response Time Series Data for Model Identification and Feature Extraction.* Vibration 2, 25-46, 2019

[243] M. Didonna, M. Stender, A. Papangelo, F. Fontanela, M. Ciavarella and N. Hoffmann: *Reconstruction of Governing Equations from Vibration Measurements for Geometrically Nonlinear Systems.* Lubricants 7, 64, 2019

and Table 4.1 gives an overview on the application of sparse regression techniques to mechanical vibrations for system identification.

System identification from experimental data is a wide research field in structural dynamics [1]. Most of the approaches originate in a-priori assumptions about the structure of the underlying mathematical description. Depending on how strict these assumptions are, the system identification can reduce to basic parameter fitting. Driven by computational fluid dynamics, distilling differential equations from time series

Table (4.1) Overview on the application case 1: sparse regression on time series data for direct model discovery and parameter identification

data	• time series data $\in \mathbb{R}^{m \times n_t}$, m states, n_t time steps
pre-processing	• noise filtering • regularized numerical differentiation
transformation	• phase space reconstruction through embedding
model	• sparse regression through l_1 penalization • polynomial ansatz function space
explain	• identification of underlying differential equations • quantification of stiffness and damping values • test for nonlinearity • time series classification

data has become a rapidly developing field in recent years [132, 179, 244]. The main objective is to infer information on the mathematical representation of the underlying system. There are two conceptual modes of identification: identification of a mathematical model that describes the data as precise as possible, and identification of the underlying dynamical system that actually created the data. The second mode is clearly more challenging, since there might be various models that can describe the current data, while the actual underlying system is unique. When the identification of a mathematical structure is successful, the analysis of the equations of motion can explain variable interactions, parametric dependencies and stability. Governing equations in physics are often sparse in the sense that few terms suffice to capture complex behavior, such as in the case of the Navier-Stokes equations. Hence, the major goal of system identification approaches from time series data is to arrive at a mathematical model that balances complexity and accuracy. Early works go back to 1980 [245], the introduction of the *flow method* by Crutchfield et al. in 1986 [246] and the *trajectory method* by Eisenhammer et al. in 1991 [247]. However, these approaches were computationally intractable at that time and sensitive to noise. In 2007, Bongard, Lipson et al. [248, 249] introduced a method to derive sparse differential equations from time series data using symbolic regression and genetic programming, which are compute-intensive algorithms. Other approaches make use of compressive sensing [133] as summarized in the review by Daniels et al. [250]. Brunton et al. [132] propose to combine the system discovery with sparsity promoting regression techniques under the term of *sparse identification of nonlinear dynamics* (SINDy). Besides the formulation for ordinary differential equations, there are also extensions to partial differential equations [251, 252] and nonlinear rational functions [253] that build on the sparse

regression idea described hereafter. SINDy makes use of nonlinear models that are parametrized by linear terms which allows solving the regression problem with classical and sparsity-promoting methods. The SINDy procedure is outlined in the following with a notation that is closely adapting the one by Brunton et al. [132, 254]. Dynamical systems

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}(t)) \tag{4.1}$$

with states $\boldsymbol{x}(t) = [x_1(t), x_2(t), \ldots, x_m(t)]^\top$ are considered. SINDy seeks to identify an analytical representation of $\boldsymbol{f} : \mathbb{R}^m \mapsto \mathbb{R}^m$ from noisy and discrete measurements of $\boldsymbol{x}(t), t = t_1, \ldots, t_n$ using a set of candidate functions. The measurements $\boldsymbol{x}(t_i)$ and time derivatives $\dot{\boldsymbol{x}}(t_i)$ are collected in $\boldsymbol{X}$ and $\dot{\boldsymbol{X}}$

$$\boldsymbol{X} = \begin{bmatrix} x_1(t_1) & x_2(t_1) & \cdots & x_n(t_1) \\ x_1(t_2) & x_2(t_2) & \cdots & x_n(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(t_m) & x_2(t_m) & \cdots & x_n(t_m) \end{bmatrix}, \quad \dot{\boldsymbol{X}} = \frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t} \quad . \tag{4.2}$$

A set of p candidate functions for representing $\boldsymbol{f}$ is specified in $\{\phi_i(\boldsymbol{x}) : i = 1, 2, \ldots, p\}$, such as polynomials or sinusoidal terms. A library matrix of nonlinear functions is obtained by evaluating the candidate functions on the measurements $\boldsymbol{x}(t_i)$

$$\boldsymbol{\Phi}(\boldsymbol{X}) = \begin{bmatrix} | & | & & | \\ \phi_1(\boldsymbol{X}) & \phi_2(\boldsymbol{X}) & \cdots & \phi_p(\boldsymbol{X}) \\ | & | & & | \end{bmatrix} \tag{4.3}$$

where $\phi_i(\boldsymbol{X})$ denotes applying a single candidate function i to each row of $\boldsymbol{X}$. Collecting all the potentially nonlinear functions in $\boldsymbol{\Phi}(\boldsymbol{X})$ transforms the approximation of a nonlinear function $\boldsymbol{f}$ into a regression problem for the linear combinations of the library functions $\phi_i(\boldsymbol{x})$

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) \mapsto \dot{\boldsymbol{X}} = \boldsymbol{\Phi}(\boldsymbol{X})\,\boldsymbol{\Xi} \tag{4.4}$$

where the jth column $\boldsymbol{\xi}_j$ of $\boldsymbol{\Xi}$ collects the coefficients of the library functions ϕ_i that constitute the representation of $\dot{\boldsymbol{x}}_j$

$$\dot{\boldsymbol{x}}_j = \boldsymbol{f}_j(\boldsymbol{x}) = \boldsymbol{\Phi}\left(\boldsymbol{x}^\top\right)\boldsymbol{\xi}_j \quad . \tag{4.5}$$

The full representation of the dynamical system is given by

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{\Xi}^\top \left(\boldsymbol{\Phi}\left(\boldsymbol{x}^\top\right)\right)^\top \tag{4.6}$$

when exact measurements of the states $\boldsymbol{x}$ are assumed to be available. In practice, noise perturbs the measurements, such that the dynamical system is only approximated by the solution of

$$\dot{\boldsymbol{X}} \approx \boldsymbol{\Phi}(\boldsymbol{X})\,\boldsymbol{\Xi} \quad . \tag{4.7}$$

Classical solution strategies, such as least squares, would result in a fully populated matrix Ξ. To enforce sparsity in the coefficients $\boldsymbol{\xi}_i$, the regression problem in Equation (4.4) is formulated as a minimization problem

$$\min_{\Xi} \left(\frac{1}{2} \left| \dot{\boldsymbol{X}} - \boldsymbol{\Phi}(\boldsymbol{X}) \Xi \right|_{\mathrm{F}}^2 + (\Xi)_{l_1} \right) \tag{4.8}$$

using the Frobenius norm $|\cdot|_{\mathrm{F}}$ with l_1 penalty regularization. A sequential thresholded least-squares algorithm [132, 255] is employed to solve the minimization problem. This scheme alternates between an unregularized least squares solution for each column of Ξ and subsequent removal of candidate function coefficients that are smaller than the sparsification threshold λ. The selection of this sparsification parameter is crucial for obtaining accurate approximations that are sparse, i.e. the common motive of balancing model quality with model complexity. The regression problem is typically highly over-determined because the number of time steps n is much larger than the number of candidate functions in the library $\boldsymbol{\Phi}$. Figure 4.1 depicts a schematic of

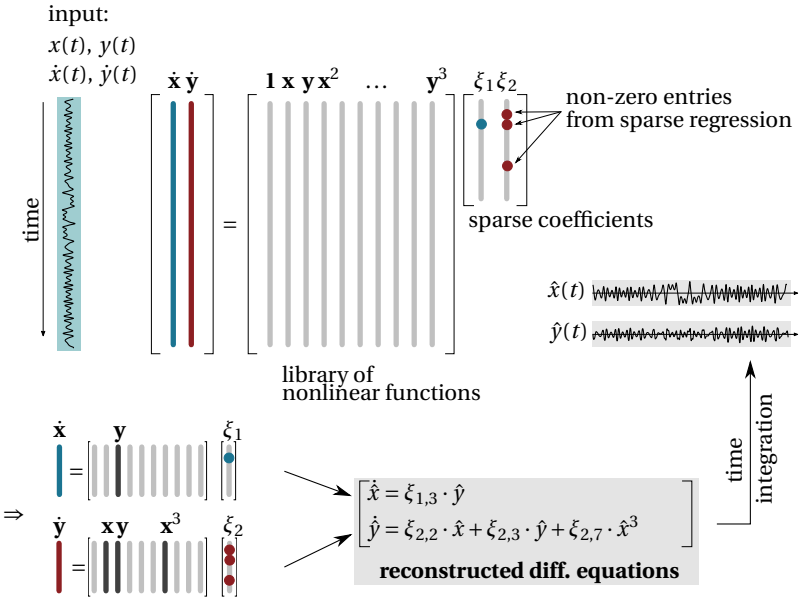


Figure (4.1) Concept of SINDy adapted from [132] and [167]. Time-marching solutions of the identified system $\dot{\boldsymbol{x}} = \hat{f}(\boldsymbol{x})$ can be used to arrive at error measures with respect to the input signals $\boldsymbol{x}$. Adapted from Figure 1 from [167]

the SINDy procedure that derives differential equations for describing the observed dynamics. Once a mathematical description of the dynamics is obtained, it can be integrated in time through time marching schemes starting from the first state $\boldsymbol{x}(t_0)$ for validation against the measured dynamics. Furthermore, stability and bifurcation analysis are possible based on the identified equations of motion. As an example, the linear mechanical oscillator $\ddot{x} + 2\zeta\omega_0\dot{x} + \omega_0^2 x = 0$ in first-order form

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -2\zeta\omega_0 x_2 - \omega_0^2 x_1 \end{aligned} \tag{4.9}$$

is considered. If the candidate functions are selected from the polynomial ansatz space $\phi_i = \{1, x_1, x_2, x_1x_2, x_1^2, x_2^2\}$, the result of the sparse regression (4.6) reads

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -\omega_0^2 & -2\zeta\omega_0 & 0 & 0 & 0 \end{bmatrix}}_{\boldsymbol{\Xi}^\top} \underbrace{\begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1x_2 \\ x_1^2 \\ x_2^2 \end{bmatrix}}_{(\boldsymbol{\Phi}(\boldsymbol{x}^\top))^\top} \tag{4.10}$$

given that complete and exact measurements of $x_1(t)$ and $x_2(t)$ are available. The application cases illustrated hereafter use variations of this mechanical oscillator to examine capabilities and limitations of SINDy for mechanical vibrations especially for scenarios with limited data availability.

First, a novel nonlinear optimization scheme is presented that allows obtaining parsimonious models that have a better balance of complexity and precision than the models obtained from the original formulation of SINDy. Second, different levels of data limitations are discussed. Third, canonical cases illustrate how SINDy can be used for model identification, for nonlinearity testing and as a feature generator for a data science process. As emphasized in Section 3.2, a quality metric must be chosen to evaluate the performance of the data science process. In case of SINDy and direct system identification of underlying differential equations, there are two possibilities for defining the quality metric: If the true dynamical system is a-priori known, the actual differences between the correct and the reconstructed set of differential equation may be evaluated. However, this scenario is only encountered in academic studies. In experimental studies, the underlying system is at least partially unknown. Hence, a quality metric is chosen that is based on time series data. In the following studies, only (damped) periodic motions are studied such that an instance-bases distance metric is appropriate. The absolute difference between the observed trajectory

$\boldsymbol{x}(t_j), j = 1,\ldots,n$ and the one stemming from the time integration of the identified dynamical system $\hat{\boldsymbol{x}}(t_j)$ is computed

$$J = \sum_{i=1}^{m} \sum_{j=1}^{n} \left| \frac{x_i(t_j)}{\max\left(|\mathrm{x_i}|\right)} - \frac{\hat{x}_i(t_j)}{\max\left(|\mathrm{x_i}|\right)} \right| \tag{4.11}$$

taking into account the absolute value range of each state $x_i, i = 1,\ldots,m$ by normalization. The error measure J takes every state x_i equally into account. For other types of dynamics, such as chaos, the error measure may be formulated in terms of dynamical invariant measures.

pre-processing: Time series data sampled at high rates is required as input to SINDy because the sparse regression is computed on the time series data directly. Therefore, noise should be low, or reduced by filtering. The time derivative $\dot{\boldsymbol{x}}$ is required for the regression, but is most often not available through direct measurement. Hence, numerical derivatives with respect to time need to be computed. However, the inevitable noise contamination in measurement data amplifies through classical numerical differentiation schemes. Specialized schemes, such as the total variation regularized numerical differentiation (TVRegDiff) [256], can help to regularize this process in order to arrive at a high-quality approximation of the derivative. The differentiation step can be validated by integrating the numerical derivative along time and comparing the result to the original time series [167]. If the time-integrated version of the numerical derivative is consistent with the original data, the differentiation process is valid.

transformation: Measurement data for mechanical vibrations are typically sparse, see Section 3.2. Especially, only a small number of measurement points are available, such that only a fraction of all active degrees of freedom can be captured. However, SINDy requires time series data from m states to reconstruct a m-dimensional system of ordinary differential equations (ODE). The required number of degrees of freedom depends on the type of dynamics observed: Two active states suffice to represent a periodic motion, while at least three states are required for chaotic dynamics. Hence, embedding procedures, such as the time delay embedding, see Equation (3.2), can be utilized to generate m time series from an univariate vibration time series measurement.

modeling: SINDy is a l_1-constrained regression of $\dot{\boldsymbol{X}} = \boldsymbol{\Phi}(\boldsymbol{X})\boldsymbol{\Xi}$ on the measured signals $\boldsymbol{X}$ and the nonlinear function library $\boldsymbol{\Phi}$. The selection of the nonlinear candidate functions is a crucial ingredient to the success of SINDy. A-priori domain knowledge is useful for selecting a family of functions ϕ_i. For mechanical vibrations, one may start with monomials of some low order. The more candidate functions are considered, the higher is the risk of overfitting the data with too complex models.

explain: The case study on SINDy for structural vibrations covers three aspects: a) identification of differential equations from limited data, b) a novel optimization approach for higher accuracy and simpler models and c) nonlinearity tests and feature extraction through sparse regression.

The linear mechanical oscillator in Equation 4.9 is studied in the damped configuration for $\omega_0 = 8$ and $\zeta = 0.2$, for the initial conditions $[x_1(t=0), x_2(t=0)]^\top = [1,0]^\top$ and a time history of 500 time steps sampled at $\Delta t = 0.01$. Hence, the original set of differential equations reads

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -3.2x_2 - 64x_1 \end{aligned} \tag{4.12}$$

For a proof of concept, the complete data of $\boldsymbol{x}(t)$ and $\dot{\boldsymbol{x}}(t)$ is provided to SINDy and polynomial library functions up to a degree of $p = 3$, i.e. $\phi_i : \{x_1, x_2, x_1x_2, \ldots, x_2^3\}$, are used. A set of two differential equations is reconstructed from the data, and the time integration of those equations for initial conditions $[1,0]^\top$ yields the time signals $\hat{\boldsymbol{x}} = [\hat{x}_1(t), \hat{x}_2(t)]^\top$. The error J is evaluated according to Equation (4.11) and the model sparsity can be measured by the fraction of non-zero elements (NZE) in $\boldsymbol{\Xi}$ which directly depends on the sparsification threshold λ. Figure 4.2 depicts the reconstruction error J and the model sparsity as a function of the threshold λ. The objective, according to the proposed data science process, is to find a parsimonious model with low error, which is typically referred to as the *pareto front*. In this case, vanishing error values are obtained for threshold values lower than $\lambda = 1.0$, which is the smallest coefficient in the regression in Equation 4.10. Although higher-order polynomials were provided as candidates in $\boldsymbol{\Phi}$, the differential equations (4.12) are identified exactly.

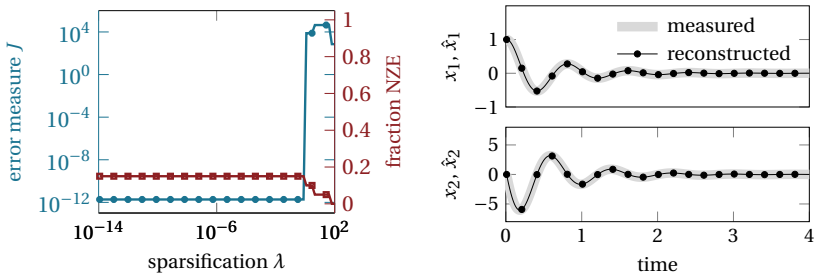


Figure (4.2) (a) model error as a function of the sparsification λ and the fraction of non-zero elements (NZE) in $\boldsymbol{\Xi}$. (b) time series $\boldsymbol{x}$ used as input to SINDy the response $\hat{\boldsymbol{x}}$ of the reconstructed dynamical system using $\lambda = 0.5$. Adapted from Figure 6 from [167]

Next, the derivatives $\dot{\boldsymbol{x}}(t)$ are assumed to not be available. Numerical differentiation of $\boldsymbol{x}(t)$ is performed using TVRegDiff with a regularization parameter value $\alpha = 1.87 \cdot 10^{-9}$ which gives the minimal error between the original states and the time-integrated version of the numerical derivatives. The lowest error J of the SINDy reconstruction is obtained for a sparsification of $\lambda = 0.3583$ using polynomials up to order $p = 3$. The time evolution is captured well, but in the reconstructed differential equations

$$\begin{aligned} \dot{\hat{x}}_1 &= 0.999\hat{x}_2 \\ \dot{\hat{x}}_2 &= -64.8124\hat{x}_2 - 2.9204\hat{x}_1 - 0.5125\hat{x}_1^3 \end{aligned} \tag{4.13}$$

an artificial stiffness term can be observed. In fact, the cubic stiffness term is introduced by SINDy to create larger restoring forces and compensate for the lack of damping with respect to the original system (4.12). If the present SINDy reconstruction was used as a test for nonlinearity, one would conclude that the underlying system features some cubic stiffness effects. However, the thresholded iterative least squares procedure for the highly over-determined system may not have found the simplest model in the given function space. This situation is addressed by a novel nonlinear optimization procedure appended to SINDy. The central idea is to formulate an optimization problem for lower reconstruction error that promotes sparsity of the model. Based on the assumption that SINDy has found the correct locations of non-zero coefficients, but not their optimal value, only the non-zero coefficients ξ_i are subjected to the optimization. The lower and upper value ranges ξ_-, ξ^+ are selected relatively to the ξ_i value using a scale parameter $\gamma > 1$. Importantly, the zero is included in the optimization search space, such that further sparsification of the model is allowed. The bounded optimization for better coefficients ξ_i

$$\min_{\mathrm{NZE}(\Xi)} \left(J\left(\xi_i\right)\right), \quad \xi_-^+ = \left[0, \xi_i \cdot \gamma\right] \tag{4.14}$$

is carried out using a sequential quadratic programming algorithm implemented in the Matlab procedure *fmincon*. As an example, the optimization boundaries for the SINDy reconstruction given in Equation (4.13) are $\xi_- = -0.77$ and $\xi^+ = 0$ for $\gamma = 1.5$ applied to the cubic stiffness term $\xi = -0.5125$. Overall, the four non-zero coefficients for $\hat{x}_2$ and $\hat{x}_1$, $\hat{x}_2$, $\hat{x}_1^3$ in (4.13) are subjected to the optimization. Here, the optimization finishes after 48 iterations, see Figure 4.3 (a), and returns the optimal coefficients with respect to the cost function in Eq. (4.11) and the imposed constraints on the coefficients ξ_-^+. The optimized equations

$$\begin{aligned} \dot{\hat{x}}_1 &= 0.99\hat{x}_2 \\ \dot{\hat{x}}_2 &= -3.19\hat{x}_2 - 64.43\hat{x}_1 - 0.0\hat{x}_1^3 \end{aligned} \tag{4.15}$$

yield an error reduction by 82%. In fact, the cubic stiffness term was eliminated by the optimization and the linear stiffness and damping terms were increased. Fig-

ure 4.3 (b) depicts the original input time series, the SINDy reconstruction and the optimized reconstruction which features a sparser model in the chosen basis and a smaller error in time domain. If the amount of available data is further restricted,

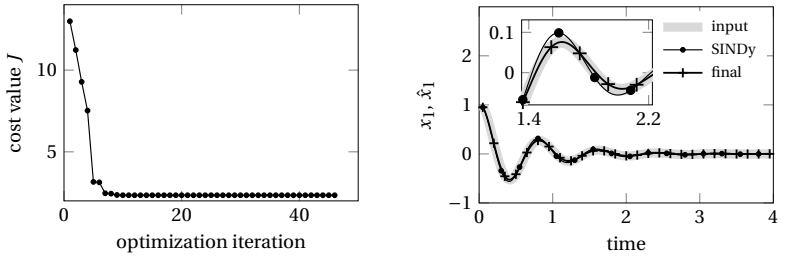


Figure (4.3) (a) optimization iterations for the optimal coefficient values in equation (4.13) using (4.14). (b) depicts the input series, the SINDy result (4.13) and the final reconstruction (4.15) obtained from the optimization result. Adapted from Figure 8 from [167]

one can assume to have knowledge only about the first state $x_1(t)$. Then, the $m = 2$ dimensional state space can be reconstructed as $\left[q_1(t), q_2(t)\right]^\top$ using time delay embedding (3.2) with delay $\tau = 18$ and the derivatives can be obtained via TVRegDiff. The embedding transforms the dynamical system into another coordinate system, such that the reconstructed differential equations for q_1 and q_2 cannot be compared directly to the ones of the mechanical oscillator in x_1 and x_2. Still, the character of the equations and the non-zero candidate function terms can be compared qualitatively. Table 4.2 reports the error value J and the non-zero coefficients found through SINDy and subsequent optimization for increasing polynomial orders p in the library of nonlinear candidate functions. Parsimonious mathematical models are reconstructed from the time delay-embedded signals. The general structure of the equations is independent from the function library because the linear coupling terms between $\dot{\hat{q}}_{1,2}$ are very similar for each reconstruction. To decrease the error, small nonlinear terms are introduced for higher polynomial degrees p.

SINDy is not only useful for system identification and nonlinearity testing as described in more detail in [167], but it can also be used as a feature extractor. The following study illustrates how SINDy can be employed to derive a small set of discriminative features that can be used for time series classification. An nonlinear adaptation of the linear oscillator (4.12)

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -2\zeta\omega_0 x_2 - \omega_0^2 x1 - k_{\mathrm{nl}} x_1^3 \end{aligned} \tag{4.16}$$

Table (4.2) Non-zero entries of the coefficient matrices for the optimized SINDy reconstruction for maximal polynomial degrees p and resulting reconstruction error J. The states q_1 and q_2 are obtained from the time delay-embedded state x_1 of the linear oscillator

order	error	reconstructed differential equation
$p=1$,	$J=1.88$	$\dot{\hat{q}}_1 = -0.74\hat{q}_1 - 5.84\hat{q}_2$ $\dot{\hat{q}}_2 = -0.008 + 10.60\hat{q}_1 - 2.45\hat{q}_2$
$p=2$,	$J=1.82$	$\dot{\hat{q}}_1 = -0.75\hat{q}_1 - 5.82\hat{q}_2 - 0.06\hat{q}_1\hat{q}_2$ $\dot{\hat{q}}_2 = -0.007 + 10.62\hat{q}_1 - 2.45\hat{q}_2$
$p=3$,	$J=1.63$	$\dot{\hat{q}}_1 = -0.81\hat{q}_1 - 5.84\hat{q}_2 - 0.13\hat{q}_1\hat{q}_2 + 0.51\hat{q}_1^2\hat{q}_2 + 0.16\hat{q}_2^3$ $\dot{\hat{q}}_2 = 10.60\hat{q}_1 - 2.39\hat{q}_2$
$p=4$,	$J=1.64$	$\dot{\hat{q}}_1 = -0.85\hat{q}_1 - 5.86\hat{q}_2 + 0.7080\hat{q}_1^2\hat{q}_2 - 0.96\hat{q}_1^3\hat{q}_2 - 1.44\hat{q}_1\hat{q}_2^3$ $\dot{\hat{q}}_2 = 10.54\hat{q}_1 - 2.36\hat{q}_2$

is used to generate 40 time series for the linear ($k_{\mathrm{nl}} = 0$) and 40 time series for the nonlinear ($k_{\mathrm{nl}} = 50$) configuration by a linear variation of ω_0^2 in the range of $\omega_0^2 \in \{36, 400\}$. Figure 4.4 (a) depicts the resulting time series which are hardly distinguishable by visual inspection.

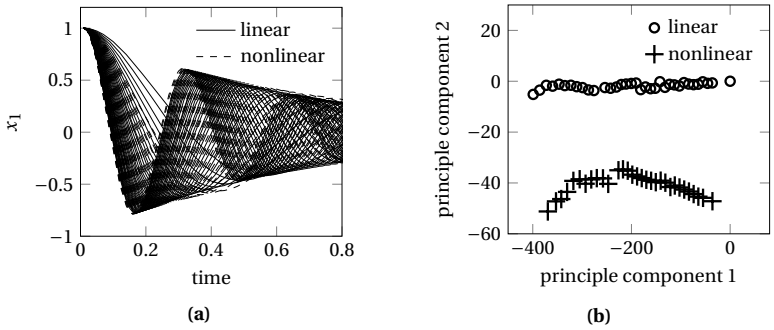


Figure (4.4) (a) 40 time series for the linear and nonlinear configuration of the oscillator (4.16) for varying frequencies ω_0. (b) Projection of the candidate function coefficients from the reconstructions onto their first two principle components for the time series classification purpose. Adapted from Figure 14 from [167]

Especially for high linear stiffness values, the nonlinear effects are small and possibly hard to detect using classical methods such as decay curve fitting. SINDy can be employed to realize the classification of those signals into the classes of *linear* and

nonlinear vibrations. Governing equations are reconstructed from the $[x_1, x_2]$ signals using polynomials of third order such that the coefficient matrix $\boldsymbol{\Xi}$ has 20 entries. These candidate function coefficients are treated as a feature set for describing the observed trajectories. Most of these coefficients will be zero anyway. Thus, not all of those 20 features may be essential for the classification of the 80 signals into the two classes of linear and nonlinear dynamics. Principle component analysis is used to reduce the set of 20 features for each of the 80 observations to a two-dimensional representation. The projection of the full feature vector onto the first two principal components is depicted in Figure 4.4 (b). Here, two distinct clusters can be observed which allow for a perfect separation of the linear and nonlinear trajectories. The large variation in the first component corresponds to the linear stiffness variation, while the two clusters along the second component carry the signature of nonlinearity, i.e. coefficients of nonlinear terms in the differential equations reconstructed by SINDy. If a new and unknown time series was subjected to the analysis, the features derived by SINDy could be used to classify whether the underlying dynamical system was linear or nonlinear.

Concluding, the sparse identification of nonlinear dynamical systems can also be used as an alternative feature extraction approach that is highly interpretable and has a controllable complexity. Given some a-priori knowledge about the system dimensionality and appropriate building blocks for the underlying differential equations, SINDy is a promising tool for direct system identification and feature extraction. The presented studies are performed on synthetic data from a simplistic model to evaluate the possibilities and limitations before considering realistic vibration measurements. In a second study [243], SINDy was used to identify the differential equations governing the dynamics of a base-excited cantilever beam in a laboratory experiment. It turned out that the signal pre-processing and filtering step is crucial for a successful system identification. The periodic response of the beam was captured well by the reconstructed differential equations which allowed inferring modal stiffness and damping values from the coefficient matrix.

4.2 Vibration characterization through nonlinear time series analysis

The objective of this study is to characterize and quantify the friction-induced vibrations in a seemingly simplistic setup of two blocks in dry frictional contact. Table 4.3 gives an overview on conceptual setting following the data science process introduced in Chapter 3. Parts of the work and figures presented hereafter have been published in the publication

[257] M. Stender, M. Di Bartolomeo, F. Massi and N. Hoffmann: *Revealing transitions in friction-excited vibrations by nonlinear time-series analysis.* Nonlinear Dynamics 47, 209, 2019.

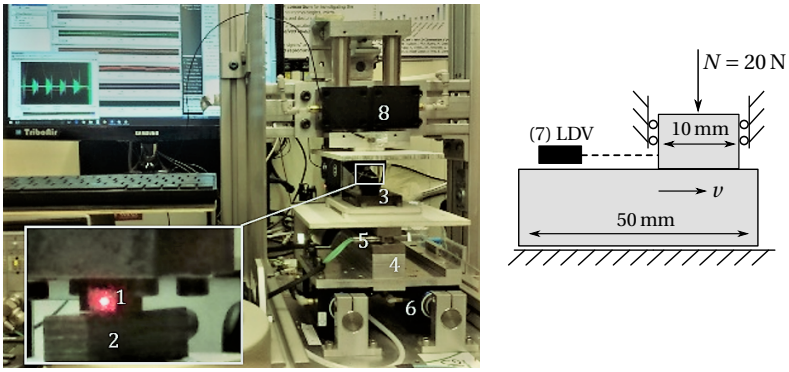
Table (4.3) Overview on the application case 2: characterization and transition detection in friction-induced vibrations

data	• univariate time series data $\in \mathbb{R}^{1\times n_t}$, n_t time steps
pre-processing	• temporal upsampling
transformation	• time-frequency analysis • state space reconstruction through embedding • recurrence plot quantification analysis
model	• exploratory data analysis
explain	• detect dynamical transitions • characterize regular and irregular vibrations • study stationarity in friction-induced vibrations

obtain: In the experimental setup *'TriboAir'*[1] [258] a fixed block of size [10 × 10 × 10] mm is in dry contact with a moving sample of size [50 × 20 × 10] mm. The translation velocity profile quickly ramps up to 20 mm/s and then linearly decreases to zero during a time of 4 seconds. The contact normal force of $N = 20$ N is applied through the upper (fixed) sample and the translation is induced through a voice-coil motor attached to the lower sample. All moving parts are supported by air bearings to decouple the specimen from the testing environment. Additionally, parasitic vibrations and the impact of geometrical boundary conditions are reduced. The vibration velocity of the fixed sample is monitored by laser doppler vibrometry (LDV) at 1.5 mm distance from the friction interface at a sampling rate of $f_s = 100$ kHz. Furthermore,

[1] The measurements were performed at La Sapienza University of Rome by Mariano Di Bartolomeo and Franceso Massi

the tangential force on the sliding specimen and the normal force are measured. Figure 4.5 depicts a photograph of the set-up and a schematic illustration of the experiment. Several samples of different material combinations were tested in this setup, but only the results for one friction pair are reported here. The specimens are made of composite carbon/carbon materials that are used in high-performance brake systems and clutches. The material features a matrix structure made of carbon fibers, fillers and additives. The frictional pair exhibits a stationary friction coefficient with respect to the sliding velocity. As the relative sliding velocity decreases during the experiment, friction-induced vibrations emerge and a plethora of dynamical transitions can be observed. This study aims at characterizing and quantifying the strongly nonlinear and potentially chaotic vibrations of the frictional system.



(a) (b)

Figure (4.5) (a) photograph of the 'TriboAir' setup: upper sample with laser measurement point 1, lower sample 2, supporting bar 3, moving base 4, tri-axial force transducers 5, horizontal air bearing 6, laser vibrometer 7, vertical air bearings 8. (b) schematic illustration of the fixed upper sample and the moving lower sample. Adapted from Figure 1 from [257]

pre-processing: Upsampling of the vibration signal to a frequency of $f_s = 1$ MHz was conducted to fully unfold the high-frequency dynamics. As the measurement noise in this setup is exceptionally low, no impact was observed on the results presented hereafter.

transform: Fourier analysis, nonlinear time series analysis (NTSA) and recurrence plot quantification analysis (RPQA) are employed to characterize the self-excited vibrations and detect dynamical transitions. Figure 4.6 depicts the short-time Fourier transform (STFT) for the laser vibration measurement of the fixed sample, representative time signals at various time instants, the corresponding Fourier transforms, time

delay embedding state space reconstructions according to Equation (3.2), the resulting recurrence plots (RP) as well as series of recurrence quantifiers from Table B.1. The instantaneous embedding parameters τ_{AMI} and m are estimated by the auto-mutual information function (AMI) and the false-nearest neighbor (FNN) algorithm, respectively, for the state space reconstruction. The fixed recurrence rate RR= 5% guarantees a consistent formulation of the recurrence plots for qualitatively different dynamics. All sequential analysis is derived from a sliding window processing with window size $w = 2000$ samples and 87.5% overlap. For consistency of the RPQA, the trajectories are over-embedded into $m = 5$ dimensions and the minimal line lengths l_{min}, v_{min} are set to 10. As the qualitative nature of the vibrations is of interest, all signals are normalized to unit maximum amplitudes. The STFTs display the amplitude spectrum in logarithmic scale.

exploratory data analysis: As the sliding velocity decreases, friction-induced vibrations (FIV) arise after $t = 1.56$ s during steady sliding. Stick-slip cycles set in at $t = 3.13$ s. Throughout the complete measurement, the friction coefficient remains approximately constant [257]. There exists a window of purely periodic dynamics for $2.28 \leq t \leq 2.40$ s where the spectral energy is confined to the main frequency at 9.62 kHz and two higher harmonics. An annularly shaped attractor ca be observed in the reconstructed state space which indicates trajectories on a limit cycle. Long diagonal lines measured by L and L_{max} confirm regularity of the dynamics. The spectral content flattens along time until several frequencies at 7.6, 9.5 and 11.4 kHz and a broad background spectrum are excited at $t = 2.92$ s. The RP exhibits many short and some longer diagonals which indicate weakly irregular motion. Shortly after at $t = 2.95$ s, the dynamics transit to a multi-periodic window with the fundamental frequency at 2.29 kHz. Many higher harmonics are excited and exhibit higher energy shares than the fundamental frequency. As subharmonics are typically only known in the context of forced vibrations, this observation remains an interesting phenomenon. The RP and the reconstructed attractor exhibit similarity to quasi-periodic motion. After the next transition at $t = 3.01$ s, a frequency of 7.57 kHz dominates the spectrum with three higher harmonics. The recurrence plot and its quantifiers indicate regular motion on a limit cycle. Even if the measurements are highly precise, there is some low-frequency modulation in the signals which breaks some of the lines in the RPs. From $t = 3.13$ s on, stick-slip motion exists. Clear sticking and slipping phases are observable in the time signal. The sticking phase represents intermittent dynamics that form block-like structures in the RP and longer vertical lines V_{max}.

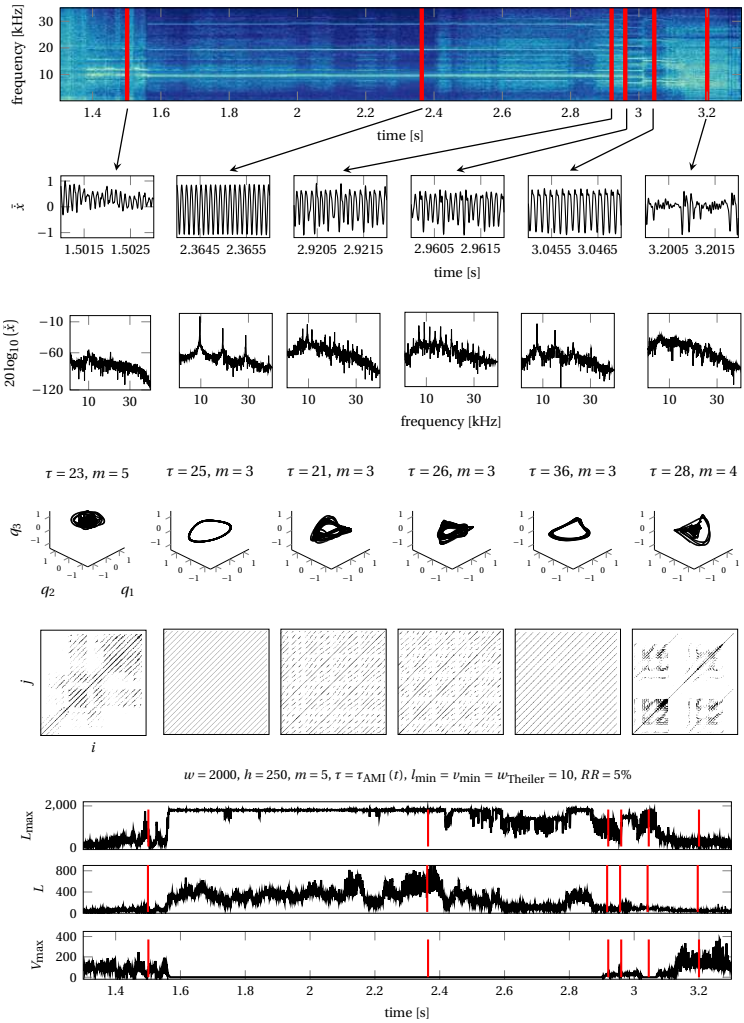


Figure (4.6) From top to bottom: short-time Fourier transform of the normalized vibration velocity signal, representative time segments and their Fourier transforms, state space reconstructions, recurrence plots and three RPQA metrics. Adapted from Figure 9 from [257]

Figure 4.7 depicts a detailed RPQA analysis of the transition phase between high-frequency steady sliding FIV to stick-slip motion. Only three RPQA metrics are shown to indicate regular-irregular transitions (by means of L and $L_{\max}$) and chaos-chaos transitions (by means of $V_{\max}$). The vibration characteristics can be studied visually in the time-frequency analysis. However, the state space reconstruction and subsequent recurrence plot analysis provide a much clearer picture of stationary regimes and sudden dynamical transitions. Furthermore, the nonlinear analysis allows to quantify the observed dynamics by means of the recurrence quantifiers. Contrary to spectral properties, the RPQA metrics carry information about the fundamental properties of the dynamics, i.e. invariant measures. These properties are encoded in few scalar values that form a small and highly descriptive feature set. This case study may motivate a careful selection of observables and transformations for revealing several aspects of mechanical vibration from measurement data.

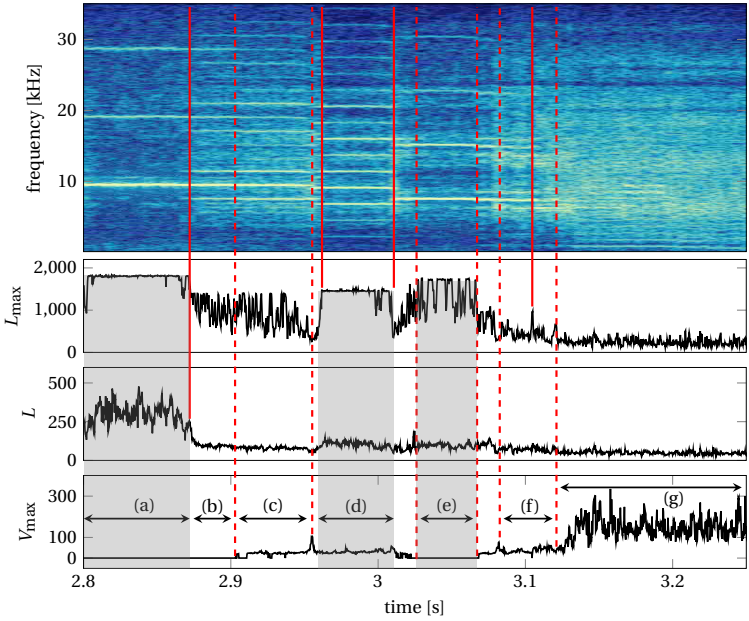


Figure (4.7) Transition from high-frequency FIV during steady sliding to stick-slip motion (detail view of Figure 4.6): (a) limit cycle, (b) chaotic transition, (c) weakly chaotic regime with dominant periodicities, (d) quasi-periodic dynamics, (e) limit cycle, (f) chaotic transition, (g) intermittent stick-slip. Regular dynamics are indicated by gray background for the RQA metrics for diagonal (L) and vertical (V) line lengths. Adapted from Figure 14 from [257]

Concluding, various signatures of regular and irregular motion are found here. The transitions and the dynamics are complex even if the frictional system may appear simple. The friction-induced vibrations are changing significantly faster than the imposed sliding velocity. Therefore, one can assume the dynamics to be stationary only for very limited time spans. One would rather denote the dynamics as non-stationary with respect to the small duration of the experiment and the multitude of different dynamical behaviors. This study indicates how nonlinear time series analysis and recurrence plots can help to characterize experimental vibration data. These techniques can be used in cooperation with classical Fourier analysis to study aspects of stationarity, regularity and transition behavior. Recurrence quantification analysis can be used as a physics-informed transformation to derive meaningful features from time series data. As the large measurement recordings are transformed into few descriptors, i.e. scalar values, simpler models could be selected for classification or identification tasks using machine learning.

4.3 Complex brake system vibrations

The following case, as well as the ones in Sections 4.5, 4.6 and 4.7, cover different aspects of brake system vibrations. The relevant data sources, formats and peculiarities are introduced here in order to provide a broader perspective. The various analysis steps will cover several facets of brake vibrations in more detail.

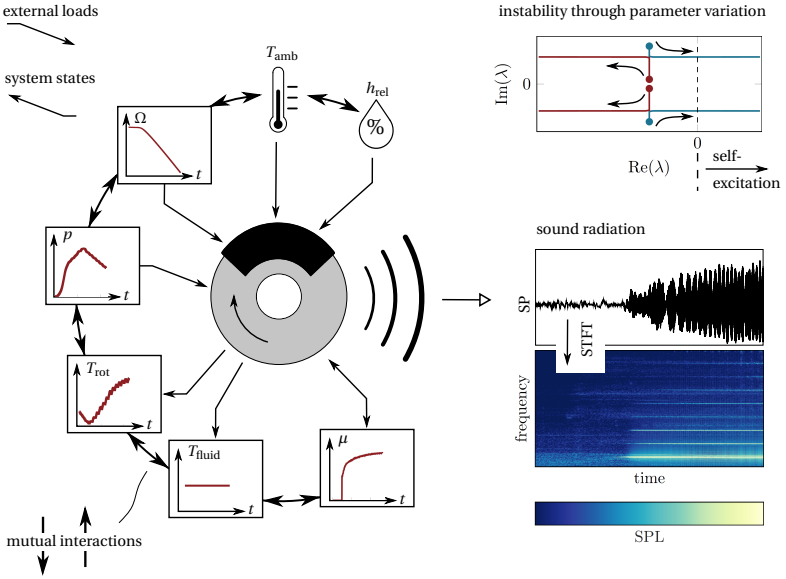


Figure (4.8) Data acquisition on disk brake systems: several loading conditions are monitored during the operation, and the emerging vibrations are recorded. The nonlinear structure is subjected to multiple transient external loads and multi-physics effects that are strongly interconnected. Once the system enters an instability regime trough parameter changes, high-intensity vibrations occur. A data set is displayed that stems from a single *stop braking* which exhibits high-frequency friction-induced vibrations. Adapted from Figure 1 from [259]

Commercial brake systems are not only safety-critical for the vehicle, but also experience increasing requirements for calm and smooth, that is vibration-free, operation. In the automotive industry, high-intensity friction-induced vibrations represent a major challenge owing to increasing customer demands. Several self-excited phenomena in the audible frequency range up to 16 kHz have been studied experimen-

tally and numerically for decades [53, 54, 56, 60, 66, 260]. The experimental measurement data used in this thesis stem from automotive disk brake testing on a NVH (noise, vibration, harshness) dynamometer. The test procedure follows the SAE-J2521 [261] protocol, and the data were provided by an industrial partner. The test matrix is composed of different types of brakings, i.e. stop and drag brakings. Several variations of the rotational velocity, brake line pressure, brake system temperature, ambient temperature and humidity are defined in the matrix test, see Figure 4.9. These parameters are referred to as *loadings* and *operational conditions* in the remainder of this thesis.

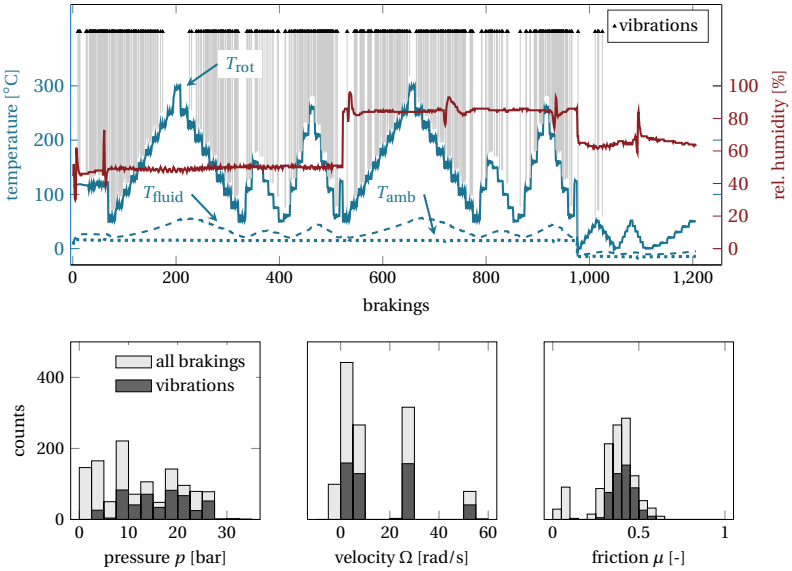


Figure (4.9) NVH matrix test featuring 1206 brakings according to the SAE-J2521 protocol [261]. Several temperature ramps can be observed in the upper panel with a high-temperature section in the beginning and a low-temperature section at the end of the test. Brakings that exhibit high-intensity vibrations above 1 kHz are marked. The lower panel depicts the histograms of the maximal pressure, velocity and friction value per braking. Essentially, vibrations can be observed for the complete range of those loading parameters. Adapted from Figure 2 from [259]

A microphone is placed in 50 cm distance to the brake disk to record the radiated sound pressure (SP). Structure-borne vibrations are recorded by an accelerometer placed on the brake caliper. While the vibrations are recorded at sampling rates of

f_s = 51.2 kHz, the loading conditions are monitored at lower rates of f_s = 100 Hz. Hence, there exist two *fast* vibration time series and several *slow* loading time series for each braking. Given a test matrix with more than 1000 brakings, the data acquisition system accumulates large amounts of data during testing. As an example, Figure 4.8 depicts a single braking and some of the recorded time series. Particularly, the following eight loading quantities are measured during testing: the disk rotation Ω, the brake line pressure p, the disk surface temperature T_{rot}, the brake fluid temperature T_{fluid}, the friction coefficient μ, the brake torque M, the ambient air temperature T_{amb} and the relative air humidity h_{rel}. Some quantities change rather fast, such as the friction coefficient, disk temperature and brake pressure, while other quantities are approximately constant, such as the air temperature. The brake fluid temperature can be thought of as capturing the long-term load history, and the disk surface temperature is a measure for the instantaneous friction energy input. Overall, the set of multivariate time series data is complex in terms of multi-physics loading conditions, temporally multi-scale effects and nonlinear relations between loads and system responses.

Section 4.4 presents two approaches for data-driven vibration detection in the fast system response data. Section 4.5 extracts linear damping properties of the dynamical system from the instantaneous amplitude growth at the onset of vibrations. Section 4.6 studies the vibration response to measure the stability margin by means of nonlinear invariant metrics. Finally, Section 4.7 illustrates the use of deep neural networks to predict the instantaneous vibrational response from the multivariate loading conditions. The data sets used in those studies do not stem from the same brake system. However, the same data acquisition setups were used to record the data in a consistent way following the SAE-J2521 protocol.

4.4 Vibration detection and characterization through deep learning

Detecting vibrations in measurements is an often-encountered task in the early stage of research projects in structural dynamics. Typically, several vibration phenomena need to be localized in time and characterized or classified by their salient spectral properties. However, time series measurements from mechanical machines are often contaminated with parasitic vibrations and noise from the environment. Furthermore, the elusive character of friction-induced vibrations in a non-stationary dynamic environment renders the detection of vibration patterns even more challenging. Therefore, trivial amplitude criteria may not suffice to detect and characterize vibration patterns in such scenarios. This study illustrates two approaches for highly-automated data-driven vibration detection. Parts of the following discussions and figures have been published in

[259] M. Stender, D. Schoepflin, D. Spieler, M. Tiedemann, N. Hoffmann and S. Oberst: *Deep learning for brake squeal: Brake noise detection, characterization and prediction*. Mechanical Systems and Signal Processing, 149, 2021

and Table 4.4 summarizes the conceptual framework for the vibration detection through deep learning.

Table (4.4) Overview on the application case 3: vibration detection and characterization using deep neural networks and computer vision techniques

data	• univariate time series vibration data $\in \mathbb{R}^{1\times n_t}$, n_t time steps
pre-processing	• manual labeling of vibrations for the training phase
transformation	• sequence to image $\mathbb{R}^1 \mapsto \mathbb{R}^2$: short-time Fourier transform
model	• deep convolutional neural networks for object detection
explain	• highly automated detection of vibration patterns • characterization and classification of vibration recordings • spectral and temporal vibration localization

The general idea is to transform the time series vibration measurement into a two-dimensional representation using time-frequency analysis. Here, computer vision techniques can be utilized to localize various vibration patterns. The detection results can be interpreted in the classical time and frequency dimensions. First, a conventional vibration detection method is proposed for periodic vibrations. Then, com-

puter vision techniques are revised and a more flexible deep learning approach is presented for a range of different vibration patterns.

Conventionally, time series measurements of mechanical vibrations are analyzed for their spectral properties by Fourier analysis. Hence, there is an underlying assumption that the relevant dynamics can be represented by a finite series of harmonics. Additionally, one inherently assumes stationarity over the time span used for the Fourier analysis. One class of major importance in brake system vibrations is a tonal high-intensity vibration in frequency regimes between 1 and 16 kHz. A tonal sound exhibits strong periodicities with the majority of vibrational energy being confined to one or few frequency bands. Figure 4.10 (a) depicts an exemplary microphone mea-

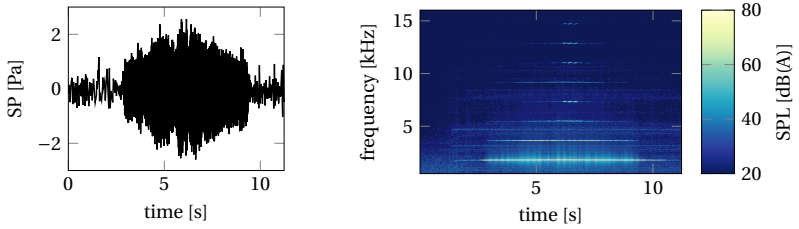


Figure (4.10) (a) microphone measurement of the sound pressure (SP) radiated during a single braking on a brake noise dynamometer. (b) corresponding spectrogram indicating the dB(A)-weighted sound pressure level of tonal and high-intensity friction-excited vibrations

surement of the sound pressure recorded during a braking. It becomes obvious that the sound pressure signal is highly non-stationary. Low-frequency modulations and high-frequency vibrations can be observed that span multiple time scales including noise artefacts from the measurement environment. The corresponding spectrogram from the Gabor transformation (4096 FFT lines, Hamming window) is displayed in Figure 4.10 (b). A high-intensity tonal sound can be observed in the time span $2.5 \leq t \leq 9.5$ s. At least six higher harmonics are excited besides the fundamental frequency of 1.8 kHz, which gives a clear evidence for a nonlinear vibration[2]. Only odd higher harmonics would be excited for a cubic nonlinearity. Hence, also other nonlinearities must be at play in this particular situation. Spectral hard-coded algorithms are conventional approaches [262, 263] to detect the onset of such kind of self-excited vibrations and characterize them by their dominant frequency and amplitude.

[2] Appendix A.5 illustrates how higher harmonics result from a cubic stiffness nonlinearity in a mechanical oscillator

In the following, a conventional detection algorithm is described for tonal periodicities. Successive Fourier transforms are computed in a sliding window approach. Peaks in the spectrum are possible candidates for tonal sounds, such that the elevation of the peak over the surrounding frequency band is used as a measure for the peak's sharpness. A heuristic sharpness threshold value discriminates *tonal* sounds from more broad-banded sounds. All tonal peaks are collected along time and the corresponding frequencies can be clustered to find those tonal sounds that exhibit a constant frequency for some time span. The final detection and vibration characterization involve a definition of a minimal peak amplitude and a minimal duration to actually form a tonal sound. Figure 4.11 depicts a schematic for this workflow. Here, a candidate peak requires a relative elevation of 50% of the peak over the surrounding spectrum in a 1 kHz frequency band. The peaks at $f = 10.3$ kHz are displayed along time. Given a minimal sound pressure level of 55 dB(A), the final tonal sound event is detected in the time range $2.0 \leq t \leq 10.2$ s with a maximal amplitude of 98 dB(A) at 10.3 kHz. Typically, brake systems exhibit self-excited dynamics at distinct clusters of frequencies, often referred to as *dominant instabilities* from a modal viewpoint.

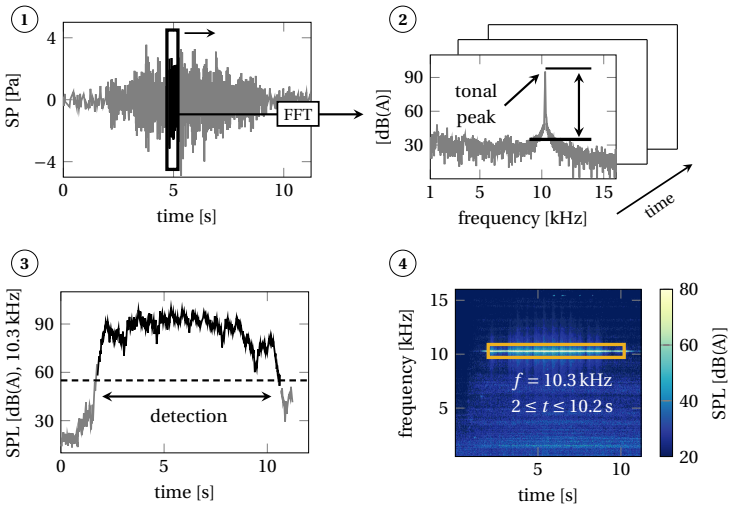


Figure (4.11) Conventional detection of dominant tonal periodicities using spectral methods: successive spectra are computed through a sliding window approach and candidate peaks are detected by their elevation above the surrounding spectrum. The amplitude evolution for those peaks is studied along time to find the regime where the periodicity exceeds a prescribed intensity level, here 55 dB(A). Finally, the detected tonal vibration is highlighted in the spectrogram. Adapted from Figure 4 from [259]

The detection of high-intensity tonal periodicities as described above can be considered robust and state-of-the-art today. However, owing to the plethora of dynamic processes involved in the friction interface and the complete system, brake systems can radiate various additional sounds. Particularly, those sounds can hardly be described by tonal characteristics. As they equally contribute to a disturbed NVH behavior for the customer, their detection is of high importance. Figure 4.12 shows examples for such non-tonal sounds that are frequently observed in disk brake systems. The time series are corrupted by multi-scale fluctuations and noise. Furthermore, the

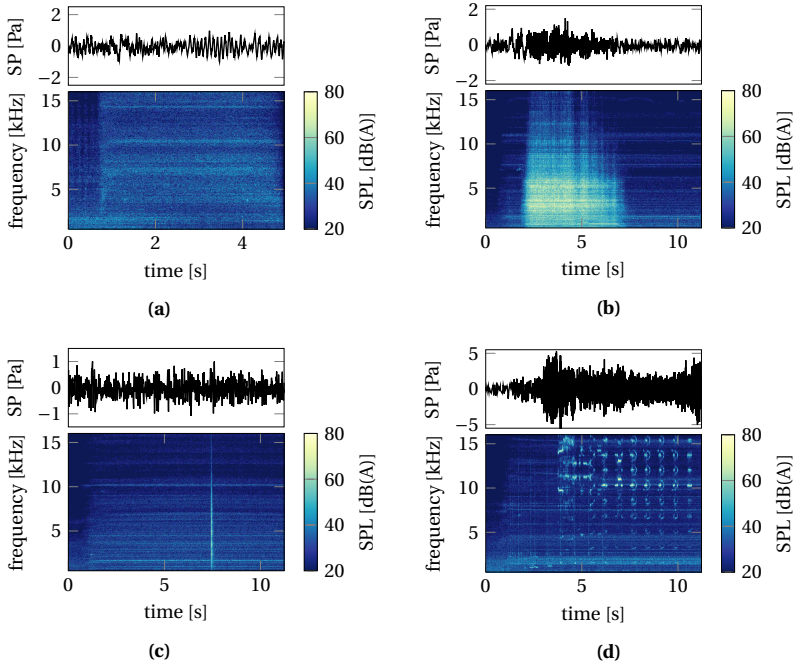


Figure (4.12) Microphone measurements and spectrograms of commonly encountered friction-excited disk brake vibrations: (a) *quiet* braking without a vibration event, (b) *noise* artefact from the environment, (c) short impulse-like *click* and (d) so-called *wirebrush* sound. Adapted from Figure 3 from [259]

spectral properties are not as pronounced as in the case of tonal sounds. Hence, the described detection approach is not suited for those vibration patterns. Yet, experienced NVH engineers can easily differentiate those sound by the visual inspection of the spectrograms and the characteristic patterns. This situation motivates the use

of computer vision techniques for designing a flexible detection algorithm for multiple vibration patterns. No specific signal **pre-processing** is required for the vibration measurements. The central idea of this approach is to **transform** the time domain data into the time-frequency domain to unravel the characteristic features of different sounds. This transformation is obtained through the short-time Fourier transform that derives a three-dimensional array of time, frequency and amplitude information from the measurement data. Deep neural networks are employed to identify patterns in the resulting arrays that are treated as images. In the proposed OPTME data science process for structural dynamics, this approach relates to finding a better observable through an appropriate transformation for the given data, and selecting a proper model for the transformed data.

Computer vision (CV) is a vast research field, such that central ideas and methodologies are revisited shortly. In CV there are the tasks of image classification and object detection among others. Image classification assigns one or multiple labels to a given image, while object detection returns the coordinates of an object and its label. As multiple vibration phenomena may occur during a single braking, and the frequency-time characteristics are of high importance, object detection techniques are used in this work. CV has developed several approaches to object detection before the rise of neural networks. Those approaches rely on increasingly complex handcrafted features, such as the milestones of cascades of low-level features [264] in 2001, oriented gradients [265] in 2005, deformable templates [266] in 2010 and multi-resolution image features [267] in 2014. However, the year 2012 defines the probably most important landmark in CV. In the ImageNet [268] challenge, Krizhevsky et al. [269] propose a deep convolutional network *AlexNet* that outperforms classical CV techniques for the first time. Since then, neural networks have demonstrated their superior classification, detection and segmentation quality in the successive ImageNet, PASCAL [270], COCO [271] and various other challenges. Tremendous improvements in prediction quality and computation time have been observed until today. Conceptually, deep learning CV techniques for object detection are required to solve two tasks in parallel: region proposal for the object location and classification of the object into predefined classes. Two methodological branches have grown in recent years, namely region proposal based methods and regression based methods [272]. Region proposal based methods regress the object locations first, and then classify the objects in each proposal region in the second step. Exemplary network architectures for this approach are R-CNN [273], Fast R-CNN [274], Faster R-CNN [275] and R-FCN [276]. This iterative two-step approach suffers from significantly higher computational efforts compared to the second branch of regression based methods. Here, the location regression and object classification tasks are addressed at once. Recent networks of this class are G-CNN [277], YOLO [278] and SSD [279]. All of the aforementioned neural networks are composed of convolutional layers as core building blocks.

model: To assess the chances and limitations of object detection networks for the task of vibration detection, region proposal networks are studied in this application. Region based convolutional networks (R-CNN) generate a large number of region proposals through selective search algorithms in the first step. In the second step, each warped image segment from each region proposal is fed to a CNN for computing a feature map from which a classifier can learn a class label. In parallel, a regressor learns how to adapt the region, i.e. its location and size, for obtaining a better prediction score. Variants of Fast R-CNN networks extract a single feature map from the complete input image using CNN layers, and then feed the subset of features per region proposal to the classification-regression unit. For even better performance, Faster R-CNNs replace the selective search algorithm by region proposal networks (RPN) that extract regions of interest directly from the CNN feature map that is computed in Fast R-CNN architectures. Region based fully convolutional networks (R-FCN) derive location-sensitive features maps using RPNs to account for the observation that characteristic features often appear in specific locations in the image.

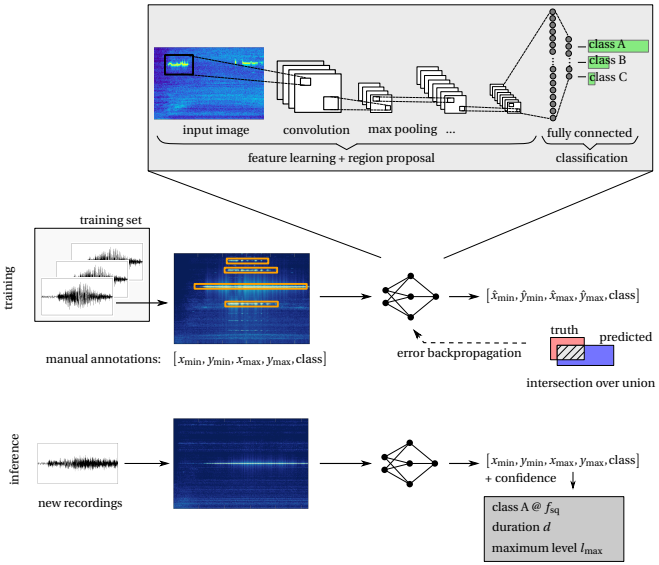


Figure (4.13) STFT transformed signals are treated as images for the object detection network. Labeled instances are used for training, and unseen vibrations measurements can be analyzed with validated model. Convolutional layers are schematically illustrated in the sketch of a simple network architecture in the upper part. Adapted from Figure 6 from [259]

The complete work-flow for deep learning enabled vibration detection is depicted in Figure 4.13. Labeled spectrograms are required for training and model validation. Particularly, a ground truth training instance comprises the spectrogram, the object location in terms of bounding box coordinates and the corresponding class label. The object detection network will make a prediction in terms of the location and object label. The error between prediction and ground truth is derived from the intersection over union (IoU) metric, which calculates the coverage ratio between the bounding boxes for a given class label. This error metric is used to adapt the network parameters during training and also to assess the generalization properties on the validation data set. After training, the object detection network has practically become a vibration detection network which can locate several vibration patterns in spectrograms derived from new and unlabeled vibration measurements. Additionally, the network outputs a confidence score for the prediction, which results from the softmax activation function (see Appendix B.8) in the last fully connected (FC) layer.

In this study, four classes of vibration events are specified, see Figure 4.12. The vibration detection network has to detect and locate these phenomena in the vibration recordings. *Tonal* high-intensity vibrations exhibit strong periodicities in the frequency range of $1 \leq f \leq 16\,\mathrm{kHz}$ with sound pressure levels above $50\,\mathrm{dB(A)}$ and a minimal duration $0.5\,\mathrm{s}$. *Click* sounds arise from sudden vibro-impacts and rearrangements of the brake pad in the caliper. *Wirebrush* sounds are characterized by multiple short pulses across the complete frequency band similar to chirps. *Artefacts* are given by high-intensity and broad-banded noise from the testing environment, the main dynamometer drive train or auxiliary engines. 3267 brake vibration measurements are available for building the detection model and for evaluating it. Manual labeling is performed for the bounding box locations and the class annotations. A set of 290 representative sounds is used for the model evaluation, such that 2977 samples remain for training and validation of the network. A stratified $80-20$ split is used, such that the training set $\mathscr{D}_{\text{train}}$ consists of 2382 images. Compared to the testing reality, the minority classes (click, wirebrush, artefacts) are over-represented to facilitate the network training.

Pre-trained networks are used as a starting point for the design of the deep learning vibration detector. The pre-trained object detection networks have already learned to abstract lines, edges and other features from images, such that training on a new set of images will converge faster than for a randomly initialized network structure. This concept is referred to as *transfer learning* [201]. Here, instances of the region proposal-based Faster-RCNN and R-FCN models are employed which have been trained on the COCO [271] data set for detecting common objects in contexts. For the Faster R-CNN two different model architectures from the *inception* and *resnet* configurations [201] are compared, such that the three models Faster R-CNN (inter-

ception), Faster R-CNN (resnet) and R-FCN are compared. Each model is trained for 100.000 steps with unit batch size. Data augmentation (flipping, randomly cropping, randomly padding and randomly adding black patches) is applied to the input images to increase the data set size and its variance. A fixed learning rate of 0.0003 is used for the momentum optimizer.

explain: As for quality metrics, the correct label and the correct bounding box location need to be evaluated in the object detection task. The performance of the vibration detection network is evaluated in two steps: first, the classification of a complete measurement is discussed, and then also the location of the objects is taken into account. The training data $\mathscr{D}_{\text{train}}$ is set up to contain only single-class images, such that a single label per image can be evaluated for pure classification. This set-up allows to compare the conventional spectral detector against the deep learning models. Besides the conventional confusion matrix entries per class, also the precision-recall curve (PRC) is assessed. Here, cumulative precision and recall metrics are computed for all predictions sorted by the confidence score. The area under curve (AUC) measures the average precision $\text{AP} = \int \text{PRC}$ for a single class. For m classes, the mean average precision $\text{mAP} = {}^{1}/_{\text{m}} \sum \text{AP}_i,\ i = 1, \ldots, m$ summarizes the overall performance of the multi-object classifier. To obtain confidence scores for the spectral detector, a synthetic measure based on maximum intensity and duration of the sound is introduced as discussed in [259]. Table 4.5 summarizes the performance of three network architectures on the evaluation set $\mathscr{D}_{\text{val}}$ which contains 200 tonal sounds, 50 wirebrush sounds, 25 click sounds, 15 noise artefacts and 200 quiet brakings. The data set is specifically set-up to replicate typical class distributions encountered in testing, although the quiet sounds are massively under-represented compared to reality. Based on several studies for optimal performance, a minimum confidence value for reporting an object is set to $C_{\text{min}} = [0.9, 0.84, 0.88]$ for model 1, model 2 and model 3, respectively.

Table 4.5 reports the classification task results. The spectral detector, serving as a baseline model, has a high detection rate for the tonal sounds, i.e. those vibration patterns for which it has been designed. Conceptually, it is not able to detect the remaining classes. Most of the other vibrations result in misclassifications, such that more false positives than true positives can be observed in the baseline model. For the class of tonal sounds, all deep learning detectors achieve slightly higher AP metrics. Furthermore, as they are able to detect also other vibration patterns, the false positive detections reduce significantly. All models exhibit relatively high false negative detections for the wirebrush class and are too sensitive to tonal sounds. The R-FCN performs significantly worse, which may stem from the inherent structure of the network: if specific objects do not appear in similar locations throughout all images, performance will drop. Overall, the R-CNN in resnet configuration achieves the

Table (4.5) Classification performance of the spectral detector and three deep learning models evaluated on the reference data set containing 200 quiet brakings, 200 tonal sounds, 50 wirebrush sounds, 25 click sounds and 15 broad-banded noise artefacts. The quality metrics F_1, average precision AP and mean average precision mAP (computed from the PRC) are evaluated

classifier	category	TP	FP	FN	TN	F_1	AP	mAP
ML model 1 (faster R-CNN incep)	tonal	149	68	7	266	0.8	0.68	0.73
	wirebrush	33	1	17	439	0.79	0.69	
	click	23	8	2	457	0.82	0.88	
	artefact	10	0	5	475	0.80	0.68	
ML model 2 (faster R-CNN resnet)	tonal	149	76	7	258	0.78	0.72	0.80
	wirebrush	38	3	12	437	0.84	0.78	
	click	22	6	3	459	0.83	0.87	
	artefact	12	0	3	475	0.89	0.81	
ML model 3 (R-FCN resnet)	tonal	146	76	10	258	0.77	0.68	0.55
	wirebrush	26	1	24	439	0.68	0.57	
	click	21	0	4	465	0.91	0.85	
	artefact	1	0	14	475	0.12	0.1	
spectral	tonal	156	177	0	157	0.64	0.63	–
	wirebrush	–	–	–	440	–	–	
	click	–	–	–	465	–	–	
	artefact	–	–	–	475	–	–	

highest performance scores and a mean average precision of 0.8 which represents a substantial improvement over the spectral detector. To evaluate the location of vibration patterns in terms of time and frequency, the object detection quality is evaluated based on the IoU measure, see Table 4.6.

Depending on a minimal IoU value, the detection results vary. Trivially, the prediction quality decreases when high IoU values are required for a true positive. Analogously to the previous result, the R-CNN models outperform the R-FCN model. Again, wirebrush sounds pose a challenge to the networks, while rather high detection rates are observed for the remaining classes. Overall, the shape of the vibration patterns in the spectrograms vary strongly in size and shape, such that inter-class comparisons must be undertaken carefully. To visualize the deep learning vibration detector output, Figure 4.14 depicts two vibration recordings from the evaluation data set with the detected objects and confidence scores.

This application case does not strive for a highly-optimized model architecture, but is set up to illustrate the benefits of transforming vibrations into the domain of computer vision to make use of deep learning. As such, the studies reveal the potential

Table (4.6) Object detection performance measured by average precision per class (AP) at IoU levels 50%, 75% and 90% and resulting mean average precision (mAP) for each deep learning (DL) detector

classifier	category	$AP_{.50}$	$AP_{.75}$	$AP_{.90}$	$mAP_{.50}$	$mAP_{.75}$	$mAP_{.90}$
DL model 1 (F. R-CNN incep)	tonal	0.66	0.66	0.60	0.66	0.65	0.59
	wirebr.	0.57	0.54	0.36			
	click	0.74	0.74	0.72			
	artefact	0.67	0.67	0.67			
DL model 2 (F. R-CNN resnet)	tonal	0.68	0.68	0.62	0.69	0.67	0.55
	wirebr.	0.61	0.51	0.19			
	click	0.80	0.80	0.71			
	artefact	0.69	0.67	0.55			
DL model 3 (R-FCN resnet)	tonal	0.66	0.65	0.60	0.43	0.39	0.34
	wirebr.	0.44	0.30	0.16			
	click	0.55	0.55	0.55			
	artefact	0.07	0.07	0.07			

of data-driven approaches to overcome the limitations of hard-coded spectral algorithms. Further optimized models, more training data and cross-validating the results with the spectral detector may leverage further performance gains. This study may support the hypothesis given in the previous Chapter 3 which states that the data transformation is the most crucial ingredient towards a successful adaptation of recent deep learning methods in mechanical vibrations.

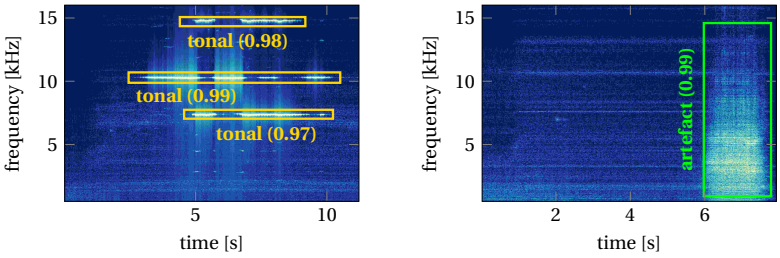


Figure (4.14) Output of the deep learning vibration detector (model 1) for two samples of the evaluation data set. Tonal sounds (a) and a broad-banded noise artefact (b) are correctly detected at high confidence scores. Adapted from Figure 8 from [259]

4.5 Instantaneous damping identification from time series data

This study shows how a structure's instantaneous damping properties can be extracted from time series data. The instantaneous growth rate, i.e. the equivalent to an numerical eigenvalue's real part, is extracted from brake system vibrations at the onset of self-excited vibrations. Thus, better insights into quasi-linear system properties are obtained from existing data by appropriate filtering and modeling without requiring additional experimental resources. An overview following the OPTME data science process for structural dynamics is given in Table 4.7. The results and figures presented hereafter have already been published in

[280] M. Stender, M. Tiedemann, L. Hoffmann and N. Hoffmann: *Determining growth rates of instabilities from time-series vibration data: Methods and applications for brake squeal.* Mechanical Systems and Signal Processing, 129: 250-264, 2019

[281] L. Hoffmann: *Calculation of growth rates for self-excited dynamical systems on the example of brake squeal noise.* Projektarbeit, Technische Universität Hamburg, 2018 .

Table (4.7) Overview on the application case 4: extracting growth rates from time series data

data	• univariate time series data $\in \mathbb{R}^1$
pre-processing	• vibration detection • heuristic detection of exponentially growing amplitudes
transformation	• spectral filtering (short-time Fourier transform) • scale selection (inverse continuous wavelet transform) • envelope fitting through the Hilbert transform
model	• linear regression for exponential fit
explain	• instability growth rates measures for negative damping

Damping parametrization of numerical models is one of the most crucial and yet one of the most uncertain ingredient for predictive modeling as discussed in Section 2.3. The complex eigenvalue analysis derives the eigenvalues $\lambda = \delta + \mathrm{i}\omega$ from a linearized model configuration. Model updating approaches [282] mostly focus on component properties and mode shapes, i.e. on steady-state behavior that can be described given the assumption of stationarity. For self-excited systems, the forced response may significantly differ from the self-excited response. Hence, the self-excited vibration response only allows to study the operational deflection shapes, but not the associated

structural damping. As a result, modal updating strategies for those systems mostly rely on component properties [75]. The central idea of the novel approach is to extract the regime of exponentially growing amplitudes from vibration measurements at the onset of self-excitation. Here, the growth rate $\hat{\delta}$ is a direct measure of the instantaneous damping in the structure and corresponds to the instability's real part δ.

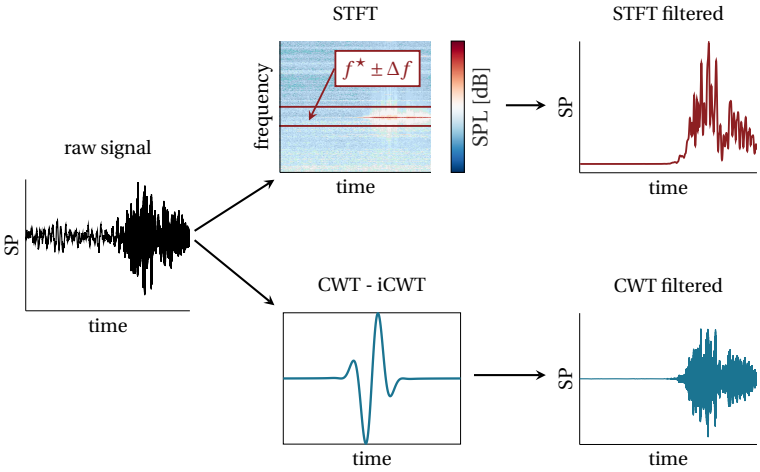


Figure (4.15) Vibration data pre-processing for isolating the unstable frequency $f^\star$ in the noisy sound pressure (SP) measurement: spectral bandwidth filtering using the STFT and the inverse CWT. Exponentially growing amplitudes can be observed at the onset of the high-intensity vibrations. Adapted from Figure 3 from [280]

pre-processing and transformation: Extensive signal processing is required to isolate the unstable frequency and time range of interest in an experimental vibration measurement. Automotive disk brake system vibrations are studied which exhibit strong multi-scale and non-stationary effects, see Figure 4.15. The objective is to study the exponential growth of the unstable mode $f^\star$ through filtering of the vibration measurement $\boldsymbol{s}(t)$. Two filtering approaches are utilized: spectral filtering using the short-time Fourier transform (STFT) and scale filtering using the continuous wavelet transform (CWT) and its inverse. The amplitude time series for a spectral bandwidth $f^\star - \Delta f \leq f \leq f^\star + \Delta f$ centered around $f^\star$ is obtained from the root-mean-squared STFT amplitudes at each time step t_i. In the second approach, filtered time domain signals are obtained from the inverse continuous wavelet transform using Morse mother wavelets. During the derivation of the inverse transform,

only the relevant frequency ranges, i.e. scales, are considered. Resulting amplitude series (STFT) and vibration series (CWT) $\hat{\mathbf{s}}(t)$ are depicted in Figure 4.15 for the exemplary brake vibration signal. Envelopes Γ are created for both signals using the Hilbert transform[3]. A heuristic bandwidth value of $\Delta f = 150$ Hz is utilized in this study.

In a second pre-processing step, the time epochs of exponentially growing amplitudes must be identified. While the spectral vibration detection approach from Section 4.4 was used to identify high-frequency vibrations in the measurements, the time epochs of exponential growth require further analysis which is displayed in Figure 4.16. Particularly, the time span of exponential growth is rather short before nonlinear effects set in owing to large vibration amplitudes, see the prototypical Van der Pol study in Figure A.1. For the present data, heuristic thresholds are defined for identifying the transient zone. As a reference point, the first inflection point of the filtered signal's envelope is chosen in the transient regime. For most observations, this point corresponds to approximately the first half of the transient zone before the exponential growth saturates. Relative to this reference point, the start t_{start} and end points t_{end} of the transient zone are defined. The exponential growth rate $\hat{\delta}$ is estimated from the data in this zone.

model: The envelope of the filtered signal $\left[\hat{s}(t_{\text{start}}), \hat{s}(t_{\text{end}})\right]$ is subjected to a fitting procedure for an exponential function $\hat{s}(t) \approx \hat{s}_0 \mathrm{e}^{\hat{\delta} t}$ that minimizes the root-mean-squared difference between the measured signal and the fit. Particularly, the fit is obtained from a linear regression in a semi-logarithmic relation of the amplitudes to the time axis. As the growth rate estimate crucially depends on the selected transient phase, a variation of the start and the end points is introduced. These variations are again formulated in a relative manner with respect to the reference point given by the envelope's inflection point. Each start and end point is varied independently, such that for n variations n^2 growth rate estimates result from individual regression models. Then, the most frequent value out of the best 20% fits, i.e. having minimal fitting error, of all $\hat{\delta}_i, i = 1, \ldots, n^2$ values is selected as final growth rate estimate $\hat{\delta}$. To be consistent with theory, only those vibration recordings are considered which contain a single and dominant periodicity, i.e. a single mode that becomes unstable. More complex vibration patterns with frequency modulations are omitted. Overall, for a single transient phase there are three growth rate estimates: one estimate from the STFT amplitude series plus the lower and upper envelope of the CWT-filtered vibration signal. For the sake of clarity, only the mean of CWT-based upper and lower estimates are displayed as $\hat{\delta}$ in the following discussions, since they turned out to be more robust and consistent than the STFT-based estimates [280] for the present data.

[3] Practically, the one-sided FFT is computed by removing all coefficients with negative frequency before the inverse transform. The *analytic signal* provides the amplitude envelope function

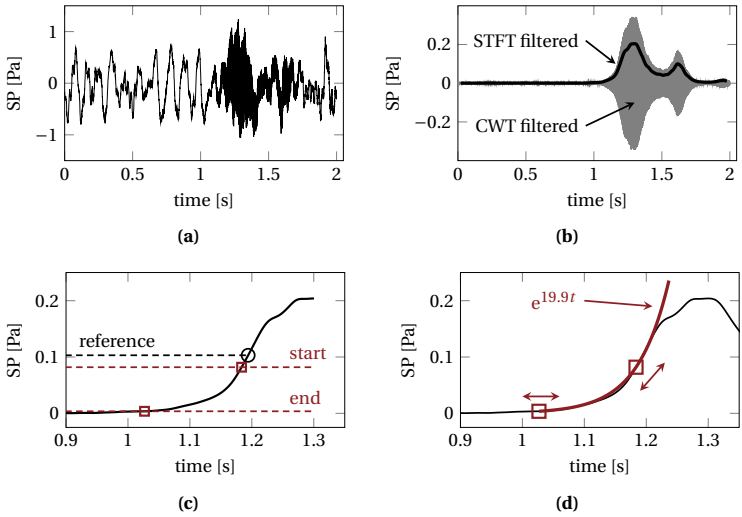


Figure (4.16) Determination of the growth rate estimate $\hat{\delta}$ from a time series recording: (a) raw sound pressure signal and (b) pre-filtering using the STFT and the CWT for a frequency bandwidth around the vibration event. (c) defines the reference point as the signal's first envelope inflection point and the transient phase between t_{start} and t_{end}. (d) the best exponential fit is $\hat{\delta} = 19.9\,\text{Hz}$. Adapted from Figure 4 from [280]

explain: Two data sets $\mathscr{D}_A$ and $\mathscr{D}_B$ from commercial brake testing are studied. Data set $\mathscr{D}_A$ contains an extended test cycle of 4754 brakings, and data set $\mathscr{D}_B$ contains 1891 brakings. Only severe tonal vibrations with a minimal average sound pressure level of 60 dB(A), a minimal duration of 1 s and a dominant vibration frequency between 1 and 16 kHz are considered in this study. Data A and B contain 323 and 197 of such events, respectively. Figure 4.17 displays the NVH behavior and sensitivities with respect to the disk rotation Ω and the brake line pressure p. System A exhibits vibrations at 5.8 and 13.6 kHz with a wide range of sound intensity levels. System B shows instabilities at 3.0, 8, 11 and 14.2 kHz with maximum sound pressure levels of 106 dB(A). Both systems show dependency on the loading conditions, such that system B is sensitive to low brake pressure values and low disk velocities. System A exhibits instabilities across the full parameter range which is also observed for other parameters such as the friction coefficient (not displayed here).

The objective of the following studies is to find out whether a specific system instability can be characterized by a constant growth rate, and how this measure of instantaneous damping depends on the loading conditions. Growth rates are derived from

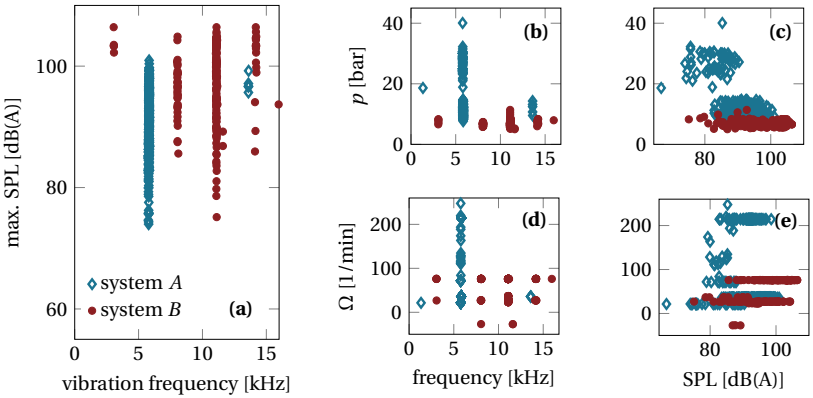


Figure (4.17) Tonal sound characteristics of the vibrations observed during NVH testing: (a) vibration frequencies and corresponding sound pressure levels, and (b) to (e) dependence of the acoustic characteristics on the maximal brake line pressure p and the disk rotation Ω during braking. Adapted from Figure 6 from [280]

the microphone recordings as described before. To increase credibility, samples are removed for which the CWT estimates from the upper and lower envelope series differ by more than 5 Hz and for which less than 120 variations of the transient range were obtained. In total, 242 growth rates $\hat{\delta}$ are determined for system A and 86 for system B. The growth rates vary in the range of 5 to 60 Hz and are displayed in Figure 4.18.

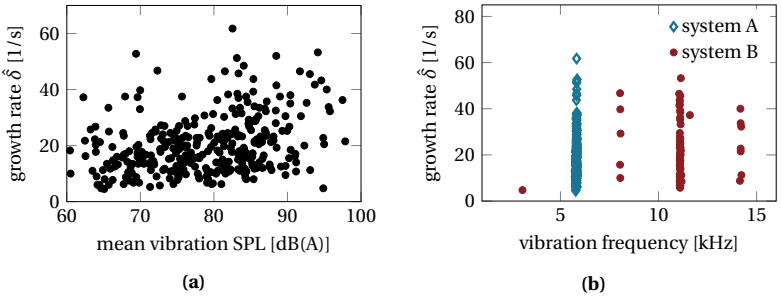


Figure (4.18) (a) 328 experimentally estimated growth rates $\hat{\delta}$ of systems A and B and the sound pressure level (SPL) of the fully developed nonlinear vibration. (b) Relation of the growth rates to the unstable vibration frequencies. Adapted from Figure 8 from [280]

Firstly, there is no theoretical relation between the growth rate being a linear system property and the final vibration amplitude which is governed by nonlinear effects. The experimental findings confirm this understanding: the estimated growth rates are not correlated with the vibration amplitude as depicted in Figure 4.18 (a). This is the first observation of that kind in friction-excited experimental systems. Secondly, the results indicate that the instantaneous damping of the unstable mode varies massively as a function of the loading conditions. Thirdly, the growth rates of the same instability, i.e. same dominant vibration frequency, vary strongly. Hence, vibrations with similar periodicity may still stem from the same mode becoming unstable, but the system's damping configuration at the point of instability changes significantly during the operation. This observation contradicts classical modeling approaches in a modal framework, where a fixed modal damping value is assigned to an individual mode. In such a scenario, the experimental growth rates would form dense clusters for a specific vibration frequency. For these data, such behavior cannot be observed. Therefore, the damping variations must stem from parameter variations through the brake operation and secondary, e.g. temperature or wear, effects.

As the growth rates are found to be independent from the resulting vibration characteristics, it is interesting to study if vibrations self-excited at similar loading conditions grow at the same rate. If that is the case, the recorded loads constitute major dimensions of the bifurcation parameter space of the dynamical system. In the following, a single instability is selected: 242 instabilities of system *A* with a vibration frequency of 11 kHz. Although these self-excited vibrations carry the same spectral characteristics, they grow at strongly varying rates. To measure similarity between loading scenarios, the similarity index for vectors $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$

$$\rho = \mathrm{sim}\left(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2\right) = \frac{1}{1 + \sqrt{\sum_{i=1}^{n} \left(\xi_{1,i} - \xi_{2,i}\right)^2}} \tag{4.17}$$

is introduced based on the Euclidean distance. $\rho = 1$ represents equality, and $\rho \to 0$ indicates orthogonality. Particularly, the individual loading parameters ξ are given by the rotational speed of the disk, brake line pressure, deceleration, friction coefficient, brake disk temperature, brake fluid temperature, ambient air temperature and relative air humidity plus the time derivative of pressure, deceleration, friction value and disk temperature, each measured at the onset of FIV. Thus, the operational state at vibration onset is described by the twelve-dimensional vector $\boldsymbol{\xi} = [\xi_1, \ldots, \xi_{12}]^\top$. Certainly, this is only a subspace of all parameters that are relevant for describing the instantaneous state and resulting damping level of the structure. The instantaneous load vector $\boldsymbol{\xi}_i$ for sample i is compared in pairwise fashion against all $j = 1, \ldots, 242, j \neq i$ samples to find the similarities ρ_{ij}. Analogously, the differences in growth rates $\Delta\hat{\delta}_{ij} = |\hat{\delta}_i - \hat{\delta}_j|$ and the differences in average sound intensity $\Delta l_{ij} =$

$|l_i - l_j|$ are computed. In a classical scenario for nonlinear structural dynamics, there exists a bifurcation point at which the equilibrium position loses stability to a limit cycle. Potentially, multiple attractors may co-exist for different parameter ranges, such that the system may jump to either of those solutions. In this study, instabilities of the same frequency were chosen, such that one may assume the structure to have ended up on similar attractors. Here, the loading vector $\boldsymbol{\xi}$ is considered to constitute major dimensions of the bifurcation parameter space. Hence, if the difference in loading conditions is small, e.g. $\rho > 0.8$, the difference in vibration level and the difference in the growth rate values would be expected to be small as well. If that is the case, the load vector will capture major system parameters, and the system will repeatably converge to a periodic solution characterized by a constant frequency and amplitude of vibration. Such rather simplistic system understanding is depicted in Figure 4.19 (a) with relations to Δl, $\Delta\hat{\delta}$ and ρ.

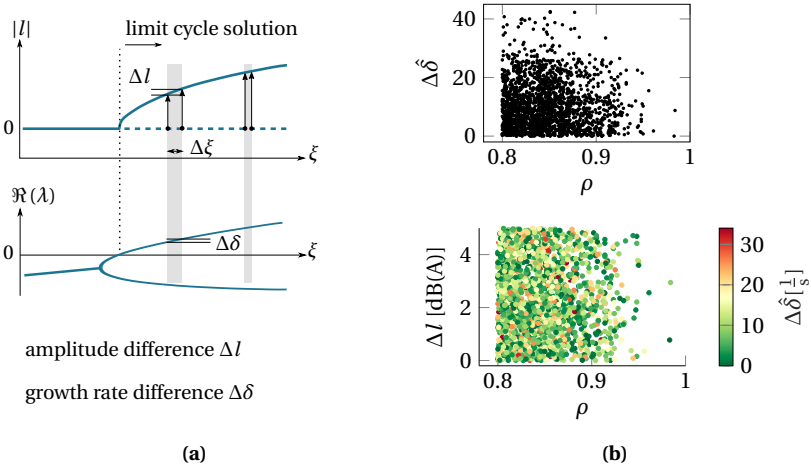


Figure (4.19) (a) illustration of a prototypical bifurcation diagram where a system becomes unstable through a flutter-type instability and a limit cycle solution exists for larger parameter values ξ. (b) experimental results for the 11 kHz instability of system *A*. ρ indicates the similarity of the loading condition ξ, $\hat{\delta}$ is the growth rate estimate and l the vibration intensity. Adapted from Figures 9 and 11 from [280]

However, the data analysis shows that for the present system no such correlation can be found. The growth rates differ by up to 40 Hz even for high loading similarity values, see Figure 4.19 (b, upper panel). Vibrations excited in similar system conditions grow at variable rates. Furthermore, even if both the loads and the growth rates are

similar, the final sound intensity, which can be thought of as a proxy for the size of the attractor, varies strongly, see Figure 4.19 (b, lower panel). Thus, for the present data it must be concluded that the dynamics involved are highly complex in terms of parameter dependencies and nonlinear steady-state behavior. Even if the vibrations appear to be rather simple and similar to a limit cycle oscillation, the linear system properties at instability and the corresponding system configurations do not fall into distinct clusters. Hence, the vibration behavior cannot be replicated for seemingly equal system configurations, which underlines the well-known elusive character [283] of FIV. There is a wide range of possible reasons for this observation, such as continuously changing conditions, wear and tribological aspects, hysteretic effects in joints, overall non-stationary behavior, such that attractors of fixed size cannot exist in a strict sense, and weakly chaotic dynamics that introduce sensitive dependence on initial conditions. As the growth rates were found to vary strongly for various loading conditions, CEA studies should consider a wide range of system configurations for a better correspondence with experimental results.

Concluding, the methodology for extracting growth rates from vibration data represents a novel approach to a better system understanding. For the first time, numerical CEA results can be compared with experimental data in a scientifically sound fashion without requiring additional testing resources. From the data science perspective, the data pre-processing requires deep domain knowledge about nonlinear vibrations and self-excitation. After the rather extensive signal filtering and data segmentation, the modeling process collapses to a simple regression in form of an exponential fit. The results of the regression can be directly interpreted as damping measures and may be studied to find parametric sensitivities or may be used for enhancing model updating procedures.

4.6 Stability margin extraction from time series data

In the third analysis of the brake vibration data, the strongly nonlinear regime of self-excited vibrations is approached. In this regime, the vibration amplitudes are limited by nonlinearities after the instantaneous growth phase that was investigated in the previous section. Similarly to the study of the block-on-block system in Section 4.2, nonlinear time series and recurrence plot analysis techniques (RPQA) are utilized to characterize and quantify the observed dynamics. Parts of the work and figures presented hereafter have been published in

[162] M. Stender, S. Oberst, M. Tiedemann and N. Hoffmann: *Complex machine dynamics: systematic recurrence quantification analysis of disk brake vibration data*. Nonlinear Dynamics, 267, 105, 2019

and the conceptual summary is given in Table 4.8. Vibration data from commercial NVH testing, see Section 4.3, are studied. In contrast to previous studies of FIV by means of nonlinear time series analysis [37, 68], this study considers large amounts of data in a highly systematic manner. Vibration measurements from seven different brake systems are subjected to the analysis in order to obtain a more general picture. In total, 156 minutes of vibrations sampled at $f_s = 51.2\,\text{kHz}$ are investigated here.

Table (4.8) Overview on the application case 5: nonlinear time series analysis of brake system vibrations

data	• univariate time series measurements $\in \mathbb{R}^1$
pre-processing	• vibration detection • wavelet bandpass filtering • sliding windowing segmentation
transformation	• time-delay embedding for state space reconstruction • recurrence plot quantification analysis
model	• exploratory data analysis
explain	• study regularity and stationarity in brake system vibrations • analyze dimensionality and complexity of the dynamics • measure the stability margin by nonlinear invariants

pre-processing: The vibration measurements are bandpass filtered by the (inverse) wavelet transformation for a frequency range of $1 \leq f \leq 20\,\text{kHz}$ to remove low-frequency modulations and noise from the environment. Nevertheless, the signals are highly transient and do not exhibit stationary amplitudes along the vibration du-

ration even after filtering. Generally, two signal segmentations are performed for the nonlinear time series analysis. First, the vibration signals are divided into *quiet* and *vibrating* segments based on the tonal sound detection presented in Section 4.4. Particularly, only those recordings with a single tonal vibration phenomenon are considered in this study with a vibration duration longer than 0.5 s. Time series are omitted if they contain different vibration phenomena such as noise artefacts, wirebrush or others. Second, the time series measurements are subjected to a sliding window seg-

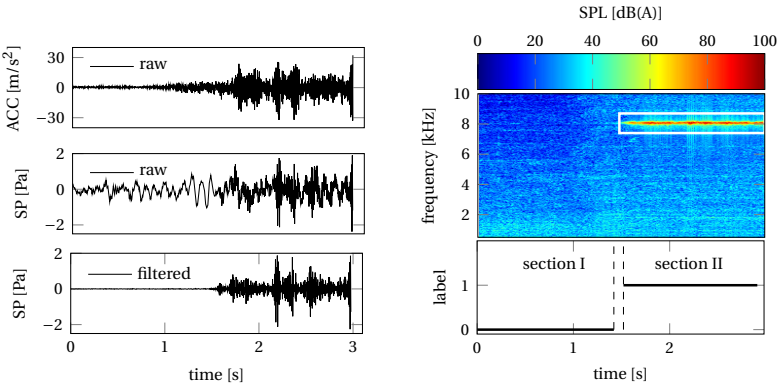


Figure (4.20) Left panel: raw vibration measurements of the structure-borne vibrations from an accelerometer (ACC), the air-borne vibrations and the wavelet bandpass filtered signal of the sound pressure (SP). Right panel: time-frequency analysis by the short-time Fourier transform and the segmentation into *quiet* and *vibrating* sections I and II, respectively. Adapted from Figures 1 and 2 from [162]

mentation within each section. This approach is motivated by the requirement of stationary dynamics to be compliant with Takens' theorem [27] when reconstructing the state space. Depending on the window length w, dynamics on different scales are captured. The window length is chosen such that the embedding dimension estimated by the FNN algorithm does not change if longer windows were chosen. Such behavior is obtained for $w = 1500$ samples[4], i.e. 0.0293 s. The overlap for the sliding window processing for embedding is set to 75% to enrich the data base and provide some smoothing behavior along time. To fulfill stationary assumptions for the RPQA, the sliding window size is set to 2000 for the recurrence analysis [162].

[4]The detailed study is given in the supplementary electronic information accompanying [162]

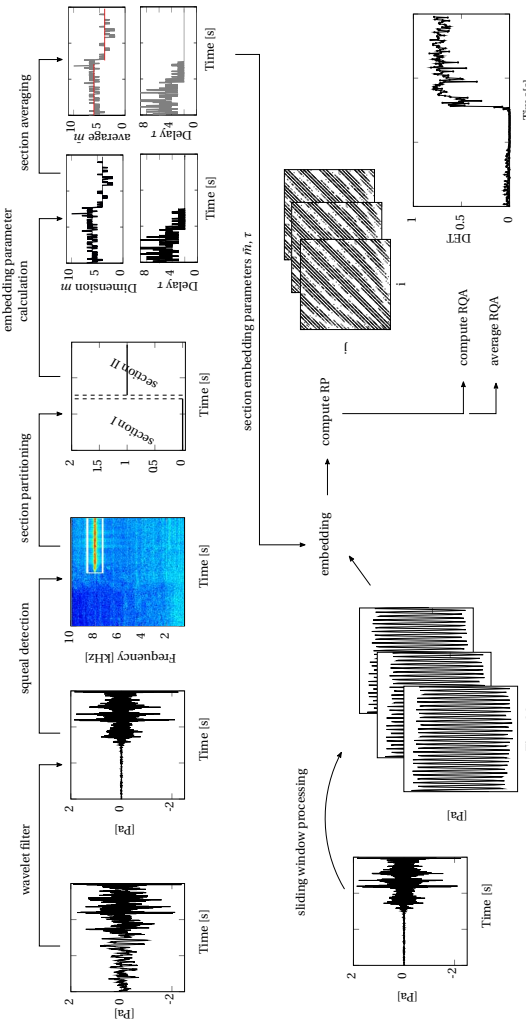


Figure (4.21) Workflow for the nonlinear time series and recurrence plot analysis of large brake vibration data sets. Adapted from Figure 4 from [162]

transformation: The trajectories are unfolded into a state space by time delay embedding Eq. (3.2). The *instantaneous embedding parameters* are computed for each window using the auto-mutual information function and the FNN algorithm. The section-averaged embedding dimension $\bar{m}$ is derived from the 90% percentile of the instantaneous dimensions. $\bar{m}$ is used to compute the state space reconstruction for all windowed sequences in the current section[5]. Recurrence plots are computed from the reconstructed trajectories. The RP norm based on a fixed amount of neighbors FAN= 6 is chosen to allow for a consistent comparison between qualitatively different dynamics [285]. In a last transformation step, the RPQA metrics are evaluated for each RP and hence for each windowed vibration sequence. The resulting time series of RPQA values can be evaluated to characterize the observed dynamics by means of regularity, complexity and stationarity. Furthermore, representative values can be derived per section which allows to compare multiple brakings in a consistent and highly automated fashion, as depicted in the complete workflow in Figure 4.21.

explain: The signals displayed in Figure 4.20 are used as a reference example for illustrating the pre-processing, transformation and analysis phase for a single braking. Figure 4.22 depicts the instantaneous embedding parameters derived for sound pressure level and acceleration measurements. Generally, non-stationary behavior can be observed with significant differences between the microphone and the accelerometer measurements. A larger variance in the parameters is exhibited by the sound signal. Here, the time delay decays from $6 \leq \tau \leq 15$ to $\tau = 2$ as the regime of friction-induced vibrations is approached. Large embedding dimensions $10 \leq m \leq 15$ can be observed in the quiet regime. The dimensions drop to $2 \leq m \leq 4$ during the vibrations. For the accelerations, the delay remains almost constant at $2 \leq \tau \leq 5$ throughout the complete signal. Compared to the sound signal, smaller dimensions can be observed in the quiet section and larger dimensions can be observed in the vibrating section.

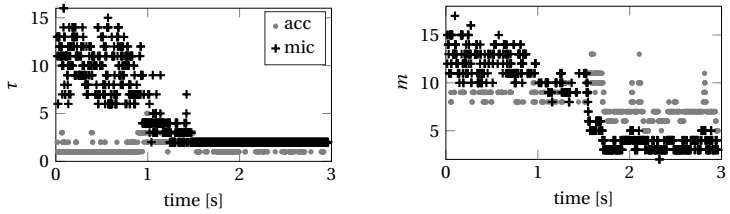


Figure (4.22) Instantaneous embedding parameters delay τ and dimension m for the signals in Figure 4.20. Results are displayed for the filtered microphone (mic) and accelerometer (acc) measurement. Adapted from Figure 6 from [162]

[5] Over-embedding is harmless [284, 198], while under-embedding may heavily affect the results

Figure 4.23 illustrates the reconstructed state space for three representative time instants of the reference signal. The quiet regime, the transient regime at onset of vibrations and the regime of fully developed high-intensity vibrations are shown by projections into a three-dimensional representation. The corresponding recurrence plots and actual embedding parameters are given as well. The unstructured point cloud

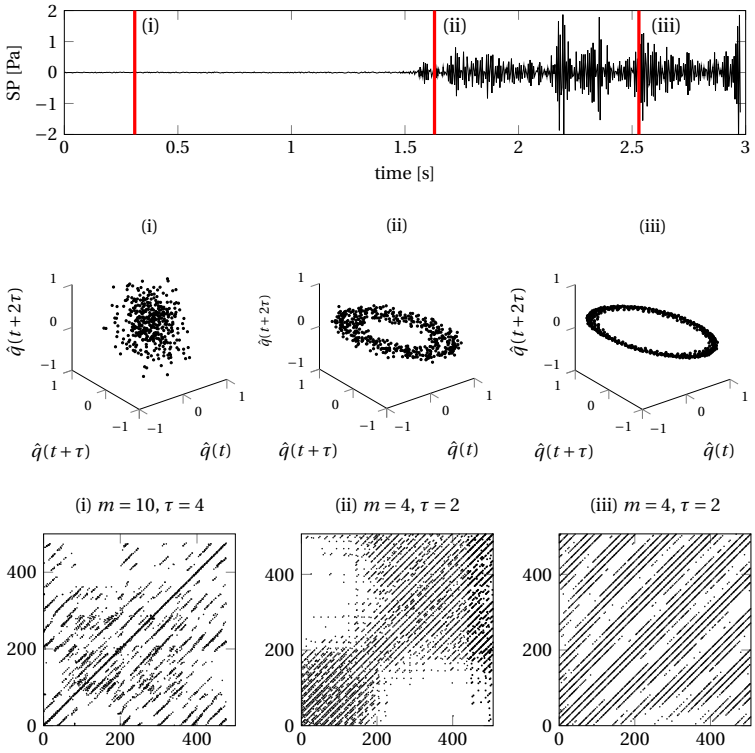


Figure (4.23) Selected state space representations and corresponding recurrence plots for the microphone reference signal displayed in Figure 4.20. Adapted from Figure 8 from [162]

in the quiet regime creates some diagonal line structures in the RP and many individual points that indicate low determinism. Clear diagonals can be observed as the vibrations grow. Long diagonal lines can be observed in the fully developed vibration regime. The high vibration frequency causes only small vertical distance between the diagonal lines. The repetitive white patterns in the RP stem from low-frequent am-

plitude modulations on a slower time scale. However, also the reconstructed trajectories indicate motion on a limit cycle that is slightly perturbed. The corresponding recurrence quantification metrics, see Appendix B.5 are displayed in Figure 4.24. The determinism DET, the fraction of vertical lines structures LAM, the maximal diagonal line length $L_{\max}$, the Shannon entropy ENTR and the clustering coefficient Clust are studied. Clust is a metric rooted in complex network theory that measures the probability of two neighbors of any state being also neighbors. Large values indicate periodic or laminar dynamics [182]. Here, LAM is constantly zero which indicates the absence of vertical lines, i.e. the absence of intermittent dynamics. All other measures show a two-fold behavior that correspond to the quiet and vibrating section of the measured signal. The diagonal lines are very short in the first section, such that DET and $L_{\max}$ exhibit vanishing values. In combination with low Clust and ENTR values, the RPQA metrics indicate highly chaotic or random dynamics in the quiet section. As the FNN algorithm finds low fraction of false nearest neighbors for high embedding dimensions, random dynamics must be assumed to play only a minor role. This observation confirms the findings of Wernitz et al. [68, 69, 286] who state that considerable higher sampling rates are necessary to resolve the dynamics during steady sliding. Qualitatively different dynamics are observed in the second section

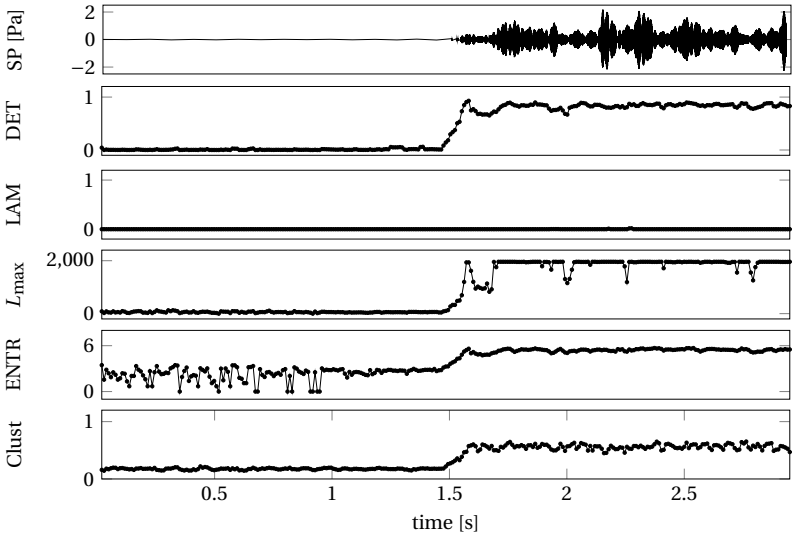


Figure (4.24) Recurrence quantification metrics derived for the reference signal of the microphone recording. High-intensity FIV set in at 1.5 s. Adapted from Figure 9 from [162]

where FIV exist. The nonlinear quantifiers increase, such that high DET and $L_{\max}$ values indicate deterministic and regular dynamics during for the FIV. Interestingly, the RPQA values show a rather stationary behavior whereas the amplitudes of the time series modulate strongly. Hence, the underlying dynamics turn out to be stationary with respect to the nonlinear quantifiers. This characteristic is not visible in the time domain representation of this signal.

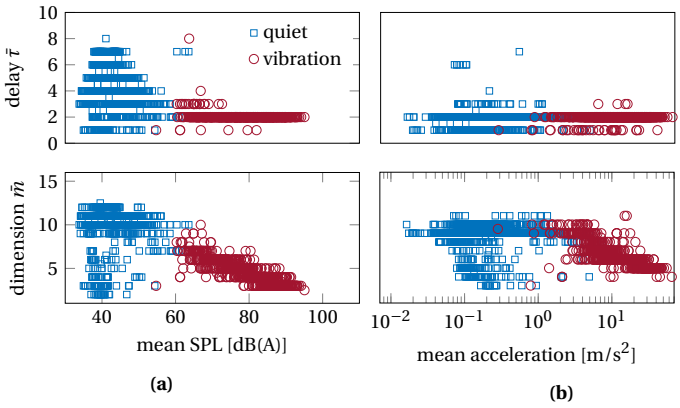


Figure (4.25) Section-median embedding parameters as a function of the mean sound pressure level (a) and the mean acceleration (b). Quiet and vibrating sections are indicated by square and circle markers, respectively. Adapted from Figure 7 from [162]

In a second step, the quantifiers are studied for many brakings. Representative values are derived by averaging (mean, median) the RPQA values per section. The data set contains 1564 quiet sections and 777 sections with FIV. The integer section-averaged (median) instantaneous embedding parameters values $\bar{m}$ and $\bar{\tau}$ are displayed in Figure 4.25 as a function of the mean sound pressure and vibration level per section. Following the NVH sound detection, quiet sections are given for mean sound pressure levels below 60 dB(A) and mean accelerations below 1.2 m/s^2. The delay values scatter significantly for the microphone measurement, while the delay can be considered to be rather constant for the accelerations irrespective of the system response. Generally, larger dimensions can be observed in the quiet sections. The higher the vibration intensity is, the smaller the required embedding dimension become. However, there are also quiet sections that exhibit low-dimensional dynamics. The values match well with the ones reported by Oberst and Lai [37]. Overall, one can find lower dimensions in the microphone measurements than in the acceleration measurements of FIV. The sound radiation to the surrounding air may act like a high pass

filter while the structure-borne vibrations encode large structural vibrations as well as highly chaotic fluctuations from the interface [68, 152].

9.6×10^8 recurrence plots are computed overall for the sliding window processing of the complete brake vibration data set. The observations made for the reference signal are confirmed by the massive data analysis of the complete data set. The corresponding plots are given in Appendix B.5. Essentially, the recurrence quantifiers allow to distinguish random from deterministic and regular from irregular dynamics. High entropy values are observed for the vibrating sections while low values are observed for the quiet sections. The entropy increases linearly with the sound pressure level for values above 60 dB(A) which is a clear indicator for growing levels of determinism in the high-intensity FIV. Stochasticity cannot be ruled out completely for the quiet sections, but the vanishing fraction of false nearest neighbors in high dimensions is a signature of deterministic dynamics. The determinism metric and the diagonal line lengths can give further insights into the level of regularity. Again, those measures increase with larger vibration intensity. Some deviations from those general trends can be observed for all RPQA values. These samples will be treated in the last part of this section. Diagonal lines spanning the complete RP and DET values above 0.9 can be observed for sound pressure levels above 75 dB(A). Therefore, those high-intensity tonal vibrations can be considered to be mostly regular or only weakly irregular. Low determinism and short diagonal lines can be observed for low-intensity FIV and the quiet sections which indicates highly irregular motion. However, the multiscale character of the measurements acquired in harsh testing conditions pose serious challenges to a definite classification into regular or irregular dynamics. Overall, there is a smooth transition from the low-intensity vibrations to the high-intensity vibrations above 60 dB(A). This value seems to be the noise threshold for the microphone measurements. On the other hand, the dynamics do not change qualitatively above a sound pressure level of 75 dB(A). These results seem to be consistent with previous work of Wernitz and Oberst. However, this study is not confined to few measurements, but reports statements that are built on the analysis of complete brake NVH matrix tests with more than 2.5 hours of vibration measurements.

Up to date, the stability of the self-excited system can only be assessed in a boolean fashion during experiments. Either, the system is in a stable sliding condition, or the system exhibits self-excited vibrations. However, the distance of a seemingly stable system to the stability border remains unknown [146]. In this analysis step, the nonlinear quantifiers are related to the operational conditions of the brake system to measure the nonlinear stability margin. It has been observed in the previous studies that the transition from the quiet sections to the vibrating sections is smooth in terms of the RPQA quantifiers, and therefore in terms of qualitative vibration characteristics. Hence, those quantifiers can be used to indicate the likelihood of friction-induced

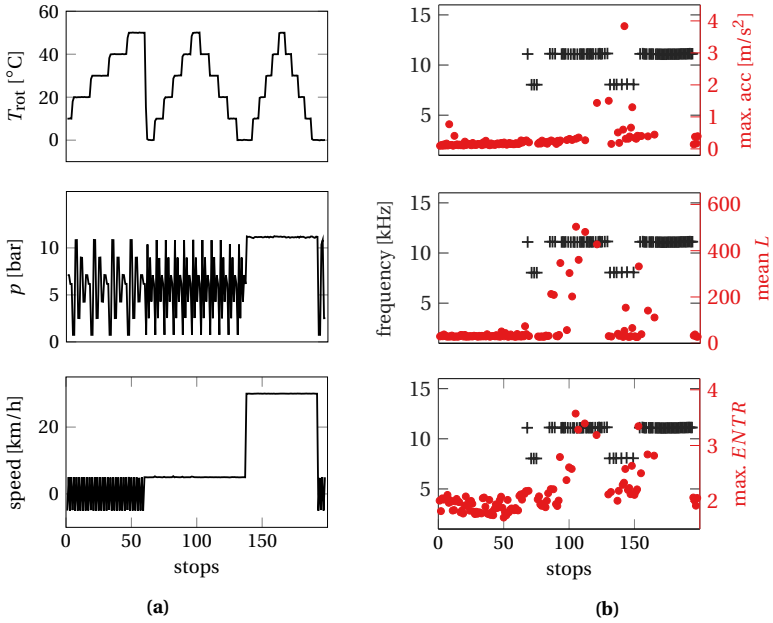


Figure (4.26) Loading conditions (a), acceleration amplitudes and RQA metrics (b) for a segment of a NVH matrix test. The vibration frequency is displayed by cross markers for those brakings that exhibit FIV. The RQA metrics are shown by circles only for those brakings that do *not* exhibit FIV. Adapted from Figure 14 from [162]

vibrations even if vibrations are not indicated in time domain by high vibration amplitudes. Thus, the quantifiers essentially can measure the distance of the current system state to the stability border after which high-intensity vibrations occur. To illustrate this finding, a segment of an NVH matrix test is displayed in Figure 4.26. The loading conditions in terms of disk temperature, brake line pressure and equivalent vehicle speed vary according to the test protocol. The system becomes susceptible to FIV after stop 70 where instabilities at $f = 8$ and $f = 11$ kHz can be observed. RPQA metrics are displayed only for those stops that *do not* exhibit FIV. Low average line lengths and low entropy values can be observed up to stop 70. Hereafter, a gentle increase of the quantifiers can be observed. Especially in the regimes of high vibration occurrence, the non-vibrating brakings exhibit significantly larger RPQA values for ENTR and L than in the beginning of the test. Hence, one can assume that even

if those brakings do not show FIV, they are close to the stability boundary [6]. Only few increased vibration amplitudes can be observed in the acceleration measurements, such that a simple metric based on the vibration amplitudes would not suffice as a continuous measure for stability. On the contrary, the transformation by state space reconstruction and recurrence plot analysis can help to carve out some more pronounced features of the vibration measurements that indicate the distance to the stability border of the dynamical system.

Concluding, the nonlinear analysis of brake vibration measurements indicates strong evidence for deterministic dynamics. High-dimensional irregular motion can be observed during steady sliding and lower-dimensional dynamics are observed during friction-induced vibrations. High-intensity vibrations take the form of regular attractors, and mostly period-1 limit cycles. While these results confirm the findings of other authors, the massive data base analyzed by the proposed data science process is a novel contribution. Additionally, it can be shown that the seemingly non-stationary vibration signals exhibit rather stationary nonlinear invariant measures. These values can be employed as an indicator for the stability margin of the dynamical system during steady sliding.

[6] In fact, those brakings cause the deviations from the general trends mentioned in the RPQA analysis before.

4.7 Instability prediction using deep learning

Realistic dynamical structures are often subjected to several transient loads and variable environmental conditions. Friction-induced vibrations of brake systems are elusive as reported by multiple authors [140, 145, 257, 287] with a lack of reproducibility for seemingly equal system and loading configurations. In most situations it is unknown *if* a system will show FIV and *when* they will occur. In this study, deep learning techniques are employed to find patterns in the operational conditions that give rise to FIV of real automotive disk brake systems. Multivariate loading sequences are mapped to the vibration behavior in an *end-to-end* learning approach, i.e. without specific data transformation steps. Generally, this study presents a first prove of concept and in this spirit, extensive search for the best model configuration using parameter fine-tuning is not pursued.

Table (4.9) Overview on the application case 6: learning the instability behavior of complex friction-excited systems

data	• multivariate time series of loading conditions $\in \mathbb{R}^{m \times n_t}$ • univariate vibration measurements $\in \mathbb{R}^{n_t}$
pre-processing	• vibration detection and sequence labeling • sliding windowing for equal-length sequences
transformation	-
model	• deep recurrent neural networks • long-short-term memory cells, fully connected layers
explain	• determinism in the emergence of FIV • learning instability conditions from experimental data

An overview following the OPTME data science process for structural dynamics is given in Table 4.9. Some of the results and figures presented hereafter have already been presented in

[259] M. Stender, M. Tiedemann, D. Spieler, D. Schoepflin, N. Hoffmann and S. Oberst: *Deep learning for brake squeal: Brake noise detection, characterization and prediction.* Mechanical Systems and Signal Processing, 149, 2021

[288] D. Schoepflin: *Development of a Virtual Twin to Predict Brake System Vibrations Using Recurrent Neural Networks.* Master thesis, Hamburg University of Technology, 2019 .

Figure 4.8 illustrates the system understanding motivating the current study. The mechanical system is subjected to multiple [7] loads and environmental conditions. Some of the loads change rather fast, such as the brake line pressure p, while other conditions, such as the ambient temperature T_{amb} can be considered constant during a single braking. The system interacts with the loads, such that the complete system description is composed of multi-physics effects interacting on multiple time scales in a non-stationary fashion. Through parameter changes in this setting, the structure may undergo a bifurcation and enter an instability regime and exhibit FIV. The loads are understood as proxies that drive dominant damping and stiffness variations in the system. The dynamic behavior of the structure may be learned by a model $\mathscr{M}$ by taking the load signals as multivariate input sequences $\boldsymbol{S}$, and the vibrational response as output sequence $\boldsymbol{s}$

$$\mathscr{M}: \quad \underbrace{\boldsymbol{S} \in \mathbb{R}^{m \times n_{\mathrm{t}}}}_{\text{multivariate input sequences}} \quad \mapsto \quad \underbrace{\boldsymbol{s} \in [0,1]^{(n_{\mathrm{t}})}}_{\text{binary vibration response}} \quad . \tag{4.18}$$

The number of input dimensions, i.e. loading conditions monitored during testing, is denoted by m, and the sequences are sampled at n_{t} equidistant time instants. From a machine learning perspective, time series classification [290] is the conceptual framework for this learning task. In this work, $m = 8$ inputs are considered: the disk rotation Ω,the brake line pressure p, the disk surface temperature T_{rot}, the brake fluid temperature T_{fluid}, the friction coefficient μ, the brake torque M, the ambient air temperature T_{amb} and relative humidity h_{rel}. Certainly, there are more factors that affect the brake system dynamics, such as brake disk and pad surface topologies [291, 292, 293], wear [64, 294] and others. However, the available measurements are sparse in a sense that only a limited number of quantities can be measured. For the output, microphone measurements of the sound pressure are considered. Recurrent neural networks (RNN) are selected for the classification model $\mathscr{M}$. These networks take the sequential character of the input explicitly into account. Just as for classical multi-layer perceptrons, error backpropagation is the most common scheme for parameter updating during the training phase of RNNs[8]. For longer sequences, backpropagation through time can result in the *vanishing gradient* issue [296, 297] of recurrent neural networks: as the error is propagated in time, derivatives get smaller and smaller, such that model parameters related to the first entries of the input sequence will not be affected by the learning process. To overcome this issue, Hochreiter and Schmidhuber [296] developed so-called long-short-term memory networks (LSTM) in 1997 that eliminate the vanishing gradient problem by adding a cell state to the recurrent

[7] Until today, the brake vibration behavior was mostly studied in univariate correlation studies [289].

[8] Backpropagation through time (BPTT) has been introduced by Werbos [295] to account for the sequential structure of the models.

units. LSTM layers allow to identify patterns and temporal correlations in longer sequences. Since 1997, several adaptations for LSTMs and other recurrent model architectures, such as the gated recurrent units (GRU) [298] have been proposed. In this study, conventional LSTM units are used to build deep learning models that can extract dominant patterns in the loading sequences that cause brake vibrations. Two classification tasks are investigated: First, a single label is used as binary output $[0,1]$, such that the model predicts whether the multivariate loading sequences $\boldsymbol{S}$ will cause vibrations (1) or not (0). This task is referred to as *sequence-to-scalar* prediction. Second, a sequential output $\boldsymbol{s} \in [0,1]^{(n_\mathrm{t})}$ is considered, such that the model predicts a binary label at each time step $t_i, i = 1,\dots,n_\mathrm{t}$. Hence, the output of this *sequence-to-sequence* classifier not only predicts if vibrations will be excited by a set of loads, but also at which time instants they will occur.

pre-processing: Data recorded on a noise dynamometer during automotive disk brake testing, see Section 4.3, is studied in the following. The loading conditions are sampled at $f_\mathrm{s} = 100\,\mathrm{Hz}$ while the system response is sampled at 51.2 kHz. Using vibration detections from Section 4.4, binary output sequences are created that encode the observed dynamics into the binary classes of *quiet* (0) and *vibrating* (1) behavior. Following the specifications in Section 4.4, the relevant vibrations are characterized by a tonal periodicity in the range of 1 to 16 kHz with a minimal sound pressure level of 55 dB(A) and a minimal duration of 0.5 s. The data set studied in the following is comprised of 1206 brakings which exhibit vibrations in 487 cases. More than 75% of the vibration durations are longer than 2.5 s and half of the vibrations are longer than 5.6 s. As the brakings have different durations depending on the braking type and braking performance, basic data pre-processing is necessary. Even though LSTM networks can handle different input sequence lengths, training can be eased with sequence data of equal length n_t. Hence, sliding window segmentation of the input and output sequences is introduced for a window length w and overlap Δw. It is physically unclear, if or which time history of the loading conditions has to be considered to predict instability at a time instant t_i, such that the sliding window parameters cannot be fixed a-priori. Therefore, they are part of a hyperparameter study which aims at finding optimal network parameters to achieve a high classification scores for the present data. Obviously, these parameters may change for different systems that exhibit different instability behavior and load history effects.

modeling: Hyperparameters of the classification model include the number of hidden LSTM layers, the number of LSTM units per layer and the batch size used during training. Generally, the models are composed of n_layers LSTM layers featuring n_units units for the $m = 8 \times n_\mathrm{t}$-dimensional input sequences, a fully connected (FC) layer with ReLu activation to flatten the LSTM outputs and the FC output layer with sigmoid activation that has either one output for the sequence-to-scalar task or n_t

outputs for the sequence-to-sequence task. A constant dropout rate of 0.1 is used in the LSTM layers to increase generalization and prevent overfitting. The model architectures considered during the hyperparameter studies for both classification tasks are summarized in Table 4.10. The adam [208] optimizer is used along with the *binary cross-entropy* loss. To account for the strongly imbalanced class distributions, Matthews correlation coefficient MCC from Equation (3.5) is employed as classification score. To obtain more representative results for the rather small data set, threefold cross-validation is used during the hyperparameter study. Models are trained for 300 epochs and the average MCC score on the validation sets is used to find the optimal model architecture. Although further performance increase would be possible through longer training, the chosen setup allows to obtain a first overview on the required model complexity.

Table (4.10) Model configurations studied in the hyperparameter search for the sequence input length $w = n_\mathrm{t} \in [200,400]$ with overlap $\Delta w \in [0,25,50\%]$, number of LSTM units $n_\mathrm{unit} \in [64,128,256]$ and number of LSTM layers $n_\mathrm{layers} \in [1,2]$ using batch sizes of $n_\mathrm{batch} \in [16,64,256]$. m is the number of measurement dimensions in the input sequences, i.e. the number of sensors

model	layer	input shape	output shape	configuration
sequence to scalar	LSTM #1	$[m \times n_\mathrm{t} \times n_\mathrm{units}]$	$[m \times n_\mathrm{t} \times n_\mathrm{units}]$	dropout= 0.1
	LSTM #2	$[m \times n_\mathrm{t} \times n_\mathrm{units}]$	$[m \times n_\mathrm{units}]$	dropout= 0.1
	FC	$[m \times n_\mathrm{units}]$	$[n_\mathrm{units}]$	ReLu activ.
	FC	$[n_\mathrm{units}]$	$[1]$	sigmoid activ.
sequence to sequence	LSTM #1	$[m \times n_\mathrm{t} \times n_\mathrm{units}]$	$[m \times n_\mathrm{t} \times n_\mathrm{units}]$	dropout= 0.1
	LSTM #2	$[m \times n_\mathrm{t} \times n_\mathrm{units}]$	$[m \times n_\mathrm{t}]$	dropout= 0.1
	FC	$[m \times n_\mathrm{t}]$	$[n_\mathrm{t}]$	sigmoid activ.

explain: All results of the hyperparameter study are reported in Appendix C.2, while this paragraph summarizes the most relevant findings. First of all, the results of the hyperparameter study indicate that it is in fact possible to predict the brake vibration response from the set of loading parameters available in the present data set: Classification score significantly larger than MCC = 0 are observed[9]. Hence, the recurrent neural network models are able to predict the correct output at significantly better accuracy than random guessing. This has not been certain in the beginning, as the load parameters might have been completely uncorrelated with the system response. Hence, the general possibility of predicting the system dynamics is the first,

[9] The MCC classification score is zero for a classifier that is not better than random and one for perfect classification.

and maybe most important, result of this study. In the hyperparameter study for the sequence-to-scalar classifier, average cross-validation scores of MCC = 0.51 to 0.74 can be observed. Overall, the classifiers are correctly predicting the emergence of vibrations in most cases, even though there is still room for improvement. The hyperparameter study suggests to select wide models with 256 LSTM units per layer. Particularly, the training batch size has a major impact on the classification score, while the number of LSTM layers has a vanishing effect. In combination with sequence pre-processing using $n_t = 200$ samples with 25% overlap, these models achieve a high classification score. Even though the hyperparameter search grid is rather coarse, the final sequence-to-scalar model architecture ($n_{layers} = 1$, $n_{units} = 256$, $w = n_t = 200$, $\Delta w = 25\%$, $n_{batch} = 256$) is selected from those preliminary results.

For the final evaluation, such a network is trained using a stratified 70-30 data split for 500 epochs. To obtain a broader picture, three additional data sets are studied, hereafter referred to as data sets *B*, *C* and *D*. The initially studied data are denoted by set *A*. The model architectures that were found for system *A* are re-used for the new data, such that no hyperparameter studies are performed for sets *B* to *D*. The physical braking systems of data *A* and *B* are very similar, and so are the systems for data *C* and *D*. Class imbalance is observed for all systems: *A* (1206 brakings / 487 with vibrations), *B* (1206 / 227), *C* (1206 / 347), *D* (1889 / 237). Ten individual classifiers are generated for each data set. Table 4.11 reports the sequence-to-scalar classification scores on the main diagonal. High scores up to MCC = 0.78 can be observed for systems *A* and *B*, while lower scores are obtained for the two other brake systems which are physically different to the first two systems. The ultimate objective of research on these dynami-

Table (4.11) Evaluation of the sequence-to-scalar classifiers: the main diagonal reports the validation results in terms of the average MCC score. The off-diagonal entries report the MCC scores for the cross-evaluation study where a model trained on one data set is used to make predictions on another data set. Additionally, the results for a model trained and evaluated on all data at once is shown

		model trained on				
		A	*B*	*C*	*D*	all
model evaluated on	*A*	**0.78**	0.45	-0.08	-0.09	-
	B	0.46	**0.72**	0.01	0.01	-
	C	-0.02	0.16	**0.65**	0.47	-
	D	-0.2	-0.01	0.43	**0.64**	-
	all	-	-	-	-	**0.24**

cal systems is to find out which underlying mechanisms trigger the instability. Hence, if a single network was able to predict instability for various different brake systems, the network can be thought of as having discovered dominant patterns in the data that dictate the vibration behavior irrespective of geometric properties of the individual brake system. Following up on this reasoning, cross-evaluation is performed for the four models and the four data sets. A model $\mathscr{M}_i$ trained on data set $i = 1,\dots,4$ is used to make predictions on data set $\mathscr{D}_j$ $j = 1,\dots,4; j \neq i$. The resulting classification scores (i,j) are reported in the off-diagonal entries of Table 4.11. Interestingly, intra-family cross-evaluation[10] shows quality scores that can still be considered better than random in the range of $0.43 \leq \text{MCC} \leq 0.47$. Inter-family cross-evaluation results in scores that are close to zero, therefore indicating completely random model predictions. Even though the classification scores are not exceptionally high, it seems that the members of the same brake system family share some instability patterns that can be learned by machine learning models. However, these patterns are individual to the family of brake systems, such that virtual twins developed for one brake system family cannot be used for a different family (at least in this study). Finally, all data from all systems are merged and a single classification model is built. The resulting score of MCC = 0.24 supports the previous findings: in the given scenario, the network cannot identify a strong common instability pattern that governs all brake systems irrespective of their geometry and performance.

The same procedure is repeated for the sequence-to-sequence classifiers. The hyperparameter study reported in Appendix C.2 indicates that the best model architecture ($n_{\text{layers}} = 1$, $n_{\text{units}} = 256$, $w = n_{\text{t}} = 400$, $\Delta w = 25\%$, $n_{\text{batch}} = 256$) achieves a score of MCC = 0.78 for data set A. Therefore, the instantaneous vibration behavior can be predicted at reasonable accuracy at each time step, which is a promising result. It must be noted that direct comparison between the sequence-to-scalar and sequence-to-sequence classification scores has to be treated with caution. For the sequential output, the MCC is computed based on the prediction per time step. If the vibration duration is not predicted perfectly, i.e. some time steps show prediction errors, a reduced MCC value results. On the contrary, in the scalar prediction only the overall behavior, i.e. vibrations during braking or not, needs to be predicted correctly.

Overall, the sequence-to-sequence classification results resemble those of the sequence-to-scalar classification, as reported in Table 4.12. The highest scores are obtained for systems A and B, and intra-family cross-evaluation yields better-than-random predictions. The chosen models cannot be used to predict the vibrational behavior of different brake system families, and one common instability pattern cannot be extracted when using all the available data for building a meta-model for all brake systems for which data are available in this study.

[10] i.e. A versus B, and C versus D

Table (4.12) Evaluation of the sequence-to-sequence classifiers: the main diagonal reports the validation results in terms of the average MCC score. The off-diagonal entries report the MCC scores for the cross-evaluation study where a model trained on one data set is used to make predictions on another data set. Additionally, the results for a model trained and evaluated on all data at once is shown

		model trained on				
		A	*B*	*C*	*D*	all
model evaluated on	*A*	**0.78**	0.46	-0.04	-0.04	-
	B	0.51	**0.75**	0.04	-0.01	-
	C	0.03	0.15	**0.62**	0.41	-
	D	-0.08	-0.01	0.37	**0.50**	-
	all	-	-	-	-	**0.23**

Figure 4.27 illustrates the capabilities of the neural network-based virtual twin for the NVH behavior of brake systems: given the sequential loading parameters as input, the validated model predicts the onset and duration of a high-intensity vibration event with high precision. Prediction errors can be observed only in the time steps of the onset of vibrations, which then result in a classification score of MCC = 0.85 for this single braking. This value may guide the reader to put the previously reported classification scores into perspective. The confidence value in Figure 4.27 indicates that at most time steps the model is very confident in its predictions. Only at the onset of vibrations and at two instants in the vibrating phase the confidence value drops shortly. These two points correlate to the sliding window processing, which segmented this braking into three sub-epochs. At the beginning of each window, the confidence score is low because the model does not have many prior time steps to take into consideration for the prediction.

The results of this study are not only relevant and promising for research on structural dynamics. In commercial development departments, training virtual NVH twins for brake systems may reduce the amount of testing. Once a prediction model has achieved the required prediction quality, no more hardware-based testing may be necessary. Instead, other loading scenarios, generated synthetically or recorded in historic test campaigns, can be fed through the model to evaluate the NVH performance virtually. These models are not predictive in a sense of numerical modeling where no a-priori hardware-based testing is required, but comparatively small amounts of testing may suffice for the training of predictive deep learning models. The end-to-end learning approach performed in this study does not require a many

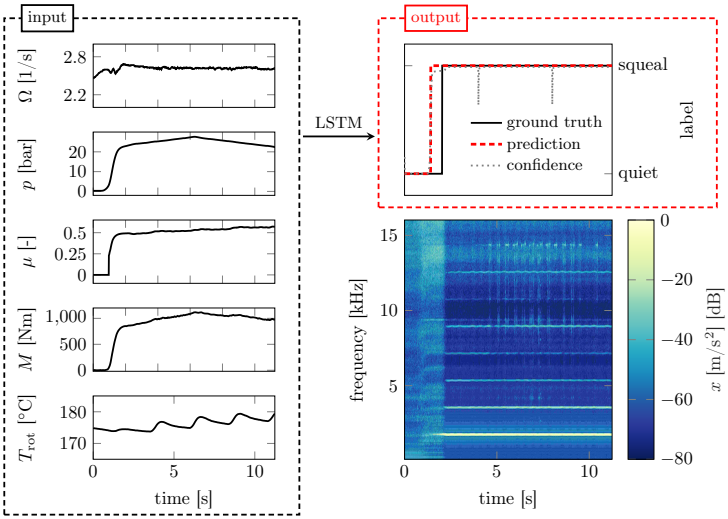


Figure (4.27) Output of the sequence-to-sequence classification model for a validation sample: left panel displays the loading signals used as input sequences for the model. The top right panel depicts the model prediction as well as the model's confidence value given by the deviation of the dotted line from the dashed line. Adapted from Figure 11 from [259]

data transformations or feature engineering phase. As a downside of black-box-like models that employ end-to-end training, the level of interpretability is very weak. The deep learning models do not offer a simplistic way to explain their decision making process. Hence, there is no direct way to derive countermeasures for the actual mechanical design of the brake system, i.e. by changing the friction characteristic if $\mu(t)$ was found to be a driver for instability. Future work will thus focus on explainability-preserving models or model-agnostic diagnosis [164, 214] of the virtual NVH twins developed in this work.

Concluding, the findings illustrate a first prove of concept that may have far-reaching benefits in the braking industry. For the scientific disciplines involved in tribology and structural dynamics, the results indicate that instabilities of real world brake systems are in fact following deterministic dependencies encoded in the complex loading scenarios. The vibration behavior can be predicted by deep recurrent neural networks trained on rather small data sets. Their classification performance is good, even though there remains some room for improvements. For the present data sets it was not possible to find models that generalize for multiple brake systems. Hence,

the instability behavior was learned only locally for single brake system families. To close the gap to very high prediction quality, there are two main pathways: first, other model architectures may leverage the full potential of the recurrent networks used here. Furthermore, performance gains may be achieved through more data, and specifically more measurement dimensions. Particularly, the friction interface is well-known to strongly affect the overall dynamics [140], such that the global friction value measurement is certainly capturing only some of the processes involved in the highly complex interface between brake pad and disk. However, the range of potential system parameters that act as bifurcation parameter can be very large: load-history effects, joints, dependency on initial conditions, and synchronization may play a role in a specific system. As these are mostly hard to measure, especially in a mechanical engineering-compliant environment, sparsity in the measurements will always remain a factor. Lastly, there will certainly be some sort of stochastic effects that affect the system stability. All of these factors put together may be root causes for the performance gap to perfect prediction results.

5 Conclusion and perspectives

This thesis discusses chances and limitations of data-driven methods for the analysis of complex mechanical vibrations. The first part targets sharpening the view of practitioners and researchers with regards to several aspects of nonlinear, and especially chaotic, dynamics. In fact, the comparison of nonlinear structural dynamics to concepts of complexity from other scientific disciplines reveals many common salient features. Those scientific disciplines have developed a long track in evidence-based, i.e. top-down or data-driven, methods for studying complex dynamics of the climate, the human brain and other systems. On the contrary, mechanical systems and their dynamics are typically studied via bottom-up approaches, i.e. through numerical models built on first principles. Challenges in nonlinear structural dynamics, and mostly those affected by friction, call for novel data-driven techniques. This thesis proposes several system identification approaches that can make use of the exponentially growing amount of data that has become accessible in recent years.

The second part of this thesis proposes a novel data science process denoted as 'OPTME'. The process is adapted from classical data science processes with special consideration of the peculiarities of complex structural dynamics. Here, special focus is put on physics-conform and interpretable methods for knowledge discovery. The proposed data science process may serve as a guideline to consider various aspects of complex dynamics, such as non-stationarity, multi-scale behavior, or irregularity, before selecting appropriate observables and machine learning models for a given task. Furthermore, the explicit consideration of dynamics-specific data characteristics in the model building and evaluation phase may help to overcome limitations of current approaches to structural dynamics. Particularly, this thesis indicates how to treat limited data availability and highlights potential pitfalls when working with small and unbalanced data sets. This kind of data naturally results from measurements of nonlinear structural vibrations. Furthermore, limitations to physical interpretability of data-driven techniques are discussed.

The third part demonstrates the usage of the novel data science process for the analysis and identification of complex mechanical systems and their dynamics. In six case studies, novel data-driven system identification, characterization, and prediction methods are presented. Using the OPTME process, governing equations of motions are distilled from vibration measurements, instantaneous damping properties

are extracted from unstable vibrations, and a purely data-driven stability estimation for the steady sliding state of brake systems is presented. The initial hypothesis of multi-scale, non-stationarity, and aperiodic dynamics in complex friction-affected systems is confirmed by the analysis of analytical models, small laboratory experiments, and automotive brake systems. The seemingly erratic friction-induced vibrations of these systems exhibit rather stationary nonlinear invariant measures, which can be used as evidence-based warning signs for the onset of self-excited vibrations. Furthermore, deep learning techniques are shown to enable instability prediction from system loading conditions. Hence, the underlying mechanisms can be assumed to be mostly deterministic, even if vibrations appear to be elusive. Machine learning models enable pattern recognition in large, high-dimensional and multi-variate data streams, and hence to explain the elusive vibrations. It is found that the identified instability patterns are characteristic to a single brake system, but common patterns are identified for families of geometrically similar systems.

In future work, aspects of random motion can be incorporated into the data science process to account for uncertainty and aspects of stochasticity. Explainability-preserving modeling approaches will attract increasing attention in the future to extract physically meaningful knowledge from data-driven black-box approaches. Moreover, hybrid approaches that link data-driven methods with conventional physics-based models will be able to leverage the full potential of both approaches to system identification and modeling in structural dynamics.

Bibliography

[1] G. Kerschen, K. Worden, A. F. Vakakis, and J.-C. Golinval, Past, present and future of nonlinear system identification in structural dynamics, *Mechanical Systems and Signal Processing*, vol. 20, no. 3, pp. 505–592, 2006.

[2] I. Newton and E. Halley, *Philosophiae naturalis principia mathematica*. Diapositivas (Biblioteca Histórica UCM), Jussu Societatis Regiae ac typis Josephi Streater, prostant venales apud Sam. Smith, 1687.

[3] H. Poincaré, *Les méthodes nouvelles de la mécanique céleste*. Gauthier-Villars, 1892.

[4] E. N. Lorenz, Deterministic nonperiodic flow, *Journal of the Atmospheric Sciences*, vol. 20, no. 2, pp. 130–141, 1963.

[5] M. J. Feigenbaum, Quantitative universality for a class of nonlinear transformations, *Journal of Statistical Physics*, vol. 19, no. 1, pp. 25–52, 1978.

[6] B. Mandelbrot, *The fractal geometry of nature*. San Francisco: Freeman, (1982).

[7] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, vol. 42 of *Applied Mathematical Sciences*. New York, NY and s.l.: Springer New York, 1983.

[8] A. H. Nayfeh and D. T. Mook, *Nonlinear Oscillations*. Wiley, 1995.

[9] F. Verhulst, *Nonlinear Differential Equations and Dynamical Systems*. Universitext, Berlin and Heidelberg: Springer, second, revised and expanded edition ed., 1996.

[10] S. H. Strogatz, *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering*. Studies in Nonlinearity, Cambridge, Mass.: Perseus Books, 2. print ed., 2001.

[11] K. Worden and G. R. Tomlinson, *Nonlinearity in structural dynamics: Detection, identification and modelling*. Bristol: Institute of Physics Publ, 2001.

[12] D. Ruelle and F. Takens, On the nature of turbulence, *Les rencontres physiciens-mathématiciens de Strasbourg-RCP25*, vol. 12, pp. 1–44, 1971.

[13] L. M. Pecora, T. L. Carroll, G. A. Johnson, D. J. Mar, and J. F. Heagy, Fundamentals of synchronization in chaotic systems, concepts, and applications, *Chaos (Woodbury, N.Y.)*, vol. 7, no. 4, pp. 520–543, 1997.

[14] J.-P. Eckmann, Roads to turbulence in dissipative dynamical systems, *Reviews of Modern Physics*, vol. 53, no. 4, pp. 643–654, 1981.

[15] M. J. Feigenbaum, The onset spectrum of turbulence, *Physics Letters A*, vol. 74, no. 6, pp. 375–378, 1979.

[16] M. J. Feigenbaum, The transition to aperiodic behavior in turbulent systems, *Communications in Mathematical Physics*, vol. 77, no. 1, pp. 65–86, 1980.

[17] M. Giglio, S. Musazzi, and U. Perini, Transition to chaotic behavior via a reproducible sequence of period-doubling bifurcations, *Physical review letters*, vol. 47, no. 4, pp. 243–246, 1981.

[18] M. J. Feigenbaum, The universal metric properties of nonlinear transformations, *Journal of Statistical Physics*, vol. 21, no. 6, pp. 669–706, 1979.

[19] Y. Pomeau and P. Manneville, Intermittent transition to turbulence in dissipative dynamical systems, *Communications in Mathematical Physics*, vol. 74, no. 2, pp. 189–197, 1980.

[20] A. V. Krysko, J. Awrejcewicz, I. V. Papkova, and V. A. Krysko, Routes to chaos in continuous mechanical systems: Part 2. modelling transitions from regular to chaotic dynamics, *Chaos, Solitons & Fractals*, vol. 45, no. 6, pp. 709–720, 2012.

[21] S. Newhouse, D. Ruelle, and F. Takens, Occurrence of strange axiom a attractors near quasiperiodic flows on T^m, $m \geq 3$, *Comm. Math. Phys.*, vol. 64, no. 1, pp. 35–40, 1978.

[22] C. Grebogi, E. Ott, and J. A. Yorke, Crises, sudden changes in chaotic attractors, and transient chaos, *Physica D: Nonlinear Phenomena*, vol. 7, no. 1-3, pp. 181–200, 1983.

[23] H. G. Schuster and W. Just, *Deterministic chaos: An introduction*. John Wiley & Sons, 2006.

[24] C. E. Shannon, A mathematical theory of communication, *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[25] Y. G. Sinai, On the notion of entropy of a dynamical system, in *Dokl. Akad. Nauk. SSSR*, vol. 124, p. 768, 1959.

[26] S. M. Pincus, Approximate entropy as a measure of system complexity, *Proceedings of the National Academy of Sciences*, vol. 88, no. 6, pp. 2297–2301, 1991.

[27] F. Takens, Detecting strange attractors in turbulence, in *Dynamical systems and turbulence, Warwick 1980*, pp. 366–381, Springer, 1981.

[28] H. V. Koch, Sur une courbe continue sans tangente, obtenue par une construction géométrique élémentaire, *Arkiv for Matematik, Astronomi och Fysik*, vol. 1, pp. 681–704, 1904.

[29] P. Grassberger and I. Procaccia, Characterization of strange attractors, *Physical review letters*, vol. 50, no. 5, pp. 346–349, 1983.

[30] A. Rényi, *Probability theory*. Dover Books on Mathematics, Mineola, New York: Dover Publications Inc, 2007.

[31] J. A. Skjeltorp, Scaling in the norwegian stock market, *Physica A: Statistical Mechanics and its Applications*, vol. 283, no. 3-4, pp. 486–528, 2000.

[32] H. Berry, D. Gracia Pérez, and O. Temam, Chaos in computer performance, *Chaos (Woodbury, N.Y.)*, vol. 16, no. 1, p. 013110, 2006.

[33] J. L. Hudson and J. C. Mankin, Chaos in the Belousov–Zhabotinskii reaction, *The Journal of Chemical Physics*, vol. 74, no. 11, pp. 6171–6177, 1981.

[34] A. Babloyantz and A. Destexhe, Low-dimensional chaos in an instance of epilepsy, *Proceedings of the National Academy of Sciences*, vol. 83, no. 10, pp. 3513–3517, 1986.

[35] J. Awrejcewicz, Chaos in simple mechanical systems with friction, *Journal of Sound and Vibration*, vol. 109, no. 1, pp. 178–180, 1986.

[36] K. Popp and P. Stelter, Stick-slip vibrations and chaos, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 332, no. 1624, pp. 89–105, 1990.

[37] S. Oberst and J. Lai, Chaos in brake squeal noise, *Journal of Sound and Vibration*, vol. 330, no. 5, pp. 955–975, 2011.

[38] V. Pilipchuk, P. Olejnik, and J. Awrejcewicz, Transient friction-induced vibrations in a 2-dof model of brakes, *Journal of Sound and Vibration*, vol. 344, pp. 297–312, 2015.

[39] D. L. Turcotte, *Fractals and chaos in geology and geophysics*. Cambridge u.a.: Cambridge Univ. Press, 2. ed., 1997.

[40] J. H. Argyris, G. Faust, M. Haase, and R. Friedrich, *Die Erforschung des Chaos: Eine Einführung in die Theorie nichtlinearer Systeme*. Springer complexity, Berlin: Springer, völlig neu bearb. und erw. 2. aufl. ed., 2010.

[41] C. Grebogi, E. Ott, S. Pelikan, and J. A. Yorke, Strange attractors that are not chaotic, *Physica D: Nonlinear Phenomena*, vol. 13, no. 1-2, pp. 261–268, 1984.

[42] S. C. Lluís Alsedà, On the definition of strange nonchaotic attractor, *Fundamenta Mathematicae*, vol. 206, no. 1, pp. 23–39, 2009.

[43] R. Gopal, A. Venkatesan, and M. Lakshmanan, Applicability of 0-1 test for strange nonchaotic attractors, *Chaos (Woodbury, N.Y.)*, vol. 23, no. 2, p. 023123, 2013.

[44] S. Bograd, P. Reuss, A. Schmidt, L. Gaul, and M. Mayer, Modeling the dynamics of mechanical joints, *Mechanical Systems and Signal Processing*, vol. 25, no. 8, pp. 2801–2826, 2011.

[45] S. Kruse, M. Tiedemann, B. Zeumer, P. Reuss, H. Hetzler, and N. Hoffmann, The influence of joints on friction induced vibration in brake squeal, *Journal of Sound and Vibration*, vol. 340, pp. 239–252, 2015.

[46] M. R. Brake, *The Mechanics of Jointed Structures*. Cham: Springer International Publishing, 2018.

[47] S. Carrino, F. Nicassio, and G. Scarselli, Subharmonics and beating: A new approach to local defect resonance for bonded single lap joints, *Journal of Sound and Vibration*, vol. 456, pp. 289–305, 2019.

[48] M. Brake, C. W. Schwingshackl, and P. Reuß, Observations of variability and repeatability in jointed structures, *Mechanical Systems and Signal Processing*, vol. 129, pp. 282–307, 2019.

[49] L. Pesaresi, L. Salles, R. Elliott, A. Jones, J. S. Green, and C. W. Schwingshackl, Numerical and experimental investigation of an underplatform damper test rig, in *Dynamics and Control of Technical Systems II* (J. Balthazar, P. B. Gonçalves, and J. Náprstek, eds.), Applied Mechanics and Materials Vol. 849, (Zurich), pp. 1–12, Trans Tech Publications Limited, 2016.

[50] R. D. Blevins, Flow-induced vibration, *New York, Van Nostrand Reinhold Co.*, p. 377, 1977.

[51] C. Weiss, P. Gdaniec, N. P. Hoffmann, A. Hothan, G. Huber, and M. M. Morlock, Squeak in hip endoprosthesis systems: An experimental study and a numerical technique to analyze design variants, *Medical engineering & physics*, vol. 32, no. 6, pp. 604–609, 2010.

[52] N. H. Fletcher and T. D. Rossing, *The physics of musical instruments*. Springer Science & Business Media, 2012.

[53] R. A. Ibrahim, Friction-induced vibration, chatter, squeal, and chaos–Part I: Mechanics of contact and friction, *Applied Mechanics Reviews*, vol. 47, no. 7, p. 209, 1994.

[54] R. A. Ibrahim, Friction-induced vibration, chatter, squeal, and chaos–Part II: Dynamics and modeling, *Applied Mechanics Reviews*, vol. 47, no. 7, p. 227, 1994.

[55] F. C. Moon and T. Kalmár-Nagy, Nonlinear models for complex dynamics in cutting materials, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 359, no. 1781, pp. 695–711, 2001.

[56] N. M. Kinkaid, O. M. O'Reilly, and P. Papadopoulos, Automotive disc brake squeal, *Journal of Sound and Vibration*, vol. 267, no. 1, pp. 105–166, 2003.

[57] M. North, Disc brake squeal. braking of road vehicles, automobile division of the institution of mechanical engineers, *Mechanical Engineering Publications Limited, London, England*, 1976.

[58] N. Hoffmann, M. Fischer, R. Allgaier, and L. Gaul, A minimal model for studying properties of the mode-coupling type instability in friction induced oscillations, *Mechanics Research Communications*, vol. 29, no. 4, pp. 197–205, 2002.

[59] Q. S. Nguyen, Instability and friction, *Comptes Rendus Mécanique*, vol. 331, no. 1, pp. 99–112, 2003.

[60] R. T. Spurr, A theory of brake squeal, *Proceedings of the Institution of Mechanical Engineers: Automobile Division*, vol. 15, no. 1, pp. 33–52, 1961.

[61] I. L. Singer and H. M. Pollock, *Fundamentals of Friction: Macroscopic and Microscopic Processes*, vol. 220 of *NATO ASI Series, Series E.* Dordrecht: Springer, 1992.

[62] M. Urbakh, J. Klafter, D. Gourdon, and J. Israelachvili, The nonlinear nature of friction, *Nature*, vol. 430, p. 525, 2004.

[63] A. Saha, M. Wiercigroch, K. Jankowski, P. Wahi, and A. Stefański, Investigation of two different friction models from the perspective of friction-induced vibrations, *Tribology International*, vol. 90, pp. 185–197, 2015.

[64] A. Cabboi and J. Woodhouse, Validation of a constitutive law for friction-induced vibration under different wear conditions, *Wear*, vol. 396-397, pp. 107–125, 2018.

[65] R. Stribeck, *Die wesentlichen Eigenschaften der Gleit- und Rollenlager*. Mitteilungen über Forschungsarbeiten auf dem Gebiete des Ingenieurwesens, insbesondere aus den Laboratorien der technischen Hochschulen, Berlin: Springer, 1903.

[66] N. Hoffmann and L. Gaul, Friction induced vibrations of brakes: Research fields and activities, SAE Technical Paper Series, SAE International, Warrendale, PA, United States, 2008.

[67] N. Hoffmann, Linear stability of steady sliding in point contacts with velocity dependent and lugre type friction, *Journal of Sound and Vibration*, vol. 301, no. 3-5, pp. 1023–1034, 2007.

[68] B. Wernitz and N. Hoffmann, Recurrence analysis and phase space reconstruction of irregular vibration in friction brakes: Signatures of chaos in steady sliding, *Journal of Sound and Vibration*, vol. 331, no. 16, pp. 3887–3896, 2012.

[69] N. K. Vitanov, N. Hoffmann, and B. Wernitz, Nonlinear time series analysis of vibration data from a friction brake: Ssa, pca, and mfdfa, *Chaos, Solitons & Fractals*, vol. 69, pp. 90–99, 2014.

[70] S. Oberst and J. Lai, A statistical approach to estimate the lyapunov spectrum in disc brake squeal, *Journal of Sound and Vibration*, vol. 334, pp. 120–135, 2015.

[71] S. Oberst and J. Lai, Nonlinear transient and chaotic interactions in disc brake squeal, *Journal of Sound and Vibration*, vol. 342, pp. 272–289, 2015.

[72] J. Awrejcewicz and P. Olejnik, Friction pair modeling by a 2-dof system: numerical and experimental investigations, *International Journal of Bifurcation and Chaos*, vol. 15, no. 06, pp. 1931–1944, 2005.

[73] C. W. Schwingshackl, E. P. Petrov, and D. J. Ewins, Effects of contact interface parameters on vibration of turbine bladed disks with underplatform dampers, *Journal of Engineering for Gas Turbines and Power*, vol. 134, no. 3, p. 032507, 2012.

[74] B. Khojasteh, M. Janko, and Y. Visell, Complexity, rate, and scale in sliding friction dynamics between a finger and textured surface, *Scientific reports*, vol. 8, no. 1, p. 13710, 2018.

[75] M. Tiedemann, S. Kruse, and N. Hoffmann, Dominant damping effects in friction brake noise, vibration and harshness: The relevance of joints, *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 229, no. 6, pp. 728–734, 2015.

[76] J. Machowski, J. W. Bialek, and J. R. Bumby, *Power system dynamics: Stability and control*. Chichester: Wiley, 2. ed., 2008.

[77] R. M. May, Thresholds and breakpoints in ecosystems with a multiplicity of stable states, *Nature*, vol. 269, no. 5628, pp. 471–477, 1977.

[78] M. Hirota, M. Holmgren, E. H. van Nes, and M. Scheffer, Global resilience of tropical forest and savanna to critical transitions, *Science (New York, N.Y.)*, vol. 334, no. 6053, pp. 232–235, 2011.

[79] P. J. Menck, J. Heitzig, N. Marwan, and J. Kurths, How basin stability complements the linear-stability paradigm, *Nature Physics*, vol. 9, no. 2, pp. 89–92, 2013.

[80] A. Papangelo, M. Ciavarella, and N. Hoffmann, Subcritical bifurcation in a self-excited single-degree-of-freedom system with velocity weakening–strengthening friction law: analytical results and comparison with experiments, *Nonlinear Dynamics*, vol. 90, no. 3, pp. 2037–2046, 2017.

[81] A. Papangelo, A. Grolet, L. Salles, N. Hoffmann, and M. Ciavarella, Snaking bifurcations in a self-excited oscillator chain with cyclic symmetry, *Communications in Nonlinear Science and Numerical Simulation*, vol. 44, pp. 108–119, 2017.

[82] A. Grolet and F. Thouverez, Computing multiple periodic solutions of nonlinear vibration problems using the harmonic balance method and groebner bases, *Mechanical Systems and Signal Processing*, vol. 52-53, pp. 529–547, 2015.

[83] O. V. Astakhov, S. V. Astakhov, N. S. Krakhovskaya, V. V. Astakhov, and J. Kurths, The emergence of multistability and chaos in a two-mode van der pol generator versus different connection types of linear oscillators, *Chaos (Woodbury, N.Y.)*, vol. 28, no. 6, pp. 63–118, 2018.

[84] A. Papangelo, F. Fontanela, A. Grolet, M. Ciavarella, and N. Hoffmann, Multistability and localization in forced cyclic symmetric structures modelled by weakly-coupled duffing oscillators, *Journal of Sound and Vibration*, vol. 440, pp. 202–211, 2019.

[85] A. D. Dimarogonas, S. A. Paipetis, and T. G. Chondros, *Analytical Methods in Rotor Dynamics: Second Edition*, vol. v.9 of *Mechanisms and Machine Science*. Dordrecht: Springer, 2012.

[86] M. Krack, L. Panning-von Scheidt, and J. Wallaschek, A method for nonlinear modal analysis and synthesis: Application to harmonically forced and self-excited mechanical systems, *Journal of Sound and Vibration*, vol. 332, no. 25, pp. 6798–6814, 2013.

[87] M. Krack, L. Salles, and F. Thouverez, Vibration prediction of bladed disks coupled by friction joints, *Archives of Computational Methods in Engineering*, vol. 24, no. 3, pp. 589–636, 2017.

[88] M. Jahn, M. Stender, S. Tatzko, N. Hoffmann, A. Grolet, and J. Wallaschek, The extended periodic motion concept for fast limit cycle detection of self-excited systems, *Computers & Structures*, pp. 106–139, 2019.

[89] D. Dudkowski, S. Jafari, T. Kapitaniak, N. V. Kuznetsov, G. A. Leonov, and A. Prasad, Hidden attractors in dynamical systems, *Physics Reports*, vol. 637, pp. 1–50, 2016.

[90] H. E. Nusse, J. A. Yorke, and E. J. Kostelich, Basins of attraction, in *Dynamics: Numerical Explorations* (H. E. Nusse, J. A. Yorke, and E. J. Kostelich, eds.), vol. 101 of *Applied Mathematical Sciences*, pp. 269–314, New York, NY: Springer, 1994.

[91] A. Pikovsky, M. Rosenblum, J. Kurths, and R. C. Hilborn, Synchronization: A universal concept in nonlinear science, *American Journal of Physics*, vol. 70, no. 6, p. 655, 2002.

[92] Y. Kuramoto and D. Battogtokh, Coexistence of coherence and incoherence in nonlocally coupled phase oscillators, *Nonlinear Phenom Complex Syst*, vol. 5, pp. 380–385, 2002.

[93] P. W. Anderson, Absence of diffusion in certain random lattices, *Physical Review*, vol. 109, no. 5, pp. 1492–1505, 1958.

[94] F. Fontanela, A. Grolet, L. Salles, A. Chabchoub, and N. Hoffmann, Dark solitons, modulation instability and breathers in a chain of weakly nonlinear oscillators with cyclic symmetry, *Journal of Sound and Vibration*, vol. 413, pp. 467–481, 2018.

[95] D. M. Abrams and S. H. Strogatz, Chimera states for coupled oscillators, *Physical review letters*, vol. 93, no. 17, p. 102, 2004.

[96] D. Avitabile, D. J. B. Lloyd, J. Burke, E. Knobloch, and B. Sandstede, To snake or not to snake in the planar swift–hohenberg equation, *SIAM Journal on Applied Dynamical Systems*, vol. 9, no. 3, pp. 704–733, 2010.

[97] C. Beaume, A. Bergeon, and E. Knobloch, Homoclinic snaking of localized states in doubly diffusive convection, *Physics of Fluids*, vol. 23, no. 9, pp. 94–102, 2011.

[98] A. R. Champneys, Homoclinic orbits in reversible systems and their applications in mechanics, fluids and optics, *Physica D: Nonlinear Phenomena*, vol. 112, no. 1-2, pp. 158–186, 1998.

[99] O. BATISTE, E. Knobloch, A. ALONSO, and I. MERCADER, Spatially localized binary-fluid convection, *Journal of Fluid Mechanics*, vol. 560, p. 149, 2006.

[100] A. F. Vakakis, L. I. Manevitch, Y. V. Mlkhlin, V. N. Pilipchuk, and A. A. Zevin, *Normal Modes and Localization in Nonlinear Systems*. Wiley, 1996.

[101] A. F. Vakakis, *Normal Modes and Localization in Nonlinear Systems*. Dordrecht: Springer Netherlands, 2001.

[102] Rosenblum, Pikovsky, and Kurths, Phase synchronization of chaotic oscillators, *Physical review letters*, vol. 76, no. 11, pp. 1804–1807, 1996.

[103] A. Agarwal, N. Marwan, M. Rathinasamy, B. Merz, and J. Kurths, Multi-scale event synchronization analysis for unravelling climate processes: A wavelet-based approach, *Nonlinear Processes in Geophysics Discussions*, pp. 1–20, 2017.

[104] L. D. Avendaño-Valencia and S. D. Fassois, Damage/fault diagnosis in an operating wind turbine under uncertainty via a vibration response gaussian mixture random coefficient model based framework, *Mechanical Systems and Signal Processing*, vol. 91, pp. 326–353, 2017.

[105] P. S. Addison, Wavelet transforms and the ecg: A review, *Physiological measurement*, vol. 26, no. 5, pp. 99–155, 2005.

[106] F.-Y. Ouyang, B. Zheng, and X.-F. Jiang, Intrinsic multi-scale dynamic behaviors of complex financial systems, *PloS one*, vol. 10, no. 10, p. 420, 2015.

[107] N. E. Huang and Z. Wu, A review on hilbert-huang transform: Method and its applications to geophysical studies, *Reviews of Geophysics*, vol. 46, no. 2, p. 705, 2008.

[108] A. Agarwal, R. Maheswaran, N. Marwan, L. Caesar, and J. Kurths, Wavelet-based multiscale similarity measure for complex networks, *The European Physical Journal B*, vol. 91, no. 11, p. 157, 2018.

[109] D. Looney, A. Hemakom, and D. P. Mandic, Intrinsic multi-scale analysis: A multi-variate empirical mode decomposition framework, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 471, no. 2173, p. 709, 2015.

[110] V. P. Nguyen, Martijn Stroeven, and L. J. Sluys, Multiscale continuous and discontinuous modelling of heterogeneous materials: A review on recent developments, *Journal of Multiscale Modelling*, vol. 3, no. 4, pp. 229–270, 2011.

[111] J. Armand, L. Pesaresi, L. Salles, and C. W. Schwingshackl, A multiscale approach for nonlinear dynamic response predictions with fretting wear, *Journal of Engineering for Gas Turbines and Power*, vol. 139, no. 2, p. 916, 2017.

[112] A. I. Vakis, V. A. Yastrebov, J. Scheibert, L. Nicola, D. Dini, C. Minfray, A. Almqvist, M. Paggi, S. Lee, G. Limbert, J. F. Molinari, G. Anciaux, R. Aghababaei, S. Echeverri Restrepo, A. Papangelo, A. Cammarata, P. Nicolini, C. Putignano, G. Carbone, S. Stupkiewicz, J. Lengiewicz, G. Costagliola, F. Bosia, R. Guarino, N. M. Pugno, M. H. Müser, and M. Ciavarella, Modeling and simulation in tribology across scales: An overview, *Tribology International*, vol. 125, pp. 169–199, 2018.

[113] I. A. Sever, Nonlinear vibration phenomena in aero-engine measurements, in *Proceedings of the 34th IMAC, a Conference and Exposition on Structural Dynamics 2016* (M. Allen, R. L. Mayes, and D. Rixen, eds.), Conference Proceedings of the Society for Experimental Mechanics Series, pp. 241–252, Cham: Springer, 2016.

[114] S. Jain and P. Tiso, Model order reduction for temperature-dependent nonlinear mechanical systems: A multiple scales approach, *Journal of Sound and Vibration*, vol. 465, p. 115, 2020.

[115] M. Mitchell, *Complexity: A guided tour*. Oxford and New York: Oxford University Press, 2009.

[116] W. Weaver, Science and complexity, in *Facets of Systems Science* (G. J. Klir, ed.), International Federation for Systems Research International Series on Systems Science and Engineering, pp. 449–456, Boston, MA and s.l.: Springer US, 1991.

[117] S. Lloyd, Measures of complexity: a nonexhaustive list, *IEEE Control Systems*, vol. 21, no. 4, pp. 7–8, 2001.

[118] B. Goswami, N. Boers, A. Rheinwalt, N. Marwan, J. Heitzig, S. F. M. Breitenbach, and J. Kurths, Abrupt transitions in time series with uncertainties, *Nature communications*, vol. 9, no. 1, p. 48, 2018.

[119] E. Bradley, M. Easley, and R. Stolle, Reasoning about nonlinear system identification, *Artificial Intelligence*, vol. 133, no. 1-2, pp. 139–188, 2001.

[120] T. Söderström and P. Stoica, *System identification*. Prentice-Hall International series in systems and control engineering, New York: Prentice-Hall, 1989.

[121] D. J. Ewins, *Modal testing: Theory and practice*, vol. 15. Research studies press Letchworth, 1984.

[122] W. Heylen, S. Lammens, and P. Sas, *Modal analysis theory and testing*. Heverlee, Belgium: Katholieke Univ. Leuven Dep. Werktuigkunde, 1997.

[123] D. E. Adams and R. J. Allemang, Survey of nonlinear detection and identification techniques for experimental vibrations, in *Proceedings of the International Conference on Noise and Vibration Engineering*, pp. 269–281, 1998.

[124] P. Hagedorn and J. Wallaschek, On equivalent harmonic and stochastic linearization for nonlinear shock-absorbers, in *Nonlinear Stochastic Dynamic Engineering Systems* (F. Ziegler and G. I. Schuëller, eds.), vol. 48 of *IUTAM Symposium, International Union of Theoretical and Applied Mechanics*, pp. 23–32, Berlin and Heidelberg: Springer, 1988.

[125] S. F. Masri, H. Sassi, and T. K. Caughey, Nonparametric identification of nearly arbitrary nonlinear systems, *Journal of Applied Mechanics*, vol. 49, no. 3, pp. 619–628, 1982.

[126] G. Box, Box and jenkins: Time series analysis, forecasting and control, in *Very British Affair* (T. C. Mills, ed.), pp. 161–215, London: SPRINGER NATURE, 2013.

[127] I. J. Leontaritis and S. A. Billings, Input-output parametric models for nonlinear systems part i: deterministic non-linear systems, *International Journal of Control*, vol. 41, no. 2, pp. 303–328, 1985.

[128] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[129] N. E. Huang, M.-L. C. Wu, S. R. Long, S. S. P. Shen, W. Qu, P. Gloersen, and K. L. Fan, A confidence limit for the empirical mode decomposition and hilbert spectral analysis, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 459, no. 2037, pp. 2317–2345, 2003.

[130] J. S. Bendat, *Nonlinear system techniques and applications*. New York: Wiley, 1997.

[131] J. K. Hammond and P. R. White, The analysis of non-stationary signals using time-frequency methods, *Journal of Sound and Vibration*, vol. 190, no. 3, pp. 419–447, 1996.

[132] S. L. Brunton, J. L. Proctor, and J. N. Kutz, 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*, vol. 113, no. 15, pp. 3932–3937, 2016.

[133] W.-X. Wang, R. Yang, Y.-C. Lai, V. Kovanis, and C. Grebogi, Predicting catastrophes in nonlinear dynamical systems by compressive sensing, *Physical review letters*, vol. 106, no. 15, p. 154101, 2011.

[134] S. L. Brunton, B. W. Brunton, J. L. Proctor, E. Kaiser, and J. N. Kutz, Chaos as an intermittently forced linear system, *Nature communications*, vol. 8, no. 1, p. 19, 2017.

[135] J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Deylon, P.-Y. Glorennec, H. Hjalmarsson, and A. Juditsky, *Nonlinear black-box modeling in system identification: a unified overview*. Linköping University, 1995.

[136] A. Juditsky, H. Hjalmarsson, A. Benveniste, B. Delyon, L. Ljung, J. Sjöberg, and Q. Zhang, Nonlinear black-box models in system identification: Mathematical foundations, *Automatica*, vol. 31, no. 12, pp. 1725–1750, 1995.

[137] S. Chen, S. A. Billings, and P. M. Grant, Non-linear system identification using neural networks, *International Journal of Control*, vol. 51, no. 6, pp. 1191–1214, 1990.

[138] S. A. Billings, H. B. Jamaluddin, and S. Chen, Properties of neural networks with applications to modelling non-linear dynamical systems, *International Journal of Control*, vol. 55, no. 1, pp. 193–224, 1992.

[139] A. R. AbuBakar and H. Ouyang, Complex eigenvalue analysis and dynamic transient analysis in predicting disc brake squeal, *International Journal of Vehicle Noise and Vibration*, vol. 2, no. 2, p. 143, 2006.

[140] F. Massi, L. Baillet, O. Giannini, and A. Sestieri, Brake squeal: Linear and nonlinear numerical approaches, *Mechanical Systems and Signal Processing*, vol. 21, no. 6, pp. 2374–2393, 2007.

[141] D. Laxalde and F. Thouverez, Complex non-linear modal analysis for mechanical systems: Application to turbomachinery bladings with friction interfaces, *Journal of Sound and Vibration*, vol. 322, no. 4-5, pp. 1009–1025, 2009.

[142] N. Hoffmann and L. Gaul, Effects of damping on mode-coupling instability in friction induced oscillations, *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, vol. 83, no. 8, pp. 524–534, 2003.

[143] O. N. Kirillov and F. Verhulst, Paradoxes of dissipation-induced destabilization or who opened whitney's umbrella?, *ZAMM*, vol. 90, no. 6, pp. 462–488, 2010.

[144] F. Cantone and F. Massi, A numerical investigation into the squeal instability: Effect of damping, *Mechanical Systems and Signal Processing*, vol. 25, no. 5, pp. 1727–1737, 2011.

[145] T. Butlin and J. Woodhouse, Sensitivity of friction-induced vibration in idealised systems, *Journal of Sound and Vibration*, vol. 319, no. 1-2, pp. 182–198, 2009.

[146] H. Ouyang, W. Nack, Y. Yuan, and F. Chen, On automotive disc brake squeal part ii: Simulation and analysis, in *SAE Technical Paper Series*, SAE Technical Paper Series, SAE International, Warrendale, PA, United States, 2003.

[147] A. Nobari, H. Ouyang, and P. Bannister, Statistics of complex eigenvalues in friction-induced vibration, *Journal of Sound and Vibration*, vol. 338, pp. 169–183, 2015.

[148] K.-J. Bathe and E. L. Wilson, Numerical methods in finite element analysis, *Prentice-Hall*, 1976.

[149] L. Gaul and R. Nitsche, The role of friction in mechanical joints, *Applied Mechanics Reviews*, vol. 54, no. 2, p. 93, 2001.

[150] M. Stender, M. Tiedemann, and N. Hoffmann, Characterization of complex states for friction-excited systems, *PAMM*, vol. 17, no. 1, pp. 45–46, 2017.

[151] U. von Wagner, D. Hochlenert, and P. Hagedorn, Minimal models for disk brake squeal, *Journal of Sound and Vibration*, vol. 302, no. 3, pp. 527–539, 2007.

[152] M. Stender, M. Tiedemann, N. Hoffmann, and S. Oberst, Impact of an irregular friction formulation on dynamics of a minimal model for brake squeal, *Mechanical Systems and Signal Processing*, vol. 107, pp. 439–451, 2018.

[153] M. Stender, M. Jahn, N. Hoffmann, and J. Wallaschek, Hyperchaos co-existing with periodic orbits in a frictional oscillator, *Journal of Sound and Vibration*, vol. 472, pp. 115–203, 2020.

[154] A. A. Andronov, *Theory of bifurcations of dynamic systems on a plane*. A Halsted Press book, 1973.

[155] M. Krack, Nonlinear modal analysis of nonconservative systems: Extension of the periodic motion concept, *Computers & Structures*, vol. 154, pp. 59–71, 2015.

[156] O. E. Rössler, An equation for continuous chaos, *Physics Letters A*, vol. 57, no. 5, pp. 397–398, 1976.

[157] Y.-C. Lai and T. Tél, *Transient Chaos: Complex Dynamics on Finite Time Scales*, vol. 173 of *Applied Mathematical Sciences*. New York, NY: Springer Science+Business Media LLC, 2011.

[158] N. Gräbner, Analyse und Verbesserung der Simulationsmethode des Bremsenquietschens, *Technische Universität Berlin*, 2016.

[159] G.-P. Ostermeyer and K. Bode, On dynamic friction phenomena in brake systems, *Friction, Wear and Wear Protection*, pp. 301–306, 2008.

[160] M. Graf and G.-P. Ostermeyer, Friction-induced vibration and dynamic friction laws: Instability at positive friction–velocity-characteristic, *Tribology International*, vol. 92, pp. 255–258, 2015.

[161] M. A. Bucci, O. Semeraro, A. Allauzen, G. Wisniewski, L. Cordier, and L. Mathelin, Control of chaotic systems by deep reinforcement learning, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 475, no. 2231, p. 351, 2019.

[162] M. Stender, S. Oberst, M. Tiedemann, and N. Hoffmann, Complex machine dynamics: systematic recurrence quantification analysis of disk brake vibration data, *Nonlinear Dynamics*, vol. 267, no. 1, p. 105, 2019.

[163] C. Shearer, The crisp-dm model: the new blueprint for data mining, *Journal of data warehousing*, vol. 5, no. 4, pp. 13–22, 2000.

[164] M. T. Ribeiro, S. Singh, and C. Guestrin, Why should i trust you?, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (B. Krishnapuram, M. Shah, A. Smola, C. Aggarwal, D. Shen, and R. Rastogi, eds.), (New York, New York, USA), pp. 1135–1144, ACM Press, 2016.

[165] Z. C. Lipton, The mythos of model interpretability, *Queue*, vol. 16, no. 3, pp. 30:31–30:57, 2018.

[166] R. Roscher, B. Bohn, M. F. Duarte, and J. Garcke, Explainable machine learning for scientific insights and discoveries, *Preprint*, 2019.

[167] M. Stender, S. Oberst, and N. Hoffmann, Recovery of differential equations from impulse response time series data for model identification and feature extraction, *Vibration*, vol. 2, no. 1, pp. 25–46, 2019.

[168] S. V. Vaseghi, *Advanced digital signal processing and noise reduction*. Chichester, West Sussex: Wiley, 4th ed., 2008.

[169] J. Gao, H. Sultan, J. Hu, and W.-W. Tung, Denoising nonlinear time series by adaptive filtering and wavelet shrinkage: A comparison, *IEEE Signal Processing Letters*, vol. 17, no. 3, pp. 237–240, 2010.

[170] M. Han, Y. Liu, J. Xi, and W. Guo, Noise smoothing for nonlinear time series using wavelet soft threshold, *IEEE Signal Processing Letters*, vol. 14, no. 1, pp. 62–65, 2007.

[171] Kaplan and Glass, Direct test for determinism in a time series, *Physical review letters*, vol. 68, no. 4, pp. 427–430, 1992.

[172] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*. Cambridge: Cambridge University Press, 2003.

[173] H. D. I. Abarbanel, R. Brown, J. J. Sidorowich, and L. S. Tsimring, The analysis of observed chaotic data in physical systems, *Reviews of Modern Physics*, vol. 65, no. 4, pp. 1331–1392, 1993.

[174] Baron Fourier, Jean Baptiste Joseph, *The analytical theory of heat*. The University Press, 1878.

[175] D. Gabor, Theory of communication, *Journal of the Institution of Electrical Engineers - Part I: General*, vol. 94, no. 73, p. 58, 1947.

[176] S. G. Mallat, A theory for multiresolution signal decomposition: The wavelet representation, *IEEE transactions on pattern analysis and machine intelligence*, vol. 11, no. 7, pp. 674–693, 1989.

[177] S. Mallat, *A wavelet tour of signal processing*. San Diego, Calif.: Academic Press, 2. ed., 1999.

[178] E. M. Stein and R. Shakarchi, *Fourier Analysis: An Introduction*. New Jersey: Princeton University Press, 2011.

[179] S. L. Brunton and J. N. Kutz, *Data-Driven Science and Engineering*, vol. 237. Cambridge University Press, 2019.

[180] B. D. Fulcher and N. S. Jones, Highly comparative feature-based time-series classification, *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 12, pp. 3026–3037, 2014.

[181] N. Marwan, M. C. Romano, M. Thiel, and J. Kurths, Recurrence plots for the analysis of complex systems, *Physics Reports*, vol. 438, no. 5-6, pp. 237–329, 2007.

[182] N. Marwan, J. F. Donges, Y. Zou, R. V. Donner, and J. Kurths, Complex network approach for recurrence analysis of time series, *Physics Letters A*, vol. 373, no. 46, pp. 4246–4254, 2009.

[183] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. Doyne Farmer, Testing for nonlinearity in time series: The method of surrogate data, *Physica D: Nonlinear Phenomena*, vol. 58, no. 1-4, pp. 77–94, 1992.

[184] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, Determining lyapunov exponents from a time series, *Physica D: Nonlinear Phenomena*, vol. 16, no. 3, pp. 285–317, 1985.

[185] H. Kantz, A robust method to estimate the maximal lyapunov exponent of a time series, *Physics Letters A*, vol. 185, no. 1, pp. 77–87, 1994.

[186] T. Schreiber, Measuring information transfer, *Physical review letters*, vol. 85, no. 2, pp. 461–464, 2000.

[187] J. Theiler, Estimating fractal dimension, *Journal of the Optical Society of America A*, vol. 7, no. 6, p. 1055, 1990.

[188] D. S. Broomhead and G. P. King, Extracting qualitative dynamics from experimental data, *Physica D: Nonlinear Phenomena*, vol. 20, no. 2-3, pp. 217–236, 1986.

[189] R. Hegger, H. Kantz, and T. Schreiber, Practical implementation of nonlinear time series methods: The tisean package, *Chaos (Woodbury, N.Y.)*, vol. 9, no. 2, pp. 413–435, 1999.

[190] E. Bradley and H. Kantz, Nonlinear time-series analysis revisited, *Chaos (Woodbury, N.Y.)*, vol. 25, no. 9, p. 097610, 2015.

[191] J.-P. Eckmann and D. Ruelle, Fundamental limitations for estimating dimensions and lyapunov exponents in dynamical systems, *Physica D: Nonlinear Phenomena*, vol. 56, no. 2-3, pp. 185–187, 1992.

[192] J. C. Sprott, *Chaos and time-series analysis*. Oxford Univ. Press, 1. ed., 2003.

[193] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, Geometry from a time series, *Physical review letters*, vol. 45, no. 9, pp. 712–716, 1980.

[194] A. Provenzale, L. A. Smith, R. Vio, and G. Murante, Distinguishing between low-dimensional dynamics and randomness in measured time series, *Physica D: Nonlinear Phenomena*, vol. 58, no. 1-4, pp. 31–49, 1992.

[195] A. M. Fraser and H. L. Swinney, Independent coordinates for strange attractors from mutual information, *Physical Review A*, vol. 33, no. 2, pp. 1134–1140, 1986.

[196] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, Determining embedding dimension for phase-space reconstruction using a geometrical construction, *Physical Review A*, vol. 45, no. 6, pp. 3403–3411, 1992.

[197] M. Paluš and I. Dvořák, Singular-value decomposition in attractor reconstruction: Pitfalls and precautions, *Physica D: Nonlinear Phenomena*, vol. 55, no. 1-2, pp. 221–234, 1992.

[198] N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, and J. Kurths, Recurrence-plot-based measures of complexity and their application to heart-rate-variability data, *Physical review. E, Statistical, nonlinear, and soft matter physics*, vol. 66, no. 2.2, p. 702, 2002.

[199] J.-P. Eckmann, S. O. Kamphorst, and D. Ruelle, Recurrence plots of dynamical systems, *Europhysics Letters (EPL)*, vol. 4, no. 9, pp. 973–977, 1987.

[200] N. Marwan, S. Schinkel, and J. Kurths, Recurrence plots 25 years later —gaining confidence in dynamical transitions, *EPL (Europhysics Letters)*, vol. 101, no. 2, p. 20007, 2013.

[201] Y. LeCun, Y. Bengio, and G. Hinton, Deep learning, *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[202] L. Breiman, Random forests, *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[203] T. G. Dietterich, An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization, *Machine learning*, vol. 40, no. 2, pp. 139–157, 2000.

[204] F. Rosenblatt, The perceptron: a probabilistic model for information storage and organization in the brain, *Psychological review*, vol. 65, no. 6, pp. 386–408, 1958.

[205] Y. LeCun, Bernhard E. Boser, John S. Denker, Donnie Henderson, R. E. Howard, Wayne E. Hubbard, and Lawrence D. Jackel, Handwritten digit recognition with a back-propagation network, in *Advances in Neural Information Processing Systems 2* (D. S. Touretzky, ed.), pp. 396–404, Morgan-Kaufmann, 1990.

[206] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Learning representations by back-propagating errors, *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[207] L. Bottou, Large-scale machine learning with stochastic gradient descent, in *Proceedings of COMPSTAT'2010* (Y. Lechevallier and G. Saporta, eds.), vol. 10, pp. 177–186, Heidelberg: Springer-Verlag Berlin Heidelberg, 2010.

[208] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, *Preprint*, 2014.

[209] S. Russell, P. Norvig, and A. Intelligence, Artificial intelligence: A modern approach, *Artificial Intelligence. Prentice-Hall, Egnlewood Cliffs*, vol. 25, p. 27, 1995.

[210] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning*, vol. 1. Springer series in statistics New York, 2001.

[211] C. M. Bishop, *Pattern recognition and machine learning*. Information science and statistics, New York NY: Springer, 2006.

[212] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An introduction to statistical learning*, vol. 112. Springer, 2013.

[213] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[214] M. T. Ribeiro, S. Singh, and C. Guestrin, Model-agnostic interpretability of machine learning, *Preprint*, 2016.

[215] S. M. Lundberg and S.-I. Lee, A unified approach to interpreting model predictions, in *Advances in Neural Information Processing Systems 30* (I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, eds.), pp. 4765–4774, Curran Associates, Inc, 2017.

[216] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, Learning deep features for discriminative localization, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2921–2929, 2016.

[217] I. Guyon and A. Elisseeff, An introduction to variable and feature selection, *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, 2003.

[218] Y. Saeys, I. Inza, and P. Larrañaga, A review of feature selection techniques in bioinformatics, *Bioinformatics (Oxford, England)*, vol. 23, no. 19, pp. 2507–2517, 2007.

[219] Y. Benjamini, D. Yekutieli, *et al.*, The control of the false discovery rate in multiple testing under dependency, *The Annals of Statistics*, vol. 29, no. 4, pp. 1165–1188, 2001.

[220] M. Christ, N. Braun, J. Neuffer, and A. W. Kempa-Liehr, Time series feature extraction on basis of scalable hypothesis tests (tsfresh – a python package), *Neurocomputing*, vol. 307, pp. 72–77, 2018.

[221] I. A. Gheyas and L. S. Smith, Feature subset selection in large dimensionality domains, *Pattern Recognition*, vol. 43, no. 1, pp. 5–13, 2010.

[222] R. Caruana and D. Freitag, Greedy attribute selection, in *Machine Learning Proceedings 1994*, pp. 28–36, Elsevier, 1994.

[223] L. Kellner, M. Stender, R. U. F. von Bock und Polach, H. Herrnring, S. Ehlers, N. Hoffmann, and K. V. Høyland, Establishing a common database of ice experiments and using machine learning to understand and predict ice behavior, *Cold Regions Science and Technology*, vol. 162, pp. 56–73, 2019.

[224] B. D. Fulcher and N. S. Jones, hctsa: A computational framework for automated time-series phenotyping using massive feature extraction, *Cell systems*, vol. 5, no. 5, pp. 527–531.e3, 2017.

[225] I. Nun, P. Protopapas, B. Sim, M. Zhu, R. Dave, N. Castro, and K. Pichara, Fats: Feature analysis for time series, *Preprint*, 2015.

[226] A. Bagnall, J. Lines, J. Hills, and A. Bostrom, Time-series classification with cote: The collective of transformation-based ensembles, *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 9, pp. 2522–2535, 2015.

[227] X. Jin, B. W. Wah, X. Cheng, and Y. Wang, Significance and challenges of big data research, *Big Data Research*, vol. 2, no. 2, pp. 59–64, 2015.

[228] S. Daskalaki, I. Kopanas, and N. Avouris, Evaluation of classifiers for an uneven distribution problem, *Applied Artificial Intelligence*, vol. 20, no. 5, pp. 381–417, 2006.

[229] Y. SUN, A. Wong, and M. Kamel, Classificationof imbalanced data: a rview: Classificationof imbalanced data: a rview, *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 23, no. 04, pp. 687–719, 2009.

[230] S. A. A. Menon, H. Narasimhan and S. Chawla, On the statistical consistency of algorithms for binary classification under class imbalance, in *Proceedings of the 30th International Conference on Machine Learning*, vol. 28 of *Proceedings of Machine Learning Research*, pp. 603–611, PMLR, 2013.

[231] G. M. Weiss and F. Provost, The effect of class distribution on classifier learning: an empirical study, *Technical Report, Department of Computer Science, Rutgers University*, 2001.

[232] Y. Sun, M. S. Kamel, A. K. Wong, and Y. Wang, Cost-sensitive boosting for classification of imbalanced data, *Pattern Recognition*, vol. 40, no. 12, pp. 3358–3378, 2007.

[233] P. K. Chan, W. Fan, A. L. Prodromidis, and S. J. Stolfo, Distributed data mining in credit card fraud detection, *IEEE Intelligent Systems*, vol. 14, no. 6, pp. 67–74, 1999.

[234] B. W. Matthews, Comparison of the predicted and observed secondary structure of t4 phage lysozyme, *Biochimica et Biophysica Acta (BBA) - Protein Structure*, vol. 405, no. 2, pp. 442–451, 1975.

[235] S. Boughorbel, F. Jarray, and M. El-Anbari, Optimal classifier for imbalanced data using matthews correlation coefficient metric, *PloS one*, vol. 12, no. 6, p. e0177678, 2017.

[236] Z. Feng, M. Liang, and F. Chu, Recent advances in time–frequency analysis methods for machinery fault diagnosis: A review with application examples, *Mechanical Systems and Signal Processing*, vol. 38, no. 1, pp. 165–205, 2013.

[237] G. Litak, M. Iwaniec, J. Iwaniec, M. Borys, Z. Czyż, K. Falkowicz, J. Kujawska, M. Kulisz, and M. Szala, Milling gate vibrations analysis via hilbert-huang transform, *ITM Web of Conferences*, vol. 15, p. 04008, 2017.

[238] A. Syta, J. Jonak, Ł. Jedliński, and G. Litak, Failure diagnosis of a gear box by recurrences, *Journal of Vibration and Acoustics*, vol. 134, no. 4, p. 87, 2012.

[239] Y. Qu, D. He, J. Yoon, B. van Hecke, E. Bechhoefer, and J. Zhu, Gearbox tooth cut fault diagnostics using acoustic emission and vibration sensors–a comparative study, *Sensors (Basel, Switzerland)*, vol. 14, no. 1, pp. 1372–1393, 2014.

[240] M. Krishnan, R. Jin, I. A. Sever, and P. A. Tarazaga, Data based modeling of aero engine vibration responses, in *Sensors and instrumentation, aircraft/aerospace, energy harvesting & dynamic environments testing* (C. Walber, P. Walter, and S. Seidlitz, eds.), vol. 58 of *Conference Proceedings of the Society for Experimental Mechanics Series*, pp. 365–368, [S.l.]: Springer nature, 2019.

[241] J. M. Nichols, C. J. Nichols, M. D. Todd, M. Seaver, S. T. Trickey, and L. N. Virgin, Use of data-driven phase space models in assessing the strength of a bolted connection in a composite beam, *Smart Materials and Structures*, vol. 13, no. 2, p. 241, 2004.

[242] B. I. Epureanu and A. Hashmi, Parameter reconstruction based on sensitivity vector fields, *Journal of Vibration and Acoustics*, vol. 128, no. 6, pp. 732–740, 2006.

[243] M. Didonna, M. Stender, A. Papangelo, F. Fontanela, M. Ciavarella, and N. Hoffmann, Reconstruction of governing equations from vibration measurements for geometrically nonlinear systems, *Lubricants*, vol. 7, no. 8, p. 64, 2019.

[244] M. Raissi and G. E. Karniadakis, Hidden physics models: Machine learning of nonlinear partial differential equations, *Journal of Computational Physics*, vol. 357, pp. 125–141, 2018.

[245] S. A. Billings, Identification of nonlinear systems–a survey, in *IEE Proceedings D (Control Theory and Applications)*, vol. 127, pp. 272–285, 1980.

[246] J. P. Crutchfield and B. S. McNamara, Equations of motion from a data series, *Complex systems*, vol. 1, no. 417-452, p. 121, 1987.

[247] T. Eisenhammer, A. Hübler, N. Packard, and J. A. Kelso, Modeling experimental time series with ordinary differential equations, *Biological cybernetics*, vol. 65, no. 2, pp. 107–112, 1991.

[248] J. Bongard and H. Lipson, Automated reverse engineering of nonlinear dynamical systems, *Proceedings of the National Academy of Sciences*, vol. 104, no. 24, pp. 9943–9948, 2007.

[249] M. Schmidt and H. Lipson, Distilling free-form natural laws from experimental data, *Science (New York, N.Y.)*, vol. 324, no. 5923, pp. 81–85, 2009.

[250] B. C. Daniels and I. Nemenman, Automated adaptive inference of phenomenological dynamical models, *Nature communications*, vol. 6, p. 8133, 2015.

[251] S. H. Rudy, S. L. Brunton, J. L. Proctor, and J. N. Kutz, Data-driven discovery of partial differential equations, *Science advances*, vol. 3, no. 4, p. 614, 2017.

[252] H. Schaeffer, Learning partial differential equations via data discovery and sparse optimization, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 473, no. 2197, p. 20160446, 2017.

[253] N. M. Mangan, S. L. Brunton, J. L. Proctor, and J. N. Kutz, Inferring biological networks by sparse identification of nonlinear dynamics, *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, vol. 2, no. 1, pp. 52–63, 2016.

[254] B. d. Silva, D. M. Higdon, S. L. Brunton, and J. N. Kutz, Discovery of physics from data: Universal laws and discrepancy models, *Preprint*, 2019.

[255] P. Zheng, T. Askham, S. L. Brunton, J. N. Kutz, and A. Y. Aravkin, A unified framework for sparse relaxed regularized regression: Sr3, *IEEE Access*, vol. 7, pp. 1404–1423, 2019.

[256] R. Chartrand, Numerical differentiation of noisy, nonsmooth data, *ISRN Applied Mathematics*, vol. 2011, no. 1–4, pp. 1–11, 2011.

[257] M. Stender, M. Di Bartolomeo, F. Massi, and N. Hoffmann, Revealing transitions in friction-excited vibrations by nonlinear time-series analysis, *Nonlinear Dynamics*, vol. 47, no. 7, p. 209, 2019.

[258] A. Lazzari, D. Tonazzi, G. Conidi, C. Malmassari, A. Cerutti, and F. Massi, Experimental evaluation of brake pad material propensity to stick-slip and groan noise emission, *Lubricants*, vol. 6, no. 4, p. 107, 2018.

[259] M. Stender, M. Tiedemann, D. Spieler, D. Schoepflin, N. Hoffmann, and S. Oberst, Deep learning for brake squeal: Brake noise detection, characterization and prediction, *Mechanical Systems and Signal Processing*, vol. 149, p. 107181, 2021.

[260] A. Akay, Acoustics of friction, *The Journal of the Acoustical Society of America*, vol. 111, no. 4, pp. 1525–1548, 2002.

[261] SAE International, Disc and drum brake dynamometer squeal noise test procedure j2521, *SAE International*, 2012.

[262] G. Mauer and M. Haverkamp, Measurement and assessment of noise caused by vehicle brake systems, in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, vol. 2007, pp. 2488–2497, 2007.

[263] Verband der Automobilindustrie, VDA 305: Datenaustauschformat, *Verband der Automobilindustrie, Berlin*, 2013.

[264] P. Viola and M. Jones, Rapid object detection using a boosted cascade of simple features, in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, pp. I–511–I–518, IEEE Comput. Soc, 8-14 Dec. 2001.

[265] N. Dalal and B. Triggs, Histograms of oriented gradients for human detection, in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, pp. 886–893, IEEE, 20-26 June 2005.

[266] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, Object detection with discriminatively trained part-based models, *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 9, pp. 1627–1645, 2010.

[267] P. Dollár, R. Appel, S. Belongie, and P. Perona, Fast feature pyramids for object detection, *IEEE transactions on pattern analysis and machine intelligence*, vol. 36, no. 8, pp. 1532–1545, 2014.

[268] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, Imagenet large scale visual recognition challenge, *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.

[269] Alex Krizhevsky, I. Sutskever, and G. E. Hinton, Imagenet classification with deep convolutional neural networks, in *Advances in Neural Information Processing Systems 25* (F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, eds.), pp. 1097–1105, Curran Associates, Inc, 2012.

[270] M. Everingham, S. M. A. Eslami, L. van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, The pascal visual object classes challenge: A retrospective, *International Journal of Computer Vision*, vol. 111, no. 1, pp. 98–136, 2015.

[271] T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick, and P. Dollár, Microsoft COCO: Common Objects in Context, *Preprint*, 2014.

[272] Z.-Q. Zhao, P. Zheng, S.-T. Xu, and X. Wu, Object detection with deep learning: A review, *IEEE transactions on neural networks and learning systems*, 2019.

[273] R. Girshick, J. Donahue, T. Darrell, and J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.

[274] R. Girshick, Fast R-CNN, in *The IEEE International Conference on Computer Vision (ICCV)*, 2015.

[275] S. Ren, K. He, R. Girshick, and J. Sun, Faster R-CNN: Towards real-time object detection with region proposal networks, in *Advances in Neural Information Processing Systems 28* (C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, eds.), pp. 91–99, Curran Associates, Inc, 2015.

[276] J. Dai, Y. Li, K. He, and J. Sun, R-fcn: Object detection via region-based fully convolutional networks, in *Advances in Neural Information Processing Systems 29* (D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, eds.), pp. 379–387, Curran Associates, Inc, 2016.

[277] M. Najibi, M. Rastegari, and L. S. Davis, G-CNN: An iterative grid based object detector, in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[278] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, You only look once: Unified, real-time object detection, in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[279] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, Ssd: Single shot multibox detector, in *Computer Vision – ECCV 2016* (B. Leibe, J. Matas, N. Sebe, and M. Welling, eds.), vol. 9905 of *Lecture Notes in Computer Science*, pp. 21–37, Cham: Springer International Publishing, 2016.

[280] M. Stender, M. Tiedemann, L. Hoffmann, and N. Hoffmann, Determining growth rates of instabilities from time-series vibration data: Methods and applications for brake squeal, *Mechanical Systems and Signal Processing*, vol. 129, pp. 250–264, 2019.

[281] L. Hoffmann, Calculation of growth rates for self-excited dynamical systems on the example of brake squeal noise, *Technische Universität Hamburg, Arbeitsgruppe Strukturdynamik, Projektarbeit*, 2018.

[282] Y. Dai and T. C. Lim, Suppression of brake squeal noise applying finite element brake and pad model enhanced by spectral-based assurance criteria, *Applied Acoustics*, vol. 69, no. 3, pp. 196–214, 2008.

[283] T. Butlin and J. Woodhouse, Friction-induced vibration: Quantifying sensitivity and uncertainty, *Journal of Sound and Vibration*, vol. 329, no. 5, pp. 509–526, 2010.

[284] J. Gao and H. Cai, On the structures and quantification of recurrence plots, *Physics Letters A*, vol. 270, no. 1-2, pp. 75–87, 2000.

[285] K. H. Kraemer, R. V. Donner, J. Heitzig, and N. Marwan, Recurrence threshold selection for obtaining robust recurrence characteristics in different embedding dimensions, *Chaos (Woodbury, N.Y.)*, vol. 28, no. 8, p. 085720, 2018.

[286] B. Wernitz and N. Hoffmann, New approaches to signal analysis of friction noise and vibration, in *Proceedings of the IMechE international conference of braking, York*, 2009.

[287] A. Akay, O. Giannini, F. Massi, and A. Sestieri, Disc brake squeal characterization through simplified test rigs, *Mechanical Systems and Signal Processing*, vol. 23, no. 8, pp. 2590–2607, 2009.

[288] D. Schoepflin, Development of a virtual twin to predict brake system vibrations using recurrent neural networks, *Technische Universität Hamburg, Arbeitsgruppe Strukturdynamik, Master Thesis*, 2019.

[289] G. X. Chen, J. Z. Lv, Q. Zhu, Y. He, and X. B. Xiao, Effect of the braking pressure variation on disc brake squeal of a railway vehicle: Test measurement and finite element analysis, *Wear*, vol. 426-427, pp. 1788–1796, 2019.

[290] A. Bagnall, J. Lines, A. Bostrom, J. Large, and E. Keogh, The great time series classification bake off: A review and experimental evaluation of recent algorithmic advances, *Data Mining and Knowledge Discovery*, vol. 31, no. 3, pp. 606–660, 2017.

[291] M. Eriksson, F. Bergman, and S. Jacobson, Surface characterisation of brake pads after running under silent and squealing conditions, *Wear*, vol. 232, no. 2, pp. 163–167, 1999.

[292] H. Hetzler and K. Willner, On the influence of contact tribology on brake squeal, *Tribology International*, vol. 46, no. 1, pp. 237–246, 2012.

[293] A. Lazzari, D. Tonazzi, and F. Massi, Squeal propensity characterization of brake lining materials through friction noise measurements, *Mechanical Systems and Signal Processing*, vol. 128, pp. 216–228, 2019.

[294] A. M. Lazim, M. Kchaou, M. A. Hamid, and A. A. Bakar, Squealing characteristics of worn brake pads due to silica sand embedment into their friction layers, *Wear*, vol. 358-359, pp. 123–136, 2016.

[295] P. J. Werbos, Backpropagation through time: what it does and how to do it, *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1550–1560, 1990.

[296] S. Hochreiter and J. Schmidhuber, Long short-term memory, *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[297] R. Pascanu, T. Mikolov, and Y. Bengio, On the difficulty of training recurrent neural networks, in *International conference on machine learning*, pp. 1310–1318, 2013.

[298] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, Empirical evaluation of gated recurrent neural networks on sequence modeling, *Preprint*, 2014.

[299] A. Dhooge, W. Govaerts, and Y. A. Kuznetsov, Matcont, *ACM Transactions on Mathematical Software*, vol. 29, no. 2, pp. 141–164, 2003.

[300] R. L. Devaney and J.-P. Eckmann, An introduction to chaotic dynamical systems, *Physics Today*, vol. 40, no. 7, p. 72, 1987.

[301] H.-O. Peitgen, H. Jürgens, and D. Saupe, *Chaos and Fractals*. New York, NY: Springer New York, 2004.

[302] C. Ross, M. Odell, and S. Cremer, The shadow-curves of the orbit diagram permeate the bifurcation diagram, too, *International Journal of Bifurcation and Chaos*, vol. 19, no. 09, pp. 3017–3031, 2009.

[303] N. Marwan, How to avoid potential pitfalls in recurrence plot based data analysis, *International Journal of Bifurcation and Chaos*, vol. 21, no. 04, pp. 1003–1017, 2011.

[304] S. Kullback and R. A. Leibler, On information and sufficiency, *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, 1951.

A Analytical oscillator models

The general form of the type of dynamical systems covered in this thesis read

$$\boldsymbol{M}\ddot{\boldsymbol{x}} + (\boldsymbol{C}+\boldsymbol{G})\,\dot{\boldsymbol{x}} + (\boldsymbol{K}+\boldsymbol{N})\,\boldsymbol{x} + \boldsymbol{f}_{\mathrm{nl}} + \boldsymbol{f}_{\mathrm{ext}} = \boldsymbol{0} \tag{A.1}$$

where $\boldsymbol{M}$ denotes the positive mass matrix, $\boldsymbol{C}$ the symmetric damping terms, $\boldsymbol{G}$ the anti-symmetric gyroscopic terms, $\boldsymbol{K}$ the symmetric stiffness terms, $\boldsymbol{N}$ the anti-symmetric circulatory terms, $\boldsymbol{f}_{\mathrm{nl}}$ nonlinear forces and $\boldsymbol{f}_{\mathrm{ext}}$ external forces. As the minimal model studies are intended to illustrate qualitative dynamical behavior, variable units are discarded for system states and parameter values. All bifurcation diagrams were created using the Matlab continuation toolbox MATCONT [299] that employs orthogonal collocation schemes.

A.1 Bi-stable single-degree-of-freedom oscillator

The bi-stability of the frictional oscillator is studied closely following the work by Papangelo et al. [80]. The single-degree-of-freedom system $M\ddot{x} + c\dot{x} + kx = F$, see Figure 2.2 (a), is given by the equations of motion

$$\begin{aligned} F &= -N\mu\left(v_{\mathrm{rel}}\right)\mathrm{sign}\left(v_{\mathrm{rel}}\right), \quad v_{\mathrm{rel}} \neq 0, \quad v_{\mathrm{rel}} = \dot{x} - v \\ |F| &< \mu_{\mathrm{st}} N, \qquad\qquad\qquad v_{\mathrm{rel}} = 0 \pm \xi, \\ \mu(v_{\mathrm{rel}}) &= \mu_{\mathrm{sl}} + \left(\mu_{\mathrm{st}} - \mu_{\mathrm{sl}}\right)\exp\left(-\frac{\left|v_{\mathrm{rel}}\right|}{v_0}\right). \end{aligned} \tag{A.2}$$

The system experiences friction-induced vibrations (FIV) owing to the friction formulation that features a velocity-dependent weakening behavior. The equations of motion are non-dimensionalized according to [80] and the parameters read $\mu_{\mathrm{sl}} = 0.5$, $\mu_{\mathrm{st}} = 1$, $\xi = 0.05$, $N = 1$ and $\tilde{v}_0 = 0.5$.

A.2 3DOF minimal model with cubic stiffness

The friction oscillator with three degrees of freedom x, y, z is described by the state vector $\boldsymbol{x} = \left[x, y, z, \dot{x}, \dot{y}, \dot{z}\right]^{\top}$ and the system definition follows the set of equations of

motion in first-order form according to

$$
\begin{aligned}
\boldsymbol{C}+\boldsymbol{G} &= \begin{bmatrix} c_\mathrm{x}+\frac{c_\mathrm{lin}}{2} & \frac{c_\mathrm{lin}}{2} & -\frac{c_\mathrm{lin}}{\sqrt{2}} \\ \frac{c_\mathrm{lin}}{2} & c_\mathrm{y}+\frac{c_\mathrm{lin}}{2} & -\frac{c_\mathrm{lin}}{\sqrt{2}} \\ -\frac{c_\mathrm{lin}}{\sqrt{2}} & -\frac{c_\mathrm{lin}}{\sqrt{2}} & c_\mathrm{z}+c_\mathrm{lin} \end{bmatrix}, \\
\boldsymbol{K}+\boldsymbol{N} &= \begin{bmatrix} k_\mathrm{x}+\frac{k_\mathrm{lin}}{2} & -\mu k_\mathrm{y}+\frac{k_\mathrm{lin}}{2} & -\frac{k_\mathrm{lin}}{\sqrt{2}} \\ \frac{k_\mathrm{lin}}{2} & k_\mathrm{y}+\frac{k_\mathrm{lin}}{2} & -\frac{k_\mathrm{lin}}{\sqrt{2}} \\ -\frac{k_\mathrm{lin}}{\sqrt{2}} & -\frac{k_\mathrm{lin}}{\sqrt{2}} & k_\mathrm{z}+k_\mathrm{lin} \end{bmatrix}, \\
\boldsymbol{M} &= \begin{bmatrix} M & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & M \end{bmatrix}, \\
\boldsymbol{f}_\mathrm{nl} &= \begin{bmatrix} {}^{-1}/_{\sqrt{2}} \\ {}^{-1}/_{\sqrt{2}} \\ 1 \end{bmatrix} k_\mathrm{nl} \cdot \left(z - \frac{x+y}{\sqrt{2}} \right)^3, \quad \boldsymbol{f}_\mathrm{ext} = \begin{bmatrix} 0 \\ -N(t) \\ 0 \end{bmatrix} \quad .
\end{aligned}
\tag{A.3}
$$

Figure 2.9 displays $x(t)$ and $N(t)$ for the parameters $M = 1$, $\mu = 0.65$, $k_\mathrm{x} = 11$, $k_\mathrm{y} = 20$, $k_\mathrm{z} = 100$, $k_\mathrm{lin} = 0$, $k_\mathrm{nl} = 5$, $c_\mathrm{x,y,z,lin,nl} = 0$ and $N = 10 \cdot \sin(0.1t)$ with initial conditions $\boldsymbol{x}_0 = [0, 0.1, 1, 0, 6, 6]^\top$ for $t \in [0, 50]$. Figure 2.4 displays the bifurcation diagram for the parameter set $m = 1$, $\mu = 0.65$, $k_\mathrm{y} = 20$, $k_\mathrm{z} = 100$, $k_\mathrm{lin} = 10$, $k_\mathrm{nl} = 5$ for different damping levels $c = c_\mathrm{x,y,z,lin,nl}$ along the bifurcation parameter k_x. Initial conditions on the stable isola solution in Figure 2.4 are given by $\boldsymbol{x}_0\,(k_\mathrm{x} = 23, c = 0.002) = [8.62, -10.52, -4.64, 1.97, 7.38, -9.34]^\top$. $\hat{x}$ denotes the maximum value of the state x along one cycle of vibration.

A.3 4DOF minimal model with cubic stiffness

The system with four degrees of freedom and the state vector $\boldsymbol{x} = \left[x, y, z_1, z_2, \dot{x}, \dot{y}, \dot{z}_1, \dot{z}_2\right]^\top$ is given by

$$
\begin{aligned}
\boldsymbol{C}+\boldsymbol{G} &= \begin{bmatrix}
c_\mathrm{x} + \frac{c_\mathrm{lin1}}{2} + \frac{c_\mathrm{lin2}}{2} & \frac{c_\mathrm{lin1}}{2} + \frac{c_\mathrm{lin2}}{2} & -\frac{c_\mathrm{lin1}}{\sqrt{2}} & -\frac{c_\mathrm{lin2}}{\sqrt{2}} \\
\frac{c_\mathrm{lin1}}{2} + \frac{c_\mathrm{lin2}}{2} & c_\mathrm{y} + \frac{c_\mathrm{lin1}}{2} + \frac{c_\mathrm{lin2}}{2} & -\frac{c_\mathrm{lin1}}{\sqrt{2}} & -\frac{c_\mathrm{lin2}}{\sqrt{2}} \\
-\frac{c_\mathrm{lin1}}{\sqrt{2}} & -\frac{c_\mathrm{lin1}}{\sqrt{2}} & c_\mathrm{z1} + c_\mathrm{lin1} & 0 \\
-\frac{c_\mathrm{lin2}}{\sqrt{2}} & -\frac{c_\mathrm{lin2}}{\sqrt{2}} & 0 & c_\mathrm{z2} + c_\mathrm{lin2}
\end{bmatrix}, \\
\boldsymbol{K}+\boldsymbol{N} &= \begin{bmatrix}
k_\mathrm{x} + \frac{k_\mathrm{lin1}}{2} + \frac{k_\mathrm{lin2}}{2} & \frac{k_\mathrm{lin1}}{2} + \frac{k_\mathrm{lin2}}{2} - \mu k_\mathrm{y} & -\frac{k_\mathrm{lin1}}{\sqrt{2}} & -\frac{k_\mathrm{lin2}}{\sqrt{2}} \\
\frac{k_\mathrm{lin1}}{2} + \frac{k_\mathrm{lin2}}{2} & k_\mathrm{y} + \frac{k_\mathrm{lin1}}{2} + \frac{k_\mathrm{lin2}}{2} & -\frac{k_\mathrm{lin1}}{\sqrt{2}} & -\frac{k_\mathrm{lin2}}{\sqrt{2}} \\
-\frac{k_\mathrm{lin1}}{\sqrt{2}} & -\frac{k_\mathrm{lin1}}{\sqrt{2}} & k_\mathrm{z1} + k_\mathrm{lin1} & 0 \\
-\frac{k_\mathrm{lin2}}{\sqrt{2}} & -\frac{k_\mathrm{lin2}}{\sqrt{2}} & 0 & k_\mathrm{z2} + k_\mathrm{lin2}
\end{bmatrix}, \\
\boldsymbol{M} &= \begin{bmatrix} M & 0 & 0 & 0 \\ 0 & M & 0 & 0 \\ 0 & 0 & M & 0 \\ 0 & 0 & 0 & M \end{bmatrix}, \quad \boldsymbol{f}_\mathrm{nl} = \begin{bmatrix} {}^{-1}/_{\sqrt{2}} & {}^{-1}/_{\sqrt{2}} \\ {}^{-1}/_{\sqrt{2}} & {}^{-1}/_{\sqrt{2}} \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} k_\mathrm{nl1} \cdot \left(z1 - \frac{x+y}{\sqrt{2}}\right)^3 \\ k_\mathrm{nl2} \cdot \left(z2 - \frac{x+y}{\sqrt{2}}\right)^3 \end{bmatrix}, \quad \boldsymbol{f}_\mathrm{ext} = \boldsymbol{0}.
\end{aligned}
\tag{A.4}
$$

Figures 2.5 (a,b) display the bifurcations for the parameter set $M = 1$, $\mu = 0.65$, $k_\mathrm{y} = 20$, $k_\mathrm{z1} = 100$, $k_\mathrm{z2} = 20$, $k_\mathrm{lin1,2} = 10$, $k_\mathrm{nl1,2} = 5$, $c_\mathrm{x,y,z1,2,lin1,2,nl1,2} = 0.02$ and the state space at $k_\mathrm{x} = 11.2$ for which the initial conditions lying on the periodic orbits read

$$
\begin{aligned}
\boldsymbol{x}_\mathrm{LC1} &= [10.97, -17.27, -0.72, -5.12, -3.93, 7.47, -2.16, 0.58]^\top \\
\boldsymbol{x}_\mathrm{LC2} &= [14.89, -21.84, -3.11, -6.89, 7.44, -6.20, -6.70, -5.36]^\top \\
\boldsymbol{x}_\mathrm{LC3} &= [18.32, -22.30, -3.12, -4.42, 8.41, -3.22, -11.36, -3.86]^\top \\
\boldsymbol{x}_\mathrm{LC4} &= [21.39, -23.98, -2.77, -3.10, 6.12, 7.91, -20.45, -3.18]^\top \\
\boldsymbol{x}_\mathrm{LC5} &= [18.75, -19.89, -3.26, -2.46, -57.67, 65.36, 1.48, 5.33]^\top \quad .
\end{aligned}
$$

In Figures 2.5 (c, d) the damping values are reduced to $c = 0.002$ and the initial conditions on the periodic orbits at $k_\mathrm{x} = 11.0$ read

$$\boldsymbol{x}_{\mathrm{LC1}} = [11.17, -17.37, -0.69, -5.04, -0.36, 0.58, -0.14, 0.03]^\top$$
$$\boldsymbol{x}_{\mathrm{LC2}} = [15.29, -21.98, -3.12, -6.73, -1.72, 1.61, 0.76, 0.50]^\top$$
$$\boldsymbol{x}_{\mathrm{LC3}} = [18.42, -22.45, -3.26, -4.58, -1.97, 1.85, 0.66, 0.32]^\top$$
$$\boldsymbol{x}_{\mathrm{LC4}} = [12.93, -12.63, -2.86, -1.37, -92.67, 101.32, 7.16, 9.02]^\top$$
$$\boldsymbol{x}_{\mathrm{LC5}} = [21.85, -24.11, -3.39, -3.25, -2.52, 2.84, 0.01, 0.20]^\top \quad .$$

A.4 Wobbling disk model

The wobbling disk model in Figure 2.6 is governed by the following set of differential equations

$$\begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \ddot{q}_1 \\ \ddot{q}_2 \\ \dot{\mu} \\ \dot{T} \end{bmatrix} = \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ {}^1\!/_\Theta \left\{ \left({}^1\!/_2 \mu N_0 {}^{h^2}\!/_{r\Omega} + 2cr^2 + c_\mathrm{t} \right) \dot{q}_1 + 2\Theta\Omega \dot{q}_2 + \left(k_\mathrm{t} + 2k_\mathrm{nl} r^2 + N_0 h \right) q_1 + \left({}^1\!/_2 \mu N_0 {}^{h^2}\!/_r \right) q_2 \right\} \\ {}^1\!/_\Theta \left\{ \left(-2\Theta\Omega - \mu chr \right) \dot{q}_1 + c_\mathrm{t} \dot{q}_2 + \left(-\mu \left(k_\mathrm{nl} hr + 2\mu N_0 r \right) \right) q_1 + \left(k_\mathrm{t} + \left(1 + \mu^2 \right) N_0 h \right) q_2 \right\} \\ -A_\mu \left(\mu - \mu_0 - C_\mu T \right) \\ -B_\mathrm{T} \left(T - T_0 - D_\mathrm{T} \left| N v_\mathrm{rel} \right| \right) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ A_\mu \mu_\mathrm{fluc} \\ 0 \end{bmatrix} \cdot \frac{x_\mathrm{L}}{18.75} \tag{A.5}$$

where x_L is the first state of the co-simulation of the Lorenz system in Equation (2.2) and the reference parameter values are disk thickness $h = 0.02$, disk radius $r = 0.13$, normal pre-load $N_0 = 5$, rotational damper $c_\mathrm{t} = 0.2$, rotational stiffness $k_\mathrm{t} = 2$, contact damper $c = 0.2$, quadratic contact stiffness $k_\mathrm{nl}(u) = 10 + 10u^2$ and central moment of inertia $\Theta = 30$. The time constants of the friction model are $A_\mu = 1$, $B_\mathrm{T} = 0.03$ and the gains $C_\mu = 0.002$ and $D_\mathrm{T} = 100$ for the initial values $\mu_0 = 0.4$ and $T_0 = 20$ with $\mu_\mathrm{fluc} = 0.2$.

A.5 Higher harmonics

Consider a Duffing-type oscillator with cubic stiffness k_3

$$m\ddot{x} + kx + k_3 x^3 = F\sin(\omega t) \tag{A.6}$$

and try to obtain the time evolution using the classical ansatz $x(t) = A\sin(\omega t)$. Substituting the ansatz and its second derivative yields

$$-m\omega^2 A\sin(\omega t) + Ak\sin(\omega t) + k_3 A^3 \sin^3(\omega t) = F\sin(\omega t) \tag{A.7}$$

where a third-order sinusoidal term arises to create the third-order higher harmonic. This term can be re-written as

$$\sin^3(\omega t) = \frac{1}{4}\left(3\sin(\omega t) + \underbrace{\sin(3\omega t)}_{3^{\text{rd}}\text{ harmonic}}\right) \tag{A.8}$$

such that one will have to add $\sin(3\omega t)$ to the ansatz, e.g. $x(t) = A\sin(\omega t) + B\sin(3\omega t)$, in order to capture the complete dynamics. Inserting this ansatz will create terms of 5^{th}, 7^{th} and 9^{th} order. Repeated extention of the original ansatz by those terms will create even more higher order harmonic terms. Finally, one will obtain an infinite series of harmonics with the specific terms depending on the type of nonlinearity. In this case, the cubic stiffness creates only odd multiples of the fundamental frequency.

A.6 Limitations of linear stability analysis

The van der Pol oscillator

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \epsilon\left(1 - x_1^2\right)x_2 - x_1 \end{aligned} \tag{A.9}$$

is a prototypical self-excited system for visualizing the interplay of linear instability, nonlinearity and damping. The parameter ϵ controls the amount of damping which depends quadratically on the first state's amplitude. For small amplitudes $x_1^2 < 1$ the damping is negative, i.e. destabilizing, while for large amplitudes $x_1^2 > 1$ the damping is stabilizing. The system is linearly unstable for $\epsilon > 0$ such that vibrations grow exponentially, see Figure A.1. The nonlinear damping limits the amplitudes by balancing out the energy flux along one cycle of vibration and resulting trajectories end up on a limit cycle. For larger ϵ values the eigenvalues' real parts increase, but the resulting limit cycles exhibit an almost constant amplitude. Hence, this small self-excited system illustrates the conceptual limitations of linear stability considerations and the impact of the nonlinearity on the steady-state system response.

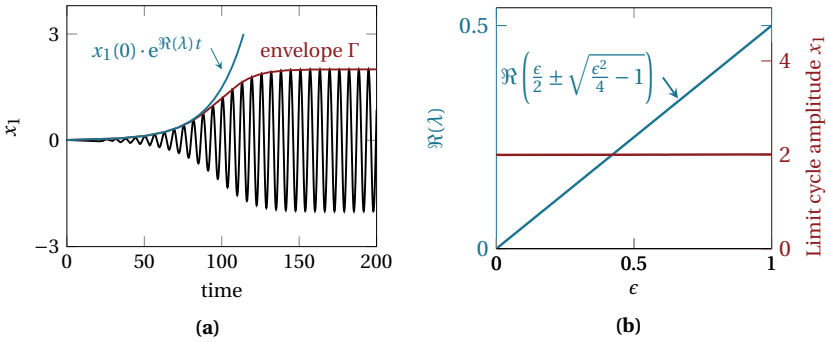


Figure (A.1) (a) time series of the first state x_1 of the van der Pol oscillator (A.9) for $\epsilon = 0.1$ starting from $[0.01, 0.0]^\top$. Amplitudes grow exponentially according to the eigenvalue's real part $\Re(\lambda)$ before the nonlinearity sets in to balance the system energy flux on a limit cycle. (b) variation of the ϵ value and corresponding eigenvalue real parts that are strictly increasing. The resulting limit cycle solutions have an approximately constant amplitude. Adapted from Figure B.12 from [280]

B Methods

B.1 Numerical approximation of Lyapunov exponents

The Lyapunov spectrum of differential equations can be calculated analytically by solving the perturbed variational problem associated with the equations of motion. Theoretically, the spectrum Λ is given by the divergence of the Jacobian. As the Jacobian is time-variant for nonlinear dynamical systems, the classical Wolff scheme [184] is not a practical option for systems with more than a few degrees-of-freedom. Since the Lyapunov exponents measure the divergence of trajectories that start nearby, the growth of slightly disturbed states is followed along the actual trajectory of the dynamical system. Mathematically, this approach takes the form of a variational problem $\dot{\boldsymbol{Y}} = \boldsymbol{J}\boldsymbol{Y}$. $\boldsymbol{Y}$ denotes the matrix of small perturbations of each state into each dimension of the phase space, and $\boldsymbol{J}$ is the Jacobian of the dynamical system. Then, the original set of differential equations $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x})$ is augmented by the variational problem to obtain the larger system $\left[\dot{\boldsymbol{x}}, \dot{\boldsymbol{Y}}\right]^\top = \left[\boldsymbol{f}(\boldsymbol{x}), \boldsymbol{J}\boldsymbol{Y}\right]^\top$. The augmented system is subjected to a classical time integration for a short time span. During this time, the perturbations $\boldsymbol{Y}$ have exponentially grown away from the system's trajectory $\boldsymbol{x}$. Most importantly, the perturbations align along the direction of fastest divergence, i.e. according to the largest positive Lyapunov exponent. The distance from the original trajectory in every dimension gives an estimate of the Lyapunov spectrum, the so-called *instantaneous spectrum*. A QR decomposition of $\boldsymbol{Y}$ yields both the instantaneous spectrum, as well as the re-initialization of the perturbations. This process is iterated in time until both the trajectories of the dynamical system as well as the spectrum have reached a quasi-steady state. The converged values of the instantaneous spectrum then form the Lyapunov spectrum. Owing to the stiff character of the variational problem, this procedure can be computationally demanding.

B.2 Orbit diagram calculation

The study of bifurcation diagrams can be considered one of the main tasks in applied nonlinear dynamics. However, classical bifurcation diagrams depict only fixed points and periodic orbits. Hence, quasi-periodic and irregular solutions are difficult to display in a coherent fashion. To overcome this issue, the Poincaré section is the best-

known representation. However, the Poincaré section typically requires to define a projection hyperplane, which most often is formulated by the period of an external excitation force. Hence, the states are studied at every integer multiple of the period. However, when a dominant period is absent, such as in self-excited vibrations, the definition of a Poincaré hyperplane can be unclear. The *orbit diagram*, as introduced and discussed by [10, 300, 301, 302] as a generalized form of the Poincaré section, allows to study the qualitative and quantitative character of both regular and irregular solutions from numerical modeling and experimental testing in a single diagram. Starting from a signal of the time evolution of a system state, all local maxima are collected and this set of amplitudes is displayed as dots along the bifurcation parameter. The local maxima of fixed points and mono-periodic motions coincide at a single amplitude value, while period-n cycles create distinct clusters of points. Chaotic motion creates a scattering range of points. As an illustrative example, Figure B.1 depicts the

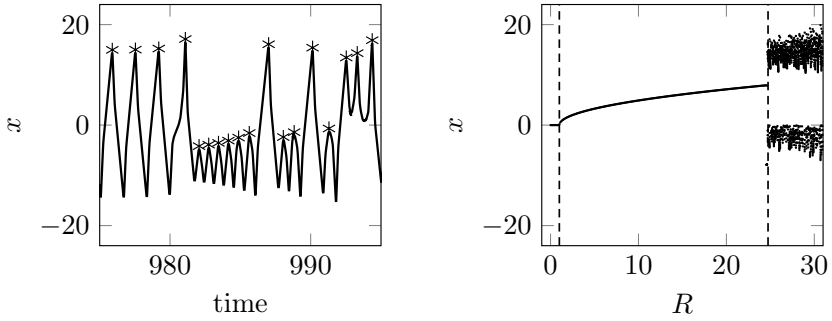


Figure (B.1) Time series for the initial conditions $[0,1,0]^\top$ of the first Lorenz system (2.2) state for $\sigma = 10$, $\beta = 8/3$ and $R = 25$ with marked local maxima. Constituting the orbit diagram, the maxima are depicted on the right for a variation of the parameter R. The pitchfork bifurcation of fixed points at $R = 1$ and the Hopf bifurcation at $R = 24.74$ giving rise to chaotic solutions are clearly visible

orbit diagram for the Lorenz system (2.2) for a parameter variation of the Rayleigh number R. The equilibrium solution undergoes two bifurcations, and chaos is born at $R \approx 24.74$. The orbit diagram is able to depict this route to chaos graphically. Naturally, unstable branches will not be displayed in an orbit diagram that is determined from a steady-state trajectory. Note that for feature generation purposes, an unsupervised machine learning approach can be utilized to quantify the structure of the point cloud arising for a single system configuration from the orbit diagram. Different clusters are formed, such as points collapsing in a ϵ-sphere for equilibria, two distinct clusters for a period-2 cycle, and so on.

B.3 Hilbert-Huang transform

The HHT is defined by an empirical algorithm that can be outlined as follows. An intrinsic mode function (IMF) is defined by two motives:

1. For the complete IMF time series data, the number of local extrema and the number of zero-crossings must either be equal or differ by one, and
2. the mean value of the upper envelope (given by the local maxima) and the lower envelope (given by the local minima) is zero

which ensures a well-behaved Hilbert transform of the IMF. Envelopes Γ^{up} and Γ^{low} are constructed by fitting a cubic spline to the local extrema. The IMF is obtained via a sifting process from the original data $\boldsymbol{s}(t)$ as follows. The mean $m_1 = \Gamma^{\text{up}} - \Gamma^{\text{low}}$ of the envelopes is subtracted from the signal to obtain the first proto-IMF h_1

$$h_{1,0} = \boldsymbol{s}(t) - m_1 \quad . \tag{B.1}$$

As $h_{1,0}$ may not define an IMF, the sifting process continues to treat $h_{1,0}$ as the input signal, construct envelopes and subtract the mean for k times

$$h_{1,k-1} - m_{1k} = h_{1k} \tag{B.2}$$

until a stopping criterion is met that ensures to have found the first IMF $c_1 = h_{1k}$. Several stopping criteria have been proposed, such as a Cauchy convergence test or a simpler threshold method. The first IMF is subtracted from the original signal to compute the first residual $r_1 = \boldsymbol{s}(t) - c_1$. To find the second IMF c_2, the residual is treated as input data to the sifting process. The overall process is finalized once the last residual becomes a monotonic function that cannot form an IMF by definition. Using the IMFs, the original signal is decomposed

$$\boldsymbol{s}(t) = \sum_{j=1}^{n} c_j + r_n \tag{B.3}$$

into n IMFs that capture the characteristic oscillatory modes of the signal. The instantaneous frequency content of the IMFs can be computed using the Hilbert spectral analysis. Well-known limitation of the HHT are narrow-band signals, mode-mixing and end effects.

B.4 Fourier and wavelet transforms

To stay consistent with the classical notation, various symbols are re-defined in the following paragraphs that have already been used in the main part of this work in a different context.

The **Fourier transform** [174] of a signal $\boldsymbol{s}(t)$ is given by

$$\hat{\boldsymbol{s}}(\omega) = \int_{-\infty}^{\infty} \boldsymbol{s}(t) \cdot \mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{d}t \quad . \tag{B.4}$$

The short-time Fourier transform (STFT) or *Gabor* transform [175]

$$\hat{\boldsymbol{S}}(\tau,\omega) = \int_{-\infty}^{\infty} \boldsymbol{s}(t)\,\mathrm{e}^{-\mathrm{i}\omega t} \cdot \boldsymbol{w}(t-\tau)\,\mathrm{d}t \tag{B.5}$$

can resolve frequency along time using the window function $\boldsymbol{w}$.

The **wavelet transform** (CWT) [176, 177]

$$\mathcal{W}_{\psi}[\boldsymbol{s}](a,b) = \int_{-\infty}^{\infty} \boldsymbol{s}(t) \cdot \overline{\psi}_{a,b}(t)\,\mathrm{d}t, \quad \psi(t) = \frac{1}{\sqrt{2}}\left(\frac{t-b}{a}\right) \tag{B.6}$$

is defined by the complex conjugate $\overline{(\cdot)}$ of the mother wavelet $\psi_{a,b}(t)$ for the translation parameter b and scale parameter a. As a central property, wavelets must have a compact support and fulfill the boundedness condition

$$C_{\psi} = \int_{-\infty}^{\infty} \frac{\left|\hat{\psi}(\omega)\right|^2}{|\omega|}\,\mathrm{d}\omega < \infty \tag{B.7}$$

where $\hat{\psi}$ denotes the Fourier transform of the wavelet. Then, the inverse wavelet transform (iCWT) is given by

$$\boldsymbol{s}(t) = \frac{1}{C_{\psi}} \int \int_{-\infty}^{\infty} \frac{1}{a^2} \left(\mathcal{W}_{\psi}[\boldsymbol{s}](a,b)\right) \cdot \psi_{a,b}(t)\,\mathrm{d}a\,\mathrm{d}b \quad . \tag{B.8}$$

Time and frequency localization properties of the CWT depend on the choice of mother wavelet. Time and frequency localization is given by the second moment measuring the energy around some mean time $t^{\star}$ and some frequency $\omega^{\star}$

$$\sigma_t^2 = \int_{-\infty}^{\infty} \left(t-t^{\star}\right)^2 \left|\psi(t)\right|^2 \mathrm{d}t, \quad \sigma_{\omega}^2 = \int_{-\infty}^{\infty} \left(\omega-\omega^{\star}\right)^2 \left|\hat{\psi}(\omega)\right| \mathrm{d}\omega \quad . \tag{B.9}$$

When computing the wavelet transform of some measured signal, the discrete wavelet transform (DWT)

$$\psi_{j,k}(t) = \frac{1}{\sqrt{a_0^j}} \psi\left(\frac{t - k b_0 a_0^j}{a_0^j}\right) \tag{B.10}$$

is obtained through sampling of the scale and translation parameters a_0 and b_0, respectively. The standard choice [105] is $a_0 = 2$, $b_0 = 1$, which corresponds to a dyadic grid arrangement. The discrete wavelet transform values at the scale location grid index j and k

$$D_{j,k}(\boldsymbol{s}) = \int_{-\infty}^{\infty} \boldsymbol{s}(t)\,\psi_{j,k}(t)\,\mathrm{d}t \tag{B.11}$$

are denoted as *wavelet coefficients* or *detail coefficients*. The *approximation coefficients* $A_{j,k}$

$$A_{j,k}(\boldsymbol{s}) = \int_{-\infty}^{\infty} \boldsymbol{s}(t)\,\phi_{j,k}(t)\,\mathrm{d}t \tag{B.12}$$

at scale j are obtained from the convolution of the signal with the scaling function

$$\phi_{j,k}(t) = 2^{-\frac{j}{2}}\phi\left(2^{-j}t - k\right) \quad \text{for } a_0 = 2, b_0 = 1 \tag{B.13}$$

which is orthogonal to the translation, but not to the dilation. Practically, the signal is fed through a high pass filter and a low pass filter, each of which is unique to the wavelet family. The result of the low pass filter is then fed through the next filter bank at the next lower scale. Each high pass yields detail coefficients $D_{j,k}$, and each low pass yields approximate coefficients $A_{j,k}$ at level j and location k. The scale, i.e. the frequency range corresponding to the levels, is specific to the wavelet family and the DWT sampling. Multi-resolution analysis and time scale separation can now be achieved through replacing all detail coefficients by zero except for level j and computing the inverse DWT. As a result, one time series $\boldsymbol{s}_{\mathrm{DWT},j}$ per level is obtained with a specific frequency bandwidth as illustrated in Figure 3.6 (a).

B.5 Recurrence plot quantification analysis

Table (B.1) Overview on the metrics from recurrence quantification analysis [181, 198]. $P(l)$ denotes the histogram of the lengths l of diagonals, $P(v)$ denotes the histogram of the lengths v of vertical lines

metric	definition	description
RR	$\mathrm{RR}=\frac{1}{N}\sum_{i,j=1}^{N^2} R_{i,j}$	recurrence rate: fraction of recurrence points in the complete RP
DET	$\mathrm{DET}=\frac{\sum_{l=l_{\min}}^{N} lP(l)}{\sum_{l=1}^{N} lP(l)}$	determinism: ratio of recurrences that form diagonals in the RP
LAM	$\mathrm{LAM}=\frac{\sum_{v=v_{\min}}^{N} vP(v)}{\sum_{v=1}^{N} vP(v)}$	laminarity: ratio of recurrences that form vertical lines in the RP
L	$L=\frac{\sum_{l=l_{\min}}^{N} lP(l)}{\sum_{l=l_{\min}}^{N} lP(l)}$	average length of all diagonal lines
TT	$\mathrm{TT}=\frac{\sum_{v=v_{\min}}^{N} vP(v)}{\sum_{v=v_{\min}}^{N} vP(v)}$	average length of all vertical lines
$L_{\max}$	$L_{\max}=\max\left(\{l_i\}\right), i=1,\ldots,N_{\mathrm{l}}$	length of the longest diagonal line
$V_{\max}$	$L_{\max}=\max\left(\{v_i\}\right), i=1,\ldots,N_{\mathrm{v}}$	length of the longest vertical line
DIV	$\mathrm{DIV}=\frac{1}{l_{\max}}$	divergence: related to the sum of positive Lyapunov exponents and the KS entropy
ENTR	$\mathrm{ENTR}=-\sum_{l=l_{\min}}^{N} p(l)\ln\left(p(l)\right)$	Shannon entropy of the probability distribution of the diagonal line lengths $p(l)$

The computation of recurrence plots requires the selection of several parameters that can crucially effect the result. The norm $|\cdot|$ can be chosen freely as well as the ϵ neighborhood. Several rules of thumb [303] propose to select the threshold such that the recurrence rate is not larger than 10%. Alternatively, an adaptive threshold can be selected that picks a fixed amount of nearest neighbors (FAN). This norm is especially useful for a qualitative comparison [181] between signals that may stem from very different sources. Most quantification metrics report distributions of the diagonal lines l and the vertical lines v. Hence, the lower limits for counting a structure as a line has to be set, i.e. $l_{\min} > 1$, $v_{\min} > 1$. Table B.1 lists most of the common RPQA metrics.

B.6 Performance metrics for machine learning

The confusion matrix for a binary classification problem is given in Figure B.2. A selection of metrics derived from this matrix is given in Table B.2. While *accuracy* gives

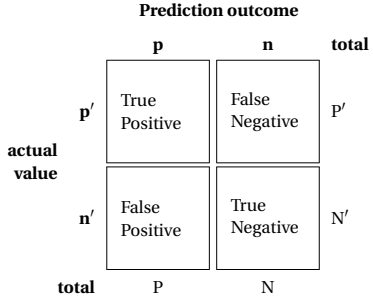


Figure (B.2) Confusion matrix for a binary classification (positive, **p**, negative **n**). P′ positive and N′ negative samples exist in the ground truth data, while the model predicts P and N

the overall averaged fraction of true positives, *recall* measures the sensitivity, also denoted as *true positive rate*. *Precision* measures only the performance on the positive class. The F_1 score balances recall and precision. The *ROC* collects the number of true positives over the number of false positives for increasing confidence levels. A large *area under curve* (AUC) value will therefore identify low error rates. In most scenarios, a false positive prediction is not as severe as a false negative prediction since a model should rather be too sensitive than too insensitive to the positive class.

Table (B.2) Overview on classification metrics, adapted from [235]

metric	definition
accuracy	$\frac{TP+TN}{TP+FP+FN+TN}$
precision	$\frac{TP}{TP+FP}$
recall	$\frac{TP}{TP+FN}$
F_1	$2 \cdot \frac{precision \cdot recall}{precision+recall}$
AUC	$\int ROC$

B.7 Loss functions for machine learning

Measuring the prediction quality of some model requires the definition of an error measure, or the so-called *loss* that is typically denoted by J. There are various loss definitions in the data sciences [211], from which the most important ones are reported in Table B.3

Table (B.3) Loss functions J used in data science. The expected probability distribution of the ground truth is denoted by $\boldsymbol{p} = \left[p_1, \ldots, p_N\right]^\top$ and the predicted probability is given by $\hat{\boldsymbol{y}}$. N denotes the number of class labels that can be predicted, i.e. $N = 2$ for a binary classification task

loss	definition
linear regression squared error	$J = \sum_{n=1}^{N} \left(y_n - \hat{y}_n\right)^2$
cross entropy	$J = -\frac{1}{N} \sum_{n=1}^{N} H\left(\boldsymbol{y}, \hat{\boldsymbol{y}}\right) = -\sum_{n=1}^{N} y_n \log\left(\hat{y}_n\right)$
binary cross entropy	$J = -\frac{1}{N} \sum_{n=1}^{N} y_n \log\left(\hat{y}_n\right) + \left(1 - y_n\right) \log\left(1 - \hat{y}_n\right)$

For regression tasks, the linear regression squared error assumes the error to be distributed by means of a Gaussian. For classification tasks, cross entropy is typically used to express the loss. It builds on the Shannon entropy [24] $H\left(\boldsymbol{p}\right) = -\sum_i p_i \log_2\left(p_i\right)$ from information theory. From a data science perspective, the probability distribution $\boldsymbol{p}$ reflects the ground truth and the probability distribution $\boldsymbol{q}$ represents the prediction. Then, the cross entropy defined by the sum of the entropy of $\boldsymbol{p}$ and the Kullback-Leibler divergence $D_{\mathrm{KL}}\left(\boldsymbol{p}, \boldsymbol{q}\right)$ [304] is given by $H\left(\boldsymbol{p}, \boldsymbol{q}\right) = -\sum_i p_i \log_2\left(q_i\right)$ which measures the distribution difference, i.e. how far the model prediction is off. $H\left(\boldsymbol{p}, \boldsymbol{q}\right)$ is not symmetrical. The binary cross entropy loss is used for classification tasks where the output is either true (1) or false (0), i.e. the cross entropy sum unrolled for $N = 2$. The loss for a complete training set is computed by averaging the cross-entropy across all examples.

B.8 Activation functions for machine learning

The activation function is the ingredient that adds nonlinearity to neural networks and most deep learning methods. Essentially, the summation of weighted inputs is fed through the activation function σ to compute the output of the current neuron. Several classical use cases are indicated for different activation functions in Table B.4. Especially relevant for the thesis at hand is the softmax activation for classification problems. This function is used for the output layer of networks in order to derive class-vise probabilities that add up to unity for all classes.

Table (B.4) Overview on common activation functions σ used for input x

function	expression	use case
Logistic, (sigmoid)	$\sigma(x) = \frac{1}{1+e^{-x}}$	binary classification
Rectified linear unit (ReLu)	$\sigma(x) = \begin{cases} 0, & \text{for } x \leq 0 \\ x, & \text{for } x > 0 \end{cases}$	faster deep learning
Softmax	$\sigma_i(x) = \frac{e^{x_i}}{\sum_{i=1}^{J} e^{x_i}}, i = 1, \ldots, J$	multi-class classification

C Additional information for application cases

C.1 Nonlinear time series analysis of brake vibrations

Figure C.1 depicts the RPQA metrics averaged per quiet and vibrating section.

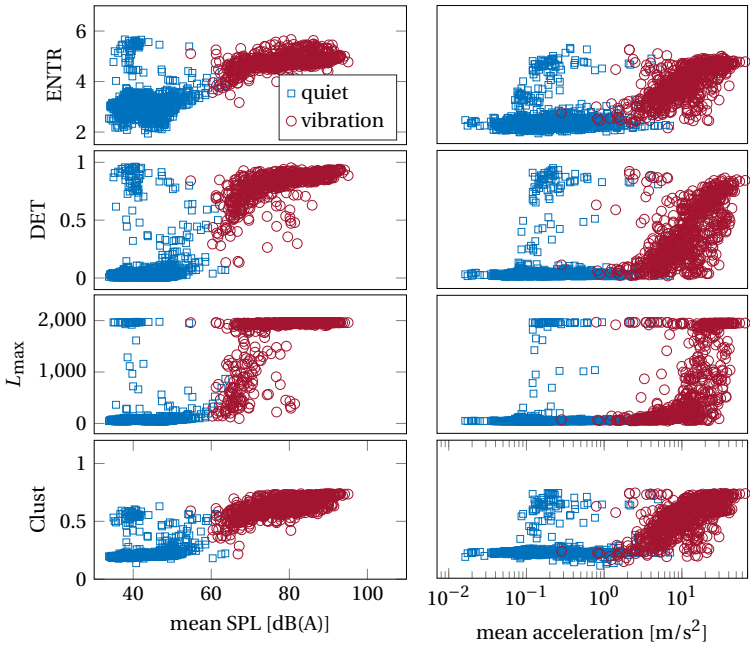


Figure (C.1) Section-averaged RPQA quantifiers for quiet and vibrating sections of the brake vibrations data set. Adapted from Figures 11, 12, 13 from [162]

C.2 Instability prediction using deep learning

Table (C.1) Hyperparameter studies for the sequence-to-scalar configuration. Each model was trained for 200 epochs for a 3-fold cross validation on data set *A*. The validation scores are reported by the MCC

batch size	layers	units	window length	rel. shift size	mean MCC	std MCC
16	1	64	200	0.5	0.54	0.03
16	1	64	200	0.75	0.58	0.02
16	1	64	200	1.0	0.61	0.04
16	1	64	400	0.5	0.62	0.01
16	1	64	400	0.75	0.6	0.08
16	1	64	400	1.0	0.6	0.02
16	1	128	200	0.5	0.59	0.0
16	1	128	200	0.75	0.64	0.01
16	1	128	200	1.0	0.65	0.03
16	1	128	400	0.5	0.65	0.02
16	1	128	400	0.75	0.58	0.05
16	1	128	400	1.0	0.64	0.04
16	1	256	200	0.5	0.56	0.01
16	1	256	200	0.75	0.61	0.02
16	1	256	200	1.0	0.64	0.0
16	1	256	400	0.5	0.62	0.01
16	1	256	400	0.75	0.61	0.03
16	1	256	400	1.0	0.69	0.04
16	2	64	200	0.5	0.55	0.01
16	2	64	200	0.75	0.57	0.02
16	2	64	200	1.0	0.58	0.05
16	2	64	400	0.5	0.58	0.01
16	2	64	400	0.75	0.62	0.04
16	2	64	400	1.0	0.57	0.08
16	2	128	200	0.5	0.51	0.06
16	2	128	200	0.75	0.58	0.02
16	2	128	200	1.0	0.63	0.02
16	2	128	400	0.5	0.58	0.04
16	2	128	400	0.75	0.62	0.04
16	2	128	400	1.0	0.64	0.02
16	2	256	200	0.5	0.51	0.07
16	2	256	200	0.75	0.58	0.02
16	2	256	200	1.0	0.62	0.03
16	2	256	400	0.5	0.62	0.02
16	2	256	400	0.75	0.6	0.03
16	2	256	400	1.0	0.66	0.01
64	1	64	200	0.5	0.68	0.03
64	1	64	200	0.75	0.67	0.03
64	1	64	200	1.0	0.61	0.02
64	1	64	400	0.5	0.65	0.02

64	1	64	400	0.75	0.61	0.08
64	1	64	400	1.0	0.66	0.04
64	1	128	200	0.5	0.68	0.02
64	1	128	200	0.75	0.7	0.02
64	1	128	200	1.0	0.69	0.01
64	1	128	400	0.5	0.67	0.05
64	1	128	400	0.75	0.62	0.09
64	1	128	400	1.0	0.67	0.03
64	1	256	200	0.5	0.7	0.01
64	1	256	200	0.75	0.69	0.01
64	1	256	200	1.0	0.7	0.01
64	1	256	400	0.5	0.7	0.02
64	1	256	400	0.75	0.71	0.02
64	1	256	400	1.0	0.65	0.08
64	2	64	200	0.5	0.63	0.01
64	2	64	200	0.75	0.64	0.03
64	2	64	200	1.0	0.63	0.03
64	2	64	400	0.5	0.63	0.03
64	2	64	400	0.75	0.63	0.06
64	2	64	400	1.0	0.64	0.03
64	2	128	200	0.5	0.68	0.02
64	2	128	200	0.75	0.71	0.02
64	2	128	200	1.0	0.65	0.04
64	2	128	400	0.5	0.67	0.02
64	2	128	400	0.75	0.63	0.03
64	2	128	400	1.0	0.62	0.06
64	2	256	200	0.5	0.7	0.03
64	2	256	200	0.75	0.69	0.02
64	2	256	200	1.0	0.68	0.02
64	2	256	400	0.5	0.68	0.01
64	2	256	400	0.75	0.67	0.04
64	2	256	400	1.0	0.66	0.04
256	1	64	200	0.5	0.7	0.01
256	1	64	200	0.75	0.71	0.03
256	1	64	200	1.0	0.65	0.05
256	1	64	400	0.5	0.68	0.0
256	1	64	400	0.75	0.65	0.06
256	1	64	400	1.0	0.65	0.02
256	1	128	200	0.5	0.68	0.01
256	1	128	200	0.75	0.73	0.02
256	1	128	200	1.0	0.69	0.03
256	1	128	400	0.5	0.72	0.01
256	1	128	400	0.75	0.63	0.07
256	1	128	400	1.0	0.69	0.05
256	1	256	200	0.5	0.73	0.01
256	**1**	**256**	**200**	**0.75**	**0.74**	**0.01**
256	1	256	200	1.0	0.72	0.01
256	1	256	400	0.5	0.72	0.02

256	1	256	400	0.75	0.74	0.02
256	1	256	400	1.0	0.73	0.04
256	2	64	200	0.5	0.65	0.02
256	2	64	200	0.75	0.71	0.02
256	2	64	200	1.0	0.66	0.02
256	2	64	400	0.5	0.69	0.02
256	2	64	400	0.75	0.66	0.01
256	2	64	400	1.0	0.65	0.01
256	2	128	200	0.5	0.68	0.02
256	2	128	200	0.75	0.7	0.03
256	2	128	200	1.0	0.69	0.04
256	2	128	400	0.5	0.71	0.01
256	2	128	400	0.75	0.74	0.03
256	2	128	400	1.0	0.65	0.05
256	2	256	200	0.5	0.69	0.01
256	2	256	200	0.75	0.71	0.01
256	2	256	200	1.0	0.7	0.02
256	2	256	400	0.5	0.72	0.02
256	2	256	400	0.75	0.73	0.04
256	2	256	400	1.0	0.68	0.03

Table (C.2) Hyperparameter studies for the sequence-to-sequence configuration. Each model was trained for 200 epochs for a 3-fold cross validation on data set *A*. The validation scores are reported by the MCC

batch size	layers	units	window length	rel. shift size	mean MCC	std MCC
16	1	64	200	0.5	0.56	0.01
16	1	64	200	0.75	0.62	0.01
16	1	64	200	1.0	0.62	0.04
16	1	64	400	0.5	0.64	0.03
16	1	64	400	0.75	0.64	0.03
16	1	64	400	1.0	0.65	0.04
16	1	128	200	0.5	0.56	0.03
16	1	128	200	0.75	0.61	0.01
16	1	128	200	1.0	0.66	0.03
16	1	128	400	0.5	0.68	0.01
16	1	128	400	0.75	0.67	0.03
16	1	128	400	1.0	0.67	0.05
16	1	256	200	0.5	0.58	0.04
16	1	256	200	0.75	0.61	0.02
16	1	256	200	1.0	0.66	0.04
16	1	256	400	0.5	0.66	0.03
16	1	256	400	0.75	0.69	0.01
16	1	256	400	1.0	0.69	0.03
16	2	64	200	0.5	0.58	0.02
16	2	64	200	0.75	0.62	0.02

16	2	64	200	1.0	0.64	0.01
16	2	64	400	0.5	0.62	0.04
16	2	64	400	0.75	0.67	0.03
16	2	64	400	1.0	0.65	0.06
16	2	128	200	0.5	0.57	0.03
16	2	128	200	0.75	0.62	0.01
16	2	128	200	1.0	0.65	0.03
16	2	128	400	0.5	0.65	0.03
16	2	128	400	0.75	0.67	0.0
16	2	128	400	1.0	0.7	0.04
16	2	256	200	0.5	0.59	0.03
16	2	256	200	0.75	0.62	0.03
16	2	256	200	1.0	0.67	0.04
16	2	256	400	0.5	0.65	0.04
16	2	256	400	0.75	0.69	0.03
16	2	256	400	1.0	0.7	0.02
64	1	64	200	0.5	0.7	0.02
64	1	64	200	0.75	0.68	0.04
64	1	64	200	1.0	0.68	0.02
64	1	64	400	0.5	0.69	0.02
64	1	64	400	0.75	0.71	0.01
64	1	64	400	1.0	0.71	0.02
64	1	128	200	0.5	0.7	0.0
64	1	128	200	0.75	0.7	0.01
64	1	128	200	1.0	0.7	0.01
64	1	128	400	0.5	0.72	0.02
64	1	128	400	0.75	0.71	0.02
64	1	128	400	1.0	0.71	0.02
64	1	256	200	0.5	0.7	0.01
64	1	256	200	0.75	0.72	0.02
64	1	256	200	1.0	0.69	0.02
64	1	256	400	0.5	0.71	0.02
64	1	256	400	0.75	0.69	0.01
64	1	256	400	1.0	0.7	0.02
64	2	64	200	0.5	0.7	0.01
64	2	64	200	0.75	0.7	0.02
64	2	64	200	1.0	0.67	0.02
64	2	64	400	0.5	0.66	0.04
64	2	64	400	0.75	0.7	0.0
64	2	64	400	1.0	0.73	0.01
64	2	128	200	0.5	0.72	0.01
64	2	128	200	0.75	0.7	0.01
64	2	128	200	1.0	0.69	0.02
64	2	128	400	0.5	0.73	0.03
64	2	128	400	0.75	0.68	0.02
64	2	128	400	1.0	0.71	0.01
64	2	256	200	0.5	0.71	0.02
64	2	256	200	0.75	0.71	0.01

64	2	256	200	1.0	0.71	0.02
64	2	256	400	0.5	0.71	0.04
64	2	256	400	0.75	0.7	0.04
64	2	256	400	1.0	0.7	0.02
256	1	64	200	0.5	0.72	0.02
256	1	64	200	0.75	0.73	0.02
256	1	64	200	1.0	0.72	0.01
256	1	64	400	0.5	0.73	0.02
256	1	64	400	0.75	0.75	0.0
256	1	64	400	1.0	0.74	0.02
256	1	128	200	0.5	0.72	0.01
256	1	128	200	0.75	0.75	0.02
256	1	128	200	1.0	0.74	0.0
256	1	128	400	0.5	0.75	0.02
256	1	128	400	0.75	0.75	0.0
256	1	128	400	1.0	0.76	0.01
256	1	256	200	0.5	0.71	0.01
256	1	256	200	0.75	0.74	0.02
256	1	256	200	1.0	0.74	0.01
256	1	256	400	0.5	0.76	0.0
256	**1**	**256**	**400**	**0.75**	**0.77**	**0.01**
256	1	256	400	1.0	0.77	0.0
256	2	64	200	0.5	0.72	0.01
256	2	64	200	0.75	0.73	0.0
256	2	64	200	1.0	0.73	0.0
256	2	64	400	0.5	0.74	0.01
256	2	64	400	0.75	0.72	0.03
256	2	64	400	1.0	0.75	0.01
256	2	128	200	0.5	0.71	0.01
256	2	128	200	0.75	0.74	0.03
256	2	128	200	1.0	0.73	0.01
256	2	128	400	0.5	0.74	0.02
256	2	128	400	0.75	0.75	0.0
256	2	128	400	1.0	0.74	0.01
256	2	256	200	0.5	0.71	0.02
256	2	256	200	0.75	0.72	0.03
256	2	256	200	1.0	0.74	0.01
256	2	256	400	0.5	0.73	0.03
256	2	256	400	0.75	0.77	0.01
256	2	256	400	1.0	0.76	0.01