Institute of Physics and Astronomy

Towards a robust framework for Recurrence Analysis
Automated state space reconstruction, optimal parameter selection and correction schemes

---

by
Kai Hauke Krämer

May, 2021

A cumulative dissertation for the degree
doctor rerum naturalium (Dr. rer. nat.) in
Theoretical Physics

Submitted to
The Faculty of Mathematics and Natural Sciences at the
University of Potsdam
and the
Potsdam Institute for Climate Impact Research

Supervisors

Prof. Dr. Dr. h.c. mult. Jürgen Kurths
University of Potsdam, Institute of Physics and Astronomy.
Research Domain Complexity Science, Potsdam Institute for Climate Impact Research (PIK),
Member of the Leibniz Association, Telegrafenberg, Potsdam, Germany.
Institute of Physics, Humboldt Universität zu Berlin, Germany.

Dr. habil. Norbert Marwan
Research Domain Complexity Science, Potsdam Institute for Climate Impact Research (PIK),
Member of the Leibniz Association, Telegrafenberg, Potsdam, Germany.
University of Potsdam, Institute of Earth and Environmental Science.

Referees:

Prof. Dr. Dr. h.c. mult. Jürgen Kurths
University of Potsdam, Institute of Physics and Astronomy.
Research Domain Complexity science, Potsdam Institute for Climate Impact Research (PIK),
Member of the Leibniz Association, Telegrafenberg, Potsdam, Germany.
Institute of Physics, Humboldt Universität zu Berlin, Germany.

Dr. habil. Norbert Marwan
Research Domain Complexity Science, Potsdam Institute for Climate Impact Research (PIK),
Member of the Leibniz Association, Telegrafenberg, Potsdam, Germany.
University of Potsdam, Institute of Earth and Environmental Science.

Prof. Dr. Holger Kantz
Max Planck Institute for the Physics of Complex Systems, Dresden, Germany.

"An expert is a person who has made all the mistakes that can be made in a very narrow field."

(–**Niels Bohr**)

Copyright

Towards a robust framework for Recurrence Analysis: Automated state space reconstruction, optimal parameter selection and correction schemes

Abstract

In our daily life, recurrence plays an important role on many spatial and temporal scales and in different contexts. It is the foundation of learning, be it in an evolutionary or in a neural context. It therefore seems natural that recurrence is also a fundamental concept in theoretical dynamical systems science. The way in which states of a system recur or develop in a similar way from similar initial states makes it possible to infer information about the underlying dynamics of the system. The mathematical space in which we define the state of a system (state space) is often high dimensional, especially in complex systems that can also exhibit chaotic dynamics. The recurrence plot (RP) enables us to visualize the recurrences of any high-dimensional systems in a two-dimensional, binary representation. Certain patterns in RPs can be related to physical properties of the underlying system, making the qualitative and quantitative analysis of RPs an integral part of nonlinear systems science. The presented work has a methodological focus and further develops recurrence analysis (RA) by addressing current research questions related to an increasing amount of available data and advances in machine learning techniques. By automatizing a central step in RA, namely the reconstruction of the state space from measured experimental time series, and by investigating the impact of important free parameters this thesis aims to make RA more accessible to researchers outside of physics.

The first part of this dissertation is concerned with the reconstruction of the state space from time series. To this end, a novel idea is proposed which automates the reconstruction problem in the sense that there is no need to preprocesse the data or estimate parameters *a priori*. The key idea is that the *goodness* of a reconstruction can be evaluated by a suitable objective function and that this function is minimized in the embedding process. In addition, the new method can process multivariate time series input data. This is particularly important because multi-channel sensor-based observations are ubiquitous in many research areas and continue to increase. Building on this, the described minimization problem of the objective function is then processed using a machine learning approach.

In the second part technical and methodological aspects of RA are discussed. First, we mathematically justify the idea of setting the most influential free parameter in RA, the recurrence threshold ε, in relation to the distribution of all pairwise distances in the data. This is especially important when comparing different RPs and their quantification statistics and is fundamental to any comparative study. Second, some aspects of recurrence quantification analysis (RQA) are examined. As correction schemes for biased RQA statistics, which are based on diagonal lines, we propose a simple method for dealing with border effects of an RP in RQA and a skeletonization algorithm for RPs. This results in less biased (diagonal line based) RQA statistics for flow-like data. Third, a novel type of RQA characteristic is developed, which can be viewed as a generalized non-linear powerspectrum of high dimensional systems. The spike powerspectrum transforms a spike-train like signal into its frequency domain. When transforming the diagonal line-dependent recurrence rate (τ-RR) of a RP in this way, characteristic periods, which can be seen in the state space representation of the system can be unraveled. This is not the case, when Fourier transforming τ-RR.

Finally, RA and RQA are applied to climate science in the third part and neuroscience in the fourth part. To the best of our knowledge, this is the first time RPs and RQA have been used to analyze lake sediment data in a paleoclimate context. Therefore, we first elaborate on the basic formalism and the interpretation of visually visible patterns in RPs in relation to the underlying proxy data. We show that these patterns can be used to classify certain types of variability and transitions in the Potassium record from six short (< 17m) sediment cores collected during the Chew Bahir Drilling Project. Building on this, the long core ($\sim$290 m composite) from the same site is analyzed and two types of variability and transitions are

identified and compared with ODP Site 967 wetness index from the eastern Mediterranean. Type 1 variability likely reflects the influence of precessional forcing in the lower latitudes at times of maximum values of the long eccentricity cycle (400 kyr) of the earth's orbit around the sun, with a tendency towards extreme events. Type 2 variability appears to be related to the minimum values of this cycle and corresponds to fairly rapid transitions between relatively dry and relatively wet conditions.

In contrast, RQA has been applied in the neuroscientific context for almost two decades. In the final part, RQA statistics are used to quantify the *complexity* in a specific frequency band of multivariate EEG (electroencephalography) data. By analyzing experimental data, it can be shown that the complexity of the signal measured in this way across the sensorimotor cortex decreases as motor tasks are performed. The results are consistent with and complement the well known concepts of motor-related brain processes. We assume that the thus discovered features of neuronal dynamics in the sensorimotor cortex together with the robust RQA methods for identifying and classifying these contribute to the non-invasive EEG-based development of brain-computer interfaces (BCI) for motor control and rehabilitation.

The present work is an important step towards a robust analysis of complex systems based on recurrence.

Zusammenfassung

In unserem täglichen Leben spielt die Rekurrenz auf vielen räumlichen und zeitlichen Skalen und in verschiedenen Kontexten eine bedeutende Rolle. Es ist die Grundlage des Lernens, sei es in einem evolutionären oder in einem neuronalen Kontext. Es erscheint daher selbstverständlich, dass Rekurrenz auch ein grundlegendes Konzept in der dynamischen Systemwissenschaft ist. In diesem Zusammenhang ermöglicht die Art und Weise, wie sich Zustände eines Systems wiederholen oder sich auf ähnliche Weise aus ähnlichen Anfangszuständen entwickeln, Informationen über die zugrunde liegende Dynamik des Systems abzuleiten. Der mathematische Raum, in dem wir den Zustand eines Systems definieren (Zustandsraum), ist häufig hochdimensional, insbesondere in komplexen Systemen, die darüberhinaus auch eine chaotische Dynamik aufweisen können. Der Rekurrenzplot (RP) ermöglicht es uns, die Rekurrenzen beliebiger hochdimensionaler Systeme in einer zweidimensionalen, binären Darstellung zu visualisieren. Bestimmte Muster in RPs können mit physikalischen Eigenschaften des zugrunde liegenden Systems in Beziehung gesetzt werden, wodurch die qualitative und quantitative Analyse von RPs ein integraler Bestandteil der nichtlinearen Systemwissenschaft wird. Die vorgestellte Arbeit hat einen methodischen Schwerpunkt und entwickelt die Rekurrenzanalyse (RA) weiter, indem sie sich mit aktuellen Forschungsfragen befasst, die sich auf eine zunehmende Menge verfügbarer Daten und Fortschritte beim maschinellen Lernen beziehen. Durch die Automatisierung eines zentralen Schritts in der RA, nämlich der Rekonstruktion des Zustandsraums aus gemessenen experimentellen Zeitreihen, und durch die Untersuchung der Auswirkungen wichtiger freier Parameter soll die RA für Forscher außerhalb der Physik zugänglicher gemacht werden.

Der erste Teil dieser Dissertation befasst sich mit der Rekonstruktion des Zustandsraums aus Zeitreihen. Hierzu wird eine neue Idee vorgeschlagen, die das Rekonstruktionsproblem so automatisiert, dass weder die Daten vorverarbeitet noch *a priori* Parameter geschätzt werden müssen. Die Schlüsselidee ist, dass die *Güte* einer Rekonstruktion durch eine geeignete Kostenfunktion evaluiert werden kann und diese Funktion im Einbettungsprozess minimiert wird. Darüber hinaus kann die neue Methode multivariate Zeitreihen-Eingabedaten verarbeiten. Das ist insbesondere deshalb von großer Bedeutung, da mehrkanalige sensorgestützte Beobachtungen in vielen Forschungsbereichen allgegenwärtig sind und weiterhin zunehmen. Darauf aufbauend wird dann das beschriebene Minimierungsproblem der Kostenfunktion mit einem Ansatz des maschinellen Lernens bearbeitet.

Im zweiten Teil werden einige technische und methodische Aspekte der RA erörtert. Zunächst begründen wir mathematisch die Idee, den einflussreichsten freien Parameter in der RA, den Rekurrenzgrenzwert ε, in Bezug auf die Verteilung aller paarweisen Abstände in den Daten festzulegen. Dies ist insbesondere dann wichtig, wenn verschiedene RPs und ihre Quantifizierungsstatistiken verglichen werden, und ist für jede vergleichende Studie von grundlegender Bedeutung. Zweitens werden einige Aspekte der Rekurrenzquantifizierungsanalyse (RQA) untersucht. Als Korrekturschemata für verzerrte RQA-Statistiken, welche auf diagonalen Linien basierenden, schlagen wir eine einfache Methode zum Umgang mit Randeffekten von RPs in der RQA und einen Skeletonisierungsalgorithmus für RPs vor. Dies sorgt in der Folge zu weniger verzerrten (auf diagonalen Linien basierenden) RQA-Statistiken für hoch abgetastete Daten. Drittens wird eine neuartige RQA-Charakteristik entwickelt, die als verallgemeinertes, nichtlineares Leistungsspektrum hochdimensionaler Systeme angesehen werden kann. Das *Spike-Powerspectrum* transformiert ein Spike-train-ähnliches Signal in seinen Frequenzbereich. Wenn die diagonallinienabhängige Rekurrenzrate (τ-RR) eines RP auf diese Weise transformiert wird, können charakteristische Perioden, die in der Zustandsraumdarstellung des Systems erkennbar sind, entschlüsselt werden. Dies ist nicht der Fall, wenn die τ-RR Fourier-transformiert wird.

Schließlich werden RA und RQA im dritten Teil auf Paläoklima-Seesedimentdaten und im vierten Teil auf EEG-Daten (Elektroenzephalographie) angewendet. Nach unserem besten Wissen ist dies das erste Mal, dass RPs und RQA für die Analyse von Seesedimentdaten in einem Paläoklima-Kontext verwendet wurden. Daher wird zunächst an dem grundlegenden Formalismus und der Interpretation visuell sichtbarer Muster in RPs in Bezug auf die zugrunde liegenden Proxy-Daten gearbeitet. Wir zeigen, dass diese Muster verwendet werden können, um bestimmte Arten von Variabilität und Übergängen im Kaliumdatensatz von sechs kurzen (< 17m) Sedimentkernen zu klassifizieren, die während des Chew Bahir-Bohrprojekts gesammelt wurden. Darauf aufbauend wird der lange Kern (∼290 m composite) desselben Standorts analysiert und zwei Arten von Variabilität und Übergängen werden identifiziert und mit dem Feuchtigkeitsindex des ODP-Standorts 967 aus dem östlichen Mittelmeerraum verglichen. Die Variabilität vom Typ 1 spiegelt wahrscheinlich den Einfluss des Präzessionsantriebs in den unteren Breiten zu Zeiten mit Maximalwerten des langen Exzentrizitätszyklus (400 kyr) der Erdumlaufbahn um die Sonne wider, wobei die Tendenz zu extremen Ereignissen besteht. Die Variabilität vom Typ 2 scheint mit den lokalen Minima dieses Zyklus verbunden zu sein und entspricht ziemlich schnellen Übergängen zwischen relativ trockenen und relativ nassen Bedingungen.

Im Gegensatz dazu wird RQA seit fast zwei Jahrzehnten im neurowissenschaftlichen Kontext angewendet. Im letzten Teil werden RQA-Statistiken zur Quantifizierung der Komplexität in einem bestimmten Frequenzband multivariater EEG-Daten verwendet. Durch die Analyse experimenteller Daten kann gezeigt werden, dass die Komplexität des auf diese Weise über den sensomotorischen Kortex gemessenen Signals abnimmt, wenn motorische Aufgaben ausgeführt werden. Die Ergebnisse stimmen mit den bekannten Konzepten motorischer Gehirnprozesse überein und ergänzen diese. Wir nehmen an, dass die so entdeckten Merkmale der neuronalen Dynamik im sensomotorischen Kortex zusammen mit den robusten RQA-Methoden zur Identifizierung und Klassifizierung dieser zu der nicht-invasiven EEG-basierten Entwicklung von Gehirn-Computer-Schnittstellen (BCI) und zur motorischen Steuerung und Rehabilitation beitragen werden.

Die vorliegende Arbeit ist ein wichtiger Schritt zu einer robusten Analyse komplexer Systeme basierend auf Rekurrenz.

List of publications

This dissertation is largely based on the following publications.

Main contributions

(P1) **Kraemer, K.H.**, Datseris, G., Kurths, J., Kiss, I. Z., Ocampo-Espindola, J. L. and Marwan, N. (2021). A unified and automated approach to attractor reconstruction. New Journal of Physics, 23 (3), 133107. doi: 10.1088/1367-2630/abe336. [162]

(P2) **Kraemer, K. H.**, Donner, R. V., Heitzig, J. and Marwan, N. (2018). Recurrence threshold selection for obtaining robust recurrence characteristics in different embedding dimensions. Chaos: An Interdisciplinary Journal of Nonlinear Science, 28 (8), 085720. doi: 10.1063/1.5024914. [163]

(P3) **Kraemer, K. H.** and Marwan, N. (2019). Border effect corrections for diagonal line based recurrence quantification analysis measures. Physics Letters A, 383 (34), 125977, doi: 10.1016/j.physleta.2019.125977. [164]

(P4) Trauth, M. H., Asrat, A., Duesing, W., Foerster, V., **Kraemer, K. H.**, Marwan, N., Maslin, M. A., Schaebitz, F. (2019). Classifying past climate change in the Chew Bahir basin, southern Ethiopia, using recurrence quantification analysis. Climate Dynamics 53(5), 2557–2572, doi: 10.1007/s00382-019-04641-3. [304]

(P5) Trauth, M. H., Asrat, A., Duesing, W., Foerster, V., **Kraemer, K. H.**, Marwan, N., Maslin, M. A., Schaebitz, F. (2021). Recurring types of variability and transitions in the ~620 kyr record of climate change from the Chew Bahir basin, southern Ethiopia. Quaternary Science Reviews 266, 106777, doi: 10.1016/j.quascirev.2020.106777. [310]

(P6) Pitsik, E., Frolov, N., **Kraemer, K. H.**, Grubov, V., Maksimenko, V., Kurths, J. and Hramov, A. (2020). Motor execution reduces EEG signals complexity: recurrence quantification analysis study. Chaos: An Interdisciplinary Journal of Nonlinear Science, 30, 023111. doi: 10.1063/1.5136246. [245]

(P7) **Kraemer, K.H.**, Gelbrecht, M., Pavithran, I., Sujith, R. I. and Marwan, N.(2021). Optimal state space reconstruction via Monte Carlo Decision Tree Search. *Accepted* for Nonlinear Dynamics. [167]

(P8) **Kraemer, K. H.**, Hellmann, F., Kurths, J. and Marwan, N. (2021). Recurrence powerspectra (in prep.)

Further related papers of the author in collaboration

(P9) Düsing, W., Berner, N., Deino, A. L., Foerster, V., **Kraemer, K. H.**, Marwan, N. and Trauth, M. H. (2021). Multiband wavelet age modeling (MUBAWA) for a ~293 m (~600 kyr) sediment core from Chew Bahir basin, southern Ethiopian Rift. Frontiers in Earth Science 9, doi: 10.3389/feart.2021.594047. [75]

(P10) Foerster, V., Deocampo, D. M., Asrat, A., Günter, C., Junginger, A., **Kraemer, K. H.**, Stroncik, N. A. and Trauth, M. H. (2018). Towards an understanding of climate proxy formation in the Chew Bahir basin, southern Ethiopian Rift. Palaeogeography, Palaeoclimatology, Palaeoecology, 501, 111-123. doi: 10.1016/j.palaeo.2018.04.009. [91]

Author contributions

(P1) K. Hauke Kraemer: Conceptualization, Methodology, Software, Formal analysis, Writing–Original Draft; George Datseris: Software, Validation, Writing–Review; Jürgen Kurths: Validation, Writing–Review; Istvan z. Kiss: Experimental analysis; Jorge L. Ocampo-Espindola: Experimental analysis; Norbert Marwan: Validation, Methodology, Software, Visualization, Writing–Original Draft.

(P2) K. Hauke Kraemer: Conceptualization, Methodology, Software, Formal analysis, Writing–Original Draft; Reik V. Donner: Methodology, Writing–Original Draft; Jobst Heitzig: Mathematical derivation; Norbert Marwan: Validation, Methodology, Visualization, Writing–Original Draft.

(P3) K. Hauke Kraemer: Conceptualization, Methodology, Software, Formal analysis, Writing–Original Draft; Norbert Marwan: Validation, Methodology, Visualization, Writing–Original Draft.

(P4) Martin H. Trauth: Conceptualization, Methodology, Software, Formal analysis Writing–Original Draft, Funding acquisition; Asfawossen Asrat: Conceptualization, Writing–Review & Editing, Funding acquisition; Walter Duesing: Validation, Writing–Review & Editing; Verena Foerster: Investigation, Writing–Review & Editing; K. Hauke Kraemer: Conceptualization, Methodology, Software, Formal analysis, Visualization, Writing–Original Draft; Mark A. Maslin: Investigation, Writing–Review & Editing; Norbert Marwan: Methodology, Software, Formal analysis, Visualization, Writing–Original Draft; Frank Schäbitz: Conceptualization, Writing–Review & Editing, Funding acquisition.

(P5) Martin H. Trauth: Conceptualization, Methodology, Software, Formal analysis Writing–Original Draft, Funding acquisition; Asfawossen Asrat: Conceptualization, Writing–Review & Editing, Funding acquisition; Andrew S. Cohen: Conceptualization, Writing–Review & Editing, Funding acquisition; Walter Duesing: Validation, Writing–Review & Editing; Verena Foerster: Investigation, Writing–Review & Editing; Stefanie Kaboth-Bahr: Validation, Investigation, Writing–Review & Editing; K. Hauke Kraemer: Methodology, Software, Formal analysis, Visualization, Writing–Original Draft; Henry F. Lamb: Conceptualization, Writing–Review & Editing, Funding acquisition; Norbert Marwan: Methodology, Software, Formal analysis, Writing–Original Draft; Mark A. Maslin: Investigation, Writing–Review & Editing; Frank Schäbitz: Conceptualization, Writing–Review & Editing, Funding acquisition.

Acknowledgements

I also thank Dr. George Datseris for introducing me into the world of software engineering and for answering every question I raised within no time.

Furthermore I would like to thank Dr. Joscha Hausam, Maximilian Gelbrecht, Tobias Braun, Walter Düsing and Jan-Eike Krämer for proofreading parts of this thesis. In particular Joscha, you kept going until the very end and I am deeply indebted to your remarkable effort.

Of course, my parents, my brother and my friend Tilman cannot be missing here either. It's nice that we have such a beautiful network of relationships, which gives me enormous stability in my life. I thank my brother for intellectual discussions about things that often have nothing to do with intellect. I am indebted to Tilman in many ways, but here I want to emphasize that he is there for my son Kolja, which in turn gave him and also myself time and a good feeling.

As always, the best comes at the end. Without you Marie-Luise, probably nothing would have happened. Or it wouldn't have been so nice. Or I would have perished alone. You and our son Kolja really made this project possible and you supported me no matter what. You gave me support in the difficult phases. You gave me meaning and strength, especially in the last – COVID-19 embossed – year. Thank you for everything, but above all for bringing me closer to the importance of the emotional level in human relationships of all kinds.

Contents

Contents

Contents

List of Figures

List of Tables

List of Tables

List of abbreviations and symbols

α	significance level
Λ_τ	a delay pre-selection statistic for obtaining potential delay value in an embedding cycle
Δt	sampling time
$\mathcal{D}$	embedding cycle
$\mathfrak{D}$	distance range: connected areas of distances less than ε in a distance Matrix
ε	the recurrence threshold for obtaining a RP from a distance Matrix $\mathbf{A}$ (neighborhood size in state space)
$\langle \varepsilon^\star \rangle (\tau)$	continuity statistic
Γ	objective function; cost-function
λ	Lyapunov exponent
Λ_τ	embedding delay pre-selection statistic
$\mathcal{L}_\ell$	a diagonal line of length ℓ
$\mathcal{T}$	measure for recurrence network quantification: transitivity
τ	a time delay value for a delay embedding procedure; in the context of the τ-recurrence rate (Chapter 6) it denotes the distance to the main diagonal of the RP
τ-RR	measure for recurrence quantification: diagonal-wise recurrence rate τ steps apart from the main diagonal of the RP
$\mathbf{A}$	adjacency matrix
AHP	African Humid Period
BCI	brain computer interface
d	a distance under a chosen metric
$\mathbf{d}$	distance matrix storing all pairwise distances $d_{i,j}$ between trajectory points i, j

DET	measure for recurrence quantification: determinism
D_B	box-counting dimension
DO	Dansgaard-Oeschger events/cycles
CBDP	Chew Bahir Drilling Project
CHB	Chew Bahir
CS	continuity statistic
EEG	electroencephalography
EMG	electromiogram
ENTR	measure for recurrence quantification: diagonal line length entropy
ENSO	El Niño-Southern Oscillation
ERD	event-related desynchronization
FAN	fixed amount of nearest neighbors
FNN	false nearest neighbors
HEs	Heinrich Events
HSPDP	Hominin Sites and Paleolakes Drilling Project
JR	joint recurrence plot
JRRF	joint recurrence rate fraction (accordance between two RPs)
k	number of nearest neighbors
K	potassium
K_2	correlation entropy
KL, KLD	Kullback-Leibler, Kullback-Leibler divergence
ℓ	length of a diagonal line in a RP
LAM	measure for recurrence quantification: laminarity
LGM	Last Glacial Maximum
L_2	Euclidean norm
L_∞	Maximum norm
m	embedding dimension
MCDTS	Monte Carlo Decision Tree Search
MDOP	an embedding approach proposed by Nichkawde [231]
MFFN	mutual false nearest neighbors
MIS	marine isotope stage

MRT	measure for recurrence quantification: mean recurrence time (average of the white vertical lines)
MSE	mean squared prediction error
$P(\ell)$	histogram of diagonal line lengths
$\mathbb{R}$	set of real numbers
RA	Recurrence Analysis
RMS	root mean square prediction error
RN	recurrence network
RQA	recurrence quantification analysis
RR	measure for recurrence quantification: recurrence rate
$\mathbf{R}$, RP	recurrence matrix, recurrence plot
RTE	measure for recurrence quantification: recurrence time entropy (entropy of white vertical lines)
$s(t)$	a scalar time series
SST	sea-surface temperature
T	prediction horizon; in Chapter 5 a parameter for the isodirectional RP
TDE	time delay embedding
$\vec{v}(t)$	state space (reconstruction) vectors
w	window length
ws	step size in a running window analysis
YD	Younger Dryas

Part

Introduction

Chapter 1

Introduction

"Everything goes, everything comes back; eternally rolls the wheel of being.
Everything dies, everything blossoms again; eternally runs the year of being.
Everything breaks, everything is joined anew; eternally the same House of Being is built.
Everything parts, everything greets every other thing again;
eternally the ring of being remains faithful to itself."

Friedrich Nietzsche

1.1 Background and motivation

The cyclic understanding of time is a basic concept which constituted many ideas in Friedrich Nietzsches philosophy ("*Ewige Wiederkunft*"). Not only does recurrence influence philosophers, it is an ubiquitous phenomenon in our daily life. Obviously, day and night are recurrent. But also states of complex chaotic systems like the weather or more abstract things like a collective mood on a festival with thousands of people seem to recur. Life on earth, and probably elsewhere, has developed under recurrent external conditions. This type of regularity is fundamental to learning, be it in an evolutionary or in a neural context associated with living beings. Some situations occur over and over again, and other situations do not occur in exactly the same way, but rather are similar in their course to previously observed situations. Both phenomena allow to make predictions of the future based on past observations. This applies both to regular deterministic systems such as the motion of the earth in the orbit of the sun, which obeys Newton's laws of gravity, and to deterministic chaotic systems like a current in a river, which follows the Navier-Stokes equation. The difference is the time scale of the prediction horizon and the observation time needed for a certain prediction accuracy [13, 233]. Nonlinear, complex systems often exhibit chaotic behavior and long term predictions will fail due to the sensitivity to initial conditions. But even then short-term predictions are possible [146]. The common feature is their deterministic structure, reflected in the governing equations. These (partial nonlinear) differential equations describe the dynamical evolution of the systems they describe and, hence, their recurring features.

Recurrence is indeed a fundamental property of many dynamical systems. Mathematically, this is formalized in the Poincaré recurrence theorem, which states that an autonomous, Hamiltonian and bounded system "*recurs infinitely many times as close as one wishes to its initial state*"

[246]. This statement seems to contradict the second law of thermodynamics, which postulates that the thermodynamic Boltzmann entropy is a monotonic function in closed systems. Ernst Zermelo argued that a function (the entropy as a function of time) can not be both, recurrent and monotonic [233]. This led Ludwig Boltzmann to respond in a rather statistical, probabilistic view of the second law, which eventually incorporates the anthropic principle. On a cosmological scale Carroll and Chen [41] argued that the *system universe* is not bounded and, thus, Poincaré's recurrence theorem would not be applicable. Despite these (ongoing) discussions on a cosmological scale, the assumptions can be fulfilled for a wide range of problems.

Due to its fundamental nature, recurrence can be used in the description, analysis and modeling of systems [43, 209]. Within the emerging field of chaos and complexity science in the last third of the past century, recurrence analysis (RA) occupied a niche in nonlinear physics and time series analysis. Applications of RA can be found in various fields of research ranging from Astrophysics, Engineering, Geo- and Earthscience to Neuroscience, Biology or Economics. It has been used for classification tasks, testing stationarity, prediction, detection of causality and synchronization and for chaos control[1]. However, RA is not a trivial tool to use in any arbitrary statistical analysis. Recall that Poincaré's recurrence theorem only holds for Hamiltonian systems, which implies an appropriate representation of a system in its phase or *state space*. This is typically not what experimenters measure directly. Moreover, for most systems this state space is high dimensional, which raises the question of how to keep track of recurrences occurring in that space.
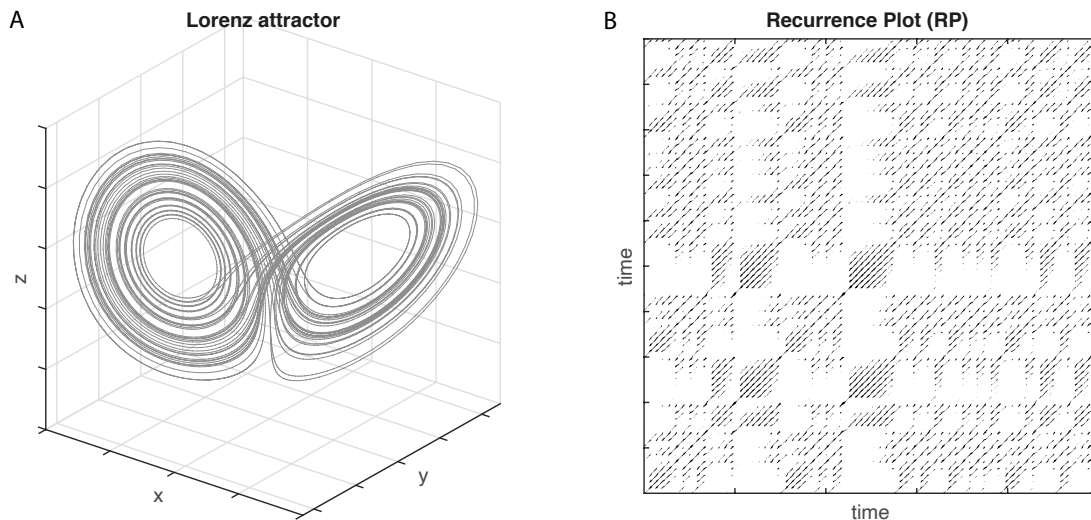
Figure 1.1: A The famous "*Butterfly*" Lorenz attractor (Eq. (B.1) with $\sigma = 10$, $\rho = 28$, $\beta = 8/3$) in state space and **B** its recurrence plot representation (5 % global recurrence rate).

Recurrence plots (RPs), introduced by Eckmann et al. [77], provide a versatile tool for visualizing and quantitatively analyzing the succession of dynamically similar states of a system. Even complex and chaotic systems can be transformed into such a 2-dimensional representation as shown in Figure 1.1 for the famous Lorenz butterfly. RPs are the foundation of RA, which is typically concerned with their analysis. This can be a visual inspection of the encoded structures

---

[1] A comprehensive and profound overview of published articles related to RA can be found at `http://www.recurrence-plot.tk/bibliography.php`

or a quantification of those, the so called *Recurrence Quantification Analysis* (RQA). By recalling the above described fundamental nature of recurrence and the idea of plotting recurrences in a simple, binary way it seems natural to ask the following questions:

- Can RA be used for any sort of system we want to analyze? In particular, can the state space of the system be obtained from only one or a few measurement time series?

- Can complex systems such as the earth's climate or the human brain be better understood by using RA?

- Is it possible to physically interpret structures and patterns in a recurrence plot? Is there some sort of recipe that relates these patterns to the properties of the underlying state space trajectory or time series?

The answers to these questions are clearly Yes *and* No. RA is an active field of research and this dissertation seeks to answer some of the many open questions that arise from these questions. A non-negligible point that motivated me to further explore RA is the fact that even extremely complicated-looking time series can be embedded in a RP, which impresses with its simplicity and beauty with a wealth of information at the same time. The question of how the topology of occasional high dimensional state spaces is encoded in quantifiable patterns in a RP is especially appealing to me.

This work focuses on the methodological aspect of RA, which I will elaborate on in the following section and which will eventually lead to the formulation of the guiding research questions of this thesis in Section 1.3. Yet, it is also interdisciplinary and motivated by direct applications in Paleoclimatology (the Chew Bahir Drilling Project), Engineering and Neuroscience. Emerging questions from applications in these fields have strongly influenced and steered the methods developed in the framework presented here.

1.2 The basics

In order to visualize *similar states of a system* one has to define what similarity actually means in this context. For this purpose, dynamical similarity is measured in terms of some metric distance $d_{i,j} = \|\vec{x}_i - \vec{x}_j\|$ defined in the underlying system's d-dimensional state space. The vector time series $\{\vec{x}_i\}_{i=1}^N$ (with $\vec{x}_i = \vec{x}(t_i)$) can be obtained from direct measurements of the system under study or from embedding a measured subset of observables.

Based on the resulting distance matrix $\mathbf{d} = (d_{i,j})$, a recurrence matrix $\mathbf{R} = (\mathbf{R}_{i,j})$ is defined by thresholding $\mathbf{d}$ such that its entries assume values of 1, if the distance between the two associated state vectors is smaller than or equal to a threshold ε, and 0 otherwise. It is, thus, a binary, square matrix:

$$\mathbf{R}_{i,j}(\varepsilon) = \begin{cases} 1: & d_{i,j} \leq \varepsilon \\ 0: & d_{i,j} > \varepsilon, \end{cases} \qquad i,j = 1,...,N \tag{1.1}$$

Equivalently, we can write

$$\begin{aligned} \mathbf{R}_{i,j}(\varepsilon) &= \Theta\left(\varepsilon - \|\vec{x}_i - \vec{x}_j\|\right) \\ &= \Theta(\varepsilon - d_{i,j}), \qquad i,j = 1,...,N, \qquad \vec{x} \in \mathbb{R}^d, \end{aligned} \tag{1.2}$$
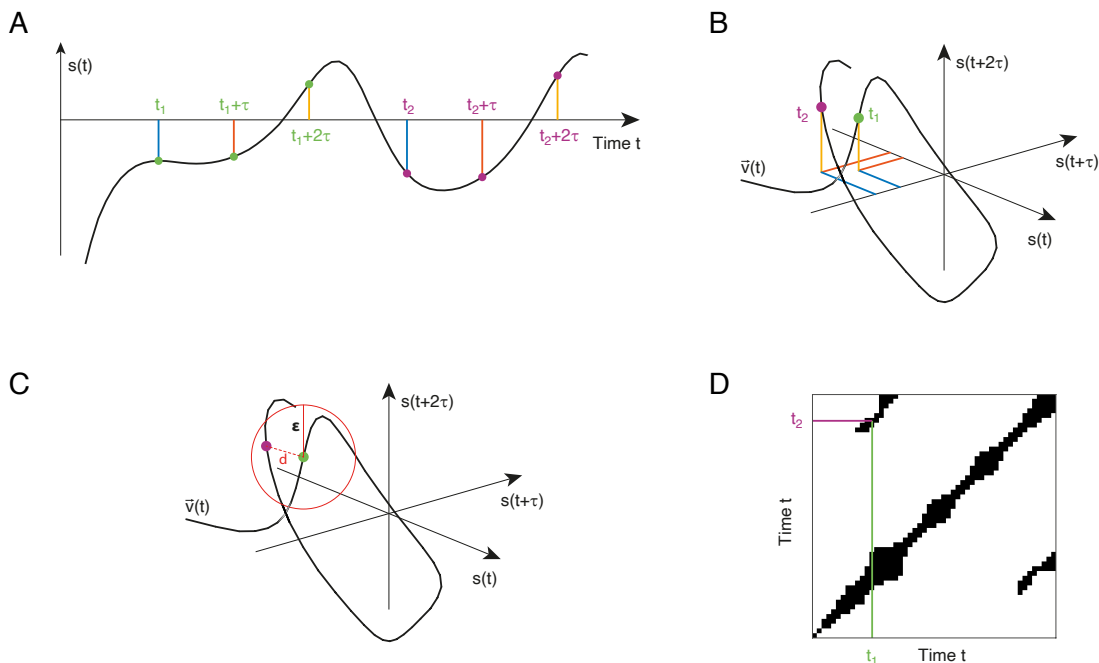
Figure 1.2: Principle of recurrence plot. One way to untangle the dynamics of a multi-dimensional system from a one-dimensional time series $s(t)$ (panel **A**) is by time-delay embedding. The embedding of the time series $s(t)$ in a three-dimensional coordinate system for example (a reconstructed *state* or *phase space*), means that three successive values $s(t)$, $s(t + \tau)$ and $s(t + 2\tau)$ with a temporal separation of τ are represented by a single point $\vec{v}(t)$ within the phase space [138, 235] (panel **B**). We picked a three dimensional embedding here for didactic and visualization reasons. Estimating the correct embedding dimension is an active research topic. Recurrence plots (RPs, panel **D**), first introduced by Eckmann et al. [77], are graphic displays of recurring states within a system, computed from the distance (e.g. the Euclidean distance d, shown in panel **C**) between all pairs of phase space vectors $\vec{v}(t_i)$ and $\vec{v}(t_j)$, below a threshold value ε (also shown in panel **C**) [209].

where $\Theta(\cdot)$ is the Heaviside function. In this definition, the threshold ε is fixed with respect to all pairwise distances contained in **d**. An alternative definition of the recurrence matrix [77, 209] replaces the global, fixed recurrence threshold ε applied to all state vectors $\vec{x}_i$ by an adaptive local one. Here the idea is that the number of recurrences (i.e., close state vectors) is the same for each $\vec{x}_i$ (fixed amount of nearest neighbors - FAN), which leads to a constant local (each column) and global (whole RP) recurrence rate. This procedure is useful, when there are severe differences in the density of points in state space. In this case lower dense populated areas in the state space are not well resolved in the RP, when ε is not adaptively chosen. Of course, this can be simply a result of under-sampling the system. On the other hand, strong density variations in phase space can be a meaningful information or it is necessary to ensure a fixed threshold, when comparing RPs stemming from different systems, or from the same system, but at different dynamics. The research question and the quality of the data decides whether to use an adaptive recurrence threshold (e.g., Chapters 7, 8) or a fixed one (e.g., Chapter 9).

In many, if not most, real world cases the vector time series $\vec{x}$ describing the system's trajectory in state space is not directly measurable and there is only access to one measured variable. The most common way to untangle the dynamics of a multi-dimensional system from a one-dimensional

time series $s(t)$ is by time-delay embedding, where the key idea is to use lagged values of the available time series as components of the reconstruction vector $\vec{v}(t)$.

$$\vec{v}(t) = (s(t + \tau_1), s(t + \tau_2), \ldots, s(t + \tau_m)). \tag{1.3}$$

For sufficient choices of the parameters $\tau_{1,\ldots,m}$ this representation preserves the dynamic characteristics of the true, but unknown system [235] (Fig. 1.2). The reason why an entire system needs to be reconstructed from a single variable is that information about the system and the factors affecting its state variables is contained in a one-dimensional time series, due to the coupling of the ordinary differential equations describing the system. The embedding of the time series $s(t)$ in a three-dimensional coordinate system (a reconstructed *phase* or *state space*), for example, means that three successive values $s(t)$, $s(t + \tau)$ and $s(t + 2\tau)$ with a temporal separation of τ are represented by a single point within the state space (Fig. 1.2). The geometric representation of the embedded time series of observations as trajectories $\vec{v}(t)$ within the reconstructed phase space is called a phase portrait. The reconstructed state space is not exactly the same as the original phase space, but its topological properties are preserved provided the embedding dimension is sufficiently large [235, 291].

1.3   Scope and objectives

The briefly discussed basics in the previous section together with the general questions related to the method of Recurrence Analysis raised in Section 1.1 define the scope of the current research presented in this thesis. We investigate the following research questions:

(R1) Are the existing approaches to state space reconstruction, using delayed shifted versions of the time series, robust, extendable to multivariate input and automatable?

(R2) Can we better understand abrupt transitions and regime changes to provide measures for early warning by using Recurrence Analysis (RA)?

(R3) How can results from recurrence quantification analysis (RQA) be evaluated from a statistical point of view and how can we ensure comparability when using RQA on different samples/ systems?

(R4) Can RA be used for classifying climate states, and therefore also climate transitions?

(R5) Can RA be used for classifying neuronal activity related to certain motor tasks?

Obviously, these five research questions are strongly interwoven and imply many related questions. Therefore, they will also be commonly addressed in the parts and chapters of this work.

1.4   Organization of the thesis

The main contribution of this thesis is based on eight articles, which make up the single chapters. Five articles are published, one is accepted and two are about to be submitted. Each article is a stand-alone contribution with an introduction, a methods-, results- and conclusion part. Different from the published versions of the articles, common definitions are summarized and presented in

Appendix A, the formatting and design were adjusted to the style of this thesis and a common bibliography has been compiled. The work is divided into four main parts, which consist of the mentioned articles (P1)-(P8) as chapters, illustrated in Figure 1.3, and a summarizing part at the end.

Part I focuses on the problem of state space reconstruction from observable time series, (R1). In Chapter 2 a novel approach is presented, the PECUZAL algorithm, which automates the reconstruction problem in the sense that neither the data has to be pre-processed nor parameters have to be estimated *a priori*. The key idea is that the "goodness" of a reconstruction can be assessed by an appropriate objective function and this function is minimized through the embedding process. In addition, it can handle multivariate time series input data, which we consider a valuable feature as multi-channel sensor-based observations are ubiquitous in many research areas. In Chapter 3 we take up the PECUZAL idea directly and tackle the minimization problem of the objective function with a machine learning approach. While PECUZAL can get stuck in a local minimum of the objective function, the embedding problem is formalized as a Monte Carlo Decision Tree Search, in order to reach the global minimum.

In Part II some technical and methodological aspects of Recurrence Analysis are discussed, in order to shed light to (R2) and (R3). Chapter 4 mathematically justifies the idea of fixing the recurrence threshold ε with respect to the distribution of all pairwise distances in the data. This is especially important, when comparing different RPs and their quantification statistics, as was done in an application presented in Chapter 9. This way of selecting the most important free parameter in Recurrence Analysis is fundamental for any comparative study. However, the concrete value of ε, i.e. the percentile of the distance distribution, depends on the research question. Chapter 5 and 6 are concerned with some aspects of RQA. While correction schemes for the biased diagonal line based RQA statistics are presented in Chapter 5, a new type of powerspectrum, the spike powerspectrum, is introduced in Chapter 6. As correction schemes we propose a simple way of dealing with border effects of an RP in RQA and a skeletonization algorithm for RPs, which leads to less biased (diagonal line based-) RQA statistics for flow-like data. The spike powerspectrum, on the other hand, adequately transforms a spike-train like signal into its frequency domain. When transforming the diagonal line dependent recurrence rate (τ-RR) of a RP in that way, characteristic periods that are apparent in the state space representation of the system can be unraveled. Hence, it can be regarded as a generalized, non-linear powerspectrum based on a system's state space trajectory. This is not the case, when Fourier transforming τ-RR. Consequently, the spike powerspectrum of a system's τ-RR can be utilized for the detection of transitions and bifurcations.

Recurrence Analysis and RQA are applied to paleoclimate lake sediment data in Part III and face (R4). To our best knowledge this is the first time RPs and RQA have been used for the analysis of lake sediment data in a paleoclimate context. Therefore Chapter 7 can be seen as an introductory paper discussing the basic formalism and the interpretation of visually apparent patterns in RPs with respect to the underlying proxy data (which is a Potassium time series in this case). We have shown that these patterns can be used to classify certain types of variability and transitions in the Potassium record from six short (< 17m) sediment cores collected during the Chew Bahir Drilling Project. Building on that, the long ($\sim$290 m composite) core from

Towards a robust framework for Recurrence Analysis

I **II** **III/IV**

State space reconstruction RP/RQA methodology Application

Climate Science

(P1) A unified and automated approach to attractor reconstruction.
Kraemer, K. H., Datseris, G., Kurths, J., Kiss, I. Z., Ocampo-Espindola, J. L. and Marwan, N.
New Journal of Physics 23(3), (2021)

(P7) Optimal state space reconstruction via Monte Carlo Decision Tree Search.
Kraemer, K.H., Gelbrecht, M., Pavithran, I., Sujith, R. I. and Marwan, N.
(accepted for Nonlinear Dynamics)

Lorem ipsum

(P2) Recurrence threshold selection for obtaining robust recurrence characteristics in different embedding dimensions.
Kraemer, K. H., Donner, R. V., Heitzig, J. and Marwan, N.
Chaos: An Interdisciplinary Journal of Nonlinear Science, 28 (8), (2018)

(P3) Border effect corrections for diagonal line based recurrence quantification analysis measures.
Kraemer, K. H. and Marwan, N.
Physics Letters A, 383 (34), (2019)

(P8) R**ecurrence powerspectra.**
Kraemer, K. H. , Hellmann, F., Kurths, J. and Marwan, N.
(in prep. 2021)

(P4) Classifying past climate change in the Chew Bahir basin, southern Ethiopia, using recurrence quantification analysis.
Trauth, M. H., Asrat, A., Duesing, W., Foerster, V., Kraemer, K. H., Marwan, N., Maslin, M. A., Schaebitz, F.
Climate Dynamics 53(5), (2019)

(P5) Recurring types of variability and transitions in the ~620 kyr record of climate change from the Chew Bahir basin, southern Ethiopia.
Trauth, M. H., Asrat, A., Duesing, W., Foerster, V., Kraemer, K. H., Marwan, N., Maslin, M. A., Schaebitz, F.
accepted for Quaternary Science Reviews, (2021)

Neuroscience

(P6) Motor execution reduces EEG signals complexity: recurrence quantification analysis study.
Pitsik, E., Frolov, N., Kraemer, K. H., Grubov, V., Maksimenko, V., Kurths, J. and Hramov, A.
Chaos: An Interdisciplinary Journal of Nonlinear Science, 30, (2020)

Figure 1.3: Overview of the sub-projects in this thesis divided into the three main parts *State space reconstruction* (left column), *RP/RQA methodology* (middle column) and *Application* (right column, divided into Climate Science and Neuroscience).

the same site is analyzed in Chapter 8. Two types of variability and transitions have been distinguished and were compared to the ODP Site 967 wetness index from the eastern Mediterranean.

In contrast, RQA has been applied in the neuroscientific context for almost two decades. In Chapter 9 of Part IV RQA statistics have been used for quantifying *complexity* in a certain frequency band of multivariate EEG-data (electroencephalography) and therefore able to classify motor-related brain processes. The active experiment setup allowed for addressing (R3) and (R5) simultaneously.

In Part V, Chapter 10 summarizes the main contributions of this thesis and discusses several open issues that have arisen in the conduct of this research. The Appendix (Part VI) includes a short overview of Recurrence Analysis and RQA (Appendix A), the models used in the different Chapters (Appendix B) as well as supplementary material (Appendices C, D, E, F, G H).

Part I

State space reconstruction

Chapter 2

A unified an automated approach to attractor reconstruction

Abstract

We present a fully automated method for the optimal state space reconstruction from univariate and multivariate time series. The proposed methodology generalizes the time delay embedding procedure by unifying two promising ideas in a symbiotic fashion. Using non-uniform delays allows the successful reconstruction of systems inheriting different time scales. In contrast to the established methods, the minimization of an appropriate cost function determines the embedding dimension without using a threshold parameter. Moreover, the method is capable of detecting stochastic time series and, thus, can handle noise contaminated input without adjusting parameters. The superiority of the proposed method is shown on some paradigmatic models and experimental data from chaotic chemical oscillators.

2.1 Introduction

State space reconstruction from observational time series often marks the first and basic step in nonlinear time series analysis. Several methods addressed the reconstruction problem, but none of them allow for a fully automatized and reliable way of embedding a uni- or multivariate set of observed time series with no, or at least very few, free parameters. The aim of this Chapter is to provide such a technique. The embedding theorems of Whitney [329], Mañé [196], and Takens [291] among with their extension by Sauer et al. [268] allow several approaches to tackle the reconstruction problem. Among using derivative coordinates [27, 197], PCA [235] or Legendre coordinates [111], uniform- and nonuniform time delay embedding [235] is by far the most commonly used technique, due to its appealing simplicity. However, since Takens' theorem [291] is based on noise-free and infinitely long data, it does not give any guidance to choose
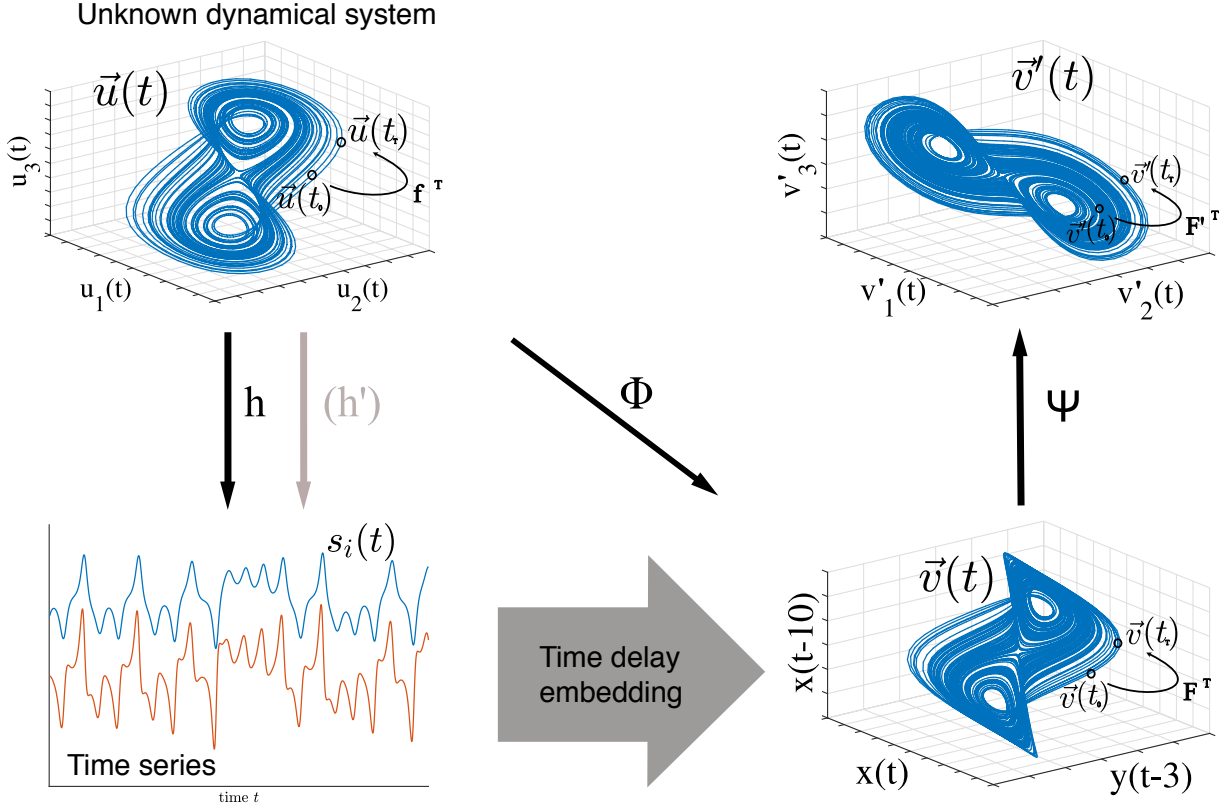
Figure 2.1: Schematic representation of the embedding/reconstruction procedure. See the text for details. This figure is inspired by Casdagli et al. [45] and Uzal et al. [315].

the proper time delay(s) τ in practice. Together with the unknown box-counting dimension D_B of the observed, but unknown system, which is needed to fulfill the embedding dimension criterion $m \geqslant 2D_B + 1$, the majority of the published articles propose ideas to infer estimates for τ and the reconstruction dimension m from data, usually a univariate time series (univariate embedding). The reconstruction problem starts with the unknown system $\vec{u}(t)$ with a mapping $f : \mathbb{R}^{D_B} \to \mathbb{R}^{D_B}$, which is observed via a measurement function h and lead to M observational time series $\{s_i(t)|i = 1,\ldots,M\}$ (Fig. 2.1). There can be different measurement functions h' forming the multivariate dataset $s_i(t)$ and the combination of Whitney's and Takens' embedding theorems allow for constructing $\vec{u}(t)$ from more than one time series (multivariate embedding) [67, 268]. One then tries to find optimal embedding parameters m and τ's (the delays can be integer multiple of some constant, uniform time delay embedding *UTDE*, or different for each embedding dimension, non-uniform time delay embedding *NUTDE*) in order to build reconstruction vectors $\vec{v}(t)$ and, thus, a mapping $F : \mathbb{R}^m \to \mathbb{R}^m$. These can be furthermore transformed by Ψ into $\vec{v}'(t)$ and F', preserving the diffeomorphism to $\vec{u}(t)$. For a detailed introduction into the reconstruction problem we refer to Casdagli et al. [45], Gibson et al. [111], Uzal et al. [315] or Nichkawde [231].

In time delay embedding the key idea is to use lagged values of the available time series as components of the reconstruction vector

$$\vec{v}(t) = (s_{i_1}(t - \tau_1), s_{i_2}(t - \tau_2), \ldots, s_{i_m}(t - \tau_m)). \tag{2.1}$$

Here the delays τ_j, $j = 1, \ldots, m$ are multiples of the sampling time Δt and the indices $i_1, i_2, \ldots, i_m$ each denote the time series index $i \in [1, \ldots, M]$, which has been chosen in the 1st, 2nd, $\ldots$, m^{th} embedding cycle. The delays τ_j and the corresponding time series s_{i_j} should be chosen, such that the resulting components of the vectors forming the reconstructed state space $\vec{v}(t)$ are as independent as possible [235, 268], but at the same time preserve the correlation structure of the system to a certain extent. These two competing objectives are also known as the problems of *redundancy* (delays should not be too small) and *irrelevance* (delays should not be too large) [45, 79, 260, 315] and its optimization is the goal to any proposed time delay embedding procedure.

Despite of its lack of a sound theoretical foundation for higher dimensional reconstructions ($m > 2$) [94, 119, 317], in a univariate scenario (i.e., $s_{i_1} = \ldots = s_{i_m} = s(t)$ in Eq. (2.1)), the approach to choose τ_2 from the first minimum of the auto-mutual information [95, 186] is most common. τ_1 is usually set to zero, i.e. the unlagged time series constitutes the first component of the reconstruction vectors. The embedding dimension m is then separately determined, usually by a false nearest neighbor approach [39, 125, 153, 154, 170] or some other neighborhood-preserving-idea [10, 49] and all delays up to τ_m are simply integer multiples of τ_2 (UTDE). Other approaches for an appropriate choice of τ_2 are possible [5, 8, 9, 27, 40, 152, 173, 260]. We refer to this as standard *time delay embedding (TDE)* in the following. More sophisticated ideas [33–35, 102, 103, 141, 152, 155, 185, 241, 284, 317], some including non-uniform delays and the extension to multivariate input data [104, 105, 123, 131, 132, 231, 240] have been presented [123, 130], but it seems their use is rather limited and not very popular. This could be due to their occasionally very complex nature and the lack of high quality open source implementation in the most commonly used programming languages. Another reason could be the fact that standard TDE performs surprisingly well in a range of examples; but still, its limitations should not be neglected, in particular when it comes to noisy time series, systems exhibiting multi-time scale behavior, or multivariate input data. The latter are becoming of increasing interest in the near future, since acquisition costs for sensors and data collection decrease rapidly. Moreover, the application of complex systems approaches and nonlinear dynamics in different scientific disciplines receives increasing popularity.

Here we propose a fully automated method for an appropriate state space reconstruction of uni- or multivariate time series input, which utilizes two concepts, the *continuity statistic* [240] and the *L-statistic* [315]. We briefly review the basic ideas we will use in our proposed method (Section 2.2) in order to illustrate their specific utilization in the algorithm (Section 2.3), before applying it to simulated and experimental data (Section 2.4).

2.2 Review of used concepts

In order to ensure comprehensibility of our proposed method in Section 2.3 we explain the two main concepts of it in the following. In Section 2.2.1 we review the *continuity statistic* [240] rather detailed, while the *L-statistic* is described only briefly in Section 2.2.2 and extensively in the Appendix C.4.

2.2.1 Continuity statistic by Pecora et al.

In the *continuity statistic*, the problem of finding an optimal state space reconstruction with respect to *redundancy* and *irrelevance* is addressed by a statistical determination of functional
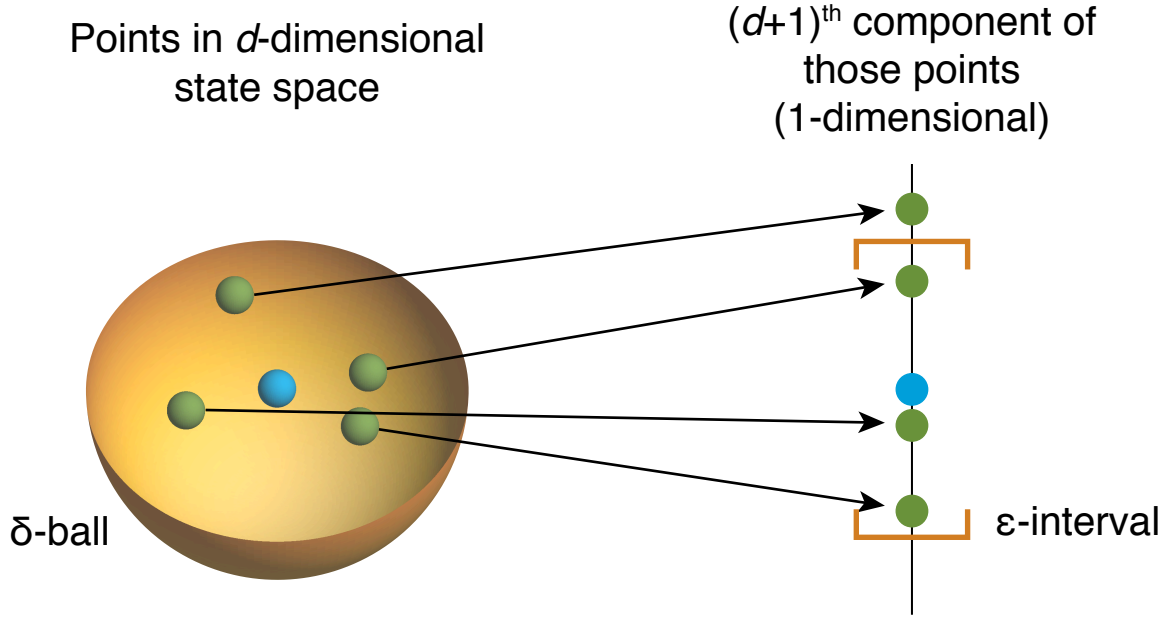
Figure 2.2: Fiducial point (blue) and its $k = 4$ nearest neighbors (green) in the d-dimensional δ-ball (left panel). Arrows indicate the mapping $f : \mathbb{R}^d \to \mathbb{R}^1$ (right panel), which is the potential $(d + 1)^{\text{th}}$ component of each of the points in the left panel (according to a specific delay τ_{d+1} and time series $s_{j_{d+1}}$), Eq. (2.2). To decide whether this ε-interval size accepts or rejects the null hypothesis on a significance level α the cumulative binomial distribution for getting at least $l = 3$ points in the ε-interval with probability p is used (modified after Pecora et al. [240]).

independence among the components of the reconstruction vector [240]. Let $\{s_i(t)|i = 1, \ldots, M\}$ be a multivariate dataset consisting of M time series, equally sampled at time instants $t = 1, \ldots, N$. Suppose we have already chosen some delays τ_k to build our temporal reconstruction vector $\vec{v}(t)$ of dimension d. This is, $\vec{v}(t) = (s_{j_1}(t + \tau_1), s_{j_2}(t + \tau_2), \ldots, s_{j_d}(t + \tau_d))$, with $j_k \in \{1, \ldots, M\}$ being the choices of the different time series and τ_k the according delays, which can be – and most often are – different. Then for a new potential component of $\vec{v}(t)$ we ask if this new potential component can be expressed as a function of the existing components. Mathematically speaking, the equality

$$s_{j_{d+1}}(t + \tau_{d+1}) \stackrel{?}{=} f\left(s_{j_1}(t + \tau_1), s_{j_2}(t + \tau_2), \ldots, s_{j_d}(t + \tau_d)\right) \tag{2.2}$$

needs to be tested in an appropriate way, i.e., a sensible choice for $f : \mathbb{R}^d \to \mathbb{R}^1$ has to be made. This choice can be based on the property of continuity [239].

The practical implementation of Eq. (2.2) would start with mapping k nearest neighbors, $\vec{v}_k(t)$, of a fiducial point $\vec{v}_{\text{fid}(t)}$ from $\mathbb{R}^d \to \mathbb{R}^1$, as illustrated in Fig. 2.2. That is, for each of the $(k + 1)$ d-dimensional points in the left panel a potential $(d+1)$th component $s_{j_{d+1}}(t + \tau_{d+1})$ is considered and drawn onto the number line (right panel). The *continuity statistic* now asks whether these $k + 1$ points on the 1-dimensional number line fall within a certain ε-interval size by chance, or due to the fact that there is a functional relationship between the d and the $(d + 1)$th component of a potential reconstruction vector $\vec{v}(t)$. The according null hypothesis is that l of the $k + 1$ points landed in the ε-interval by chance, with probability p on the basis of the binomial distribution. When the number of observed neighbors, which are mapped into the ε-interval, is larger than

the expected number from the binomial distribution for a selected α, i.e. l points, then the null can be rejected and, thus, a functional relationship can be assumed. The number of considered nearest neighbors k (i.e., the size of the δ-ball in Fig. 2.2) also determines the acceptable number of $k + 1 - l$ points falling outside the chosen ε-interval for a given probability parameter p of the binomial distribution and a given α. For each candidate delay τ_{d+1} and each time series $s_{j_{d+1}}$ for each k (at a given p and α) there is a minimal spatial scale $\varepsilon^{\star\prime}$ for which the null hypothesis can be rejected, i.e., a minimal size of the ε-interval in the right panel of Fig. 2.2. For the sake of avoiding *redundancy* while choosing the right delay, an $\varepsilon^{\star\prime}$ for each potential τ_{d+1} has to be found. This is simply the distance from the fiducial point to its lth-nearest neighbor. By averaging over a range of fiducial points we eventually get the final continuity statistic $\langle \varepsilon^{\star} \rangle (\tau)$ as a function of considered delay values τ (Fig. 2.3).

The final idea for achieving an optimal embedding is a sequential one. For each embedding cycle $\mathcal{D}_d$, i.e. for trying to find an appropriate additional component to build a reconstruction vector $\vec{v}(t)$ of dimension $d + 1$, initially starting with a 1-dimensional time series $\vec{v}(t) = \{s_i(t) | i = 1, \ldots, M\}$, the $\langle \varepsilon^{\star} \rangle (\tau)$ values for a range of possible delay values τ and for each time series $s_i(t)$ gets computed. The τ of the highest relative maximum of $\langle \varepsilon^{\star} \rangle (\tau)$ is selected as the optimal delay for this embedding dimension d. This delay is used to build up the temporal reconstruction vectors $\vec{v}(t)$ with the according time series. From here the next embedding cycle $\mathcal{D}_{d+1}$ gets executed and the process gets repeated until some break criterion terminates the procedure, i.e., when a sufficiently high embedding dimension m is reached.

Even though the idea of the continuity statistic is indeed promising, in this approach several unanswered questions remain, making the proposed idea not suitable for a fully automated embedding approach.

i) The choice of $p = 0.5$ has been made plausible and also our tests support this idea, while $\alpha = 0.05$ or $\alpha = 0.01$ is standard in science (see Figs. C.2, C.3, C.4 in Appendix C.3), so we can safely fix them to these values. What is not so clear, but highly relevant for the method, is how to choose the optimal delay τ from the continuity statistic. Specifically, we might ask what "relative maximum" exactly means and if there is any objective criteria for that choice. Moreover, it is also not clear how to obtain the continuity statistic in the first place with respect to the size of the neighborhood, i.e. the size of the δ-ball in Fig. 2.2. We propose to vary k from $k = 8$ to some higher value, like $k = 14$, for each considered delay τ and take the minimum of all trials $\varepsilon^{\star\prime}$ (and finally average over all fiducial points in order to obtain $\langle \varepsilon^{\star} \rangle$). This is allowed, because there is no preferred choice of k, but a lower bound (see Table 1 in Pecora et al. [240]), and generally the choice depends on the amount of available data and its sampling rate.

ii) In the original study, it was proposed that the continuity statistic on its own provides a breaking criterion for the method, namely, when "$\langle \varepsilon^{\star} \rangle$ remains small out to large τ, we do not need to add more components." [240] However, this is no objective criterion and introduces a statistic, which would quantify *small*, and also a threshold, which determines when *small* is *small enough*. Due to folding and stretching of the attractor for high delay values τ, especially in case of chaotic systems, we expect $\langle \varepsilon^{\star} \rangle$ to increase with higher τ, anyway. For these cases a (computationally intensive) irrelevance measure, the *undersampling statistic*, has been proposed [240]. Nevertheless, even though the undersampling statistic prevents the choice of *irrelevant* delays, it does not tell which of the local maxima of the continuity statistic we should pick and when to stop adding more components to the reconstruction vectors [2]. As an alternative to assess the irrelevance,

---

[2]Implicitly the undersampling statistic could be used as a break criterion, when all of the considered delays (local maxima in the continuity statistic) are above the chosen significance level β in the corresponding undersampling

the *L-statistic* has been suggested [315] which will be later used for our automated approach to attractor reconstruction.

2.2.2 *L*-statistic by Uzal et al.

The *L-statistic* is an objective cost function, which quantifies the goodness of a reconstruction, independent of the reconstruction method [315]. It has two free parameters, k and T_M. The approach uses ideas of *noise amplification* and minimization of the complexity of the reconstruction, which lead to a variable σ, and combines it with an irrelevance measure α. Specifically, the method estimates the conditional variance of neighborhoods consisting of k nearest neighbors as the relative dispersal of these neighborhood points with respect to the center of mass of that neighborhood T time steps ahead. Eventually this conditional variance estimate gets averaged over a range of prediction horizons T up to a maximum value T_M and is normalized with respect to the original neighborhood size, thus defining σ. The irrelevance measure α is basically the averaged inter-point distance, which depends on the sampling. The final statistic is then defined as

$$L_k = \log_{10}(\alpha_k \sigma_k), \tag{2.3}$$

where the index k indicates the dependence on the chosen number of nearest neighbors. A detailed description can be found in Appendix C.4. The authors showed, that the *L*-statistic converges for any $k \geq 3$. Our analysis (Fig. C.2 in Appendix C.3) supports this assumption and, thus, we can fix $k = 3$. However, the second free parameter T_M will alter the resulting *L*-statistic at any value. Particularly in the way we want to utilize the concept of this cost function in our automated embedding scheme, we need to tackle this parameter dependency (see Section 2.3). It is worth mentioning that the *L*-statistic inherits the minimization of a mean squared prediction error (MSE) (here computed using a local constant model based on the first k neighbors) and the FNN-statistic proposed by Kennel et al. [154] (when $k = 1$ and $T_M = \tau$).

2.3 New reconstruction method

The *L*-statistic (Section 2.2.2, Appendix C.4) on its own could guide the reconstruction problem on finding the optimal delay values and a sufficiently high embedding dimension, when used in a brute-force-approach, i.e., scanning all possible delay values of all available time series in every single possible combination. It is not clear a priori how to set the parameter T_M for obtaining the *L*-statistic, so the described procedure has to be repeated for a range of values for T_M. In most cases, this is not computationally feasible and, therefore, not suitable for a fully automated embedding approach. We propose to combine the *continuity statistic* (Section 2.2.1) and the *L-statistic* (Section 2.2.2). The continuity statistic $\langle \varepsilon^\star \rangle$ guides the preselection of potential delays τ and time series $\{s_i(t)|i = 1, \ldots, M\}$ in each embedding cycle $\mathcal{D}_d$, whereas the *L*-statistic decides which one to pick.

This algorithm works recursively. An embedding cycle $\mathcal{D}_d$ determines the optimal time delay and its corresponding time series and enables us to build the actual reconstruction vectors $\vec{v}_{\mathrm{act}}$

---

statistic. But this will not hold for non-chaotic systems and leads to very high embedding dimensions in these cases.

from this, having dimension $d + 1$. Algorithm 1 and Fig. 2.3 explains the method in detail, which we refer to as "PECUZAL" algorithm in honour of its roots.
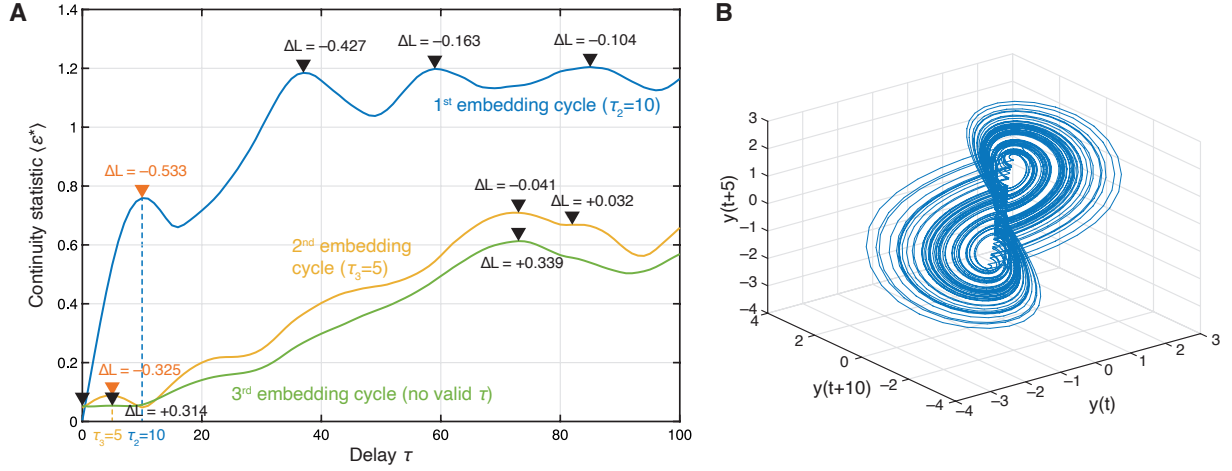
Figure 2.3: Illustration of the proposed embedding procedure for a univariate case, using the y-component of the Lorenz system (Appendix B.1). **A** Blue, yellow, and green lines represent the continuity statistics $\langle \varepsilon^\star \rangle$ for the three embedding cycles the Algorithm 1 computes. Triangles identify the τ values corresponding to local maxima of $\langle \varepsilon^\star \rangle$. Then, the local maximum which corresponds to the maximum decrease of the L-statistic with respect to the actual reconstruction vectors $\vec{v}_{\mathrm{act}}$, denoted as ΔL (orange triangle) is chosen as a delay time in each embedding cycle. In the third embedding cycle the cost-function cannot be minimized any further, i.e. all peaks lead to a non-negative ΔL. In this case no delay value is chosen and the algorithm breaks. **B** We end up with a 3-dimensional embedding and lags $\tau_1 = 0, \tau_2 = 10, \tau_3 = 5$.

(1) For the actual reconstruction vectors $\vec{v}_{\mathrm{act}}$, in each embedding cycle $\mathcal{D}_d$, $\langle \varepsilon^\star \rangle_i(\tau)$ is computed for all available M time series $s_i(t)$. We comment on the procedure in the first embedding cycle $\mathcal{D}_1$ further below.

(2) We consider all those delays τ_j for each $\langle \varepsilon^\star \rangle_i(\tau)$, which correspond to local maxima $\hat{\tau}_j$ in $\langle \varepsilon^\star \rangle_i(\tau)$. These delays τ_j (and their corresponding time series $s_{i_j}(t)$) are used to build temporary reconstruction vectors $\vec{v}_{\mathrm{temp}}(\tau_j, s_{i_j})$, by concatenating $\vec{v}_{\mathrm{act}}$ with the τ_j-lagged time series $s_{i_j}(t)$.

(3) For each $\vec{v}_{\mathrm{temp}}(\tau_j, s_{i_j})$ and the actual reconstruction vectors $\vec{v}_{\mathrm{act}}$ the L-statistic is simultaneous computed for many parameters T_M (c.f. Fig. 2.4), and we denote them as $L_{\vec{v}_{\mathrm{temp}}(\tau_j, s_{i_j})}(T_M)$ and $L_{\vec{v}_{\mathrm{act}}}(T_M)$. We compute the maximum L-decrease for $\vec{v}_{\mathrm{temp}}(\tau_j, s_{i_j})$ with respect to $\vec{v}_{\mathrm{act}}$ as $\Delta L_{\mathrm{temp}(\tau_j, s_{i_j})} = \min_{T_M}[L_{\vec{v}_{\mathrm{temp}}(\tau_j, s_{i_j})}(T_M) - L_{\vec{v}_{\mathrm{act}}}(T_M)]$. This way T_M is not a free parameter anymore.

(4) The delay-time series combination $(\tau_{j'}, s_{i_{j'}}(t))$, which yields the minimum ΔL value will be picked for this embedding cycle $\mathcal{D}_d$, if $\Delta L < 0$, and is used to construct the actual reconstruction vectors $\vec{v}_{\mathrm{act}} = \vec{v}_{\mathrm{temp}}(\tau_{j'}, s_{i_{j'}})$ of dimension $d + 1$

(5) $\vec{v}_{\mathrm{act}}$ is passed into the next embedding cycle $\mathcal{D}_{d+1}$.

(6) We repeat steps (1) to (5) until we cannot minimize the L-statistic any further, i.e. when each considered potential embedding $\vec{v}_{\mathrm{temp}}$ will yield a positive ΔL. In this case the reconstruction cannot get any better and we stop. $\vec{v}_{\mathrm{act}}$ constitutes the final reconstruction. Thus, the

L-statistic provides a break criterion, without the introduction of any threshold parameter. Each embedding cycle ensures the maximum possible decrease of the cost-function.

Algorithm 1 Pecuzal Embedding

1: **Input:** A uni- or multivariate dataset consisting of M time series s_i with same length and sampling and a range of possible delay values $\tau = 0 : \tau_{max}$
2: Normalize all M time series to zero mean and unit variance
3: Set $\Delta L_{\min} = -\inf$
4: **while** $\Delta L_{\min} < 0$ **do**
5:      **if** 1st embedding cycle $\mathcal{D}_1$ **then**
6:          Compute $\langle \varepsilon^\star \rangle_{ik}(\tau)$ for all M^2 pairwise combinations of s_i, s_k for the given τ's
7:          **for** each peak $\hat{\tau}_j$ in each $\langle \varepsilon^\star \rangle_{ik}(\tau)$ **do**
8:              Create $\vec{v}_{\text{temp}}$ by appending the time series s_i with the τ_k-lagged time series s_k
9:              Compute the L-statistics for $\vec{v}_{\text{temp}}$ and s_i for a range of parameter values $T_M = 2 : \tau_{max}$, denote them as $L_{\text{temp}}(T_M)$ and $L_{s_i}(T_M)$
10:              Compute $\Delta L_{\text{temp}} = \min_{T_M}[L_{\text{temp}}(T_M) - L_{s_i}(T_M)]$
11:          **end for**
12:          From all ΔL_{temp_j} take the τ_j, which corresp. to the peak with minimum ΔL, $\Delta L_{\min}$
13:          Save $\Delta L_{\min}$ and $\vec{v}_{\text{temp}}$
14:      **else if** $\mathcal{D}_d$ **then**
15:          Compute $\langle \varepsilon^\star \rangle_i(\tau)$ for $\vec{v}_{\text{act}}$ and all s_i for the given τ's
16:          **for** each peak $\hat{\tau}_j$ in each $\langle \varepsilon^\star \rangle_i(\tau)$ **do**
17:              Create $\vec{v}_{\text{temp}}$ by appending $\vec{v}_{\text{act}}$ with the τ_i-lagged time series s_i
18:              Compute the L-statistics for $\vec{v}_{\text{temp}}$ and $\vec{v}_{\text{act}}$ for a range of parameter values $T_M = 2 : \tau_{max}$, denote them as $L_{\text{temp}}(T_M)$ and $L_{\text{act}}(T_M)$
19:              Compute $\Delta L_{\text{temp}} = \min_{T_M}[L_{\text{temp}}(T_M) - L_{\text{act}}(T_M)]$
20:          **end for**
21:          From all ΔL_{temp_j} take the τ_j, which corresp. to the peak with minimum ΔL, $\Delta L_{\min}$
22:          Save $\Delta L_{\min}$ and $\vec{v}_{\text{temp}}$
23:      **end if**
24:      **if** $\Delta L_{\min} < 0$ **then**
25:          Set $\vec{v}_{\text{act}} = \vec{v}_{\text{temp}}$
26:      **end if**
27: **end while**
28: Set $\vec{v}_{\text{final}} = \vec{v}_{\text{act}}$
29: **Output:** The final reconstruction vectors $\vec{v}_{\text{final}}$

We give some remarks on the proposed idea:

i) In case of the very first embedding cycle the actual reconstruction vectors $\vec{v}_{\text{act}}$ are not yet defined. Roughly speaking, the algorithm needs to find a time series to start with and which can act as the first component of the final reconstruction vectors. As explained in Algorithm 1, each of the available time series serves as $\vec{v}_{\text{act}}$ in the first run and consequently M^2 continuity statistics get computed in the first step, i.e., for each combination (i, k) of the given time series $s_i(t)$. Note that we always use the unlagged available time series $s_i(t)$, which will constitute the first component of the reconstruction vectors, i.e. we set $\tau_1 = 0$. The continuity statistic

reveals the correlation structure of its input, meaning that a lagged initial time series would lead to different consecutive delay values. However, the difference of the finally chosen delay values as well as the total time window of the reconstruction $Tw = \max(\tau_1, \tau_2, \ldots)$ would be identical, because the correlation structure does not change, at least in case of infinite data. Practically, any $\tau_1 \neq 0$ only reduces the amount of data available and searching for the optimal τ_1 in the sense of a minimal ΔL for the first embedding cycle $\mathcal{D}_1$ would increase the computation time tremendously.

ii) The L-statistic can not serve as an absolute measure, due to its dependence on the parameter T_M. Consider a time series and a potential 2-dimensional embedding consisting of this time series and a τ-lagged version of it. This corresponds to the first embedding cycle $\mathcal{D}_1$ in Algorithm 1. Figure 2.4 shows the L-statistic for a range of parameters T_M for the single time series and for the potential 2-dimensional reconstruction for two deterministic systems (panels A, B) and in case of the time series being uniformly distributed random numbers (panel C). Whenever the L-statistic of the 2-dimensional reconstruction (orange graph) is lower than the one from the single time series (blue graph) an embedding would be beneficial (gray shaded areas in Fig. 2.4). But this is not always the case throughout the course of T_M (see panel B). The conclusion is, that it is not meaningful to judge a reconstruction on a single L-value, gained from a fixed T_M. A reconstruction is always related to a function $L(T_M)$ of the parameter T_M and it is sensible to look at relative L-discrepancies between two consecutive embedding cycles, namely ΔL. It turns out that some stochastic signals will yield a negative ΔL for $T_M = 1$. In panel C of Fig. 2.4 this is not the case, but the proximity of the two curves only for $T_M = 1$ is decernible. This is comprehensible by recalling that the L-statistic inherits the mean squared prediction error (see Eq. (C.5)), and a one-sample-prediction horizon is simply too short, to properly distinguish deterministic from stochastic sources. As a consequence we compute the L-statistics in each embedding cycle for T_M-values starting with $T_M = 2$. Thus, for any embedding cycle each peak of the continuity statistic does not receive a certain absolute L-value, but rather a maximum possible decrease of L, ΔL, with respect to the actual embedding (Fig. 2.3). Then one simply picks the peak, which yields the largest decrease. We can not rule out the possibility that we could obtain a lower overall decrease in L for all embedding cycles by taking a different "path", i.e. not go for the maximum decrease of L in each embedding cycle. This would correspond to achieving a local minimum of the cost function in the parameter and embedding cycle space.

iii) We propose to stop the embedding process, when $\Delta L > 0$ for all considered temporary reconstruction vectors $\vec{v}_{\text{temp}}$. One could think about incorporating a threshold, a small negative number, e.g. $\Delta L > -0.05$, to avoid only tiny decreases of the cost function encountered in an embedding cycle. Throughout all our computations this has not been necessary and, therefore, we dispense on such an additional parameter.

iv) The continuity statistic $\langle \varepsilon^\star \rangle$ itself contains information about the correlation structure of the data (cf. Section 2.2.1), which makes it valuable in the context of an automated embedding procedure as proposed here, especially for multivariate input. Not only that the first maximum most often coincides with the value we would obtain from the first minimum of the mutual information, but the continuity statistic of two time series "levels off" at a certain value range. The absolute value of $\langle \varepsilon^\star \rangle$ represents the correlation structure of the data we are looking at and quantifies the independence from each other for a specific time lag. This fact allows our method to pick only time series from a multivariate set, which belong together, and, consequently, in combination with the corresponding decrease of the L-statistic, ΔL, avoid stochastic signals (cf. Fig 2.4C, Table 2.1).
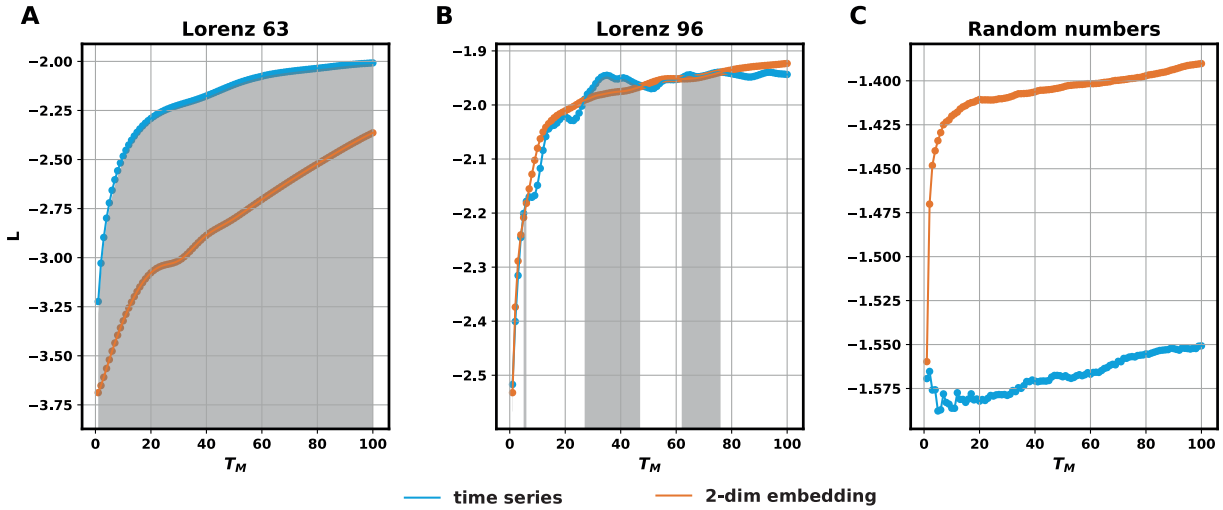
Figure 2.4: Illustration of the determination of ΔL within the embedding procedure for the first embedding cycle in a univariate case, using **A** the y-component of the Lorenz system (Appendix B.1), **B** the time series of the second node of a $N = 8$ Lorenz 96 setup (Appendix B.2) and **C** 1,000 uniformly distributed random numbers. Shown are the L-statistics for the single time series (blue graphs) and a 2-dimensional embedding with a τ-lagged version of itself (orange graphs) for a range of parameter values T_M. We set the number of nearest neighbors, which constitute a neighborhood, necessary for computing the L-statistic, to $k = 3$ and estimate τ from the first minimum of the corresponding auto-mutual information. The possible decrease of the L-statistic for this hypothetical embedding cycle with the chosen τ is simply $\Delta L(T_M) = L_{\text{orange}}(T_M) - L_{\text{blue}}(T_M)$. When $\Delta L(T_M) < 0$ (shaded areas) the additional reconstruction vector component does increase the quality of the reconstruction, whereas when $\Delta L(T_M) > 0 \ \forall \ T_M$ a further embedding is not beneficial. As expected that is the case for the stochastic signal in panel **C**. The Algorithm 1 automatically picks the first minimum ΔL over all T_M, which has also been the global minimum with respect to all T_M throughout all examples we have considered so far.

v) The time complexity of the proposed method is $\mathcal{O}(N \log N)$ as illustrated in Fig. C.1 (Appendix C.1).

2.4 Application, comparison & results

We apply the proposed method to a range of different systems, artificial and experimental data, exemplifying its advantage over established methods. Specifically, we compare our method to the standard *time delay embedding* (TDE). We estimate the uniform delay τ by means of the first minimum of the auto-mutual information [95] and estimate the appropriate embedding dimension by using Cao's method [39]. Specifically, we automatically select the appropriate embedding dimension, when the E_1-statistic reaches the saturation value within a given threshold of the slope (we picked a slope of < 0.05). We also look at two more sophisticated methods, which can also handle multivariate input, namely the method proposed by Garcia & Almeida (G&A) [104, 105][3] and Nichkawde's method (MDOP) [231]. The latter mentioned methods do not come with a predefined way to terminate the embedding process. In order to terminate them we use Hegger's [125] method of obtaining the optimal embedding dimension. Specifically, we set a FNN-threshold of 0.05, i.e., the algorithm breaks when the normalized fraction of FNN's fall below this threshold in order to allow the algorithm to give meaningful results in the presence of noise. The threshold for the tolerable relative increase of the distance between the nearest neighbors, when increasing the embedding dimension is set to 2, as suggested in [125, 154]. The threshold, which defines the factor of tolerable stretching for the d_{E_1}-statistic in case of G&A's method is set to 10, as suggested by the authors. We estimate the decorrelation time by using the first minimum of the auto-mutual information and use it as the Theiler window [292] in all approaches. In the multivariate input cases, we pick the maximum from all first minima of all auto-mutual information statistics. For distance computation, the Euclidean norm is used.

2.4.1 Reconstruction evaluation statistics

In order to compare our approach with the established methods we need to quantify the goodness of the embedding. For this, we will consider six statistics. In addition to the overall decrease of the L-statistic, that is $\Delta L = \sum_{i=1}^{m-1} \Delta L_i$, where m is the embedding dimension and ΔL_i the corresponding L-decreases in the encountered embedding cycles, we use two other statistics, which also reflect the neighborhood relations of the reconstruction compared to the reference. One is the mutual false nearest neighbor statistic (MFNN) [262]. Instead of estimating the coupling strength/ degree of synchrony of two coupled oscillators, we use the statistic for assessing the similarity between the reference (time series gained from numerical integration) and the reconstruction:

$$\text{MFNN} = \frac{1}{N} \sum_{i=1}^{N} \quad \frac{\sum_{k=1}^{K} |\vec{v}_i - \vec{v}_{i_k^{\text{ref}}}|}{\sum_{k=1}^{K} |\vec{u}_i - \vec{u}_{i_k^{\text{ref}}}|} \quad \frac{\sum_{k=1}^{K} |\vec{u}_i - \vec{u}_{i_k^{\text{rec}}}|}{\sum_{k=1}^{K} |\vec{v}_i - \vec{v}_{i_k^{\text{rec}}}|}, \tag{2.4}$$

---

[3]Note that we were not able to reproduce the results shown in the papers from Garcia & Almeida[104, 105]. Two of the authors, K.H.K and N.M as well as another experienced researcher independently implemented this method and got the exact same results. An email to Prof. Garcia explaining this issues and seeking for help remains unanswered. We improved the method to be at least capable of producing acceptable results. First we implemented a Theiler window, which has not been discussed by the authors. Second we introduced the forward time step in order to produce the d_{E_2} statistic as a free parameter. The authors only discuss the case of a forward time step of 1. Throughout all our computations we set this parameter to the same value as the Theiler window, i.e. the value of the first minimum of the auto-mutualinformation. In the multivariate input cases, we picked the maximum from all first minima of all auto-mutualinformation. The reader is welcome to follow the implementation process on https://github.com/JuliaDynamics/DelayEmbeddings.jl/pull/38.

where $\vec{u}_i$ are the vectors of the reference/original system, $\vec{v}_i$, $i = 1, \ldots, N$ the reconstruction vectors, i_k^{ref} the indices of the K-nearest neighbors of index i in the original system and i_k^{rec}, $k = 1, \ldots, K$ the corresponding indices measured in the reconstruction. We propose the comparison of K nearest neighbors instead of just focusing on the first nearest neighbor, in order to receive more robust results in the presence of noise. By sampling the data sufficiently high we allow the precise determination of quite large neighborhoods of a fiducial point, so we set $K = 10$ in our computations. The results vary in their absolute values with different choices of K, but the order of rank for the different test methods and their relative difference remain approximately constant. MFNN $= 1$ corresponds to an ideal Afraimovich diffeomorphism [4, 239], higher values mark worse reconstructions. The other strict criterion we propose quantifies the degree of neighborhood-relation conservation: the Joint Recurrence Rate Fraction (JRRF). It is based on the concept of a recurrence plot (RP), which is a 2-dimensional representation of a dynamical system as a binary matrix [78, 209] (Appendix A, Eqs. (1.1), (1.2)). JRRF measures the accordance of the recurrence plot of the reference system, $\mathbf{R}^{\text{ref}}$, and the RP of the reconstruction, $\mathbf{R}^{\text{rec}}$.

$$\text{JRRF} = \frac{\sum_{i,j}^{N} JR_{i,j}}{\sum_{i,j}^{N} R_{i,j}^{\text{ref}}}, \quad \text{JRRF} \in [0\ 1] \tag{2.5}$$

$$\mathbf{JR} = \mathbf{R}^{\text{ref}} \circ \mathbf{R}^{\text{rec}}. \tag{2.6}$$

We compute both, $\mathbf{R}^{\text{ref}}$ and $\mathbf{R}^{\text{rec}}$, by fixing the recurrence threshold corresponding to a global recurrence rate of 8%. This is also to ensure comparability of the recurrence quantifiers described below [163]. Results are fairly similar for a wide range of choices of the recurrence rate we tried and the particular choice (in our case 8%) is not so important, since we apply them to all RP's we compare. It is, of course, also possible to compare different recurrence plot quantifiers gained from $\mathbf{R}^{\text{rec}}$ to the ones derived from $\mathbf{R}^{\text{ref}}$ [209]. We here choose the determinism (DET), Eq. (A.6), the diagonal line length entropy (ENTR), Eq. (A.9), and the recurrence time entropy (RTE), Eq. (A.10) (Appendix A). The latter two are related to the Kolmogorov-Sinai-Entropy [15, 199], but do not serve as straight forward estimators, when necessary corrections on the RP and its quantifiers are ignored [164].

2.4.2 Paradigmatic examples

We investigate three paradigmatic chaotic systems, the Rössler system in the funnel regime (Appendix B.3), a driven Duffing oscillator in regular motion (Appendix B.4) and the Mackey-Glass delay equation (Appendix B.5). For all systems we compare the mean values of the evaluation statistics from ensembles of 1,000 trajectories with different initial conditions. Table C.1 in Appendix C.2 summarizes all results, also including uncertainties and results for 10% additive measurement noise. In order to easily compare the evaluation statistics, we use the relative deviation from the reference (e.g., $|\text{DET}^{\text{rec}} - \text{DET}^{\text{ref}}|\big/\text{DET}^{\text{ref}}$), except for the MFNN and the ΔL-statistic, where we use the relative deviation from the best score (i.e., $|\text{MFNN}^{\text{rec}} - \text{MFNN}^{\text{best}}|\big/\text{MFNN}^{\text{best}}$), For a concise visual presentation we use spider plots in the following and plot *1 - rel. deviation*, i.e. the closer to unity the value gets, the better the accordance to the reference or the best achieved value is.

(i) For the chaotic Rössler system, we reconstruct the state space for the univariate case (using the y-component, in order to allow TDE having the best chances [184]) and for the multivariate case (using the x- and y-component). An overview over the results are shown in Figure 2.5. For
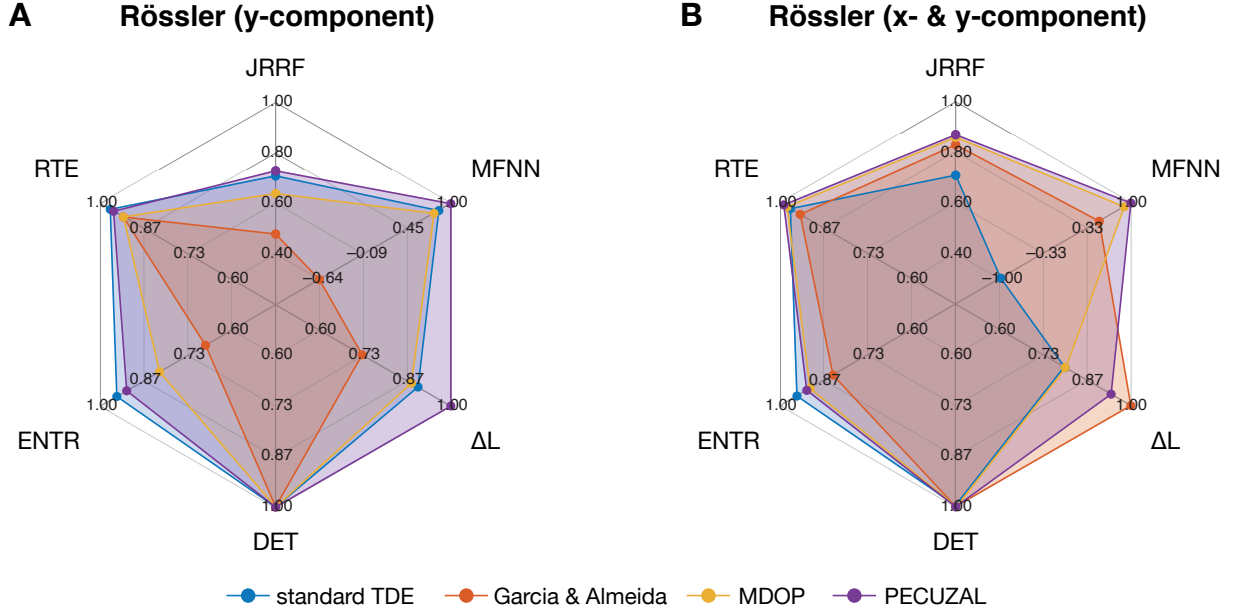
Figure 2.5: (Relative) deviation of reconstruction by standard time delay embedding (TDE), the Garcia & Almeida (G&A), Nichkawde's (MDOP), and PECUZAL methods, comparing the accordance of the RPs of the reconstructed attractor and the reference attractor (JRRF), mutual false nearest neighbors (MFNN), ΔL-statistic, as well as the recurrence quantifiers determinism (DET), diagonal line length entropy (ENTR), and the recurrence time entropy (RTE). **A** Univariate case using the y-component of the numerically integrated Rössler system (Appendix B.3) and **B** multivariate case using the x- and y-values of the Rössler system. Since TDE cannot handle multivariate input we take the values from the univariate case here for illustrative reasons, which result in different relative values in case of MFNN and the ΔL-statistic. For these measures we plot the *1- relative deviations to the best score*, which increases in the multivariate case. For the other statistics we plot *1- relative deviations to the reference score*, i.e. the closer to unity the value gets, the better the accordance to the reference or the best achieved value is.

TDE in the multivariate case we take the results from the univariate example, because TDE cannot handle multivariate input. Note that this leads to different relative values for MFNN and ΔL, since we plot the deviation to the best score in these cases. The PECUZAL method performs best in the univariate as well as in the multivariate scenario, with improved outcomes for the multivariate one, as expected. This also holds in case of applied measurement noise, where our method even expands its lead for most measures (Table C.1). Surprisingly, TDE also provides very good results in the univariate case and in the multivariate setup Garcia & Almeida's method yields an overall larger decrease of the L-statistic than PECUZAL, but only in the noise free case (see Table C.1). This lower ΔL is not reflected in better performance of the other evaluation statistics. Specifically, the diagonal line length entropy values differ from the true reference value in the double-digit percentage range for G&A's method.

(ii) The overall rating also holds for the driven Duffing oscillator (Fig. 2.6), but there are severe differences. In contrast to the chaotic Rössler system here it seems that the additional given time series for the multivariate scenario do not improve the reconstructions significantly. The *JRRF* do not get better and remain on a quite high level of 80% (G&A) up to 84% (PECUZAL) accordance with the reference recurrence plot $\mathbf{R}^{\mathrm{ref}}$. This is also reflected in the ΔL statistic, which also
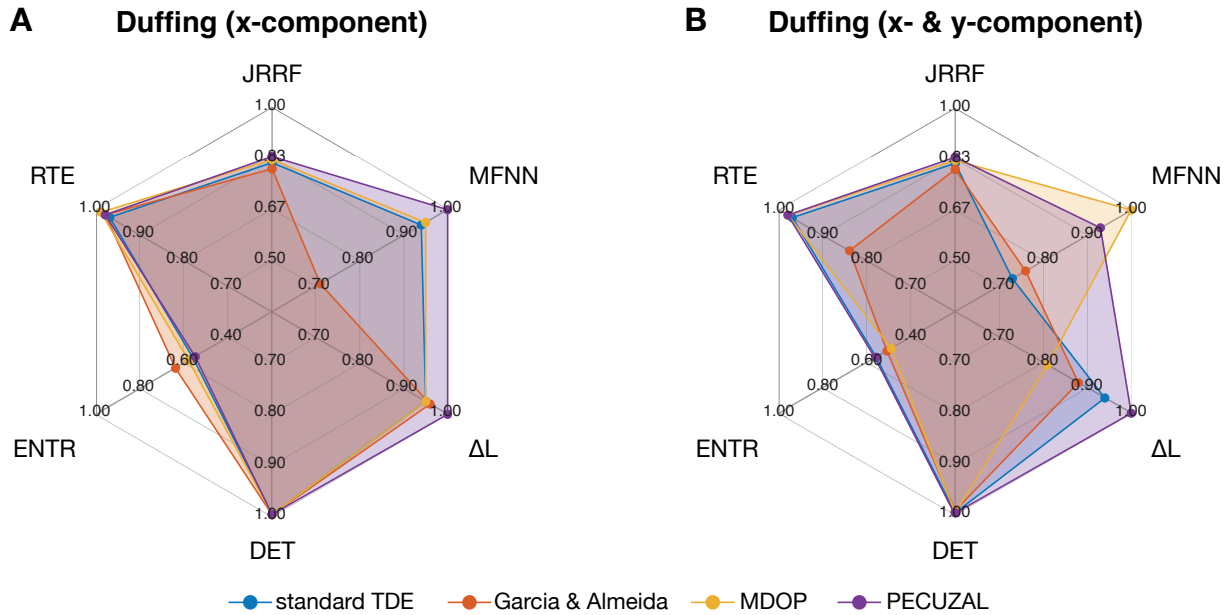
Figure 2.6: Same as Fig. 2.5, here for the Duffing system (Appendix B.4).

does not improve in the multivariate case for G&A and MDOP, and only slightly decreases for PECUZAL. The same story is told by looking at the other evaluation statistics summarized in Table C.1. It is important to note that under noise our proposed method outperforms the others in almost all cases, but in principle we notice that the signal to noise ratio seems to be very low and biases the results of all methods more than in case of the Rössler example. Specifically in case of RTE deviations to the true reference value increase from the low single digit percentage range up to 62% in the noisy case. The reason is that in regular motion we expect a near zero value of RTE and noise easily blurs the diagonal lines in the recurrence plot, leading to a broader distribution of recurrence times and, thus, to randomly elevated RTE values. The very same problem make the $ENTR$ values deteriorate for all methods, here already apparent in the noise free case. In a regular motion system $ENTR$ is biased the most when no correction of the diagonal lines is performed, which has been shown by Kraemer and Marwan [164] (Chapter 5). The proposed skeletonizaton of the recurrence plots for the purpose of reducing the bias in the diagonal line based recurrence plot quantifiers, such as $ENTR$ and RTE, lead to way better results in the one digit percentage range even for noise, as expected. Due to the computational complexity of the skeletonization algorithm [164] we did not apply this correction scheme to all of the 1,000 runs in this experiment, but rather tried it on small samples not shown here, in order to understand the bad performances for all methods in case of $ENTR$ and RTE.

(iii) In contrast to the above systems, the Mackey-Glass system (Appendix B.5) is infinite dimensional and we have do deal with a univariate time series from the numerical integration, which is why we do not have a "reference" value we could base our computations on. Therefore, we can only use the ΔL-statistic (Table C.1). The proposed PECUZAL method performs significantly better than the other methods with Garcia & Almeida's method coming second, especially in the presence of noise. Here all methods but the proposed one yield reconstructions with corresponding positive ΔL-values, i.e. the suggested embeddings do not decrease the overall L-statistic. Admittedly, PECUZAL also comes with a very small negative ΔL value, with 0 falling

within the 1σ-interval. This finding indicates a too low signal to noise ratio, which we comment on below. Farmer [85] conjectured a linear relation between the delay chosen in the Mackey-Glass equation and the corresponding dimensionality of the attractor (c.f. Table I in [85]). A linear fit to the data of that table suggests an attractor-dimension of $d_A \approx 5.3$ with the 95% confidence interval being between ≈ 2.1773 and ≈ 8.2888. The studied methods give 5 ± 0 ($6.7 \pm .5$) (TDE), 3 ± 0 (4.6 ± 1) (G&A), 4 ± 0 (2 ± 0) (MDOP) and 7 ± 0 (2.2 ± 1.7) (PECUZAL). The bracketed values correspond to the case of 10% additive noise. While all methods meet Farmer's conjecture in the noise free scenario, this does not hold for the MDOP method an the proposed method in the noisy case. Both methods suggest too low embedding dimensions. This is due to the fact that the signal to noise ratio is apparently too low and PECUZAL treats the signal as a stochastic source for some realizations where it does not embed the data at all, while it did not do it in case of the Rössler and Duffing system, despite the same variance of the white Gaussian noise. Results from G&A and TDE do fall in the 95% confidence interval, which is large, because of the weak data basis given in Ref.[85] and the resulting uncertain fit. We find the time window of the embedding, i.e., the total time span covered by a reconstruction vector, decreasing with increasing noise level throughout our experiments. This is very much in line with the findings of Ragwitz and Kantz [251].

We finally look at two made up, ill-defined multivariate datasets, in order to see how the G&A, MDOP, and the PECUZAL method cope with redundant data and with stochastic signals.

(i) First we construct a dataset consisting of six time series (*Fooling dataset I*). The first two time series are the x- and y-component of the Rössler system (Appendix B.3), the third and fourth time series are the x- and y-component of the Lorenz system (Appendix B.1), the fifth time series corresponds to $x_{\text{Rössler}}^2 + y_{\text{Rössler}}$, whereas the sixth time series is set to $x_{\text{Lorenz}} \cdot y_{\text{Lorenz}} + y_{\text{Lorenz}}$. Our proposed method does not mix time series from both systems and sticks to one system (Lorenz in this case), as shown in Table. 2.1. It suggests a 3-dimensional embedding and also does not need the redundant information stored in the fifth and the sixth time series of the input dataset. In contrast, G&A and MDOP fail here, suggesting a 3-dimensional and a 6-dimensional embedding, respectively, mixing up the different systems yielding a useless reconstruction (Fig. 2.7).

(ii) The second made-up dataset (*Fooling dataset II*) consists of three time series of length 5,000, with the first one being an auto-regressive process of order 1 with parameters (0, 0.9) and initial condition $u_0 = 0.2$, to mimic colored noise. The second and third time series are Gaussian and uniform random numbers. While G&A and MDOP embed the non-deterministic time series, our proposed algorithm suggests no embedding and throws an error (Tab. 2.1). The reason is, that the L-statistic is a monotonically increasing function of the embedding dimension for stochastic data for any prediction horizon parameter T_M, i.e., the algorithm cannot minimize L already in the first embedding cycle. For the sake of completeness we have to stress that this particular example should not be read as a claim of a generalizable behavior of our proposed method to deal with auto-regressive processes of arbitrary order p. In the case of higher-order AR processes, PECUZAL often suggests an embedding with a dimension that corresponds to the order of the AR process, as we would expect it theoretically. We have noticed, however, that the embedding depends heavily on the length of the time series used, which is in line with the findings of Holstein and Kantz [132], but also of the choice of the particular AR-parameters and the order of the AR process under study. A systematic consideration of PECUZAL's embedding suggestions for this class of processes is beyond the scope of the work presented here.

A   **Garcia & Almeida**
delays: [0 1 1]
ts: [1 4 1]

B   **MDOP**
delays: [0 20 16 20 0 23]
ts: [6 2 1 3 3 4]

C   **PECUZAL**
delays: [0 3 0]
ts: [3 4 4]
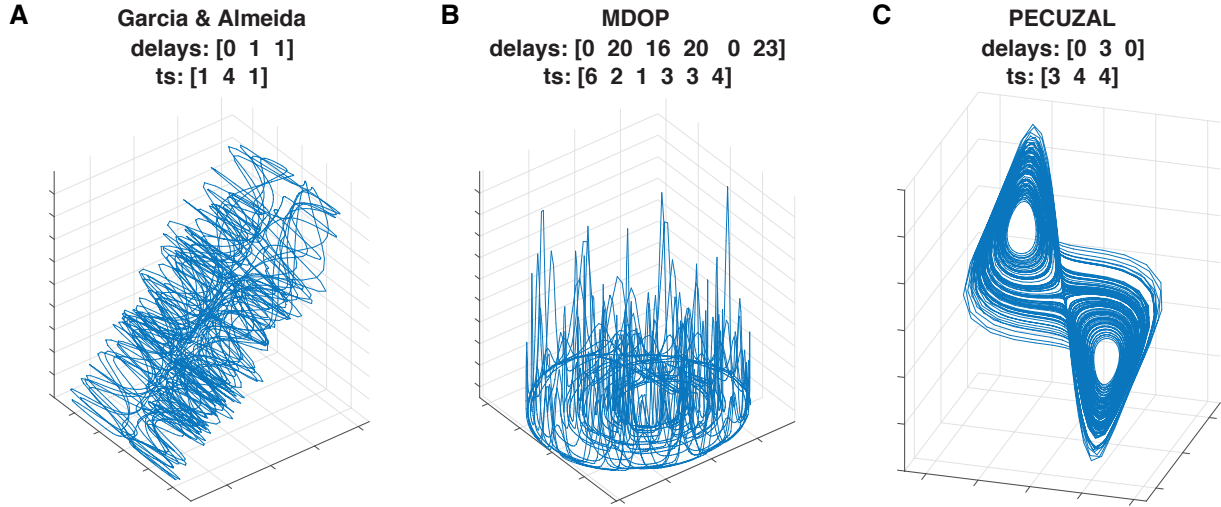
Figure 2.7: Reconstructions of the *Fooling dataset I* (see text for details). In case of the MDOP method (panel **B**), we plot the first three components of the 6-dimensional trajectory. While G&A and MDOP methods mix time series from different systems, leading to nonsensical reconstructions (panels **A** and **B**), our proposed method (panel **C**) sticks to time series from one system (Lorenz in this case).

Table 2.1: Embedding dimension, the accordingly chosen time series and corresponding time lags for the fooling datasets, mimicking mixed deterministic data from different systems and redundant time series as well as stochastic time series. We compare Garcia & Almeida's (G&A), Nichkawde's (MDOP), and our proposed PECUZAL method.

	m			chosen time series			Delay's		
System	G&A	MDOP	PECUZAL	G&A	MDOP	PECUZAL	G&A	MDOP	PECUZAL
Fooling dataset I	3	6	3	1, 4, 1	6, 2, 1, 3, 3, 4	3, 4, 4	0, 1, 1	0, 20, 16, 20, 0, 23	0, 3, 0
Fooling dataset II	6	6	—	1, 2, 2, 1, 1, 1	3, 3, 1, 2, 2, 1	—	0, 1, 2, 3, 2, 1	0, 1, 0, 1, 0, 1	-

2.4.3   Experimental data

We will now utilize the PECUZAL embedding method on experimental data. Specifically, we look at a chaotic time series from electrochemical oscillations. The experiment was performed with the chaotic electrodissolution of nickel in sulfuric acid [157]. A standard three-electrode electrochemical cell was used with a 1-mm diameter nickel wire as working, a Pt counter, and a Hg/Hg_2SO_4/sat. K_2SO_4 reference electrode. The electrolyte was 4.5 M H_2SO_4 at 10°C. The nickel wire was connected to the working point of the potentiostat through an individual resistance (R_{ind}), and a potentiostat (Gill AC, ACM Instruments) applied a constant circuit potential (V_0, with respect to the reference electrode). At a given circuit potential, the rate of the metal dissolution, measured as the current, can exhibit chaotic oscillations due the hidden negative differential resistance and additional nonlinear processes related to the passivation kinetics [330]. About 500 current oscillations were recorded at a data acquisition rate of 200Hz, which corresponds to about 200 points per cycle and a time series length of $N = 100,000$. There are two primary bifurcation parameters in the experiment: the individual resistance, which affects the charging time constant of the electrical double layer, and the circuit potential, which drives the dissolution.

We consider a setting with $R_{\text{ind}} = 1.5$ kΩ and $V_0 = 1{,}360$ mV.
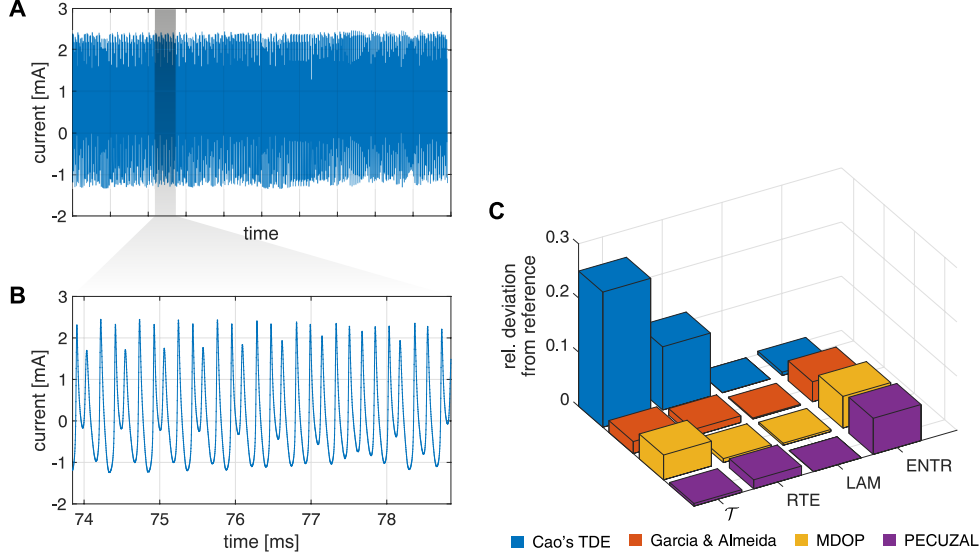
Figure 2.8: A Entire z-standardized and down-sampled time series from electrochemical oscillations of length $N = 50{,}000$ and **B** a sub-sample of length $\hat{N} = 2{,}500$. **C** The relative deviation of the medians of the distributions of the RQA-quantifiers ENTR, LAM, RTE and $\mathcal{T}$ obtained from RPs of $1{,}000$ sub-samples to their reference values obtained from RPs of the entire time series for the four reconstruction methods.

We demonstrate the ability of the proposed embedding method to cope with rather small to intermediate sized experimental datasets. We first down-sample the time series to $N = 50{,}000$ (Fig. 2.8A). Reconstructions for the four methods (TDE, G&A, MDOP, PECUZAL) are then computed from which we obtain RPs (with a fixed recurrence rate of 8%, in order to guarantee comparability) and the corresponding RQA-quantifiers diagonal line length entropy (ENTR), the laminarity (LAM), the recurrence time entropy (RTE) and the recurrence network measure transitivity ($\mathcal{T}$), see Appendix A. We denote each of these values for each of the reconstruction method as its reference values. We then repeat the described procedure for $1{,}000$ sub-samples of length $\hat{N} = 2{,}500$ drawn from the time series at random (shown exemplary in Fig. 2.8B), i.e. for each of the $1{,}000$ sub-samples we compute the reconstruction for each of the four methods, its corresponding RP and the RQA-quantifiers. This will result in distributions for ENTR, LAM, RTE and $\mathcal{T}$ for each reconstruction method. Finally we compare the medians of these distributions to their reference values and plot the relative deviation $\frac{|\text{RQA}_{\text{median(distr.)}} - \text{RQA}_{ref}|}{\text{RQA}_{ref}}$ in Fig. 2.8C. The capability of the four methods to allow for satisfying estimates from short time series samples differs strongly for the different RQA-quantifiers. The largest discrepancies to the reference can be noted for TDE in case of $\mathcal{T}$ and RTE. For LAM all methods estimate the reference very well from the sub-samples. While our proposed method slightly comes last in case of ENTR, it yields the best results for $\mathcal{T}$ and LAM and is performing well in case of RTE. The example shown here can not be generalized, but it underpins our claim that PECUZAL provides robust state space reconstructions for a very broad range of processes under different conditions, which are often better, but always at least equally well than the ones obtained from the established methods.

2.5 Conclusions

A fully automated approach for state space reconstruction from uni- or even multivariate time series has been proposed and compared to established methods. The algorithm works iteratively and appends the reconstruction vectors in each embedding cycle with an appropriate time delay and an according time series until a cost function cannot be minimized further. Its core functionality is based on identifying potential time delays and its corresponding time series in each embedding cycle by using the continuity statistic. For each of those delays, temporary reconstruction vectors are build and the cost function is computed. The delay value, which yields the maximum decrease of the cost function is selected. If none of the considered delay values yields a decrease of the cost function the reconstruction can not get any better and the final embedding is obtained without the need of setting any threshold parameter. Usually the time delays chosen that way are not simply multiples of each other, but rather reflect even complex correlation structures within the data. This is why the algorithm is also able to detect time series stemming from a stochastic source, which it will not embed. Except from computing the decorrelation time of the data for providing a valid Theiler window for the nearest neighbor search, and providing a range of possible delay values the algorithm shall encounter, there are neither any data preprocessing steps necessary, nor any free parameters need to be adjusted before using the proposed routine. The approach has been demonstrated on a variety of exemplary systems as well as on experimental data stemming from chaotic chemical oscillators. We find that it provides often better, but always at least equally well reconstructions than the established methods. It is furthermore capable of providing meaningful reconstructions for rather short time series, which particularly holds for the case of multivariate input. The additional computational effort in comparison to standard time delay embedding is manageable and justified. Since the proposed method works automatically, is basically parameter free, and can handle multivariate input without mixing data originating from different systems, we can think of a wide range of potential applications. This is especially true for scenarios, where multiple sensors or channels of a detector monitor real world processes, which are not isolated observables, i.e., in engineering, earth- and life science contexts. The provided software (Appendix C.1) in three common coding languages will facilitate the use of the presented method.

Chapter 3

Optimal state space reconstruction via Monte Carlo Decision Tree Search

Abstract

A novel idea for an optimal time delay state space reconstruction from uni- and multivariate time series is presented. The entire embedding process is considered as a game, in which each move corresponds to an embedding cycle and is subject to an evaluation through an objective function. This way the embedding procedure can be modeled as a tree, in which each leaf holds a specific value of the objective function. By using a Monte Carlo ansatz the proposed algorithm populates the tree with many leafs by computing different possible embedding paths and the final embedding is chosen as that particular path, which ends at the leaf with the lowest achieved value of the objective function. The method aims to prevent getting stuck in a local minimum of the objective function and can be used in a modular way, enabling practitioners to choose a statistic for possible delays in each embedding cycle as well as a suitable objective function themselves. The proposed method guarantees the optimization of the chosen objective function over the parameter space of the delay embedding as long as the tree is sampled sufficiently. As a proof of concept, we demonstrate the superiority of the proposed method over the classical time delay embedding methods using a variety of application examples. We compare recurrence plot based statistics inferred from reconstructions of a Lorenz-96 system and highlight an improved forecast accuracy for map-like model data as well as for palaeoclimate isotope time series. Finally we utilize state space reconstruction for the detection of causality and its strength between observables of a gas turbine type thermoacoustic combustor.

3.1 Introduction

The famous embedding theorems of Whitney [329], Mañé [196], and Takens [291] together with their enhancement by Sauer et al. [268] allow a high dimensional state space reconstruction from

(observed) uni- or multivariate time series. Computing dynamical invariants [117, 118, 127, 147, 150] from the observed system, making meaningful predictions even for chaotic or stochastic systems [44, 86, 136, 146, 237, 250, 289], detecting causal interactions [88, 288, 333] or non-linear noise reduction algorithms [149, 216] all rely explicitly or implicitly on (time delay) embedding [235] the data into a reconstructed state space. Other ideas rather than time delay embedding (TDE) are also possible [27, 111, 190, 197, 235, 236], but due to its simple use and its proficient outcomes in a range of situations, TDE is by far the most common reconstruction technique. Suppose there is a multivariate dataset consisting of M time series $s_i(t)$, $i = 1, \ldots, M$. The basic idea is to use lagged values of the available time series as components of the reconstruction vector

$$\vec{v}(t) = (s_{i_1}(t - \tau_1), s_{i_2}(t - \tau_2), \ldots, s_{i_m}(t - \tau_m)). \tag{3.1}$$

Here the delays τ_j are multiples of the sampling time Δt and the indices $i_1, i_2, \ldots, i_m$ each denote the time series index $i \in [1, \ldots, M]$, which has been chosen in the 1st, $2^{\text{nd}}, \ldots, m^{\text{th}}$ embedding cycle. The total number of delays τ_j, $j = [1, \ldots, m]$, i.e., the embedding dimension m, its values and the corresponding time series s_{i_j}, $i_j \in [1, \ldots, M]$ need to fulfill certain criteria to guarantee the equivalence to the unknown true attractor, e.g., the embedding dimension must suffice $m \geqslant 2D_B + 1$, with D_B being the unknown box-counting dimension (see Casdagli et al. [45], Gibson et al. [111], Uzal et al. [315] or Nichkawde [231] for a profound overview of the problem). Picking optimal embedding parameters τ_j and m comes down to make the resulting components of the reconstruction vectors $\vec{v}(t)$ as independent as possible [235, 268], but at the same time not too independent, in order to keep sufficient information of the correlation structure of the data [45, 79, 173, 260, 315]. Besides some unified approaches [33–35, 102, 104, 105, 152, 162, 185, 284, 312], which tackle the estimation of the delays τ_j and the embedding dimension m simultaneously, most researchers use two different methods to perform the reconstruction.

(1) A statistic determines the delays τ_j, we call it Λ_τ throughout this paper. Usually $\tau_1 = 0$, i.e., the first component of $\vec{v}(t)$ is the unlagged time series s_{i_1} in Eq. (3.1). For embedding a univariate time series, $s_{i_1} = \ldots = s_{i_m} = s(t)$, the approach to choose τ_2 from the first minimum of the auto-mutual information [95, 186] is most common. All consecutive delays are then simply integer multiples of τ_2. Other ideas based on different statistics like the auto-correlation function of the time series have been suggested [5, 8, 9, 27, 40, 152, 173, 260]. However, by setting $\tau_j, j > 2$ to multiples of τ_2, one ignores the fact that this "measure" of independence strictly holds only for the first two components of reconstruction vectors ($m = 2$) [94, 119], even though in practice it works fine for most cases. More sophisticated ideas, like high-dimensional conditional mutual information [140, 317] and other statistics [36, 103, 155, 217, 241, 317], some of which include non-uniform delays and the extension to multivariate input data [104, 105, 123, 130–132, 140, 142, 231, 240, 317], have been presented.

(2) A statistic, we call it Γ throughout this Section, which serves as an objective function and quantifies the goodness of a reconstruction, given that delays τ_j have been estimated. The embedding process is thought of as an iterative process, starting with an unlagged (given) time series s_{i_1}, i.e., $\tau_1 = 0$. In each embedding cycle $\mathcal{D}_d, [d = 1, \ldots, m]$ a time series s_{i_d} lagged by τ_d, gets appended to obtain the actual reconstruction vectors $\vec{v}_d(t) \in \mathbb{R}^{d+1}$ and these are compared to the reconstruction vectors $\vec{v}_{d-1}(t)$ of the former embedding cycle (if $d = 1$, $\vec{v}_{d-1}(t)$ is simply the time series s_{i_1}). This "comparison" is usually achieved by the amount of false nearest neighbors (FNN) [39, 125, 153, 154, 170], some other neighborhood-preserving-idea [10, 49], or

more ambitious ideas [231, 315].

We have recently proposed an algorithm [162] (Algorithm 1 in Section 2.3), which minimizes the L-statistic [315] (the objective function, Appendix C.4) in each embedding cycle $\mathcal{D}_d$ over possible delay values in this embedding cycle determined by a continuity statistic [240] (Section 2.2.1). Nichkawde [231] minimizes the FNN-statistic in each embedding cycle over time delays given by a statistic, which maximizes directional derivatives of the actual reconstruction vectors. However, it cannot be ruled out that these approaches result in achieving a local minimum of the corresponding objective function, rather than attaining the global minimum.

Here we propose a Monte Carlo Decision Tree Search (MCDTS) idea to ensure the reach of a global minimum of a freely selectable objective function Γ, e.g., the L- or FNN-statistic or any other suitable statistic, which evaluates the goodness of the reconstruction with respect to the task. A statistic Λ_τ, which guides the pre-selection of potential delay values in each embedding cycle (such as the continuity statistic, described in Section 2.2.1, or conditional mutual information) is also freely selectable and can be tailored to the research task. This modular construction might be useful for practitioners, since it has been pointed out that optimal embedding parameters – thus also the used statistics to approximate them – depend on the research question, e.g., computing dynamical invariants or prediction [24, 108, 132, 142, 325]. Thus, the proposed method is neither restricted to the auto-mutual information, in order to measure the independence of consecutive reconstruction vector components, nor does it necessarily rely on the ubiquitous false nearest neighbor statistic. Independently from the chosen statistic for potential time delays and from the chosen objective function, the proposed method computes different embedding pathways in a randomized manner and structures these paths as a tree. Consequently it is able to reveal paths through that tree – if there are any – which lead to a lower value of the objective function than paths, which strictly minimize the costs in each embedding cycle. Given a sufficiently high number of samplings, MCDTS guarantees to optimize the chosen objective function Γ over the (delay embedding-) parameter space. In Section 3.2 we describe this method before we apply it to paradigmatic examples in Section 3.3, which include Recurrence Analysis, nearest-neighbor-based time series prediction and causal analysis based on convergent cross mapping.

3.2 Method

When embedding a time series, in each embedding cycle a suitable delay, and for multivariate data a suitable time series, has to be chosen. While the final embedding vector is invariant to the order of chosen components, the embedding process, and the used statistics and methods to suggest suitable delays, generally depend on all the previous embedding cycles [4]. It seems therefore natural to visualize all possible embedding cycles in a tree-like hierarchical data structure as shown in Figure 3.1. The initial time series s_{i_1} with delay $\tau_1 = 0$ forms the root of the tree and each possible embedding cycle $\mathcal{D}_d$ is a leaf or node of the tree. With the large amount of possible delays and time series to choose from, this decision tree becomes too large to fully compute it. At the same time, aforementioned statistics like the continuity statistic or conditional mutual information can guide us in pre-selecting potentially suitable delay values and an objective function like the L- or FNN-statistic can pick the most suitable delay value of the pre-selection by quantifying the quality of the reconstruction in each embedding cycle. Throughout this paper we

---

[4] The continuity statistic $\langle \varepsilon^\star \rangle (\tau)$ is one example for such a statistic that depends on all previous embedding cycles.
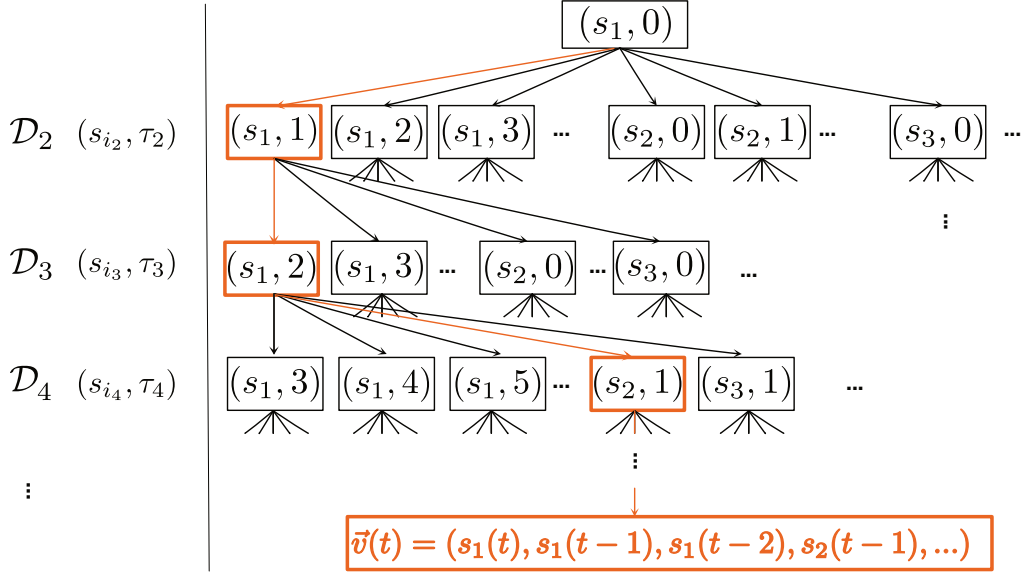
Figure 3.1: All possible embeddings of a time series visualized by a tree. Each leaf of the tree symbolizes one embedding cycle $\mathcal{D}_d$ using one selected time series s_{i_d} from the multivariate data set and delay τ_d. Marked in orange is one chosen full embedding.

denote a statistic, which pre-selects potential delay values as Λ_τ and the objective function as Γ. The task to embed a time series can then be interpreted as minimizing $\Gamma(i_1, i_2, .., i_m, \tau_1, \tau_2, ..., \tau_m)$. Visualizing this with a tree as in Fig. 3.1, we actually perform a tree search to minimize Γ. However, always choosing the leaf of the tree that decreases Γ the most, might lead only to a local minimum.

As we strive to find a global minimum and cannot compute the full embedding tree, we proceed by sampling the tree. This approach is inspired by the Monte Carlo Tree Search algorithms that were originally envisioned to master the game of Go [59]. Ultimately computer programs based on these algorithms were able to beat a reigning world champion, a feat that was long thought to be impossible for computer programs [283]. Adapting this idea to the embedding problem we proceed as follows. We randomly sample the full tree, for each embedding cycle we compute the change in the objective function Γ and pick for the next embedding cycle preferably those delays that decrease Γ further. Each node $\mathcal{N}_d$ of the tree encodes one possible embedding cycle and holds the time series used $[s_{i_1}, \ldots, s_{i_d}]$, the delays used until this node $[\tau_1, \ldots, \tau_d]$, i.e., the current *path* through the tree up to node $\mathcal{N}_d$, and a value of the objective function Γ_d. We sample the tree N_{trial}-times in a two-step procedure:

- **Expand:** Starting from the root, for each embedding cycle $\mathcal{D}_d$, possible next steps $(s_{i_j}, \tau_j, \Gamma_j)$ are either computed using suitable statistics Λ_τ and Γ or, if there were already previously computed ones, they are looked up from the tree. We consider the first embedding cycle $\mathcal{D}_2$ and use the continuity statistic $\langle \varepsilon^\star \rangle(\tau)$ for Λ_τ. Then, for each time series s_i the corresponding local maxima of all $\langle \varepsilon^\star \rangle(\tau)$ (for a univariate time series there will only be one $\langle \varepsilon^\star \rangle(\tau)$) that determines the set of possible delay values τ_2 (see the rows in Figs. 3.1, 3.2 corresponding to $\mathcal{D}_2$). Then, one of the possible τ_2's is randomly chosen with probabilities computed with a softmax of the corresponding values of Γ_j. Due to its normalization, the softmax function is able to convert all possible values of Γ_j to probabilities with $p_j = \exp(-\beta \Gamma_j) / \sum_k \exp(-\beta \Gamma_k)$.

This procedure is repeated (consecutive rows for $\mathcal{D}_3 \dots$ etc. in Figs. 3.1, 3.2) until the very last computed embedding cycle $\mathcal{D}_{m+1}$. This is, when the objective function Γ_{m+1} cannot be further decreased for any of the τ_{m+1}-candidates. Figure 3.2 visualizes this procedure.

- **Backpropagation:** After the tree is expanded, the final value Γ_m is backpropagated through the taken *path* of this trial, i.e., to all leafs (previous embedding cycles d), that were visited during this expand, updating their Γ_d values to that of the final embedding cycle.

With this two-step procedure, we iteratively build up the part of the tree that leads to embedding with the smallest values for the objective function. The following two refinements are made to improve this general strategy: in case of multivariate time series input, the probabilities are chosen uniformly random in the zeroth embedding cycle $\mathcal{D}_1$. This ensures an even sampling over the given time series, which can all serve as a valid 1st component of the final reconstruction vectors. Additionally, as soon as a Γ_j is found that is smaller than the previous global minimum, this embedding cycle is directly chosen and not randomized via the softmax function. This also means that for the very first trial always the smallest value of Γ_j is chosen, resulting in a good starting point for the further Monte Carlo search of the tree. In case the continuity statistic $\langle \varepsilon^\star \rangle(\tau)$ is used as the delay pre-selection statistic Λ_τ and the ΔL-statistic [315] as the objective function Γ, the first sample thus is identical to the PECUZAL algorithms [162] and every further sample improves upon this embedding further minimizing ΔL. Aside from the choice of Λ_τ and Γ, the two hyperparameters of the method are the number of trials N_{trials} and the β parameter of the probability distribution chosing the next delay value. The parameter β governs how likely it is that the minimum of all Γ_i is chosen, i.e. in the extreme cases for $\beta = 0$ the possible delay times are chosen uniformly random and for $\beta \to \infty$ always the smallest Γ_i is chosen. For the tree search algorithms this means that β governs how "wide" the tree search is, larger β values search the tree more along the already found previously found minima, whereas for smaller values the tree search will stress previously unvisted paths through tree stronger. The default value for β which is used in all shown results is $\beta = 2$.

The computational complexity of this algorithm obviously scales with the number of trials N_{trials}, even though already computed embedding cycles are not computed again in later trials. When sampling the tree many times, the path through the tree of the first few embedding cycles will likely often be the same as that of previous trials. In these cases computing the delay-preselection and objective function will be identical to that of previous trials. All the values of possible delays and values of the objective function that are computed in previous trials are saved during the tree search, and are reused when the same embedding cycle needs to be computed again.

Otherwise the complexity depends on the chosen delay pre-selection function Λ_τ and the objective function Γ. It has to be clear that the algorithm is computationally much more demanding than a classical TDE. However, once an embedding is computed for a specified system it can be reused in later applications.
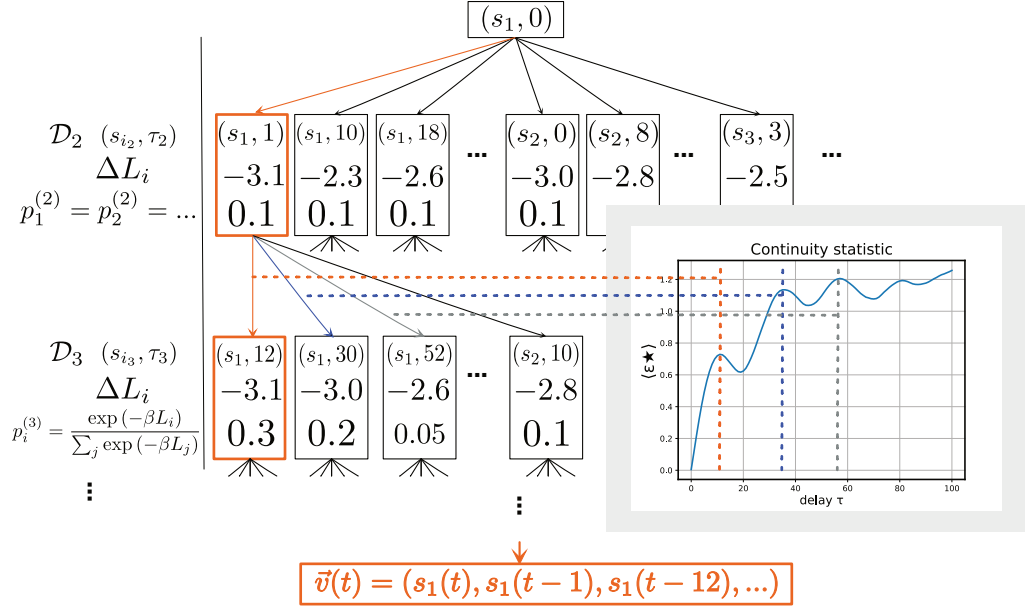
Figure 3.2: Visualization of the expand step of the MCDTS algorithm. Here we exemplary use the continuity statistic $\langle \varepsilon^\star \rangle(\tau)$ as the delay pre-selection statistic Λ_τ and the ΔL-statistic [315] as the objective function Γ, as it has been utilized in the recently proposed PECUZAL algorithm [162], Chapter 2.

Table 3.1: The different implementations of the MCDTS algorithm used in this chapter and their choice of delay-preselection and objective function, which is minimized through the tree search. Here the abbreviations *CS: continuity statistic* and *KLD: Kullback-Leibler divergence* are used.

Notation	delay pre-selection method Λ_τ	objective function Γ
MCDTS-C-FNN	maxima of CS [239, 240]	FNN [125, 153, 154]
MCDTS-C-L	maxima of CS	$L/\Delta L$-statistic [162, 315]
MCDTS-C-MSE	maxima of CS	mean squared prediction error
MCDTS-C-MSE-KL	maxima of CS	mean KLD of true and predicted
MCDTS-C-CCM	maxima of CS	negative CCM-correlation coefficient [288]
MCDTS-R-MSE-KL	given range of delay values	mean KLD of true and predicted
MCDTS-R-MSE	given range of delay values	mean squared prediction error

3.3 Applications

In this section, we present the potential of the proposed MCDTS method by various applications. Here, we aim to provide suggestions and show that there are a number of state-space based applications that directly benefit from our method or provide better results than with the state of the art embedding techniques. A variety of applications are presented to support the fact that different research questions elicit different embedding behavior and that our proposed method is able to optimize the embedding with respect to different study objectives. In particular, we investigate the influence of the state space reconstruction parameters on a recurrence analysis of the chaotic Lorenz-96 system (Section 3.3.1), a nearest-neighbor time series prediction for the chaotic Hénon map and for a palaeoclimate dataset (Sections 3.3.2, 3.3.3), and last but not least, a causal analysis of two physical observables of a combustion process (Section 3.3.4). The selected applications cover many areas of nonlinear time series analysis and it is not our intention here to propose new techniques for prediction or causal analysis which are necessarily superior to other, alternative approaches. We rather chose well established state-space based methods and use them to show how our proposed method optimizes results with respect to the chosen embedding.

3.3.1 Recurrence properties of the Lorenz-96 system

At first, we consider a potentially higher dimensional nonlinear dynamical system and compare the recurrence properties of its dynamics as derived from the original set of system variables with such by applying the different embedding approaches. We utilize the Lorenz-96 system [188], a set of N ordinary first-order differential equations

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F, \tag{3.2}$$

with x_i being the state of the system of node $i = 1, \ldots, N$ and it is assumed that the total number of nodes is $N \geq 4$. We can think of this system as a ring-like structure of N coupled oscillators – each representing some atmospheric quantity – all connected to the same forcing. The forcing constant F serves as the control parameter. Here we vary F from $F = 3.7$ to 4.0 in steps of 0.002 covering limit cycle dynamics as well as chaos. We set $N = 8$, randomly choose the initial condition to $u_0 = [0.590; 0.766; 0.566; 0.460; 0.794; 0.854; 0.200; 0.298]$, and use a sampling time of $\Delta t = 0.1$. By discarding the first 2,500 points of the integration as transients, we get time series consisting of 5,000 samples for each of the encountered values of F. We focus on two scenarios: (1) only the time series of the 2nd node (univariate embedding) and (2) three time series of nodes 2, 4, and 7 are used to mimic a uni- and a multivariate embedding case. For each of these time series we perform an embedding, using three classic time delay approaches as proposed by Kennel et al. [154] (5%-threshold), Cao [39] (slope threshold of 0.2), and Hegger and Kantz [125] (5%-threshold) with a uniform delay value estimated as the first minimum of the auto mutual information (only applicable to the univariate case) and the recently proposed PECUZAL algorithm [162], Chapter 2. For our proposed MCDTS approach we embedd the data using the continuity statistic $\langle \varepsilon^\star \rangle(\tau)$ as the delay pre-selection statistic Λ_τ. For the objective function Γ we try two different approaches, namely the ΔL-statistic [315] (MCDTS-C-L) as well as the FNN-statistic [125] (MCDTS-C-FNN). In all approaches we discard serially correlated points from the nearest neighbor search by setting a Theiler window [292] to the first minimum of the mutual information. An overview over all MCDTS implementations and abbreviations is given in Table 3.1.
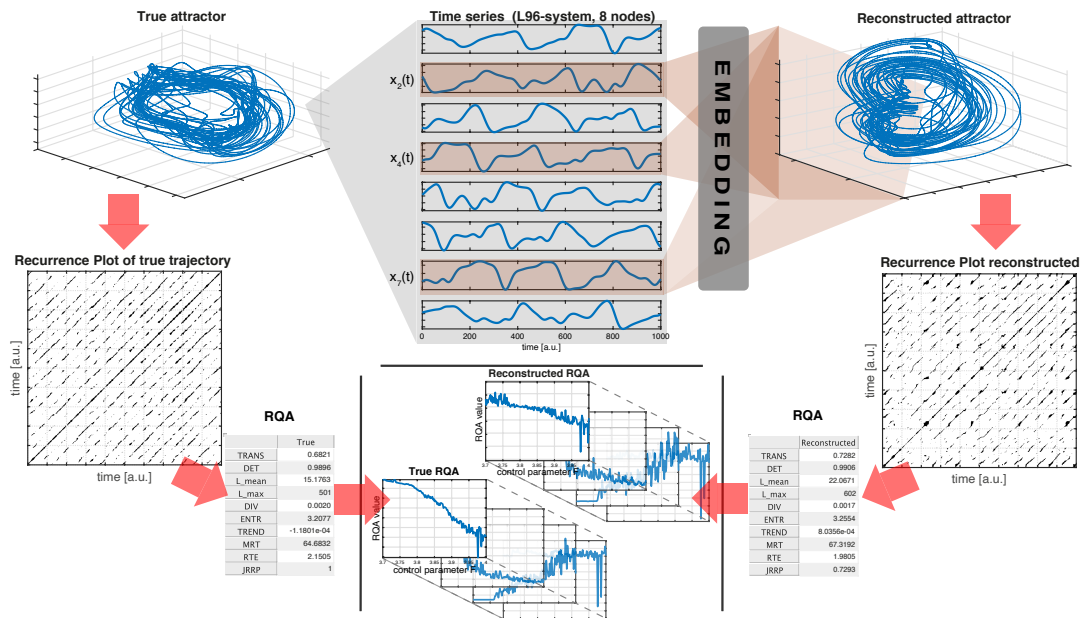
Figure 3.3: Schematic visualization of the data analysis for the Lorenz 96 system, Eq. (3.2) (see text for details). In case of the univariate approach the $x_2(t)$-time series gets embedded by all considered reconstruction methods, for the multivariate approach, three time series ($x_2(t)$, $x_4(t)$ and $x_7(t)$) are passed to the reconstruction algorithms. From the reconstructed attractors we obtain a recurrence plot and quantify it (RQA) by using ten different quantifiers. The same is done for the reference trajectory gained from all 8 time series from the numerical integration. Repeating the analysis for time series corresponding to varying values of the control parameter F of the system, we finally obtain time series of the RQA-quantifiers for each reconstruction method as well as for the *true* trajectory.

By varying the control parameter F, the system varies its dynamics which is well represented by a change in the recurrence behaviour [151]. In previous work we have demonstrated that recurrence quantification analysis (RQA) can be used to qualitatively characterize the typical dynamical properties of the Lorenz-96 system such as chaotic or periodic dynamics [205]. We, therefore, compare the recurrence properties of all reconstructed trajectories to recurrence properties of the true trajectory (obtained from the numerical integration) by using recurrence quantification analysis (RQA). For a sound definition of the used RQA statistics see Appendix A.

For our purpose of comparing different aspects of recurrence properties of original and reconstructed trajectories, we use the transitivity (*TRANS*, Eq. (A.14)) of the ε-RN, the *determinism* (*DET*, Eq. (A.6)), the mean diagonal line length (ℓ_{mean}, Eq. (A.7)), the the maximal diagonal line length (ℓ_{max}) and its reciprocal (*DIV*, Eq. (A.8)), the entropy of diagonal line lengths (*ENTR*, Eq. (A.9)), the *TREND* (Eq.(58) in [209]), mean recurrence time (*MRT*, Eq. (A.12)), the recurrence time entropy (*RTE*, Eq. (A.11)) and the joint recurrence rate fraction (*JRRF*, Eq. (2.6)). As already utilized in Chapter 2, JRRF measures the accordance of the recurrence plot of the (true) reference system, $\mathbf{R}^{\mathrm{ref}}$ with the RP of the reconstruction, $\mathbf{R}^{\mathrm{rec}}$. We compute both, $\mathbf{R}^{\mathrm{ref}}$ and $\mathbf{R}^{\mathrm{rec}}$, by fixing the recurrence threshold corresponding to a global recurrence rate (RR) of 5% in order to ensure comparability [163] (Chapter 4). Although the quantification measures depend crucially on the chosen recurrence threshold, the particular choice we make here is not so important, since we apply it to all RPs we compare. $RR = 5\%$ ensures a proper resolution of the

inherent structures to be quantified by the ten aforementioned measures.

The described procedure is schematically illustrated in Figure 3.3. For each reconstruction method and for each of the ten RQA-statistics the mean squared error (MSE) with respect to the RQA-statistics of the true reference trajectory is computed (normalized to the reference RQA-values). The pairwise comparison of the MSEs is evaluated as the percentage of the ten RQA-MSEs, which take a lower MSE (Fig. 3.4). For instance, a value of 70% in the table indicates that for seven out of the ten considered RQA-quantifiers the normalized mean squared error for the reconstruction method displayed on the *y*-axis is *lower* than for the reconstruction method displayed on the *x*-axis. The *m*-notation indicates the multivariate embedding approach, where three instead of one time series have been passed to the reconstruction methods ($x_2(t)$, $x_4(t)$, and $x_7(t)$, see Fig. 3.3). Since the classic TDE algorithms from Cao, Kennel et al., and Hegger & Kantz are not able to handle multivariate input data, only PECUZAL and the proposed MCDTS-idea combined with the *L*-statistic and with the FNN-statistic are considered in the multivariate scenario. The superiority over the three classic TDE methods is discernible in values $> 50\%$ for PECUZAL and MCDTS in the first three columns. While we would expect a better reconstruction for the multivariate cases – because we simply provide more information – our proposed method also performs better in the univariate case when the FNN-statistic is used as an objective function. When using MCDTS with the *L*-statistic, there is hardly any improvement discernible, while the computational costs are magnitudes higher. Here PECUZAL reveals better results, even though it uses the same statistics. However, combined with the FNN-statistic our proposed idea performs very well in the univariate case and reveals excellent results for the multivariate case.

3.3.2 Short time prediction of the Hénon map time series

In the following, a state space reconstruction $\vec{v}(t)$ of a single time series $s(t)$ is used to further predict its course. Besides a very recent idea [68] to train neural ordinary differential equations on a reconstructed trajectory, which then allows prediction, several attempts have been published [44, 86, 136, 146, 237, 250, 289] which more or less rely on the same basic idea. For the last vector of the reconstructed trajectory, denoted with a time-index l, $\vec{v}(t_l)$, a nearest neighbor search is performed. Then these neighbors are used to predict the future value of this point T time steps ahead, $\vec{v}(t_{l+T})$. Knowledge of the used embedding, which led to the reconstruction vectors $\vec{v}(t)$, then allows to read the prediction of the time series $s(t_l + T)$ from the predicted reconstruction vector $\vec{v}(t_{l+T})$. Usually $T = 1$, i.e., the forecast is iteratively build by appending $\vec{v}(t_{l+T})$ to the trajectory $\vec{v}(t_i)$, $i = 1, \ldots, l$, and this procedure is repeated N times, in order to obtain an N-step prediction. The aforementioned approaches differ from the way they construct a local model of the dynamics based on the nearest neighbors. For instance, Farmer and Sidorowich [86] proposed a linear approximation, i.e., a linear polynomial is fitted to the pairs $(\vec{v}(t_{nn_i}), \vec{v}(t_{nn_i+T}))$, where nn_i denotes the ith nearest neighbor time-index. Sugihara and May [289] used a simplex with minimum diameter to select the nearest neighbor indices nn_i and projected this simplex T steps into the future. The prediction is then being made by computing the location of the original predictee $\vec{v}(t_l)$ within the range of the projected simplex, "giving exponential weight to its original distances from the relevant neighbors". Here a much simpler idea is considered: a *zeroth-order approximation* of the local dynamics. The prediction is simply the projection of the nearest neighbor of $\vec{v}(t_l)$, denoted by the index nn_1, $\vec{v}(t_{l+T}) = \vec{v}(t_{nn_1+T})$. It is clear that the performance of all prediction approaches based on an approximation of the local dynamics by making use of nearest neighbors will crucially depend on the length of the training set. By training set we mean the time series

Comparison of RQA-results with ground truth (Lorenz96-system, 8 nodes)
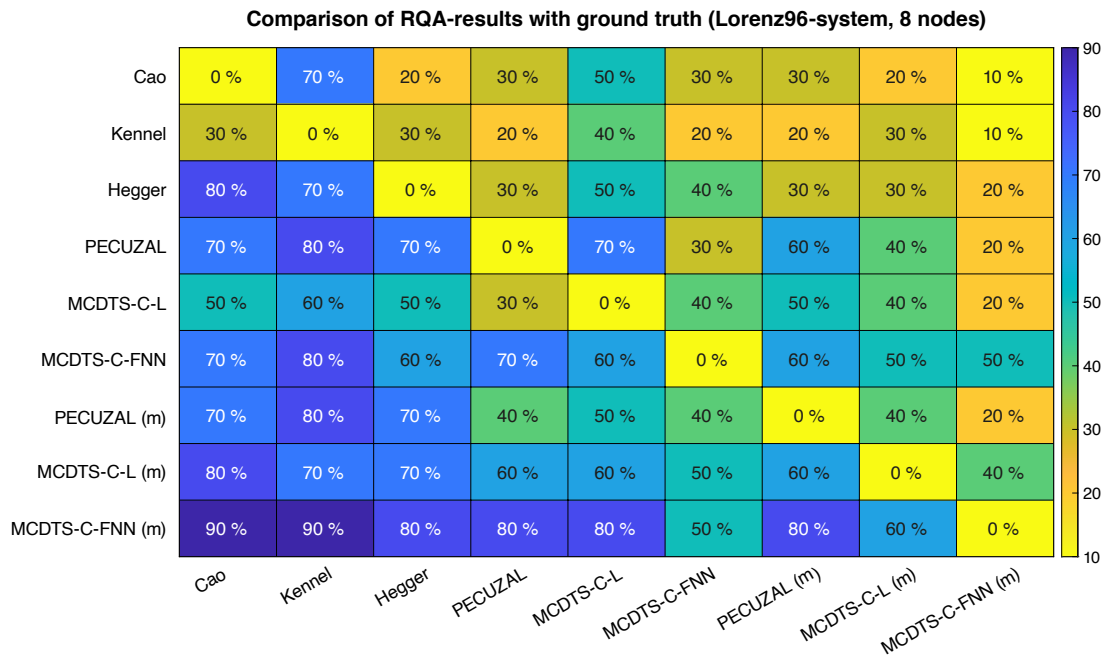
Figure 3.4: Results of the analysis of the Lorenz 96 system with varying control parameter and for all considered reconstruction approaches (see Table 3.1 for notations). Shown is the pairwise comparison of the normalized mean squared error of all considered ten RQA-quantifiers with respect to the truth RQA-time series (see text for details). For instance, a value of 70% in the table indicates that for seven out of the ten considered RQA-quantifiers the normalized mean squared error for the reconstruction method displayed on the y-axis is *lower* than for the reconstruction method displayed on the x-axis.

$s(t)$, which has been used to construct the trajectory $\vec{v}(t)$. We hypothesize that the accuracy of such a prediction will also depend on the reconstruction method, especially when the training set is rather short (Small and Tse [284] and also Bradley and Kantz [24]). In particular, Garland and Bradley [108] have shown that accurate predictions can be achieved with the aforementioned *zeroth-order approximation* when using an incomplete embedding of the data, i.e., reconstructions that do not satisfy the theoretical requirements on the embedding dimension in Takens' sense.

As a proof of concept we now use the described nearest-neighbor prediction method to predict the x-time series of the Hénon map [126], even though other simple models like low order polynomial models might be superior for such noise-free and pure deterministic dynamics (we provide a more challenging example in Section 3.3.3). The time series $x_{i+1} = y_i + 1 - ax_i^2$ and $y_{i+1} = bx_i$, with standard parameters $a = 1.4$, $b = 0.3$ and 100 randomly chosen different initial conditions are used. For each of those 100 samples x- and y-time series of length $N = 10,030$ are obtained (transients removed). The first $10,000$ points of the time series are used for state space reconstruction (both time series for the multivariate cases, only the x-time series in the univariate case), while the last 30 points are the prediction test set (only the x-time series is predicted). The same reconstruction methods as in Section 3.3.1 are used, but for MCDTS we try two different delay pre-selection statistics Λ_τ. Rather than only considering the continuity-statistic (denoted as C in the model description) we also look at a whole range of delay values $\tau = 0, \ldots, 50$ (denoted as R in the model description).

For the objective function Γ we try

- the ΔL-statistic (denoted as L in the model description),

- the FNN-statistic (denoted as *FNN* in the model description),

- the root mean squared in-sample one-step prediction error on the first component of the reconstruction vectors, i.e., the x-time series (denoted as *MSE* in the model description), and finally

- the mean Kullback-Leibler-distance of the in-sample one-step prediction and the "true" trajectory points (denoted as *MSE-KL* in the model description).

By "in-sample" we mean the training set, which is used for the reconstruction. For all MCDTS implementations and abbreviations see again Table 3.1. The accuracy of the prediction is evaluated by the normalized root-mean-square forecast error (*RMS*),

$$
e_{\mathrm{rms}}(T) = \frac{\sqrt{\left\langle [x_{\mathrm{pred}}(T) - x_{\mathrm{true}}(T)]^2 \right\rangle}}{\sqrt{\left\langle [x_{\mathrm{true}}(T) - \langle x_{\mathrm{true}}(T) \rangle]^2 \right\rangle}}
$$

with index *true* denoting the test set values. This way $e_{\mathrm{rms}}(T) = 0$ indicates a perfect prediction, whereas $e_{\mathrm{rms}}(T) \approx 1$ means that the prediction is not better than a constant mean-predictor of the test set. Figure 3.5 shows the mean forecast accuracy for the traditional TDE methods (Cao, Kennel et al., Hegger & Kantz) and two selected MCDTS approaches as a function of the prediction time. The largest Lyapunov exponent is estimated to $\lambda_1 \approx 0.419$ and we display *Lyapunov times* on the x-axis, i.e., units of $1/\lambda_1$. As in Section 3.3.1, m indicates the multivariate case, in which both, x- and y-time series are fed into the reconstruction algorithms. The results for all discussed reconstruction methods can be found in Appendix D.1 (Fig. D.1). As expected, the forecast accuracy is worse in case of added white noise (Fig. 3.5B) and the predictions based on multivariate reconstructions perform slightly better. The MCDTS-based forecasts perform significantly better than the forecasts based on the traditional TDE methods. Even though the continuity statistic constitutes a reasonable delay pre-selection statistic with a clear physical meaning, when utilized in our MCDTS approach (MCDTS-C-) it performs not as good as if we would not pre-select delays on the basis of some statistic, but try delays in a whole range of values ($\tau \in [0, 50]$, MCDTS-R-). At least this statement holds for this example of the Hénon map time series.

A Wilcoxon rank sum test is applied to underpin the better performance of the MCDTS-approaches in comparison to the classical time delay methods. Therefore we define a threshold $\zeta = 0.1$ and compute the prediction times for which $e_{\mathrm{rms}}(T)$ first exceeds ζ for all trials and for all considered reconstruction methods. These distributions of prediction times for each method are used for the statistical test with the null hypothesis that two considered distributions have equal medians. The tests complement the visual analysis of Figs. 3.5 and D.1. A significantly better forecast performance (α=0.01) than the classic time delay embedding methods for PECUZAL and all considered MCDTS-based approaches, but the ones combined with the FNN-statistic (MCDTS-FNN), can be verified for the noise free case. In the case of the noise corrupted time series PECUZAL (m), all MCDTS-MSE-approaches and MCDTS-C-L (m) achieve a significantly better prediction performance than the classical time delay methods.
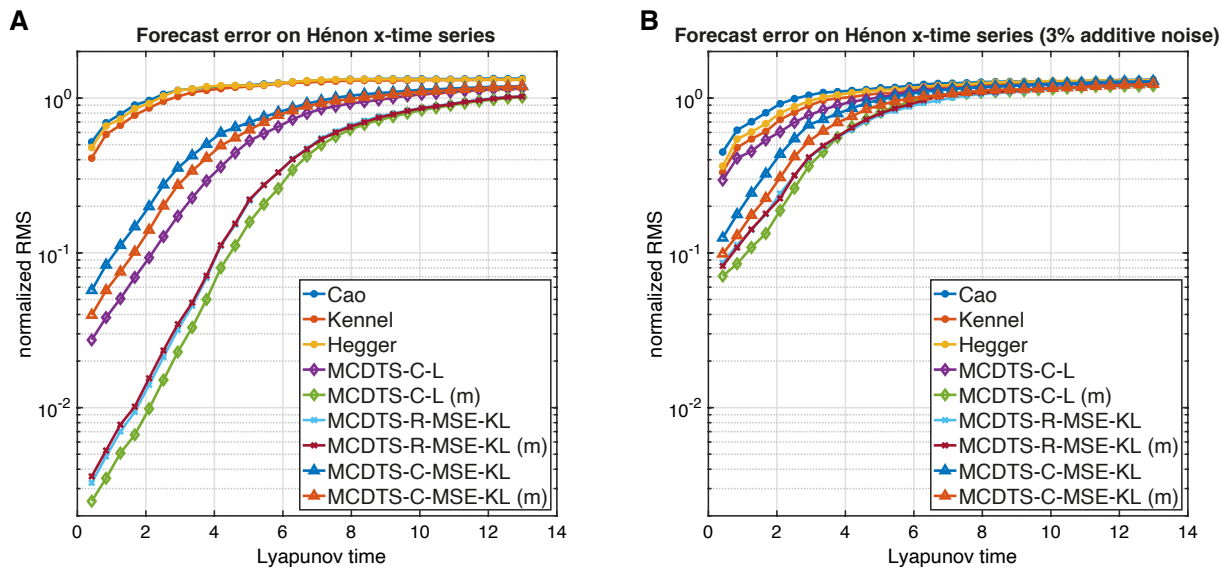
Figure 3.5: A Normalized root-mean-square prediction errors (*RMS*) for the Hénon *x*-time series and for selected reconstruction methods (see Fig. D.1 for all mentioned approaches and Table 3.1) as a function of the prediction time. Shown are mean values of a distribution of 100 trials with different initial conditions. For the prediction we used a one step ahead zeroth-order approximation on the nearest neighbor of the last point of the reconstructed trajectory and iteratively repeated that procedure 30 times in order to obtain a prediction of 31 samples in total for each trial. **B** Same as in **A** but with 3% additive white noise.

Some remarks: Together with PECUZAL (m) and MCDTS-R-MSE (m), MCDTS-C-L (m) achieves the overall best results (Fig. D.1). The choice of the threshold ζ is obviously subjective, but a range of thresholds gave similar results and the "grouping" of the results according to the different techniques is clearly discernible already when looking at the mean (Figs. 3.5, D.1). We have to mention that we could not achieve results as shown here for continuous systems like the Lorenz-63 or the Rössler model. In those cases the difference in the prediction accuracy was not as clear as it is in the Hénon example and not significant, for both, noise-free and noise corrupted time series. We also investigated the influence of the time series length of the training sets, but the results did not change much. All reconstruction methods gave similar prediction results. We could, however, observe that simple and incomplete embeddings, i.e., a too low embedding dimension, often – but not always – led to similarly good prediction results, when compared to "full" embeddings. This was true for the continuous examples (not shown in this work), but this also holds for the Hénon example shown here, where the MCDTS-C-L approach does not yield the best results in the univariate case, although it targets the total minimum of the *L*-objective-function, which the authors consider to be a suitable cost-function for a good/full embedding. These observations are in line with the findings of Garland and Bradley [108] and the fact that our reconstruction methods tend to suggest higher dimensional embeddings with smaller delays in the presence of noise support the findings of Small and Tse [284]. The FNN-statistic does not seem to be useful in the prediction application shown here, since all approaches which make use of it (including classic TDE) perform clearly worse compared to the other methods used.

3.3.3 Improved short time predictions for CENOGRID

To demonstrate that the prediction procedure from the preceding section works for real, noisy data, we apply it to the recently published CENOzoic Global Reference benthic foraminifer carbon and oxygen Isotope Dataset (CENOGRID) [328]. The temperature-dependent fractionation of carbon and oxygen isotopes in benthic foraminifera is an important means to reconstruct past global temperatures and environmental conditions. Moreover, the Cenozoic is interesting, because it provides an analogue of future greenhouse climate and how and which regime shifts in large-scale atmospheric and ocean circulation can be expected in the future warming climate. Predicting these data may be unrealistic and not motivated by an actual research question. However, this task shall serve as a proof of concept. The non-stationarity and noise level of CENOGRID make prediction particularly difficult.
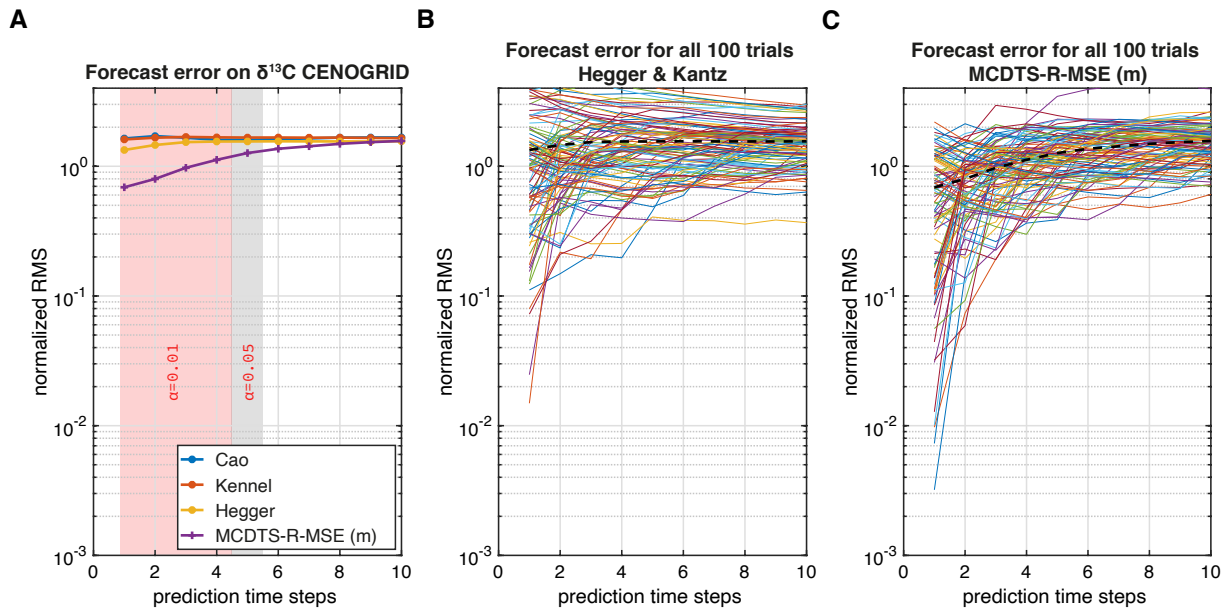
Figure 3.6: A Mean normalized root mean square prediction error for four selected reconstruction methods on the δ^{13}C CENOGRID record. **B** Prediction error for all 100 trials for the classic TDE method of Hegger and Kantz [125] (yellow line in panel A). **C** Prediction error for all 100 trials for the MCDTS-R-MSE (m) method (purple line in panel A). The forecasts based on this method are significantly better than for all three classic TDE methods (up to 4 prediction time steps under a significance level $\alpha = 0.01$ and up to 5 prediction time steps under a significance level $\alpha = 0.05$).

The dataset consists of a detrended δ^{18}O and a detrended δ^{13}C isotope record with a total length of $N = 13,421$ samples and a sampling period of $\Delta t = 5,000$yrs (Figure D.2 in Appendix D.2). Here we made predictions on the δ^{13}C isotope record. The first $13,311$ samples have been used as a training set, from which state space reconstructions were obtained. The remaining 110 samples of the δ^{13}C record acted as the test set. For 100 different starting points in the test set we have made 10-step-ahead predictions for each reconstruction method by using the embedding parameters gained from the training and with the iterative *zeroth-order approximation* prediction procedure described in Section 3.3.2. This way we have simulated different initial conditions for the prediction and obtained a distribution of forecasts for each reconstruction method. We again

used a Wilcoxon rank sum test on these distributions in order to see whether predictions based on some reconstruction method are significantly better than the predictions obtained from classic TDE (Cao, Kennel et al., Hegger & Kantz). Only one of the applied reconstruction methods (listed in Table 3.1), MCDTS-R-MSE (m), scored significantly better predictions (highly significant for prediction horizons up to $4\Delta t$ and significant for prediction horizon up to $5\Delta t$). Figure 3.6A shows the mean normalized root mean square prediction error gained from the 100 predictions for the classic TDE and the mentioned MCDTS-R-MSE (m). The distribution of all prediction trials for the best performing classic TDE method (Hegger & Kantz) and for MCDTS-R-MSE (m) are shown in panels B, C. Even though the multivariate approach MCDTS-R-MSE (m) could have been used both, the $\delta^{18}O$ and the $\delta^{13}C$ time series for the reconstruction, it only used $\delta^{13}C$ lagged by 1 and 2 samples in a 3-dimensional reconstruction. The classic TDE methods and all other reconstruction methods (listed in Table 3.1, not shown in Fig. D.2) yielded higher dimensional embeddings (Table D.1). Yet, all these higher dimensional reconstructions gave poor prediction results, except for MCDTS-C-MSE-KL (m), which gave significant better predictions ($\alpha = 0.05$) than the classic TDE methods at least for the one-step-ahead prediction.

3.3.4 Estimating causal relationship of observables of a thermoacoustic system

As a final proof of concept we utilize state space reconstruction for detecting causality between observables X and Y in a turbulent combustion flow in a gas turbine. It is possible to infer a causal relationship between two (or more) time series $x(t)$ and $y(t)$ via *convergent cross mapping (CCM)* [55, 288, 333], which – in contrast to Granger causality [114] – also works for time series stemming from non separable systems, i.e., deterministic dynamical systems. The CCM method "tests for causation by measuring the extent to which the historical record of Y values can reliably estimate states of X. This happens only if X is causally influencing Y."[288] This also incorporates the embedding theorems [196, 291, 329] in a sense that a state space reconstruction based on $x(t)$ is diffeomorphic to a reconstruction of $y(t)$, if $x(t)$ and $y(t)$ describe the same dynamical system *and* the embedding parameters have been chosen correctly. To check for a causal relationship from $X \rightarrow Y$, a state space reconstruction of $y(t)$ yields a trajectory $\vec{v}_y(t) \in \mathbb{R}^m$, with m denoting the embedding dimension, which is then used for estimating values of $x(t)$, namely $\hat{x}(t)$. It is said that $\vec{v}_y(t)$ *cross-maps* $x(t)$, in order to get estimates $\hat{x}(t)$. Technically, this is done by first searching for $m + 1$ nearest neighbors of a point corresponding to a time index $t' \in t$, i.e., find the $m + 1$ time indices $t'_{NN_i}, i = 1, \ldots, m + 1$ of the nearest neighbors of $\vec{v}_y(t')$. Further, these time indices t'_{NN_i} are used to "identify points (neighbors) in X (a putative neighborhood) to estimate $x(t')$ from a locally weighted mean of the $m + 1$ $x(t'_{NN_i})$ values"[288]:

$$\hat{x}(t') = \sum w_i x(t'_{NN_i}), \qquad i = 1, \ldots, m + 1, \tag{3.3}$$

with the weighting w_i based on the nearest neighbor distance to $\vec{v}_y(t')$.

$$w_i = u_i / \sum u_j, \qquad j = 1, \ldots, m + 1 \tag{3.4}$$

$$u_i = \exp\left[-\|\vec{v}_y(t') - \vec{v}_y(t'_{NN_i})\| \,/\, \|\vec{v}_y(t') - \vec{v}_y(t'_{NN_1})\| \right] \tag{3.5}$$

with $\|\cdot\|$ a norm (we used Euclidean distances). Finally, the agreement of the cross-mapped estimates $\hat{x}(t')$ with the true values $x(t')$ is quantified for all considered $t' \in t$, e.g., by computing a linear Pearson correlation ρ_{CCM}, which has been done in this study. The clou is that the
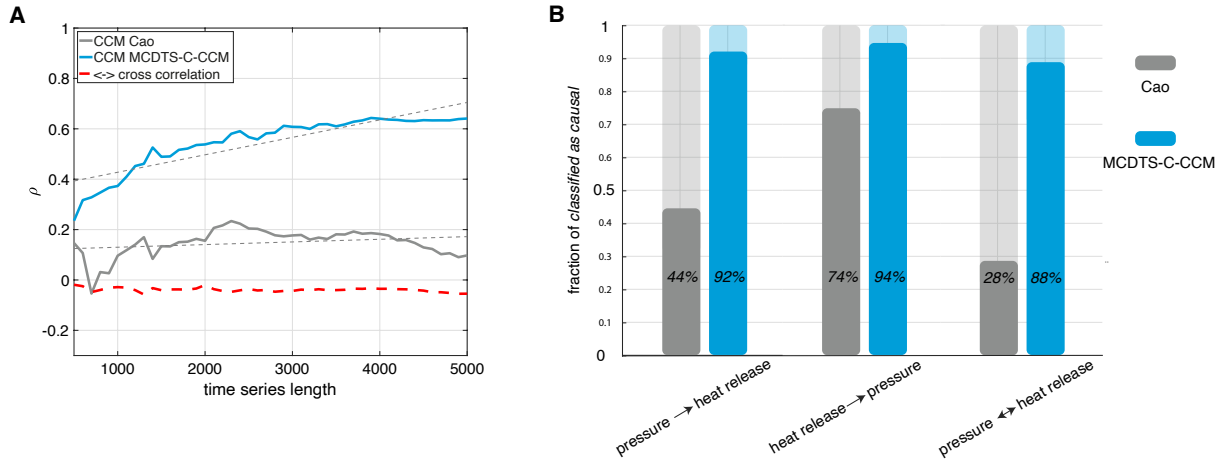
Figure 3.7: A Linear correlation coefficient of convergent cross mapping (CCM) *heat release → pressure* as a function of the considered time series length for Cao's embedding method (gray) and the proposed MCDTS embedding (blue) exemplary shown for on one out of 50 drawn sub-samples of length $N = 5,000$ from the entire time series (Fig. D.3, c.f. Table 3.1 for abbreviations). While the dashed black lines show the linear trend for both CCM correlations, the dashed red line shows the Pearson linear correlation between the heat release and the pressure time series, indicating no influence. We ensured convergence of the cross mapping, and, thus, a true causal relationship, if there was a positive trend in the CCM-correlation over increasing time series length (slope of the dashed black lines) and when the last point of the CCM-correlation (i.e., longest considered time series length) exceeded a value of 0.2 (in the shown case Cao's method did not detect a causal influence of the heat release to the pressure). We tested this on all 50 sub-samples for both causal directions. **B** True classified causal relationships as a fraction of all sub-samples based on the embedding of each time series using Cao's method and our proposed MCDTS method.

estimation skill, here represented by ρ_{CCM}, increases with the considered amount of data used, if X indeed causally influences Y. This is because the attractor – represented by the reconstruction vectors $\vec{v}_y(t)$ – gets resolved better with increasing time series length, resulting in closer nearest neighbors and therefore a better concordance of $\hat{x}(t)$ and $x(t)$, i.e., an increase of ρ_{CCM} with increasing time series length. This *convergence* of the estimation skill based on cross-mapping is a necessary condition for causation, not only a high value of ρ_{CCM} itself (Fig. 3.7A). Although the embedding process is key to a successful application of CCM to data, its influence has not been discussed by Sugihara et al. [288]. However, Schiecke et al. [270] discussed the impact of the embedding parameters on CCM briefly and we hypothesize that the embedding method can play a crucial role, when analyzing real world data. Therefore, we utilize the MCDTS framework in the following way. As a delay pre-selection method Λ_τ, we use the reliable continuity statistic $\langle\varepsilon^\star\rangle(\tau)$ [239, 240]. As a suitable objective function Γ, we use the negative of the corresponding ρ_{CCM}, i.e., MCDTS optimizes the embedding with respect to maximizing ρ_{CCM} of two given time series. According to our abbreviation-scheme given in Table 3.1 we will refer to this approach as MCDTS-C-CCM.

We apply the CCM-method to time series data that spans the different dynamical regimes of a thermoacoustic system. Here, we investigate the mutual causal influence of two recorded variables of the thermoacoustic system, namely the pressure and the heat release rate fluctuations (Fig. D.3). The original experiments were performed on a turbulent combustor with a rectangular combustion chamber (length 700 mm, cross-section 90 mm × 90 mm, Fig. D.4). In such a

combustion experiment, a fixed vane swirler is used to stabilize the flame and a central shaft that supports the swirler injects the fuel through four radial injection holes. The fuel used is liquefied petroleum gas (60% butane and 40% propane). The airflow enters through the inlet to the combustion chamber. The partially premixed reactant mixture is ignited using a spark plug. Once the flame is established in the combustor, we continuously varied the control parameter (mass flow rate of air, which, in turn, varies the Reynolds number[5] and the equivalence ratio[6]) to observe the dynamical transitions in the system. Acoustic pressure fluctuations were measured using a piezoelectric transducer (PCB103B02) and heat release rate using a photomultiplier tube (Hamamatsu H10722-01) at a sampling rate of 4 kHz.

The interactions between the turbulent flow, the unsteady fluctuations of the flame due and the acoustic field of the chamber lead to different dynamical states. As the airflow rate increases, the system transitions from a state of stable operation (which comprises high dimensional chaos having low amplitude [302]) to intermittency, a state that comprises bursts of periodic oscillations amid epochs of aperiodicity [228], and then to limit cycle [290]. The self-sustained limit cycle oscillations represent a state of oscillatory instability, known as thermoacoustic instability [144]. When the flow rate of air is further increased, the flame loses its stability inside a combustor and blows out. The pressure and heat release rate data capture the transition through all these dynamical states in sequence. In the many different dynamical regimes recorded in the time series, we expect the strength of causal interference between the heat release and the pressure to vary. But in all dynamics we expect a mutual causal interaction between heat release and pressure. Moreover, since a possible asymmetric bi-directional coupling between heat release and pressure has been discovered in a stationary setup of a very similar experiment [112] we would also expect that the heat release rate has a slightly stronger causal influence on pressure than vice versa.

In short, the goal here is twofold:

1. Prove the expected mutual causal relationship between heat release rate and pressure as well as

2. the hypothesized asymmetry in its strengths by applying MCDTS-C-CCM on a range of time series, sampled from the entire record (Fig. D.3).

We compare it to results obtained from using the CCM method with the classical embedding approach of Cao [39]. Specifically, we set up the following workflow for this analysis:

1. 50 time indices $t' \in t$ are drawn randomly, where t covers the entire record.

2. For each of these indices t', time series of length $N = 5,000$ for pressure and heat release are obtained and standardized to zero mean and unit variance (Fig D.3).

3. Both time series samples (of full length $N = 5,000$) each are embedded using Cao's method as a classical reference and our proposed framework MCDTS-C-CCM with 100 trials (Table 3.1). Based on the obtained reconstructions $\rho_{\text{CCM-Cao}}$ and $\rho_{\text{CCM-MCDTS}}$ are computed for both directions as a function of increasing time series length as exemplary shown in Fig. 3.7A.

---

[5]Reynolds number is $\frac{\rho U D}{\mu}$, where ρ is the density, U is a characteristic velocity, D is a characteristic dimension (the diameter) and μ is the viscosity.
[6]Equivalence ratio is the ratio between the actual fuel-air ratio to the stoichiometric fuel-air ratio.

CCM-ρ difference of both causal directions (heat release →pressure & pressure →heat release)
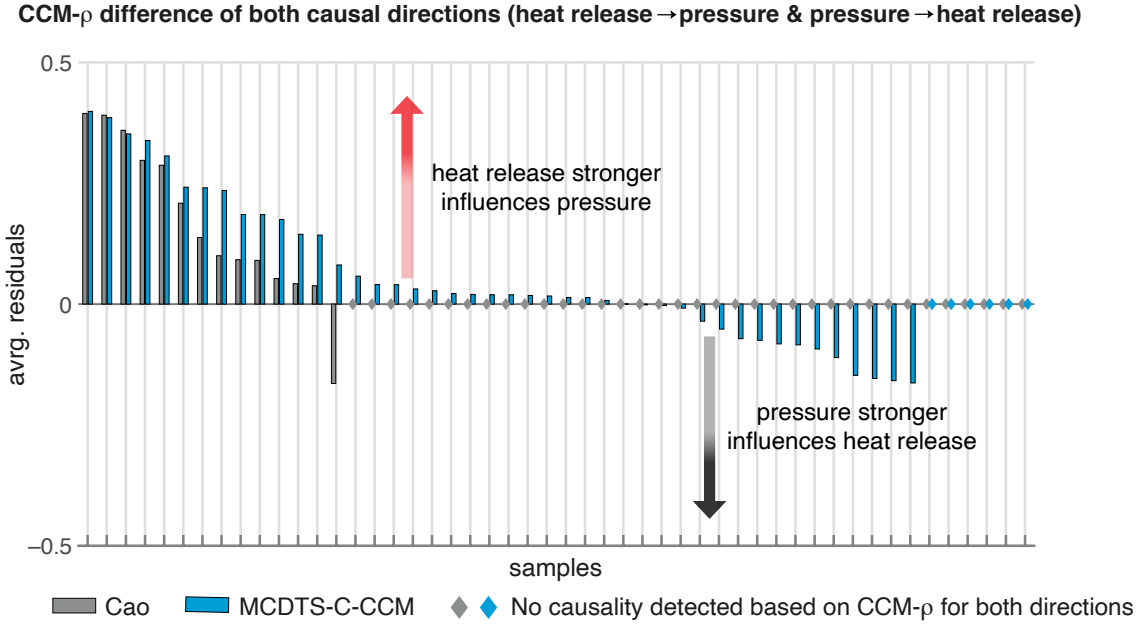
Figure 3.8: Average pointwise difference of the CCM-correlation coefficients for the direction *heat release →
pressure* and vice versa for both underlying reconstruction approaches. For a better visualization
we sorted these values here separately for both methods. A positive value indicates that the heat
release has a stronger causal influence on pressure than vice versa, which is the expectation value.
Diamonds indicate cases, where we could not deduce a causal relationship for both directions in
one sample. As also shown in the right panel of Fig. 3.7B, MCDTS-C-CCM was able to correctly
detect a mutual causal relationship in 88% of all considered samples (only 12 % marked with blue
diamonds in this Figure), whereas in the case of Cao's reconstruction approach, we could only
detect this in 28% of all cases (72 % marked with gray diamonds in this Figure).

4. To ensure convergence in the CCM-sense we fit a linear model to ρ_{CCM} (dashed black lines
   in Fig. 3.7A) and whenever that model gives a positive slope *and* the last value of ρ_{CCM}
   (i.e., for the longest considered time series of length $N = 5,000$) exceeds a value of 0.2, we
   infer a true causal relationship.

5. When we can detect a causal relation simultaneously in both directions, we compute the
   average of the pointwise difference $\rho_{\text{CCM }heat \rightarrow pressure} - \rho_{\text{CCM }pressure \rightarrow heat}$

The minimum considered value of 0.2 for ρ_{CCM} is an arbitrary and subjective choice and we
could have made other choices. But since this procedure is applied to $\rho_{\text{CCM-Cao}}$ and $\rho_{\text{CCM-MCDTS}}$
at the same time, we think this is reasonable and it prevents samples to be accounted for as
"true causal" when there is near-0 ρ_{CCM}, but a positive linear trend. Results do only change
slightly when varying this value in some interval [0.2 0.3]. Figure 3.7B summarizes the results
obtained for both considered embedding methods. Shown are the classification results for correctly
deducing a causal influence of pressure on heat release (left panel) and of heat release on pressure
(middle panel) based on our definition (item 4 in the list above). Thus, in this first step, we do
not measure the strength of the causal relationship, but rather test whether such a relationship
actually exists. While MCDTS-C-CCM maintains a correct classification in 92% of all cases
considered (50 samples) for *pressure → heat release* and 94% for *heat release → pressure*, Cao's

method is only able to correctly classify 44% and 74%, respectively. These results themselves already demonstrate a clear advantage of our proposed method, but recall that we expect a causal relationship between heat release *and* pressure simultaneously for each sample. The right panel of Figure 3.7B reveals that in 88% of all cases considered, MCDTS-C-CCM is able to detect a mutual causal relationship, while Cao's method managed to do so in only 28% of the cases.

Furthermore we try to validate a hypothesis made by Godavarthi et al. [112] that heat release has a stronger effect on pressure than vice versa for most of the considered dynamics. The problem of measuring the strength of a causal relationship is two-fold: First, the experiment considered here exhibits a number of different dynamics due to the continuously changing control parameter. The hypothesis of an asymmetry in the strength of the interaction was made for stationary cases and four considered dynamics the authors investigated. Second, in describing the CCM method, Sugihara et al. [288] merely described that in the case of a stronger causal effect of X on Y, cross-mapping X with $\vec{v}_y$ converges *faster* than the other way around. Thus, we would have to define what *faster* means with respect to our experimental curves like the ones shown in Figure 3.7A. That would mean introducing some parameters on which the results would depend too much. Here we pursue a simpler idea in order to detect the strength of a causal interaction. For samples where a causal relation in *both* directions has been detected, we compute the average of the pointwise difference of the CCM-correlation coefficients, i.e., $\Delta\rho_{\mathrm{CCM}} = \rho_{\mathrm{CCM}\ Heat\rightarrow pressure} - \rho_{\mathrm{CCM}\ pressure\rightarrow heat}$. When this difference is positive, we claim that heat release stronger effects pressure in a causal sense than vice versa. Our analysis reveals that the proposed method is able to reflect the hypothesized stronger causal effect of the heat release on pressure data. Figure 3.8 shows that for 29 of the 50 samples ($\sim 58\%$) $\Delta\rho_{\mathrm{CCM}}$ is indeed positive. Using the Cao method, we were able to derive such a result in only $\sim 26\%$ of all samples. In this case, however, only $\sim 28\%$ of the samples were found to be mutually causally related at all (cf. Fig. 3.7B). Within the group of mutually causally related samples the assumed asymmetry is reflected very well (13 of 14 mutually causally related samples had a positive $\Delta\rho_{\mathrm{CCM}}$).

The proposed MCDTS reconstruction approach shows a clear advantage when using it together with the CCM method. Not only is the general classification ability remarkable, but the MCDTS reconstructions also allow verification of an assumed asymmetric causal interaction, which would be limited by the classical time delay method.

3.4 Conclusions

A novel perspective of the embedding process has been proposed, in which the state space reconstruction from single time series can be treated as a game, in which each move corresponds to an embedding cycle and is subject to an evaluation through an objective function. It is possible to model different embeddings, i.e., different choices of delay values and time series (if there are multivariate data at hand) in the embedding cycles, in a tree like structure. Consequently our approach randomly *samples* this tree, in order to ensure the finding of a global minimum of the chosen objective function. We leave it to practitioners which state space evaluation statistic, i.e., objective function, they use, since different research questions require different reconstruction approaches. There is also a free choice of a delay pre-selection method for each embedding cycle, e.g., using the minima of the auto-mutual information statistic. We recommend the combination of the continuity statistic of Pecora et al. [240] as a delay pre-selection method together with the L-statistic of Uzal et al. [315] as an objective function as a very good "all-rounder" for many research questions in nonlinear time series analysis, as already shown by Kraemer et al. [162]

(PECUZAL algorithm), Chapter 2. Since the sampling of the tree is a random procedure the proposed idea only yields converging embedding parameters for a sufficient sampling size N_{trial}. In our numerical investigations, $N_{\text{trial}} = 50$ usually led to satisfying results for univariate cases and $N_{\text{trial}} = 80$ for multivariate embedding scenarios. Our proposed method initializes in a local minimum of the objective function, which is achieved by minimizing the objective function in each embedding cycle to the maximum extend. So in practice, even setting N_{trial} too low would lead to similar – but never worse – results as the state of the art methods. Moreover, the proposed method is not limited to delay pre-selection and objective functions that take into account certain physical constraints. It would also optimize the reconstruction vectors of the state space for research questions such as classification, where we could speak of a feature or latent space instead of the state or phase space notation associated with statistical physics. We exemplified the use of such a modular algorithm by combining different objective- and delay pre-selection functions. Its superiority to classical time delay embedding methods has been demonstrated for a recurrence analysis of the Lorenz-96 system, a prediction of the x-time series of the chaotic Hénon map and the δ^{13}C CENOGRID record as well as on studying causal interactions between variables in a combustion process.

With these applications we showed the advantage MCDTS brings for any kind of method that utilizes an embedding such as recurrence analysis, embedding-based predictions of time series, or causal analysis with convergent cross mapping. It, thus, has potential in many applications and disciplines, everywhere where such phase space based approaches are used but an automatic phase space reconstruction is required. The latter is of increasing interest, e.g., for big data analysis, analysis with highly reliable requirements (e.g., in medical applications), and also for deep learning based frameworks.

Part II

Recurrence plot related investigation

Chapter 4

Recurrence threshold selection for obtaining robust recurrence characteristics in different embedding dimensions

Abstract

The appropriate selection of recurrence thresholds is a key problem in applications of recurrence quantification analysis (RQA) and related methods across disciplines. Here, we discuss the distribution of pairwise distances between state vectors in the studied system's state space reconstructed by means of time-delay embedding as the key characteristic that should guide the corresponding choice for obtaining an adequate resolution of a recurrence plot. Specifically, we present an empirical description of the distance distribution, focusing on characteristic changes of its shape with increasing embedding dimension. Our results suggest that selecting the recurrence threshold according to a fixed percentile of this distribution reduces the dependence of recurrence characteristics on the embedding dimension in comparison with other commonly used threshold selection methods. Numerical investigations on some paradigmatic model system with time-dependent parameters support these empirical findings.

4.1 Introduction

Recurrence plots (RPs) provide an intuitive tool for visualizing the (potentially multi-dimensional) trajectory of a dynamical system in state space. In case only univariate observations of the system's overall state are available, time-delay embedding has become a standard procedure for qualitatively reconstructing the dynamics in state space. The selection of a threshold distance ε,

which distinguishes close from distant pairs of (reconstructed) state vectors, is known to have a substantial impact on the recurrence plot and its quantitative characteristics, but its corresponding interplay with the embedding dimension has not yet been explicitly addressed. Here, we point out that the results of RQA and related methods are qualitatively robust under changes of the (sufficiently high) embedding dimension only if the full distribution of pairwise distances between state vectors is considered for selecting ε, which is achieved by consideration of a fixed recurrence rate.

A vector time series $\{\vec{x}_i\}_{i=1}^N$ (with $\vec{x}_i = \vec{x}(t_i)$) provides an approximation of a specific trajectory of a given dynamical system in finite-time and (for time-continuous dynamical systems) finite-resolution. In many real-world applications, however, inferring complete dynamical information from observations is hampered by the fact that only some of the dynamically relevant variables are directly observable. In such cases, it has been demonstrated [291] that it is possible to qualitatively reconstruct representations of the unobserved components of a higher-dimensional system by means of embedding techniques applied to a suitably chosen individual component [184]. Specifically, time-delay embedding has become a widely utilized method in nonlinear time series analysis, where a series of univariate observations $\{x_i\}$ (the actual time series at hand) is unfolded into a sequence of m-dimensional state vectors $\{\vec{x}_i\}$ [235, 291] defined as $\vec{x}_i = (x_i, x_{i-\tau}, \ldots, x_{i-(m-1)\tau})^T$, where m and τ denote the chosen embedding dimension and embedding delay, respectively. There are more sophisticated methods for reconstructing the state space, which can also handle multivariate time series input. An overview is given in Section 2.1.

According to the definition of a RP (Eq. (1.1)), for a given time series the recurrence matrix depends on the chosen recurrence threshold ε together with the selected norm $\|\cdot\|$ used for defining pairwise distances between the state vectors. In this work, we will restrict ourselves to two of the most commonly used norms: the Euclidean (L_2) and maximum (L_∞, supremum, Chebychev) norms. Specifically, we will study how the distributions of pairwise L_2 and L_∞ distances depend on the embedding dimension.

Previous studies have provided various complementary suggestions for (i) selecting the right method of determining the recurrence threshold and (ii) choosing its actual value in some automatic way based on the specific properties of the system under study. Corresponding approaches include the spatial extent of the trajectory in the (reconstructed) state space [221, 336], signal to noise ratio [272, 295, 336, 337], the specific dynamical system underlying the time series under investigation [215, 337], or properties of the associated recurrence network [74, 204] with adjacency matrix $\mathbf{A}_{i,j} = \mathbf{R}_{i,j} - \delta_{i,j}$ (with $\delta_{i,j}$ being the Kronecker symbol) like the percolation threshold [73, 139], second smallest eigenvalue of the graph's Laplacian [80], breakdown of ε^{-1} scaling of the average path length [73], or information-theoretic characteristics [331]. In practice, the appropriate choice of the method for determining the recurrence threshold, as well as its resulting value itself, can depend on the specific problem under study and take any of the above criteria or even some multiple-objective considerations based on different criteria into account. To this end, a general solution to the second problem of selecting a specific value of ε has not yet been obtained, and we will also not address this problem specifically in the course of this Chapter. Instead, we are attempting to provide some further insights into the first, more conceptual problem setting, i.e., which type of approach for selecting recurrence thresholds should be taken in case of varying situations such as different embedding dimensions [73, 74, 204, 331].

As we will further detail in the course of this chapter, some previously suggested approaches [160, 221, 272, 336] to link a recurrence threshold to a certain percentage of the maximal or mean
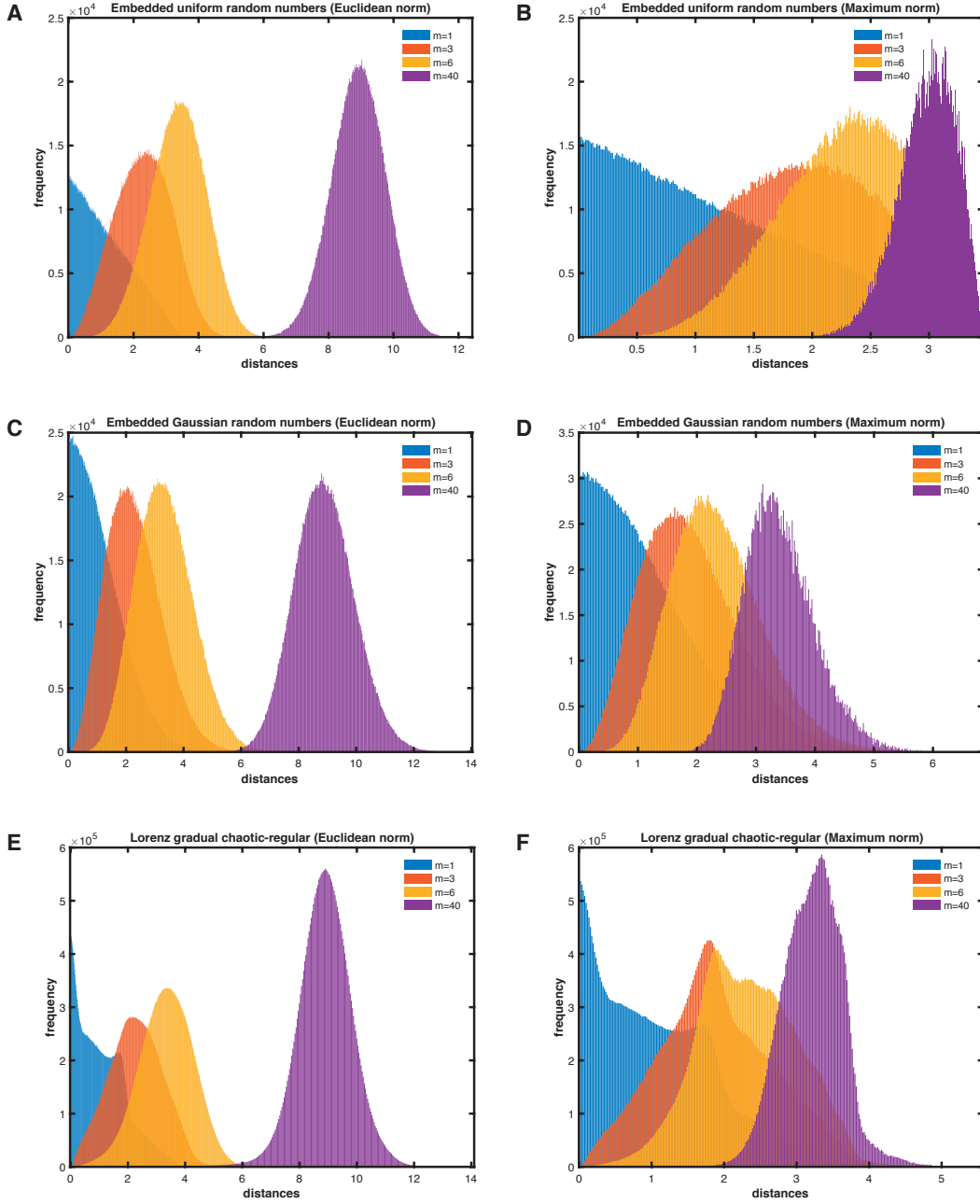
Figure 4.1: Selected histograms of the L_2 (**A**,**C**,**E**) and L_∞ (**B**,**D**,**F**) distances of $N = 1,500$ independent random numbers with uniform **A**+**B** and Gaussian **C**+**D** distribution as well as **E**+**F** for the y component of the Lorenz-63 system (Eq. (B.1), $N = 6,000$, see Section 4.3) with control parameters $\sigma = 10$, $\beta = 8/3$ and r linearly increasing from 180 (chaotic regime) to 210 (periodic regime), for different embedding dimensions m.

distance of all pairwise distances of state vectors (i.e., a given fraction of the attractor's diameter in the reconstructed state space) cause the resulting recurrence characteristics to strongly depend on the embedding dimension. The reason for this behavior is that in addition to a general increase of distances – depending on the chosen norm – [160, 339], the shape of the distance distribution also changes with increasing embedding dimension (see Fig. 4.1 and further discussions in Section 4.2).

It should be noted that embedding a time series with $m \sim \mathcal{O}(10^1)$ or even larger can become necessary when the correlation dimension D_2 of the attractor is rather large. This is due to the fact that Takens' theorem (and several extensions thereof) guarantee the existence of a diffeomorphism between the original and the reconstructed attractor if m satisfies $m \geqslant 2D_2 + 1$ [209, 268, 291]. Hegger et al. [124] emphasize that it is also advisable to choose a rather high value of m when dealing with time series originating from a D-dimensional deterministic system that is driven by P slowly time dependent parameters. An appropriate yet conservative choice for m then fulfills $m \geqslant 2(D + P)$. Concerning practical applications of nonlinear time series analysis, one commonly deals with signals originating from complex, non-stationary systems and, therefore, high embedding dimensions can become necessary, requiring threshold selection methods which lead to robust results of RQA and related state space based techniques that are robust under different choices of the embedding dimension.

In the following Section 4.2, we study the influence of an increasing embedding dimension on the shape of the distance distribution in more detail. We deduce that, in order to avoid problems arising due to an unfavorable fixed recurrence threshold when varying m, we could choose ε as a certain percentile of the distance distribution rather than a certain percentage of the maximum or mean phase space diameter. Successively, Section 4.3 presents a numerical example of a classical Lorenz-63 system with a time-dependent parameter, illustrating that the changes in some recurrence characteristics with varying embedding dimension are particularly small under a fixed recurrence rate in comparison with other strategies. Finally, the main results of this study are summarized in Section 4.4.

4.2 Influence of embedding dimension on the distance distribution

Let us consider a univariate time series $\{x_i\}$ of length N. As an overarching question, we study the effect of time-delay embedding on the distribution of all pairwise distances of its reconstructed state vectors. The variation of this distribution with increasing embedding dimension m is expected to depend on the chosen norm used for the calculation of distances. Note that the effective number of state vectors $N_{\text{eff}}(m) = N - (m-1)\tau$ available for estimating the probability distribution of distances in m dimensions will decrease with m. In order to avoid sample size effects in comparing the results for different m, we therefore choose N sufficiently large so that $1 - N_{\text{eff}}(m_f)/N \ll 1$, where m_f is the largest considered embedding dimension.

4.2.1   Maximum norm

Numerical results for different types of systems demonstrate (see Appendix E.1) that the largest of all pairwise L_∞ distances, $d_{max}^{(\infty)}$, stays constant with increasing embedding dimension, whereas the mean of all pairwise L_∞ distances, $d_{mean}^{(\infty)}$, monotonically increases with m (Fig. E.1). In order

to understand this observation, recall that the L_∞ distance between two embedded state vectors $\vec{x}_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,m})^T$ and $\vec{x}_j = (x_{j,1}, x_{j,2}, \ldots, x_{j,m})^T$ is

$$\|\vec{x}_i - \vec{x}_j\|_\infty = \max_{k=1,\ldots,m} |x_{i,k} - x_{j,k}| = d_{i,j}^{(\infty)}(m) \tag{4.1}$$

For $m = 1$ (i.e., no embedding), the distance between two observations at times t_i and t_j therefore is simply $d_{i,j}^{(\infty)}(1) = |x_i - x_j|$. For $m = 2$, we find

$$\begin{aligned} d_{i,j}^{(\infty)}(2) &= \max\{|x_i - x_j|, |x_{i+\tau} - x_{j+\tau}|\} \\ &= \max\left\{d_{i,j}^{(\infty)}(1), |x_{i+\tau} - x_{j+\tau}|\right\} \geqslant d_{i,j}^{(\infty)}(1). \end{aligned} \tag{4.2}$$

By induction, we can easily show that

$$d_{i,j}^{(\infty)}(m) = \max\left\{d_{i,j}^{(\infty)}(m-1), \left|x_{i-(m-1)\tau} - x_{j-(m-1)\tau}\right|\right\}$$

and therefore

$$d_{i,j}^{(\infty)}(m) \geqslant d_{i,j}^{(\infty)}(m-1) \qquad \forall\, m > 1. \tag{4.3}$$

Hence, considering all possible pairs of state vectors $(\vec{x}_i, \vec{x}_j)$ from the time series, the largest L_∞ distance

$$d_{max}^{(\infty)}(1) = \max_{i,j}[d_{i,j}^{(\infty)}(1)] = \max_{i,j}[d_{i,j}^{(\infty)}(m)] = d_{max}^{(\infty)}(m), \forall\, m$$

cannot change with m, since the largest maximum distance will already appear for $m = 1$. The mean distance

$$d_{mean}^{(\infty)}(m) = \frac{1}{N_{\text{eff}}^2(m)} \sum_{i,j=1}^{N_{\text{eff}}(m)} d_{i,j}^{(\infty)}(m),$$

however, necessarily increases with m or stays at most constant. More specifically, as m increases, smaller distances systematically disappear, so that the entire distribution is systematically shifted towards its (constant) maximum, thereby becoming narrower and exhibiting an increasing mean along with decreasing variance. We conjecture that, for large m, the distribution of $d^{(\infty)}(m)$ will converge to a limiting distribution (see below) possibly depending on the embedding delay τ.

4.2.2 Euclidean norm

In case of the L_2 (Euclidean) norm, both mean and maximum of all pairwise distances ($d_{mean}^{(2)}(m)$ and $d_{max}^{(2)}(m)$, respectively) monotonically increase with rising m (Appendix E.1, Fig. E.1B). This can be understood as follows: The L_2 distance between two points in an m-dimensional state space, $\vec{x}_i$ and $\vec{x}_j$, is given as

$$\|\vec{x}_i - \vec{x}_j\|_2 = \left(\sum_{k=1}^m |x_{i,k} - x_{j,k}|^2\right)^{\frac{1}{2}} = d_{i,j}^{(2)}(m) \tag{4.4}$$

For the squared L_2 distance, this implies:

$$\left[d_{i,j}^{(2)}(1)\right]^2 = (x_i - x_j)^2$$

$$\left[d_{i,j}^{(2)}(2)\right]^2 = (x_i - x_j)^2 + (x_{i-\tau} - x_{j-\tau})^2$$

$$= \left[d_{i,j}^{(2)}(1)\right]^2 + (x_{i-\tau} - x_{j-\tau})^2$$

$$\geqslant \left[d_{i,j}^{(2)}(1)\right]^2 \tag{4.5}$$

$$\vdots$$

$$\left[d_{i,j}^{(2)}(m+1)\right]^2 \geqslant \left[d_{i,j}^{(2)}(m)\right]^2 \geqslant \ldots \geqslant \left[d_{i,j}^{(2)}(1)\right]^2, \tag{4.6}$$

which explains the observed behavior of both mean and maximum distance using the L_2 norm. Specifically, unlike for L_∞, the maximum L_2 distance between two points is not bound by the largest pairwise distance in one dimension.

In a similar way, we may argue for all L_p distances ($p \in (0, \infty)$) defined as

$$\|\vec{x}_i - \vec{x}_j\|_p = \left(\sum_{k=1}^{m} |x_{i,k} - x_{j,k}|^p\right)^{\frac{1}{p}} = d_{i,j}^{(p)}(m) \tag{4.7}$$

that, by the same argument as above,

$$\left[d_{i,j}^{(p)}(m+1)\right]^p \geqslant \left[d_{i,j}^{(p)}(m)\right]^p, \tag{4.8}$$

implying again a monotonic increase of mean and maximum distances with rising embedding dimension (recall the positive semi-definiteness of distances and p).

4.2.3   Changing shape of distance distribution with increasing embedding dimension

Building upon our previous considerations and numerical results, a mathematically more specific yet challenging question is how exactly an increasing embedding dimension m is affecting the shape of the distribution of all pairwise distances rather than just its central tendency (mean).

For the maximum norm, one may argue that the individual components of each embedded state vector are commonly constructed such that they are as independent as possible [95]. Accordingly, for a system without serial correlations (i.e., uncorrelated noise), the absolute differences $d = d^{(\infty)}(1)$ between the components of two state vectors are also independent, identically distributed (i.i.d.) and lie within the interval $[0, d_{max}]$. In such case, for sufficiently large m, the pairwise L_∞ distance between two state vectors can be interpreted as the maximum of m i.i.d. variables that are bounded from above, which should follow a reversed Weibull distribution according to the Fisher-Tippett-Gnedenko theorem from extreme value statistics. Note, however, that this expectation is valid only if m is sufficiently large and the i.i.d. assumption is (approximately) fulfilled, both of which do not necessarily have to be the case for real-world time series. Moreover, it is not guaranteed that the given distance distribution in one dimension lies within the domain of attraction of the reversed Weibull class [177], which calls for further theoretical investigation in each specific case.

For other L_p norms including the Euclidean norm, the aforementioned considerations do not apply. Instead, for any L_p norm with $p < \infty$,

- the pairwise distances d^p are of the form $(\sum_i z_i^p)^{1/p}$ $(i = 1, ..., m)$ as given in Eq. (4.7) with approximately i.i.d. variables z_i.

- From the central limit theorem it follows that the distribution of d^p is approximately a normal distribution with mean and standard deviation growing proportionally with m and $\sqrt{m}$, respectively, for large m.

- The coefficient of variation of d^p thus declines approximately as $\sim 1/\sqrt{m}$.

- For large m, also $d = (d^p)^{1/p}$ is approximately normally distributed with mean and standard deviation growing approximately as $\sim m^{1/p}$ and $\sim \sqrt{m}\frac{dz^{1/p}}{dz}|_{z=m} \sim \sqrt{m}m^{1/p-1} = m^{1/p-1/2}$.

- The coefficient of variation of d thus behaves approximately as $\sim m^{1/p-1/2}/m^{1/p} = 1/\sqrt{m}$, just as for d^p.

- As a consequence, the relative variability of d narrows in the same fashion for all $p < \infty$ as m grows, and only the growth of the absolute scale of d with m depends on p ("curse of dimensionality" [339]).

The considerations made above do explain the numerical results in Fig. 4.1, showing histograms of the distances of three different time series for selected values of the embedding dimension m and for the L_2 and L_∞ norms. In addition to time series fulfilling the i.i.d. assumption (Fig. 4.1 A,B,C,D), here we are also interested in deterministic systems. As an illustrative example, we choose the Lorenz-63 system (Eq. (B.1), Fig. 4.1 E,F) in some non-stationary (drifting parameter) setting, which will be further studied in Section 4.3.

In this regard, it is confirmed that the expectation value of the distance distribution takes higher values with increasing m. The probability to find small distances therefore decreases. In case of the L_∞ norm (Fig. 4.1 B,D,F), this growth is bounded and we can identify a convergence of the distribution, in some cases eventually towards the aforementioned reversed Weibull distribution. In turn, for the L_2 norm (Fig. 4.1 A,C,E) the convergence towards a normal distribution is discernible. Considering the Lorenz-63 time series (Fig. 4.1 E,F), the empirical expectations are approximately met by the observations, even though the distribution of L_∞ distances exhibits a slightly more complex (i.e., less symmetric) shape than for the two noise series. Specifically, for the L_2 norm the resulting distance distribution is left-skewed with a pronounced lower tail (see Fig. 4.1 E), whereas for the L_∞ norm we observe a disturbed Weibull-like shape. Notably, the i.i.d. assumption is violated when dealing with such a deterministic dynamical system. For a more detailed characterization of the shape of the empirically observed pairwise distance distributions shown in Fig. 4.1, see Appendix E.2.

In general, we emphasize that it is not straightforward to analytically describe the shape of the distance distribution of an embedded time series stemming from an arbitrary dynamical system with potentially nontrivial serial correlations. Regarding our overarching question how we could automatically choose a recurrence threshold such that the resulting recurrence characteristics are as independent as possible of the embedding dimension and chosen norm, we need to consider both,

(i) the general increase of distances together with their successive concentration and

(ii) the varying shape of the distribution of distances

with increasing embedding dimension. The first aspect could be accounted for by relating the threshold selection to the spatial extent of the state space object (attractor), similar as, for instance, suggested by Abarbanel [1] in the context of the false nearest neighbor algorithm. However, our findings suggest that accounting for the second point is key to an appropriate recurrence threshold selection method that relieves the effects of the embedding dimension on the recurrence properties as much as possible. As a simple possible solution, we recommend to use a numerical estimate of a certain (sufficiently low) percentile of the distance distribution as threshold [73, 74, 204, 331]. This approach considers both above mentioned effects and leads to a constant global recurrence rate (which equals the chosen percentile). As a result, the recurrence properties become much less dependent on the embedding dimension and chosen norm than when using other methods, as we will exemplify in the following section.

We emphasize that in addition, by conserving the recurrence rate, possible dependencies of RQA characteristics on the density of recurrences for different m are omitted, and corresponding residual changes of these measures upon varying m could rather point to either insufficiently low embedding dimension (missing essential factors contributing to the system's dynamics, in a similar spirit as, e.g., for the false nearest neighbor method) or spurious recurrence structures arising from over-embedding [294]. These ideas should be further studied in future work.

4.3 Numerical example

In this section, we will demonstrate the effect of the varying shape of the distance distribution with increasing embedding dimension on different threshold selection approaches working with a globally fixed value of ε. In order to mimic a practically relevant test case of a non-stationary low-dimensional dynamical system, where we should use some higher embedding dimension (following Hegger et al. [124]) instead of a more moderate choice, we consider the classical Lorenz-63 system, Eq. (B.1). Depending on the parameters σ, β and r, the system exhibits either regular or chaotic dynamics. Here, we consider a transitory setting, where the parameter r gradually increases from 180 to 210 while keeping $\beta = 8/3$ and $\sigma = 10$ fixed. In this case, the system undergoes a transition from a chaotic regime into a regular (limit cycle) phase as r rises before it exhibits again a chaotic behavior. Note again that instead of studying the stationary Lorenz-63 system for different values of r, we intentionally employ a gradual parameter change leading to a non-stationary system which calls for a systematic over-embedding when performing nonlinear time series analysis [124]. Specifically, we implement a linear variation of r as

$$r(t_{\mathrm{is}}) = 180 + 2.5 \cdot 10^{-2} t_{\mathrm{is}}. \tag{4.9}$$

For numerically solving this system of equations, we use a fourth-order Runge-Kutta integrator with an integration step of $t_{\mathrm{is}} = 0.001$ and a total of 1,300,000 iterations. Therefore, we simulate the system's evolution over 1,300 time units (t.u.). By using a sampling interval of $\delta t = 0.2$ t.u. we obtain 6,500 samples forming our time series for the three components x, y and z. We remove the first 500 points ($\widehat{=}$100 t.u.) that could be affected by transient dynamics and retain the remaining 6,000 points ($\widehat{=}$1200 t.u.) of the y component for further analysis.

We integrate the Lorenz-63 equations, Eq. (B.1), with the linear parameter change, Eq. (4.9), 1,000 times with randomly chosen initial conditions, embed the y component time series using a delay $\tau = 4$, consistent with the first local minimum of the mutual information [95], and assess the

resulting RPs. For each of these 1,000 RPs, we use a running window along their main diagonal with a window size of $w = 400$ and mutual shift of $ws = 40$ data points, i.e., 90% overlap between consecutive windows, to quantitatively study the time-dependence of the resulting recurrence characteristics. We repeat this procedure for embedding dimensions ranging from $m = 3$ to $m = 10$ and for four different threshold selection methods: (i) a fixed percentile of the distance distribution (as recommended by our theoretical considerations in Section 4.2) as well as some fixed percentage of the (ii) maximum, (iii) mean and (iv) median pairwise distances between all state vectors in the reconstructed state space, respectively.

Since we are aiming to study the change of recurrence properties associated with a transition between chaotic and periodic dynamics and vice versa, we choose the *recurrence time entropy (RTE)*, Eq. (A.11) in Appendix A. We choose the actual recurrence threshold for each threshold selection method (i)-(iv) such that a global recurrence rate of $RR \approx 4\%$ is achieved in all four cases for $m = 3$. Therefore, for each embedding dimension we obtain a distribution of 1,000 RTE time series and show the mean (blue lines in Fig. 4.2) together with the two-sided 90% confidence interval ($[5\%, 95\%]$, gray shaded areas). In order to put these time dependent RTE estimates of the non-stationary Lorenz-63 system into a context, we consider a reference reflecting the time-dependent RTE values directly computed from the true three-dimensional state vectors without embedding, using otherwise the same analysis strategy (window size and overlap) as for the embedding scenario. Thus, for each point in time we obtain 1,000 reference measurements and consider the mean (red line) and the two-sided 90% confidence interval (red shaded area in Fig. 4.2).

The robustness of the observed time-dependence of RTE with respect to the chosen embedding dimension when using a fixed percentile of the distance distribution (i.e., a fixed recurrence rate) is shown in Fig. 4.2A (here we used the L_2 norm, but the results are similar when using the L_∞ norm). For any embedding dimension larger than $m = 4$, the variations of the RTE estimates originating from the embedding procedures match the red reference time series within its uncertainties for times $t \gtrsim 200$. For adequately revealing the chaotic regime in the first part until $t \approx 160$, an embedding dimension larger than $m = 7$ seems to be inevitable, whereas results from any embedding dimension coincide with the reference estimate within its uncertainties at the limit cycle regime ($1,000 \leq t \leq 1,080$). In case of not using the recommended threshold selection method, this robustness is clearly lost (Fig. 4.2B,C,D), and only the limit cycle regime (plus some shorter sections before) are properly revealed by the estimates obtained in the reconstructed state space.

Considering the results of Section 4.2, the reason for the failure of the methods based on individual location parameters (maximum, mean, median) of the pairwise distance distribution between state vectors for higher embedding dimensions is the change in the shape of that distribution beyond its characteristic location and range parameters. Appendix E.3 demonstrates this effect on the RPs in some more detail. Hence, we argue that selecting the recurrence threshold at some percentile of the distance distribution is to be preferred if we aim to obtain stable results for a broad range of embedding dimensions, which is the case if we wish to automatically choose fixed recurrence thresholds for the analysis of arbitrary complex systems.

We note that the presented example has focused on a recurrence characteristic that is particularly well suited for detecting transitions between chaotic and periodic dynamics and is linked to a dynamical invariant. Other recurrence characteristics, like classical RQA measures or recurrence network characteristics, have been found to exhibit less stable variations with changing embedding dimension (not shown) and are therefore not further discussed here. Clarifying the reasons for the
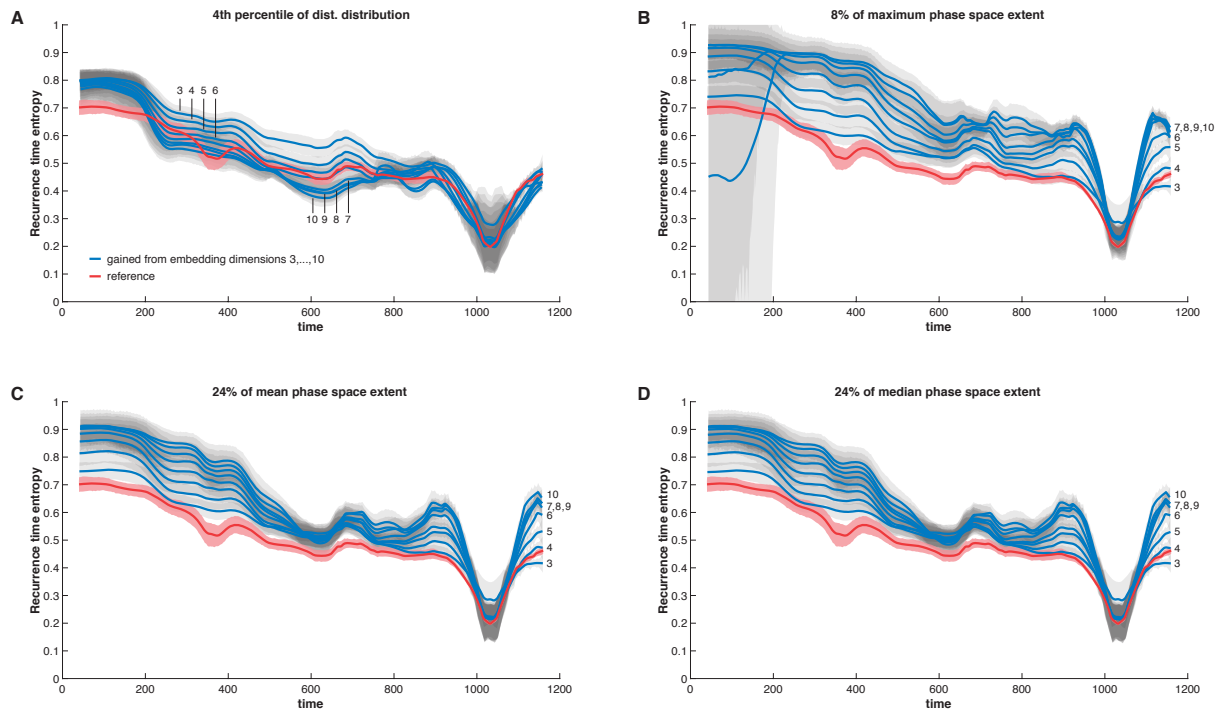
Figure 4.2: Time-dependence of *RTE* (ensemble means and two-sided 90% confidence intervals from 1,000 independent realizations) based on the y component of the non-stationary Lorenz-63 system (see text for details) using the L_2 norm. The blue lines show the results for time-delay embedding with different embedding dimensions ($m = 3, \ldots, 10$) and for four different methods to select the recurrence threshold according to **A** a certain percentile of the distance distribution and some percentage of the **B** maximum, **C** mean and **D** median distance between state vectors on the reconstructed attractor. The actual threshold values (4th percentile, 8%, 24% and 24%, respectively) have been chosen such that the global recurrence rate of approximately 4% is achieved for each method in the embedding scenario with $m = 3$. The red line shows the reference time series gained from 1,000 independent realizations of the non-stationary Lorenz-63 system by randomly choosing initial conditions and using all three components as state variables. Shaded areas (gray and red) indicate the two-sided 90% confidence intervals estimated from the respective ensembles.

different behaviors of different recurrence characteristics will be an important subject of future work.

4.4 Conclusions

We have discussed the changing shape of the distribution of pairwise distances between state vectors obtained by time delay embedding with increasing embedding dimension and its implications for different methods for selecting temporally fixed recurrence thresholds. While specific values of the recurrence threshold should always be chosen based upon a multitude of criteria ranging from time series length to different topological and/or geometric characteristics of the associated trajectory, we have provided both empirical arguments and numerical indications that selecting the recurrence threshold at a prescribed percentile of the distance distribution (i.e., conserving the global recurrence rate) results in quantitative recurrence characteristics that are more stable

under changes of the embedding dimension than when using alternative approaches. In the latter context, we have demonstrated that measures from RQA and related frameworks may exhibit a crucial dependence on the embedding dimension when selecting the recurrence threshold according to a certain percentage of the mean or maximum state space diameter, as sometimes suggested in other works [272, 336]. This also indicates that some alternative approaches, such as normalizing the time series and applying a uniform threshold independent of the embedding dimension and the considered norm [139], are not likely to perform well for any kind of data, when neglecting the effect on the distance distribution with increasing embedding dimension.

At the conceptual level, our general discussion of the changing shape of distance distributions with embedding dimension has led to some interesting follow-up questions associated with the convergence properties of these distributions at high embedding dimensions, which should be further addressed in future studies. Notably, the relationship between the distribution of L_∞ distances and extreme value statistics clearly deserves further investigations to fully understand the emerging shape of the distributions as the embedding dimension becomes large. As a cautionary note, we emphasize that the considerations presented in this work relate exclusively to the concept of time delay embedding as the most widely applied embedding technique, but not necessarily to methodological alternatives like derivative embedding [180], for which the metric properties of different components of the embedding vector cannot be easily related to each other.

Taken together, the results presented in this work are important for automatizing the problem of data-adaptive recurrence threshold selection, which is key for further widening the scope of applications of recurrence plots, recurrence quantification analysis and related techniques across scientific disciplines. Especially in the context of long time series originating from non-stationary systems, which frequently appear in many fields of science, a generally applicable approach is crucial for obtaining reliable and easily interpretable results.

Chapter 5

Border effect corrections for diagonal line based recurrence quantification analysis measures

Abstract

Recurrence Quantification Analysis (RQA) defines a number of quantifiers, which base upon diagonal line structures in the recurrence plot (RP). Due to the finite size of an RP, these lines can be cut by the borders of the RP and, thus, bias the length distribution of diagonal lines and, consequently, the line based RQA measures. In this chapter we investigate the impact of the mentioned border effects and of the thickening of diagonal lines in an RP (caused by tangential motion) on the estimation of the diagonal line length distribution, quantified by its entropy. Although a relation to the Lyapunov spectrum is theoretically expected, the mentioned entropy yields contradictory results in many studies. Here we summarize correction schemes for both, the border effects and the tangential motion and systematically compare them to methods from the literature. We show that these corrections lead to the expected behavior of the diagonal line length entropy, in particular meaning zero values in case of a regular motion and positive values for chaotic motion. Moreover, we test these methods under noisy conditions, in order to supply practical tools for applied statistical research.

5.1 Introduction

Recurrence quantification analysis (RQA) is a powerful tool for the identification of characteristic dynamics and regime changes [209, 321]. This approach is successfully applied in many scientific disciplines [84, 121, 122, 129, 161, 213, 223, 227, 266, 287, 322]. Several measures of complexity are defined on geometric features (such as diagonal and vertical lines) in the recurrence plot (RP), which represents time points j when a state $\vec{x}_i$ at time i recurs [209, 213, 321, 336]. These line structures represent typical dynamical behavior and are related to certain properties of the dynamical system, e.g., chaotic or periodic dynamics. Therefore, their quantitative study by

the RQA measures within sliding windows is a frequently used task for the detection of regime changes [71, 82, 311, 321]. However, as some RQA measures rely on the probability distribution of the lengths of the diagonal lines in an RP, the artificial alteration of these lines due to border effects [48, 80], insufficient embedding [203, 209], or a certain sampling setting [83, 292] can have significant impact on these measures. A few ideas have been suggested to overcome such problems [48, 277, 326]. Here we review these ideas, propose novel correction schemes, and systematically compare them.

5.2 Recurrence quantification analysis and border effects

A recurrence plot $\mathbf{R}$ (Eq. (1.2)) consists of small-scale structures, such as single points and diagonal and vertical lines, which characterize important dynamical properties of the system. A diagonal line is a sequence of pairs of time points $\mathcal{L} := \{(i,j),(i+1,j+1),\ldots,(i+\ell-1,j+\ell-1)\}$ where $R_{i,j} \equiv 1$ for all index pairs in $\mathcal{L}$. Diagonal lines in the RP represent the temporal duration that two distinct parts of the phase space trajectory run parallel (Figs. 5.1 and 5.2). The histogram $P(\ell)$ of the lengths of diagonal lines (Fig. 5.3) characterizes the dynamics [87, 199, 297] and can be and has been used to quantitatively distinguish between RPs, the underlying dynamics, or to identify regime transitions [161, 213, 223, 227, 311].

For uncorrelated noise, the probability to find a line $\mathcal{L}$ of exact length ℓ decays exponentially [293] (Fig. 5.3A), i.e., the RP consists only of very short diagonal lines, if there are any lines at all (Fig. 5.2A). In contrast, for chaotic dynamics, the RP contains diagonal lines of different lengths (Fig. 5.2C), resulting in a broad distribution $P(\ell)$ (Fig. 5.3C). The RP for a periodic system contains continuous, non-interrupted diagonal lines, virtually of infinite length (Fig. 5.2B). In principal, we would expect a discrete line length distribution with a peak at line length infinity. However, the lines are cut at the begin and end of the RP, such that an uncorrected conventional line length measurement results in a discrete distribution $P(\ell)$ with uniform characteristics (Fig. 5.3B).

The RP is a discrete matrix. Therefore, the creation of the histogram $P(\ell)$ appears to be trivial. But it is not as simple as it looks at the first glance. Diagonal lines can be quite long and – as already mentioned – can exceed the finite size of the RP. In practice, this is a very common problem, particularly when a sliding window method is applied. How to count such diagonal lines? As we will see later, for some measures, it can be important to have the correct length of the lines, for other measures it does not play any role. In the original definition, the lines are also counted even if they were cut by the RP border [209, 323, 336].

Several quantification measures for RP analysis have been introduced which use $P(\ell)$ (see also Appendix A). The firstly introduced measure was the determinism DET (Eq. (A.6)) [323]. This measure is defined as the fraction of recurrence points that form diagonal lines and considers lines which have at least length $\ell_{\min}$, which in principle is a free parameter, but often set to 2. Nevertheless the choice of the minimal line length can be crucial for the correct estimation of some RQA measures and we come back to that in Section 5.6.1. More details about this can be found in Marwan et al. [209]. Since RPs of uncorrelated noise have mainly single points and only few and short diagonal lines, for such dynamics DET has rather low values (although embedding can result in artificially high DET values, see discussion in [203, 325]). In contrast, RPs for deterministic dynamics contain many diagonal lines, resulting in elevated values of DET, with the special case of $DET = 1$ for periodic and quasi-periodic dynamics. As this measure only quantifies
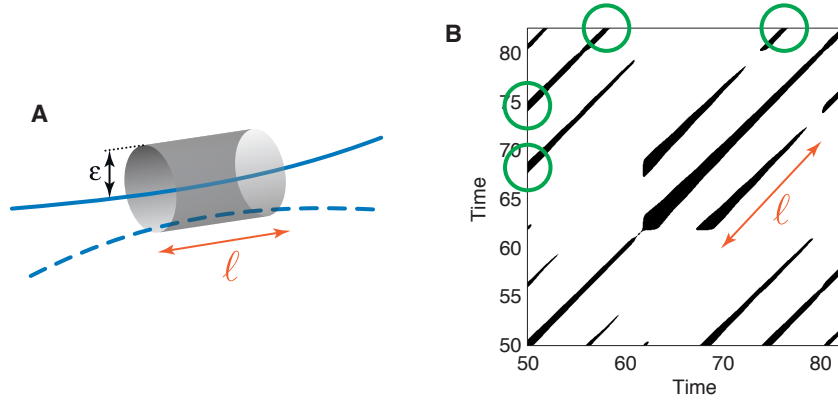
Figure 5.1: Parallel and close parts of a phase space trajectory **A** correspond to diagonal lines of length ℓ in an RP **B**. Diagonal lines can be cut by the border of the RP (green circles).

whether a recurrence point is on a diagonal line or not, the actual length of a diagonal line is not important (i.e., whether the line crosses the RP border or not).

Another idea is to look at the average and maximal length of the detected diagonal lines (related to prediction time and Lyapunov exponent, resp.[209]). The average, of course, depends on the actual line lengths and will be biased when diagonal lines cross the RP borders.

Because the shape of $P(\ell)$ differs for different dynamics, the Shannon entropy of the probability distribution $p(\ell) = P(\ell)/\sum_\ell P(\ell)$ to find a diagonal line of exact length ℓ was suggested (Eq. (A.9)) [323]. In the following we denote this entropy as *ENTR*. This measure was introduced in a pragmatic way to quantify the visual line structures in the RP and has been interpreted as the "information content of the trajectories" [335]. Here, the choice of the minimal line length $\ell_{\min}$ has a significant effect, since it discards parts of the line length histogram and therefore alters its shape. For uncorrelated noise, *ENTR* has low values, because $p(\ell)$ is exponentially decaying. For chaotic dynamics, $p(\ell)$ is a broad distribution, resulting in relatively large *ENTR* values. However, for periodic signals $p(\ell)$ has more similarity with a uniform distribution if the mentioned border effects are not accounted for. Therefore, *ENTR* is not low for periodic signals, although we would expect it, but rather large, even larger than for chaotic dynamics. Here, the effect of the sliced lines at the RP border has the strongest and remarkable effect, which is why we focus on this measure only in this chapter.
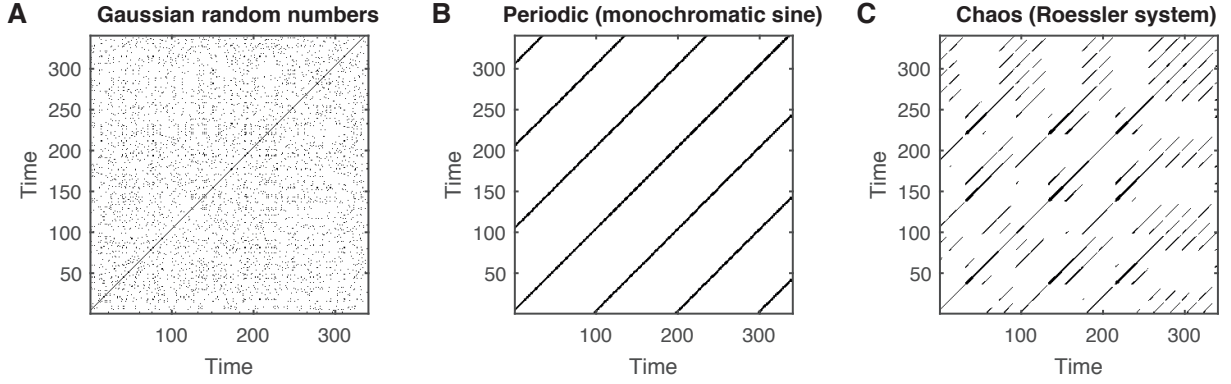
Figure 5.2: RPs of **A** standard normal Gaussian random numbers, **B** time-delay embedded sinusoidal with an oscillation period $T = 100$ time units ($m = 2$, $\tau = T/4$), and **C** the Rössler system ($a = 0.15$, $b = 0.2$, $c = 10$) (only subsets shown). RPs were constructed from time series of 2,000 samples (in case of the Rössler system we removed transients) using a constant global recurrence rate of 4% with a fixed threshold and Euclidean norm.
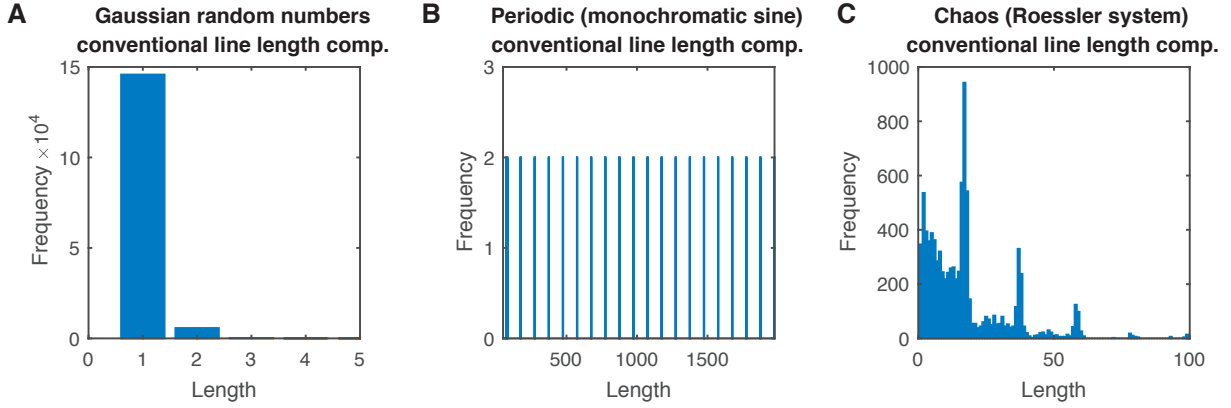
Figure 5.3: Diagonal line length distributions of the different systems types described in Fig. 5.2, gained from the conventional line counting.

5.3 Correction schemes for counting diagonal lines

In this section we show two ways of overcoming the problem of biased diagonal line based measures due to the border effect. Either we manipulate the histogram of the diagonal lines (Section 5.3.1) or we change the shape of the RP in order to avoid a bias in the first place (Section 5.3.2).

5.3.1 Alternative ways of counting line lengths

Let $\mathbf{R}$ be a $N \times N$ recurrence matrix, Eq. (1.2), and $P(\ell)$ the histogram of the diagonal lines contained in $\mathbf{R}$. We now substantiate the definition of a diagonal line in an RP from Section 5.2. A diagonal line $\mathcal{L}$ of length ℓ is a set of ℓ index tupels $(\cdot, \cdot)_{k=1,...,\ell}$:

$$\mathcal{L}_\ell := \{(i + k, j + k) \, | \forall k = 0, ..., \ell - 1 \; : \; (1 - R_{i-1,j-1})(1 - R_{i+\ell,j+\ell}) R_{i+k,j+k} \equiv 1\}. \quad (5.1)$$

The length ℓ of a line is usually the cardinality of this set $|\mathcal{L}|$.

We denote any diagonal line which starts and ends at the border of $\mathbf{R}$ as a *border diagonal*, e.g., in case of the lower triangle of the RP, when starting at $(i, 1)$ in the first column and ending at $(N, N - i + 1)$ in the last row:

$$
\begin{aligned}
\mathcal{L}_{\text{border}} :=& \{(i + k, 1 + k) \mid \forall k = 0, \ldots, N - i \ : \ R_{i+k,1+k} \equiv 1 \quad \vee \\
& (1 + k, j + k) \mid \forall k = 0, \ldots, N - j \ : \ R_{1+k,j+k} \equiv 1\}.
\end{aligned} \tag{5.2}
$$

Any diagonal of length ℓ, which starts or ends at the border of $\mathbf{R}$ and has an end or start point within the recurrence matrix, we call *semi border diagonal*:

$$
\begin{aligned}
\mathcal{L}_{\text{semi border}} :=& \{(i + k, j + k) \mid \forall k = 0, \ldots, \ell - 1 \wedge (j = 1 \vee i + \ell - 1 = N) \ : \ R_{i+k,j+k} \equiv 1 \quad \vee \\
& (i + k, j + k) \mid \forall k = 0, \ldots, \ell - 1 \wedge (i = 1 \vee j + \ell - 1 = N) \ : \ R_{i+k,j+k} \equiv 1\}.
\end{aligned} \tag{5.3}
$$

Discard border diagonals from histogram (dibo correction)

The real length of the border diagonals is unknown. Therefore, we are not able to assign their true length to them and, hence, one option to deal with the missing length regarding the line length histogram is setting their length to zero. That is, we simply discard all (semi-)border diagonals from $P(\ell)$ and, thus, avoid the broad line length distribution as exemplary shown in Fig. 5.3B. As desired, this results in a lowered entropy value, but also has some drawbacks. In case of a perfectly sampled stationary periodic signal (without any noise contamination) this method would empty the histogram $P(\ell)$ completely, leaving an undefined entropy (Figs. 5.8, F.1, F.2) and a mean and maximum line length of zero. In the following, we refer to this approach as *dibo correction* (DIscard BOrder diagonals).

Assign maximum line length to all border diagonals (Censi correction)

To avoid an empty diagonal line histogram Censi et al. [48] suggested to assign all border diagonals the length of the main diagonal of the RP (line of identity). Sticking to the aforementioned example of a perfectly sampled and uncontaminated stationary periodic signal, this modification results in a delta peak in $P(\ell)$ (Fig. 5.4C), and therefore a sound defined entropy value of zero as well as meaningful mean and maximal line length estimate (Figs. 5.8, F.1, F.2). For deterministic chaotic processes this correction scheme could underestimate the entropy, if the RP is smaller than the average length scale of the diagonal lines. Especially in a running window approach, this effect is assumed to be significant. We refer to this approach as *Censi correction*.

Keeping just the longest border diagonal (kelo correction)

In alternative to the correction in Section 5.3.1, all (semi-)border diagonals from $P(\ell)$ are discarded, but the longest one. This approach would also avoid the broad line length distribution shown in Fig. 5.3B, but would leave a valid definition of the entropy, since $P(\ell)$ is not an empty set (cf. Figs. 5.8, F.1, F.2). The resulting entropy for the aforementioned example would be low. In contrast to the Censi correction, this approach would avoid the bias for deterministic chaotic processes when a windowing approach is applied. In the following, we refer to this approach as *kelo correction* (KEeping the LOngest border diagonal).
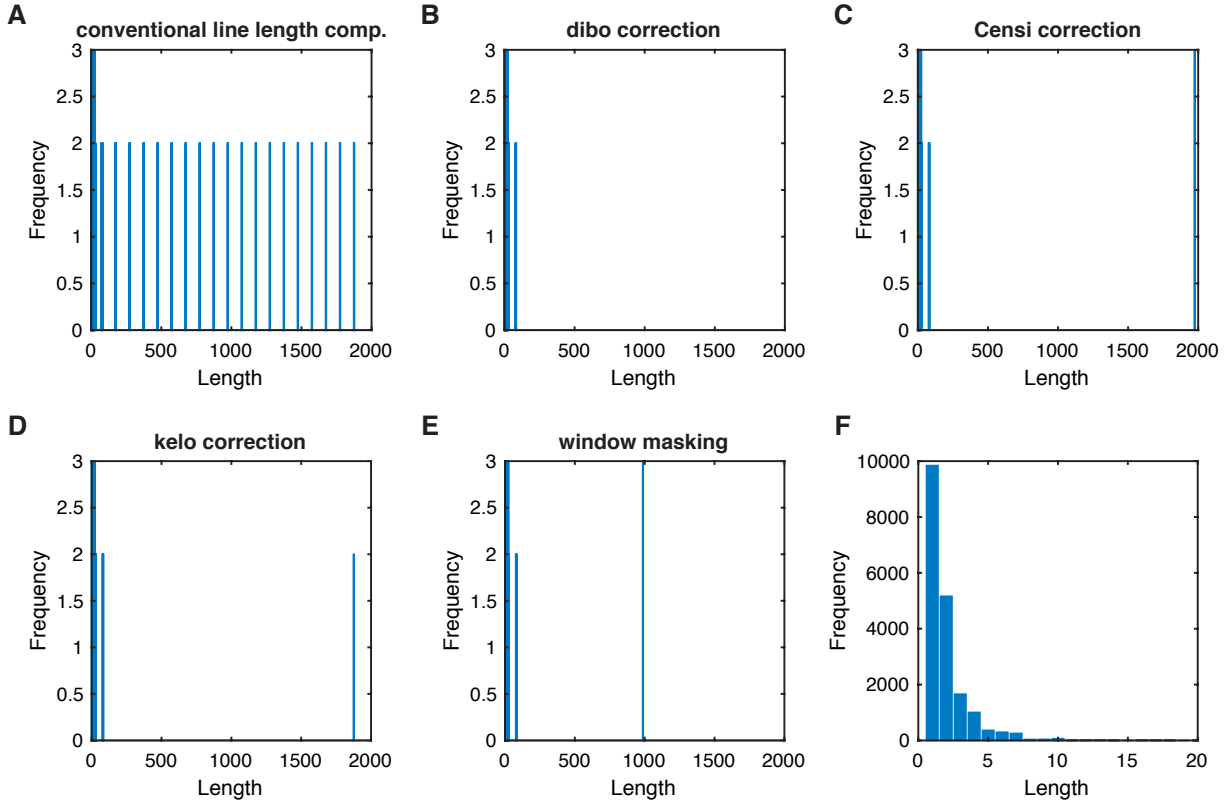
Figure 5.4: Diagonal line length histograms of the conventional line length computation **A** and of the presented correction schemes **B**-**E** for a monochromatic time-delay embedded sinusoidal with an oscillation period $T = 100$ time units ($m = 2$, $\tau = T/4$, same as in Figs. 5.2B and 5.3B). An enlargement of the histograms from panels **A** to **D**, focusing on the shorter line lengths, is presented in panel **F**. A corresponding enlargement of panel **E** does qualitatively look the same, but with reduced frequencies, due to the smaller effective window size (see text for details). For a better visibility we enlarged single bars in panels **B** to **E** and limited the view to a frequency range $[0\ 3]$ in panels **A** to **E** (in **F** the full range is used).

5.3.2 Alternative RP window shapes (window masking)

The origin of the border diagonals is related to geometric difference between the RP and the diagonals. Therefore, a further approach to avoid the length bias of border diagonals is to apply a specific window to the RP which has the same geometric orientation as the diagonals. One realization of such a window is a $45°$ rotated cutout from the original RP with an edge length s (Fig. 5.5). Conventionally counting the lines of this cutout preserves a delta peak distribution in $P(\ell)$ for a periodic signal (Fig. 5.4E). However, with this shape we loose $w^2 - 2s^2 = w^2 - \frac{1}{2}w^2 = \frac{1}{2}w^2$ data points with respect to the original RP. Note that s and w in Fig. 5.5 imply a number of data points, meaning hypotenuse and catheti of an isosceles triangle have the same length ($w = \frac{s}{2}$). We argue that this approach could be rather useful in a running window approach over a global RP, where the size of the alternative shape could be chosen such that it contains as many data points as the *classic*, non-rotated, window. We refer to this approach as *window masking*. Another alternative would be a parallelogram with the top and bottom sides having the $45°$ direction [70].
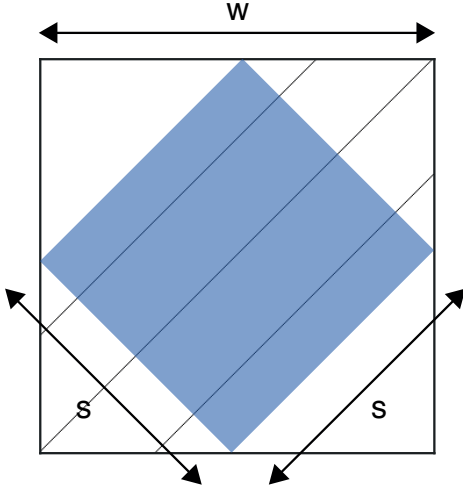
Figure 5.5: Blue shaded alternative window shape with edge length s of a $w \times w$ recurrence plot. s and w imply the number of RP matrix elements covered by the window shapes.

5.4 Tangential motion in recurrence quantification analysis

Even though the considerations made in the preceding sections are valid and useful, the correction schemes presented in Section 5.3 most likely do not give the expected correction for the entropy of diagonal line lengths for experimental data, unless the data has been properly preprocessed. There are three reasons why the correction of the border diagonals in the diagonal line histogram $P(\ell)$ is not sufficient: (*i*) temporal correlations in the data, especially when highly sampled flow data is used, (*ii*) noise, and (*iii*) insufficient embedding of the time series at hand (if needed) combined with the effect of discretization and an inadequate choice of parameters needed to construct the RP (recurrence threshold method, recurrence threshold size, norm).

Temporal correlation means that states $\vec{x}_j$ preceding or succeeding a state $\vec{x}_i$ (or a recurring state $\vec{x}_k$ of $\vec{x}_i$), are very similar to this one and, hence, falling into the neighborhood of $\vec{x}_i$ (or $\vec{x}_k$) and to be considered as recurrences, i.e., $\mathbf{R}_{i,j} := 1$ for $j = [i - m, \ldots, i + n]$ or for $j = [k - m, \ldots, k + n]$ when $\mathbf{R}_{i,k} := 1$ (Fig. 5.6A). This results in vertically extended sequences in the RP, i.e., thickening its diagonal lines. The thickening leads to an artificially enlarged number of diagonal lines, thus effecting the distribution $P(\ell)$, and is often referred to as *tangential motion* [101, 209]. Moreover, the thickening is not evenly distributed along a diagonal line (Fig. 5.6B). For border diagonals, this means that there are not only additional border diagonals (which could be handled by applying correction schemes as described in Section 5.3), but additional shorter diagonal lines, again leading to a broadening of the line length distribution $P(\ell)$ and an elevated entropy *ENTR*.

Additive noise causes the already thickened lines in the RP to appear more diffus (Fig. 5.6C,E). Technically speaking, the noise alters the phase space trajectory, causing the pairwise distances to randomly scatter about their true/noise-free values and, thus, the histogram $P(\ell)$ gets enriched with small line lengths (cf. Fig. 5.4F, [295]). This eventually biases the RQA measures discussed in Section 5.2.
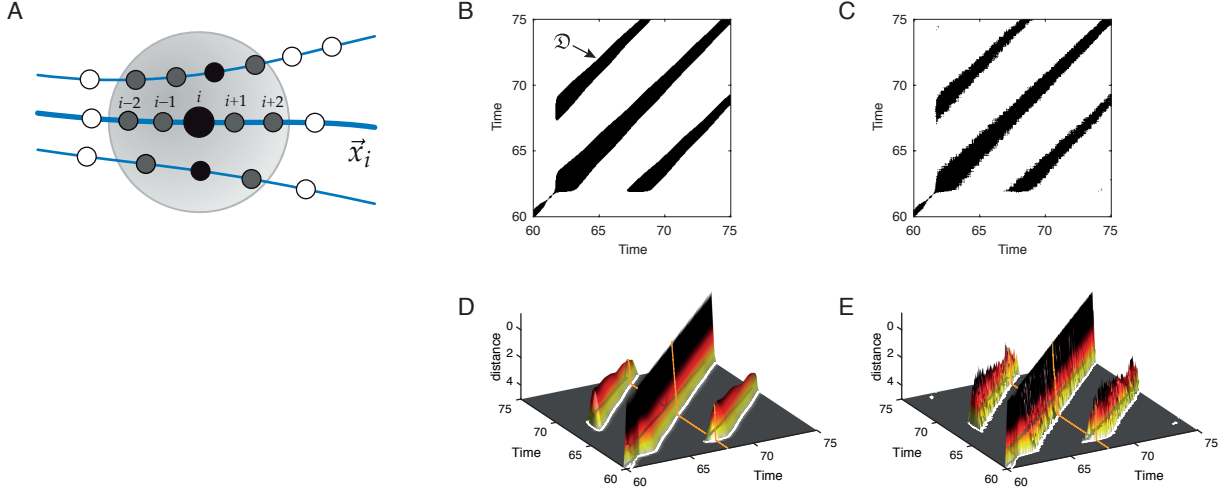
Figure 5.6: A Tangential motion, i.e., points of a trajectory preceding and succeeding a (recurring) state (gray), cause thickening of diagonal lines in the RP **B**, **C**. The thickening of diagonal lines can vary, e.g., as in this example of the Rössler system (noise free case in **B** and additive noise in **C**). The diagonal lines are more thick at the beginning and become less thick with time. A diagonal line in an RP (**B**, **C**) denotes a range of distances in the distance matrix falling under a the recurrence threshold ε. Panels **D** and **E** show three "distance ranges" (we call such a range $\mathfrak{D}$ in the text) corresponding to the three lines in **B**, **C** respectively. Shown is a color-coded, thresholded distance matrix with reversed z-axis for a better visibility (increasing distances from top to bottom). The colormap encodes zero distance as black and the distance corresponding to the recurrence threshold as grey.

5.5 Correction schemes for reducing the effects of tangential motion

5.5.1 Perpendicular RP

A straightforward way to reduce the thickening of the diagonal lines from a theoretical perspective is the *perpendicular RP*, suggested by Choi et al. [54]

$$\mathbf{R}_{i,j}^{\perp}(\varepsilon) = \Theta\left(\varepsilon - \|\vec{x}_i - \vec{x}_j\|\right) \cdot \delta\left(\dot{\vec{x}}_i \cdot (\vec{x}_i - \vec{x}_j)\right), \qquad \vec{x} \in \mathbb{R}^d. \tag{5.4}$$

This RP contains only those points $\vec{x}_j$ that fall into the neighborhood of $\vec{x}_i$ and lie in the $(d-1)$-dimensional subspace of $\mathbb{R}^d$ that is perpendicular to the phase space trajectory at $\vec{x}_i$. Although theoretically there is no need for an additional parameter in order to construct a perpendicular RP, in practical situations almost no points in $\mathbb{R}^d$ phase space end up on the mentioned $(d-1)$-dimensional subspace of $\vec{x}_i$ (Poincaré section), due to limited resolution (discretization) of the data. Hence, it is reasonable to introduce an additional threshold parameter φ, which allows points $\vec{x}_j$ to be considered as perpendicular to $\vec{x}_i$, if

$$\arccos \frac{\dot{\vec{x}}_i \cdot (\vec{x}_i - \vec{x}_j)}{|\dot{\vec{x}}_i| \cdot |(\vec{x}_i - \vec{x}_j)|} \in \left[\left(\frac{\pi}{2} - \varphi\right), \left(\frac{\pi}{2} + \varphi\right)\right]. \tag{5.5}$$

Thus, Eq. (5.4) transforms to

$$\mathbf{R}_{i,j}^{\perp}(\varepsilon, \varphi) = \Theta\left(\varepsilon - \|\vec{x}_i - \vec{x}_j\|\right) \cdot \Theta\left(\varphi - \left|\arccos\frac{\dot{\vec{x}}_i \cdot (\vec{x}_i - \vec{x}_j)}{|\dot{\vec{x}}_i| \cdot |(\vec{x}_i - \vec{x}_j)|}\right| - \frac{\pi}{2}\right), \qquad \vec{x} \in \mathbb{R}^d. \qquad (5.6)$$

Figure 5.7B shows a perpendicular RP for a Rössler system (with parameters $a = 0.15$, $b = 0.2$, $c = 10$, transients removed). For the estimation of the tangential at each point in phase space we used the reference point, its predecessor and its successor. We set the angle threshold to $\varphi = \frac{\pi}{12}$ ($= 15°$).

5.5.2 Isodirectional RP

Requiring less computational effort, the *iso-directional RP* suggested by Horai et al. [133] also promises to cope with the tangential motion, but also inherits two additional parameters T and ε_2 (Fig. 5.7C). In this approach two points in phase space are denoted recurrent, if their mutual distance falls within the recurrence threshold ε and the distance of their trajectories throughout T consecutive time steps falls within a recurrence threshold ε_2

$$\mathbf{R}_{i,j}^{\rightarrow}(\varepsilon, \varepsilon_2, T) = \Theta(\varepsilon - \|\vec{x}_i - \vec{x}_j\|) \cdot \Theta(\varepsilon_2 - \|(\vec{x}_{i+T} - \vec{x}_i) - (\vec{x}_{j+T} - \vec{x}_j)\|), \qquad \vec{x} \in \mathbb{R}^d. \quad (5.7)$$

We achieved reasonable results when choosing T in the size of the decorrelation time (e.g., first minimum of the mutual information) and the second recurrence threshold as half of the size of the recurrence threshold ε, which determines the parent RP.

5.5.3 True recurrence point RP (TRP)

Inspired by the work of Gao [100], Ahlstrom et al. [6] compute a normal RP, Eq. (1.2), but only accept those points to be recurrent, which "first" enter the ε-neighborhood shown in Fig. 5.6A. To ensure this, they first identify all points which fall into an ε-neighborhood of a certain point $\vec{x}_i$

$$\zeta_i \equiv \{\vec{x}_{j_1}, \vec{x}_{j_2}, \dots | R_{i,j_k} := 1\}, \qquad (5.8)$$

i.e., all points j_k in column i of the RP. The time difference of two consecutive recurrence points $\vec{x}_{j_k}, \vec{x}_{j_{k+1}}$ is $\{T_k^{(1)} = j_{k+1} - j_k\}_{k \in \mathbb{N}}$ in units of the sampling time (recurrence times of first type, [100]) and these correspond to the vertical distances between these points in column i of the RP. They now discard all points from the RP, whose vertical distance to its neighboring point in a column is 1 and leaving all points with recurrence time larger than 1,

$$\zeta_i^* \equiv \{\vec{x}_{j_1}, \vec{x}_{j_2}, \dots | R_{i,j_k} := 1, \ T_k^{(1)} > T_{\min}\}, \qquad T_{\min} = 1. \qquad (5.9)$$

The authors call this modified RP a *true recurrence point recurrence plot (TRP)*. This is different than simply discarding all points from the computations of Eq. (1.2) which fall within a certain time range w_{Theiler} of the reference point (*Theiler window*, [292])

$$\mathbf{R}_{i,j}(\varepsilon) = \Theta\left(\varepsilon - \|\vec{x}_i - \vec{x}_j\|\right), \qquad |i - j| > w_{\text{Theiler}}, \ \vec{x} \in \mathbb{R}^d. \qquad (5.10)$$

To obtain a TRP, we suggest to discard all recurrence points with recurrence times greater than w_{Theiler}, i.e., $T_{\min} = w_{\text{Theiler}}$ in Eq. (5.9). The Theiler window should be set in the order of
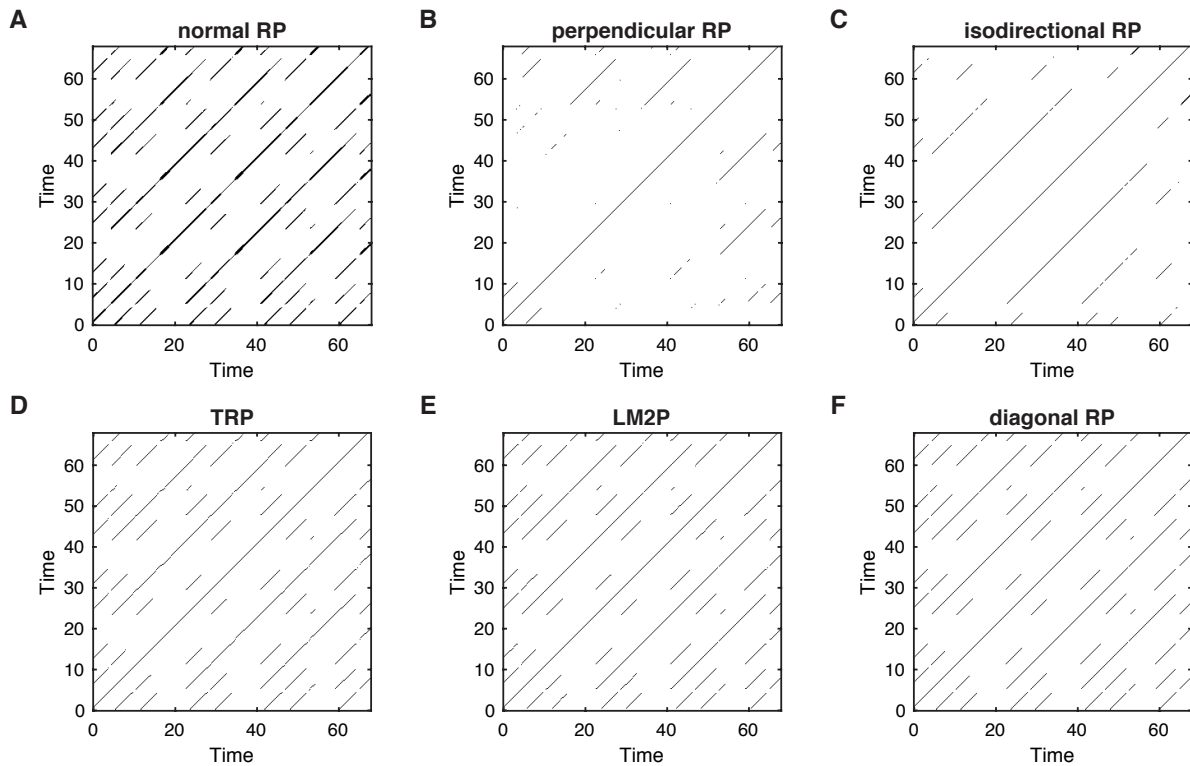
Figure 5.7: Different approaches for avoiding the effect of tangential motion in a recurrence plot (RP), exemplary shown for the Rössler system (with parameters $a = 0.15$, $b = 0.2$, $c = 10$, sampling time $\Delta t = 0.2$). **A** Normal RP with fixed recurrence threshold ensuring 4% global recurrence rate as a basis to all other RPs shown in this figure. **B** Perpendicular RP with angle threshold $\varphi = 15°$, **C** isodirectional RP with $T = 5$ [sampling units] and $\varepsilon_2 = \varepsilon/2$, **D** true recurrence point RP (TRP) with $T_{\min} = 5$ [sampling units], which coincides with the first minimum of the mutual information, **E** thresholded local minima approach with two parameters (LM2P) and $\tau_m = 5$, and **F** diagonal RP.

the decorrelation time or the delay, if time delay embedding is used for reconstructing the phase space vectors from time series.

However, the TRP most often leads to disjoint, deviated diagonal line structures (Fig. 5.7D), which correspond to the white embraced lines in Fig. 5.6D, E.

An alternative would be to use the mid-points of the recurrence sequences. This would also correspond to recurrence times as discussed in Ngamga et al. [230]. In Subsection 5.5.5, we will develop another correction scheme which is motivated by these mid-point based "true recurrences".

5.5.4 RP by means of local minima

Another approach for reducing the effect of tangential motion which shares the basic idea from the TRP approach was introduced by Schultz et al. [277], who track the local minima of the distance matrix (corresponding to the maxima in Fig. 5.6D, E). Wendi and Marwan [326] then extended this idea in order to make the method more robust against noise. However, such local-minima based RP can contain bended or disrupted diagonal line structures. The key idea is to look

for local minima in each column of the distance matrix, illustrated as an orange cross section in Fig. 5.6D, E. If such a local minimum is smaller than the recurrence threshold, then it is a recurrence (LocalMinimaThresholded, LMT). In the two-parameter approach (LM2P) [326] shown in Fig. 5.7E, there is an additional constraint for two consecutive local minima to be displaced by at least τ_m time steps.

5.5.5 Diagonal RP

We now propose an additional approach to cope with the tangential motion, which does not need any additional parameters and leads to an RP of straight, unbended diagonal line structures (Fig. 5.7F). We call this approach the *diagonal RP*, since it generates an RP with only diagonal line structures that are just one point thick.

A diagonal line in a RP corresponds to a connected region in the distance matrix with distances smaller than the recurrence threshold ε (Fig. 5.6D, E, white embraced region). We call such region a "distance range" $\mathfrak{D}$. Typically, the larger ε the larger the $\mathfrak{D}_i$'s in the RP. Moreover, tangential motion, noise and insufficient embedding affect the shape and width of the $\mathfrak{D}_i$'s. For the diagonal line based RQA measures we are interested in these ranges to be represented by single, connected diagonal lines in the corresponding RP. We choose the longest line of each $\mathfrak{D}_i$ to be its adequate representative in the RP. We define the "distance ranges" $\mathfrak{D}_i$ of an RP recursively as a set of adjacent diagonal lines $\mathcal{L}_{\ell_m}^{(m)}$ of length ℓ_m (cf. Eq. (5.1)), initializing with the longest line $\mathcal{L}_{\ell_k}^{(k)}$, for which $\ell_k = \max(\ell \; : \; P(\ell) > 0)$.

$$\mathfrak{D}_i := \{\mathcal{L}_{\ell_k}^{(k)}, \mathcal{L}_{\ell_m}^{(m)} \mid \mathcal{L}_{\ell_m}^{(m)} \curvearrowleft \mathcal{L}_{\ell_{m-1}}^{(m-1)} \curvearrowleft \mathcal{L}_{\ell_{m-2}}^{(m-2)} \curvearrowleft ... \curvearrowleft \mathcal{L}_{\ell_k}^{(k)} \vee$$
$$\mathcal{L}_{\ell_m}^{(m)} \curvearrowright \mathcal{L}_{\ell_{m-1}}^{(m-1)} \curvearrowright \mathcal{L}_{\ell_{m-2}}^{(m-2)} \curvearrowright ... \curvearrowright \mathcal{L}_{\ell_k}^{(k)}\} \tag{5.11}$$

with the line-neighbor-relations $\curvearrowleft$ and $\curvearrowright$ defined by

$$\exists p \in [1, ..., \ell_m] \; \exists q \in [1, ..., \ell_k] :$$
$$(i_m, j_m)_p := \begin{cases} (i_k + 1, j_k)_q \vee (i_k, j_k + 1)_q, & \text{if } \mathcal{L}_{\ell_m}^{(m)} \curvearrowleft \mathcal{L}_{\ell_k}^{(k)} \\ (i_k - 1, j_k)_q \vee (i_k, j_k - 1)_q, & \text{if } \mathcal{L}_{\ell_m}^{(m)} \curvearrowright \mathcal{L}_{\ell_k}^{(k)} \end{cases} \tag{5.12}$$

where $(i_m, j_m)_{p=[1,...,\ell_m]}$ denote the index tuples corresponding to lines $\mathcal{L}_{\ell_m}^{(m)}$ and $(i_k, j_k)_{q=[1,...,\ell_k]}$ denote the index tuples corresponding to the longest line $\mathcal{L}_{\ell_k}^{(k)}$. We then delete all lines contained in $\mathfrak{D}_i$ from the histogram $P(\ell)$ and define the next distance range $\mathfrak{D}_{i+1}$ with a new $\mathcal{L}_{\ell_{k'}}^{(k')}$ from the histogram and so on until $P(\ell)$ is an empty set.

We construct the *new* RP by keeping the longest line of each $\mathfrak{D}_i$ (all the $\mathcal{L}_{\ell_k}^{(k)}$'s). Denote the set of index tuples (i, j) corresponding to the set of longest lines gained from the $\mathfrak{D}_i$'s as $\mathfrak{S}$, then

$$R_{i,j}^{\nearrow} = \begin{cases} 1, & \text{if } (i, j) \in \mathfrak{S} \\ 0, & \text{otherwise} \end{cases} \tag{5.13}$$

Note that this algorithm constricts clusters of adjacent recurrence points to a single diagonal line, representing this "distance range" $\mathfrak{D}$ (skeletonization). Although this method impresses with the absence of additional parameters, caution in its use is advised concerning the choice of the embedding parameters and the recurrence threshold. A wrong setup, specifically a too

high recurrence threshold and/or a "wrong" time delay, could lead to an overall connected RP, which in turn would cause a *diagonal RP* consisting of just one single line in each triangle (if the main diagonal is discarded). However, concerning the sensitivity to the choice of the recurrence threshold, our numerical investigations suggest a rather low risk of this special case and a broad range of threshold values, which do work well (cf. Section 5.6, Appendix. F and figures therein).

5.6 Results: Efficiency of correction schemes

We now apply the correction schemes for counting diagonal lines (Section 5.3) and suppressing tangential motion (Section 5.5) on a time-discrete as well as time-continuous example, in order to test their ability to give valid estimates for the entropy of diagonal line lengths, Eq. (A.9) in Appendix A. In case of the former we choose the Logistic map $x_{n+1} = rx_n(1 - x_n)$ with control parameter $r = 3.5$, leading to regular limit cycle behavior, and control parameter $r = 3.8$, where a chaotic regime is obtained. For the latter we show diagonal line length entropies of RPs of the Rössler system (Eq. (B.3)) in two parameter configurations, also leading to regular limit cycle behavior ($a = 0.15, b = c = 10$) and chaotic motion ($a = 0.15, b = 0.2, c = 10$) [16]. The results shown in this section are based on ensembles of 100 realizations of each parameter setting for the Rössler system and on ensembles of 1,000 realizations of each parameter setting for the Logistic map, gained from randomly chosen initial conditions out of a uniformly distributed interval $x_0 \in [0, 0.5]$ (Logistic map), $x(0), y(0), z(0) \in [0, 2]$ (Rössler system). We numerically integrate the Rössler equations using the explicit Runge-Kutta (4,5) formalism (Dormand-Prince pair) as provided by the ode45-solver in MATLAB [281] with a fixed sampling time of $\Delta t = 0.2$. For both systems we discard the first 2,500 data points as transients, keeping 1,000 (Logistic map) and 2,000 (Rössler) data points as the time series we base our further computations on. For estimating the entropy, we use the Maximum-Likelihood-estimator $p(\ell) \cong \hat{p}(\ell) = \frac{\#\text{number of lines of length } \ell}{\#\text{number of all lines in the RP}}$ for the probabilities.

Generally, we expect (near-)zero entropy values for the regular regime setups and high(er) values for the chaotic regime setups for both considered examples in the noise free case (Section 5.2). Moreover, we expect the correction schemes for counting diagonal lines (Section 5.3) to perform well in case of the Logistic map examples, due to the absence of tangential motion. For the flow data in the Rössler examples, we expect a combination of these correction schemes with the correction schemes for tangential motion described in Section 5.5 to give reasonable results. In order to validate our results, we compute the diagonal line length entropy analytically for the mentioned cases. March et al. [199] gave an expression for this:

$$\text{ENTR}_{\text{theoretical}} = K_2 \left(\frac{1}{\gamma} - 1\right) - \ln\gamma \quad , \tag{5.14}$$

with $\gamma = (1 - e^{-K_2})$ and K_2 the correlation entropy. Practically we compute the largest Lyapunov exponent for our experimental settings [62] and use Pesin's identity to get the Kolmogorov entropy. Because the correlation entropy is a lower bound for the Kolmogorov entropy [116], the reference values computed from Eq. (5.14) should be slightly overestimated.

The results confirm our expectations (Fig. 5.8). While the conventional way of counting diagonal lines, where border effects are not taken into consideration, lead to counterintuitive behavior, all the described correction schemes are able to distinguish chaotic from regular regimes in both exemplary systems. In this laboratory, noise free conditions, the entropy estimates in case of
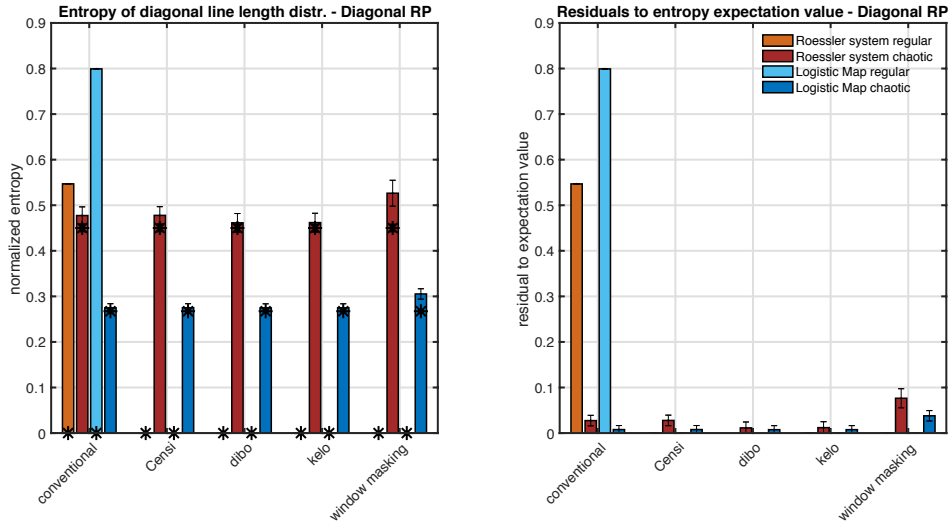
Figure 5.8: Diagonal line length entropy of the proposed diagonal recurrence plot $R^\nearrow$ (cf. Section 5.5.5) of the Rössler system (reddish) and the Logistic map (bluish) in a regular limit cycle regime (bright) as well as in a chaotic regime (dark). Shown are medians of the diagonal line length entropies gained from 1,000 realizations of the Logistic map and 100 realizations of the Rössler example, respectively, for the different line counting correction schemes described in Section 5.3. Errorbars indicate two standard deviations of these distributions. Black stars show medians of ensembles of 1,000 analytically computed values derived from Eq. (5.14) (its errorbars, as two standard deviations of the ensemble distribution, are barely visible and smaller than markers used). Firstly, RPs were obtained with a fixed recurrence threshold corresponding to 19% recurrence rate in case of the Rössler examples and a fixed recurrence threshold corresponding to $1/10$ of the range of the underlying time series in case of the Logistic map examples (for noise free map data the ε-adjustment with respect to the global recurrence rate does not work properly). Then our proposed, parameter free correction scheme leading to the diagonal recurrence plot $R^\nearrow$ was applied. Results for a range of recurrence thresholds and for all tangential motion RP-correction schemes are shown in Fig. F.1 and Fig. F.2 in the Appendix F.

the regular limit cycle regimes are zero (or in case of the *dibo*-correction scheme not defined, due to the absence of any diagonal line). For *dibo* and *kelo* the estimated values for the chaotic Rössler regime fall within the two standard deviation margin of the theoretical values, whereas Censi's correction scheme comes very close and the windowshape correction scheme misses it by approx. 5%. Again, we have to stress that we expect the expectation values to be underestimated, i.e. we assume Censi's method and the window masking do also perform well. Let us stick to the *kelo* correction scheme for now and look how the different correction schemes for tangential motion perform (Fig. 5.9). First of all we have to mention that we were not able to produce any kind of reasonable estimates while using the perpendicular recurrence plot $R^\perp$, regardless of the angle threshold parameter. This straightforward approach is extremely sensitive to any kind of noise and to the sampling time of the system under observation. It needs a fairly high density of state space points, in order to yield a non empty RP and, thus, any meaningful diagonal line length entropy estimate. Hence, we skip this approach in our further analysis, especially the dependence of the shown results to the choice of the recurrence threshold and additive noise, but will discuss the performance of the perpendicular RP for a high sampled Rössler setup in
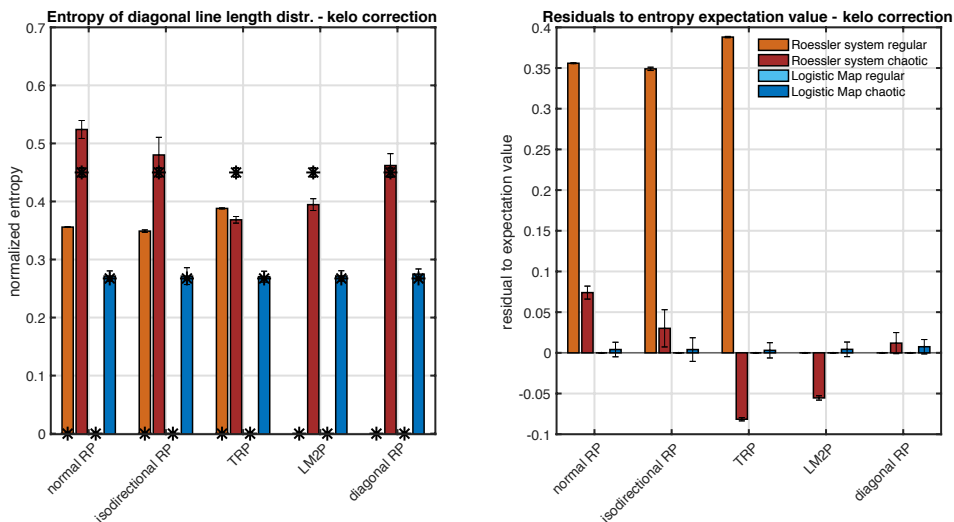
Figure 5.9: Diagonal line length entropy based on the proposed line counting correction scheme *kelo* (cf. Section 5.3.1) for the Rössler system (reddish) and the Logistic map (bluish) in a regular limit cycle regime (bright) as well as in a chaotic regime (dark). Shown are medians of the diagonal line length entropies gained from 1,000 realizations of the Logistic map and 100 realizations of the Rössler example, respectively, for all the different tangential motion correction schemes described in Section 5.4, but the perpendicular recurrence plot $R^\perp$. Errorbars indicate two standard deviations of these distributions. Black stars show medians of ensembles of 1,000 analytically computed values derived from Eq. (5.14) (its errorbars, as two standard deviations of the ensemble distribution, are barely visible and smaller than markers used). The normal RP with a fixed recurrence threshold corresponding to 19% recurrence rate in case of the Rössler examples and a fixed recurrence threshold corresponding to $1/10$ of the range of the underlying time series in case of the Logistic map examples (for noise free map data the ε-adjustment with respect to the global recurrence rate does not work properly) serves as a basis for the RP correction schemes shown here. Results for a range of recurrence thresholds and for all tangential motion RP-correction schemes are shown in Fig. F.1 and Fig. F.2 in the Appendix F.

the next subsection. For a general use, we cannot recommend the application of perpendicular RPs. Coming back to the results (Fig. 5.9), solely the LM2P approach and the diagonal RP perform as expected (zero-values in case of the regular regime setups and higher values for the chaotic regimes, clearly distinguishable). Only the proposed diagonal RP is able to give estimates within the errorbars of the theoretical values (which is why only this approach was selected for Fig. 5.8). Note that the reference values slightly underestimate the "true" value and we cannot quantitatively correct for this bias. As in Fig. 5.7, we set the parameters T, $T_{\min}$ and τ_{m} to the corresponding first minimum of the auto mutual information and the second recurrence threshold for the isodirectional RP was again set to $\varepsilon_2 = \varepsilon/2$, but we tried many parameter configurations.

5.6.1 Results for highly sampled data and the effect of noise

For the sake of completeness and in order to investigate the behavior of our proposed methods under more realistic conditions, we now look at the noise corrupted Rössler system in the two dynamical regimes (Section 5.6), but with an increased sampling frequency (sampling time $\Delta t = 0.08$) and with total lengths of the three numerically integrated time series of 10,000
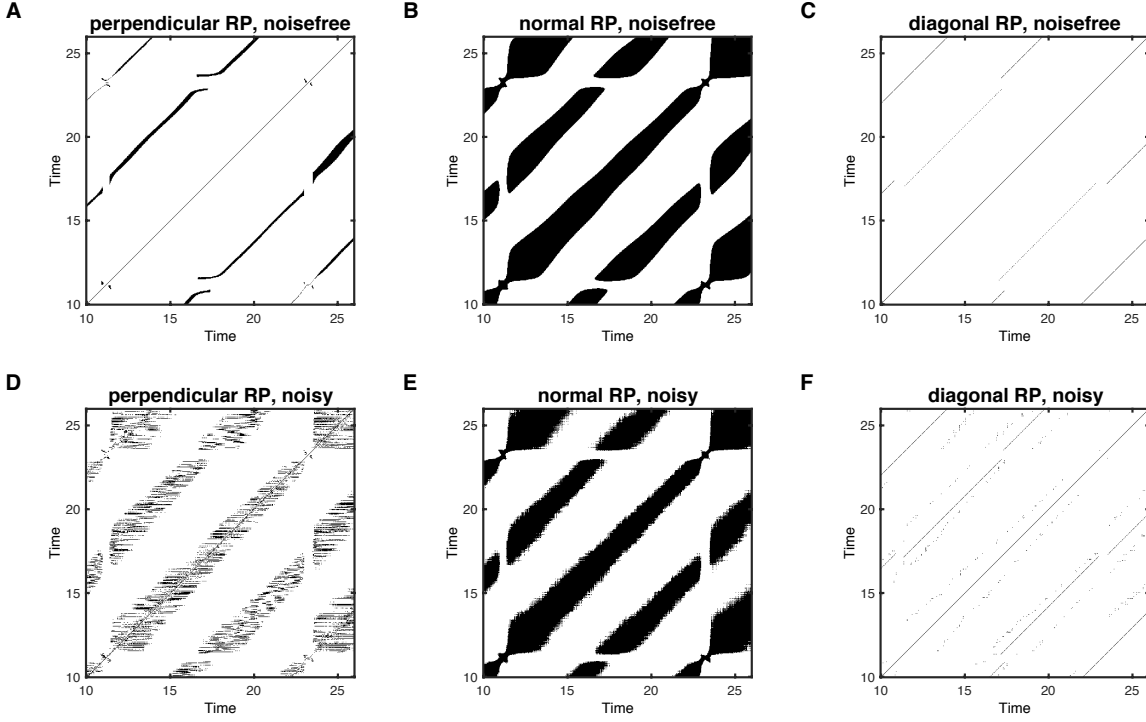
Figure 5.10: Cut outs of **A**+**D** the perpendicular recurrence plot $R^\perp$, **B**+**E** normal RP, and **C**+**F** the diagonal recurrence plot $R^\nearrow$ of the highly sampled Rössler system in chaotic regime (see text for details). Top panels **A**-**C** show noise free cases, bottom panels **D**-**F** show their noise contaminated counterparts. Computations have been carried out by using a fixed recurrence threshold corresponding to 35% recurrence rate and an angle threshold $\varphi = 15°$ for $R^\perp$.

(transients already removed). In this setup the perpendicular recurrence plot $R^\perp$ (Section 5.5.1) yields meaningful results (Fig. 5.10), and we compare its utility with respect to the estimation of the diagonal line length entropy to the normal RP and the novel diagonal recurrence plot $R^\nearrow$ (Section 5.5.5).

Figure 5.11 illustrates the capability of $R^\nearrow$ to cope with tangential motion, especially under noise. Due to a too high computational effort we did not compute an ensemble in this case as we did in the lower sampled cases, so the errorbars are missing. Here we added an auto regressive (AR) process of second order with an amplitude corresponding to 20% of the mean standard deviation of the multivariate signal.

$$x_t = a_1 x_{t-1} + a_2 x_{t-2} + \varepsilon_t \quad , \tag{5.15}$$

with parameters $a_1 = 0.7$, $a_2 = 0.15$ and ε_t denotes a white noise process with zero mean and constant variance of unity. Outcomes for the normal RP and the perpendicular recurrence plot $R^\perp$ can be found in the Appendix (Figs. F.5, F.6). Additive white noise of the same magnitude gave similar results to the ones discussed here.

As expected from the examples in the last section, the diagonal RP approach performs well under noise free conditions and all, but the conventional line counting algorithms yield zero-value entropy estimates for the regular regime (panel B) and clearly non-zero entropies in case of the

A

B

C

D
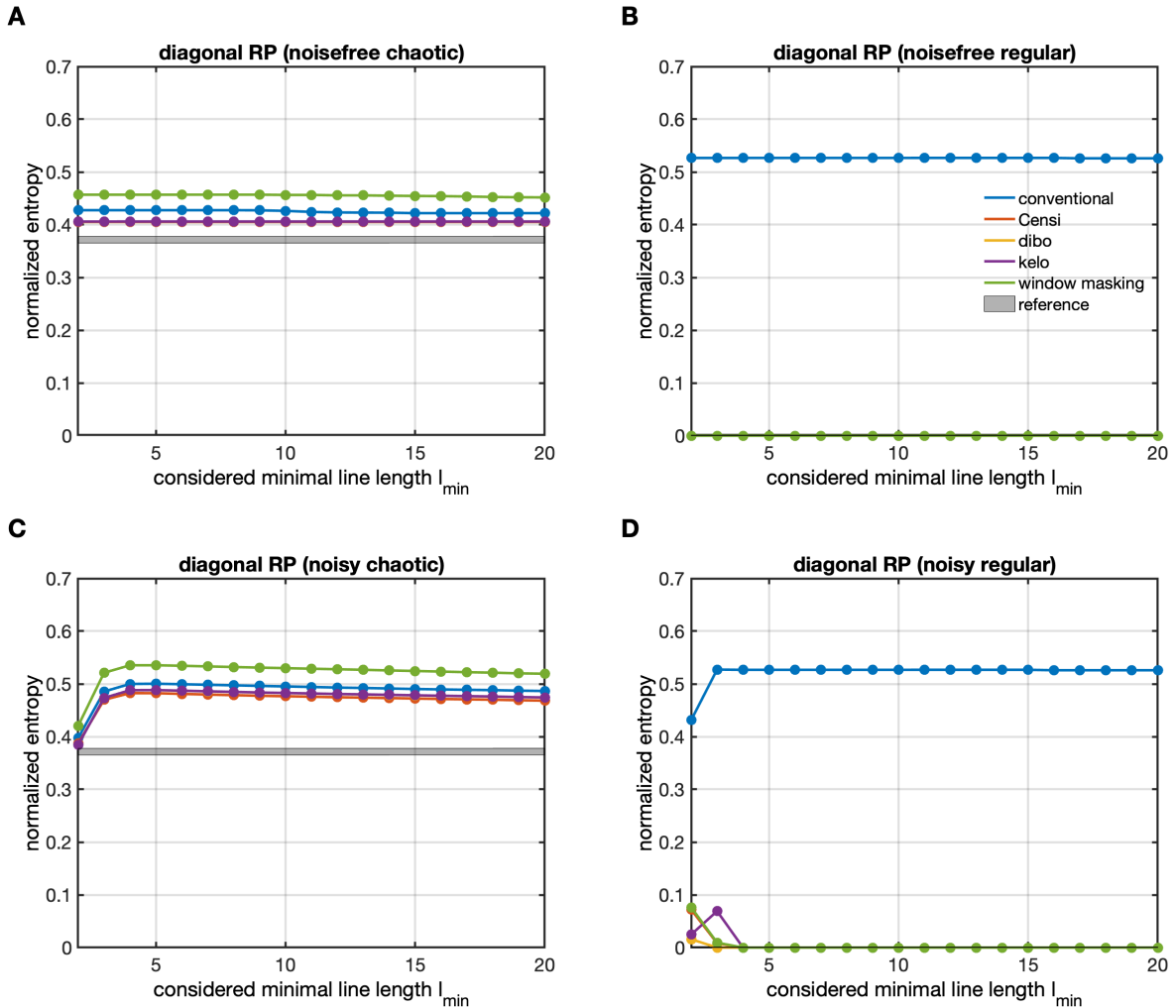
Figure 5.11: Normalized diagonal line length entropy estimates for all described correction schemes for counting diagonal lines (Section 5.3) based on the diagonal recurrence plot $R^\nearrow$ (Section 5.5.5) of the high sampled Rössler system as a function of the chosen minimal line length ℓ_{min}. The top panels (**A** - chaotic motion , **B** - regular motion) show the noisefree case and in the bottom panels (**C** - chaotic motion, **D** - regular motion) the results for noise corrupted data are shown. We added noise from an auto-regressive (AR) process of second order as 20% of the mean standard deviation of the multivariate signal gained from the numerical integration (cf. Eq. (5.15)). The underlying RPs for obtaining $R^\nearrow$ were computed using a fixed recurrence threshold corresponding to 35% recurrence rate. The grey shaded areas show medians of ensembles of 1,000 analytically computed reference values for $K_1 \pm$ two standard deviations of these distributions transformed by using Eq. (5.14).

chaotic regime (panel A) close to the underestimated reference values. The perpendicular RP also performs well in noise free conditions (Fig. F.5). Even the conventional line length counting leads to the desired zero entropy estimates in case of regular motion. In the presence of noise, however, $R^\perp$ is not able to distinguish regular from chaotic behavior (Fig. F.6), whereas $R^\nearrow$ still performs well, giving almost the same results as in the noise free setup. The explanation can be found in considering the RPs (Fig. 5.10). For this noiselevel our proposed skeletonization approach (Fig. 5.10F) leaves small lines of maximum length 4 after its application to the noisy normal recurrence plot (Fig. 5.10E) as noise-leftovers. The appearance of these lines is not a result of the dynamics itself. Noise enriches the RP and its corresponding diagonal line histogram with small line lengths depending on the noiselevel ([295], Fig. 5.10, Fig. 5.2A, Fig. 5.3A). By increasing the minimum line length one gradually discards the majority of the lines contained in the histogram and, thus, increases the prominence of larger line lengths for the computation of the entropy. For a regular regime, the distribution of lines of intermediate length is broader for all the correction schemes, but the diagonal RP. Therefore an increasing minimal line length increases the entropy in the presence of noise for all the correction schemes, but the diagonal RP (cf. Fig. F.6). In case of a chaotic regime the distribution of diagonals due to the dynamics is broader anyway (Fig. 5.2C, Fig. 5.3C) leading to the same effect.

When increasing the minimal line length for the diagonal RP, the entropy estimates stay more or less constant after a certain, sufficiently high, minimum line length, which depends on the noiselevel (Fig. 5.11C, D). The offset to the underestimated reference value for the chaotic case grows for increasing noiselevels. Note that the effect of additive noise is harder to tackle for the tangential motion correction schemes for high sampled data like in this case, than it is for lower sampled examples as discussed in Section 5.6. The higher the sampling, the finer the ramification of distance ranges $\mathfrak{D}_i$ (thickened diagonal lines). Results for all correction schemes for a wide range of the recurrence thresholds and under the influence of white noise for the lower sampled situation can be found in the Appendix (Figs. F.3 and F.4).

5.7 Conclusion

In this chapter we investigated the effect of the finite size of a recurrence plot on its diagonal line length based quantification. Specifically, we showed how these border effects influence the diagonal line length entropy and proposed three new line length counting correction schemes, which take these effects into account (cf. Sections 5.3.1, 5.3.1, 5.3.2) and systematically compared them to an already proposed correction by Censi et al. [48] (Section 5.3.1). It turned out that for noise free or slightly noise corrupted map data all these correction methods solve the problem of the biased diagonal line length entropy due to lines cut by the borders of the RP. However, for flow data the effect of tangential motion has a much bigger influence on the entropy bias than the border effects. Therefore, we systematically compared already proposed ideas to handle tangential motion and proposed a new, parameter free method, the *diagonal RP* (cf. Section 5.5.5). It can properly tackle the tangential motion effects and yield, in combination with the border effect correction schemes, meaningful estimates for the diagonal line length entropy. We have to emphasize that this method, in contrast to other suggested ideas, also works for noise contaminated data, is not sensitive to the particular choice of the recurrence threshold, does not introduce any additional parameter, and is, therefore, easy to use. In case of a noise corrupted flow-like signal the diagonal line length entropy approaches its constant expectation value for sufficiently high choices of the minimal line length, when the diagonal RP together with Censi's or our propsed border effect

correction schemes is used. Fairly high recurrence thresholds ($>10\%$ recurrence rate) favour the *diagonal RP* method for intermediate or high noise levels.

Code availability

We offer a MATLAB® implementation of the proposed algorithms at *https://zenodo.org/record/ 3595503/files/hkraemer/Border-effect-corrections-for-diagonal-line-based-recurrence-quantification- analysis-measures-2.1.zip?download=1* [165]

Chapter 6

Recurrence powerspectra

Abstract

A novel kind of power spectrum is constructed, the *inter spike spectrum*, which transforms spike-train-like signals into their frequency domain. This method clearly shows the apparent cycles in the data and overcomes the problems when using the obvious idea of Fourier-transforming it. We invent this instructive approach with the idea of transforming the τ-recurrence rate of a recurrence plot (RP), which often has a spiky appearance. The τ-recurrence rate is the density of recurrence points along diagonals of the RP, which are parallel to the main diagonal with a distance of τ. In this context the inter spike spectrum can be interpreted as a nonlinear power spectrum of a potentially high dimensional system which constitutes the RP. The proposed measure is simple to compute, robust to noise and is able to detect bifurcations inducing regular-regular, regular-chaos as well as chaos-chaos transitions.

6.1 Introduction

Recurrence Plots (RPs) provide a vivid representation of complex dynamics stemming from potentially high dimensional systems, Eq. (1.1). The simple idea to track recurring states of the d-dimensional trajectory $\vec{x}_i$ of the system under study not only allows for a beneficial visualization of the dynamics, but also for its quantification, using certain structures in the RP, such as diagonal or vertical lines [209] (see Appendix A). Some of these recurrence quantification measures, the entropy of diagonal lines and the entropy of recurrence times, can be related to the Kolmogorov-Sinai entropy [15, 199]. However, these quantifiers have a free parameter, the minimal considered line length, and are usually biased, due to the finite size of the RP and thickened diagonal lines, which needs to be corrected [164] (Chapter 5). Moreover, the mentioned statistics cannot account for changing regular (non-chaotic) dynamics, such as period-doubling bifurcations.
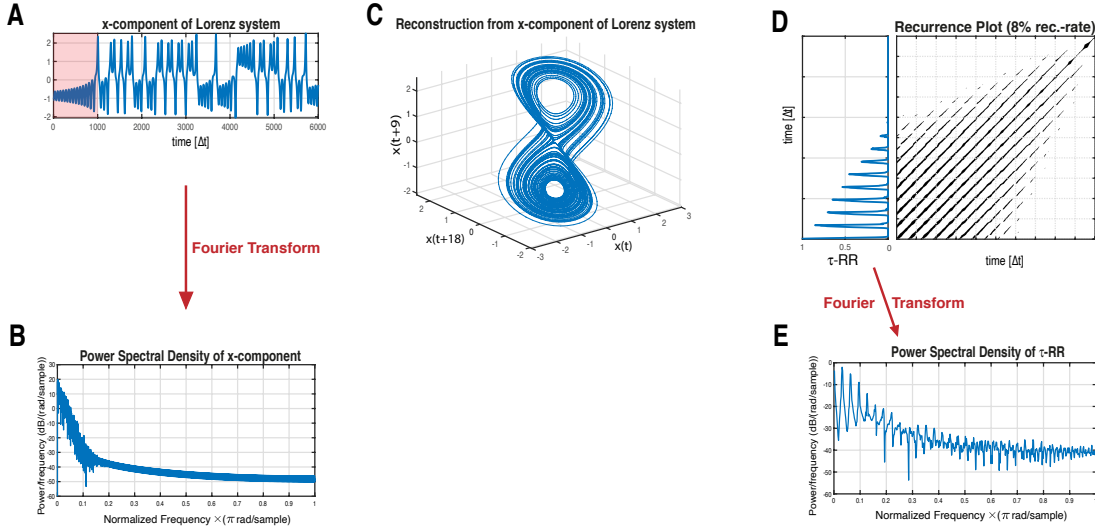
Figure 6.1: Schematic illustration of a τ-recurrence rate based spectrum. **A** x-component time series of the Lorenz63-System (Eq. (B.1)) and **B** its corresponding Fourier power spectrum. **C** Reconstructed state space portrait from the time series shown in **A** using PECUZAL time-delay embedding [162]. **D** Subset of the recurrence plot and the corresponding τ-recurrence rate obtained from the state space trajectory in **C**. The shaded interval in the time series in **A** corresponds to the shown subset. **E** Fourier Power spectrum obtained from the τ-recurrence (subset shown in panel **D**) [338].

A rather simple idea is to look at the τ-recurrence rate of the RP (τ-RR, Eq. 6.1) [206, 338]. This is the density of recurrence points along the diagonals of the recurrence matrix, as a function of the distance τ (sampling units) to the main diagonal:

$$\tau\text{-RR}(\varepsilon) = RR(\tau, \varepsilon) = \frac{1}{N - \tau} \sum_{i=1}^{N-\tau} R_{i,i+\tau}. \tag{6.1}$$

τ-RR serves as an estimator for the probability that the system recurs after time $\tau\Delta t$, with Δt being the sampling time of the trajectory $\vec{x}_i = \vec{x}(\Delta t \cdot i)$, $i = 1, \ldots, N$. Zbilut and Marwan [338] pointed out that τ-RR could be used as a plugin value for the auto-correlation function $C(\tau)$ and, hence, via the Wiener-Khinchim theorem a "generalized" powerspectrum can be obtained. This is reasonable, since the average distances for a given lag τ

$$\overline{D}(\tau) = \frac{1}{N - \tau} \sum_{i=1}^{N-\tau} D_{i,i+\tau} \tag{6.2}$$

can be directly read from the distance matrix **D** and is also preserved in its thresholded version τ-RR. There are clear advantages for a recurrence-derived powerspectrum, i.e., Fourier transforming (FT) τ-RR (Fig. 6.1D), instead of $C(\tau)$ (Fig. 6.1C): There are no assumptions for stationarity or sampling, when constructing a RP. Furthermore, the correlation structures of higher dimensional spaces can be resolved in the recurrence-derived Fourier-spectrum.

However, there is a drawback to this approach. Whenever τ-RR is a spike-train-like signal, which it is in most cases (see Fig. 6.1) especially for map-data (low-resolution data), a FT of such

a signal leads to a spike-train-like image in the frequency domain (e.g. [58, 271], see Fig. 6.1D). Thus, it is not intuitive how to extract meaningful information about dominant frequencies of the systems' state space trajectory.

For clarification, consider the signal we would like to analyze (e.g. the τ-RR of a system) to be a Dirac comb (DC) with inter-spike period T_{is}:

$$\mathrm{DC}_{\mathrm{is}}(t) = \sum_{k=-\infty}^{\infty} \delta(t - kT_{\mathrm{is}}), \tag{6.3}$$

i.e., a series of Dirac delta functions for a period T_{is}. There is only one single period – T_{is} – in this signal (Fig. 6.2A, D), so in principle we would strive for a single peak in the frequency domain of this signal at a frequency $f = 1/T_{\mathrm{is}}$. Surprisingly the Fourier spectrum does not meet this expectation and instead of a single frequency, there are exceptionally many frequencies excited (Fig. 6.2B, E). This is, because the Fourier components add constructively for every frequency $1/T_{\mathrm{is}}$ and therefore $\mathrm{DC}_{\mathrm{is}}(t)$ coincides with its own Fourier transform up to a factor $1/T_{\mathrm{is}}$.
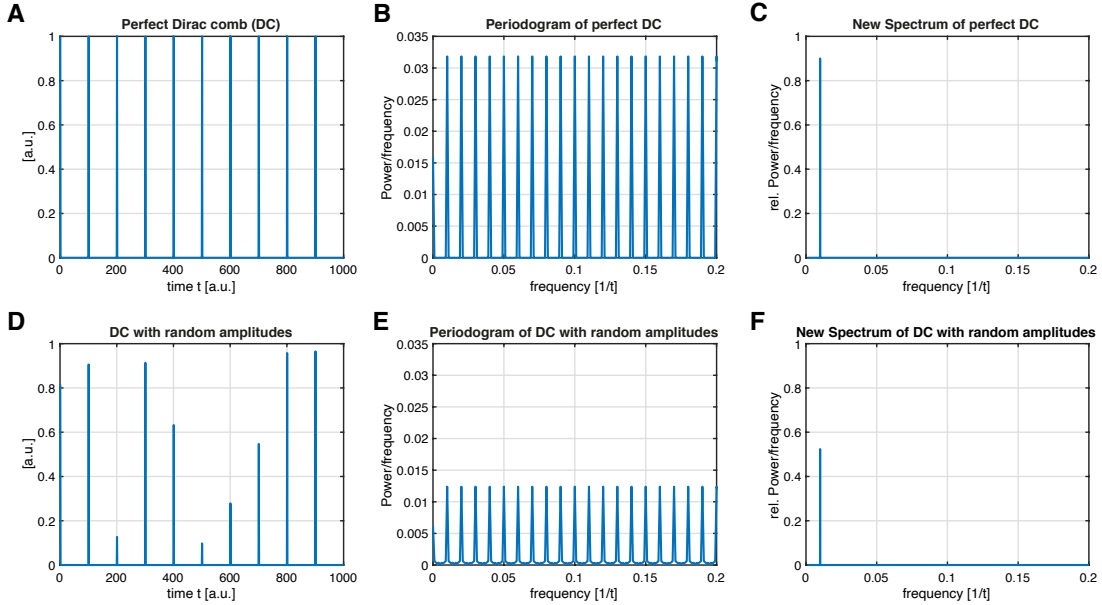
Figure 6.2: The transformation of a Dirac comb (series of Dirac delta functions) with a single inter-spike period $T_{\mathrm{is}} = 100$ ($\hat{=} f = 0.01$) into the frequency domain. **A** Dirac Comb (DC) with equal amplitudes and **B** its FFT-based powerspectrum. **C** Proposed inter spike spectrum of the signal in **A** showing a single frequency, which corresponds to the inter-spike period T_{is} ($f = 0.01$). **D** DC with randomly chosen amplitudes and same T_{is} as in **A**, and **E** its FFT-based powerspectrum. **F** Proposed inter spike spectrum of the signal in **D** showing a single frequency, which corresponds to the expected inter-spike period T_{is} ($f = 0.01$).

In this chapter we propose a new way of transforming a spike-train-like signal into its frequency domain. This novel *inter spike spectrum* does not show resonance behavior of the signal's inherent inter-spike frequencies (Fig. 6.2C, F). Section 6.2 explains the idea, which can be used to decompose any arbitrary signal and is not restricted to the τ-RR, which we exemplify here. However, the more spiky the signal is, the more outstanding our new approach is compared to a FT. In Section 6.3 we exemplify its use when transforming the τ-RR of a system under study.

In this case, the inter spike spectrum can unravel characteristic time scales of high dimensional systems, which is not possible when using a FT.

6.2 Method

The idea is as simple as it is obvious. The signal, which we would like to transform – we focus on τ-RR in this chapter, but this can be applied to any sort of signal – is decomposed into a set of appropriate basis functions. Instead of using trigonometric functions, as it is the idea in the Fourier decomposition, we use Dirac combs (DC) with different inter-spike periods as basis functions, Eq. (6.3). Let $s(t_i)$ be the normalized signal we want to transform with length N and $t_i = i \cdot \Delta t$, $i = 1, \ldots, N$, where Δt denotes the sampling time and $s(t_i) \in [0, \ 1] \ \forall \ i$. In the following we label this time series as a $(1 \times N)$-dimensional vector $\mathbf{s}$. First, N different DC's of length N are constructed with inter-spike periods $T_{\text{is}} \in [1, \ldots, N]$. Second, in order to account for possible phase shifts of these basis functions occurring in $\mathbf{s}$, each of these N different DC's also need to be shifted one step further $T_{\text{is}} - 1$ times. This leaves us with a total number of $M = \sum_{i=1}^{N} i$ basis functions which we can arrange as rows of a $(M \times N)$-sized matrix $\mathbf{X}$ (Fig. 6.3 illustrates the described procedure)

$$\mathbf{X}_{i,j} = \sum_{k=0}^{N} \delta \left(j - 1 - kT(i) - i + T(i) \right), \tag{6.4}$$

$$T(i) = n \quad , \forall n : \ \frac{n(n-1)}{2} + 1 \leq i < \frac{n(n+1)}{2} + 1, \quad n \in \mathbb{N}_{+}. \tag{6.5}$$

Due to the shifting of each of the basis functions of inter-spike period T_{is}, $\mathbf{X}$ is not linear independent anymore. Thus, an under-determined linear system

$$\mathbf{X}^T \boldsymbol{\beta} = \mathbf{s} \tag{6.6}$$

has to be solved for $\boldsymbol{\beta}$, a $(M \times 1)$-sized vector carrying the loadings we are interested in. Along a variety of algorithms, which can solve this problem, we are particularly interested in those solutions, which promote sparsity in $\boldsymbol{\beta}$, since our goal is to decompose the signal $\mathbf{s}$ into a minimal number of basis functions (for an excellent overview of the topic we refer to Brunton and Kutz [31]). Therefore, we here choose the *least absolute shrinkage and selection operator* (LASSO regression) [298]. Finally, we group loadings which correspond to basis functions having the same period T_{is} into $\boldsymbol{\beta}_f$ and obtain the *inter spike spectrum* by simply plotting $\boldsymbol{\beta}_f$ as a function of the frequency $f = T_{\text{is}}^{-1}$ (Fig. 6.2C, F).

Figure 6.3: Example for a set of basis functions for an input signal of length $N = 5$, aligned in the matrix $\mathbf{X}$.

6.3 Application

We exemplify the use of the inter spike spectrum in combination with the τ-RR as outlined in Section 6.1 on several interesting research questions. The procedure is the following:

(1) Compute a RP of the trajectory of the system, Eq. (1.1).
(2) Compute the τ-RR of that RP, Eq. (6.1).
(3) Transform the τ-RR into the proposed inter spike spectrum, see Section 6.2.

6.3.1 Period estimation for different dynamics in the Rössler system

First, we consider the Rössler system (Eq. (B.3) in Appendix B.3) in three different dynamical setups. We use the proposed inter spike spectrum to identify the type of dynamics. We set the parameters $b = 2$, $c = 4$ and analyze period-2 limit cycle dynamics ($a = 0.36$, Fig. 6.4A, D, G, J), period-3 limit cycle dynamics ($a = 0.41$, Fig. 6.4B, E, H, K) and chaotic dynamics ($a = 0.428$, Fig. 6.4C, F, I, L).

The inter spike spectra unravel the specific dynamics, which are also apparent in the state space portraits (Fig. 6.4A, B, C) and in the τ-RRs (Fig. 6.4G, H, I). The proposed idea is also robust to noise (see Fig. G.1 in the Appendix). This is because the peaks of the τ-RR are insensitive to noise. While the peak shape does change in the presence of noise, its position does not, and this is what the inter spike spectrum encrypts after all.

6.3.2 Bifurcations in the Logistic map

We consider the Logistic map $x_{n+1} = r \cdot x_n \left(1 - x_n\right)$ for changing control parameter r. We vary r from $r = 3.4$ to $r = 4$ in steps of 0.001. For each setting of r

(1) a time series of length $N = 201$ is computed with a random initial condition $u_0 \in [0, 1]$, neglecting the first $1,000$ samples as transients,
(2) 100 iterative Amplitude Adjusted Fourier Transform (iAAFT) surrogates [275, 276] are computed,
(3) the time series and its iAAFT surrogates are embedded in a 2-dimensional state space using a time delay of unity,
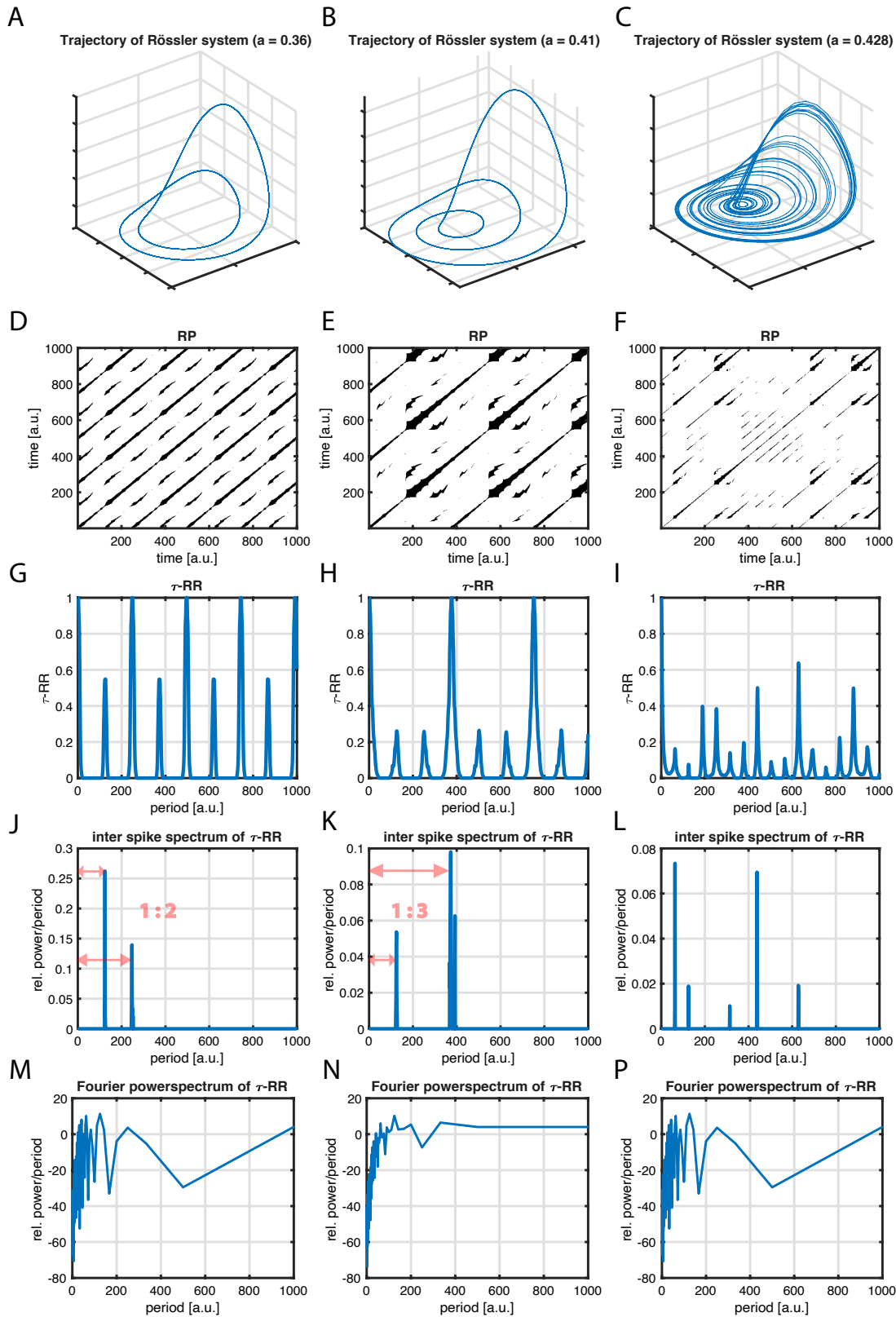
Figure 6.4: Inter spike spectra of the τ-RR of the Rössler system in three different dynamical regimes with parameters $b = 2$, $c = 4$. Trajectory of the system in a **A** period-2 (parameter $a = 0.36$), **B** in a period-3 (parameter $a = 0.41$) and **C** in a chaotic regime (parameter $a = 0.428$). **D, E, F** The corresponding RPs, obtained by using a recurrence threshold corresponding to a 10% global recurrence rate for D & E and 5% for F. **G, H, I** τ-RR's of the shown RPs. **J, K, L** The proposed inter spike spectra of the τ-RR's shown in panels G, H, I. The distance ratio of the peaks reflect the limit cycle dynamic. **M, N, P** Fourier power spectra of the τ-RR's shown in panels G, H, I.

(4) from the 2-dimensional trajectories RPs, Eq. (1.1), are computed under a threshold $\varepsilon = 0.05$,

(5) τ-RR, Eq. (6.1), is computed from the RP of the signal and from the RPs of the surrogates,

(6) spike powerspectra are obtained from τ-RR of the signal and from the τ-RRs of the surrogates, see Section 6.2, and finally,

(7) from the distribution of the surrogate spike powerspectra the 95th percentile is computed. The peaks of the inter spike spectrum of the signal which exceed this percentile are counted.

In this example, the according null hypothesis is that the data stems from a process which yields the same auto-correlation, hence the same Fourier powerspectrum, and the same amplitude distribution. We consider the number of significant peaks in the inter spike spectrum with respect to the control parameter in order to distinguish the corresponding dynamics (Fig. 6.5C). A correlation with the positive Lyapunov exponent (Fig. 6.5A) is discernible ($\rho_{\text{Pearson}} = 0.88$). Moreover, this analysis can tackle period-doubling, since it "measures" the dominant cycles via the inter spike spectrum.
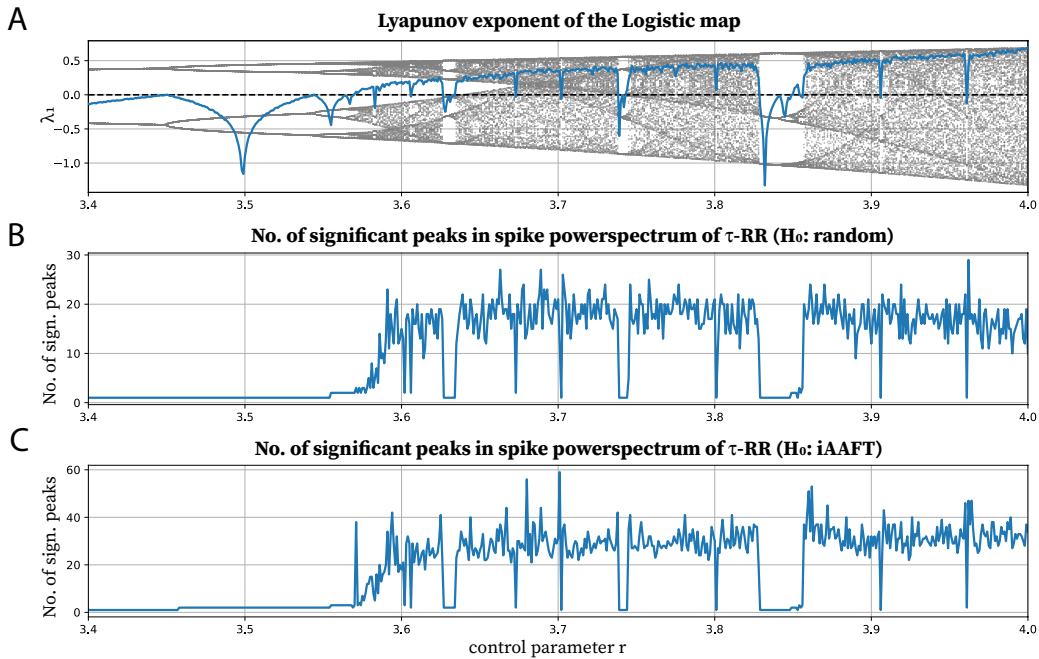
Figure 6.5: A Lyapunov exponent of the Logistic map as a function of the control parameter r. **B** Number of significant peaks ($\alpha = 0.05$) in the inter spike spectrum of the τ-RR, random surrogates. **C** Same as **B**, but for iterative Amplitude Adjusted Fourier Transform (iAAFT) surrogates [275, 276].

A less computationally intensive approach is to compute surrogates for the τ-RR analytically, rather than computing a RP and its τ-RR for each iAAFT surrogate of the time series. This translates into a null hypothesis that the τ-RR and its corresponding inter spike spectrum stems from a RP of a random signal. In this case the probability of finding a black point in the RP can be obtained from a binomial distribution with probability parameter p set to the recurrence rate of the RP of the signal. This way 100 surrogate τ-RRs are computed in step (5). The results are similar to the ones obtained from the iAAFT surrogates and nearly as good (Fig. 6.5B). The first

period doubling at $r \approx 3.458$ cannot be detected, while the correlation to the positive Lyapunov exponent is only slightly worse ($\rho_{\text{Pearson}} = 0.84$).

The described procedure does work well for map data, because most often the τ-RR for those kind of data reveals a "spiky enough" nature. On the contrary, highly sampled (flow-) data often yield not as spiky τ-RR's and, thus, the number of significant peaks in the inter spike spectrum may not be sensitive enough to detect period-doubling bifurcations. Nevertheless the according spike powerspectra is still revealing important information (Fig. 6.4) and practitioners can design appropriate quantifying statistics based on these spectra, which suit the research task.

6.3.3 Evolutionary spike spectrogram of CENOGRID isotope time series

We again utilize the CENOzoic Global Reference benthic foraminifer carbon and oxygen Isotope Dataset (CENOGRID) [328] (c.f. Chapter 3, Section 3.3.3). The time series have a total length of $N = 13,421$ samples and a sampling period of $\Delta t = 5,000$yrs. The records and their corresponding RPs are shown in Figure 6.6 together with remarkable climate transitions in the Cenozoic (detrended time series in Fig. D.2 in Appendix D.2). Among other things, the large-scale patterns in the RPs shown were used by Westerhold et al. [328] to classify major climate states: Hothouse, Warmhouse, Coldhouse and Icehouse.
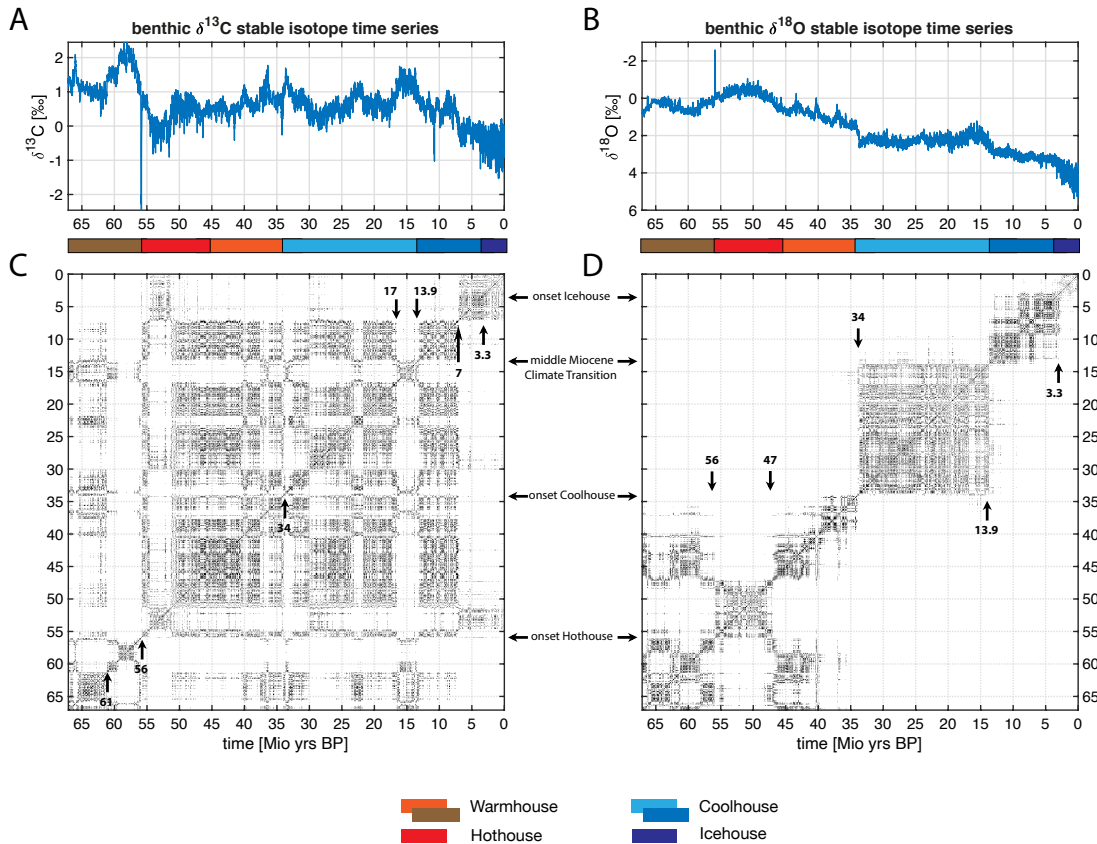
Figure 6.6: A δ^{13}C and **B** δ^{18}O isotope records and their corresponding RPs (**C+D**), obtained from the non-embedded time series under a fixed recurrence threshold corresponding to 8% global RR. These plots were used by Westerhold et al. [328] to classify the major climate states Hothouse, Warmhouse, Cold- and Icehouse.
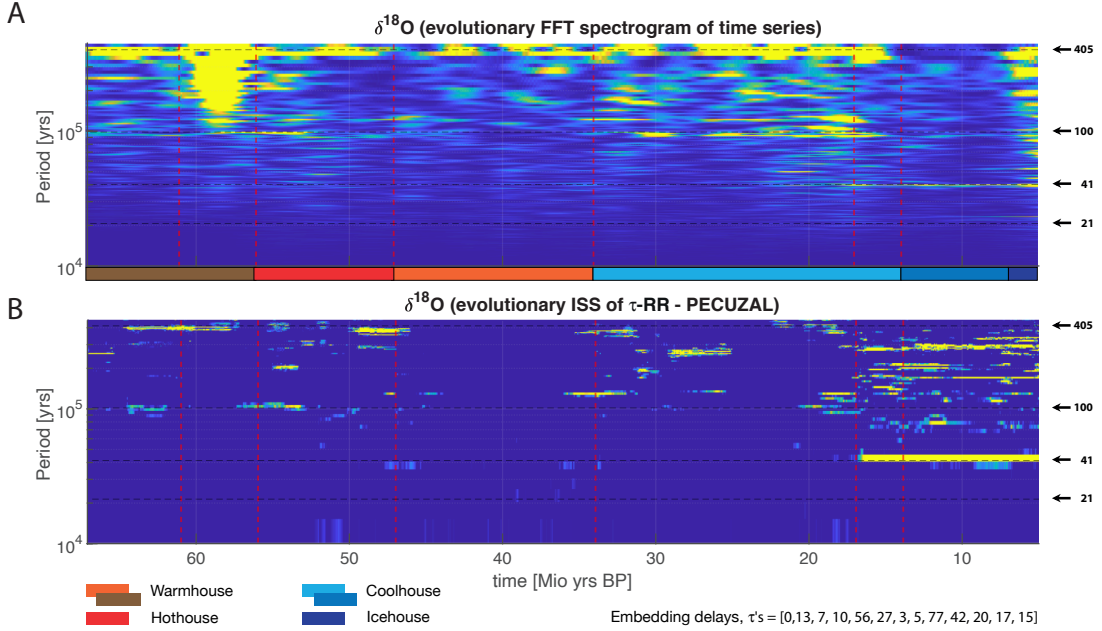
Figure 6.7: Evolutionary spectrograms based on **A** FFT powerspectra of the detrended δ^{18}O time series and **B** on the inter spike powerspectra of the τ-RR of the PECUZAL embedded δ^{18}O record. The spectrograms were computed with a $w = 1,000 (\hat{=}$ 5-Myr) window and a step $ws = 1$. We used the embedding computed in Chapter 3 (Section 3.3.3, see Table D.1) and a recurrence threshold fixed to 5% global RR, in order to ensure comparability within the different windows [163] (Chapter 4). Dashed red horizontal lines indicate major climate transitions according to Westerhold et al. [328].

While in Chapter 3, Section 3.3.3 predictions were made for the detrended δ^{13}C record, here the focus is on the analysis of the characteristic time scales in both, the detrended δ^{18}O and the detrended δ^{13}C isotope records. Specifically, we have computed inter spike spectra of the τ-RR windowed over each record. Figures 6.7, 6.8 show that major features of their linear counterparts, namely ordinary FFT spectrograms (A panels), are also present in these spike spectrograms. But there is additional information in the spike spectrograms (B panels). Firstly, overall there seem to be less frequencies excited than in the FFT spectrograms, but these excited frequencies are therefore more pronounced. This is especially apparent in the oxygen isotope record spectrogram (Fig. 6.7), in which the obliquity-related time scale ($\sim$41 kyrs) is very prominent from 17-0 Mio yrs BP in the inter spike spectrogram (panel B). Our inter spike spectrograms also support the general conclusion given by Westerhold et al. [328] that "in the Hothouse and Warmhouse climate states, as well as the first Coolhouse phase, eccentricity-related cycles ($\sim$100 & $\sim$405 kyrs) dominate the CENOGRID records, while after the increasing influence of highlatitude cooling and ice growth during the second Coolhouse phase, the obliquity-band response steadily increases after 13.9 Mio yrs BP (middle Miocene climate transition) before dominating climate dynamics by $\sim$5 Mio yrs BP (late Miocene–early Pliocene)". The inter spike spectrograms indicate a strong obliquity influence already from 17 Mio yrs BP (the onset of the Miocene Climate optimum). Moreover, in the carbon isotope record based inter spike spectrogram (Fig. 6.8B) there is a strong influence of the 100-kyr-eccentricity cycle in the first Warmhouse phase from $\sim$61-56 Mio yrs BP, which is not represented in the FFT-based spectrogram (panel A).
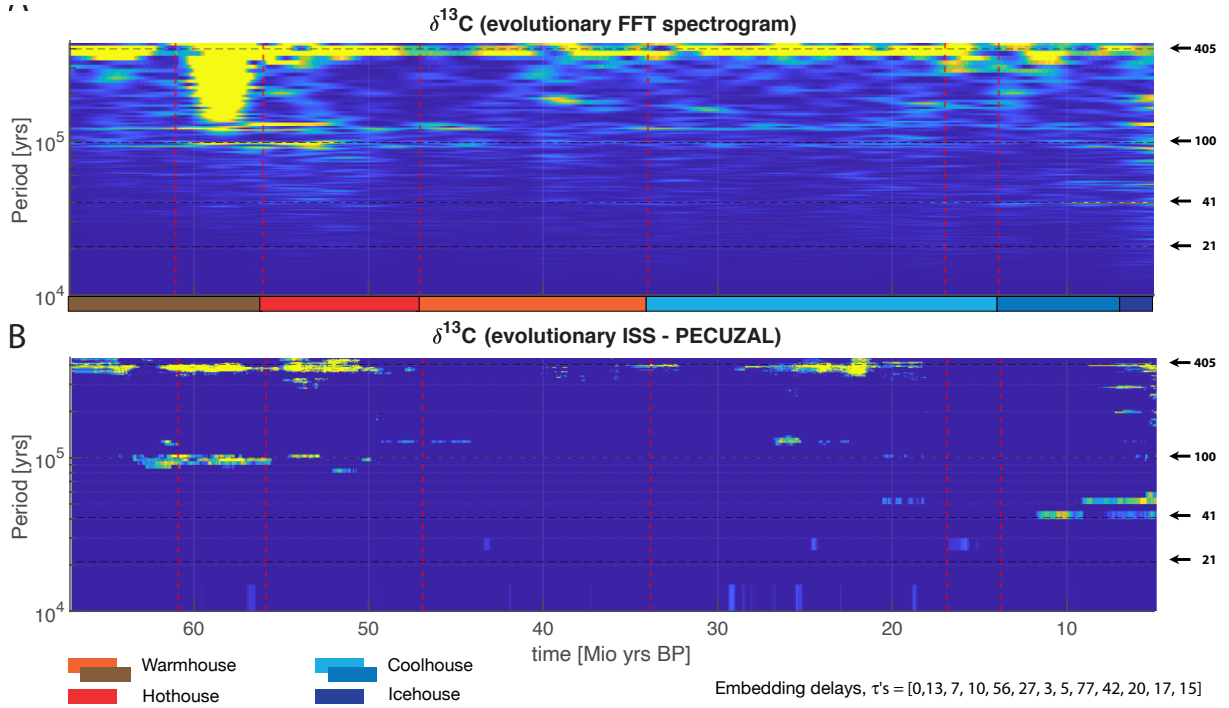
Figure 6.8: Same as in Figure 6.7, but for the detrended δ^{13}C isotope record.

The shown inter spike spectrograms in Figs. 6.7, 6.8 are based on the τ-RR of the respective records, which have been time delay embedded by using the PECUZAL algorithm [162] (Chapter 2). By construction, we would expect different inter spike spectrograms, when different embeddings are used. This is indeed the case as shown in Figures G.2, G.3, G.4, G.5 (Appendix G.2), which also show inter spike spectrograms, but for an embedding using MCDTS-R-MSE (Chapter 3, see Table D.1) and for the non-embedded time series. Even though differences are discernible, the mentioned overall conclusion given by Westerhold et al. [328] also holds for these cases.

6.4 Conclusion

A novel type of powerspectrum, the *inter spike spectrum*, has been proposed. The method decomposes any arbitrary signal into basis functions which consist of (lagged) Dirac combs (DC) of different inter-spike period. The loading for each period is obtained by a regularized regression, which promotes sparsity in its solution (we chose LASSO). Since there are $M = \sum_{i=1}^{N} i$ basis functions for a signal of length N the regression can get computationally intensive for $N > 1,000$. When plotting the computed loadings as a function of the frequency (or period) the inter spike spectrum is obtained. This novel powerspectrum is superior to an ordinary FFT-based powerspectrum, when the signal has a spike-train-like appearance. The invention of the proposed method has been motivated by the idea of transforming τ-recurrence rate signals (τ-RR's) into their frequency domain. This general idea [338] allows for a frequency analysis of high dimensional systems, because the RP is a representation of the system's state space trajectory. The τ-RR of a recurrence plot (RP) usually has a spiky shape, especially for map-like data, and the inter spike spectrum can reliably reveal the system's dominant frequencies, which is not possible when

Fourier transforming the τ-RR or the underlying signal itself. Since the position of the peaks in the τ-RR are not sensitive to noise, the corresponding inter spike spectrum also yields robust results in the presence of noise. We have successfully used the idea of transforming the τ-RR for the detection of bifurcations in the Logistic map. By constructing appropriate surrogates of the inter spike spectra, and thus a null model, the number of significant peaks in the inter spike spectrum correlated well with the positive Lyapunov exponent. This measure was also able to resolve period-doubling bifurcations. We could further demonstrate that characteristic frequencies in the CENOGRID dataset revealed by FFT-spectrograms are also apparent in the corresponding inter spike spectrograms of the τ-RR of the signals.

We could think of a broad range of applications of the proposed idea. The inter spike spectrum itself can serve as a valuable tool for the analysis of any sort of spike-train-like data. On the other hand, the inter spike spectrum of the τ-RR of a signal can serve as a generalized, nonlinear frequency analysis tool for complex systems. Uncovering the main frequencies of high-dimensional systems could be particularly advantageous for chaos control applications in an engineering context. When there is only a subset of state variables available, the state space has to be reconstructed as a pre-processing step. Recent findings [162] show that this reconstruction process can be reliably automated and applied to multivariate data as well. This would allow for a "running window" approach, in order to detect transitions. Due to the mentioned computational constraints of our proposed method, a window size $w \leq 1,000$ would possibly suffice for most data, especially when it is map-like, i.e., not highly sampled.

Part III

Application in climate science

Chapter 7

Classifying past climate change in the Chew Bahir basin, southern Ethiopia, using recurrence quantification analysis

Abstract

The Chew Bahir Drilling Project (CBDP) aims to test possible linkages between climate and evolution in Africa through the analysis of sediment cores that have recorded environmental changes in the Chew Bahir basin. In this statistical project we consider the Chew Bahir paleolake to be a dynamical system consisting of interactions between its different components, such as the waterbody, the sediment beneath lake, and the organisms living within and around the lake. Recurrence is a common feature of such dynamical systems, with recurring patterns in the state of the system reflecting typical influences. Identifying and defining these influences contributes significantly to our understanding of the dynamics of the system. Different recurring changes in precipitation, evaporation, and wind speed in the Chew Bahir basin could result in similar (but not identical) conditions in the lake (e.g., depth and area of the lake, alkalinity and salinity of the lake water, species assemblages in the water body, and diagenesis in the sediments). Recurrence plots (RPs) are graphic displays of such recurring states within a system. Measures of complexity were subsequently introduced to complement the visual inspection of recurrence plots, and provide quantitative descriptions for use in recurrence quantification analysis (RQA). We present and discuss herein results from an RQA on the environmental record from six short (< 17 m) sediment cores collected during the CBDP, spanning the last 45 kyrs. The different types of variability and transitions in these records were classified to improve our understanding of the response of the biosphere to climate change, and especially the response of humans in the area.

7.1 Introduction

The development by humans of settlement systems, artistic representations, and hunting strategies during the last 50–40 kyrs marks a "human revolution" and the emergence of behaviorally modern humans in Africa ([225, 255]). These behaviorally modern humans may have adapted to environmental changes and extreme climatic oscillations through technological, behavioral, cultural and cognitive innovation rather than through physical adaptation ([56, 158, 159]). Determining the nature and pace of changes in the environment of early modern humans is crucial to understanding the factors that influenced this human revolution. For example, different types of climate variability would have resulted in different types of climatic stress and changes to environmental boundaries ([92, 128, 318]).

There are currently ongoing discussions concerning global and regional climate fluctuations that had an effect on human habitats (e.g. [92, 107, 137, 175, 195, 306, 309]) and which did not, either because these fluctuations had little or no effect on the African climate (e.g. [301]) or because their effect was buffered by the environmental system (e.g. [61]). Among the most debated episodes of African climate (including their onset, termination, and internal variability) are the Dansgaard-Oeschger (DO) cycles and Heinrich Events (HEs) (~110–12 kyr BP, e.g. [21, 29, 107, 175, 299, 301]), the marine isotope stage 4 (MIS 4) aridification (~71 kyr BP, e.g. [301]), the Last Glacial Maximum (LGM, 23.5–18 kyr BP, e.g. [109, 280, 299]), and the African Humid Period, 15–5 kyr BP, e.g. [64, 282, 299, 300]). Such global (DO cycles, HEs, LGM) and regional (AHP) episodes may have affected the availability of water and food, spatial retreats and shelter, and migration corridors, over variable periods of time (e.g. [11, 25, 42, 46, 89, 92, 200]).

Time-series analysis provides a number of tools with which to characterize past climate change, which can be random, clustered, cyclic, or chaotic (e.g. [209, 226, 305]). The most popular methods for characterizing variations are based on Fourier or wavelet transforms, decomposing time series into a linear combination of sinusoids (e.g. [305] and references therein). Past climate change is, however, often nonlinear (i.e. there is no simple proportional relation between cause and effect) and techniques to describe nonlinear behavior have therefore become increasingly popular in recent decades, defining the scaling laws and fractal dimensions of natural processes ([146, 313, 314]) and detecting nonlinear interrelationships using methods such as transfer entropies, graphic models, and recurrence plots ([32, 88, 113, 212, 252, 258, 263, 264, 336]). The availability of long time series in the Earth sciences in recent times (for example from multi-sensor core logger and micro X-ray fluorescence results) facilitates the use of these methods and increases the reliability of the results obtained.

In this chapter we present a classification of past climate variability in the Chew Bahir basin of southern Ethiopia over approximately the last 45 kyrs using recurrence plots, which provide a graphic display of recurring states in the environmental system ([77, 209]). Quantitative descriptions (measures of complexity) have been developed to complement visual inspection of recurrence plots (RPs) and for recurrence quantification analysis (RQA) (e.g. [202, 209, 336]). We previously used RPs to identify past climate transitions during the Plio-Pleistocene in Africa, the Middle East, and East Asia (e.g. [72, 81]). Such plots enable us to detect nonlinear patterns in past climate change, helping to improve our understanding of the underlying process of climate transitions in the Chew Bahir basin by statistically describing the dynamical characteristics of the environment ([72, 81, 207, 209, 210]). We first used the method on prototypical data in order to assess its performance with typical paleoclimate transitions. We then performed an RQA on the Late Quaternary climate record from Chew Bahir cores CB01–06 because it provides one
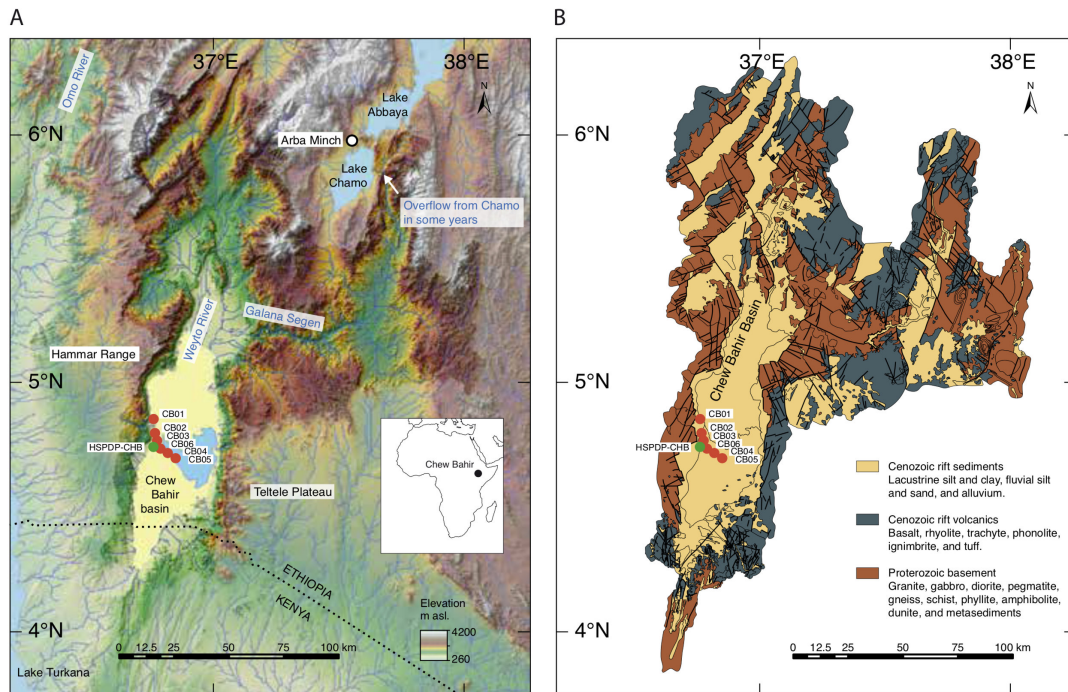
Figure 7.1: A Topographic map of the Chew Bahir basin, showing the outline of the catchment, the drainage network, the locations of the short cores in the pilot study (2009, 2010), and the 2014 HSPDP-CHB drill site. **B** Geologic map of the Chew Bahir basin, showing the three generalized rock types: Cenozoic rift sediments, Cenozoic rift volcanics, and Proterozoic basement. Compilation based on Omo River Project Map [63], Geology of the Sabarei Area, Geology of the Yabello Area, and Geology of the Agere Maryam Area. Maps are modified versions of the ones previously published in Trauth et al. [306] and Foerster et al. [91]

of the most detailed records of climate change available from the vicinity of important sites for modern human fossil ([91–93, 305, 306]).

7.2 Regional setting

The sediment cores described herein were recovered from the Chew Bahir basin in the southern Ethiopian Rift (4.1–6.3°N, 36.5–38.1°E; Figs. 7.1, 8.1) (see details in [93]). Chew Bahir is a closed basin, separated from the Turkana Basin to the west by the Hammar Range. The western part of the 32,400 km^2 catchment, drained by the perennial Segen and Weyto rivers, is mostly formed by Late Proterozoic gneisses, whereas the eastern part is dominated by Miocene basalts [63, 224]. Rainfall in the area is associated with the seasonal migration of the Intertropical Convergence Zone (ITCZ), resulting in two rainy seasons in March–May and October–November [232]. Rainfall intensity strongly depends on Atlantic and Indian Ocean sea-surface temperature (SST) variations caused by the Indian Ocean Dipole (IOD) and the El Niño-Southern Oscillation (ENSO), also explaining the recent reduction of rainfall intensities over the last decades [52, 232, 265, 278, 279].

7.3 Methods

7.3.1 The Late Pleistocene–Holocene record of the Chew Bahir basin

We reconstructed climate fluctuations in the Chew Bahir from six short ($<$17 m) sediment cores (CB01–06), collected along a ~20 km transect from Northwest to Southeast in the basin (Fig. 7.1). These six cores were collected in a pilot study for the Hominin Sites and Paleolakes Drilling Project (HSPDP; HSPDP-CHB deep coring site in Fig. 7.1) [37, 57] and the Collaborative Research Center CRC-806 at the University of Cologne, and were described in detail in earlier publications (e.g. [92, 93, 306]). We used the potassium (K) concentrations of the sediment, determined by micro X-ray fluorescence (μXRF) scanning, as a proxy of aridity [91, 93, 305]. Dynamic time warping (DTW) was employed for aligning the K records from the six cores CB01–06 [306]. The composite age model of Trauth et al. [305], based on 32 AMS ^{14}C ages derived from biogenic carbonate, fossilized charcoal and organic sediment, resulting in a very solid chronology for lake record spanning the last ~45 kyr, was then used to convert composite core depths into ages.

A principal component analysis (PCA) helped us to separate the mixed regional and local environmental signals in the K records from the six aligned cores. The first principal component (PC1) contains more than 94% of the variance of the data and was therefore interpreted to best represent regional climate. The temporal resolution of the climate proxy record in CB01, with 2,812 original measurements, has a calculated mean spacing of ~16 yrs, ranging from ~4 yrs in the upper part of the core to almost 2 kyrs in the deeper part of the core [92, 93, 305]. The K record (following DTW alignment of cores CB01 to CB06 and PCA-based unmixing) runs from 45.358 to 0 kyr BP with a mean resolution of 8 yrs (ranging from 2.6 to 30.6 yrs) and the record has therefore been interpolated to an evenly-spaced time axis running from 45.358 to 0 kyr BP at 10 year intervals, which is close to the mean intervals of the original data (~16 yrs, ranging from ~4 yrs to 2 kyrs) and in the aligned and unmixed data (~8 yrs, ranging from 2.6 to 30.6 yrs) [306].

7.3.2 Principles of recurrence plots (RPs) and recurrence quantification analysis (RQA)

We considered the Chew Bahir paleolake to represent a complex system of interacting components, such as the waterbody, the sediment beneath the lake, and the organisms living within and around the lake. Systems with properties that change over time, such as the Chew Bahir paleolake with its slowly changing geomorphologic features, much more rapidly varying climatic factors, and the possible (very recent) influence of human activities, are known as dynamical systems. The Chew Bahir multi-dimensional paleolake system has been affected by a number of factors (identified as state variables of the system) such as precipitation (with higher rainfall resulting in increased weathering and erosion within the catchment, and hence more potassium washed into the lake), evaporation (with increased evaporation producing more extreme hydrochemical conditions that enhance potassium fixation in the sediment through authigenic clay-mineral alterations, [91, 93]) and wind speed (with higher wind speeds and reduced vegetation cover resulting in more potassium-rich particles being blown into the lake).

Analysis of temporal variations in the true, but unknown, state variables of the Chew Bahir paleolake would require a record of the variables influencing those state variables over a relevant time period. Since the state variables of the Chew Bahir system (e.g., the hypothesized precipitation, evaporation and wind speed) cannot be measured directly, we used indirect indicators (known as climate proxies) measured from natural archives of environmental change such as the sediments of the Chew Bahir basin, sampled by coring. A proxy record of a multi-dimensional

dynamical system obtained by sampling a single variable is equivalent to projecting the dynamics of a complex system onto a single axis [138] (c.f. Chapter 2). In our case the sampled variable is the series $s(t)$ of potassium concentrations s along a lake sediment core, which provides a natural archive of past influences on the Chew Bahir lake system, with sediment depth d converted into time t using the age model from Trauth et al. [305].

We used time-delay embedding to reconstruct the state space from the observable $s(t)$, see Section 1.1 and Fig. 1.2 as well as Section 2.1 for a more detailed view on the topic. Since the potassium concentration in the Chew Bahir sediments has been the result of a complex interplay between unknown amplitudes of different environmental (or state-) variables (such as precipitation, evaporation, and wind velocity), an analysis of the temporal variations in this environmental proxy will help us to understand the state variables of the Chew Bahir paleolake and the time-varying interactions between its different components.

A common feature of dynamical systems is the property of recurrence [324]. Recurring patterns in the state of a system are a reflection of typical characteristics of the dynamical system. Defining these patterns can contribute significantly to our understanding of the system's dynamics. Changes in environmental (or state-) variables (such as precipitation, evaporation, and wind velocity) often follow characteristic courses (represented as trajectories $\vec{v}(t)$ in phase space, Fig. 1.2B+C) that could lead to similar (but not identical) lake characteristics (e.g. depth and area of the lake, alkalinity and salinity of the lake, species assemblage in the waterbody, or formation of authigenic minerals in the sediment). Recurrence plots (RPs), see Chapter 1 and Appendix A, are graphic displays of such recurring states within a system have been shown to be successfully applied on even non-stationary paleoclimate data [72, 328] (Figs. 1.2, 7.2).

RP analysis is complicated by trends in the mean and variance (i.e. nonstationarity and long-term variations), which have a marked effect on the recurrence rate (RR), i.e. on the density of dots in RPs, Eq. (A.4) (Appendix A). If the long-scale variations are not of interest, the RQA characteristics can be used to reveal undesired fluctuations that do not reflect the more rapid variations in the dynamics of the system. We therefore use an adaptive recurrence threshold instead of a fixed value, which is determined in such a way that all points in the state space have the same number of neighbors (fixed amount of neighbors). Thus, the RR is globally and locally fixed. This means, we expect a time-dependent behavior of RR, when performing a windowed analysis over the entire RP and a constant RR with respect to the whole RP (see description below). The corresponding more homogeneous RPs make it easier to detect rapid changes of the system dynamics at a particular point, while long-term changes can be observed at other points. Another possible problem in analyzing paleoclimate records, which typically have very low signal-to-noise ratios, is the disruption of patterns in the RP (such as diagonal lines) by noise, gaps and other disturbances.

The RQA can be carried out using moving windows in order to detect changes in the system dynamics, represented by a change in the RQA measures, (1) by developing a single RP and calculating the RQA measures for windows moving along this RP, or (2) by developing multiple RPs of individual windows from the time series and calculating the RQA measures for these RPs. If nonstationarities (e.g. trends) are not the main focus of the analysis, then approach (2) makes it easier to find transitions while ignoring any nonstationarities. However, if the detection of overall changes is of interest (e.g. to test for nonstationarity), the recurrence conditions should be kept constant over time (thus taking into account the RP of the entire time series) and approach (1) will be more appropriate [203]. A window size needs to be chosen that is small enough to ensure good temporal resolution but large enough to cover typical variations (e.g. the number of

cycles) in order to be able to detect recurrences. Since our data shows a very dominant long-term trend, we first performed the RQA on the original data and then on a high-pass filtered (and hence detrended) version of the data.

7.3.3   Synthetic examples of a recurrence quantification analysis

The investigation of synthetic data using RPs and the RQA measures described above and in Appendix A, which are then used to analyze real data, has proven to be particularly advantageous when the methods are complex and the results not immediately obvious (e.g. [209, 305] (Fig. 7.2). The use of conceptual models that mimic typical system changes helps us to understand the typical changes seen in RPs and to assign them to one or other of those changes. The first example investigated was of normally-distributed (Gaussian) noise, for which an RP and RQA measures were derived (Fig. 7.2A). The RP shows randomly distributed points, each representing times when the system randomly returned to a similar state. Similar states frequently recur in random noise but without any regularity except the states represented by the main diagonal line, the line of identity. The RR is therefore more or less constant with very low values. Since there is no systematic pattern (e.g. cyclically recurring states), the RP does not show any linear patterns and the *DET* is therefore very low.

The second example investigated was a composite signal of two sine waves, for which an RP and RQA measures were again derived. In the interval between $-1,400$ and -600 there is a positive trend in the mean (Fig. 7.2B). The RP shows long diagonal lines, diagnostic of cyclicity in time series, with shorter lines in between. The horizontal distances between these lines corresponds to the periods of the two sine waves ($T_1 = 50$ and $T_2 = 200$). Since the higher frequency is a harmonic of the lower frequency, the corresponding diagonal lines appear thicker in the RP. The RP clearly shows the effect of the trend in the mean as the diagonal lines disappear towards the upper left and lower right corners of the RP due to the trend-induced increase in the distance, leaving a blocky pattern in the middle of the RP. Because the dynamic itself does not change, neither do the *DET* values; instead they persist at their maximum values, which are unaffected by the trend. The RR is however affected, as shown by the lower density of black dots in the RP, indicating cyclic variations with a period of the same order as the dominant period in the signal ($T_2 = 200$).

The synthetic data in the third example comprised a sine wave and Gaussian noise with a signal-to-noise ratio that decreases from left to right (Fig. 7.2C). As a result the continuity of the diagonal lines decreases to the right, as do both the *DET* and the *RR* values. The synthetic data in the fourth example investigated comprised a composite signal from two sine waves with distinct trends in their frequencies (Fig. 7.2D). The distances between diagonal lines decrease as a result of increasing signal compression with time. The convergence of the lines and their degree of curvature depend on the function describing the signal compression. The increase in frequency with time results in reductions in the *DET* because the higher-frequency cycles seen on the right of the plot are no longer adequately resolved. Recurrence points in between diagonal lines are scarce due to the fact that the time delay τ chosen for the plot no longer suits the increased frequency. A higher sampling rate eliminates this phenomenon.

The results of the fifth example revealed how RPs and RQAs respond to an abrupt transition from a composite signal consisting of two sine waves ($T_1 = 300$, $T_2 = 50$) to a signal with only one sine wave ($T_3 = 60$) (Fig. 7.2E). As before, the oscillations of the signal produce diagonal lines in the RP, with horizontal distances between the lines corresponding to the periods of the signals. Two sets of diagonal lines in the lower-left corner of the RP correspond to the two frequencies
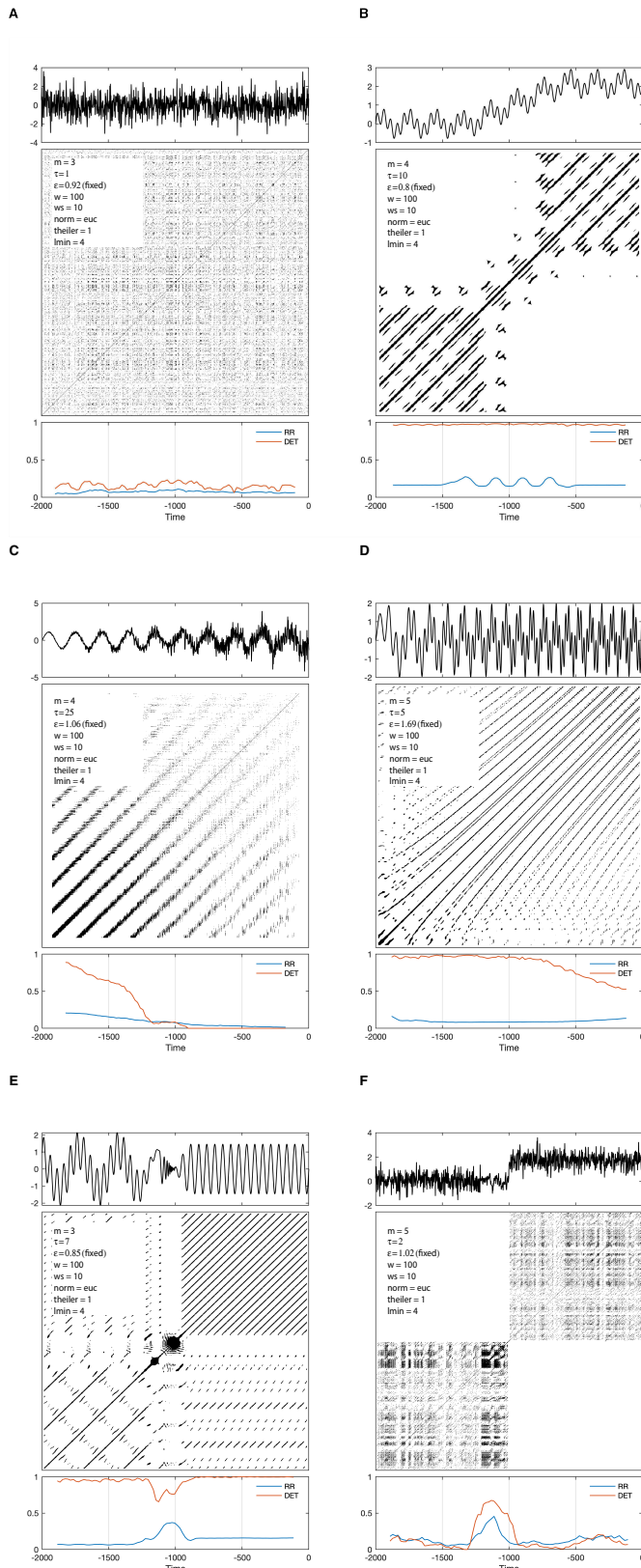
Figure 7.2: Recurrence plots (RPs) and recurrence quantification analysis (RQA) measures for synthetic data representing common types of dynamic behavior: **A** normally-distributed (Gaussian) noise. **B** Composite signal comprising two sine waves and a positive trend in the mean. **C** Composite signal comprising a sine wave and Gaussian noise with decreasing signal-to-noise ratio from left to right. **D** Composite signal comprising two sine waves and a trend in the frequencies. **E** Abrupt transition from a composite signal comprising two sine waves to a signal with only one sine wave. **F** Normally-distributed (Gaussian) noise with a stepwise transition in the mean and a change in the auto-correlation prior to this transition. The examples display the time series (upper panel), the RP (middle panel) and the RQA measures (lower panel). Parameter abbreviations are m = embedding dimension, τ = time delay, ε = recurrence threshold, w = window size, ws = window moving steps, *norm* = vector norm, *thei* = size of Theiler window, *lmin* = minimum line length, RQA measures RR = recurrence rate and DET = determinism.

of the sine waves, while there is only one set of diagonals in the upper right corner. During the transition we also let the amplitude of the signal decrease. Both the change from two sine waves to a single sine wave and the decrease in variance during the transition are clearly visible in the RP, as well as in the RQA measures. The single period oscillation has more recurrences than the two-period signal and therefore a higher RR, peaking at the transition because the lower signal variance produces a blocky pattern in the RP.

Because of the two different time scales of the two-period signal it has a more complicated phase-space trajectory on the left-hand side of the RP than the one-period signal. Trajectory segments are therefore only parallel at particular times, resulting in long diagonals for the long period signal (because T_2 is a harmonic of T_1, as in the second example above), but interrupted diagonals (and lines that are curved at times) for the short period signal. This results in slightly lower DET values before the transition to a one-period signal than after it. An effect similar to that seen in our third example is observed in the interval between $t = -1200$ and $t = -1050$ (where the actual transition from a two period signal to a one period signal occurs), which is why the DET decreases here before again increasing due to the blocky pattern in the RP, which also causes the RR to increase, as mentioned previously.

The synthetic data in the sixth example investigated was of Gaussian noise with a stepwise transition in the mean and a change in the auto-correlation prior to this transition (Fig. 7.2F). It is important to note that neither the mean nor the variances change in the pre-transition section, so this change cannot be recognized using conventional methods. The two blocky features in the RP before and after the transition look very similar to those in the RP of the first example above. However, the interval between $-1,200$ and $-1,000$ clearly shows distinctive patterns within the RP that are different from those typically occurring in RPs of pure noise. White vertical lines help to define blocky features that represent episodes with different dynamics. The RQA characteristics look similar to those in the Gaussian noise example except for a section with higher auto-correlation, which is reflected in the higher density of black dots, clear diagonal lines, and a change in the dynamics of the system. This fact can be used to detect such a change in auto-correlation as a precursor to a tipping point and to ultimately predict the tipping point itself.

7.4   Results

We have used recurrence plots (RPs), complemented by a recurrence quantification analysis (RQA), to characterize past climate change in the Chew Bahir basin over approximately the last 45 kyrs. We selected the RR and DET measures because they describe fundamental properties of the complex Chew Bahir system dynamics, while keeping the theoretical complexity within reasonable limits to facilitate interpretations. The RP and RQA approach was applied to the record of K concentrations in the sediment cores (following DTW-based alignment and linear unmixing using a PCA), which has previously been shown to be a reliable proxy for aridity in the Chew Bahir basin [91, 93, 305, 306]. The K record was embedded in state spaces with dimensions varying from $m = 5$ to $m = 6$ and time delays varying from $\tau = 3$ to $\tau = 10$ data points, equivalent to 3×10 yrs = 30 yrs and 10×10 yrs = 100 yrs, where 10 yrs is the resolution of the time series following interpolation. These choices of the embedding parameters were the results of the corresponding first minimum of the auto-mutual information (for the time lag τ) [95] and a zero (or global minimum) of the false nearest neighbors statistic (for the embedding dimension m) [154]. The RPs were calculated using a Euclidean norm, a Theiler window of thei=1
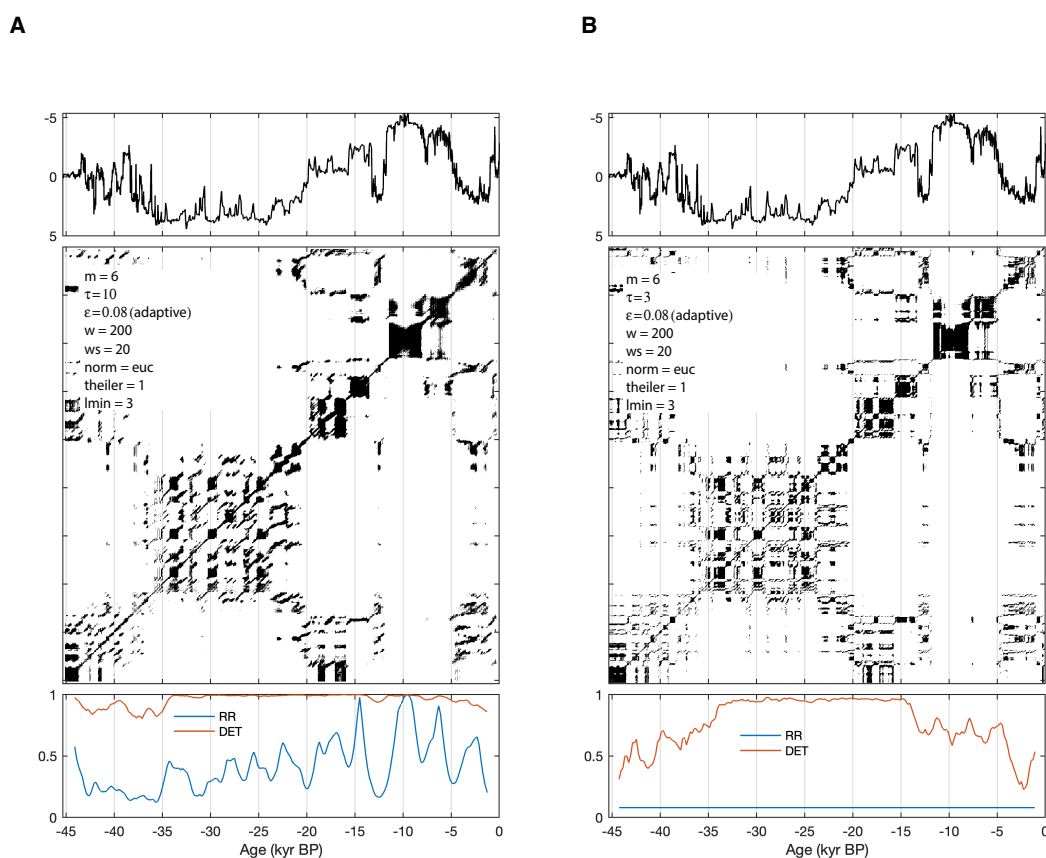
Figure 7.3: Recurrence plot (RP) and recurrence quantification analysis (RQA) measures of the complete record ($-45,358$ to 0 yrs BP) from the Chew Bahir basin: time series (upper panel), the RP (middle panel) and the RQA measures (lower panel) of moving windows determined either by **A** calculating the RQA measures for windows moving along a single (global) RP and **B** by calculating individual RPs for windows moving along the entire time series. See previous figure for the meaning of the abbreviations

data point, a minimum length of lmin=3 for the lines used to calculate *DET*, and an adaptive recurrence threshold value (FAN) set to a corresponding *RR* of 8%. The size w and the step size ws of the moving window depend on the period of time investigated and the number of data points contained therein.

In order to compare different climatic conditions we first looked at the RP of the complete, unfiltered time series documenting the long-term variations in the Chew Bahir system (Fig. 7.3). This RP reveals a clear division of the time series into sections of different lengths, indicated by square blocky features in the RP separated by white vertical lines (see also Figs. 7.2B, E and F). The first cluster of recurrence points occurs between 45.35 and 37 kyr BP, comprising both connected and isolated points. This interval is characterized by both vertical and horizontal lines, representing episodes of stability (both wet and dry) interrupted by a series of extremely wet events, indicated by white vertical lines in the RP. We observe low but gradually increasing *DET* values in this episode, suggesting increasing predictability in the system.

A second, very obvious, cluster of recurrence points occurs between 37 and 20 kyr BP, which includes the time intervals in which the Dansgaard-Oeschger (DO) cycles (~110–23 kyr BP), the
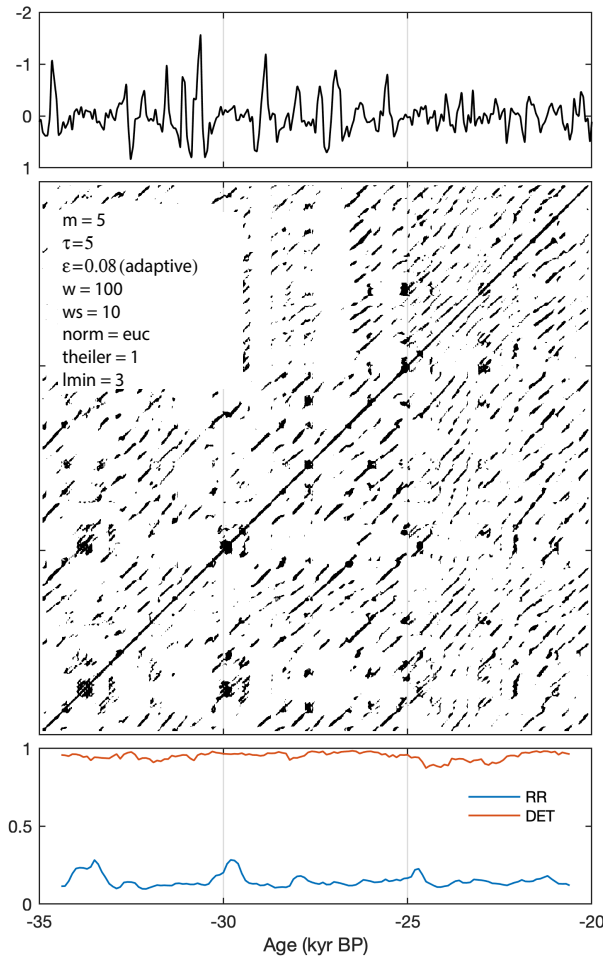
Figure 7.4: Recurrence plot (RP) and recurrence quantification analysis (RQA) measures for the Chew Bahir basin covering of the time interval between 35 and 20 kyr BP: time series (upper panel), the RP (middle panel) and the RQA measures (lower panel) of moving windows determined by calculating individual RPs for windows moving along the entire time series. See previous figure for the meaning of the abbreviations.

Heinrich Events (HE, ~60–12 kyr BP) and the Last Glacial Maximum (LGM, 23.5–18 kyr BP) affected the climate further to the north. The RQA reveals consistently high *DET* values until about 15 kyr BP, exceeding those of the previously described cluster and suggesting a much higher predictability in the system during that time. The recurrence plot for the time interval from 20 to 0 kyr BP includes the African Humid Period (AHP, ~15–5 kyr BP). This interval is characterized by three large clusters of recurrence points, interrupted by both white vertical and horizontal lines, together with fluctuating RR values and a long-term trend towards lower *DET* values. The white vertical lines again help to define blocky features marking episodes with different dynamics.

To analyze the dynamics of these individual sections, the time series was high-pass filtered with a cutoff frequency of 0.001 yrs^{-1} in order to remove any long-term trends (Figs. 7.4, 7.5, 7.6). An RP was constructed and an RQA performed using a sliding window ($w = 100$, $ws = 10$) over the entire RP, assuming stationarity within the intervals. The interval between 35 and 20 kyr BP is remarkable for numerous short, slightly curved diagonal lines, suggesting a cyclic recurrence of wetter episodes in the Chew Bahir basin within a period that had a generally dry climate (Fig. 7.4). The oscillating climatic conditions are reflected in higher *DET* values, while the low *RR* values suggest a low probability of recurring states occurring within the system over
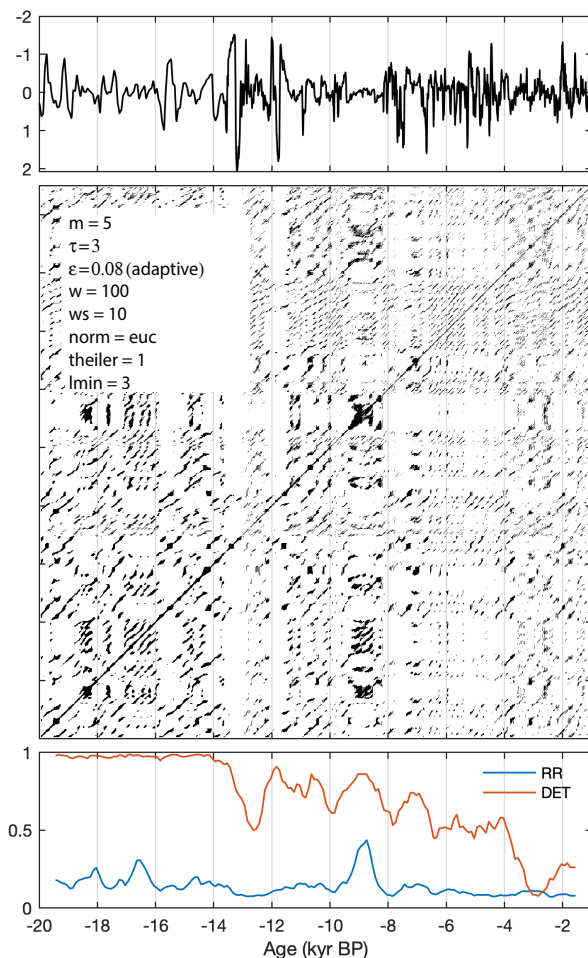
Figure 7.5: Recurrence plot (RP) and recurrence quantification analysis (RQA) measures for the Chew Bahir basin covering of the time interval between 20 and 1 kyr BP: time series (upper panel), the RP (middle panel) and the RQA measures (lower panel) of moving windows determined by calculating individual RPs for windows moving along the entire time series. See previous figure for the meaning of the abbreviations.

a particular time period [209].

The RP for the time interval from 20 to 1 kyr BP includes the AHP (~15–5 kyr BP) (Fig. 7.5). As before, having used moving windows with a length of $w = 100$ (1,000 yrs) and a step size of $ws = 10$ (100 yrs) the obtained RP reveals a series of blocky recurrence point patterns, interrupted by a series of white vertical lines. These patterns suggest distinct episodes of relative stability, both wet and dry, separated by abrupt transitions at ~13.2 kyr BP, ~11.8 kyr BP, ~7.5 kyr BP, ~5.2 kyr BP, and ~4.5 kyr BP. The overall appearance of the RP (and of the time series itself) reflects dynamics characterized by a period of higher variability followed by a period of low variability between 9.5 and 8 kyr BP, and dynamics dominated by a high-frequency cyclicity between 8 and 1 kyr BP, which is roughly similar to the fifth synthetic example above (see Fig. 7.2E). In addition to these patterns, the RP for the interval from 20 to 1 kyr BP also shows numerous short diagonal lines, suggesting a weak cyclicity. The diagonal lines, however, are very different from each other in width and in the number present. The RR values are very low except for an interval of relative stability between 10 and 8 kyr BP, which has a high probability of recurring states. The DET values document a general trend towards lower predictability in the system dynamics, but this decline exhibits a very complicated and somewhat cyclical pattern,
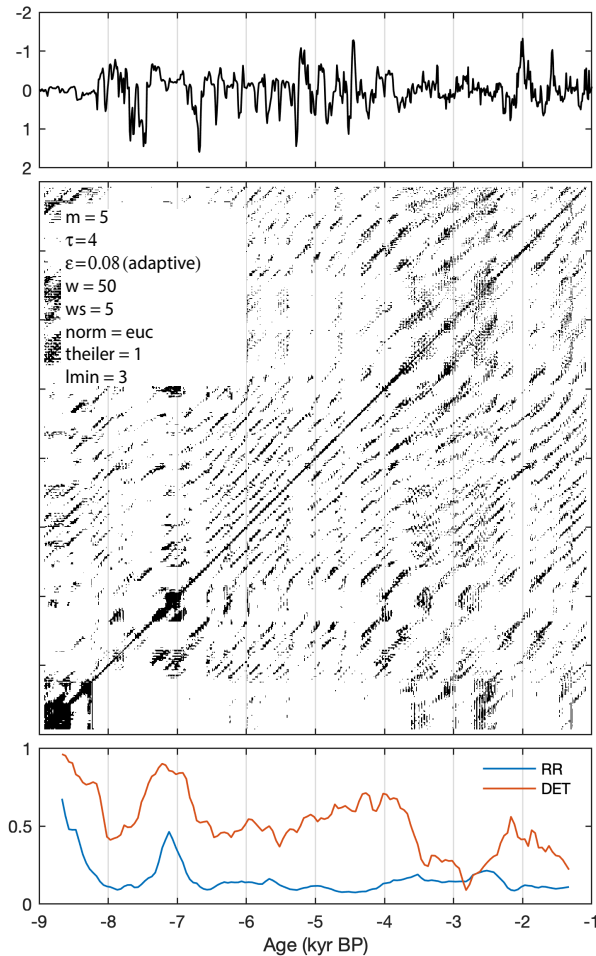
Figure 7.6: Recurrence plot (RP) and recurrence quantification analysis (RQA) measures for the Chew Bahir basin of the time interval between 9 and 1 kyr BP: time series (upper panel), the recurrence plot (middle panel) and the RQA measures (lower panel) of moving windows determined by calculating individual RPs for windows moving along the time series. See previous figure for the meaning of the abbreviations.

rather than a simple linear trend.

In contrast to the RP for the first half of the time interval from 20 to 1 kyr BP, the RP for the period between 8 and 1 kyr BP contains numerous spotty diagonals following two blocky features at about 9 and 7 kyr BP (Fig. 7.6). These blocky features indicate a slowing-down of the system dynamics and therefore a higher predictability, as also indicated by high DET values (similar to the synthetic example shown in Figs. 7.2E+F). The RR values are very low, except for one interval of relative stability. The distinctive diagonal lines after about 6.8 kyr BP suggest a marked cyclic recurrence of droughts approximately every 100–150 yrs, which lasted until stable conditions returned following the termination of the AHP. The relatively low DET values ($\sim$0.5) and therefore low predictability during this interval, however, reflects the discontinuity in the diagonal lines and the dispersion of cycles (i.e. the variability in distances between the diagonal lines, similar to the synthetic example shown in Figs. 7.2C+D), suggesting that a stochastic process is superimposed on the cyclicity. The DET values during the wet-dry transition at the end of the AHP remain moderately high in the high-pass filtered time series until stable dry conditions are established. The interval between 4 and 2 kyr BP with very low DET values (< 0.5) reflects a predominantly stochastic process.

Age (kyr)	45–37	37–20	20–16	16–10	10–8	8–4	4–now
Main climate	intermediate	dry	intermediate	wet, with YD dry event	wet	wet	dry
Environmental conditions	wet-dry trend	1500 yrs cycles	Long-term sinusoidal dry-wet-dry trend due to 20 kyr precession cycle			stable with 160 yrs cycles	stable, wet at the end
RP appearance	irregular	regular	regular	irregular	regular	regular	regular
Recurrence rate (RR)	low	intermediate	intermediate	high, with low YD event	high	high	intermediate
Determinism (DET)	low	high	high	intermediate	intermediate	intermediate	low
Extreme events	random wet	wet-dry	none	YD dry event	possibly 8.2 kyr dry event	approx. 14 dry events, 20–80 yrs long	very arid, possibly wet a er 2 kyr BP
Human habitat	extreme radom wet events	wet-dry cycles	stable	threshold	stable, except for 8.2 kyr event	threshold and cycles	aridity

Table 7.1: Compilation of the most important time periods in the Chew Bahir sediment cores, classified according to main climate, environmental conditions, recurrence plot appearance, recurrence rate *RR*, determinism *DET*, occurrence of extreme events and human habitat

7.5 Discussion

A section-by-section analysis of the RPs of the time series together with an examination of the temporal course of the RQA measures allowed us to identify and eventually classify different types of variability and transitions (Table 7.1). The classification of variability and transitions can help to improve our understanding of the response of the biosphere (including humans) to climate changes (e.g. [72, 91, 92, 305, 306]). There is a general consensus among anthropologists that both long-term trends and severe, abrupt changes resulted in significant changes to the social and socio-economic behavior of early humans [56, 89, 110, 222]. The response to a changing habitat, both subtle or dramatic, accompanied by changes in essential resources such as food and water, could be either expansion, decease, migration, or adaption. In this process, the nature of the adaptation is a function of the adaptability and the time scale (e.g. [91, 92]. Of particular interest is how the adaptability of humans has enabled them to deal with such diverse and profound environmental changes since 45 kyrs BP through behavioral changes, and what level of environmental change met the limit of resilience (e.g. [91, 92, 305, 306]). Furthermore, the different types of variability and transitions, and the corresponding response of the biosphere (including humans), will help to detect similar types of changes in the long (~280 m) sediment cores recently collected in the Chew Bahir basin within the Hominin Sites and Paleolakes Drilling Project (HSPDP) and to investigate whether or not these types are typical for and exclusive to the basin, as well as the response of the biosphere to these changes (see Chapter 8).

Our analysis clearly showed a number of different types of variability in the K record, separated by either gradual or rapid transitions. The first type of variability occurs between the beginning of the record (45 kyr BP) and about 35 kyr BP. Within this interval we observed a relatively low but gradually increasing predictability during times of relative stability (both wet and dry), interrupted by a series of extremely wet events. There is no apparent cyclicity in this interval but rather an irregular pattern of different types of variability. Both the extreme events and the rapid transitions between episodes of relative (wet and dry) stability will certainly have had an impact on humans in the area, leaving them with a range of possible responses (adapt, migrate,

starve) to a dramatically changing environment.

The second type of variability occurs between 35 and 20 kyr BP, with slightly different dynamics before and after 25 kyr BP. This interval is characterized by a millennial-scale climate variations during the last glacial cycle, which includes the time intervals in which the Dansgaard-Oeschger (DO) cycles (~110–23 kyr BP), the Heinrich Events (HE, ~60–12 kyr BP) and the Last Glacial Maximum (LGM, 23.5–18 kyr BP) affected the climate further to the north. This millennial-scale climate fluctuations are cyclic with minor variations in the period, as indicated by the slight curvature of the diagonals in the RP (see synthetic example shown in Fig. 7.2D), although this curvature could also suggest inaccuracies in the age model, rather than real variations in the cyclicity. The pronounced cyclicity is reflected in the RP and in the RQA measures by a very high predictability, but on time scales that are certainly not relevant to humans, because it is orders of magnitude longer than the lifetime of individual humans. However, the transitions between the dry and wet episodes were very rapid which has probably caused significant stress to human populations.

The third type of variability occurs between 16 and 10 kyr BP, including parts of the AHP (~15–5 kyr BP). During this interval we found episodes of relative stability, both wet and dry, separated by abrupt transitions at ~13.2 kyr BP and ~11.8 kyr BP. There is also evidence of a weak cyclicity with a general (but complicated) trend towards lower predictability. The onset of the AHP in the Chew Bahir area was relatively rapid (covering ~240 yrs, [306], which is in agreement with similar records from elsewhere, as a result of large-scale deglacial forcings (i.e. changes of Atlantic sea-surface temperatures causing meridional shifts of the African easterly jet and the monsoon belt, and changes in the atmospheric concentration of greenhouse gases causing changes in atmospheric temperatures, [282]). The onset and termination of the dry episode during the Younger Dryas (YD) were also rapid transitions (over less than 100 yrs, [306]), again very similar to other sites in N and NE Africa.

The climate variability within the AHP and the long-term transition that it represents clearly affected human communities and has fueled massive changes in the population size and structure such as the profound socio-economic transformations that have been documented for N and NE Africa (e.g. [26, 56, 200, 222]). A well-studied example is provided by demographic reconstructions that have been made for the Saharan Holocene. These reconstructions show a temporal delay between the onset of humid conditions (based on sedimentary dust flux records) and human reoccupation of former desert areas, with associated societal changes seen as a response to the environmental changes [56, 110, 198].

Having adapted to the wet climate of the AHP, humans certainly had to cope with the very rapid transition towards extreme dryness at the onset of the YD dry episode. Highly mobile groups of hunter-gatherers responded to short-term arid pulses by vertical migration as documented by the settlement patterns in what are assumed to have been refuge areas such as the SW Ethiopian Highlands (e.g. [92]). The YD is followed by a short interval of relative stability between 10 and 8 kyr BP, followed by a dry episode centered around 7.5 kyr BP. This event, which is synchronous with the prolonged pause in the Green Sahara conditions 8 kyrs ago (within the uncertainty of our age model, see [305]), coincides with a temporary abandonment of sites previously occupied by Neolithic humans [300].

The fourth type of variability occurs between 8 and 4 kyr BP. This interval includes the transition from the humid climate of the AHP to the subsequent dry climate. The termination of the AHP was a result of weaker, insolation-driven forcing (and hence more complex and time-transgressive responses) than that which produced the DO cycles, the onset of the AHP,

or the onset/termination of the YD dry episode [282]. This was nevertheless compared to those produced by more subtle changes in orbital forcing, but certainly not abrupt compared to human time scales, i.e. human lifespan or a little more, as it continued for approximately 990 years [306]. The termination of the AHP occurred at different times in other areas, suggesting a strong influence of Indian Ocean SSTs on the East African climate [282]. Most of the transition at the end of the AHP is characterized by wet conditions, interrupted by ~14 dry events that have recurred every 160 ± 40 years and lasted 20–80 years [305]. Compared to the low-frequency cyclicity of climate variability during the DO cycles, this type of cyclicity occured on time scales equivalent to a few human generations. In other words, it is very likely (albeit speculative) that people were conscious of these changes and adapted their lifestyles to the consequent changes in water and food availability [56].

An interesting aspect of this variability is the nature of the transitions close to the threshold in the system response, and how the environment switches from one stable mode to another. A rapid change of climate in response to a relatively modest change in forcing appears to be typical of tipping points in complex systems such as the Chew Bahir basin [69, 181]. If this is the case then the 14 dry events at the end of the AHP could represent precursors of an imminent tipping point that would have allowed a prediction of climate change in the Chew Bahir basin at that time. A deeper analysis of our data is however required to understand whether the wet-dry climate transition in the area was due to a saddle-node bifurcation in the structural stability of the climate, or whether it was induced by a stochastic fluctuation. The time interval after the termination of the AHP (< 4 kyr BP) leads into present-day conditions in the Chew Bahir basin.

7.6 Conclusion

We have used a recurrence quantification analysis (RQA) on environmental records from short cores collected during a pilot study within the Chew Bahir basin to characterize the Chew Bahir paleolake as a dynamical system composed of interacting components. The different types of variability and transitions in these records were classified to improve our understanding of the response of the biosphere to climate change, and especially the response of humans in the area. This classification and the corresponding responses of the biosphere will enable the detection of similar types of variability and transitions in the long (~280 m) ICDP core collected in the Chew Bahir basin within the Hominin Sites and Paleolakes Drilling Project (HSPDP), see Chapter 8, and allow us to investigate whether or not these types are typical for and exclusive to the basin.

Code availability

The MATLAB® code to calculate the RPs and to perform the RQA is available at *http://mres.uni-potsdam.de* and on request from the corresponding author.

Chapter 8

Recurring types of variability and transitions in the ~620 kyr record of climate change from the Chew Bahir basin, southern Ethiopia

Abstract

The Chew Bahir Drilling Project (CBDP) aims to test possible linkages between climate and hominin evolution in Africa through the analysis of sediment cores that have recorded environmental changes in the Chew Bahir basin (CHB). In this statistical project we used recurrence plots (RPs) together with a recurrence quantification analysis (RQA) to distinguish two types of variability and transitions in the Chew Bahir aridity record and compare them with the ODP Site 967 wetness index from the eastern Mediterranean. The first type of variability is one of slow variations with cycles of ~20 kyr, reminiscent of the Earth's precession cycle, and subharmonics of this orbital cycle. In addition to these cyclical wet-dry fluctuations in the area, extreme events often occur, i.e. short wet or dry episodes, lasting for several centuries or even millennia, and rapid transitions between these wet and dry episodes. The second type of variability is characterized by relatively low variation on orbital time scales, but significant century-millennium-scale variations with progressively increasing frequencies. Within this type of variability there are extremely fast transitions between dry and wet within a few decades or years, in contrast to those within Type 1 with transitions over several hundreds of years. Type 1 variability probably reflects the influence of precessional forcing in the lower latitudes at times with maximum values of the long (400 kyr) eccentricity cycle of the Earth's orbit around the sun, with the tendency towards extreme events. Type 2 variability seems to be linked with minimum values of this cycle. There does not seem to be a systematic correlation between Type 1 or Type 2 variability with atmospheric CO_2 concentration. The different types of variability and the transitions between those types had important effects on the availability of water, and could have transformed eastern Africa's

environment considerably, which would have had important implications for the shaping of the habitat of H. sapiens and the direct ancestors of this species.

8.1 Introduction

Some hypotheses about the relationship of climate and human evolution suggest that episodes of increased climate variability (e.g. [214, 247, 248]) or prominent transitions (e.g. [320]) may have enhanced rates of speciation, dispersal and technological innovation. Examples on long time scales are the termination of the permanent El Niño/establishment of the modern Walker/Hadley circulation between 3.5-2.0 Ma, possibly linked to the closure of the Indonesian sea way [38, 254, 307]), and the intensification of the Northern Hemisphere Glaciation at 2.75 Ma (e.g., [23, 307]), both being subject to lively discussions during the last four decades [28, 65, 66, 171, 172, 307, 308, 320]). As the most recent example of a major climate shift in the tropics, in particular in Africa, the termination of the African Humid Period (AHP, ~15-5 kyr BP) has also been intensely investigated, in particular the extent to which it was abrupt or gradual ([64, 174, 299, 306]), which is important for potential migration scenarios within and across the Sahara and cultural transformations [174].

Revived by these debates, statistical methods have recently been used to make quantitative statements about the degree of variability and character of transitions. According to their analysis, the most important transition during the long-term trend towards a more arid climate was at ~1.9 Ma, at about the time of the establishment of the modern Walker/Hadley circulation [254, 307], and not, as suggested earlier by deMenocal [65, 66] during the intensification of the Northern Hemisphere Glaciation (INHG). Similarly, the termination of the AHP at ~5 kyr BP was tested for its relative abruptness comparing observed and theoretical probability distributions of paleoclimate time series from multiple locations in and around Africa [299]. According to their analysis, the wet-dry transition occurred within centuries, which agrees with the results of Trauth et al. [306] using a change point analysis to determine a ~880 yr interval within which this important climate shift occurred.

More sophisticated approaches to classifying variability and transitions were used by Trauth et al. [304] with recurrence plots together with a recurrence quantification analysis on six short (< 17 m) sediment cores collected during the Chew Bahir Drilling Project (CBDP) from the Chew Bahir basin (CHB) in southern Ethiopia, reaching back to ~47 kyr BP. Recurrence plots (RPs) are graphic displays of recurring states in the environmental system. Quantitative descriptions (measures of complexity) have been developed to complement visual inspection of RPs and for recurrence quantification analysis (RQA) (see Chapter 1 & Appendix A). Trauth et al. [304] presented and discussed results from such an RQA on the environmental record of the CHB short cores. The different types of variability and transitions in these records were classified to shed light on our understanding of the response of the biosphere to climate change, particularly the response of humans in the area.

One of the most interesting transitions examined with the RP/RQA was once again the termination of the African Humid period [304, 306]. The rapid (~880 yr) change of climate in response to a relatively modest change in orbital forcing appears to be typical of tipping points in complex systems such as the Chew Bahir basin [69, 181]. If this is the case then 14 dry events at the end of the AHP, each of them 20–80 yrs long and recurring every 160 ± 40 yrs as documented in the Chew Bahir cores could represent precursors of an imminent tipping point which, if properly interpreted, would allow predictions to be made of future climate change in

the Chew Bahir basin [304, 306]. Compared to the low-frequency cyclicity of climate variability before and after the termination of the AHP, this type of cyclicity occurs on time scales equivalent to a few human generations. In other words, it is very likely (albeit speculative) that people were conscious of these changes and adapted their lifestyles to the consequent changes in water and food availability. A deeper analysis of our data is however required to understand whether the wet-dry climate transition in the area was due to a saddle-node bifurcation in the structural stability of the climate, or whether it was induced by a stochastic fluctuation.

Here we present a RP/RQA-based analysis of two long (~290 m) cores collected in 2014 in the Chew Bahir basin (4°45'40.55"N 36 °46'0.85"E), spanning the time from ~620 kyr to present (Figs. 8.1, 7.1). The Chew Bahir basin is situated in a transition zone between the Main Ethiopian Rift and the Omo-Turkana basin, adjacent to the Lower Omo Basin, where some of the oldest known fossils of anatomically modern humans were found [218]. According to recent archeological findings, the adjoining highlands in the area may have been a refuge area for groups of H. sapiens during times of climatic stress [11, 234, 319].

We compared the Chew Bahir record of environmental change during the past ~620 kyr with the wetness index for the wider northeastern Saharan/North Africa from Ocean Drilling Program (ODP) Site 967 [115] (Fig. 8.1). The site was drilled during ODP Leg 160 in the eastern Mediterranean (34°N, 34°E, $2,252$ m water depth), where Saharan and North African dust and Nile riverine input are the primary contributors of sediment. The ODP Site 967 wetness index is a combined run-off and dust signal in a single metric, reflecting the effects of both strengthening/northward migration (increased run-off) and weakening/southward retreat (increased dust) of the northern and northeastern African monsoon [115]. The catchment of the Nile River with its two tributaries, the White Nile and the Blue Nile, extends from southeastern Africa to parts of the northwestern Ethiopian highlands. In the very wet phases, the Lakes Abaya-Chamo-Chew Bahir-Turkana system with its connecting rivers drained into the Nile catchment [143]. Therefore, both the upper Nile catchment and the Chew Bahir catchments are in spatial proximity and are most likely exposed to similar climate fluctuations and their causes. Humid conditions, recorded in both ODP Site 967 and CHB cores, could indicate the regional significance of a wet phase. Similar patterns in the types of variability and transitions could be indicative of the effect of a similar climate dynamic.

As a contribution to the reconstruction of environmental conditions for eastern Africa based on long terrestrial sediment records, our CHB record is firstly used to classify variability down core in order to identify recurring episodes of stable wet or dry, of cyclic or more complex but deterministic variability, and of random variability. Second, we classified types of transitions, including episodes with no change, linear/gradual shifts with different rates of change, as well as different types of rapid transitions such as tipping points. As soon as a classification of variability and transitions is available, one can discuss possible reasons for the similarity, e.g. similar boundary conditions such as global ice volume, local insolation, atmospheric CO_2 levels and ocean sea-surface temperatures. Finally, we hypothesize which types of variability and transitions may have affected the biosphere including hominins.
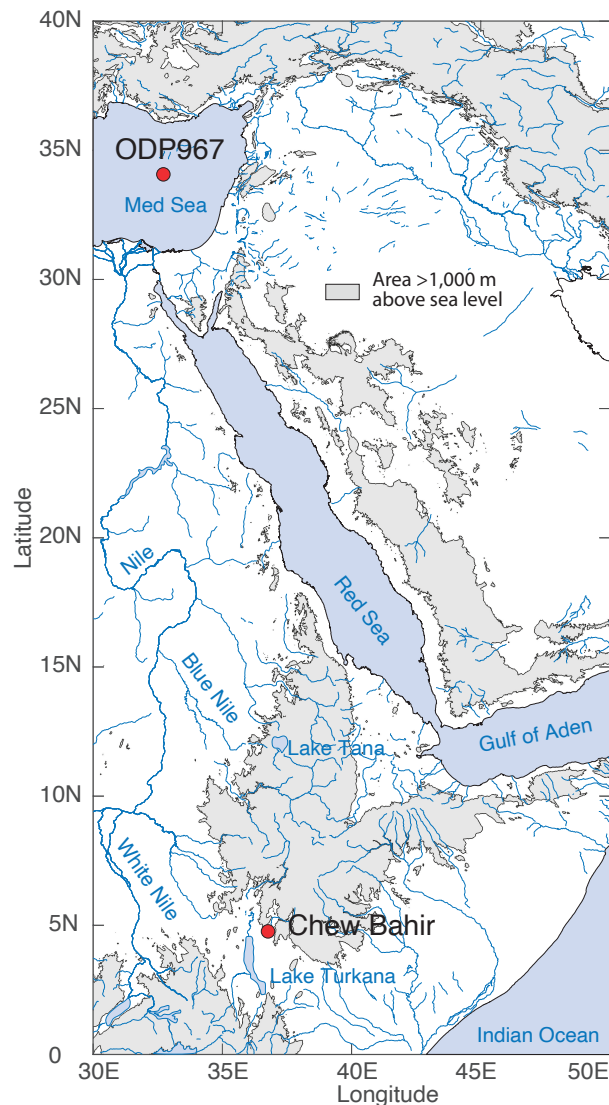
Figure 8.1: Map of northeastern Africa and adjacent areas showing the location of the Chew Bahir basin (4°45'40.55"N 36°46'0.85"E, ~500 m above sea level), the ODP Leg 160 Site 967 in the eastern Mediterranean Sea (34°4'6"N 32°43'31"E, ~2,254 m water depth), and the river Nile with its two tributaries the White and Blue Niles connecting both regions. Coastline and river polygons from the Global Self-consistent, Hierarchical, High-resolution Geography Database (GSHHG) [327]. Topography from the 1 arc-minute global relief model of the Earth's surface (ETOPO1).

8.2 Materials & methods

8.2.1 The Middle Pleistocene–Holocene paleoclimate record of the Chew Bahir basin

The sediment cores described herein were collected in the western part of the Chew Bahir basin in the southern Ethiopian Rift (4.1–6.3°N, 36.5–38.1°E, ~500 above sea level; Fig. 8.1). Chew Bahir is a tectonic basin, separated from the Lower Omo basin to the west by the Hammar Range, which is the source of most of the sediments at the coring site. This range to the west and the highlands to the north and north-east consist of Late Proterozoic granitic and mafic gneisses, whereas the eastern part of the catchment is dominated by Miocene basaltic lava flows. Oligocene basalt flows with subordinate rhyolites, trachytes, tuffs and ignimbrites cover the Precambrian basement units in the distal north-eastern, northern, and north-western parts of the catchment [63, 224]. Being a closed basin, Chew Bahir forms a terminal sink for weathering products from its $32,400$ km^2 catchment.

The present-day climate in eastern and northeastern Africa is influenced by a number of major air streams and convergence zones, with their effects superimposed on regional influences associated with topography, large lakes, and the oceans [232]. Rainfall in the Chew Bahir catchment is associated with the passage of the tropical rain belt, resulting in a strongly bimodal annual cycle. Most of the moisture reaching the Ethiopian highlands in June–August comes from the Mediterranean and Red Sea (55%), and from the Indian Ocean (31%) [316]. Short-term (annual to decadal) fluctuations in the intensity of precipitation relate to E-W adjustments in the zonal Walker circulation associated with the El Niño-Southern Oscillation (ENSO) and the Indian Ocean Dipole (IOD), possibly as a direct response to sea-surface temperature (SST) variations in the Indian and Atlantic Oceans, which are in turn affected by the ENSO and the IOD [232].

The paleoclimate of Chew Bahir was first reconstructed using six short cores, up to ~17 m long and collectively spanning ~47 kyr, which were collected in 2009–2010 [92, 93, 304–306]. In the context of the Hominin Sites and Paleolakes Drilling Project (HSPDP) to drill at key fossil hominin and archeological sites [37, 57], we collected parallel, duplicate cores: HSPDP-CHB14 2A (4°45'40.32"N 36°46'0.48"E) and 2B (4°45'40.68"N 36°46'1.20"E) in the Chew Bahir basin, 266.38 and 278.58 m long, respectively, in Nov–Dec 2014. A 292.87 m long composite core of the Chew Bahir Drilling Project (CBDP) with more than 90% recovery was created from the duplicate cores.

The composite core was developed by merging the two parallel cores 2A and 2B by core-to-core correlation using MSCL logs, core images, lithological description and XRF data sets. Radiometric age constraints were based on ^{14}C dating of ostracodes, optically stimulated luminescence (OSL) dating of fine-silt sized quartz grains, and single-crystal total-fusion (SCTF) 40Ar/39Ar dating of feldspars from tuffaceous zones within the core. In addition, a volcanic ash layer identified in the core has been correlated on the basis of major and minor element geochemistry to a dated tephra found in the outcrop at Konso, in the southern Main Ethiopian Rift, namely the Silver Tuff [257]). The ages generated are stratigraphically consistent, and Bayesian age-depth modeling incorporating ^{14}C, OSL and 40Ar/39Ar ages, and tephrochronological data has been used to build an age model for the Chew Bahir cores (age model RRMay2019, Roberts et al. [257]). The 1σ confidence intervals increase with increasing depth, ranging from < 10 kyr in the uppermost 50 m (corresponding to ~100 kyr BP) and 10-35 kyr below 50 m composite depth (~100-620 kyr) [257].

We analyzed the potassium (K) concentrations of the sediment, determined by micro X-ray fluorescence (μXRF) scanning which was previously shown to be a reliable proxy for aridity in the Chew Bahir basin [93, 305, 306]. The most likely process linking climate with K concentrations is the authigenic illitization of smectites during episodes of higher alkalinity and salinity in the closed-basin lake resulting from a drier climate [91]. After processing the μXRF data to remove coring and scanning artifacts, the data were corrected for outliers and jumps, before we applied various types of normalizations and standardizations of the data [90].

Micro XRF scanning was carried out with a resolution of 5 mm, corresponding to a mean time resolution of ~10 yrs, ranging from ~7–50 yrs [90, 257]. The K record has been interpolated to an evenly-spaced time axis with a resolution of 100 yrs, corresponding to a ~50 mm (10–70 mm) spatial resolution, using a piecewise cubic Hermite interpolating polynomial [97]. This resolution avoids the effects of the scanner's beam width as well as the effects of bioturbation on the signal (e.g. [303]), with the exception of large, discrete burrows, which we (as well as other disturbances of the core) have tracked down with the help of core photos and excluded from the interpretation. On the other hand the centennial resolution allows us to investigate the effect of millennium-scale or orbital-scale climatic fluctuations on environmental conditions as well as transitions over several

hundreds or thousands of years.

8.3 Results

Analogue to the approach in Trauth et al. [304] (see Chapter 7, Section 7.3.2) recurrence plots (RPs), together with a recurrence quantification analysis (RQA), were used to describe different types of environmental variability and transitions in the Chew Bahir (Tab. 8.1, Fig. 8.2 and Suppl. Figs. H.2, H.3, H.4, H.1) From the available RQA measures we have selected *RR* and *DET* because they describe important properties of the dynamic Chew Bahir system but are very descriptive compared to other RQA measures [209, 304] (see Appendix A). We compared the RPs and RQA measures of the Chew Bahir record with those of the wetness index for the wider northeastern Saharan/North African record from ODP Site 967 [115] (Fig. 8.3).

Table 8.1: Summary of the RP/RQA results to describe different types of environmental variability and transitions in the Chew Bahir.

	Type 1	Type 2
Occurrence/Time (kyr BP)	617-410	410-348
	348-272	272-252
	252-122	59-11
	107-59	
	11-0	
RP Texture	sparse recurrences	dense recurrences
	longer diagonal lines with larger spacing (5, 10-20 kyr)	block-like structures, short diagonal lines with smaller spacing (1-2 kyr)
Recurrence Quantification	RR low	RR intermediate to high
	DET low to intermediate (except for the oldest part, where DET is also high)	DET high

The embedding parameters m and τ have been selected by an improved false nearest neighbors approach [125] and mutual information [95] (see Figure H.1). Before embedding, it is often helpful to filter the time series to improve the signal-to-noise ratio. Here we used a simple nonlinear noise reduction method that is based on state space averaging [274]. We have, thus, used noise filtered K as a proxy for aridity, as the dominant process linking climate with K concentrations is the authigenic illitization of smectites during episodes of higher alkalinity and salinity in the closed-basin lake resulting from a drier climate [91]. For the analysis of the ~620 kyr record the K record was embedded in a state space with a dimension of $m = 6$ and temporal distances of $\tau = 10$ (Suppl. Fig. H.1), equivalent to 10×0.1 kyr $= 1.0$ kyr, where 0.1 kyr is the resolution of the time series following a piecewise cubic Hermite interpolating polynomial [97]. We used the window size $w = 500$ and the step size $ws = 50$ data points of the moving window to calculate the RQA measures (Fig. 8.2). The size w of the window corresponds to 500×0.1 kyr $= 50$ kyr and the step size is 50×0.1 kyr $= 5$ kyr. We used a minimum length of 10 points to compute DET and an adaptive recurrence threshold (fixed amount of neighbours, see Chapter 1), which corresponds to a local and global recurrence rate of 8%. To compare the RP/RQA based dynamics in the Chew Bahir record of aridity with the wetness index of ODP Site 967, we interpolated the marine record to the same time axis. Then the same embedding parameters (Suppl. Fig. H.1) and recurrence threshold selection method and size were used to obtain the RPs and ensure comparability. For the computation of the RQA measures the same window size has been selected (Fig. 8.3). Similarities in the texture of the recurrence plots of both proxy records show that the
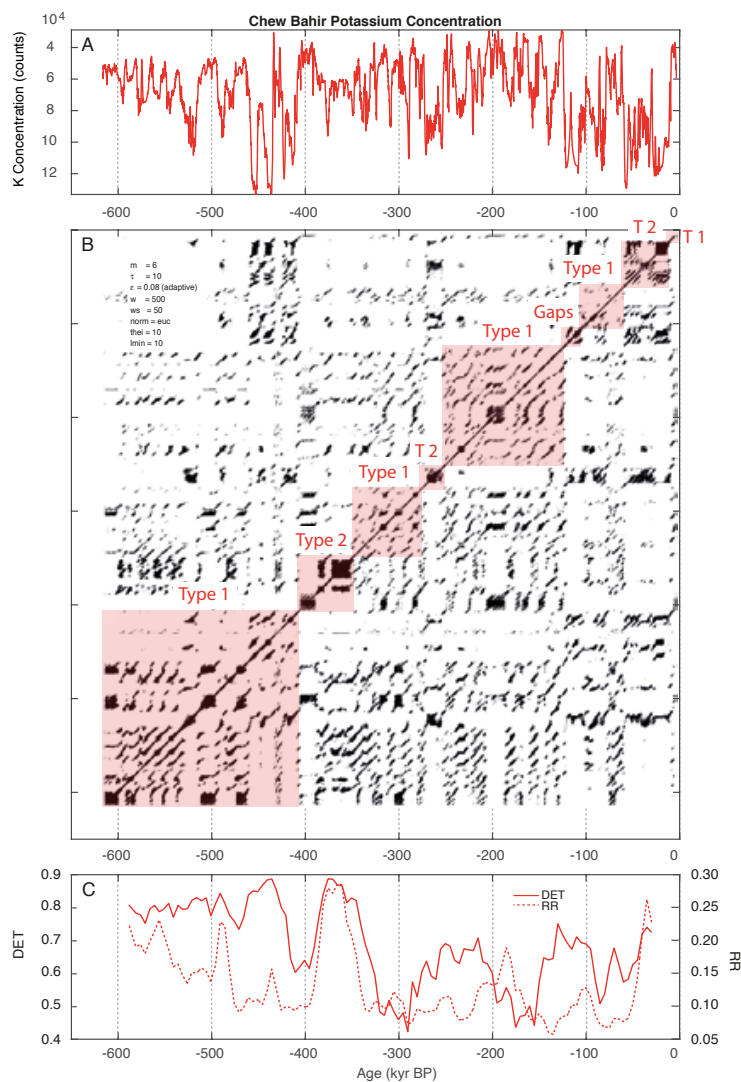
Figure 8.2: Recurrence plot (RP) and recurrence quantification analysis (RQA) measures of the potassium (K) concentrations of the sediment in Chew Bahir, southern Ethiopia: **A** the time series according to age model (3); **B** the recurrence plot and **C** the RQA measures of moving windows. Parameter abbreviations are as follows: m=embedding dimension, τ=time delay, ε=recurrence threshold, w=window size, ws=window moving steps, norm=vector norm, $thei$=size of Theiler window, $lmin$=minimum line length, RQA measures RR=recurrence rate and DET=determinism. See text and Fig. H.1 for a detailed description of the embedding parameters and RQA measures.

recurrence threshold selection and the embedding procedure provides comparable results with these parameter values.

The first series of four clusters of recurrence points occurs between 617 and 410 kyr BP in the Chew Bahir record, with the largest cluster between 617 and 462 kyr BP, then three very weak clusters at 456–435, 435–425, and 425–410 kyr BP (Fig. 8.2). These clusters are structured by a series of partly s-shaped diagonal lines or they represent rather compact blocks with high recurrence rates centered at 610, 505 and 470 kyr, which mark 1–2 kyr long episodes of relative stability. The diagonal lines are 10–15 kyr apart, reflecting a series of wet-dry episodes recurring with approximately half-precession cyclicity, initiated and terminated by relatively abrupt transitions, marked by white vertical lines. As the result of the cyclic recurrence of wet and dry conditions, the DET values are relatively high, whereas RR has moderate, but declining values. Two extremely dry episodes centered at ~455 and ~438 kyr BP are reflected in two clusters of recurrence points and high DET values within the 458–436 kyr BP interval, separated by white vertical lines and slightly lower DET values. Zooming into the interval between 617 and 410 kyr BP with higher (0.015 kyr) resolution the diagonal lines become blurred, are separated by a few
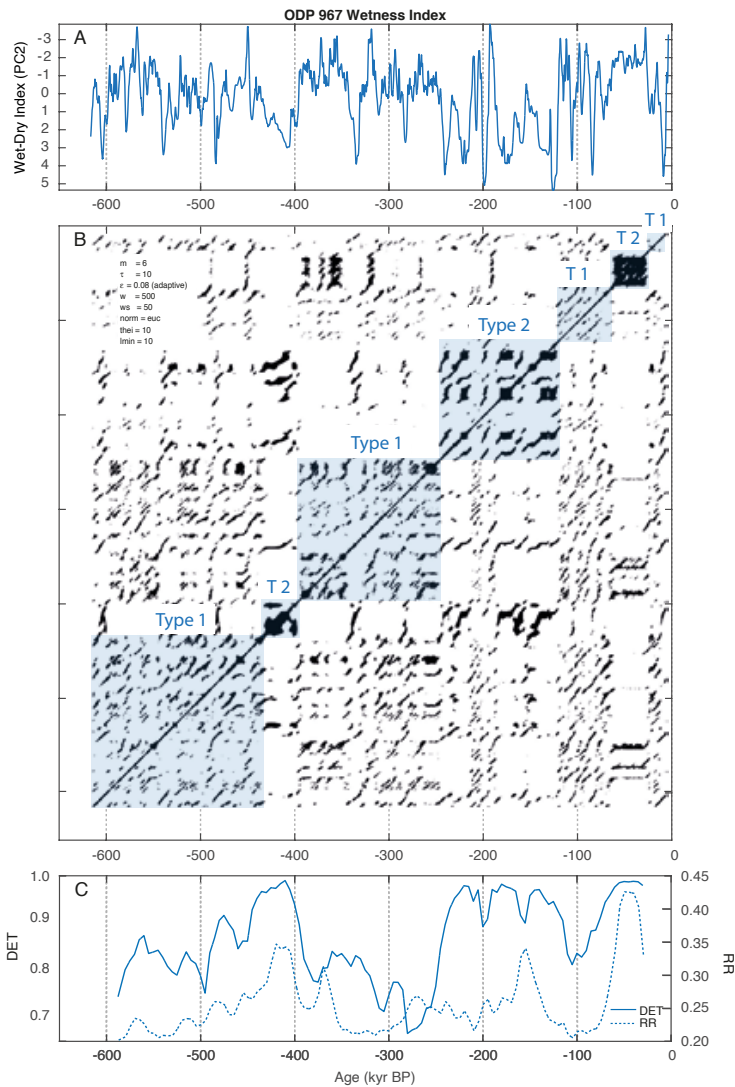
Figure 8.3: Recurrence plot (RP) and recurrence quantification analysis (RQA) measures of the ODP Site 967 wetness index according to Grant et al. [115] from the eastern Mediterranean: **A** the time series; **B** the recurrence plot and **C** the RQA measures of moving windows (lower panel). See previous figure for the meaning of the parameter abbreviations (Fig. 8.2).

thousand years (up to 10 kyr), and the RP is dominated by several small (< 10 kyr) blocks that are connected by black horizontal and vertical lines (Suppl. Fig. H.2A). This suggests that, at time scales of < 10 kyr, the system oscillates between shorter stable states, each 1–5 kyr long, with rapid transitions between them. Within the blocks, we observe diagonal lines indicating high-frequency (< 1 kyr) cyclicities. The large-scale block and line separation is dominated by 5, 10 and 20 kyr cycles. In the ODP Site 967 wetness index we find a similar type of variability in the time before 435 kyr, although in comparison diagonal structures are much less pronounced or even absent (Fig. 8.3). The RR values are at similar values to those in the Chew Bahir, except for a significant anticorrelation at ~420 kyr (Fig. 8.4). The DET values are generally lower but increase after ~500 kyr BP up to similar values as in the Chew Bahir record.

This interval ends with a rapid transition at ~410 kyr BP from dry to wet conditions. This transition is followed by two dense clusters of recurrence points between 410 and 348 kyr BP, indicating episodes of a stable wet climate with extreme RR and DET values (Fig. 8.2). The two clusters reflect relatively stable humid conditions, with the humidity level in the first cluster being higher than in the second cluster and hence the two clusters are separated by a transition towards

a slightly less wet climate at around 390 kyr BP. The second cluster is interrupted by a distinctive dry event at around 376 kyr BP. Interestingly, both clusters show an irregular pattern of diagonal lines, partly slightly curved, suggesting recurrent dry events, but with slightly variable cyclicity. Zooming into the interval between 410 and 348 kyr BP with higher (0.015 kyr) resolution we observe blurred diagonal lines with 1–2 kyr distances in the earlier part before ~390 kyr, with 3–5 kyr in the younger part, indicating a weak cyclicity on millennium time scales even within the otherwise quite stable wet episodes (Suppl. Fig. H.2B). The interval between 410 and 348 kyr BP is terminated by gradual (~10 kyr) transition towards slightly more humid and more variable conditions after ~348 kyr BP. The ODP Site 967 wetness index indicates similar wet conditions during this episode, also with relatively low variability (Fig. 8.3). As a consequence, the RR values are relatively high but decreasing whereas the DET values are also at high level but relatively low compared to before and after the event.

The next cluster of recurrence points between 348 and 272 kyr BP is marked by a series of blocks with weak internal structure and separated by white vertical lines (Fig. 8.2). This structure reflects a series of relatively stable wet conditions, interrupted by several thousand-year long dry episodes, some of which are bounded by relatively rapid transitions from wet to dry and back. The occasional appearance of diagonal lines, though rarely parallel to the main diagonal, indicates weak cyclic behavior. Within this interval the RR values are constantly low, whereas the DET values start at high values and decline until about 320 kyr BP before they remain at low values. Zooming into the interval between 348 and 272 kyr BP with higher (0.015 kyr) resolution confirms the observation of small blocks connected with black horizontal and vertical lines in the RP, as a result of a rapid change between relatively stable dry and wet conditions, with the exception of the block between 330 and 327 kyr BP. This block is merely the result of a gap that was closed by interpolation (Suppl. Fig. H.3A). The blocks themselves have little internal structure, with the exception of very weak diagonal lines with a spacing of 1–2 kyr. The interval between 348 and 272 kyr BP is terminated by a very rapid transition from wet to dry conditions at ~272 kyr. The ODP Site 967 wetness index shows a similar variability type during this episode, but with higher DET values remaining at high levels compared to those of the Chew Bahir (Fig. 8.3).

Between 272 and 252 kyr BP, we observe another dense cluster of recurrence points (Fig. 8.2). The internal structure of this block is reminiscent of similar structures of the clusters at ~400 and ~360 kyr BP, whereas the system state based on the difference in the DET values does not support this conclusion. The interval again shows very weak diagonal lines. As a result of this cyclicity, the RR and DET values are at moderate levels, with DET at a local maximum. This pattern suggests that the climate was relatively stable, but fluctuations between humid and dry occur in cycles. These changes occur more and more frequently over time, until the climate finally quickly changes to generally humid conditions at around 228 kyr BP. Zooming into the interval between 272 and 252 kyr BP with higher (0.015 kyr) resolution reveals that the curved diagonal lines are rather a sequence of short, laterally offset diagonal pieces. In the younger section of the block, the diagonal lines disappear (Suppl. Fig. H.3B). Considering the course of the time series it can be seen that the diagonals reflect increasingly shorter wet phases, and until after about 236 kyr quite stable, relatively dry conditions prevail in the Chew Bahir. The ODP Site 967 wetness index shows similar climate fluctuations during this episode, but these are too unclear to be described with a different type of variability (Fig. 8.3). The RR values are low, similar to those in the Chew Bahir, whereas the DET values rise sharply, probably because of subsequent long-wave cyclicities, in contrast to the Chew Bahir.

The episode between 252 and 122 kyr BP mirrors the earlier ones between 617 and 410 kyr

BP and between 348 and 272 kyr BP, whereby in the second interval also the state of the environmental system also recurs. The episode is unfortunately followed by two gaps due to core loss between ~122 and 107 kyr BP, which are filled by the interpolation with curves, so we get a high RR at this point that we cannot interpret (Fig. 8.2). Zooming into the interval between 252 and 122 kyr BP with higher (0.015 kyr) resolution reveals many small blocks surrounded by s-shaped curved structures rather than continuous diagonal lines which mark short (< 5 kyr) wet phases which begin and end gradually (Suppl. Fig. H.4A). Besides that, we observe longer (5–10 kyr) wet episodes with rapid onset and termination, internally structured by converging diagonals, structurally similar to the ones in the interval between 272 and 252 kyr BP, but with a different system state based on the RR and DET values. This interval is terminated by ~5 kyr long gaps after 122 kyr BP. During this episode, the ODP Site 967 wetness index shows a type of variability that is more similar to the episode of 435 and 370 kyr BP, rather than the one after 370 kyr BP (Fig. 8.3). The course of the RR values is similar to the one of Chew Bahir in the first half of the interval, but seems to be anticorrelated in the second half. The DET values show a very similar pattern to those of the Chew Bahir, although at higher levels. The increasing DET values at the beginning of the interval are about 20 kyr later for ODP Site 967, then they vary in a very similar way even including a minimum at about 160 kyr, and then decline at the end of this interval (Fig. 8.4).

Between 107 and 59 kyr BP, we find a very complex phase with fast changes between stable humid, very variable humid and very dry conditions, each separated by fast transitions (Fig. 8.3 and Suppl. Fig. H.4B). After a very dramatic transition we observe very regular climate fluctuations between ~59 kyr BP and about 11 kyr BP, when it became relatively stable dry. This interval is again characterized by a cluster of dense recurrence points, slightly converging diagonal lines which indicate increasing recurrence rates (i.e. shorter wet-dry cycles), before we see a black block of recurrence points as the result of very stable dry conditions. The last ~11 kyr are characterized by very wet conditions, interrupted with a short, about ~1 kyr long dry intervals and terminated by a transition towards a dry climate. The distinct cyclic alternation of wet and dry conditions, each of which are ~5 kyr long, results in high RR and DET values. The last ~47 kyr BP are preserved at higher resolution in the short cores from the same basin, also examined with the method of RP/RQA [304]. Here we found very similar patterns as on the long time scale, with alternating appearances of blocky structures and diagonal lines, with different transitions between episodes. The ODP Site 967 wetness index here shows a similar variability with long-periodic, high amplitude fluctuations between wet and dry between ~120 and 60 kyr BP, followed by an episode without such variations (Fig. 8.3). The time interval between 60 and 25 kyr BP is relatively stable, as compared with the high-frequency variability observed in the Chew Bahir, before we observe two long wet cycles, out of which the second one is also seen in the Chew Bahir record.

8.4 Discussion

We applied a detailed analysis of the RPs together with a RQA to distinguish between different types of climate variability, and transitions in the Chew Bahir basin [306] (Fig. 8.4). Here, climate is inferred from our key proxy, the potassium concentration of the sediment representing the relative aridity in the Chew Bahir Basin [91]. Our statistical analyses herein are a contribution to an accurate picture of environmental change in eastern Africa during the last ~620 kyr and thereby provide a quantitative, high resolution climatic component useful for investigating human-climate

interactions. The fossil and archaeological record of eastern Africa is still too limited to draw definitive conclusions with respect to current hypotheses on the relationship between climate and evolutionary patterns in humans and other animals. However, our results do allow for some initial comparisons and hypotheses (e.g. [99, 269, 286]).

Our analysis clearly shows a number of different types of variability in the K record that overlay a long-term trend towards greater aridity and variability (Figs. 8.2 and 8.4). These types of variability are separated by transitions, which are of varying types of durations and structures. Many of these types, both variability and transitions, occur multiple times during the last ~620 kyr, so it is interesting to investigate them more closely to see if they are characteristic of the Chew Bahir Basin or possibly occur even beyond the region. In addition, it is important to examine whether these types are linked to certain regional or global boundary conditions (e.g. global ice volume, atmospheric CO_2 levels, ocean sea-surface temperatures, Fig. 8.4).

We observe two basic types of variability that do not mix, but each form a series of variants. The first type of variability, occurring at 617–410 kyr BP, 348–272 kyr BP, 252–122 kyr BP, and (after a period with no data) 107–59 kyr BP, are slow variations with cycles of ~20 kyr and subharmonics of this cycle, as indicated by the occurrence of diagonal lines with 20 kyr, 10 kyr, and 5 kyr spacing (Figs. 8.2 and 8.3). In addition to these cyclical wet-dry fluctuations in the area, extreme events often occur, i.e. short wet or dry episodes, lasting for several centuries or even millennia, and rapid transitions between wet and dry episodes. This type of variability probably reflects the influence of precessional forcing in the lower latitudes at times of increased eccentricity, with a tendency towards extreme events. This also shows in comparatively low RR and DET values, suggesting a lower determinism of climate variations for this type of variability. This type of variability correlates with maximum values of the long (400 kyr) eccentricity cycle, and hence maximum variability in the precession frequency band. There does not seem to be a systematic correlation with atmospheric CO_2 concentration within this type of variability.

The second type of variability, occurring at 410–348 kyr BP, 272–252 kyr BP, and 59–11 kyr BP, is characterized by relatively low variation on orbital time scales. Instead, we observe significant century-millennium-scale variations with increasing frequency in the course of an episode with Type 2 variability, as block-like pattern with thin diagonal lines of short spacing suggest (Figs. 8.2 and 8.3). The very prominent cycles in the frequency band lead to very high DET and partly high RR values, which indicate a higher deterministic regime of climate change, much higher than at times of Type 1. Within this type of variability there are extremely fast transitions between dry and wet within a few decades or years, in contrast to those within Type 1 with transitions that last several hundreds of years. Type 2 variability seems to be linked with minimum values of the long (400 kyr) eccentricity cycle, and again there does not seem to be a link with atmospheric CO_2 levels. The first episode with Type 2 variability occurs in an interval with maximum eccentricity in the 100 kyr frequency band when the atmospheric CO_2 was quite high. The other two episodes occur during minimal eccentricity in this frequency band, with low CO_2 levels at the same time.

The ODP Site 967 wetness index shows a very similar type of variability, with RR values that highly resemble the variations in the Chew Bahir record, except for the prominent anti-correlation at ~420 and ~150 kyr BP (Fig. 8.4). The DET values fluctuating in a smaller range compared to those of the Chew Bahir, but the temporal variations of DET show obvious similarities at both sites, in particular during the last ~350 kyr. Since the age models of both sites have large uncertainties, it is not possible at this point to judge conclusively whether time shifts in the DET curves are the result of poor age control or actual differences in climate dynamics recorded in the two locations. Comparing the temporal occurrence of the different types of variability in
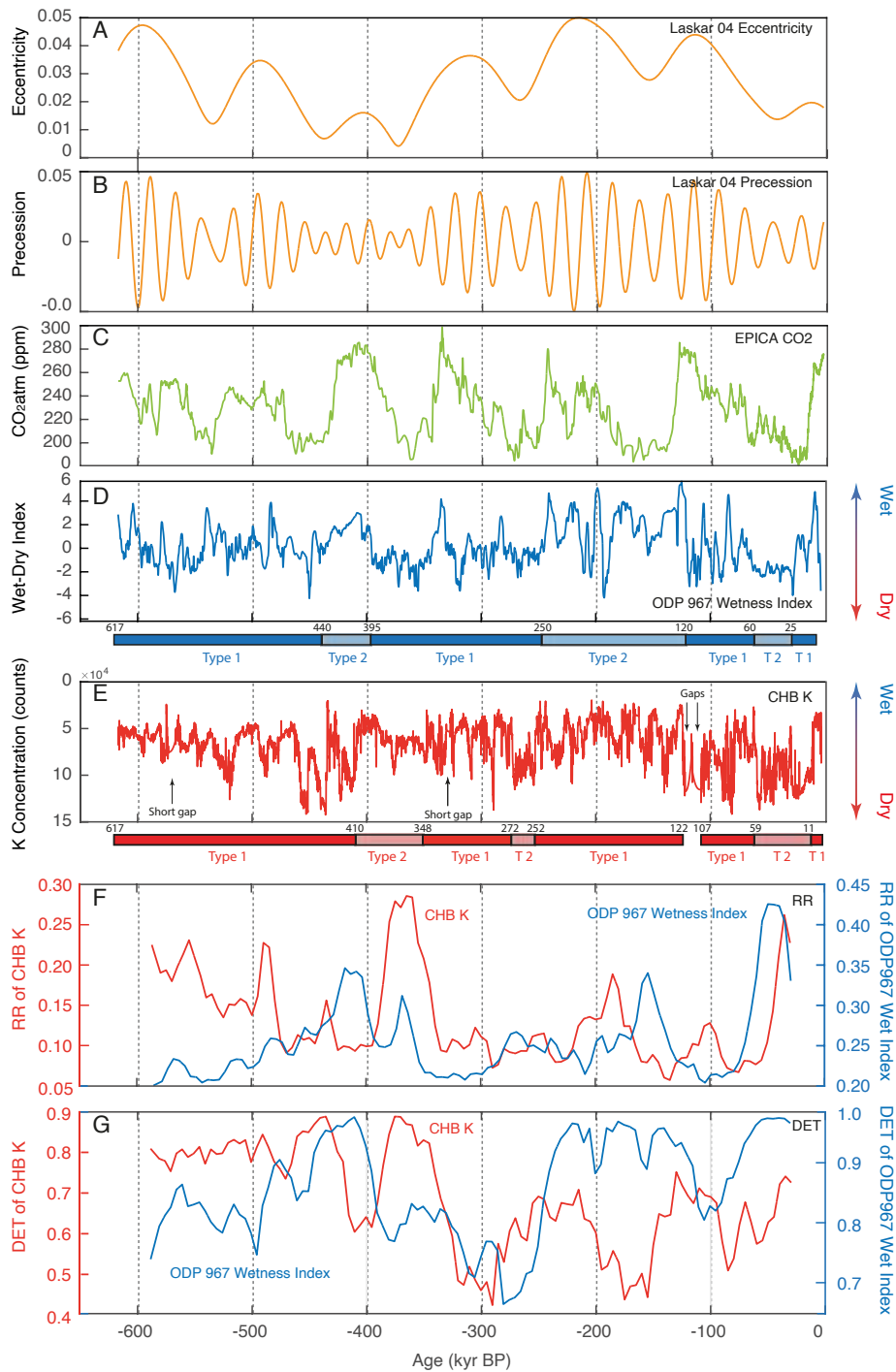
Figure 8.4: Comparison of the complex dynamics of variations in aridity in southern Ethiopia with strengthening/northward migration and weakening/southward retreat of the North African monsoon. **A**+**B** Earth's eccentricity and precession cycle [176]; **C** Antarctica EPICA Dome C atmospheric CO_2 according to Bereiter et al. [20]; **D** ODP Site 967 wetness index according to Grant et al. [115]; **E** Chew Bahir aridity derived from the potassium (K) concentrations of the sediment in Chew Bahir using age model RRMHT500 [257], note the reverse y-axis; **F**+**G** comparison of RQA measures determinism (DET) and recurrence rate (RR) of the recurrence plots (RP) of the Chew Bahir aridity record (in red) and ODP Site 967 wetness index (in blue). DET is a measure of the determinism of the system and RR describes the probability of recurring states of the system in a particular time period.

both localities, the ODP Site 967 wetness index shows a Type 2 variability approximately during the first and third episode of Type 2 variability in the Chew Bahir record, but not during the second episode of Type 2 variability in the Chew Bahir. Instead we find a Type 2 variability between ~245 and 120 kyr BP in the ODP Site 967 wetness index, all with high DET and RR values. The different types of variability, very obvious in the RPs, but differing slightly in the RQA measures, could also indicate actual differences in dynamics, but also the influence of the spline interpolation on the temporal auto-correlation.

Overall, the K curve shows a clear trend towards a drier and more variable climate, most prominently during the last 200 kyr, which is also reflected in the reducing DET and RR values (Fig. 8.4). The three episodes with Type 2 variability are about the same length (45–60 kyr) but have (according to the long-term trend) decreasing average humidity levels. The first episode is characterized by a very humid climate, while the two later episodes were rather dry. In particular, the DET values show several stepwise declines over the entire time series, which are not compensated by corresponding increases. The most striking declines are centered at approximately 480 kyr BP, 420 kyr BP, 330 kyr BP, 200 kyr BP, 180 kyr BP and 90 kyr BP. The most dramatic increases in DET are centered at approximately 460 kyr BP, 385 kyr BP, 285 kyr BP, 255 kyr BP, 150 kyr BP and 75 kyr BP. The variability also increased, especially if a few larger fluctuations between 460-410 kyr, at ~335 kyr BP and at ~245 kyr BP are disregarded. In addition, the frequency of rapid transitions from dry to wet and back increases over time. The ODP Site 967 record shows similar trends on long time scales, whereas there are large differences on shorter time scales, especially in the degree of variability.

The different types of variability and the transitions between these types have important implications for our understanding of the availability of water. Different modes of variability would have transformed eastern Africa's environment considerably, including its vegetation and fauna, and would have shaped the habitat of hominins, including archaic and modern H. sapiens, in that part of the continent. The central question, however, is to what extent the different types of fluctuations observed in the environmental record of the Chew Bahir have actually had an impact on living conditions of hominins. Overall, long-term changes (> 1 kyrs) would have formed the living environment of hominins on a time-scale that responds to human evolution and corresponds to a time window that is long enough to facilitate large-scale dispersal. The short-term fluctuations (< 1 kyrs) may have had dramatic consequences for populations including differential mortality/fertility of H. sapiens down to the level of individuals, and thus short-term changes in behavior, including evasive movements to more favorable habitats (e.g. [92]).

To assess the impact of climate variability on people, it is worth using a well-studied younger analogue of hydroclimatic transformation in eastern Africa. The most recent example of a wet-to-dry transition within Type 1 was the time-progressive termination of the African Humid Period (AHP, 15–5 kyr BP), which lasted several hundred years in most areas [282, 306]. At this time, a previously green, then yellow Sahara was largely depopulated, but this happened quite slowly and due to the time-transgressive termination rather in the form of people slowly following favorable living conditions [172, 174, 282]. During this gradual transition, climate deterioration could have fostered an important socio-economic transition, including the transition from hunter-gatherer to pastoralism [92, 106].

In contrast, the most recent example of a wet-to-dry transition within Type 2 is the onset of pronounced arid conditions during the Younger Dryas chronozone (YD, ~12.8–11.6 kyr BP) that occurred within ~45 yr at Chew Bahir [306]. Most importantly, millennial-scale transitions such as the YD happened everywhere at the same time, unlike the orbital-controlled slow changes

[282, 306]. It is implausible that such relatively rapid transitions triggered a fundamental societal transformation, similar to the one during the termination of the AHP. Instead, climate shifts of this rapidity would allow response patterns that are implementable within (less than) a life-time span, such as short-term migration towards proximal more favorable living conditions. Examples of this are movements from hot dry low-lands into still vegetation-rich high altitudes, or even the complete disappearance of entire local human populations due to scarcity of resources [92]. The possible alternative response to environmental extremes and too fast transformation is extinction/extirpation, because living conditions deteriorated so quickly that neither physical nor cultural adaptation was possible. However, short-term and short-distance mobility depends (a) on a number of socio-cultural conditions (= how flexible, how adaptable are groups, how are they organized?) and (b) mode of climatic change (= pulsed, rapid, parallel with other areas).

8.5   Conclusion

We find two types of variability in the Chew Bahir record, (1) Type 1 probably reflecting the influence of a precessional forcing in the lower latitudes at times of increased eccentricity, with the tendency towards extreme events, and (2) Type 2 with significant century-millennium-scale variations with increasing frequency. Within Type 2 of variability there are extremely fast transitions between dry and wet within a few decades or years that would have exerted a high level of climatic stress on the biosphere including humans, in contrast to those within Type 1 with transitions within several hundreds of years. As the body of archaeological evidence, including fossils and diagnostic tools, continues to increase in the future, it will be exciting to compare potential response patterns to our prosed Type 2 phases of high climatic stress and see whether changes in settlement activities, cultural innovation, or even the emergence or disappearance of populations/occupancy can be correlated with the climatic dimension of the complex framework in human-climate interactions.

Code availability

The Chew Bahir potassium record and all MATLAB$^{®}$ scripts are available at *http://mres.uni-potsdam.de.*

Part IV

Application in neuroscience

Chapter 9

Motor execution reduces EEG signals complexity: recurrence quantification analysis study

Abstract

The development of new approaches to detect motor-related brain activity is key in many aspects of science, especially in brain-computer interface (BCI) applications. Even though some well-known features of motor-related electroencephalograms (EEGs) have been revealed using traditionally applied methods, they still lack a robust classification of motor-related patterns. Here we introduce new features of motor-related brain activity and uncover hidden mechanisms of the underlying neuronal dynamics by considering event-related desynchronization (ERD) of μ-rhythm in the sensorimotor cortex, i.e. tracking the decrease of the power spectral density in the corresponding frequency band. We hypothesize that motor-related ERD is associated with the suppression of random fluctuations of μ-band neuronal activity. This is due to a lowering of the number of active neuronal populations involved in the corresponding oscillation mode. In this case we expect more regular dynamics and a decrease in complexity of the EEG signal recorded over the sensorimotor cortex. In order to support this theses we apply measures of signal complexity by means of recurrence quantification analysis (RQA). In particular, we demonstrate that certain RQA quantifiers are very useful to detect the moment of movement onset and therefore are able to classify the laterality of executed movements.

9.1 Introduction

The study of motor-related brain activity for noninvasive EEG-based brain-computer interfaces is a challenging task at the intersection of neuroscience, medicine, nonlinear physics and engineering. One active research area is, for example, neurorehabilitation in post-stroke patients with motor and cognitive impairment [285]. Another branch of actual research demanding brain motor-related

activity decryption is a mental control of robotic systems, prosthetic devises and vehicles [219]. By translating recorded signals of brain activity into control commands, brain-computer interfaces (BCIs) can provide a communication channel between the human and the external device [50, 179, 332]. Recently a considerable progress has been achieved in invasive BCIs for motor control. This is due to the principles of the invasive interfaces operation, which rely on the firing properties of individual neurons or small groups of neighboring neurons modulating their activity according to the motor tasks [178]. In this case motor-related neuronal activity patterns are pronounced and well reproduced, which allows to develop precise schemes for motor control [19, 179]. Invasive BCIs have an outstanding ability for an accurate detection and translation of brain motor commands. However, its application can pose severe challenges to patients and surgeons, since it requires complex brain surgery. It is therefore only carried out in rare cases of urgent need and is not suitable for daily purposes.

On the contrary, noninvasive BCIs are easy to apply and much more convenient in terms of usability. Among a variety of neuroimaging methods, electroencephalography (EEG) appears to be one of the most suitable ones for routine BCI applications [220]. A comprehensive review on the current state, future perspectives and remaining challenges of sensorimotor EEG-based interfaces was given by Yuan and He [334]. Methods of spatial filtering [30, 253], machine learning [53, 193] and time-frequency analysis [193, 242] are the core algorithms for feature extraction in this context. EEG simultaneously records electrical activity of a large group of neuronal populations located close to the measuring sensor [194]. Generally, distinct neuronal ensembles do not behave in coherency. Therefore, EEG signals represent a complex mixture of local neuronal activity components. The latter determines inherent critical properties of EEG signals, such as a poor signal-to-noise ratio and nonstationarity. Traditionally used methods for motor-related feature extraction, such as spatial filtering and time-frequency analysis, allow to associate motor actions with event-related desynchronization (ERD) of μ-band oscillations (8-13 Hz) over the sensorimotor cortex. However, these features (e.g. location of brain activity sources, amplitudes of spectral components) are of strong inter- and intra-subject variability [14, 120, 145, 193]. The aforementioned inherent nonstationarity and a poor signal-to-noise ratios of EEG signals strongly complicate the detection and classification of motor-related patterns in single trials. Hence, it is of high interest to find relevant features and methods that will withstand the discussed weaknesses of EEG recordings.

It is known that motor tasks block ongoing activity in the μ-band (8–13 Hz) of a EEG record, i.e. ERD takes place [242]. Motor-related ERD implies a time-locked decrease in the number of active neurons involved in μ-oscillations [243]. We hypothesize that this is equivalent to a suppression of spontaneous fluctuation of neuronal activity in the corresponding frequency band compared to the preceding background activity. Thus, we expect that motor-related neuronal dynamics should be reflected in EEG recordings by the signal's complexity reduction.

In order to find new features of the motor-related brain activity we explore EEG signals from the viewpoint of signal complexity. Complexity of EEG signals was assessed and quantified with recurrence quantification analysis (RQA). RQA is a nonlinear method, which describes the recurrence structure of a system by several quantifiers, in order to examine its complexity and uncover hidden underlying phenomena [209, 321]. It is a well known and powerful tool for the analysis of biological signals, specifically heart rate variability [2, 213], muscle activity [18, 135], sleep [3, 238, 267] and pathological EEGs [229, 256]. Early RQA studies that focused on EEG analysis demonstrated the ability of RQA measures to quantify N400 event-related potentials (ERPs) in single trials (e.g. [208]), which emphasizes the robustness of the RQA approach in the

context of the current study.

In summary, this work intends to further contribute to the non-invasive BCI research. New findings of features of motor-related brain activity support our hypothesis of a reduction of EEG signal's complexity in the μ-band when motor-tasks are performed. We test our hypothesis on the upper limb motor execution tasks and apply RQA to quantify changes of signal complexity caused by the motor task accomplishment. We seek to demonstrate that certain RQA quantifiers, namely determinism and recurrence time entropy, are sensitive to the transition from background to motor-related brain activity. This, in turn, reveals differences between left and right upper limb movements, which is valuable in the context of potential BCI applications.

This chapter is organized as follows. Section 9.2 describes the details of the experimental study, the data pre-/postprocessing and briefly the RQA method. Section 9.3 is devoted to the analysis of the time-dependent RQA measures and the inference of task vs. background differences along with differences between left and right limb movements. Finally, we summarize our results and discuss them in context of BCI development in Section 9.4.

9.2 Methods

9.2.1 Participants

Participants were recruited among the employees and students of the Innopolis University. Ten subjects (7 male, 3 female) were selected according to the following checklist: healthy, aged 18–33, right-handed, never participated in this or similar experiments before and no history of brain tumors, trauma or stroke-related medical conditions. All the participants were pre-informed about the goals and design of the experiment ans signed consent. The experimental study was performed in accordance with the Declaration of Helsinki and approved by the local research Ethics Committee of Innopolis University.

9.2.2 Data acquisition

EEG signals along with electromiograms (EMGs) from both hands were recorded using non-invasive EEG/EMG system "Encephalan-EEGR-19/26" (Medicom MTD company, Taganrog, Russian Federation). Electrocardiogram (ECG) and electro-oculogram (EOG) were also recorded for further removal of cardiac and eye-movement artifacts. All recorded signals were amplified and digitized at the sampling rate of 250 Hz. In order to record motor brain activity we used 9 EEG Ag/AgCl electrodes Fc3, Fcz, Fc4, C3, Cz, C4, Cp3, Cpz, Cp4 located over the motor cortex according to the international "10-10" system proposed by the American Electroencephalographic Society. To capture hand movements execution we placed 4 EMG electrodes as follows: 1 reference on the wrist and 1 on the forearm muscle for each hand.

9.2.3 Experimental setup

The session started with a 5-min recording of background brain activity, during which the participants were instructed to relax and listen to classical music. They were also asked not to think about anything special and to make no hand movements. During the active phase of the experiment each participant performed two types of motor actions according to the experimental
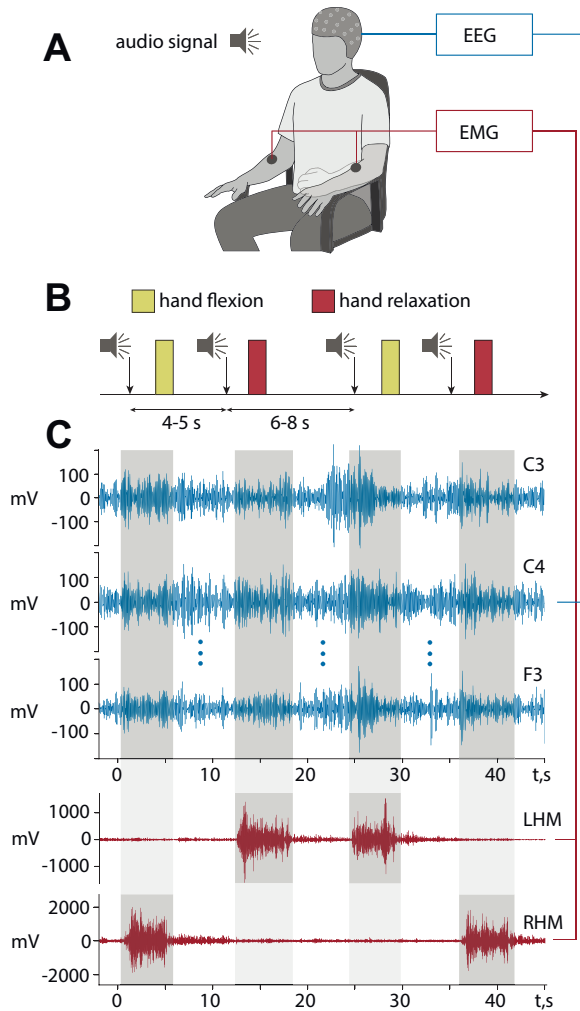
Figure 9.1: Schematic representation of the experimental procedure. **A** Subjects were sitting comfortably in the chair while performing motor actions of left and right hands on audio signal command. **B** Experimental sequence. Time intervals between the signals were chosen randomly in ranges 4–5 seconds between first and second signals for one task and 6–8 seconds from second signal of previous and first signals of the next task. **C** Examples of recorded μ-band-pass filtered EEG and EMG signals (LHM = left hand movement, RHM = right hand movement).

protocol, i.e. movements for the left and right hands (Fig. 9.1A). Each hand movement implied squeezing the hand into a fist after the first signal, holding it down to the second signal and then relaxing it. The time interval between the first and the second signals was randomly chosen for each motor task in the range 4–5 s and the time interval between the second signal of the current task and the first signal of the next task (resting period) was randomly chosen in the range 6–8 s (Fig. 9.1B). The active phase of the experiment consisted of 30 repetitions of each type of motor task (60 total) and the overall duration of the experimental procedure was approximately 18 minutes per participant, including background activity recording.

9.2.4 Data preprocessing

The following preprocessing steps were carried out to prepare raw EEG and EMG recordings for further analysis. First, cardiac and eye-movement artifacts were removed using recorded ECG and EOG signals via artifact removal method based on Gram-Schmidt process [134]. A Notch filter around 50 Hz was applied to EEG and EMG data to exclude power line effects. Second, we applied a 5th-order Butterworth band-pass filter in the range 8–13 Hz to the 18 minute multichannel EEG signals in order to extract μ-band neuronal oscillations associated
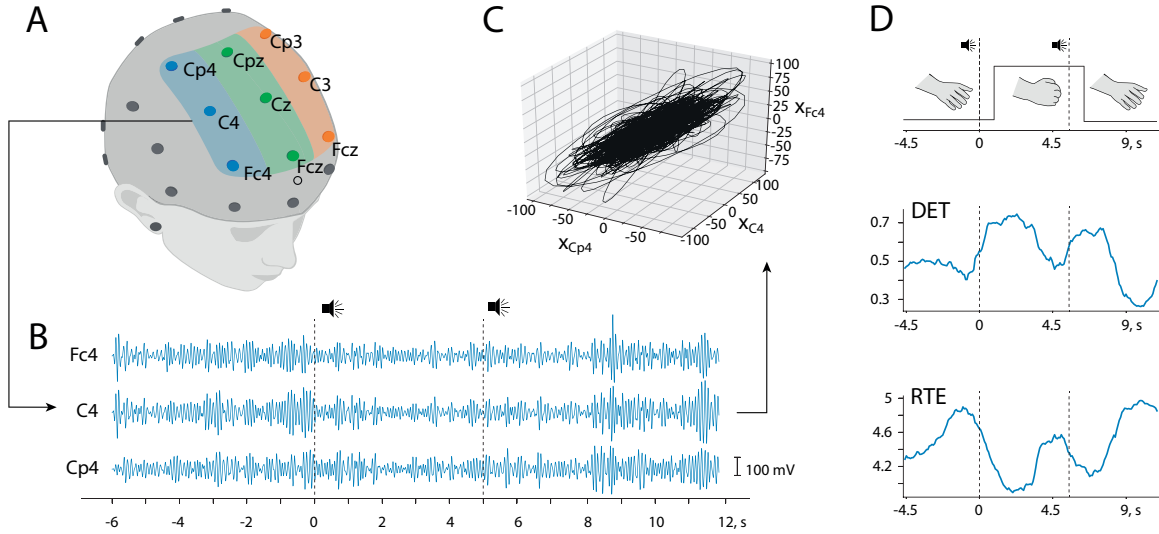
Figure 9.2: Step-by-step visualization of the EEG signal analysis. **A** EEG electrodes, located at the sensorimotor area, forming a multivariate set $\mathbf{X}(t)$: right hemisphere (subset $\mathbf{X}_R(t)$, blue area), left hemisphere (subset $\mathbf{X}_L(t)$, orange area), and longitudinal fissure (subset $\mathbf{X}_F(t)$, green area). **B** Example of a motor-related EEG trial from $\mathbf{X}_R(t)$ (left hand movement). Vertical dashed lines correspond to the first and second audio signal at 0 s and 5 s, respectively. **C** Representation of the current trial from $\mathbf{X}_R(t)$ as a trajectory in 3D phase space. **D** Illustrative scheme of the movement execution accessed from a EMG signal (top panel) and corresponding time-dependent measures of *DET* (middle panel) and *RTE* (bottom panel).

with motor-related brain activity. EMG recordings were also band-pass filtered (10–100 Hz) to capture pronounced high-frequency fluctuations of muscle activity caused by muscle tension during movement execution (Fig. 9.1C). The latter allows to determine exact times for the beginning and the end of movement executions and to study the motor-related brain activity at these specific intervals. Finally, the bandpass filtered time series (both EEG and EMG) were split into 60 trials, each lasting 18 seconds (6 seconds before and 12 seconds after the command, totaling 4500 data points), i.e. 30 attempts for the left and right hand.

The considered EEG trials represent 9-dimensional multivariate sets $\mathbf{X}(t) = (x_{Cp4}(t), x_{C4}(t), x_{Fc4}(t), x_{Cpz}(t), x_{Cz}(t), x_{Fcz}(t), x_{Cp3}(t), x_{C3}(t), x_{F3}(t))^{\mathrm{T}}$ composed of EEG signals recorded over the sensorimotor brain area (Fig. 9.2A). To describe brain dynamics in three areas of interest we separated $\mathbf{X}$ into three 3D subsets, according to their location on the scalp (Fig. 9.2B):

1. left hemisphere (LH): $\mathbf{X}_L(t) = (x_{Cp4}(t), x_{C4}(t), x_{Fc4}(t))^{\mathrm{T}}$;

2. right hemisphere (RH): $\mathbf{X}_R(t) = (x_{Cp3}(t), x_{C3}(t), x_{Fc3}(t))^{\mathrm{T}}$;

3. longitudinal fissure (F): $\mathbf{X}_F(t) = (x_{Cpz}(t), x_{Cz}(t), x_{Fcz}(t))^{\mathrm{T}}$.

Consequently, from a physical perspective, each brain area is represented by a three dimensional trajectory, treating the constitutive time series as state variables (Fig. 9.2C). Since the chosen time series forming these trajectories can possibly carry redundant information, we constructed subsets of linearly independent components $\mathbf{X}'_L(t)$, $\mathbf{X}'_R(t)$ and $\mathbf{X}'_F(t)$ by using principal component analysis (PCA) on $\mathbf{X}_L(t)$, $\mathbf{X}_R(t)$ and $\mathbf{X}_F(t)$. This way of state space trajectory construction

is convenient in the context of multivariate EEG analysis and circumvents the single variable embedding problem [47, 141, 182, 183].

Particularly noteworthy is the fact that the further analysis dealt with EEG recordings at the sensor level. This has been done in order to exclude EEG pre-pocessing steps related with source reconstruction and capture general effects of motor-related activity from the viewpoint of an overall decrease of complexity of the underlying neuronal processes in the sensorimotor cortex. The advantages and limitations of such an approach will be discussed in Section 9.3.

9.2.5 Recurrence quantification analysis

Being a fundamental property of most dynamical systems, recurrence implies that the system's state repeats itself in time [202]. It is represented as neighboring points (states) of the system's trajectory in its state space. A common way of visualizing the system's repeating states is the recurrence plot (RP), Eq. (1.2), which can show structures such as diagonal and horizontal/vertical lines and areas of different recurrence densities [209]. Certain structures are related to the system's complexity and recurrence quantification analysis (RQA) was introduced to analyze them numerically, using various measures of complexity [321] (c.f. Appendix A).

To analyze the recurrence structures in the selected brain areas, using the multivariate set X of length $N = 4500$ (Section 9.2.4), we created a binary recurrence matrix. The recurrence threshold ε determines the size of the neighborhood in state space in which states being considered as recurring [209], see also Chapter 4. When analyzing an RP one should take into account that the obtained results can crucially depend on the choice of this threshold. To provide a robust representation of the RP and ensuring comparability within the samples, i.e. data from different participants, we determined the value of the threshold ε for each sample as the 3rd percentile of the pairwise distance distribution, following Kraemer et al. [163], Chapter 4.

To access time-dependence of the estimated RQA quantifiers we used a running window along the main diagonal line of each RP with a window size of $w = 750$ data points ($3\,\text{s}$) and a shift $\delta w = 20$ data points ($0.08\,\text{s}$).

In the current study we want to quantify regularity and complexity of EEG signals affected by motor tasks execution. Therefore we picked two suitable RQA quantifiers, namely the determinism (DET), Eq. (A.6), and the recurrence time entropy (RTE), Eq. (A.6) (see Appendix A). The presence of diagonal lines in the RP is an important indicator of a deterministic process, since in this case, trajectories at different points in time evolve in a similar manner. More correlated and regularized processes are characterized by longer diagonal lines and less isolated points, i.e. higher DET values.

Along with DET, we estimated RTE, which is a complexity measure based on the "white" (non-recurrent) vertical lines in a RP, which indicate recurrence times. RTE is well suited for capturing the transitions between periodic and chaotic dynamics (and vice versa), because it is related to the Kolmogorov-Sinai entropy [15]. A regular process results in low RTE values, whereas a chaotic process increases the number of different recurrence times, thus, increases its distribution and consequently increases the corresponding RTE value.

Typical single trial time series of DET and RTE are shown in Fig 9.2D. The increase of DET and the decrease of RTE associated with two motor actions following the corresponding audio signal are clearly discernible. A detailed discussion of RQA results will be given based on the

between-subject analysis in Section 9.3. For each participant, we averaged the RQA time series over the trials and excluded the baseline level (3 s prior the first audio signal):

$$\Delta DET(t) = DET(t) - DET_{bckg},$$
$$\Delta RTE(t) = RTE(t) - RTE_{bckg}, \tag{9.1}$$

where DET_{bckg} and RTE_{bckg} are mean values of DET and RTE 3 s prior the first audio signal.

All RQA related computations were performed using the *DynamicalSystems.jl* software library for Julia programming language [62].

9.2.6 Statistical test

The motor-related changes of the RQA measures calculated at the different (area, time)-pairs have been treated as different aspects of the data with respect to which the experimental conditions (motor-task vs baseline and left vs right limb movement) will be compared. Each (area, time)-pair is tested via statistical t-test. Since we do not know exactly the locus of the possible differences in the (area,time)-domain, the multiple comparisons problem (MCP) takes place. To control family-wise error rate (FWER) and address MCP we used a non-parametric statistical test based on random partitions following Maris and Oostenveld [201].

9.3 Results and discussion

To address the main research question of the current study, namely the quantification of an expected reduction in the complexity and randomness of neuronal processes in the sensorimotor cortex in the execution of motor tasks, we considered general cross-subject effects of motor-related changes in the corresponding RQA time series.

First, we analyzed the transition from (random) background neuronal activity to brain activity in the accomplishment of motor tasks. Figure 9.3 shows the results of DET and RTE, averaged over the subjects and along with the standard error, for the movements of right (A, B) and left (C, D) hands. As noted in Section 9.2.4, we were particularly interested in differences in results regarding the right hemisphere ($\mathbf{X}_R$, blue), the left hemisphere ($\mathbf{X}_L$, orange) and the longitudinal fissure ($\mathbf{X}_F$, green). These results indicate that motion execution is associated with an increase in DET (Fig. 9.3 A,C). In addition, DET takes local maxima near start (approximately 2 s after the first audio command) and end (approximately 7 s after the first command) of the motion execution. The positions of the local maxima can be associated with the hand flex and hand relaxation that are performed after the first and second audio commands, while the DET values decrease when holding the hands in a compressed state. The execution of motor tasks in this experimental setup is thus characterized by the pronounced local increase of DET. In fact, the growth of DET implies a predictable or regular motor neuronal activity. This finding is consistent with our hypothesis that motor action is associated with the suppression of random μ-band fluctuations in the EEG that are inherent in background activity. Local peaks of DET are accompanied by a decrease of RTE (Fig. 9.3 B,D). This also shows that the underlying motor neuronal activity recorded by the EEG becomes less chaotic and complex.

Second, we applied a non-parametric statistical test in the (area,time)-domain (cf. Section 9.2.6), which revealed that the described motor changes of the RQA quantifiers compared to the background activity are significant ($p < 0.05$) for all three considered areas during the duration of the motor task execution. The occurrence of significant changes about 1.5 s before and after the
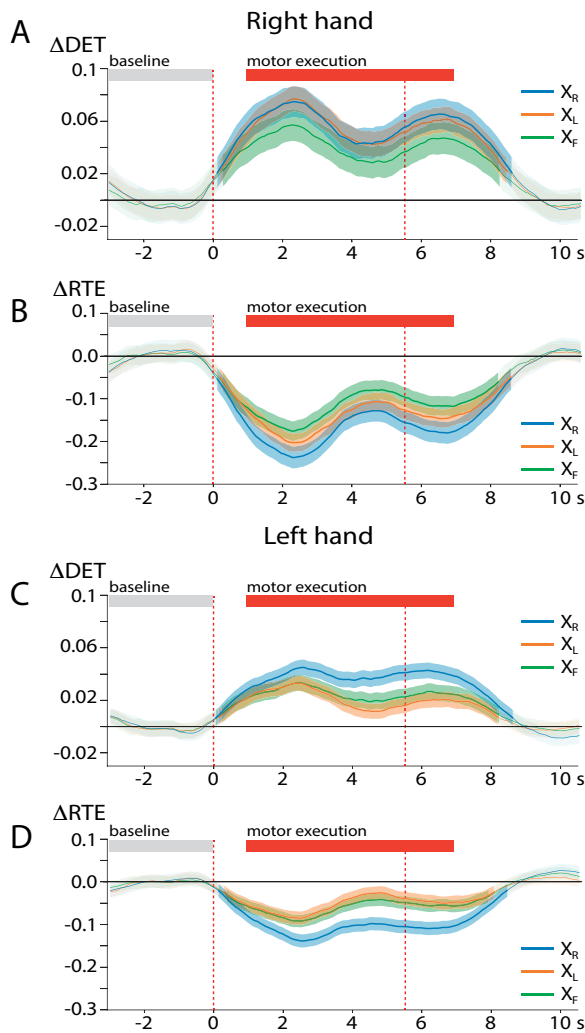
Figure 9.3: Time dependence of ΔDET and ΔRTE averaged over all subjects ($\pm$SE) for the right hemisphere ($\mathbf{X}_R$, blue), the left hemisphere ($\mathbf{X}_L$, orange) and the longitudinal fissure ($\mathbf{X}_F$, green) in case of the right (**A, B**) and left hand movements (**C, D**), respectively. Bold areas highlight the time intervals of significant divergence from the baseline level ($p < 0.05$, MCP corrected via a non-parametric statistical test). In each panel the red dashed lines indicate the moments of the first (0 s) and the second (5.5 s) audio signal and the black horizontal line corresponds to zero level. Gray boxes show 3 s baseline interval before the first audio signal and red boxes show the interval of movement execution obtained from averaged EMGs.

execution of the motor task was related to the half width of the selected window size of $w = 3$ s. It is noteworthy that the RQA measurements returned to the background level after the end of the movement, which is a clear indication of the backward transition of the neuronal dynamics into the background mode.

The RQA time series for the right and left hands had a similar qualitative time course (two maxima/minima associated with hand flexion and relaxation) but took on different absolute values. It can also be seen that the changes of *RTE* and *DET* were more pronounced for the right hand in the considered brain regions (see Fig. 9.3). The movement of the right hand almost equally affected all areas of the sensorimotor cortex without visible contralateral effects. In contrast, left hand movement reduced the complexity of neural dynamics in the right hemisphere more (blue curve in Fig. 9.3 C,D), which coincides with the known contralaterality of the brain's motor-related activity. These results may be biased by the specificity of the participant sample, as only right-handed subjects were selected to participate in the experiments. For right-handed participants, we presume that the right-hand movement is accompanied by a greater neural response compared to left-hand movement.

This leads to the following effects at right-hand movement trials: (i) greater changes in estimated
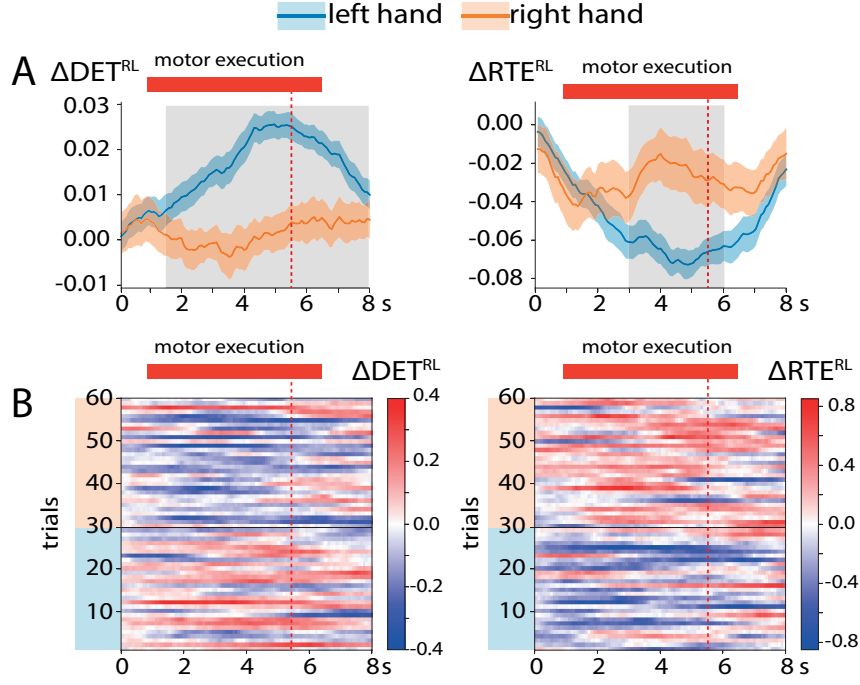
Figure 9.4: A Time dependence of ΔDET^{RL} and ΔRTE^{RL}, derived from EEG data of the right and left hemispheres (Eq. (9.2), see text for details). Measures are averaged over the subjects and displayed as mean±SE. Shaded areas mark the areas with significant differences between the time series corresponding to left and right hand movements ($p < 0.05$, MCP corrected via non-parametric statistical test) and the red boxes indicate the movement execution interval determined from averaged EMGs. **B** Exemplary representation of motor-related EEG samples analysis on an individual test level using ΔDET^{RL} and ΔRTE^{RL}. EEG experiments are arranged as follows: left-hand movements (trials 1-30, highlighted in blue) and right-hand movements (trials 31-60, highlighted in orange).

RQA measures in terms of absolute value and (ii) less pronounced interhemispheric asymmetry due to the volume conduction/field spread effect, which is critical for non-invasive measurements [273]. The latter effect may also be due to the fact that the RQA was performed in the EEG at the sensor level without additional preprocessing, which is aimed either at the reconstruction of the source or at the spatial filtering. Despite these limitations, our results suggest that analyzing the complexity of neuronal dynamics using RQA makes it possible to distinguish between the further discussed lateral types of motor actions.

Finally, we take a closer look at the differences in brain dynamics during right and left hand movements with respect to contralateral effects. For this analysis we only used the $\mathbf{X}_R$ and $\mathbf{X}_L$ records and introduced a measure of asymmetry as the differences between RQA measures in the right and left hemisphere for both hands:

$$\Delta DET^{RL}(t) = \Delta DET^{R}(t) - \Delta DET^{L}(t)$$
$$\Delta RTE^{RL}(t) = \Delta RTE^{R}(t) - \Delta RTE^{L}(t),$$

(9.2)

where superscripts R and L indicate right and left hemispheres, respectively. Figure 9.4 A

shows the course of $\Delta DET^{RL}(t)$ and $\Delta RTE^{RL}(t)$ during the motion execution, averaged over all subjects and along with the standard error. Here, the first audio signal corresponds to the time $t = 0$. In addition to the previous results, Fig. 9.4 shows that the reduction of the complexity of the neural dynamics during the execution of left and right hand movements is different. A comparison of ΔDET^{RL} and ΔRTE^{RL} for left- and right-hand movements using a non-parametric statistical test showed that both measures reflect significant differences between the types of movement (α=0.05, Section 9.2.6). Specifically, reducing the complexity of the underlying neuronal dynamics results in pronounced interlateral asymmetry during left-hand movement, which is reflected in a maximum at DET and a minimum at RTE. The movement of the right hand is in turn linked to the interlateral symmetry – both ΔDET^{RL} and ΔRTE^{RL} fluctuate around zero-level. Note that there is a broader time interval of significant differences of right- and left-hand motion related ΔDET^{RL} time series than it is for the ΔRTE^{RL} time series (Fig. 9.4A). In the case of the former, the brain dynamics associated with left and right hand movement may be discriminated based on a statistical test (cf. Section 9.2.6) at an interval that approximately covers the motion execution (1.5–8 s after the first audio signal). Using ΔRTE^{RL}, this happens at a much narrower interval (3–6 s after the first audio signal). Therefore, the determinism regarding the properties of the interlateral asymmetry seems to be more sensitive than the recurrence time entropy and thus represents a better classification measure. In fact, the disclosed properties of motor-related EEG samples associated with contralateral asymmetry are suitable for a single-trial analysis and classification. Figure 9.4 B shows the exemplary representation of RQA applied to individual EEG experiments collected from a randomly selected subject. It can be seen, that the chosen RQA quantifiers are able to clearly distinguish between left- and right-hand movements: the former are characterized by the positive $\Delta DET^{RL}(t)$ and simultaneously negative $\Delta RTE^{RL}(t)$ values and vice versa for the latter.

Our results indicate that the discussed features of motor neuronal activity detected by EEG signals at the sensor level through RQA complexity measurements are clearly observed and well reproduced in the experimental group under consideration. More importantly, the generality of the cross-subject analysis also applied to the single trial analysis (see exemplary illustration in Fig. 9.2E). The latter, together with the low computational cost of RQA algorithms, offers the prospect of their application in EEG-based BCIs for motion control and assessment.

9.4 Conclusion

We have used RQA to study features of motor-related neuronal processes that have been measured by non-invasive EEG. Our study focused on the consideration of time-dependent RQA quantifiers based on diagonal lines (determinism, DET) and non-recurrent vertical lines (recurrence time entropy, RTE). These measures have been shown to be suitable for detecting transitions between regular (periodic) and irregular (chaotic) dynamics and for quantifying the complexity of the system under study. Both quantifiers clearly showed that the direct execution of motor tasks is associated with a large increase in the regularity of the EEG signals, i.e. a reduction in the complexity of underlying motor-related neuronal processes. In other words, RQA has shown that μ-band ERD causes a reduction in random fluctuations in neuronal activity inherent in background brain activity, leading to more regular behavior of the EEG signal during motor task execution. In addition to the detection of an increase in motor-related regularity of brain dynamics, DET and RTE measurements proofed to be sensitive to indicate the difference between

two lateral types of motion due to the inherent differences in neuronal response in the group of right-handed participants. Specifically, we observed a much greater increase in regularity with right-hand movements, which was distributed almost equally across the sensorimotor areas compared to movements with the left hand. Despite the limitations of EEG analysis at the sensor level, such as volume conduction/field spreading effect, interhemispheric asymmetry was identified in left hand movements in the right-handed group of participants, which was also supported by significant differences between two types of performed movements. In the latter context, it is important to mention that the *DET* measure allowed a much clearer distinction between brain activity associated with left and right hand movements. The limitations of the used RQA are the free parameters for constructing the RP (recurrence threshold) and its quantifying statistics (minimal considered line length), as well as a robust state space reconstruction of the data. Because we used the RQA measures as relative statistics, the particular choice of the recurrence threshold was not so important, since we ensured that the same threshold selection method has been used for obtaining the corresponding RPs and the same minimal line length has been considered for the different trials (c.f. Chapter 4). In this study we circumvented the problem of state space reconstruction by treating single EEG-channels as state variables. However, recent studies indicate that the optimal state space reconstruction can be robustly achieved in an automated fashion [162] (Chapter 2), which would further reduce the effort of pre-processing the data.

Overall, the current results are consistent with and complement the well known concepts of motor-related brain processes. We suppose that the discovered features of neuronal dynamics in the sensorimotor cortex and the robust RQA methods of identification and classification will contribute to the study of non-invasive EEG-based BCI development for motor control and rehabilitation [98, 156, 179, 191].

Part V

Discussion

Chapter 10

Synthesis, conclusion and outlook

This dissertation was concerned with the further development of Recurrence Analysis (RA). The goal of the work was the automation of state space reconstruction, which is fundamental for RA, the development of correction schemes for biased RQA statistics as well as their further development and application in climate and neuroscience.

10.1 How everything fits together

The way this thesis is structured and how each article contributes to this framework are part of the main work flow in (nonlinear) time series analysis (Figure 10.1). Typically there is some data related to one or more research questions and there is a variety of statistical tools the practitioner can use, in order to answer these questions (**application**). Occasionally, however, established methods cannot provide a satisfactory answer to the relevant research questions. In this case, appropriate methods need to be developed (**methodology**) and communicated. This process is reflected in the presented work, which focuses on the methodological aspect.

In the following a synthesis of the different articles ((P1)-(P8), Fig. 10.1) is drawn along with their relations to the overarching research questions (R1)-(R5) raised in Section 1.3. A dense conclusion is then given in Section 10.2. A discussion on remaining open questions and an outlook to further work concludes the thesis in Section 10.3.

Together with other preprocessing steps like noise-/ frequency-filtering, interpolation etc., state space reconstruction most often marks the very first step in nonlinear time series analysis [1, 146]. This abstract representation of the data is key to a subsequent Recurrence Analysis (RA) and potentially affects its results. Its investigation accompanied me during the entire phase of my doctoral research and therefore it constitutes Part I of this thesis. During the years I repeatedly asked myself: When there is a lot of information from experiments, for instance, datasets stemming from micro X-ray fluorescence scanning of a drilling core (Chapters 7, 8) or a multi-sensor EEG (Chapter 9) or several sensors on an aircraft turbine measuring different physical variables, why should I not include these information in my state space model of the data? Of course, Takens' theorem guarantees that this is not necessary for infinitely long, noise free data, but apparently this is not what the researchers and engineers I met were facing. Additionally, when using the traditional embedding methods two free parameters need to be adjusted, the uniform time delay and a threshold for the false nearest neighbor statistic (respectively a slope-threshold in case of Cao's method). Please recall that for clean noise free data acting on a single fundamental
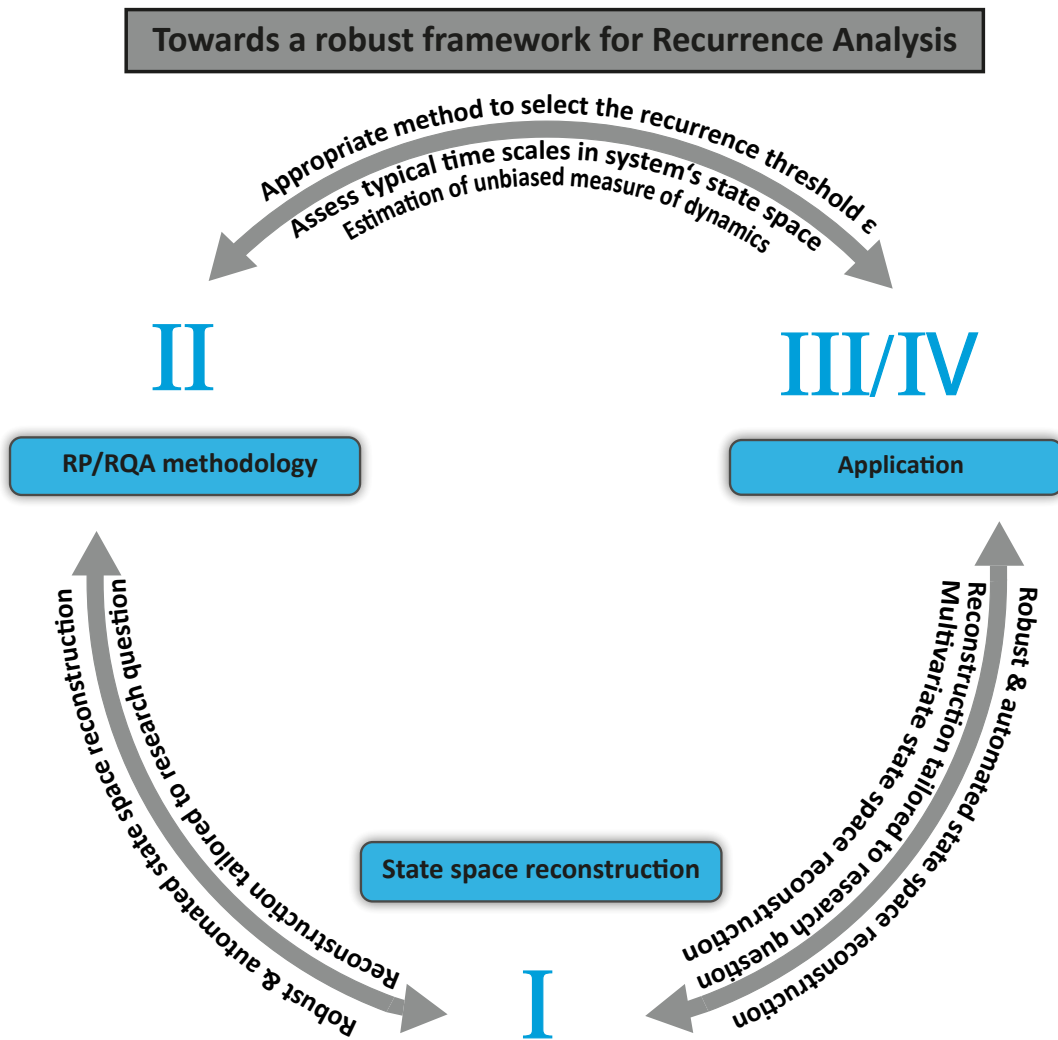
Figure 10.1: Synthesis of the major parts in this thesis. The chapters that make up the parts formulate the interactions shown (see also Fig. 1.3).

timescale, or on timescales which are integer multiples of it, the threshold for the FNN statistic is simply zero and the time lag estimates converge for the many different approaches. In all other cases, I found it advantageous to develop a method that works almost autonomously (P1, Chapter 2). One of the aims was to enable researchers from other scientific disciplines who had never heard of the concept of state or phase space nor of any false nearest neighbor statistics to use state space related analysis tools such as RA. In contrast to the established methods of estimating a single delay and using multiples of it for the the reconstruction, I was firstly looking for a statistic, which was able to quantify the functional dependence between 1-dimensional time series (the potential component in the reconstruction vectors) and higher dimensional time series (the actual reconstruction vectors). The quasi-independence of consecutive components in the reconstruction vectors is essentially what matters in the embedding process, because *redundancy* and *irrelevance* should be avoided. This way an embedding incorporating multiple time series could be included and one could account for different timescales by choosing different delay values.

The continuity-statistic by Pecora et al. [240] was the most promising candidate (Section 2.2.1). We addressed relevant open questions, i.e. parameter pre-selection and the question which peak to pick from the statistic, inter alia by incorporating an evaluation statistic, which was able to quantify the goodness of a reconstruction. We picked the *L*-statistic by Uzal et al. [315], because it is a generalization of the famous FNN statistic, which is undoubtedly a very good concept. Summarizing, our proposed PECUZAL algorithm (P1) could answer research question (R1) and supports near-automatic state space reconstruction from multivariate data. The implementation of the algorithm in three common coding languages (Appendix C.1) shall further contribute to its use beyond the physicist community.

(P7) in Chapter 3 is a straightforward generalization of the findings in (P1). We used the machine learning concept of randomly sampling a decision tree and translated the embedding problem into that framework (Figs. 3.1, 3.2). Specifically, the aim was to ensure that the objective function used for the embedding, which is the *L*-statistic in the PECUZAL algorithm, would certainly reach a global minimum in the parameter space. In this concept the different delays and time series for each embedding cycle act as parameters. In this regard we generalized the whole PECUZAL idea and allowed for different delay pre-selection statistics (in the PECUZAL algorithm this is the mentioned continuity statistic) and different objective functions. We find this modular concept very tempting, particularly because it allows to choose the mentioned statistics according to the research question (e.g., prediction). Our research is in line with findings of other researchers and underpins the need for a state space reconstruction tailored to the specific research question. Hence, (P7) further contributes to answer research question (R1) in the sense that a complete automation of the embedding process independent from the research question is not meaningful.

The recurrence threshold ε is the most important and influential parameter in RA. Consequently, (P2) in Chapter 4 investigated the different methods of setting a fixed size recurrence threshold. There are two ways of setting ε: Either ε is a fixed distance, which is applied to all pairwise distances in the distance matrix *or* ε varies for each point on the state space trajectory. That is, for each column in the distance matrix ε takes a different value, such that the number of recurrence points in each column is constant (adaptive threshold selection, fixed amount of neighbors - FAN). In the former case several ideas had been published which relate the way of setting the fixed threshold to some property of the data, e.g. to the diameter of the state space trajectory. In (P2) we elaborated on – and mathematically substantiated – why it makes sense to set ε to some percentile of the distribution of the pairwise distances (Section 4.2, Fig. 4.1). In contrast to the established approaches which can yield meaningless results for highly embedded data this works well independent of the embedding dimension. Thus, the finding of (P2) clearly justifies why practitioners should either use the suggested approach or the adaptive FAN threshold selection. A fixed threshold accentuates transitions in the RP, while the FAN method balances density fluctuations of points in state space. The appropriate method and the specific percentile to actually choose (or number of nearest neighbors, respectively) depends on the research question. Moreover, both methods ensure comparability of different RPs, since the global RR is fixed. This is of utterly importance in many RQA applications (e.g. Chapter 9) and helps to answer the questions whether RA can be used for the classification of climate states or neuronal activity related to motor task accomplishments, (R4) and (R5). It is also necessary for any meaningful RQA statistic and is therefore crucial for answering (R2).

For a better understanding of abrupt transitions and regime changes via RA (R2) it is inevitably necessary to gain knowledge of the physical meaning of the various established RQA measures. Some of them are purely heuristic, for others relationships to dynamical invariants have been shown. In Chapter 5 (P3) we focused on a deeper understanding of diagonal line based RQA statistics and examined their biases, which originate from the finite size of the considered RP and *tangential motion.* Tangential motion means the thickening of diagonal lines in a RP, which often occurs in highly sampled flow-data (Fig. 5.6). We proposed three novel correction schemes for the finite size correction and a more sophisticated parameter free skeletonization scheme for RPs, in order to tackle the tangential motion problem. A systematic comparison of these new approaches with ideas from the literature revealed the efficiency of the proposed correction schemes. Coming back to the interpretation of RQA statistics, we showed that the entropy of diagonal lines (*ENTR*) behaves in the expected manner only when the proposed correction schemes were applied. In particular, *ENTR* approached zero values for regular limit-cycle dynamics and non-zero values for chaotic dynamics (Figs. 5.8, 5.9). As a by-product the minimal line length for which *ENTR* approaches zero values in regular dynamics is an indicator of the noise level of the system under study.

To further address (R2), a novel idea based on an existing RQA characteristic, namely the diagonal-wise recurrence rate of a RP (τ-RR), has been developed in (P8). The idea of transforming the τ-RR into its frequency domain and consequently obtaining a generalized, nonlinear frequency analysis of a complex, potentially high dimensional system had been published already in 2008 [338]. In that work the idea was to Fourier transform (FT) τ-RR, which seemed to be a straight forward idea. The problem with this approach arises due to the usually spike-train like appearance of τ-RR, for which the FT also yields a spike-train like frequency portrait with many frequencies excited (Figs. 6.1D, 6.2B, E). To overcome this problem we developed a novel way of transforming any arbitrary signal into its frequency domain. Instead of decomposing a signal into trigonometric basis functions as it is done in the FT, we decomposed the signal into a set of basis functions, which are spike-trains (Dirac-combs) with different inter-spike periods. Unlike in the FT-case, this set of basis functions is not linearly independent and thus we had to solve an under-determined linear system, which we have done by applying a common machine learning regression technique (LASSO). Independently from the actual idea to transform the τ-RR of a system, we have developed a novel frequency representation for arbitrary signals. This novel idea, the spike powerspectrum, is superior to the FT in case the signal has a spike-train like shape. The spike powerspectrum gives a sparse representation of the excited frequencies in state space (Figs. 6.2C, F, 6.4J, K, L). Despite its computational limitation we think this technique by itself could be beneficial for researchers, which commonly face spiky signals (e.g. in the Neuroscience or when analyzing event-data in Hydroscience). When applied to the τ-RR of a system, this novel frequency representation showed its true potential. Single, dominant frequencies in state space can be represented in a spike powerspectrum, which is not the case when simply FT the underlying time series. We also presented the idea of incorporating surrogate τ-RRs, which enabled us to judge peaks in the obtained spike powerspectrum by a statistical test and answers (R3), see Figure 6.5. Also (R4) seems to be addressed in this work, because different dynamical regimes operate on different characteristic state space frequencies. This can also be inferred from the presented spike spectrograms (Figs. 6.7, 6.8). These share similar dominant frequencies in similar periods as the FT-based spectrograms, but give additional information.

The question whether RA can be used for classifying climate states and therefore also climate transitions (R4), is investigated in (P4) and (P5) (Chapters 7, 8). In these projects drill core data from lake sediments in the Chew Bahir basin, Ethiopia, were analyzed. (P4) matters, because it acts as an introductory paper discussing RA in the paleoclimate context and analyzed proxy data of six short test sediment cores. (P5) then extended the RA-based climate classification from (P4), see Table 7.1, in the analysis of the long sediment core from the same site. Two types of variability and transitions have been distinguished on the basis of visual appearance of the corresponding RPs and their RQA, see Table 8.1. Hence, (R4) can be answered by (P4) and (P5). Remarkably, this procedure is underpinned by a very similar approach in a recent publication on the Cenozoic Global Reference benthic foraminifer carbon and oxygen Isotope Dataset (CENOGRID) [328] (Fig. 6.6). I have to emphasize that the analysis of the Chew Bahir drill core data mainly motivated the research in (P1), (P7) and (P8). The development of a quantitative, yet intuitive, classifying RQA statistic related to frequency analysis seemed long overdue.

Last but not least, (P6) in Chapter 9 illustrates another application to real data. Here the results from RQA have been evaluated from a statistical point of view in an active neuroscientific experiment and address (R3) and (R5). Motor-task experiments were carried out and the according EEG-data were analyzed by using RA. By comparing the distributions of RQA statistics from the participants it could be shown that the complexity of the signal measured this way via the sensorimotor cortex significantly decreased when motor-tasks were performed (Figs. 9.3, 9.4). The results are consistent with and complement the well known concepts of motor-related brain processes. We suppose that the discovered features of neuronal dynamics in the sensorimotor cortex and the robust RQA methods of identification and classification will contribute to the study of non-invasive EEG-based brain-computer-interface development for motor control and rehabilitation. Hence, we showed that RA can be used for classifying neuronal activity related to certain motor tasks in a non-invasive experimental setup (R5).

In summary, the overarching research questions (R1)-(R5) could be answered. Response limitations are noted for (R2) and (R3). It was not possible to define early warning signatures on RA basis and to develop robust surrogates for recurrence plots, which would allow general significance statements of RQA statistics. In this regard, reference is made to Section 10.3.

10.2 Conclusion

This thesis entitled "Towards a robust framework for Recurrence Analysis: automated state space reconstruction, optimal parameter selection and correction schemes" holds six main advances to the science of non linear time series analysis.

1. It fosters an automated state space reconstruction from observable time series, which is capable of including multivariate data.

2. It emphasizes the importance of an optimal state space reconstruction tailored to the research question at hand.

3. It sheds light on biases of recurrence threshold selection methods and diagonal line based recurrence quantifiers.

4. It uncovers dominant frequencies of (potentially) high dimensional (chaotic) systems.

5. It paves the way for a wider application of Recurrence Analysis in (Paleo-)Climatology, Neuroscience and Engineering.

6. Last but not least it helps practitioners from various scientific disciplines to leverage the extensive methodological toolbox that nonlinear time series analysis has to offer by avoiding the caveats of sub-optimal state space reconstruction, biased recurrence quantifiers and the limitation to linear frequency analysis.

While the thesis, by its methodological nature, could make a contribution to the broader scientific society, it also raises related fundamental methodological questions, which are addressed in the next section. However, this work underscores the importance of an adequate state space reconstruction for a variety of nonlinear analysis tools even beyond those discussed here. In addition, it gives an insight into the power of Recurrence Analysis regarding very different research questions, from inferring causal relationships in complex systems to prediction or to feature-based classification.

10.3 Open questions & outlook

The preceding sections highlighted the advances made in this thesis and their ability to answer the overarching research questions (R1)-(R5). Now we highlight the limitations of our findings and give outlooks for future work.

10.3.1 The embedding problem

Although the findings in (P7) suggest that the optimal embedding varies according to the research question, the PECUZAL algorithm (P1) occasionally gives contradictory embedding suggestions. For example, the Logistic map is clearly a deterministic system and therefore the L-statistic should recommend an embedding with dimension $m > 1$. When analyzed in a chaotic regime, PECUZAL suggest no embedding and treats the input as a stochastic signal, which it does not embed. For other maps, e.g. the Ikeda or Hénon map, this is not the case. Furthermore, this does not seem to be a L-statistic or PECUZAL related problem. When running the "stochastic indicator" proposed by Cao [39], it also values the chaotic Logistic map as a stochastic source and does not suggest any embedding. A similar problem arises when analyzing map-like data in the geoscientific context. These time series are often interpolated and despite their inherent non-stationarity one could try to embed small pieces with approximately constant parameters. In many cases, ranging from drill core data under a certain age model to climate index data such as the Southern Oscillation Index (SOI) and to Earth system models of intermediate complexity (EMICs), PECUZAL does not suggest any embedding and also other stochastic indicators would treat the signals as stochastic. We formulate the following research questions which should be addressed in future work:

- How does interpolation affect the estimation of the embedding parameters?

- How does the sampling affect the estimation of the embedding parameters (flow-like vs. map-like data)?

- Countless real world processes can be described by a Langevin Equation. Yet, to my best knowledge there is no study which systematically investigates the embedding of systems described by such a stochastic differential equation.

- For a Markov process of order p an optimal embedding for prediction should yield a $> p$-dimensional embedding. A first try I made with the PECUZAL algorithm showed that the suggested embedding dimension varies with the time series length and the transition probabilities, i.e. the auto-correlation. Future work should be concerned with the embedding of such processes when using PECUZAL or an MCDTS-ansatz (Chapter 3). A comparison of the suggested embedding dimension with the ones obtained from the Akaike information criterion (AIC) [7] and from the optimal encoding description length minimization [17] would be desirable.

- Related to the last point is the question whether an optimal embedding in the PECUZAL- or MCDTS-sense could lead to improved forecasts for stochastic systems [148] (e.g. in the way Ragwitz and Kantz [250] addressed this problem).

- Since PECUZAL and the MCDTS framework allow the embedding of multivariate data, one could think of an application, which detects synchronization. The idea could be to feed all available time series from two (or more) systems to the algorithm(s). Then the time series used for embedding should differ in different states of synchronicity, because redundancy is avoided. I would expect a multivariate embedding including more available time series in case of non-synchronicity and less included time series for phase- or generalized synchronization.

- The impact of the embedding on state-space-based causality measures such as *convergent cross mapping* [288] and its extensions has not been investigated yet.

10.3.2 Recurrence analysis

The question of how to select the optimal recurrence threshold is still not answered satisfactorily. Concerning (P2), the question is which percentile to choose for a certain research question. Yet, in this problem there are almost no objective methods for any sort of research question, but very many heuristic ideas do exist. A very recent publication by Prado et al. [249] is indeed very promising. The authors estimate the optimal threshold by maximizing the diversity of structures/patterns in the RP. In this framework structures are estimated via recurrence *microstates* (small subbatches of a RP) and the diversity is assessed by maximizing the entropy of their distribution gained from sampling an entire RP. Building on this, the following idea could be addressed in future work:

- Similar to the approach of using the maximum entropy of microstates for selecting the optimal recurrence threshold, one could think of utilizing our proposed skeletonization scheme (Section 5.5.5) for that purpose. For an increasing threshold one counts the number of diagonal lines contained in the corresponding skeletonized RP. If the threshold is too high, distinct lines (*distance ranges* in the jargon of Chapter 5) merge and the skeletonization algorithm will reduce these merged distance ranges to a single diagonal line. Therefore, the optimal threshold would maximize the number of diagonal lines.

As shown in (P3) Fig. 5.11, the analytically derived relation between the diagonal line length entropy (*ENTR*, Eq. (A.9)) and the correlation entropy given by March et al. [199], Eq. (5.14),

does not seem to hold. Neither for the uncorrected, nor for the corrected *ENTR*. In experiments that I have done that are not shown in this thesis, I have not been able to verify the proposed relation regardless of the many different systems and many different parameter settings that I have tried.

- A thorough study on the impact of the correction schemes in (P3) on the estimation of dynamical invariants by using line based RQA characteristics is needed, in order to make advances in transition and bifurcation detection as well as classifying regimes.

- Furthermore, the impact of the correction schemes in (P3) on the τ-RR and its corresponding spike powerspectrum (P8) should be investigated.

In Chapter 9 we were able to make statements about the significance of the results, because the experimental design allowed the acquisition of distributions of RQA characteristics. In most passive experiment setups, as it is most often the case in Geoscience, this is not possible. It would, thus, be beneficial to construct surrogates of state space trajectories in order to obtain distributions of corresponding RQA statistics, which could then be used for statistical testing. Thiel et al. [296] proposed the method of *twin surrogates*, which constructs surrogates from 1) identifying twins in the state space trajectory (=points which share the same neighborhood) and 2) randomly jump to one of the possible futures of the existing twins. The drawback is of course, that for proper statistical testing one would seek around $1,000$ surrogates or more and in the described method this number is determined by the total number of twins, which is a property of the data. Another idea for line based RQA statistics has been proposed by Marwan et al. [210] for running window approaches. To estimate the unknown variance of the diagonal line length distribution of a RP, they bootstrapped surrogate line length distributions from the cumulative line length distribution of all windows. This idea is very promising, but the resulting confidence intervals seem to be extremely sensitive to the number of lines one bootstraps. There is no objective way to determine this number and thus there remains quite some ambiguity in this method. The novel RQA statistic proposed in (P8), the spike powerspectrum of the τ-RR, would also benefit from robust RP-surrogates. However, since it is based on the τ-RR we were able to construct RP-surrogates for a simple Null-model (=system is completely random), because in this case recurrent points in the RP follow a binomial distribution, with parameter p being the global recurrence rate of the RP.

- There is an urgent need for a robust method, that constructs RP-surrogates, which preserves basic properties (correlation structure) of the RP and/or the underlying state space trajectory. This would affect all existing RQA measures and would allow to make statements about the statistical relevance a measured RQA statistic has, even in passive experiments with single runs.

In a more general sense, the following statements could lead to research questions in future work:

- The findings from (P1), (P3), (P7) and (P8) should be applied in Climatology, Neuroscience & Engineering. In particular, Fourier based frequency analysis is common in all of these fields. The spike powerspectra and spectrograms of the τ-RR from (P8) could offer an intuitive additional source of valuable information.

- In the Climate context the findings from (P1) & (P7) would allow to make state space models from various available proxy data, which potentially describe the same system. In

the Neuroscience context one could think about constructing a state space model from various EEG-channels. Both ideas need to be investigated and compared to the established methods.

- The coincidence of the same dominant frequencies in the powerspectra of the τ-RR of two (or more) coupled chaotic systems should be a sign of phase-, generalized- and complete synchronization [211, 244, 259].

Final remarks

The previous section supports George Bernard Shaw's view that "Science never solves a problem without creating ten more". Nevertheless, in the larger context my dissertation improved the recurrence analysis framework. I hope it helps to make RA more accessible to researchers outside of physics and I am extremely curious to see how this interesting and active field of research will develop in the future.

Part VI

Appendix

Appendix A

The recurrence plot and its quantification measures

In the following we extend the brief introduction to recurrence plots (RPs) given in Chapter 1 to a more technical consideration. Introduced by Eckmann et al. [77], RPs provide a versatile tool for visualizing and quantitatively analyzing the succession of dynamically similar states in a time series (e.g., Figs. 1.2, 5.2, 7.2). For this purpose, dynamical similarity is measured in terms of some metric distance $d_{i,j} = \|\vec{x}_i - \vec{x}_j\|$ defined in the underlying system's d-dimensional state space. The vector time series $\{\vec{x}_i\}_{i=1}^N$ (with $\vec{x}_i = \vec{x}(t_i)$) can be obtained from direct measurements of the system under study or from embedding a measured subset of observables (c.f. Chapters 2, 3).

Based on the resulting distance matrix $\mathbf{d} = (d_{i,j})$, a recurrence matrix $\mathbf{R} = (\mathbf{R}_{i,j})$ is defined as a thresholded version such that its entries assume values of 1, if the distance between the two associated state vectors is smaller than or equal to a threshold ε, and 0 otherwise. It is, thus, a binary, square matrix:

$$\mathbf{R}_{i,j}(\varepsilon) = \begin{cases} 1: & d_{i,j} \leq \varepsilon \\ 0: & d_{i,j} > \varepsilon, \end{cases} \qquad i,j = 1, ..., N \tag{A.1}$$

Equivalently, we can write

$$\mathbf{R}_{i,j}(\varepsilon) = \Theta\left(\varepsilon - \|\vec{x}_i - \vec{x}_j\|\right) \tag{A.2}$$

$$= \Theta(\varepsilon - d_{i,j}), \qquad i,j = 1, ..., N, \qquad \vec{x} \in \mathbb{R}^d, \tag{A.3}$$

where $\Theta(\cdot)$ is the Heaviside function. In this definition, the threshold ε is fixed with respect to all pairwise distances contained in $\mathbf{d}$. Chapter 4 addresses the question how to select this threshold, in order to obtain robust recurrence characteristics for arbitrary systems. An alternative definition of the recurrence matrix [77, 209] replaces the global, fixed recurrence threshold ε applied to all state vectors $\vec{x}_i$ by an adaptive local one that is defined such that the number of recurrences (i.e., close state vectors) is the same for each $\vec{x}_i$ (fixed amount of nearest neighbors - FAN), leading to a constant local and global recurrence rate. This procedure is useful, when there are severe differences in the density of points in state space. In this case lower dense populated areas in the state space are not well resolved in the RP, when ε is not adaptively chosen. Of course, this can be simply a result of under-sampling the system. On the other hand, strong density variations

in phase space can be a meaningful information or it is necessary to ensure a fixed threshold, when comparing RPs stemming from different systems, or from the same system, but at different dynamics. The research question and the quality of the data decides whether to use a fixed recurrence threshold (e.g., Chapter 9) or an adaptive one (e.g., Chapters 7, 8).

Some additional approaches of defining a recurrence plot other than Eqs. (1.1), (1.2) have been proposed. The main purpose of these alternative definitions is to avoid "tangential motion", i.e. a thickening of diagonal lines in the recurrence plot, due to oversampling in flow-like data. We review these ideas in detail in Chapter 5, Section 5.5.

There are several ways to quantify the structures and dynamics encoded in $\mathbf{R}$, the so called recurrence quantification analysis (RQA). The measures can be grouped by their underlying geometrical structure, i.e. the number of points in the RP, diagonal line structures as well as vertical line structures (see Marwan et al. [209] for a detailed overview). Here we limit ourselves to a selection of quantifiers, which are used in this thesis.

The most straightforward idea is to look at the density of recurrences, the *recurrence rate*

$$\mathrm{RR} = \frac{1}{N^2} \sum_{i,j=1}^{N} \mathbf{R}_{i,j} \quad , \tag{A.4}$$

which is nothing else than the probability of a state to occur in a certain time frame, which is covered by the RP. In a mathematical sense this only holds in the limit $N \to \infty$. A natural extension would be to look at the τ-recurrence rate of the RP [206, 338], which is the density of recurrence points along the diagonals of the recurrence matrix, as a function of the distance τ (sampling units) to the main diagonal:

$$\tau\text{-}\mathrm{RR}(\varepsilon) = \mathrm{RR}(\tau, \varepsilon) = \frac{1}{N-\tau} \sum_{i=1}^{N-\tau} \mathbf{R}_{i,i+\tau}. \tag{A.5}$$

τ-RR serves as an estimator for the probability that the system recurs after time $\tau\Delta t$, with Δt being the sampling time of the trajectory $\vec{x}_i = \vec{x}(\Delta t \cdot i)$, $i = 1, \dots, N$.

Diagonal lines typically occur in recurrence plots when one segment of the trajectory runs almost parallel to another segment (e.g. within the recurrence threshold ε) representing an earlier episode in the history of the system within the state space, Fig. 5.1. Diagonal lines in recurrence plots are therefore diagnostic of cyclicity in state space; such cyclicity detected using recurrence plots is not restricted to sinusoidal structures in the underlying time series, in contrast to cyclicity in Fourier-based time-series analysis. There are many ideas of measures based on diagonal lines, which heavily vary in their physical interpretation. For example we can look at the determinism [323], which is the fraction of recurrence points that form diagonal lines of length ℓ

$$\mathrm{DET} = \frac{\sum_{\ell=\ell_{\min}}^{\ell_{\max}} \ell P(\ell)}{\sum_{\ell=1}^{\ell_{\max}} \ell P(\ell)}, \tag{A.6}$$

with $P(\ell)$ being the histogram of all diagonal line lengths ℓ in $\mathbf{R}$, $\ell_{\min}$ the considered minimal line length (set to 2, if not specified differently) and $\ell_{\max}$ the longest line found in the RP. Although deterministic dynamics does not automatically mean that the dynamics is predictable (e.g., for chaotic systems), a predictable system, however, means that the system is deterministic. Moreover,

this definition of determinism is not a strict mathematical one, but more from a heuristic point of view. Two related heuristics for estimating the predictability are the mean diagonal line length

$$\ell_{\text{mean}} = \frac{\sum_{\ell=\ell_{\min}}^{\ell_{\max}} \ell P(\ell)}{\sum_{\ell=\ell_{\min}}^{\ell_{\max}} P(\ell)} \tag{A.7}$$

and the Divergence

$$\text{DIV} = \ell_{\max}^{-1}. \tag{A.8}$$

Another quantifier is the Shannon entropy of the probability distribution $p(\ell) = P(\ell)/\sum_\ell P(\ell)$ to find a diagonal line of exact length ℓ

$$\text{ENTR} = -\sum_{\ell=\ell_{\min}}^{\ell_{\max}} p(\ell) \ln p(\ell), \tag{A.9}$$

which is related to the Kolmogorov-Sinai-entropy [199]. However, it turns out that the distribution of diagonal lines, and thus, $p(\ell)$, is biased by the finite size of $\mathbf{R}$ and by "thickened lines" due to tangential motion. Among all the line based recurrence quantifiers, *ENTR* seems to respond most strongly to the biased $p(\ell)$ and, in turn, shows counter intuitive behavior for paradigmatic examples. In Chapter 5 we investigate this fact and propose correction schemes.

It is also possible to look at the Shannon entropy of the length distribution of the white vertical lines ℓ_w, which correspond to recurrence times. In general, such recurrence times can be estimated directly from the RP in different ways [229], among which the vertical non-recurrence lines offer a particularly simple estimator. The entropy of the distribution of recurrence times, referred to as the *recurrence period density entropy* [187] or *recurrence time entropy* and originally introduced without any direct link to RPs, is given as

$$\text{RTE} = -\sum_{\ell_\text{w}=1}^{\ell_{\text{w}_{max}}} p(\ell_\text{w}) \ln p(\ell_\text{w}). \tag{A.10}$$

Often it is useful to normalize this measure with respect to the largest recurrence time $\ell_{\text{w}_{max}}$ (the same can be done in case of *ENTR*, Eq. (A.9))

$$\text{RTE} = -\frac{1}{\ln \ell_{\text{w}_{max}}} \sum_{\ell_\text{w}=1}^{\ell_{\text{w}_{max}}} p(\ell_\text{w}) \ln p(\ell_\text{w}) \in [0,1]. \tag{A.11}$$

The estimate $p(\ell_\text{w})$ can be derived from the histogram of recurrence times, $P(\ell_\text{w})$, similar to the procedure for the black diagonal lines, as $p(\ell_\text{w}) = P(\ell_\text{w})/\sum_{\ell_\text{w}} P(\ell_\text{w})$, i.e., as the probability to find a white vertical line of exactly length ℓ_w in the RP. It can be shown that *RTE* is closely linked to the Kolmogorov-Sinai (KS) entropy of the system under study [15]. It is also possible to look at the mean recurrence time *MRT*, in order to unravel the characteristic time scale in the underlying RP

$$\text{MRT} = \frac{\sum_{\ell_\text{w}=\ell_{\text{w,min}}}^{\ell_{\text{w,max}}} \ell_\text{w} P(\ell_\text{w})}{\sum_{\ell_\text{w}=\ell_{\text{w,min}}}^{\ell_{\text{w,max}}} P(\ell_\text{w})}. \tag{A.12}$$

Appendix A The recurrence plot and its quantification measures

Analogue to the definition of the determinism, the laminarity is the fraction of recurrence points that form vertical lines

$$\text{LAM} = \frac{\sum_{\ell_v=\ell_{v,\min}}^{N} \ell_v P(\ell_v)}{\sum_{\ell_v=1}^{N} \ell_v P(\ell_v)}, \tag{A.13}$$

with ℓ_v the length of a vertical line, $P(\ell_v)$ the histogram of all vertical lines in $\mathbf{R}$ and $\ell_{v,\min}$ the considered minimal line length (set to 2, if not specified differently). It is assumed that LAM quantifies the presence of laminar states in intermittent regimes and can be used for the detection of chaos–chaos transitions [213].

Finally, the RP can be considered as the adjacency matrix of an ε-recurrence network [340], $\mathbf{A} = \mathbf{R} - \mathbb{1}$ and the transitivity [22] can be defined as

$$\mathcal{T} = \frac{\sum_{i,j,k=1}^{N} A_{jk} A_{ij} A_{ik}}{\sum_{i,j,k=1}^{N} A_{ij} A_{ik}(1 - \delta_{jk})}, \tag{A.14}$$

which is a measure characterizing the geometric structure of the state space attractor [74].

Appendix B

Numerical models

B.1 Lorenz system

The classical Lorenz-63 system [189] is defined as

$$
\begin{aligned}
\dot{x} &= \sigma(y - x) \\
\dot{y} &= x(\rho - z) - y \\
\dot{z} &= xy - \beta z.
\end{aligned}
\tag{B.1}
$$

In Chapter 2 we set the initial condition to $u_0 = [1.0, 1.0, 50.0]$, used a sampling time of $\Delta t = 0.01$ and discarded the first 1,000 points of the integration as transients. For producing Fig. 2.3 we set the parameters to $\sigma = 10, \beta = 8/3, \rho = 60$ and used a time series consisting of 5,000 samples. In Figs. 2.4, C.2, C.3, C.4 and the redundancy fooling-dataset in Section 2.4.2 we use the standard parameter values $\sigma = 10$, $\beta = 8/3$, $\rho = 28$ and also a time series consisting of 5,000 samples.

B.2 Lorenz 96 system

The Lorenz 96 system [188] is defined as

$$
\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F
\tag{B.2}
$$

with x_i the state of the system for nodes $i = 1, \ldots, N$ and it is assumed that the total number of nodes is $N \geq 4$. One can think of this system as a ring-like structure of N coupled oscillators – each representing some atmospheric quantity – all connected to the same forcing. The forcing constant F serves as the control parameter. In Chapter 2 we set $N = 8$, $F = 4.472$ and the initial condition to $u_0 = [0.590; 0.766; 0.566; 0.460; 0.794; 0.854; 0.200; 0.298]$, used a sampling time of $\Delta t = 0.1$ and discarded the first 2,500 points of the integration as transients leaving us with a time series consisting of 5,000 samples. The same setup is being used in Chapter 3, but with the control parameter F varying from $F = 3.7$ to 4.0.

B.3 Rössler system

The Rössler system [261] is defined as

$$
\begin{aligned}
\dot{x} &= -y - z \\
\dot{y} &= x + ay \\
\dot{z} &= b + z(x - c).
\end{aligned}
\tag{B.3}
$$

In Chapter 2 we randomly chose the initial conditions uniformly from an interval $[0, 2]$ and discarded the first 2,000 points of the integration as transients working with time series of length $N = 10,000$ (for Fig. C.1 we set $N = 20,000$). The parameters were set to $a = 0.2925$, $b = 0.1$, $c = 8.5$ (funnel regime) for the experiment in Section 2.4.2, where the sampling time was set to $\Delta t = 0.03$ and $a = 0.2$, $b = 0.2$, $c = 5.7$ with a sampling time of $\Delta t = 0.02$ elsewhere.

For producing Figs. G.1, 6.4 in Chapter 6 the initial condition for producing panels A & B was set to $u_0 = [.7, -1, 0.4]$ with a sampling time of $dt = 0.05$ and in case of panel C $u_0 = [-0.1242, -2.5415, 0.2772]$ with a sampling time of $dt = 0.1$. The first 5,000 samples were discarded as transients and trajectories of length $N = 5,000$ were obtained from which we computed the RPs and the corresponding τ-RR's. For the spike powerspectra only the first 1,000 values of the τ-RR's were considered.

B.4 Driven Duffing oscillator

The driven Duffing/Van der Pol oscillator [12, 76] is defined as

$$
\begin{aligned}
\dot{x} &= y \\
\dot{y} &= \mu(1 - x^2)y - \alpha x - \beta x^3 + z \\
\dot{z} &= \gamma \cdot \omega \cdot \cos(\omega \cdot t),
\end{aligned}
\tag{B.4}
$$

with $\mu = 0.1$, $\alpha = 1$, $\beta = 0$, $\gamma = 0.5$, and $\omega = 2$ [60] resulting in a regular, quasi periodic motion. In Chapter 2 we randomly chose the initial conditions x_0, y_0, z_0 uniformly from the interval $[0, 0.2]$, used a sampling time of $\Delta t = 0.1$ and discarded the first 2,000 points of the integration as transients, resulting in time series of length $N = 5,000$.

B.5 Mackey-Glass equation

The Mackey-Glass equation [192] is the nonlinear time delay differential equation

$$
\dot{x} = \beta \frac{x_\tau}{1 + x^n} - \gamma x,
$$

with the lag $\tau = 44$, and the parameters $n = 10$, $\beta = 0.2$, and $\gamma = 0.1$. x_τ represents the value of x at time t-τ. In Chapter 2 we randomly chose the initial conditions uniformly from an interval $[0, 1.5]$, used a sampling time of $\Delta t = 0.5$ and discarded the first 2,000 points of the integration as transients, resulting in time series of length $N = 10,000$.

Appendix C

Supplementary material for Chapter 2: Automated attractor reconstruction

C.1 Implementation and code availability

The study that we present here is available as a fully reproducible code base [166]. In addition, we have implemented performant versions of the embedding algorithms, as well as the automated pipeline for optimal embedding. A single package is provided for Python [169] (https://pypi.org/project/pecuzal-embedding/) and a toolbox for MATLAB® [168], (https://de.mathworks.com/matlabcentral/fileexchange/86004-pecuzal-embedding-algorithm-for-matlab) while for the Julia language we have integrated the PECUZAL algorithm into the library DynamicalSystems.jl [62] (the other methods "TDE, Garcia& Almedia, MDOP" were already part of the library). The automated pipeline for optimal embedding has been further refined for better user experience and is also part of DynamicalSystems.jl.
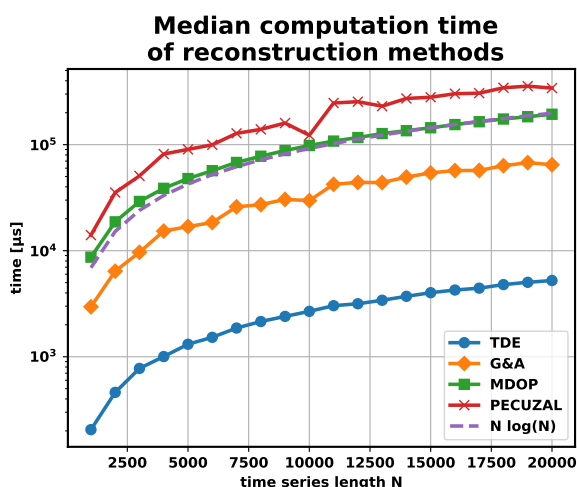
Figure C.1: Median time complexity for the state space reconstruction from the y-component of the Rössler system (Appendix B.3) for TDE, Garcia & Almeida's, MDOP, and PECUZAL method. Ensembles of function calls for each considered time series length N are computed and the median values are shown (implemented in Benchmark-Tools.jl [51]). A Theiler window is set as the first minimum of the auto-mutualinformation and the maximum encountered time delay is set to four times the Theiler window for TDE, G&A, and PECUZAL.

The Julia versions of the algorithms have been heavily optimized and are the most performant ones by orders of magnitude in parts, when compared to the Python (and MATLAB®)

implementation. Due to the nearest neighbor tree search, all compared methods have $\mathcal{O}(N \log N)$ time complexity (Fig. C.1). Our proposed method is slightly (but consistently) slower than the established methods, with standard time delay embedding performing best (Cao's method). Worth mentioning is the fact that in the MDOP case the most computational effort is put on the estimation of the maximum considered delay value and not on the computation of their β-statistic. Considering that an optimal embedding is a one-time operation, we believe that all methods shown here are practically useful with respect to their computational complexity.

C.2 Numerical results

Chaotic Rössler system, driven Duffing oscillator in regular motion, and nonlinear time-delay Mackey-Glass equation are used to evaluate the performance of the different embedding approaches (TDE, G&A, MDOP, and PECUZAL). We consider univariate and multivariate embedding (Rössler and Duffing). The y-component of the Rössler system and the x-component of the Duffing system are used for the univariate embedding, wheres the x- and y-values of the corresponding systems are used for the multivariate embedding. For the TDE only univariate embedding is possible and, thus, there are no results for the multivariate case.

Ensembles of 1,000 trajectories using different initial conditions are used. Additionally, we consider additive noise with an amplitude of 10% of the standard deviation of the corresponding signal. All RPs are computed using a fixed recurrence rate of 8% and a minimum line length of 2.

C.3 Dependency on parameters

We investigate the impact of different parameter settings on the resulting reconstruction of our proposed method for the x-component of the Lorenz system (Appendix B.1). In particular, Figure C.2) shows the sensitivity of the L-statistic value and the chosen delays with respect to the number of nearest neighbors kNN for a fixed parameter $T_M = 20$ (panels A, B) and also the dependence on the continuity statistic parameters δ-Neighborhoodsize (panels C, D), the (binomial) probability p (panels E, F) and significance-level α (panels G, H). The critical dependence of the L-statistic on the parameter T_M is discussed in Section 2.3 and Fig. 2.4. The results show very little impact on the reconstruction quality and, thus, confirm our choice of fixed parameter values for the algorithm. Specifically, the crucial qualitative course of the continuity statistic, i.e., the position of the local maxima, remains unchanged for relevant choices of α and in the vicinity of $p = 0.5$, which is the proposed fixed value (Figs. C.3 and C.4).

Table C.1: Results of the evaluation measures: accordance of the recurrence plot (JRRF), mutual false nearest neighbors (MFNN), ΔL-statistic, as well as the recurrence quantifiers determinism (DET), diagonal line length entropy (ENTR), and recurrence time entropy (RTE) for the standard time delay embedding (TDE), Garcia & Almeida (G&A), Nichkawde's (MDOP), and PECUZAL-method (see text for details), using chaotic Rössler system, driven Duffing oscillator in regular motion (both in univariate (u) and multivariate (m) case), and nonlinear time-delay Mackey-Glass equation. Mean values and standard deviation of ensembles consisting of 1,000 integrated trajectories are shown. Bracketed values present results for additive noise. The results of recurrence quantifiers are presented as relative deviations from the reference. The best results for each considered system are highlighted in black.

System	JRRF TDE	JRRF G&A	JRRF MDOP	JRRF PECUZAL	MFNN TDE	MFNN G&A	MFNN MDOP	MFNN PECUZAL	ΔL TDE	ΔL G&A	ΔL MDOP	ΔL PECUZAL
Rössler (u)	.71±.02	.48±.04	.64±.01	.73±.01	3.8±.3	8.7±2.4	4.0±.4	3.3±.2	−1.27±.09	−1.03±.17	−1.24±.08	−1.41±.08
	(.71±.02)	(.6±.03)	(.62±.01)	(.73±.01)	(4.1±.3)	(5.1±.6)	(4.8±.4)	(3.7±.2)	(−1.14±.05)	(−0.94±.08)	(−0.95±.04)	(−1.23±.05)
Rössler (m)	—	.83±.02	.86±.04	.87±.01	—	1.9±.6	1.4±.1	1.29±.08	—	−1.58±.2	−1.26±.12	−1.48±.08
	—	(.82±.04)	(.87±.01)	(.87±.01)	—	(2.4±.6)	(1.57±.06)	(1.5±.1)	—	(−1.14±.09)	(−1.04±.07)	(−1.28±.04)
Mackey-G	—	—	—	—	—	—	—	—	−.43±.0	−.57±.0	−.35±.0	−.77±.0
	—	—	—	—	—	—	—	—	(.12±.04)	(.12±.07)	($1 \times 10^7 \pm .0$)	(−0.02±.03)
Duffing (u)	.82±.01	.80±.04	.83±.02	.84±.02	4.0±1.2	4.9±1.7	3.9±.9	3.8±.8	−1.66±1.7	−1.68±.18	−1.66±.16	−1.75±.15
	(.74±.01)	(.63±.04)	(.69±.02)	(.73±.02)	(8.9±.4)	(10.2±1.6)	(8.8±.3)	(15.1±3.7)	(−.5±.05)	(−.51±.08)	(−.41±.06)	(−.69±.04)
Duffing (m)	—	.80±.06	.83±.02	.84±.02	—	3.9±1.5	3.1±.8	3.3±.9	—	−1.6±.2	−1.4±.2	−1.8±.2
	—	(.72±.04)	(.74±.02)	(.73±.01)	—	(10.0±1.8)	(7.9±.6)	(14.1±3.9)	—	(−.5±.1)	(−.3±.1)	(−.71±.04)

System	DET (×10³) TDE	DET (×10³) G&A	DET (×10³) MDOP	DET (×10³) PECUZAL	ENTR TDE	ENTR G&A	ENTR MDOP	ENTR PECUZAL	RTE TDE	RTE G&A	RTE MDOP	RTE PECUZAL
Rössler (u)	.03±.01	.71±.28	.03±.01	.01±.01	.05±.01	.32±.03	.18±.01	.08±.01	.03±.01	.07±.02	.07±.01	.04±.01
	(6.6±.8)	(7.2±1.2)	(8.0±.7)	(3.0±.4)	(.14±.02)	(.05±.03)	(.08±.02)	(.08±.01)	(.02±.01)	(.04±.01)	(.04±.01)	(.02±.01)
Rössler (m)	—	.95±.35	.19±.06	.14±.03	—	.2±.04	.09±.02	.08±.02	—	.06±.02	.02±.01	.01±.01
	—	(6.6±2.4)	(4.9±.4)	(2.8±.3)	—	(.05±.03)	(.08±.01)	(.06±.01)	—	(.02±.01)	(.01±.01)	(.01±.0)
Duffing (u)	3.9±.08	3.6±1.0	4.1±.3	4.2±.3	.44±.09	.36±.07	.42±.02	.45±.05	.03±.04	.02±.03	.01±.01	.02±.02
	(97.4±4.3)	(113.8±14.2)	(113.4±15.4)	(61.5±5.7)	(.03±.02)	(.30±.06)	(.37±.05)	(.21±.09)	(.62±.01)	(.32±.16)	(.54±.03)	(.46±.06)
Duffing (m)	—	8.3±19	4.2±.3	4.2±.3	—	.49±.12	.51±.07	.44±.06	—	.16±.57	.02±.02	.02±.03
	—	(113±18)	(103±10)	(61±4)	—	(.32±.07)	(.22±.07)	(.22±.08)	—	(.56±.09)	(.61±.01)	(.44±.06)
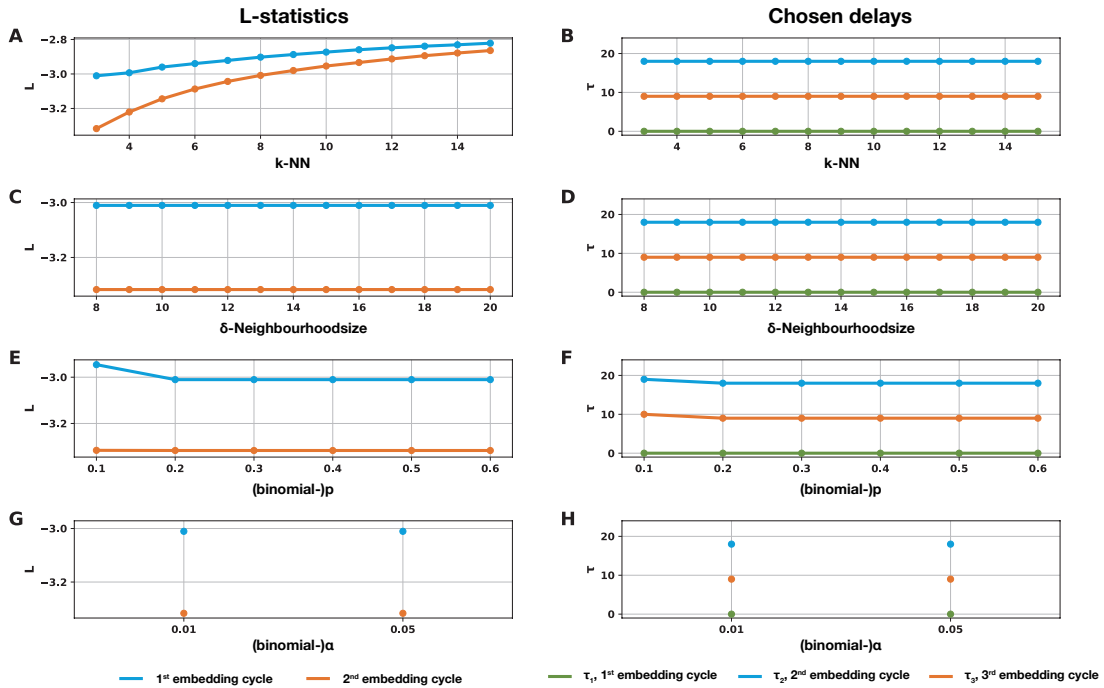
Figure C.2: L-statistic (left panels) and chosen delays (right panels) for a variety of choices of parameters for embedding the Lorenz system, relevant for the PECUZAL method. See text for details. When not varied, the other parameters were fixed to δ-neighborhoodsize$= 14$, $p = 0.5$, $\alpha = 0.05$ and $kNN = 3$, as we propose.
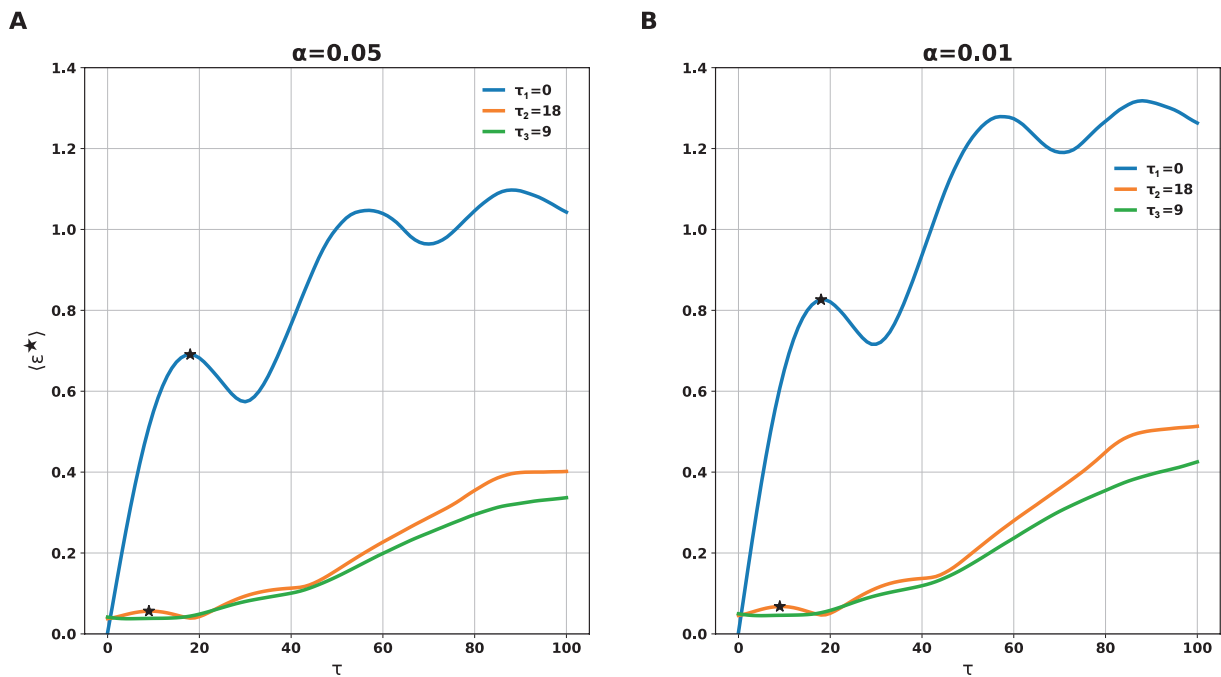
Figure C.3: Impact of different significance-level choices α on the continuity statistic. The other parameters for obtaining the continuity statistic were fixed to δ-neighborhoodsize$= 14$ and $p = 0.5$.
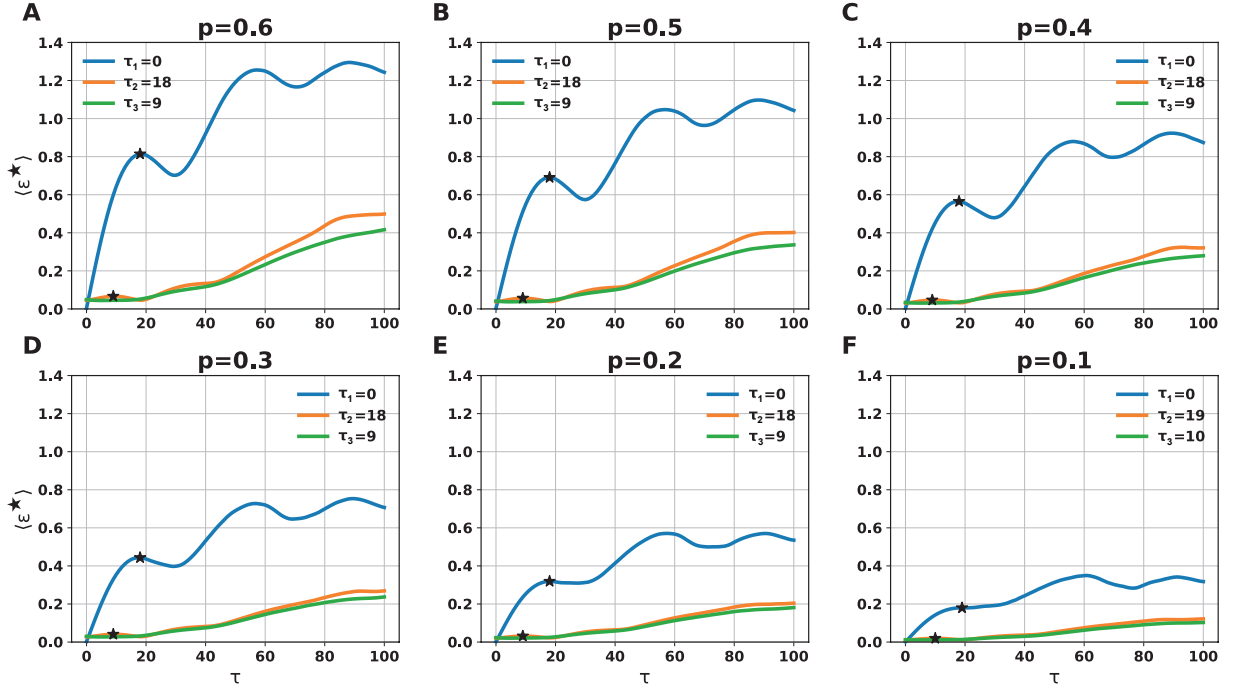
Figure C.4: Impact of different choices of the binomial probability parameter p on the continuity statistic. The other parameters for obtaining the continuity statistic were fixed to δ-neighborhoodsize$= 14$ and $\alpha = 0.05$.

C.4 Details of the L-statistic

The concept of quantifying noise amplification in the context of the validation of an attractor reconstruction has been proposed in [45]. The reconstruction process is considered in the presence of noise and the finite data availability as a modeling problem, introducing a noise amplification and estimation/prediction error for any measures on the reconstructed attractor. The variance of the conditional probability density function is a "natural criterion for assessing predictability" [45] and is used as the noise amplification for a fiducial point $\vec{v}_{\mathrm{fid}}(t)$ at a given noise level ε on the reconstructed attractor

$$\sigma_\varepsilon(T, \vec{v}_{\mathrm{fid}}(t)) = \frac{1}{\varepsilon}\sqrt{Var(\vec{v}(t+T)|B_\varepsilon(\vec{v}_{\mathrm{fid}}(t)))}, \tag{C.1}$$

with $Var(\vec{v}(t+T)|B_\varepsilon(\vec{v}_{\mathrm{fid}}(t)))$ being the conditional variance of $\vec{v}(t+T)$ for $\vec{v}_{\mathrm{fid}}(t)$ in a radius ε ball $B_\varepsilon(\vec{v}_{\mathrm{fid}}(t))$ for a prediction horizon T. Finally, the noise amplification σ

$$\sigma(T, \vec{v}_{\mathrm{fid}}(t)) = \lim_{\varepsilon \to 0} \sigma_\varepsilon(T, \vec{v}_{\mathrm{fid}}(t)) \tag{C.2}$$

averaged over all fiducial points on the attractor (and squared), $\langle \sigma(T) \rangle^2$, serves as a measure of the predictive power, with respect to the time horizon T, the reconstruction vectors $\vec{v}$ allow for (for details see Casdagli et al. [45]). Broadly speaking a low conditional variance, and thus, a low value of $\langle \sigma(T) \rangle^2$ is achieved for sufficiently unfolded attractors, because in this case noise distortions of the true trajectory are not likely to result in mixing states, which are far away from each other in true state space and consequently preserve the neighborhood relations on the

reconstructed attractor.

Uzal et al. [315] reinterpret Eqs. (C.1), (C.2) and give an approximation-recipe for the conditional variance. The authors redefine the mentioned equations as

$$\sigma_\varepsilon^2(\vec{v}_{\text{fid}}(t)) = \frac{1}{T_M} \int_0^{T_M} \sigma_\varepsilon^2(T, \vec{v}_{\text{fid}}(t)) dT \tag{C.3}$$

and consider the limit

$$\sigma(\vec{v}_{\text{fid}}(t)) = \lim_{\varepsilon \to 0} \sigma_\varepsilon(\vec{v}_{\text{fid}}(t)) \tag{C.4}$$

where ε is not related to any observational noise level anymore. The ε-ball $B_\varepsilon(\vec{v}_{\text{fid}}(t))$ is simply a tool for determining certain neighborhood relations of a fiducial point $\vec{v}_{\text{fid}}(t)$ and their changes when mapped to future states by the reconstruction function $F/\ F'$ (Fig. 2.1). It is possible to approximate the conditional variance in Eq. (C.1) by

$$E_k^2(T, \vec{v}_{\text{fid}}) \equiv \frac{1}{k+1} \sum_{\vec{v}' \in B_k(\vec{v}_{\text{fid}})} (\vec{v}'(t+T) - u_k(T, \vec{v}_{\text{fid}}))^2 \tag{C.5}$$

where $B_k(\vec{v}_{\text{fid}})$ estimates $B_\varepsilon(\vec{v}_{\text{fid}}(t))$ by the fiducial point itself and its k nearest neighbors, respecting a *Theiler window* (i.e., avoid temporal correlations in the neighbor-searching) [292]. The center of mass with respect to the chosen time horizon T and the fiducial point $\vec{v}_{\text{fid}}$ is defined as

$$u_k(T, \vec{v}_{\text{fid}}) \equiv \frac{1}{k+1} \sum_{\vec{v}' \in B_k(\vec{v}_{\text{fid}})} \vec{v}'(t+T). \tag{C.6}$$

The size of the k-neighborhood of $\vec{v}_{\text{fid}}$, $B_k(\vec{v}_{\text{fid}})$ is estimated as

$$\epsilon_k^2(\vec{v}_{\text{fid}}) \equiv \frac{2}{k(k+1)} \sum_{\substack{\vec{v}', \vec{v}'' \in B_k(\vec{v}_{\text{fid}}) \\ \vec{v}' \neq \vec{v}''}} \|\vec{v}' - \vec{v}''\|^2, \tag{C.7}$$

where $\|\cdot\|$ is a norm used for the distance computation. Finally, $E_k^2(T, \vec{v}_{\text{fid}})$ (Eq. (C.5)) is averaged over a range of T's in $[0, T_M]$ and the noise amplification estimated from k nearest neighbors is

$$\sigma_k^2(\vec{v}_{\text{fid}}) \equiv \frac{E_k^2(\vec{v}_{\text{fid}})}{\epsilon_k^2(\vec{v}_{\text{fid}})}, \tag{C.8}$$

which needs to averaged over all considered fiducial points N' on the reconstructed attractor to obtain

$$\sigma_k^2 = \sum_{i \in \{\vec{v}_{\text{fid}}\}}^{N'} \sigma_k^2(\vec{v}_{\text{i}}). \tag{C.9}$$

Since the reinterpretation of ε with the related k now acts as a neighborhood size parameter (Eqs. C.3,C.4,C.5,C.6,C.7,C.8), σ_k^2 can be normalized with respect to the averaged inter-point

distance, which depends on the sampling rate and the scale of the input data [315]. The normalization factor is

$$\alpha_k^2 = \frac{1}{\sum\limits_{i \in \{\vec{v}_{\mathrm{fid}}\}}^{N'} \epsilon_k^{-2}(\vec{v}_{\mathrm{i}})} \tag{C.10}$$

with ϵ_k^2 from Eq. (C.7). In this way, the final statistic will be able to compare attractor reconstructions stemming from different input data and also serves as an irrelevance measure, because large delays will result in large ϵ_k^2's.

Finally, Eqs. (C.9),(C.10) define the L-statistic

$$L_k = \log_{10}(\alpha_k \sigma_k), \tag{C.11}$$

which has a free parameter k and another implicit parameter T_M (Eq. (C.3)).

Appendix D

Supplementary material for Chapter 3: MCDTS for optimal embedding

D.1 Forecast of Hénon map time series
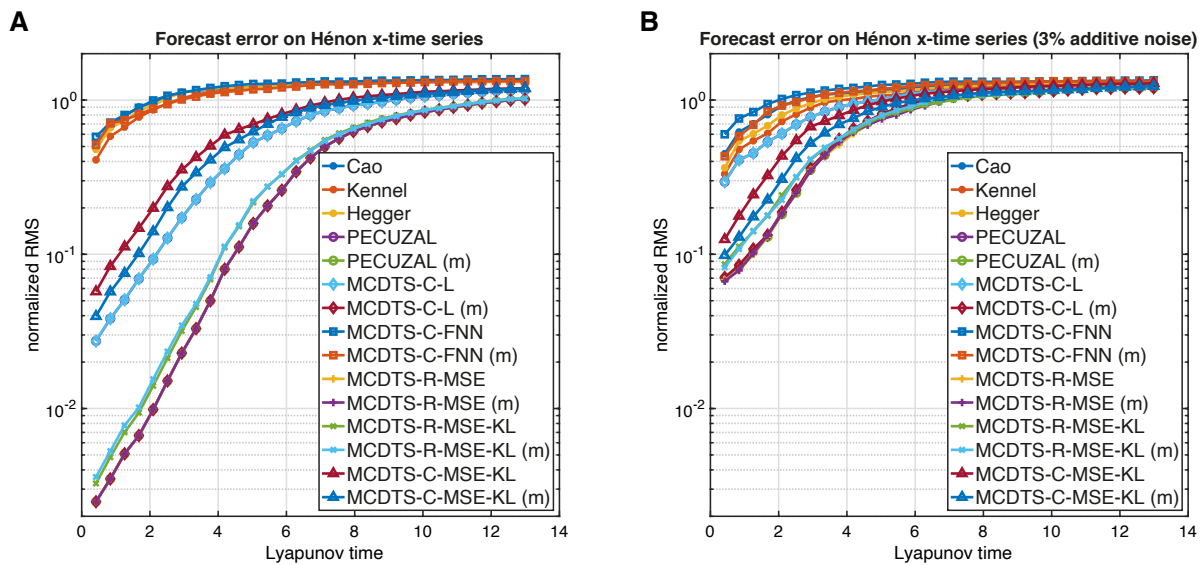
Figure D.1: A Normalized root-mean-square prediction errors (*RMS*) for the Hénon x-time series for all discussed reconstruction methods as a function of the prediction time. Shown are mean values of a distribution of 100 trials with different initial conditions. For the prediction we used a one step ahead zeroth-order approximation on the nearest neighbor of the last point of the reconstructed trajectory and iteratively repeated that procedure 30 times in order to obtain a prediction of 31 samples in total for each trial. **B** Same as in **A** but with 3% additive white noise.

D.2 CENOGRID prediction
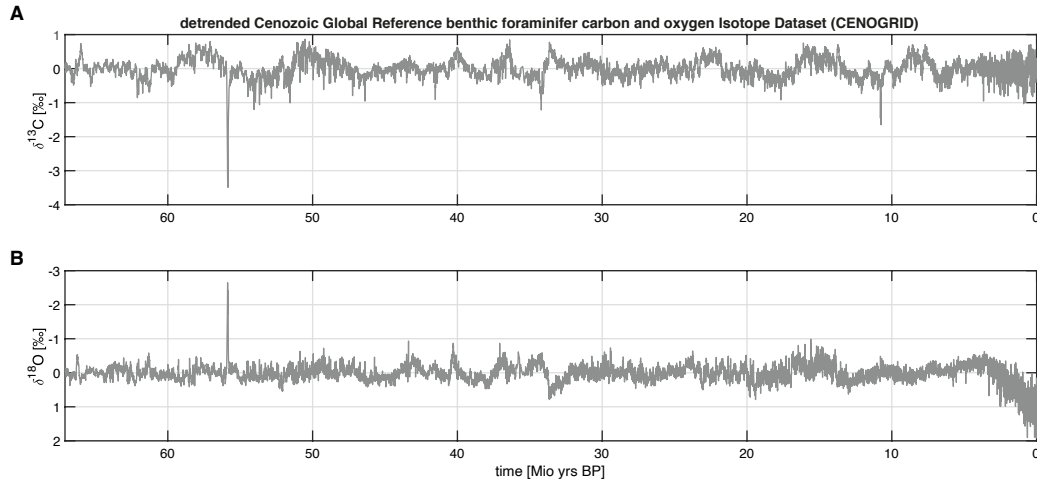
Figure D.2: A Detrended δ^{13}C and **B** δ^{18}O time series of a total length of $N = 13,421$ samples and a sampling period of $\Delta t = 5,000$yrs [328].

Table D.1: Obtained embedding parameters for the different reconstruction methods. Time series index 1 in the third column corresponds to the detrended δ^{13}C and time series index 2 to the detrended δ^{18}O record shown in Fig. D.2. For a description of the reconstruction methods see Table 3.1. The sequence of the delays (center column) and time series (right column) are a result of the embedding cycles which have been passed through in the corresponding reconstruction methods, which is why they are not necessarily ordered. For a reconstruction based on these embedding parameters it would make no difference whether delays and corresponding time series were sorted beforehand.

Reconstruction method	chosen delays [in index values]	chosen time series
Cao [39]	0,14,28,42,56,70,84	1,1,1,1,1,1,1
Kennel et al. [154]	0,14,28,42,56,70	1,1,1,1,1,1
Hegger and Kantz [125]	0,14,28,42,56,70,84,98	1,1,1,1,1,1,1,1
PECUZAL [162]	0,13,7,10,56,27,3,5,77,42,20,17,15	1,1,1,1,1,1,1,1,1,1,1,1,1
PECUZAL (m) [162]	0,13,7,10,56,27,3,5,77,42,20,17,15	1,1,1,1,1,1,1,1,1,1,1,1,1,1
MCDTS-C-FNN	0,69,84,54,37,48,63,26,12,20,16	1,1,1,1,1,1,1,1,1,1,1,1
MCDTS-C-FNN (m)	0,53,51,98,33,67,73,91,25,40,46	1,2,1,2,2,2,2,1,1,1,1
MCDTS-C-MSE-KL	0,54,117,93,126	1,1,1,1,1
MCDTS-C-MSE-KL (m)	0,109,78,198,37,155,53,95,133	1,2,1,2,2,1,1,1,1
MCDTS-R-MSE-KL	0,39,38,40,1,27	1,1,1,1,1,1
MCDTS-R-MSE-KL (m)	0,41,48,33,45,31	1,2,1,1,1,2
MCDTS-R-MSE	0,28,1,2	1,1,1,1
MCDTS-R-MSE (m)	0,1,2	1,1,1

D.3 Causal relationships in combustion process
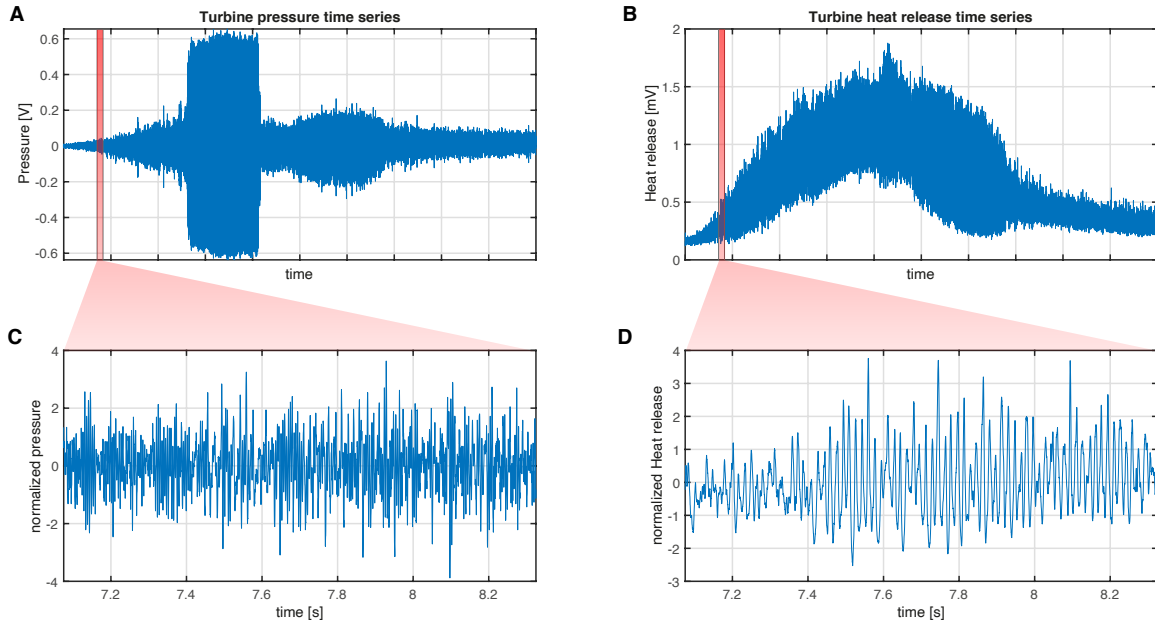
Figure D.3: Entire time series of length $N' = 400,000$ of **A** turbine pressure (measured in V) and **B** combustion heat release (measured in mV). This is a non-stationary setup with a linearly varying control parameter (air flow rate) leading to bifurcations and different dynamics. For the calculations in Section 3.3.4 we sampled both time series 50 times in subsamples of length $N = 5,000$. Panels **C, D** show one sample for both cases. Each sample has been normalized to zero mean and unit variance before we applied the embedding and the CCM.
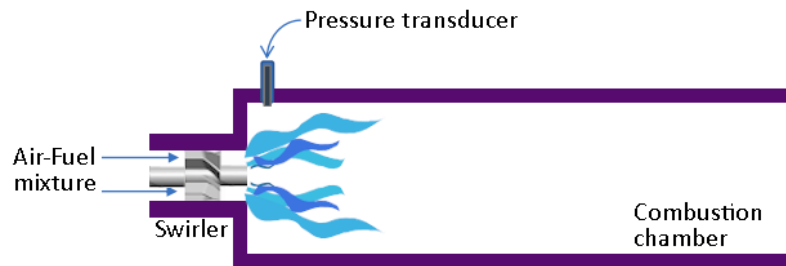
Figure D.4: Schematic experimental setup of the combustion experiment, see main text for details.

Appendix E

Supplementary material for Chapter 4: Recurrence threshold selection

E.1 Influence of embedding dimension on the variations in the maximum and mean pairwise distances

As discussed in Section 4.2, we show some numerical results illustrating the general behavior of mean and maximum L_∞ and L_2 distances for different types of systems in Fig. E.1. For a theoretical explanation of the observed changes with increasing embedding dimension, see Section 4.2.
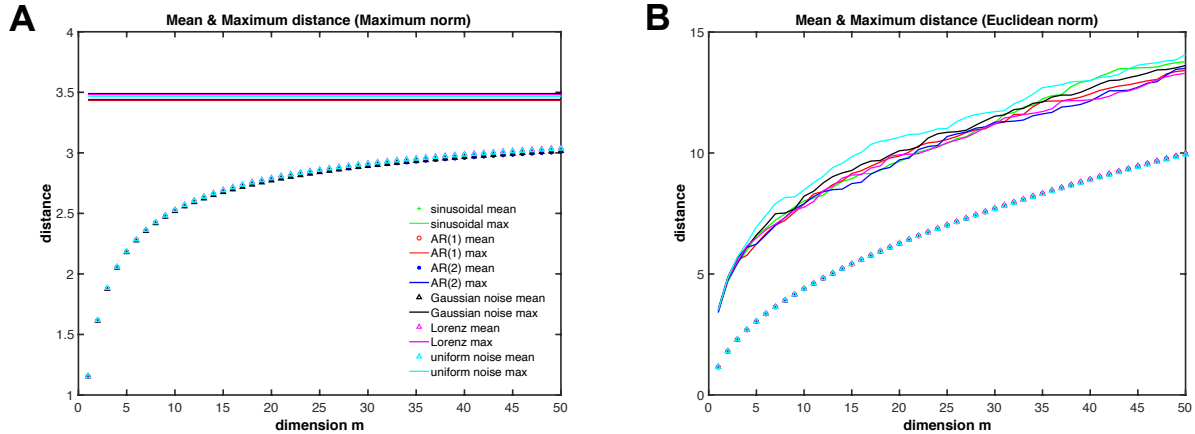
Figure E.1: **A** Mean $d_{mean}^{(\infty)}$ and maximum $d_{max}^{(\infty)}$ L_∞ distance between all pairs of state vectors as a function of the embedding dimension m for different types of time series: polychromatic harmonic oscillation with periods 3, 50 and 500; auto-regressive processes of first and second order with parameters $\varphi_1 = 0.5, \varphi_2 = 0.3$; random numbers of standard Gaussian (zero mean and unit variance) and uniform (unit variance) distributions, and y component of the Lorenz-63 system (Eq. (B.1), see Section 4.3) with control parameters $\sigma = 10$, $\beta = 8/3$ and r linearly increasing from 180 (chaotic regime) to 210 (periodic regime) as a function of the embedding dimension m. **B** Same as in panel **A**, but for L_2 distances.

E.2 Empirical shape parameters of the distance distributions for different systems

In order to further characterize the shape of the empirically observed pairwise distance distributions shown in Fig. 4.1 in more detail, we consider two standard characteristics from descriptive statistics. On the one hand, the skewness

$$\hat{s} = \frac{\frac{1}{N_d}\sum_{i=1}^{N_d}(d_i - \bar{d})^3}{\left(\sqrt{\frac{1}{N_d}\sum_{i=1}^{N_d}(d_i - \bar{d})^2}\right)^3} \tag{E.1}$$

of the distribution measures its asymmetry around the sample mean distance $\bar{d}$. On the other hand, we study the associated Shannon entropy

$$\hat{h} = -\sum_{j=1}^{N_b} p_j \frac{\log(p_j)}{\log(N_b)} \tag{E.2}$$

providing an integral measure of the heterogeneity of the distribution of d. Here, j enumerates the bins of a histogram of the values of d with N_b bins and relative frequencies p_j, and N_d is the number of pairwise distances in the sample (i.e., the number of independent entries of the distance matrix $\mathbf{d}$, $N_d = N_{\text{eff}}(N_{\text{eff}} - 1)/2$). The bin width has been selected by first computing the optimum value according to the Freedman-Diaconis rule [96] for each embedding dimension m and then averaging over all corresponding values and taking the resulting mean to keep N_b fixed for each considered setting. Specifically, for the time series drawn from the Gaussian and uniform distributions, $N_{b,L_2} = 355$ and $N_{b,L_\infty} = 286$, while for the Lorenz system, $N_{b,L_2} = 701$ and $N_{b,L_\infty} = 771$.

According to the corresponding normalization, $\hat{h}$ assumes its maximum of one in case of a uniform distribution (since then, $p_j = 1/N_b \quad \forall j = 1,...,N_b$, i.e., for each (binned) distance within $[d_{min}, d_{max}]$). In turn, the more heterogeneous (e.g., spiky or generally asymmetric) the distribution of distances gets, the lower $\hat{h}$.

Figure E.2 shows the resulting behavior of both characteristics for the L_2 (panels A,C,E) and L_∞ (panels B,D,F) distances obtained from uniform and Gaussian distributed noise as well as for the non-stationary Lorenz-63 system (Eq. (B.1), see Section 4.3) in dependence on the embedding dimension. The results complement the qualitative description based on a visual inspection of Fig. 4.1 as given in Section 4.2. In case of the L_2 norm and time series drawn from uniform and Gaussian distributions (Fig. E.2 A,C) we observe the skewness converging towards zero (symmetric Gaussian distribution) and the entropy reflecting this convergence towards a normal distribution by a downward trend until the skewness approaches zero as m further increases. Although the theoretically predicted Gaussian shape for high m is visually apparent in case of the time series from the Lorenz-63 system (see Fig. 4.1 E), the skewness takes clearly non-zero negative values while the entropy constantly decreases with increasing m, indicating an asymmetric shape (Fig. E.2 E). In case of the L_∞ norm, the considered maximum embedding dimension appears not suited for observing convergence of both shape parameters.
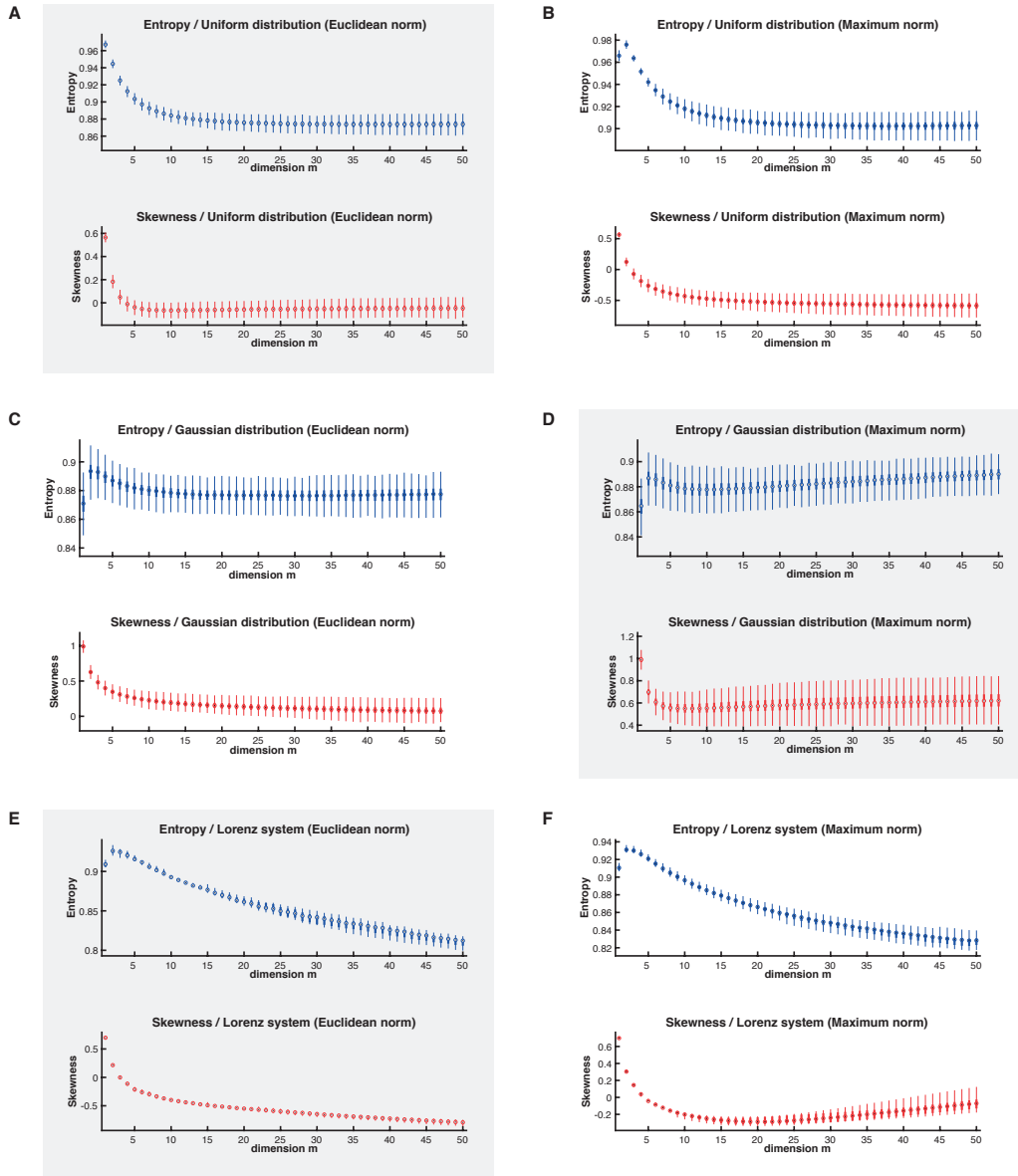
Figure E.2: Skewness (red) and Shannon entropy (blue) of the L_2 (**A**,**C**,**E**) and L_∞ (**B**,**D**,**F**) distances of $N = 1,500$ independent random numbers with uniform **A**+**B** and Gaussian **C**+**D** distribution and **E**+**F** the y component of the Lorenz-63 system (Eq. (B.1), $N = 6,000$, see Section 4.3) with control parameters $\sigma = 10$, $\beta = 8/3$ and r linearly increasing from 180 (chaotic regime) to 210 (periodic regime) as a function of the embedding dimension m. For the two noise series, box plots show the variability estimated from 1,000 independent realizations for each data set, using a random number generator. In case of the Lorenz-63 system the variability is estimated from 10 independent realizations of the non-stationary Lorenz-63 equations with randomly chosen initial conditions.

E.3   RP's and RQA for one realization of the non-stationary Lorenz system

For further illustrating the RPs resulting from the time-dependent Lorenz-63 system discussed in Section 4.3, we show here the results for just one example trajectory corresponding to a set of randomly chosen initial conditions $x(0) = 0.9649$, $y(0) = 0.1576$, $z(0) = 0.9706$. As before, we embed the y component time series and study the RP for each previously discussed threshold selection method. Then, we use a running window over each (global) RP with a window size of $w = 400$ and mutual shift of $ws = 40$ data points, i.e., 90% overlap between consecutive windows.

The RPs and the associated time-dependent recurrence characteristic RTE (Eq. (A.10)) for a "normal" three-dimensional embedding with time delay $\tau = 4$, consistent with the first local minimum of the mutual information [95], are shown in Fig. E.3, using the Euclidean norm. We compare the results for four different threshold selection methods but similar effective threshold values (corresponding to a global recurrence rate of $RR \approx 4\%$), which are thus expected to give comparable results. The left panel corresponds to the recommended method of taking a certain percentile of the distance distribution, while the other three panels are based on thresholds selected according to some percentage of the maximum, mean and median distance of state vectors on the attractor in the reconstructed state space. Comparing the different panels, as expected there are hardly any marked differences in the RPs or the temporal changes of RTE. The transition from a chaotic regime into a periodic one is well reflected by a constantly decreasing RTE, which takes its minimum for the limit cycle behavior between $t_1 \approx 1,000$ and $t_2 \approx 1,080$.

However, if choosing a higher-dimensional embedding (e.g., $m = 10$) motivated by the nonstationarity of the system, the RP becomes almost completely white if the recurrence threshold is chosen based upon the same percentages of the maximum, mean or median state space distances as used before (Fig. E.4). In this case the RTE is still able to detect the transitory limit cycle regime, but one looses information about the chaotic regime before. In contrast, we retain the same density of recurrences and, hence, resolution of the RP as for $m = 3$ when fixing the threshold according to the whole distance distribution (left panel in Fig. E.4). Here, the overall behavior of RTE from the lower-dimensional ($m = 3$) case is qualitatively retained, although the periodic regime is less well expressed than in the former case.
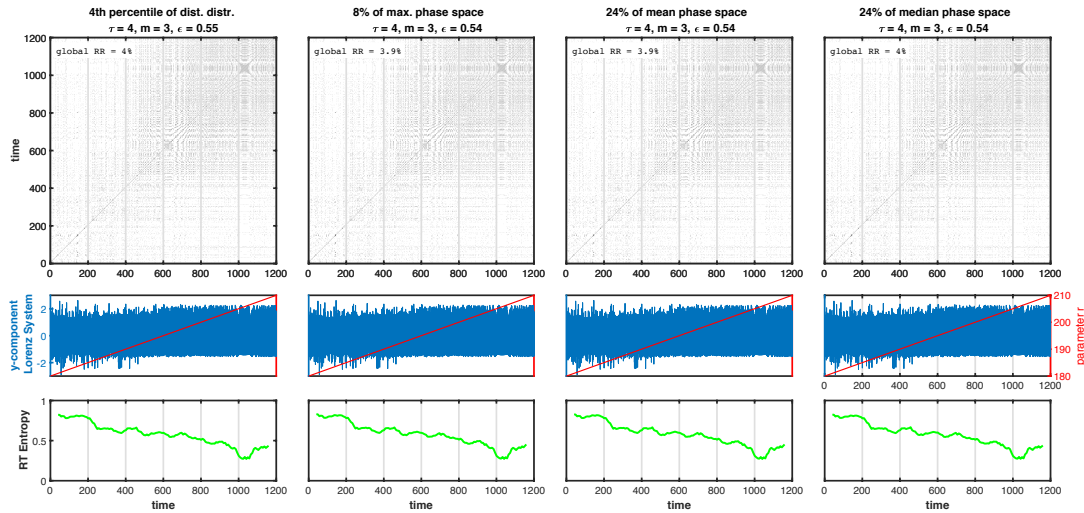
Figure E.3: RPs, according time series (blue), time-dependence of the control parameter r (red) and recurrence characteristic *RTE* (green) based on the y component of the non-stationary Lorenz-63 system (see text for details), using the L_2 norm. Shown are the results for low-dimensional embedding ($m = 3$) and for four different methods to select the recurrence threshold according to a certain percentile of the distance distribution and some percentage of the maximum, mean or median distances of state vectors on the reconstructed attractor (from left to right). The actual threshold values (4th percentile, 8%, 24% and 24%, respectively) have again been chosen such that the global recurrence rate for each method in this embedding scenario is $\approx 4\%$.
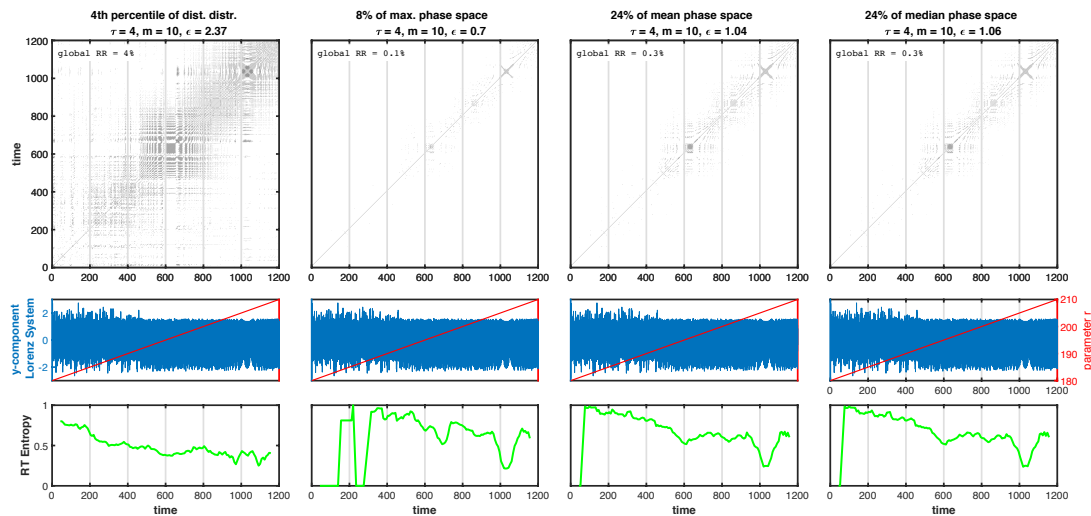
Figure E.4: Same as in Fig. E.3, but for ten-dimensional embedding. In comparison to Fig. E.3, three of the four methods lead to a marked drop in the global recurrence rate and a resulting change in the *RTE* values. Only for a recurrence threshold corresponding to the same percentile of the distance distribution, the results are qualitatively stable over the full considered time evolution.

Appendix F

Supplementary material for Chapter 5: Border effect and tangential motion correction for RQA
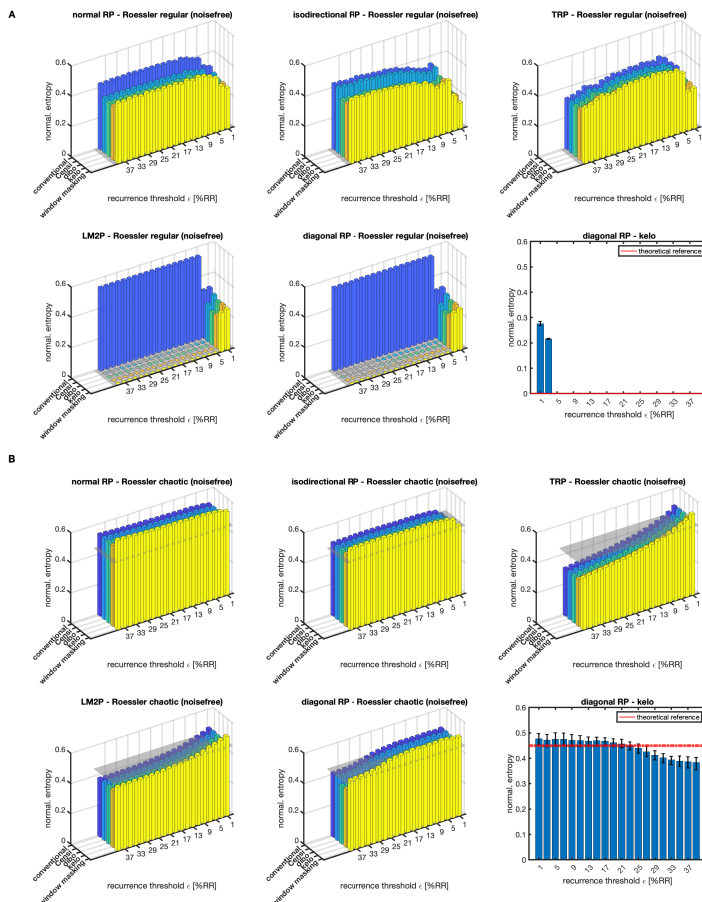
Figure F.1: Diagonal line length entropy estimates as a function of the recurrence threshold ε. Shown are results for all described correction schemes for counting diagonal lines (Section 5.3) and suppressing tangential motion (Section 5.4), except the perpendicular recurrence plot $R^{\perp}$. In the top panel **A** median diagonal line length entropy values gained from 100 realizations of the noise free regular limit cycle regime of the Rössler system are shown, whereas the bottom panel **B** shows its chaotic regime counterpart, see text in Section 5.6 for details. The grey-shaded surface denotes the theoretical expectation value (median) computed from Eq. (5.14). Results for the diagonal RP and the kelo correction scheme are shown in the bottom right subplot, which is a cutout of the orange bars in the bottom center subplot, here including errorbars as two standard deviations from the computed ensemble.

Figure F.2: Diagonal line length entropy estimates as a function of the recurrence threshold ε. Shown are results for all described correction schemes for counting diagonal lines (Section 5.3) and suppressing tangential motion (Section 5.4), except the perpendicular recurrence plot $R^{\perp}$. In the top panel **A** median diagonal line length entropy values gained from 1,000 realizations of the noise free regular limit cycle regime of the Logistic map are shown, whereas the bottom panel **B** shows its chaotic regime counterpart, see text in Section 5.6 for details. The grey-shaded surface denotes the theoretical expectation value (median) computed from Eq. (5.14). Results for the diagonal RP and the kelo correction scheme are shown in the bottom right subplot, which is a cutout of the orange bars in the bottom center subplot, here including errorbars as two standard deviations from the computed ensemble.

Figure F.3: Diagonal line length entropy estimates as a function of the recurrence threshold ε. Shown are results for all described correction schemes for counting diagonal lines (Section 5.3) and suppressing tangential motion (Section 5.4), except the perpendicular recurrence plot $R^{\perp}$. In the top panel **A** median diagonal line length entropy values gained from 100 realizations of the additive noise contaminated regular limit cycle regime of the Rössler system are shown, whereas the bottom panel **B** shows its chaotic regime counterpart, see text in Section 5.6.1 for details. Here, we added noise as 10% of the mean standard deviation of the multivariate signal gained from the numerical integration. The grey-shaded surface denotes the theoretical expectation value (median) computed from Eq. (5.14). Results for the diagonal RP and the kelo correction scheme are shown in the bottom right subplot, which is a cutout of the orange bars in the bottom center subplot, here including errorbars as two standard deviations from the computed ensemble.
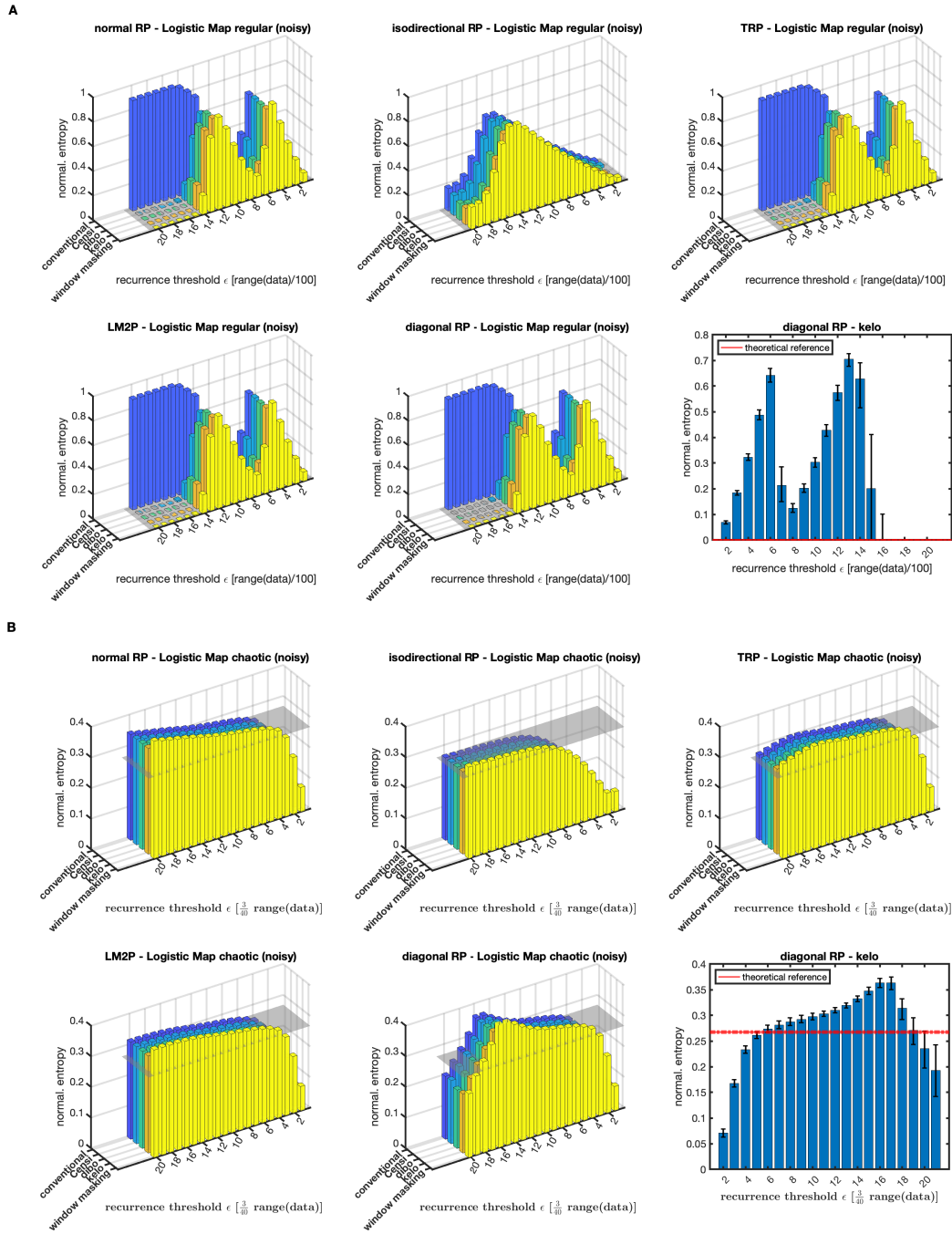
Figure F.4: Diagonal line length entropy estimates as a function of the recurrence threshold ε. Shown are results for all described correction schemes for counting diagonal lines (Section 5.3) and suppressing tangential motion (Section 5.4), except the perpendicular recurrence plot $R^{\perp}$. In the top panel **A** median diagonal line length entropy values gained from 1,000 realizations of the additive noise contaminated regular limit cycle regime of the Logistic map are shown, whereas the bottom panel **B** shows its chaotic regime counterpart, see text in Section 5.6.1 for details. Here, we added additive noise as 10% of the standard deviation of the time series. The grey-shaded surface denotes the theoretical expectation value (median) computed from Eq. (5.14). Results for the diagonal RP and the kelo correction scheme are shown in the bottom right subplot, which is a cutout of the orange bars in the bottom center subplot, here including errorbars as two standard deviations from the computed ensemble.

F.1   Sensitivity of the results to noise
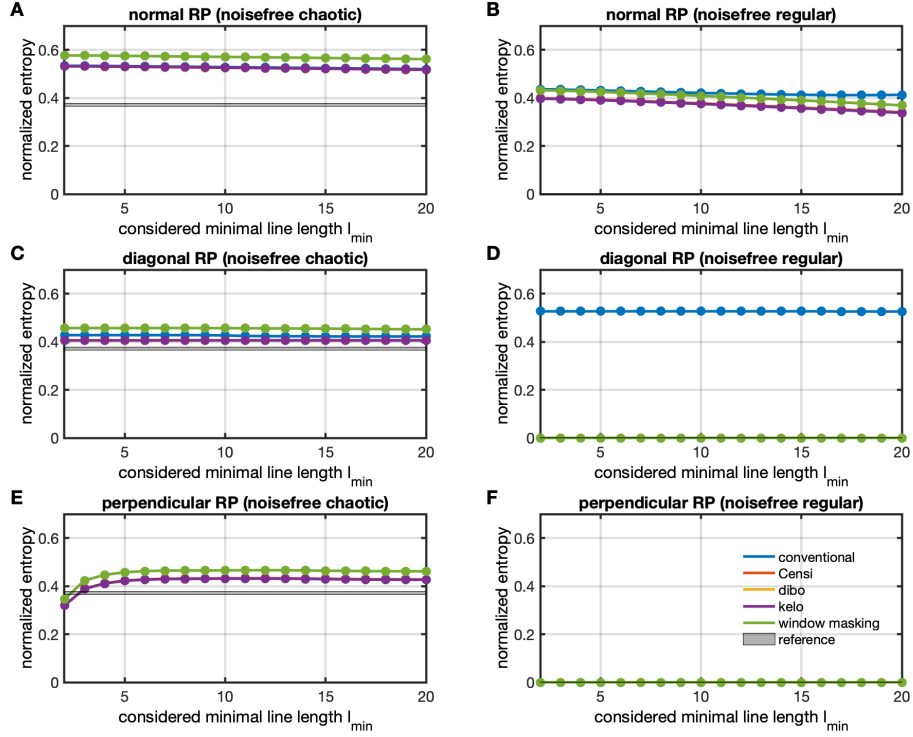
Figure F.5: Normalized diagonal line length entropy estimates as a function of the minimum line length $\ell_{\min}$ for noisefree data from the high sampled Rössler system (cf. Section 5.6.1). In the left panels (**A**, **C**, **E**) the underlying system exhibits chaotic dynamics, whereas the right panels (**B**, **D**, **F**) show their regular counterparts. The normal RPs (**A**, **B**) and the perpendicular RPs (**E**, **F**) were constructed using a fixed recurrence threshold corresponding to 35% recurrence rate. The normal RPs served as input for obtaining the diagonal RPs $R^{\nearrow}$ (**C**, **D**) and for the computation of the perpendicular RPs $R^{\perp}$ we used an angle threshold $\varphi = 15°$. The grey shaded areas show medians of ensembles of 1,000 analytically computed reference values for $K_1 \pm$ two standard deviations of these distributions transformed by using Eq. (5.14).
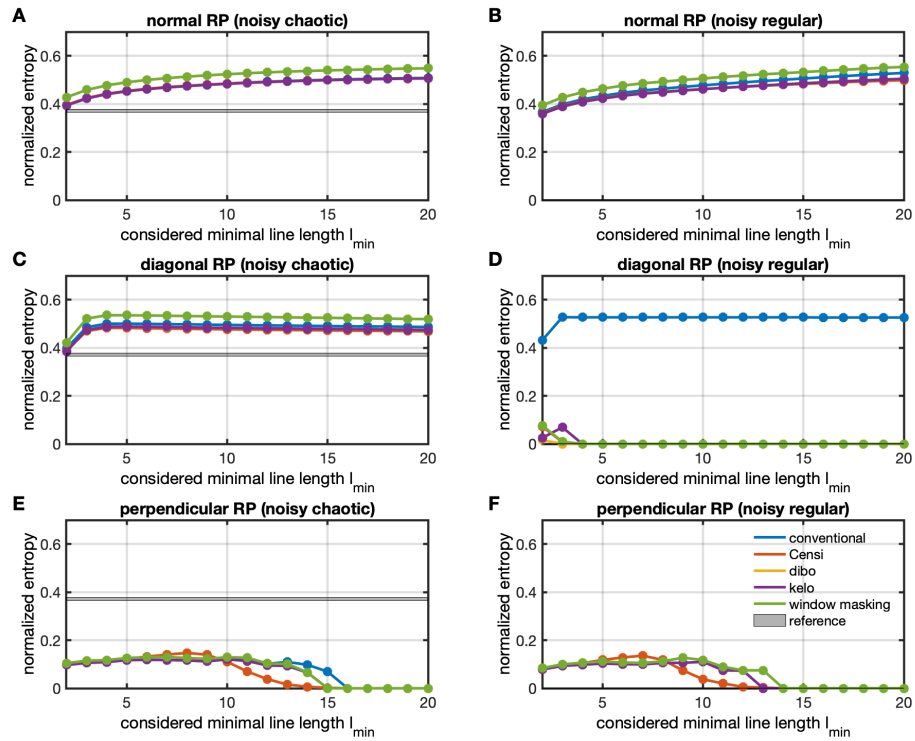
Figure F.6: Normalized diagonal line length entropy estimates as a function of the minimum line length ℓ_{min} for noise corrupted data from the high sampled Rössler system (cf. Section 5.6.1). We added noise from an auto-regressive (AR) process of second order as 20% of the mean standard deviation of the multivariate signal gained from the numerical integration (cf. Eq. (5.15)). In the left panels (**A**, **C**, **E**) the underlying system exhibits chaotic dynamics, whereas the right panels (**B**, **D**, **F**) show their regular counterparts. The normal RPs (**A**, **B**) and the perpendicular RPs (**E**, **F**) were constructed using a fixed recurrence threshold corresponding to 35% recurrence rate. The normal RPs served as input for obtaining the diagonal RPs $R^{\nearrow}$ (**C**, **D**) and for the computation of the perpendicular RPs $R^{\perp}$ we used an angle threshold $\varphi = 15°$. The grey shaded areas show medians of ensembles of 1,000 analytically computed reference values for $K_1 \pm$ two standard deviations of these distributions transformed by using Eq. (5.14).

Appendix G

Supplementary material for Chapter 6: Recurrence powerspectra

G.1 Recurrence powerspectra for noisy Rössler system
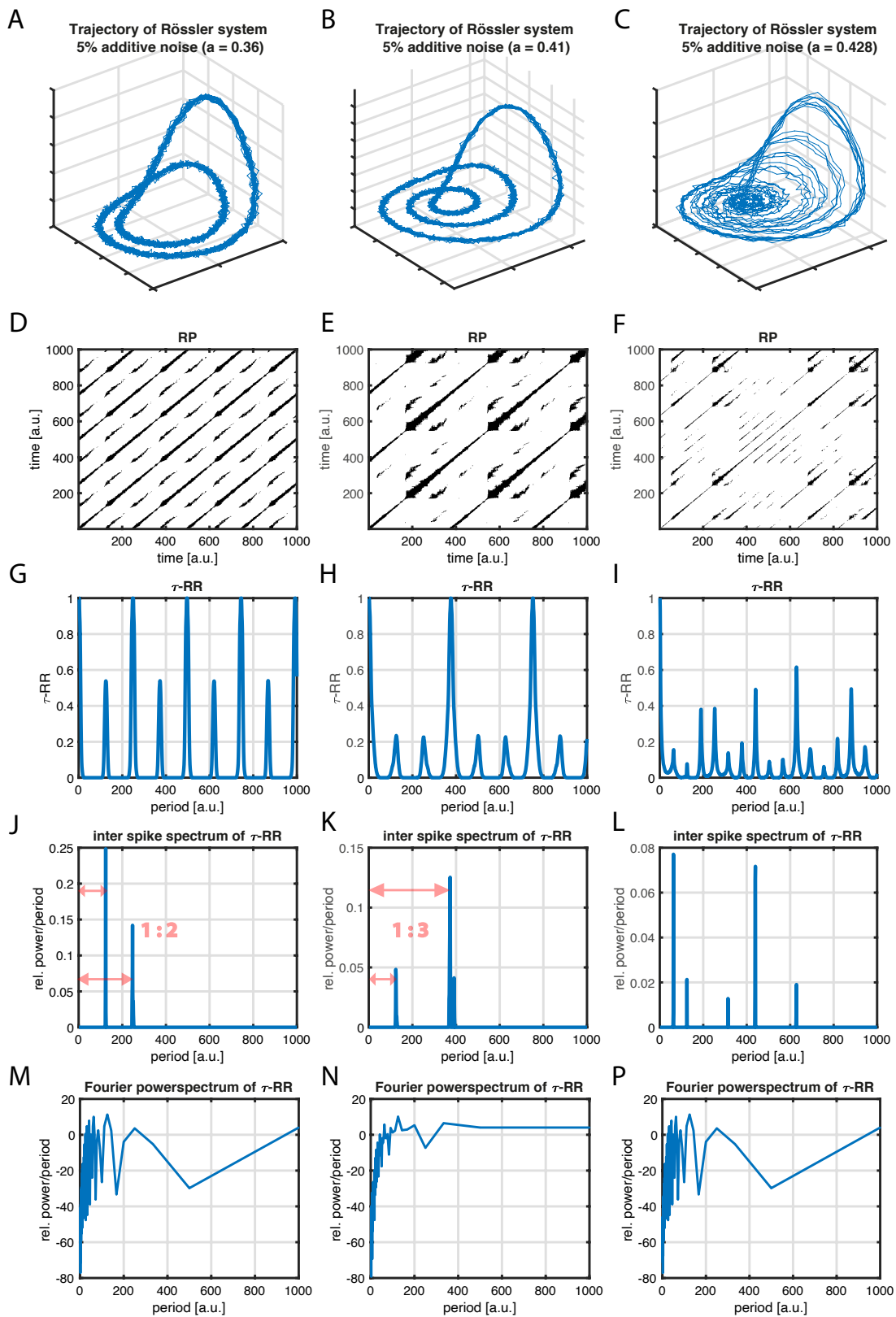
Figure G.1: Same as in Fig. 6.4, but here with 5% additive Gaussian white noise. The appearance of the inter spike spectra in **J, K, L** and the Fourier spectra in **M, N, P** are unaffected by the additive noise noise.

G.2 Evolutionary spike spectrograms for CENOGRID
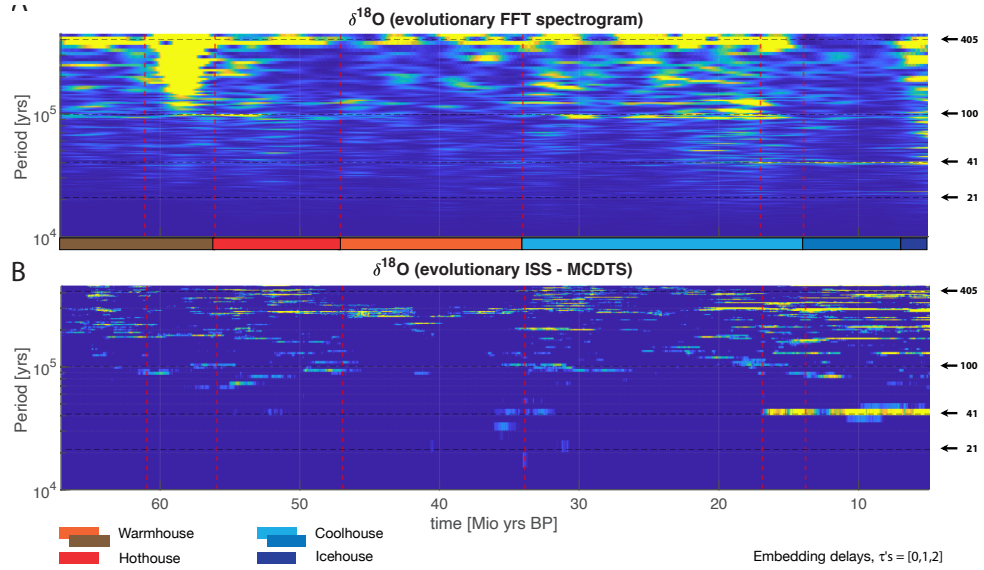
Figure G.2: Evolutionary spectrograms based on **A** FFT powerspectra of the detrended δ^{18}O time series and **B** on the inter spike spectra of the τ-RR of the MCDTS-R-MSE embedded δ^{18}O record. The spectrograms were computed with a $w = 1,000(\hat{=}$ 5-Myr) window and a step $ws = 1$. We used the embedding computed in Chapter 3 (Section 3.3.3, see Table D.1) and a recurrence threshold fixed to 5% global RR, in order to ensure comparability within the different windows [163] (Chapter 4). Dashed red horizontal lines indicate major climate transitions according to Westerhold et al. [328].
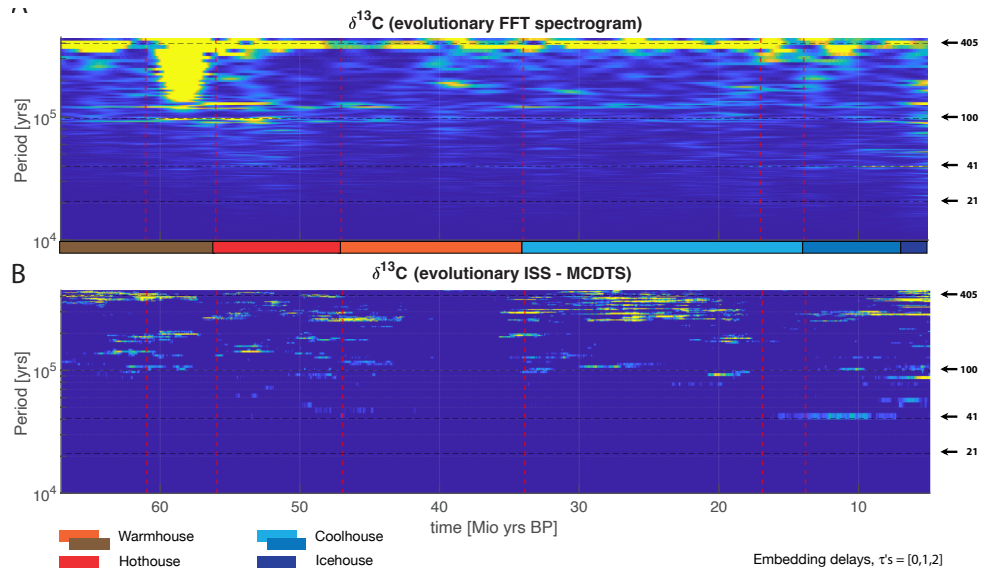
Figure G.3: Same as in Figure G.2, but for the detrended δ^{13}C isotope record.
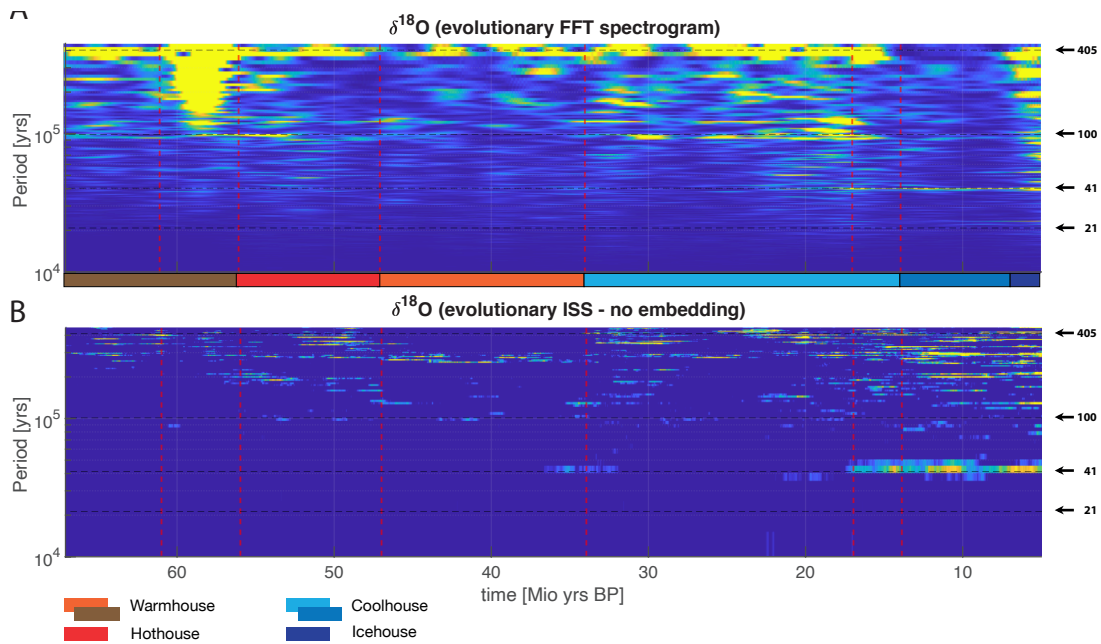
Figure G.4: Evolutionary spectrograms based on **A** FFT powerspectra of the detrended δ^{18}O time series and **B** on the inter spike spectra of the τ-RR of the non-embedded δ^{18}O record. The spectrograms were computed with a $w = 1,000 (\hat{=}$ 5-Myr) window and a step $ws = 1$. We used a recurrence threshold fixed to 5% global RR, in order to ensure comparability within the different windows [163] (Chapter 4). Dashed red horizontal lines indicate major climate transitions according to Westerhold et al. [328].
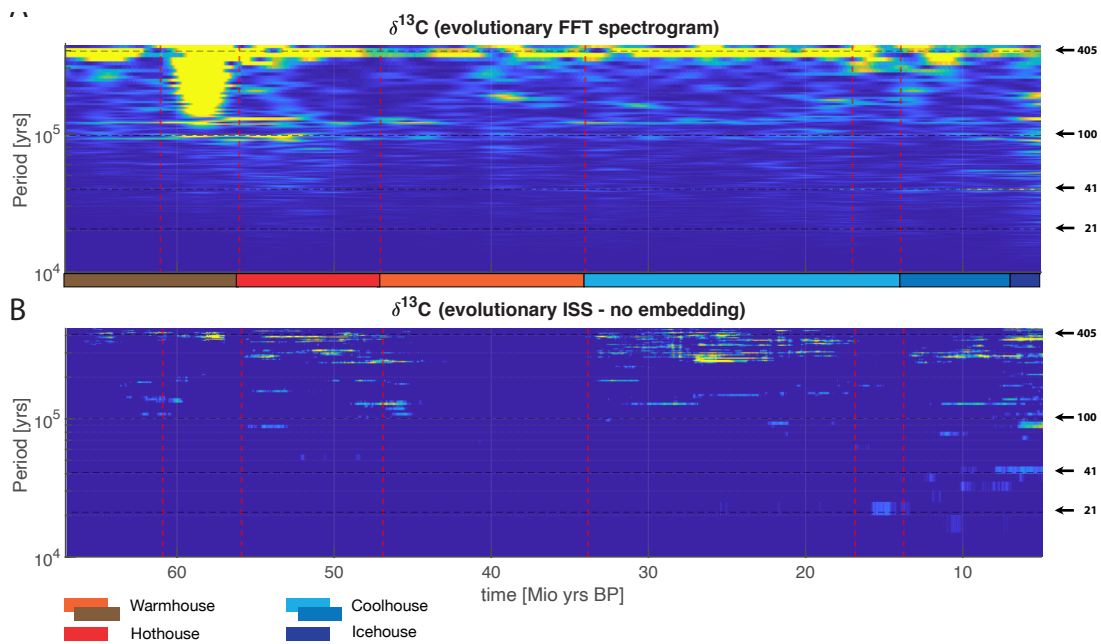
Figure G.5: Same as in Figure G.4, but for the detrended δ^{13}C isotope record.

Appendix H

Supplementary material for Chapter 8: Recurring types of paleoclimate in southern Ethiopia

A Record of potassium (K) concentrations

B Record of the filtered ODP Site 967 wetness index
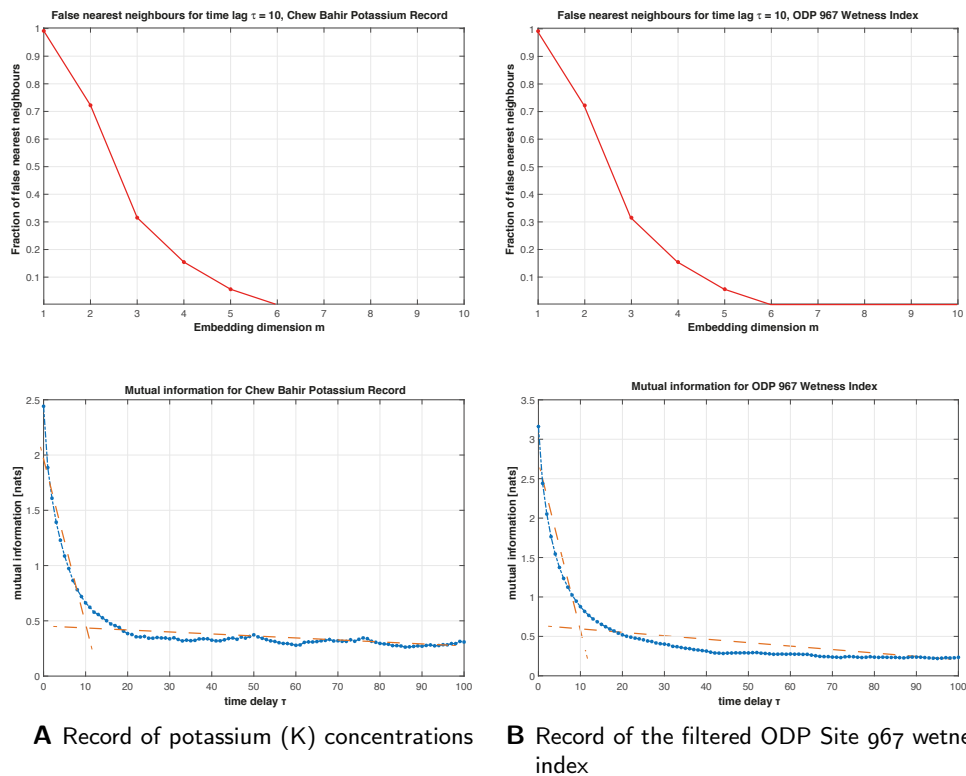
Figure H.1: Top panels: Fraction of false nearest neighbours vs. increasing embedding dimension using the method by Hegger and Kantz [125]. The embedding dimension is chosen in such a way that the fraction of the false nearest neighbors vanishes, which in our example occurs at $m = 6$. Bottom panels: Mutual information. The time delay is chosen at the point where the steep decline in mutual information changes into a flatter decline, which happens at $\tau = 10$.

A Interval between 617 and 410 kyr BP   **B** Interval between 410 and 348 kyr BP
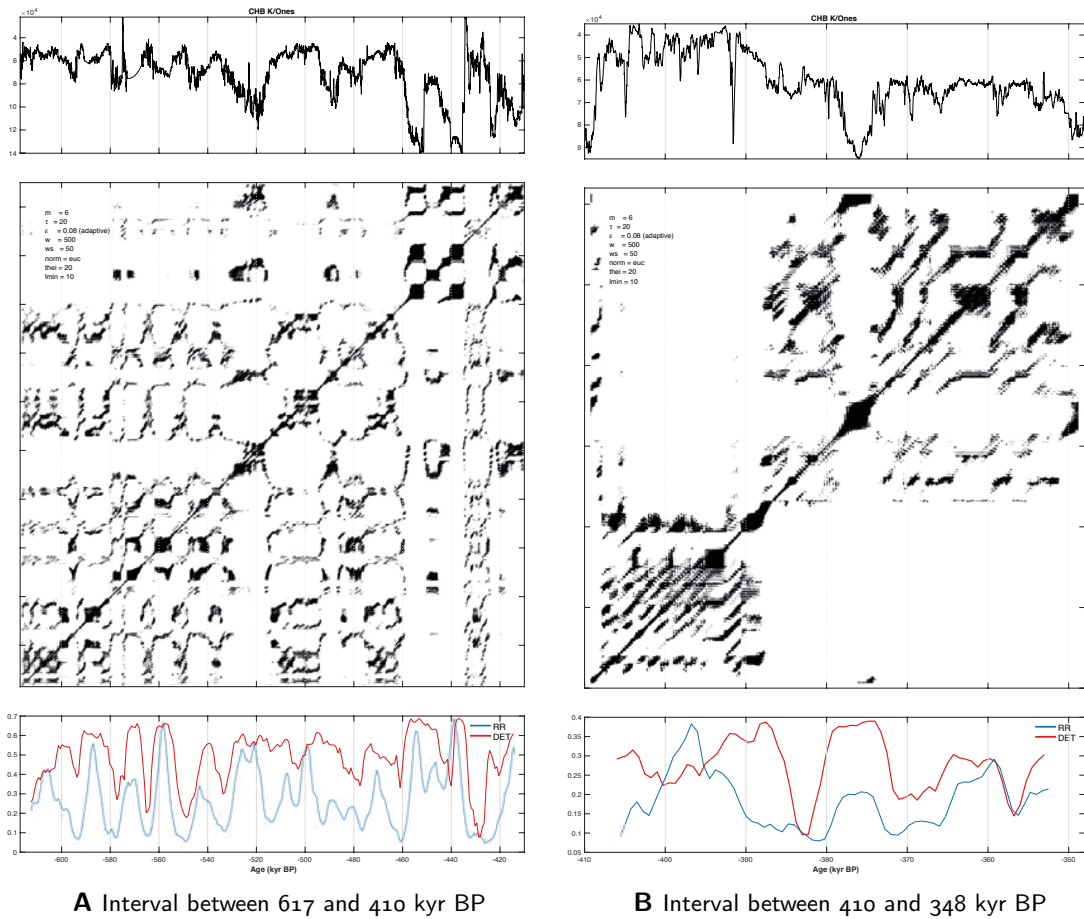
Figure H.2: Recurrence plot (RP) and recurrence quantification analysis (RQA) measures of the high resolution (0.015 kyr) potassium (K) concentrations of the sediment in Chew Bahir in southern Ethiopia: the time series according to age model (3) (upper panel), the recurrence plot (middle panel) and the RQA measures RR (blue) and DET (red) of moving windows (lower panel). Embedding parameters m=embedding dimension, τ=time delay, ε=threshold, w=window size, ws=window moving steps, norm=vector norm, thei=size of Theiler window, lmin=minimum line length, RQA measures RR=recurrence rate and DET=determinism.
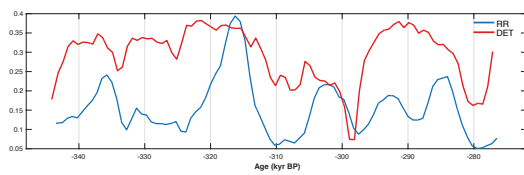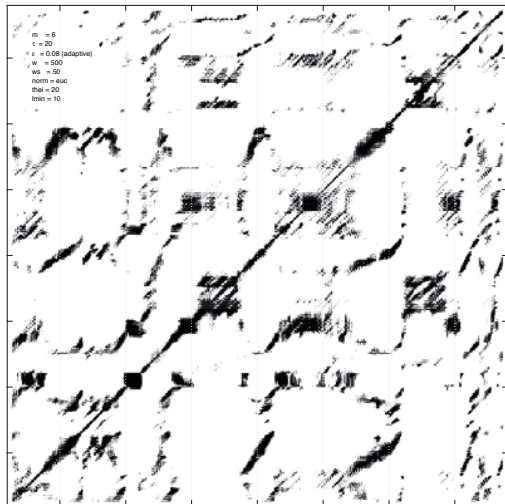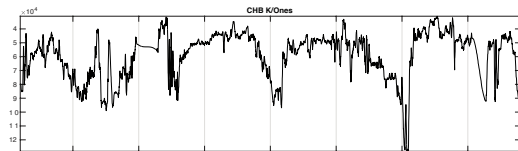
A Interval between 348 and 272 kyr BP
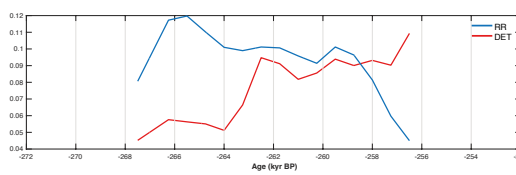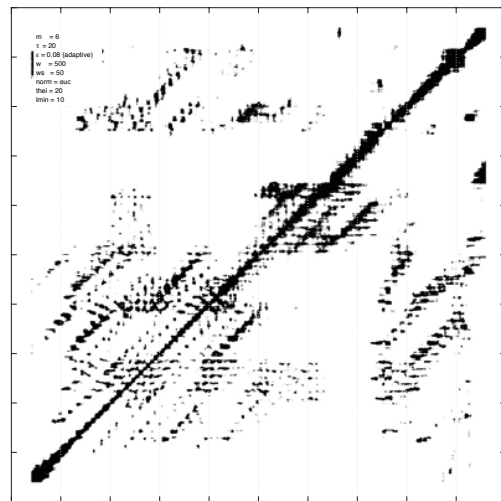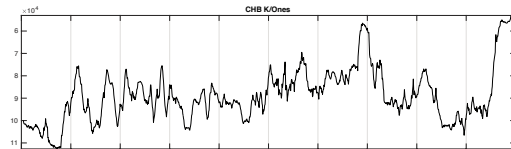
B Interval between 272 and 252 kyr BP

Figure H.3: Same as in Fig. H.2, but for different time intervals. The obvious block after 330 kyr BP is caused by the constant values in the time series at this point of time, which is the result of interpolating a larger gap in the original time series.

A Interval between 252 and 122 kyr BP **B** Interval between 107 and 3 kyr BP
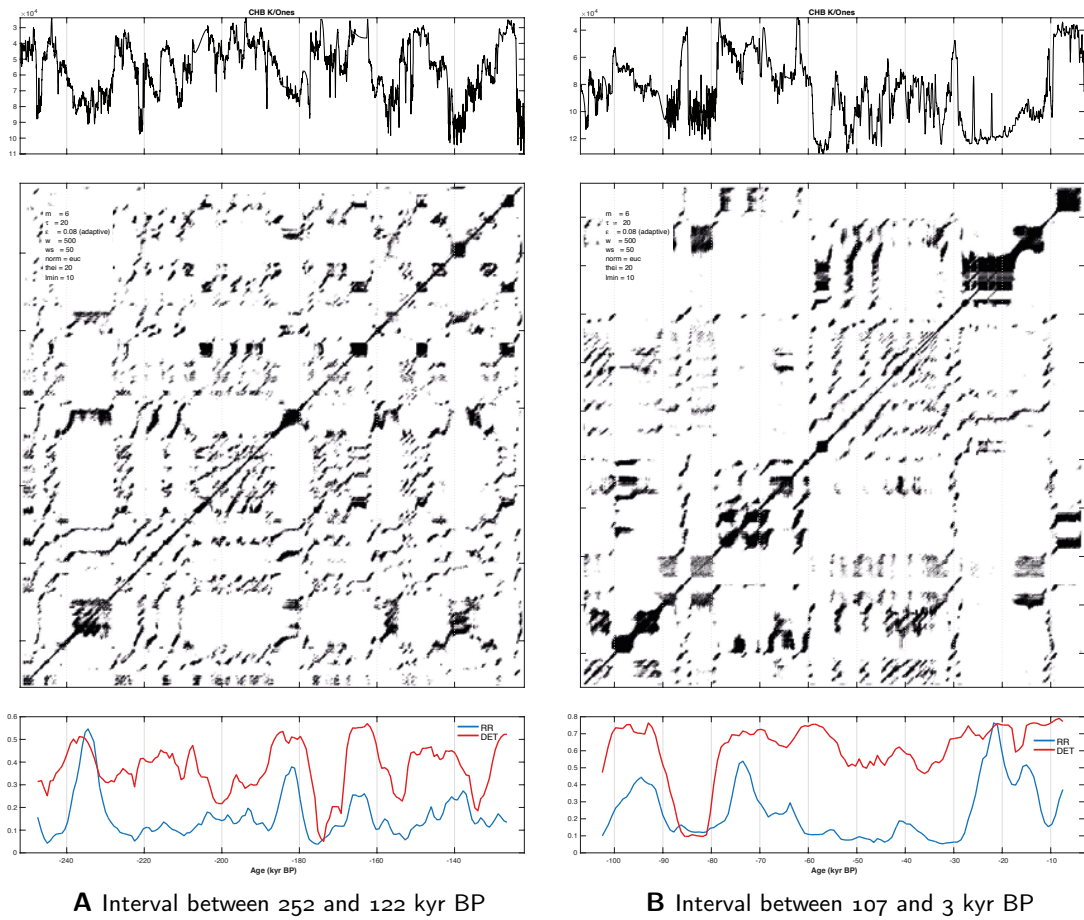
Figure H.4: Same as in Fig. H.2, but for different time intervals.

Code and data availability

All proposed algorithms are OpenSource and can be found and downloaded at `https://github.com/hkraemer/` [165, 166, 168, 169]. The code for the data analysis and modeling presented in this thesis is available from the author upon request and the references can be found in the corresponding Chapters. Please do not hesitate to contact me in case of any questions.

Bibliography

[1] Henry D.I. Abarbanel. *Analysis of Observed Chaotic Data.* Springer, New York, 1996.

[2] U Rajendra Acharya, Eric Chern-Pin Chua, Oliver Faust, Teik-Cheng Lim, and Liang Feng Benjamin Lim. Automated detection of sleep apnea from electrocardiogram signals using nonlinear parameters. *Physiological measurement,* 32 (3), (2011), 287.

[3] U. R. Acharya, V. K. Sudarshan, H. Adeli, J. Santhosh, J. E. W. Koh, and A. Adeli. Computer-Aided Diagnosis of Depression Using EEG Signals. *European Neurology,* 73 (5–6), (2015), 329–336.

[4] V. S. Afraimovich, N. N. Verichev, and M. I. Rabinovich. Stochastic synchronization of oscillation in dissipative systems. *Radiophysics and Quantum Electronics,* 29 (9), (1986), 795–803.

[5] Luis Antonio Aguirre. A nonlinear correlation function for selecting the delay time in dynamical reconstructions. *Physics Letters A,* 203 (2), (1995), 88–94.

[6] C. Ahlstrom, P. Hult, and P. Ask. Thresholding distance plots using true recurrence points. *Proceedings of the IEEE Conference on Acoustics, Speech and Signal Processing (ICASSP 2006),* 3 (1660747), (2006), III688–III691.

[7] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control,* 19 (6), (1974), 716–723.

[8] A. M. Albano, J. Muench, C. Schwartz, A. I. Mees, and P. E. Rapp. Singular-value decomposition and the Grassberger-Procaccia algorithm. *Phys. Rev. A,* 38, (6 Sept. 1988), 3017–3026.

[9] A.M. Albano, A. Passamante, and Mary Eileen Farrell. Using higher-order correlations to define an embedding window. *Physica D: Nonlinear Phenomena,* 54 (1), (1991), 85–97.

[10] Zoran Aleksić. Estimating the embedding dimension. *Physica D: Nonlinear Phenomena,* 52 (2), (1991), 362–368.

[11] Stanley H. Ambrose. Late Pleistocene human population bottlenecks, volcanic winter, and differentiation of modern humans. *Journal of Human Evolution,* 34 (6), (1998), 623–651.

[12] E.V. Appleton and Balth. van der Pol junr. XVI. On a type of oscillation-hysteresis in a simple triode generator. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science,* 43 (253), (1922), 177–193.

[13] John H. Argyris, Gunter Faust, Maria Haase, and Rudolf Friedrich. *Die Erforschung des Chaos: eine Einführung in die Theorie nichtlinearer Systeme.* Zweite. Heidelberg: Springer, 2010.

Bibliography

[14] J Asensio-Cubero, JQ Gan, and Ramaswamy Palaniappan. Multiresolution analysis over simple graphs for brain computer interfaces. *Journal of neural engineering*, 10 (4), (2013), 046014.

[15] M. S. Baptista, E. J. Ngamga, P. R. F. Pinto, M. Brito, and J. Kurths. Kolmogorov-Sinai entropy from recurrence times. *Physics Letters A*, 374 (9), (2010), 1135–1140.

[16] Roberto Barrio, Fernando Blesa, and Sergio Serrano. Qualitative analysis of the Rössler equations: Bifurcations of limit cycles and chaotic attractors. *Physica D: Nonlinear Phenomena*, 238 (13), (2009), 1087–1100.

[17] A. Barron, J. Rissanen, and Bin Yu. The minimum description length principle in coding and modeling. *IEEE Transactions on Information Theory*, 44 (6), (1998), 2743–2760.

[18] C. M. Bauer, F. M. Rast, M. J. Ernst, A. Meichtry, J. Kool, S. M. Rissanen, J. H. Suni, and M. Kankaanpää. The effect of muscle fatigue and low back pain on lumbar movement variability and complexity. *Journal of Electromyography and Kinesiology*, 33, (2017), 94–102.

[19] Alim Louis Benabid, Thomas Costecalde, Andrey Eliseyev, Guillaume Charvet, Alexandre Verney, Serpil Karakas, Michael Foerster, Aurélien Lambert, Boris Morinière, Neil Abroug, et al. An exoskeleton controlled by an epidural wireless brain–machine interface in a tetraplegic patient: a proof-of-concept demonstration. *The Lancet Neurology*, (2019).

[20] Bernhard Bereiter, Sarah Eggleston, Jochen Schmitt, Christoph Nehrbass-Ahles, Thomas F. Stocker, Hubertus Fischer, Sepp Kipfstuhl, and Jerome Chappellaz. Revision of the EPICA Dome C CO_2 record from 800 to 600kyr before present. *Geophysical Research Letters*, 42 (2), (2015), 542–549.

[21] Melissa A. Berke, Thomas C. Johnson, Josef P. Werne, Daniel A. Livingstone, Kliti Grice, Stefan Schouten, and Jaap S. Sinninghe Damsté. Characterization of the last deglacial transition in tropical East Africa: Insights from Lake Albert. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 409, (2014), 1–8.

[22] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang. Complex networks: Structure and dynamics. *Physics Reports*, 424 (4), (2006), 175–308.

[23] Raymonde Bonnefille. Cenozoic vegetation, climate changes and hominid evolution in tropical Africa. *Global and Planetary Change*, 72 (4), (2010). Quaternary and Global Change: Review and Issues Special issue in memory of Hugues FAURE, 390–411.

[24] Elizabeth Bradley and Holger Kantz. Nonlinear time-series analysis revisited. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25 (9), (2015), 097610.

[25] Steven A. Brandt, Erich C. Fisher, Elisabeth A. Hildebrand, Ralf Vogelsang, Stanley H. Ambrose, Joséphine Lesur, and Hong Wang. Early MIS 3 occupation of Mochena Borago Rockshelter, Southwest Ethiopian Highlands: Implications for Late Pleistocene archaeology, paleoenvironments and modern human dispersals. *Quaternary International*, 274, (2012). Temporal and spatial corridors of Homo sapiens sapiens population dynamics during the Late Pleistocene and Early Holocene, 38–54.

[26] Nick Brooks. Cultural responses to aridity in the Middle Holocene and increased social complexity. *Quaternary International*, 151 (1), (2006). Dark nature: responses of humans and ecosystems to rapid environmental changes, 29–49.

[27] D.S. Broomhead and Gregory P. King. Extracting qualitative dynamics from experimental data. *Physica D: Nonlinear Phenomena*, 20 (2), (1986), 217–236.

[28] Victor Brovkin and Martin Claussen. Comment on "Climate-Driven Ecosystem Succession in the Sahara: The Past 6000 Years". *Science*, 322 (5906), (2008), 1326–1326.

[29] E. T. Brown, T. C. Johnson, C. A. Scholz, A. S. Cohen, and J. W. King. Abrupt change in tropical African climate linked to the bipolar seesaw over the past 55,000 years. *Geophysical Research Letters*, 34 (20), (2007).

[30] Clemens Brunner, Muhammad Naeem, Robert Leeb, Bernhard Graimann, and Gert Pfurtscheller. Spatial filtering and selection of optimized components in four class motor imagery EEG data using independent components analysis. *Pattern recognition letters*, 28 (8), (2007), 957–964.

[31] Steven L. Brunton and J. Nathan Kutz. *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control.* Cambridge University Press, 2019.

[32] Alejandro Builes-Jaramillo, Norbert Marwan, Germán Poveda, and Jürgen Kurths. Non-linear interactions between the Amazon River basin and the Tropical North Atlantic at interannual timescales. *Climate Dynamics*, 50 (7), (2018), 2951–2969.

[33] Th Buzug, T Reimers, and G Pfister. Optimal Reconstruction of Strange Attractors from Purely Geometrical Arguments. *Europhysics Letters (EPL)*, 13 (7), (Dec. 1990), 605–610.

[34] Th. Buzug and G. Pfister. Comparison of algorithms calculating optimal embedding parameters for delay time coordinates. *Physica D: Nonlinear Phenomena*, 58 (1), (1992), 127–137.

[35] Th. Buzug and G. Pfister. Optimal delay time and embedding dimension for delay-time coordinates by analysis of the global static and local dynamical behavior of strange attractors. *Phys. Rev. A*, 45, (10 May 1992), 7073–7084.

[36] Wei-Dong Cai, Yi-Qing Qin, and Bing-Ru Yang. Determination of phase-space reconstruction parameters of chaotic time series. *Kybernetika*, 44 (4), (2008), 557–570.

[37] Christopher Campisano et al. The Hominin Sites and PaleolakesDrilling Project: High-Resolution Paleoclimate Records from the East African Rift System and Their Implicationsfor Understanding the Environmental Context of Hominin Evolution. *Paleoanthropology*, (2017), 1–34.

[38] Mark A. Cane and Peter Molnar. Closing of the Indonesian seaway as a precursor to east African aridification around 3–4 million years ago. *Nature*, 411 (6834), (2001), 157–162.

[39] Liangyue Cao. Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D: Nonlinear Phenomena*, 110 (1), (1997), 43–50.

[40] Liangyue Cao, Alistair Mees, and Kevin Judd. Dynamics from multivariate time series. *Physica D: Nonlinear Phenomena*, 121 (1), (1998), 75–88.

[41] Sean M. Carroll and Jennifer Chen. Does inflation provide natural initial conditions for the universe? *International Journal of Modern Physics D*, 14 (12), (2005), 2335–2339.

[42] Shannon L. Carto, Andrew J. Weaver, René Hetherington, Yin Lam, and Edward C. Wiebe. Out of Africa and into an ice age: on the role of global climate change in the late Pleistocene migration of early modern humans out of Africa. *Journal of Human Evolution*, 56 (2), (2009), 139–151.

Bibliography

[43] M. C. Casdagli. Recurrence plots revisited. *Physica D*, 108 (1–2), (1997), 12–44.

[44] Martin Casdagli. Nonlinear prediction of chaotic time series. *Physica D: Nonlinear Phenomena*, 35 (3), (1989), 335–356.

[45] Martin Casdagli, Stephen Eubank, J.Doyne Farmer, and John Gibson. State space reconstruction in the presence of noise. *Physica D: Nonlinear Phenomena*, 51 (1), (1991), 52–98.

[46] Isla S. Castañeda, Stefan Mulitza, Enno Schefuß, Raquel A. Lopes dos Santos, Jaap S. Sinninghe Damsté, and Stefan Schouten. Wet phases in the Sahara/Sahel region and human migration patterns in North Africa. *Proceedings of the National Academy of Sciences*, 106 (48), (2009), 20159–20163.

[47] C. J. Cellucci, A. M. Albano, and P. E. Rapp. Comparative study of embedding methods. *Phys. Rev. E*, 67, (6 June 2003), 066210.

[48] F. Censi, G. Calcagnini, and S. Cerutti. Proposed corrections for the quantification of coupling patterns by recurrence plots. *IEEE Transactions on Biomedical Engineering*, 51 (5), (2004), 856–859.

[49] A. Čenys and K. Pyragas. Estimation of the number of degrees of freedom from chaotic time series. *Physics Letters A*, 129 (4), (1988), 227–230.

[50] Ujwal Chaudhary, Niels Birbaumer, and Ander Ramos-Murguialday. Brain–computer interfaces for communication and rehabilitation. *Nature Reviews Neurology*, 12 (9), (2016), 513.

[51] Y. Chen and H. Yang. Heterogeneous recurrence representation and quantification of dynamic transitions in continuous nonlinear processes. *The European Physical Journal B*, 89 (6), (2016), 155.

[52] Wing H. Cheung, Gabriel B. Senay, and Ashbindu Singh. Trends and spatial distribution of annual and seasonal rainfall in Ethiopia. *International Journal of Climatology*, 28 (13), (2008), 1723–1734.

[53] Parth Chholak, Guiomar Niso, Vladimir A Maksimenko, Semen A Kurkin, Nikita S Frolov, Elena N Pitsik, Alexander E Hramov, and Alexander N Pisarchik. Visual and kinesthetic modes affect motor imagery classification in untrained subjects. *Scientific reports*, 9 (1), (2019), 1–12.

[54] J. M. Choi, B. H. Bae, and S. Y. Kim. Divergence in perpendicular recurrence plot; quantification of dynamical divergence from short chaotic time series. *Physics Letters A*, 263 (4–6), (1999), 299–306.

[55] Adam Thomas Clark, Hao Ye, Forest Isbell, Ethan R. Deyle, Jane Cowles, G. David Tilman, and George Sugihara. Spatial convergent cross mapping to detect causal relationships from short time series. *Ecology*, 96 (5), (2015), 1174–1181.

[56] Joanne Clarke et al. Climatic changes and social transformations in the Near East and North Africa during the 'long' 4th millennium BC: A comparative study of environmental and archaeological evidence. *Quaternary Science Reviews*, 136, (2016). Special Issue: Mediterranean Holocene Climate, Environment and Human Societies, 96–121.

[57] A. Cohen et al. The Hominin Sites and Paleolakes Drilling Project: inferring the environmental context of human evolution from eastern African rift lake deposits. *Scientific Drilling*, 21, (2016), 1–16.

[58] A. Córdoba. Dirac combs. *Letters in Mathematical Physics*, 17 (3), (1989), 191–196.

[59] Rémi Coulom. Efficient Selectivity and Backup Operators in Monte-Carlo Tree Search. In: *Computers and Games*. Ed. by H. Jaap van den Herik, Paolo Ciancarini, and H. H. L. M. (Jeroen) Donkers. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 72–83.

[60] Jifeng Cui, Wenyu Zhang, Zeng Liu, and Jianglong Sun. On the limit cycles, period-doubling, and quasi-periodic solutions of the forced Van der Pol-Duffing oscillator. *Numerical Algorithms*, 78 (4), (2018), 1217–1231.

[61] M. O. Cuthbert, T. Gleeson, S. C. Reynolds, M. R. Bennett, A. C. Newton, C. J. McCormack, and G. M. Ashley. Modelling the role of groundwater hydro-refugia in East African hominin evolution and dispersal. *Nature Communications*, 8 (1), (2017), 15696.

[62] George Datseris. DynamicalSystems.jl: A Julia software library for chaos and nonlinear dynamics. *Journal of Open Source Software*, 3 (23), (Mar. 2018), 598.

[63] A. Davidson. The Omo river project, reconnaissance geology and geochemistry of parts of Ilubabor, Kefa, Gemu Gofa and Sidamo, Ethiopia. *Ethiop. Inst. Geol. Surv. Bull.*, 2, (1983), 1–89.

[64] Peter deMenocal, Joseph Ortiz, Tom Guilderson, Jess Adkins, Michael Sarnthein, Linda Baker, and Martha Yarusinsky. Abrupt onset and termination of the African Humid Period:: rapid climate responses to gradual insolation forcing. *Quaternary Science Reviews*, 19 (1), (2000), 347–361.

[65] Peter B deMenocal. African climate change and faunal evolution during the Pliocene–Pleistocene. *Earth and Planetary Science Letters*, 220 (1), (2004), 3–24.

[66] Peter B. deMenocal. Plio-Pleistocene African Climate. *Science*, 270 (5233), (1995), 53–59.

[67] Ethan R Deyle and George Sugihara. Generalized theorems for nonlinear state space reconstruction. *PloS one*, 6 (3), (Mar. 2011), e18295–e18295.

[68] Jayesh Dhadphale, Vishnu R. Unni, Abhishek Saha, and R. I. Sujith. Neural ODE to model and prognose thermoacoustic instability. 2021.

[69] Peter D. Ditlevsen and Sigfus J. Johnsen. Tipping points: Early warning and wishful thinking. *Geophysical Research Letters*, 37 (19), (2010).

[70] J. Donath. Untersuchung alternativer Zeitfensterformen für die quantitative Rekurrenzanalyse anhand von Modellsystemen und Klimadaten. Bachelor Thesis. Humboldt Universität zu Berlin, May 2016.

[71] J. F. Donges, R. V. Donner, K. Rehfeld, N. Marwan, M. H. Trauth, and J. Kurths. Identification of dynamical transitions in marine palaeoclimate records by recurrence network analysis. *Nonlinear Processes in Geophysics*, 18, (2011), 545–562.

[72] J. F. Donges, R. V. Donner, M. H. Trauth, N. Marwan, H. J. Schellnhuber, and J. Kurths. Nonlinear detection of paleoclimate-variability transitions possibly related to human evolution. *Proceedings of the National Academy of Sciences*, 108 (51), (2011), 20422–20427.

[73] J. F. Donges, J. Heitzig, R. V. Donner, and J. Kurths. Analytical framework for recurrence network analysis of time series. *Physical Review E*, 85, (2012), 046105.

[74] R. V. Donner, M. Small, J. F. Donges, N. Marwan, Y. Zou, R. Xiang, and J. Kurths. Recurrence-based time series analysis by means of complex network methods. *International Journal of Bifurcation and Chaos*, 21 (4), (2011), 1019–1046.

Bibliography

[75] Walter Duesing, Nadine Berner, Alan L. Deino, Verena Foerster, K. Hauke Kraemer, Norbert Marwan, and Martin H. Trauth. Multiband Wavelet Age Modeling for a ~ 293 m (~ 600 kyr) Sediment Core From Chew Bahir Basin, Southern Ethiopian Rift. *Frontiers in Earth Science*, 9, (2021), 35.

[76] Georg Duffing. *Erzwungene Schwingungen bei veraenderlicher Eigenfrequenz und ihre technische Bedeutung*. F. Vieweg & Sohn, 1918.

[77] J.-P. Eckmann, S. Oliffson Kamphorst, and D. Ruelle. Recurrence Plots of Dynamical Systems. *Europhysics Letters*, 4 (9), (1987), 973–977.

[78] J.-P. Eckmann, S. Oliffson Kamphorst, and D. Ruelle. Recurrence Plots of Dynamical Systems. *Europhysics Letters*, 4 (9), (1987), 973–977.

[79] Armin Eftekhari, Han Lun Yap, Michael B. Wakin, and Christopher J. Rozell. Stabilizing embedology: Geometry-preserving delay-coordinate maps. *Phys. Rev. E*, 97, (2 Feb. 2018), 022222.

[80] D. Eroglu, N. Marwan, S. Prasad, and J. Kurths. Finding recurrence networks' threshold adaptively for a specific time series. *Nonlinear Processes in Geophysics*, 21, (2014), 1085–1092.

[81] D. Eroglu, F. H. McRobie, I. Ozken, T. Stemler, K.-H. Wyrwoll, S. F. M. Breitenbach, N. Marwan, and J. Kurths. See-saw relationship of the Holocene East Asian-Australian summer monsoon. *Nature Communications*, 7, (2016), 12929.

[82] D. Eroglu, T. K. D. Peron, N. Marwan, F. A. Rodrigues, L. d. F. Costa, M. Sebek, I. Z. Kiss, and J. Kurths. Entropy of weighted recurrence plots. *Physical Review E*, 90, (2014), 042919.

[83] A. Facchini and H. Kantz. Curved structures in recurrence plots: The role of the sampling time. *Physical Review E*, 75, (2007), 036215.

[84] A. Facchini, F. Rossi, and C. Mocenni. Spatial recurrence strategies reveal different routes to Turing pattern formation in chemical systems. *Physics Letters A*, 373 (46), (2009), 4266–4272.

[85] J. D. Farmer. Chaotic attractors of an infinite-dimensional dynamical system. *Physica D: Nonlinear Phenomena*, 4 (3), (1982), 366–393.

[86] J. Doyne Farmer and John J. Sidorowich. Predicting chaotic time series. *Phys. Rev. Lett.*, 59, (8 Aug. 1987), 845–848.

[87] P. Faure and H. Korn. A new method to estimate the Kolmogorov entropy from recurrence plots: its application to neuronal signals. *Physica D*, 122 (1–4), (1998), 265–279.

[88] J. H. Feldhoff, R. V. Donner, J. F. Donges, N. Marwan, and J. Kurths. Geometric detection of coupling directions by means of inter-system recurrence networks. *Physics Letters A*, 376 (46), (2012), 3504–3513.

[89] Pascal Flohr, Dominik Fleitmann, Roger Matthews, Wendy Matthews, and Stuart Black. Evidence of resilience to past climate change in Southwest Asia: Early farming communities and the 9.2 and 8.2 ka events. *Quaternary Science Reviews*, 136, (2016). Special Issue: Mediterranean Holocene Climate, Environment and Human Societies, 23–39.

[90] V. Foerster et al. 620,000 years of eastern African climate variability and hominin evolution. *Nature Geoscience, submitted.*, (2021).

[91]   Verena Foerster, Daniel M. Deocampo, Asfawossen Asrat, Christina Günter, Annett Junginger, Kai Hauke Krämer, Nicole A. Stroncik, and Martin H. Trauth. Towards an understanding of climate proxy formation in the Chew Bahir basin, southern Ethiopian Rift. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 501, (2018), 111–123.

[92]   Verena Foerster, Ralf Vogelsang, Annett Junginger, Asfawossen Asrat, Henry F. Lamb, Frank Schaebitz, and Martin H. Trauth. Environmental change and human occupation of southern Ethiopia and northern Kenya during the last 20,000 years. *Quaternary Science Reviews*, 129, (2015), 333–340.

[93]   Verena Foerster et al. Climatic change recorded in the sediments of the Chew Bahir basin, southern Ethiopia, during the last 45,000 years. *Quaternary International*, 274, (2012). Temporal and spatial corridors of Homo sapiens sapiens population dynamics during the Late Pleistocene and Early Holocene, 25–37.

[94]   Andrew M. Fraser. Reconstructing attractors from scalar time series: A comparison of singular system and redundancy criteria. *Physica D: Nonlinear Phenomena*, 34 (3), (1989), 391–404.

[95]   Andrew M. Fraser and Harry L. Swinney. Independent coordinates for strange attractors from mutual information. *Phys. Rev. A*, 33, (2 Feb. 1986), 1134–1140.

[96]   David Freedman and Persi Diaconis. On the Histogram as a Density Estimator: L_2 Theory. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 57, (1981), 453–476.

[97]   F. N. Fritsch and R. E. Carlson. Monotone Piecewise Cubic Interpolation. *SIAM Journal on Numerical Analysis*, 17 (2), (1980), 238–246.

[98]   Alexander A Frolov, Olesya Mokienko, Roman Lyukmanov, Elena Biryukova, Sergey Kotov, Lydia Turbina, Georgy Nadareyshvily, and Yulia Bushkova. Post-stroke rehabilitation training with a motor-imagery-based brain-computer interface (BCI)-controlled hand exoskeleton: a randomized controlled multicenter trial. *Frontiers in neuroscience*, 11, (2017), 400.

[99]   Julia Galway-Witham, James Cole, and Chris Stringer. Aspects of human physical and behavioural evolution during the last 1 million years. *Journal of Quaternary Science*, 34 (6), (2019), 355–378.

[100]   J. B. Gao. Recurrence Time Statistics for Chaotic Systems and Their Applications. *Physical Review Letters*, 83 (16), (1999), 3178–3181.

[101]   J. B. Gao and H. Q. Cai. On the structures and quantification of recurrence plots. *Physics Letters A*, 270 (1–2), (2000), 75–87.

[102]   Jianbo Gao and Zhemin Zheng. Local exponential divergence plot and optimal embedding of a chaotic time series. *Physics Letters A*, 181 (2), (1993), 153–158.

[103]   Jianbo Gao and Zhemin Zheng. Direct dynamical test for deterministic chaos and optimal embedding of a chaotic time series. *Phys. Rev. E*, 49, (5 May 1994), 3807–3814.

[104]   Sara P. Garcia and Jonas S. Almeida. Multivariate phase space reconstruction by nearest neighbor embedding with different time delays. *Phys. Rev. E*, 72, (2 Aug. 2005), 027205.

[105]   Sara P. Garcia and Jonas S. Almeida. Nearest neighbor embedding with different time delays. *Phys. Rev. E*, 71, (3 Mar. 2005), 037204.

Bibliography

[106] Yannick Garcin, Daniel Melnick, Manfred R. Strecker, Daniel Olago, and Jean-Jacques Tiercelin. East African mid-Holocene wet–dry transition recorded in palaeo-shorelines of Lake Turkana, northern Kenya Rift. *Earth and Planetary Science Letters*, 331-332, (2012), 322–334.

[107] Yannick Garcin et al. Early anthropogenic impact on Western Central African rainforests 2,600 y ago. *Proceedings of the National Academy of Sciences*, 115 (13), (2018), 3261–3266.

[108] Joshua Garland and Elizabeth Bradley. Prediction in projection. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25 (12), (2015), 123108.

[109] Françoise Gasse. Hydrological changes in the African tropics since the Last Glacial Maximum. *Quaternary Science Reviews*, 19 (1), (2000), 189–211.

[110] Maria Carmela Gatto and Andrea Zerboni. Holocene Supra-Regional Environmental Changes as Trigger for Major Socio-Cultural Processes in Northeastern Africa and the Sahara. *African Archaeological Review*, 32 (2), (2015), 301–333.

[111] John F. Gibson, J. Doyne Farmer, Martin Casdagli, and Stephen Eubank. An analytic approach to practical state space reconstruction. *Physica D: Nonlinear Phenomena*, 57 (1), (1992), 1–30.

[112] Vedasri Godavarthi, Samadhan A. Pawar, Vishnu R. Unni, R. I. Sujith, Nobert Marwan, and Jürgen Kurths. Coupled interaction between unsteady flame dynamics and acoustic field in a turbulent combustor. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28 (11), (2018), 113111.

[113] B. Goswami, N. Marwan, G. Feulner, and J. Kurths. How do global temperature drivers influence each other? – A network perspective using recurrences. *European Physical Journal – Special Topics*, 222, (2013), 861–873.

[114] C. W. J. Granger. Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37 (3), (1969), 424–438.

[115] Katharine M. Grant, Eelco J. Rohling, Thomas Westerhold, Matthias Zabel, David Heslop, Tiuri Konijnendijk, and Lucas Lourens. A 3 million year index for North African humidity/aridity and the implication of potential pan-African Humid periods. *Quaternary Science Reviews*, 171, (2017), 100–118.

[116] P. Grassberger and I. Procaccia. Estimation of the Kolmogorov entropy from a chaotic signal. *Physical Review A*, 9 (1–2), (1983), 2591–2593.

[117] Peter Grassberger and Itamar Procaccia. Estimation of the Kolmogorov entropy from a chaotic signal. *Phys. Rev. A*, 28, (4 Oct. 1983), 2591–2593.

[118] Peter Grassberger and Itamar Procaccia. Measuring the strangeness of strange attractors. *Physica D: Nonlinear Phenomena*, 9 (1), (1983), 189–208.

[119] Peter Grassberger, Thomas Schreiber, and Carsten Schaffrath. Nonlinear time sequence analysis. *International Journal of Bifurcation and Chaos*, 01 (03), (1991), 521–547.

[120] Vadim Grubov, V Yu Musatov, VA Maksimenko, AN Pisarchik, AE Runnova, and AE Hramov. Development of intelligent system for classification of multiple human brain states corresponding to different real and imaginary movements. *Cybernetics and Physics*, 6, (2017), 103–107.

[121] K. Guhathakurta, B. Bhattacharya, and A. R. Chowdhury. Using recurrence plot analysis to distinguish between endogenous and exogenous stock market crashes. *Physica A*, 389 (9), (2010), 1874–1882.

[122] Z. O. Guimarães-Filho, I. L. Caldas, R. L. Viana, J. Kurths, I. C. Nascimento, and Y. K. Kuznetsov. Recurrence quantification analysis of electrostatic fluctuations in fusion plasmas. *Physics Letters A*, 372 (7), (2008), 1088–1095.

[123] M. Han, W. Ren, M. Xu, and T. Qiu. Nonuniform State Space Reconstruction for Multivariate Chaotic Time Series. *IEEE Transactions on Cybernetics*, 49 (5), (2019), 1885–1895.

[124] R. Hegger, H. Kantz, L. Matassini, and T. Schreiber. Coping with Nonstationarity by Overembedding. *Physical Review Letters*, 84 (18), (2000), 4092–4095.

[125] Rainer Hegger and Holger Kantz. Improved false nearest neighbor method to detect determinism in time series data. *Phys. Rev. E*, 60, (4 Oct. 1999), 4970–4973.

[126] M. Hénon. A two-dimensional mapping with a strange attractor. *Communications in Mathematical Physics*, 50 (1), (1976), 69–77.

[127] H.G.E. Hentschel and Itamar Procaccia. The infinite number of generalized dimensions of fractals and strange attractors. *Physica D: Nonlinear Phenomena*, 8 (3), (1983), 435–444.

[128] Elisabeth Anne Hildebrand and Katherine M. Grillo. Early herders and monumental sites in eastern Africa: dating and interpretation. *Antiquity*, 86 (332), (2012), 338–352.

[129] Y. Hirata, Y. Shimo, H. L. Tanaka, and K. Aihara. Chaotic Properties of the Arctic Oscillation Index. *SOLA*, 7, (2011), 33–36.

[130] Yoshito Hirata and Kazuyuki Aihara. Dimensionless embedding for nonlinear time series analysis. *Phys. Rev. E*, 96, (3 Sept. 2017), 032219.

[131] Yoshito Hirata, Hideyuki Suzuki, and Kazuyuki Aihara. Reconstructing state spaces from multivariate data using variable delays. *Phys. Rev. E*, 74, (2 Aug. 2006), 026202.

[132] Detlef Holstein and Holger Kantz. Optimal Markov approximations and generalized embeddings. *Phys. Rev. E*, 79, (5 May 2009), 056202.

[133] S. Horai, T. Yamada, and K. Aihara. Determinism Analysis with Iso-Directional Recurrence Plots. *IEEE Transactions - Institute of Electrical Engineers of Japan C*, 122 (1), (2002), 141–147.

[134] Alexander E Hramov, Vladimir A Maksimenko, Svetlana V Pchelintseva, Anastasiya E Runnova, Vadim V Grubov, Vyacheslav Yu Musatov, Maksim O Zhuravlev, Alexey A Koronovskii, and Alexander N Pisarchik. Classifying the perceptual interpretations of a bistable image using EEG and artificial neural networks. *Frontiers in neuroscience*, 11, (2017), 674.

[135] S. Ikegawa, M. Shinohara, T. Fukunaga, J. P. Zbilut, and C. L. Webber, Jr. Nonlinear time-course of lumbar muscle fatigue using recurrence quantifications. *Biological Cybernetics*, 82 (5), (2000), 373–382.

[136] Jonas Isensee, George Datseris, and Ulrich Parlitz. Predicting Spatio-temporal Time Series Using Dimension Reduced Local States. *Journal of Nonlinear Science*, 30 (3), (2020), 713–735.

Bibliography

[137] Sarah J. Ivory and James Russell. Lowland forest collapse and early human impacts at the end of the African Humid Period at Lake Edward, equatorial East Africa. *Quaternary Research*, 89 (1), (2018), 7–20.

[138] J. S. Iwanski and E. Bradley. Recurrence plots of experimental data: To embed or not to embed? *Chaos*, 8 (4), (1998), 861–871.

[139] Rinku Jacob, K. P. Harikrishnan, R. Misra, and G. Ambika. Uniform framework for the recurrence-network analysis of chaotic time series. *Physical Review E*, 93, (1 2016), 012202.

[140] Ziyu Jia, Youfang Lin, Yunxiao Liu, Zehui Jiao, and Jing Wang. Refined nonuniform embedding for coupling detection in multivariate time series. *Phys. Rev. E*, 101, (6 June 2020), 062113.

[141] Kevin Judd and Alistair Mees. Embedding as a modeling problem. *Physica D: Nonlinear Phenomena*, 120 (3), (1998), 273–286.

[142] Kevin Judd and Alistair Mees. Embedding as a modeling problem. *Physica D: Nonlinear Phenomena*, 120 (3), (1998), 273–286.

[143] Annett Junginger and Martin H. Trauth. Hydrological constraints of paleo-Lake Suguta in the Northern Kenya Rift during the African Humid Period (15–5kaBP). *Global and Planetary Change*, 111, (2013), 174–188.

[144] Matthew P Juniper and Raman I Sujith. Sensitivity and nonlinearity of thermoacoustic oscillations. *Annual Review of Fluid Mechanics*, 50, (2018), 661–689.

[145] Tae-Eui Kam, Heung-Il Suk, and Seong-Whan Lee. Non-homogeneous spatial filter optimization for ElectroEncephaloGram (EEG)-based motor imagery classification. *Neurocomputing*, 108, (2013), 58–68.

[146] H Kantz and T Schreiber. *Nonlinear Time Series Analysis*. Cambridge nonlinear science series. Cambridge University Press, 2004.

[147] Holger Kantz. A robust method to estimate the maximal Lyapunov exponent of a time series. *Physics Letters A*, 185 (1), (1994), 77–87.

[148] Holger Kantz. Time series analysis in reconstructed state spaces. *Stochastics and Dynamics*, 01 (01), (2001), 85–111.

[149] Holger Kantz, Thomas Schreiber, Ingo Hoffmann, Thorsten Buzug, Gerd Pfister, Leci G. Flepp, Josef Simonet, Remo Badii, and Ernst Brun. Nonlinear noise reduction: A case study on experimental data. *Phys. Rev. E*, 48, (2 Aug. 1993), 1529–1538.

[150] Holger Kantz and Thomas Schürmann. Enlarged scaling ranges for the KS-entropy and the information dimension. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 6 (2), (1996), 167–171.

[151] A. Karimi and M. R. Paul. Extensive chaos in the Lorenz-96 model. *Chaos*, 20 (4), (2010), 043105.

[152] G. Kember and A.C. Fowler. A correlation function for choosing time delays in phase portrait reconstructions. *Physics Letters A*, 179 (2), (1993), 72–80.

[153] Matthew B. Kennel and Henry D. I. Abarbanel. False neighbors and false strands: A reliable minimum embedding dimension algorithm. *Phys. Rev. E*, 66, (2 Aug. 2002), 026209.

[154] Matthew B. Kennel, Reggie Brown, and Henry D. I. Abarbanel. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys. Rev. A*, 45, (6 Mar. 1992), 3403–3411.

[155] H.S. Kim, R. Eykholt, and J.D. Salas. Nonlinear dynamics, delay times, and embedding windows. *Physica D: Nonlinear Phenomena*, 127 (1), (1999), 48–60.

[156] Anton R Kiselev, Vladimir A Maksimenko, Nikolay Shukovskiy, Alexander N Pisarchik, Elena Pitsik, and Alexander E Hramov. Post-stroke rehabilitation with the help of brain-computer interface. In: *2019 3rd School on Dynamics of Complex Networks and their Application in Intellectual Robotics (DCNAIR)*. IEEE. 2019, pp. 83–85.

[157] István Z. Kiss, Wen Wang, and John L. Hudson. Complexity of globally coupled chaotic electrochemical oscillators. *Phys. Chem. Chem. Phys.*, 2, (2000), 3847–3854.

[158] Richard G. Klein. Anatomy, behavior, and modern human origins. *Journal of World Prehistory*, 9 (2), (1995), 167–198.

[159] Richard G. Klein and Teresa E. Steele. Archaeological shellfish size and later human evolution in Africa. *Proceedings of the National Academy of Sciences*, 110 (27), (2013), 10910–10915.

[160] M. Koebbe and G. Mayer-Kress. Use of Recurrence Plots in the Analysis of Time-Series Data. In: *Proceedings of SFI Studies in the Science of Complexity*. Ed. by M. Casdagli and S. Eubank. Vol. XXI. Redwood City: Addison-Wesley, 1992, pp. 361–378.

[161] O. Kopáček, V. Karas, J. Kovář, and Z. Stuchlík. Transition from regular to chaotic circulation in magnetized coronae near compact objects. *The Astrophysical Journal*, 722 (2), (2010), 1240.

[162] K H Kraemer, G Datseris, J Kurths, I Z Kiss, J L Ocampo-Espindola, and N Marwan. A unified and automated approach to attractor reconstruction. *New Journal of Physics*, 23 (3), (Mar. 2021), 033017.

[163] K. Hauke Kraemer, Reik V. Donner, Jobst Heitzig, and Norbert Marwan. Recurrence threshold selection for obtaining robust recurrence characteristics in different embedding dimensions. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28 (8), (2018), 085720.

[164] K. Hauke Kraemer and Norbert Marwan. Border effect corrections for diagonal line based recurrence quantification analysis measures. *Physics Letters A*, 383 (34), (2019), 125977.

[165] Kai Hauke Kraemer. *hkraemer/Border-effect-corrections-for-diagonal-line-based-recurrence-quantification-analysis-measures: MATLAB routines for RQA border effect and tangential motion correction*. Version 2.1. Dec. 2019.

[166] Kai Hauke Kraemer and George Datseris. *hkraemer/PECUZAL Julia: Reproducible code base for PECUZAL embedding*. Version 1.1.1. Jan. 2021.

[167] Kai Hauke Kraemer, Maximilian Gelbrecht, Induja Pavithran, R. I. Sujith, and Norbert Marwan. Optimal state space reconstruction via Monte Carlo Decision Tree Search. *Nonlinear Dynamics*, (2022).

[168] Kai Hauke Kraemer and Norbert Marwan. *hkraemer/PECUZAL Matlab: PECUZAL embedding algorithm*. Version 1.0.1. Jan. 2021.

Bibliography

[169] Kai Hauke Kraemer and Norbert Marwan. *hkraemer/PECUZAL python: PECUZAL embedding algorithm.* Version 1.0. Jan. 2021.

[170] Anna Krakovská, Kristína Mezeiová, and Hana Budáčová. Use of False Nearest Neighbours for Selecting Variables and Embedding Parameters for State Space Reconstruction. *Journal of Complex Systems*, 2015, (2015), 932750.

[171] S. Kröpelin, D. Verschuren, and A.-M. Lézine. Response to Comment on "Climate-Driven Ecosystem Succession in the Sahara: The Past 6000 Years". *Science*, 322 (5906), (2008), 1326–1326.

[172] S. Kröpelin et al. Climate-Driven Ecosystem Succession in the Sahara: The Past 6000 Years. *Science*, 320 (5877), (2008), 765–768.

[173] D Kugiumtzis. State space reconstruction parameters in the analysis of chaotic time series — the role of the time window length. *Physica D: Nonlinear Phenomena*, 95 (1), (1996), 13–28.

[174] Rudolph Kuper and Stefan Kröpelin. Climate-Controlled Holocene Occupation in the Sahara: Motor of Africa's Evolution. *Science*, 313 (5788), (2006), 803–807.

[175] Henry F. Lamb, C. Richard Bates, Charlotte L. Bryant, Sarah J. Davies, Dei G. Huws, Michael H. Marshall, Helen M. Roberts, and Harry Toland. 150,000-year palaeoclimate record from northern Ethiopia supports early, multiple dispersals of modern humans from Africa. *Scientific Reports*, 8 (1), (2018), 1077.

[176] Laskar, J., Robutel, P., Joutel, F., Gastineau, M., Correia, A. C. M., and Levrard, B. A long-term numerical solution for the insolation quantities of the Earth. *A&A*, 428 (1), (2004), 261–285.

[177] M. R. Leadbetter, G. Lindgren, and H. Rootzen. *Extremes and Related Properties of Random Sequences and Processes.* Springer, New York, 1983.

[178] Mikhail A Lebedev and Miguel AL Nicolelis. Brain–machine interfaces: past, present and future. *TRENDS in Neurosciences*, 29 (9), (2006), 536–546.

[179] Mikhail A Lebedev and Miguel AL Nicolelis. Brain-machine interfaces: From basic science to neuroprostheses and neurorehabilitation. *Physiological reviews*, 97 (2), (2017), 767–837.

[180] Jaqueline Lekscha and Reik V. Donner. Phase space reconstruction for non-uniformly sampled noisy time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28 (8), (2018), 085702.

[181] Timothy M. Lenton, Hermann Held, Elmar Kriegler, Jim W. Hall, Wolfgang Lucht, Stefan Rahmstorf, and Hans Joachim Schellnhuber. Tipping elements in the Earth's climate system. *Proceedings of the National Academy of Sciences*, 105 (6), (2008), 1786–1793.

[182] C Letellier, J Maquet, L Le Sceller, G Gouesbet, and L A Aguirre. On the non-equivalence of observables in phase-space reconstructions from recorded time series. *Journal of Physics A: Mathematical and General*, 31 (39), (Oct. 1998), 7913–7927.

[183] C. Letellier, I. M. Moroz, and R. Gilmore. Comparison of tests for embeddings. *Phys. Rev. E*, 78, (2 Aug. 2008), 026203.

[184] Christophe Letellier and Luis A. Aguirre. Investigating nonlinear dynamics from time series: The influence of symmetries and the choice of observables. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 12 (3), (2002), 549–558.

[185] W Liebert, K Pawelzik, and H. G Schuster. Optimal Embeddings of Chaotic Attractors from Topological Considerations. *Europhysics Letters (EPL)*, 14 (6), (Mar. 1991), 521–526.

[186] W. Liebert and H.G. Schuster. Proper choice of the time delay for the analysis of chaotic time series. *Physics Letters A*, 142 (2), (1989), 107–111.

[187] Max A. Little, Patrick E. McSharry, Stephen J. Roberts, Declan AE Costello, and Irene M. Moroz. Exploiting Nonlinear Recurrence and Fractal Scaling Properties for Voice Disorder Detection. *BioMedical Engineering OnLine*, 6 (1), (2007), 23.

[188] E.N. Lorenz. Predictability: a problem partly solved. In: *Seminar on Predictability, 4-8 September 1995*. Vol. 1. ECMWF. Shinfield Park, Reading: ECMWF, 1995, pp. 1–18.

[189] Edward N. Lorenz. Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences*, 20 (2), (Mar. 1963), 130–141.

[190] Zhixin Lu, Brian R. Hunt, and Edward Ott. Attractor reconstruction by machine learning. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28 (6), (2018), 061104.

[191] Teng Ma, Hui Li, Lili Deng, Hao Yang, Xulin Lv, Peiyang Li, Fali Li, Rui Zhang, Tiejun Liu, Dezhong Yao, et al. The hybrid BCI system for movement control by combining motor imagery and moving onset visual evoked potential. *Journal of neural engineering*, 14 (2), (2017), 026015.

[192] MC Mackey and L Glass. Oscillation and chaos in physiological control systems. *Science*, 197 (4300), (1977), 287–289.

[193] Vladimir A Maksimenko, Semen A Kurkin, Elena N Pitsik, Vyacheslav Yu Musatov, Anastasia E Runnova, Tatyana Yu Efremova, Alexander E Hramov, and Alexander N Pisarchik. Artificial neural network classification of motor-related eeg: An increase in classification accuracy by reducing signal complexity. *Complexity*, 2018, (2018).

[194] Vladimir A Maksimenko, Annika Lüttjohann, Vladimir V Makarov, Mikhail V Goremyko, Alexey A Koronovskii, Vladimir Nedaivozov, Anastasia E Runnova, Gilles van Luijtelaar, Alexander E Hramov, and Stefano Boccaletti. Macroscopic and microscopic spectral properties of brain networks during local and global synchronization. *Physical Review E*, 96 (1), (2017), 012316.

[195] Jean Maley, Charles Doumenge, Pierre Giresse, Gil Mahé, Nathalie Philippon, Wannes Hubau, Michel O. Lokonda, John M. Tshibamba, and Alex Chepstow-Lusty. Late Holocene forest contraction and fragmentation in central Africa. *Quaternary Research*, 89 (1), (2018), 43–59.

[196] Ricardo Mañé. On the dimension of the compact invariant sets of certain non-linear maps. In: *Dynamical Systems and Turbulence, Warwick 1980*. Ed. by David Rand and Lai-Sang Young. Berlin, Heidelberg: Springer Berlin Heidelberg, 1981, pp. 230–242.

[197] B.P. Mann, F.A. Khasawneh, and R. Fales. Using information to generate derivative coordinates from noisy time series. *Communications in Nonlinear Science and Numerical Simulation*, 16 (8), (2011), 2999–3004.

[198] Katie Manning and Adrian Timpson. The demographic response to Holocene climate change in the Sahara. *Quaternary Science Reviews*, 101, (2014), 28–35.

[199] T. K. March, S. C. Chapman, and R. O. Dendy. Recurrence plot statistics and the effect of embedding. *Physica D*, 200 (1–2), (2005), 171–184.

Bibliography

[200] Rob Marchant et al. Drivers and trajectories of land cover change in East Africa: Human and environmental interactions from 6000 years ago to present. *Earth-Science Reviews*, 178, (2018), 322–378.

[201] Eric Maris and Robert Oostenveld. Nonparametric statistical testing of EEG-and MEG-data. *Journal of neuroscience methods*, 164 (1), (2007), 177–190.

[202] N. Marwan. A Historical Review of Recurrence Plots. *European Physical Journal – Special Topics*, 164 (1), (2008), 3–12.

[203] N. Marwan. How to avoid potential pitfalls in recurrence plot based data analysis. *International Journal of Bifurcation and Chaos*, 21 (4), (2011), 1003–1017.

[204] 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*, 373 (46), (2009), 4246–4254.

[205] N. Marwan, S. Foerster, and J. Kurths. Analysing spatially extended high-dimensional dynamics by recurrence plots. *Physics Letters A*, 379, (2015), 894–900.

[206] N. Marwan and J. Kurths. Nonlinear analysis of bivariate data with cross recurrence plots. *Physics Letters A*, 302 (5–6), (2002), 299–307.

[207] N. Marwan and J. Kurths. Complex network based techniques to identify extreme events and (sudden) transitions in spatio-temporal systems. *Chaos*, 25, (2015), 097609.

[208] N. Marwan and A. Meinke. Extended recurrence plot analysis and its application to ERP data. *International Journal of Bifurcation and Chaos*, 14 (2), (Feb. 2004), 761–771.

[209] N. Marwan, M. C. Romano, M. Thiel, and J. Kurths. Recurrence Plots for the Analysis of Complex Systems. *Physics Reports*, 438 (5–6), (2007), 237–329.

[210] N. Marwan, S. Schinkel, and J. Kurths. Recurrence plots 25 years later – Gaining confidence in dynamical transitions. *Europhysics Letters*, 101, (2013), 20007.

[211] N. Marwan, M. Thiel, and N. R. Nowaczyk. Cross Recurrence Plot Based Synchronization of Time Series. *Nonlinear Processes in Geophysics*, 9 (3/4), (2002), 325–331.

[212] N. Marwan, M. H. Trauth, M. Vuille, and J. Kurths. Comparing modern and Pleistocene ENSO-like influences in NW Argentina using nonlinear time series analysis methods. *Climate Dynamics*, 21 (3–4), (2003), 317–326.

[213] N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, and J. Kurths. Recurrence Plot Based Measures of Complexity and its Application to Heart Rate Variability Data. *Physical Review E*, 66 (2), (2002), 026702.

[214] Mark A. Maslin and Martin H. Trauth. Plio-Pleistocene East African Pulsed Climate Variability and Its Influence on Early Human Evolution. In: *The First Humans – Origin and Early Evolution of the Genus Homo: Contributions from the Third Stony Brook Human Evolution Symposium and Workshop October 3 – October 7, 2006*. Ed. by Frederick E. Grine, John G. Fleagle, and Richard E. Leakey. Dordrecht: Springer Netherlands, 2009, pp. 151–158.

[215] L. Matassini, H. Kantz, J. A. Hołyst, and R. Hegger. Optimizing of recurrence plots for noise reduction. *Physical Review E*, 65 (2), (2002), 021102.

[216] Lorenzo Matassini, Holger Kantz, Janusz Hołyst, and Rainer Hegger. Optimizing of recurrence plots for noise reduction. *Phys. Rev. E*, 65, (2 Jan. 2002), 021102.

[217]   Mariano Matilla-García, Isidro Morales, Jose Miguel Rodríguez, and Manuel Ruiz Marín. Selection of Embedding Dimension and Delay Time in Phase Space Reconstruction via Symbolic Dynamics. *Entropy*, 23 (2), (2021).

[218]   Ian McDougall, Francis H. Brown, and John G. Fleagle. Stratigraphic placement and age of modern humans from Kibish, Ethiopia. *Nature*, 433 (7027), (2005), 733–736.

[219]   Dennis J McFarland and Jonathan R Wolpaw. Brain-computer interface operation of robotic and prosthetic devices. *Computer*, 41 (10), (2008), 52–56.

[220]   DJ McFarland and JR Wolpaw. EEG-based brain–computer interfaces. *current opinion in Biomedical Engineering*, 4, (2017), 194–200.

[221]   G. M. Mindlin and R. Gilmore. Topological analysis and synthesis of chaotic time series. *Physica D*, 58 (1–4), (1992), 229–242.

[222]   Marta Mirazón Lahr. The shaping of human diversity: filters, boundaries and transitions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371 (1698), (2016), 20150241.

[223]   V. Mitra, A. Sarma, M. S. Janaki, A. N. Sekar Iyengar, B. Sarma, N. Marwan, J. Kurths, P. K. Shaw, D. Saha, and S. Ghosh. Order to chaos transition studies in a DC glow discharge plasma by using recurrence quantification analysis. *Chaos, Solitons & Fractals*, 69, (2014), 285–293.

[224]   John M. Moore and A. Davidson. Rift structure in southern Ethiopia. *Tectonophysics*, 46 (1), (1978), 159–173.

[225]   Colin Morley and Iain Renfrew. Situating the creative explosion: universal or local? In: *Becoming human: innovation in prehistoric material and spiritual culture*. Ed. by Colin Morley and Iain Renfrew. Cambridge University Press, 2009. Chap. Situating the creative explosion: universal or local?, pp. 74–92.

[226]   Manfred Mudelsee. *Climate Time Series Analysis. Classical Statistical and Bootstrap Methods*. Springer, Cham, 2014.

[227]   V. Nair, G. Thampi, and R. I. Sujith. Intermittency route to thermoacoustic instability in turbulent combustors. *Journal of Fluid Mechanics*, 756, (2014), 470–487.

[228]   Vineeth Nair, Gireeshkumaran Thampi, and RI Sujith. Intermittency route to thermoacoustic instability in turbulent combustors. *Journal of Fluid Mechanics*, 756, (2014), 470–487.

[229]   E. J. Ngamga, S. Bialonski, N. Marwan, J. Kurths, C. Geier, and K. Lehnertz. Evaluation of selected recurrence measures in discriminating pre-ictal and inter-ictal periods from epileptic EEG data. *Physics Letters A*, 380 (16), (2016), 1419–1425.

[230]   E. J. Ngamga, D. V. Senthilkumar, A. Prasad, P. Parmananda, N. Marwan, and J. Kurths. Distinguishing dynamics using recurrence-time statistics. *Physical Review E*, 85 (2), (2012), 026217.

[231]   Chetan Nichkawde. Optimal state-space reconstruction using derivatives on projected manifold. *Phys. Rev. E*, 87, (2 Feb. 2013), 022905.

[232]   Sharon E. Nicholson. Climate and climatic variability of rainfall over eastern Africa. *Reviews of Geophysics*, 55 (3), (2017), 590–635.

Bibliography

[233] Gregoire Nicolis and Catherine Nicolis. *Foundations of Complex Systems: Nonlinear Dynamics, Statistical Physics, Information and Prediction.* World Scientific, 2007.

[234] Götz Ossendorf et al. Middle Stone Age foragers resided in high elevations of the glaciated Bale Mountains, Ethiopia. *Science,* 365 (6453), (2019), 583–587.

[235] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw. Geometry from a Time Series. *Phys. Rev. Lett.,* 45, (9 Sept. 1980), 712–716.

[236] Ulrich Parlitz. Identification of true and spurious Lyapunov exponents from time series. *International Journal of Bifurcation and Chaos,* 02 (01), (1992), 155–165.

[237] Ulrich Parlitz and Christian Merkwirth. Prediction of Spatiotemporal Time Series Based on Reconstructed Local States. *Phys. Rev. Lett.,* 84, (9 Feb. 2000), 1890–1893.

[238] VC Parro and L Valdo. Sleep-wake detection using recurrence quantification analysis. *Chaos: An Interdisciplinary Journal of Nonlinear Science,* 28 (8), (2018), 085706.

[239] Louis M. Pecora, Thomas L. Carroll, and James F. Heagy. Statistics for mathematical properties of maps between time series embeddings. *Phys. Rev. E,* 52, (4 Oct. 1995), 3420–3439.

[240] Louis M. Pecora, Linda Moniz, Jonathan Nichols, and Thomas L. Carroll. A unified approach to attractor reconstruction. *Chaos: An Interdisciplinary Journal of Nonlinear Science,* 17 (1), (2007), 013110.

[241] Alessio Perinelli and Leonardo Ricci. Identification of suitable embedding dimensions and lags for time series generated by chaotic, finite-dimensional systems. *Phys. Rev. E,* 98, (5 Nov. 2018), 052226.

[242] Gert Pfurtscheller, Clemens Brunner, Alois Schlögl, and FH Lopes Da Silva. Mu rhythm (de) synchronization and EEG single-trial classification of different motor imagery tasks. *NeuroImage,* 31 (1), (2006), 153–159.

[243] Gert Pfurtscheller and FH Lopes Da Silva. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical neurophysiology,* 110 (11), (1999), 1842–1857.

[244] Arkady Pikovsky, Michael Rosenblum, and Jürgen Kurths. *Synchronization: A Universal Concept in Nonlinear Sciences.* Cambridge Nonlinear Science Series. Cambridge University Press, 2001.

[245] Elena Pitsik, Nikita Frolov, K. Hauke Kraemer, Vadim Grubov, Vladimir Maksimenko, Jürgen Kurths, and Alexander Hramov. Motor execution reduces EEG signals complexity: Recurrence quantification analysis study. *Chaos: An Interdisciplinary Journal of Nonlinear Science,* 30 (2), (2020), 023111.

[246] H. Poincaré. Sur la probleme des trois corps et les équations de la dynamique. *Acta Mathematica,* 13, (1890), 1–271.

[247] Richard Potts. Evolution and Climate Variability. *Science,* 273 (5277), (1996), 922–923.

[248] Richard Potts. Hominin evolution in settings of strong environmental variability. *Quaternary Science Reviews,* 73, (2013), 1–13.

[249] T. L. Prado, G. Corso, G. Z. dos Santos Lima, R. C. Budzinski, B. R. R. Boaretto, F. A. S. Ferrari, E. E. N. Macau, and S. R. Lopes. Maximum entropy principle in recurrence plot analysis on stochastic and chaotic systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science,* 30 (4), (2020), 043123.

[250]   M Ragwitz and H Kantz. Detecting non-linear structure and predicting turbulent gusts in surface wind velocities. *Europhysics Letters (EPL)*, 51 (6), (Sept. 2000), 595–601.

[251]   Mario Ragwitz and Holger Kantz. Markov models from data by simple nonlinear time series predictors in delay embedding spaces. *Phys. Rev. E*, 65, (5 Apr. 2002), 056201.

[252]   Antônio M. de T. Ramos, Yong Zou, Gilvan Sampaio de Oliveira, Jürgen Kurths, and Elbert E. N. Macau. Unveiling non-stationary coupling between Amazon and ocean during recent extreme events. *Climate Dynamics*, 50 (3), (2018), 767–776.

[253]   Herbert Ramoser, Johannes Muller-Gerking, and Gert Pfurtscheller. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE transactions on rehabilitation engineering*, 8 (4), (2000), 441–446.

[254]   Ana Christina Ravelo, Dyke H. Andreasen, Mitchell Lyle, Annette Olivarez Lyle, and Michael W. Wara. Regional climate shifts caused by gradual global cooling in the Pliocene epoch. *Nature*, 429 (6989), (2004), 263–267.

[255]   Jürgen Richter, Thomas Hauck, Ralf Vogelsang, Thomas Widlok, Jean-Marie Le Tensorer, and Peter Schmid. "Contextual areas" of early Homo sapiens and their significance for human dispersal from Africa into Eurasia between 200 ka and 70 ka. *Quaternary International*, 274, (2012). Temporal and spatial corridors of Homo sapiens sapiens population dynamics during the Late Pleistocene and Early Holocene, 5–24.

[256]   Thorsten Rings, Mahmood Mazarei, Amin Akhshi, Christian Geier, M Reza Rahimi Tabar, and Klaus Lehnertz. Traceability and dynamical resistance of precursor of extreme events. *Scientific reports*, 9 (1), (2019), 1744.

[257]   H.M. Roberts et al. Using multiple chronometers to establish a long, directly dated lacustrine record: constraining >600,000 years of environmental change at Chew Bahir, Ethiopia. *Quaternary Science Reviews, submitted.*, (2021).

[258]   X. Rodó and M.-À. Rodríguez-Arias. A new method to detect transitory signatures and local time/space variability structures in the climate system: the scale-dependent correlation analysis. *Climate Dynamics*, 27 (5), (2006), 441–458.

[259]   Michael G. Rosenblum, Arkady S. Pikovsky, and Jürgen Kurths. From Phase to Lag Synchronization in Coupled Chaotic Oscillators. *Phys. Rev. Lett.*, 78, (22 June 1997), 4193–4196.

[260]   Michael T. Rosenstein, James J. Collins, and Carlo J. [De Luca]. Reconstruction expansion as a geometry-based framework for choosing proper delay times. *Physica D: Nonlinear Phenomena*, 73 (1), (1994), 82–98.

[261]   O.E. Rössler. An equation for continuous chaos. *Physics Letters A*, 57 (5), (1976), 397–398.

[262]   Nikolai F. Rulkov, Mikhail M. Sushchik, Lev S. Tsimring, and Henry D. I. Abarbanel. Generalized synchronization of chaos in directionally coupled chaotic systems. *Phys. Rev. E*, 51, (2 Feb. 1995), 980–994.

[263]   Jakob Runge, Jobst Heitzig, Norbert Marwan, and Jürgen Kurths. Quantifying causal coupling strength: A lag-specific measure for multivariate time series related to transfer entropy. *Phys. Rev. E*, 86, (6 Dec. 2012), 061121.

[264]   Jakob Runge, Vladimir Petoukhov, and Jürgen Kurths. Quantifying the Strength and Delay of Climatic Interactions: The Ambiguities of Cross Correlation and a Novel Measure Based on Graphical Models. *Journal of Climate*, 27 (2), (15 Jan. 2014), 720–739.

Bibliography

[265] N. H. Saji, B. N. Goswami, P. N. Vinayachandran, and T. Yamagata. A dipole mode in the tropical Indian Ocean. *Nature*, 401 (6751), (1999), 360–363.

[266] M. S. Santos, J. D. Szezech, A. M. Batista, I. L. Caldas, R. L. Viana, and S. R. Lopes. Recurrence quantification analysis of chimera states. *Physics Letters A*, 379 (37), (2015), 2188–2192.

[267] Gustavo Zampier dos Santos Lima, Sergio Roberto Lopes, Thiago Lima Prado, Bruno Lobao-Soares, George C do Nascimento, John Fontenele-Araujo, and Gilberto Corso. Predictability of arousal in mouse slow wave sleep by accelerometer data. *PloS one*, 12 (5), (2017), e0176761.

[268] Tim Sauer, James A. Yorke, and Martin Casdagli. Embedology. *Journal of Statistical Physics*, 65 (3), (1991), 579–616.

[269] Eleanor M. L. Scerri et al. Did Our Species Evolve in Subdivided Populations across Africa, and Why Does It Matter? *Trends in Ecology & Evolution*, 33 (8), (), 582–594.

[270] Karin Schiecke, Britta Pester, Martha Feucht, Lutz Leistritz, and Herbert Witte. Convergent Cross Mapping: Basic concept, influence of estimation parameters and practical application. In: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2015, pp. 7418–7421.

[271] Detlev Schild. An Efficient Method for the Fourier Transform of a Neuronal Spike Train. *International Journal of Neuroscience*, 17 (3), (1982), 179–182.

[272] S. Schinkel, O. Dimigen, and N. Marwan. Selection of recurrence threshold for signal detection. *European Physical Journal – Special Topics*, 164 (1), (2008), 45–53.

[273] Jan-Mathijs Schoffelen and Joachim Gross. Source connectivity analysis with MEG and EEG. *Human brain mapping*, 30 (6), (2009), 1857–1865.

[274] Thomas Schreiber. Extremely simple nonlinear noise-reduction method. *Phys. Rev. E*, 47, (4 Apr. 1993), 2401–2404.

[275] Thomas Schreiber and Andreas Schmitz. Improved Surrogate Data for Nonlinearity Tests. *Phys. Rev. Lett.*, 77, (4 July 1996), 635–638.

[276] Thomas Schreiber and Andreas Schmitz. Surrogate time series. *Physica D: Nonlinear Phenomena*, 142 (3), (2000), 346–382.

[277] A. Schultz, Y. Zou, N. Marwan, and M. T. Turvey. Local Minima-based Recurrence Plots for Continuous Dynamical Systems. *International Journal of Bifurcation and Chaos*, 21 (4), (2011), 1065–1075.

[278] Zewdu T. Segele, Peter J. Lamb, and Lance M. Leslie. Seasonal-to-Interannual Variability of Ethiopia/Horn of Africa Monsoon. Part I: Associations of Wavelet-Filtered Large-Scale Atmospheric Circulation and Global Sea Surface Temperature. *Journal of Climate*, 22 (12), (15 Jun. 2009), 3396–3421.

[279] Yilma Seleshi and Ulrich Zanke. Recent changes in rainfall and rainy days in Ethiopia. *International Journal of Climatology*, 24 (8), (2004), 973–983.

[280] Jeremy D. Shakun and Anders E. Carlson. A global perspective on Last Glacial Maximum to Holocene climate change. *Quaternary Science Reviews*, 29 (15), (2010). Special Theme: Arctic Palaeoclimate Synthesis (PP. 1674-1790), 1801–1816.

[281] L. F. Shampine and M. W. Reichelt. The MATLAB ODE Suite. *SIAM Journal on Scientific Computing*, 18, (1997), 1–22.

[282] Timothy M. Shanahan, Nicholas P. McKay, Konrad A. Hughen, Jonathan T. Overpeck, Bette Otto-Bliesner, Clifford W. Heil, John King, Christopher A. Scholz, and John Peck. The time-transgressive termination of the African Humid Period. *Nature Geoscience*, 8 (2), (2015), 140–144.

[283] David Silver et al. Mastering the game of Go with deep neural networks and tree search. *Nature*, 529 (7587), (2016), 484–489.

[284] Michael Small and C.K. Tse. Optimal embedding parameters: a modelling paradigm. *Physica D: Nonlinear Phenomena*, 194 (3), (2004), 283–296.

[285] Mary Ellen Stoykov and Sangeetha Madhavan. Motor priming in neurorehabilitation. *Journal of neurologic physical therapy: JNPT*, 39 (1), (2015), 33.

[286] Chris Stringer and Julia Galway-Witham. On the origin of our species. *Nature*, 546 (7657), (2017), 212–214.

[287] N. P. Subramaniyam and J. Hyttinen. Characterization of dynamical systems under noise using recurrence networks: Application to simulated and EEG data. *Physics Letters A*, 378 (46), (2014), 3464–3474.

[288] George Sugihara, Robert May, Hao Ye, Chih-hao Hsieh, Ethan Deyle, Michael Fogarty, and Stephan Munch. Detecting Causality in Complex Ecosystems. *Science*, 338 (6106), (2012), 496–500.

[289] George Sugihara and Robert M. May. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature*, 344 (6268), (1990), 734–741.

[290] RI Sujith and Vishnu R Unni. Complex system approach to investigate and mitigate thermoacoustic instability in turbulent combustors. *Physics of Fluids*, 32 (6), (2020), 061401.

[291] Floris Takens. Detecting strange attractors in turbulence. In: *Dynamical Systems and Turbulence, Warwick 1980*. Ed. by David Rand and Lai-Sang Young. Berlin, Heidelberg: Springer Berlin Heidelberg, 1981, pp. 366–381.

[292] James Theiler. Spurious dimension from correlation algorithms applied to limited time-series data. *Phys. Rev. A*, 34, (3 Sept. 1986), 2427–2432.

[293] M. Thiel, M. C. Romano, and J. Kurths. Analytical Description of Recurrence Plots of white noise and chaotic processes. *Izvestija vyssich ucebnych zavedenij/ Prikladnaja nelinejnaja dinamika – Applied Nonlinear Dynamics*, 11 (3), (2003), 20–30.

[294] M. Thiel, M. C. Romano, and J. Kurths. Spurious Structures in Recurrence Plots Induced by Embedding. *Nonlinear Dynamics*, 44 (1–4), (2006), 299–305.

[295] M. Thiel, M. C. Romano, J. Kurths, R. Meucci, E. Allaria, and F. T. Arecchi. Influence of observational noise on the recurrence quantification analysis. *Physica D*, 171 (3), (2002), 138–152.

[296] M. Thiel, M. C. Romano, J. Kurths, M. Rolfs, and R. Kliegl. Twin Surrogates to Test for Complex Synchronisation. *Europhysics Letters*, 75 (4), (2006), 535–541.

[297] M. Thiel, M. C. Romano, P. L. Read, and J. Kurths. Estimation of dynamical invariants without embedding by recurrence plots. *Chaos*, 14 (2), (2004), 234–243.

Bibliography

[298] Robert Tibshirani. Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58 (1), (1996), 267–288.

[299] Jessica E. Tierney and Peter B. deMenocal. Abrupt Shifts in Horn of Africa Hydroclimate Since the Last Glacial Maximum. *Science*, 342 (6160), (2013), 843–846.

[300] Jessica E. Tierney, Francesco S. R. Pausata, and Peter B. deMenocal. Rainfall regimes of the Green Sahara. *Science Advances*, 3 (1), (2017).

[301] Axel Timmermann and Tobias Friedrich. Late Pleistocene climate drivers of early human migration. *Nature*, 538 (7623), (2016), 92–95.

[302] J Tony, EA Gopalakrishnan, E Sreelekha, and RI Sujith. Detecting deterministic nature of pressure measurements from a turbulent combustor. *Physical Review E*, 92 (6), (2015), 062902.

[303] Martin H. Trauth. TURBO2: A MATLAB simulation to study the effects of bioturbation on paleoceanographic time series. *Computers & Geosciences*, 61, (2013), 1–10.

[304] Martin H. Trauth, Asfawossen Asrat, Walter Duesing, Verena Foerster, K. Hauke Kraemer, Norbert Marwan, Mark A. Maslin, and Frank Schaebitz. Classifying past climate change in the Chew Bahir basin, southern Ethiopia, using recurrence quantification analysis. *Climate Dynamics*, 53 (5), (2019), 2557–2572.

[305] Martin H. Trauth, Andreas G.N. Bergner, Verena Foerster, Annett Junginger, Mark A. Maslin, and Frank Schaebitz. Episodes of environmental stability versus instability in Late Cenozoic lake records of Eastern Africa. *Journal of Human Evolution*, 87, (2015). Environmental Variability and Hominin Dispersal, 21–31.

[306] Martin H. Trauth, Verena Foerster, Annett Junginger, Asfawossen Asrat, Henry F. Lamb, and Frank Schaebitz. Abrupt or gradual? Change point analysis of the late Pleistocene–Holocene climate record from Chew Bahir, southern Ethiopia. *Quaternary Research*, 90 (2), (2018), 321–330.

[307] Martin H. Trauth, Juan C. Larrasoaña, and Manfred Mudelsee. Trends, rhythms and events in Plio-Pleistocene African climate. *Quaternary Science Reviews*, 28 (5), (2009), 399–411.

[308] Martin H. Trauth, Mark A. Maslin, Alan L. Deino, Annett Junginger, Moses Lesoloyia, Eric O. Odada, Daniel O. Olago, Lydia A. Olaka, Manfred R. Strecker, and Ralph Tiedemann. Human evolution in a variable environment: the amplifier lakes of Eastern Africa. *Quaternary Science Reviews*, 29 (23), (2010), 2981–2988.

[309] Martin H. Trauth, Mark A. Maslin, Alan L. Deino, Manfred R. Strecker, Andreas G.N. Bergner, and Miriam Dühnforth. High- and low-latitude forcing of Plio-Pleistocene East African climate and human evolution. *Journal of Human Evolution*, 53 (5), (2007). African Paleoclimate and Human Evolution, 475–486.

[310] Martin H. Trauth et al. Recurring types of variability and transitions in the ~620 kyr record of climate change from the Chew Bahir basin, southern Ethiopia. *Quaternary Science Reviews*, 266, (2021), 106777.

[311] L. L. Trulla, A. Giuliani, J. P. Zbilut, and C. L. Webber, Jr. Recurrence quantification analysis of the logistic equation with transients. *Physics Letters A*, 223 (4), (1996), 255–260.

[312]   Anastasios A. Tsonis. Reconstructing dynamics from observables: The issue of the delay parameter revisited. *International Journal of Bifurcation and Chaos*, 17 (12), (2007), 4229–4243.

[313]   Anastasios A. Tsonis. *Advances in Nonlinear Geosciences*. Springer, New York, 2018.

[314]   Donald L. Turcotte. *Fractals and Chaos in Geology and Geophysics*. 2nd ed. Cambridge University Press, 1997.

[315]   L. C. Uzal, G. L. Grinblat, and P. F. Verdes. Optimal reconstruction of dynamical systems: A noise amplification approach. *Phys. Rev. E*, 84, (1 July 2011), 016223.

[316]   Ellen Viste and Asgeir Sorteberg. The effect of moisture transport variability on Ethiopian summer precipitation. *International Journal of Climatology*, 33 (15), (2013), 3106–3123.

[317]   Ioannis Vlachos and Dimitris Kugiumtzis. Nonuniform state-space reconstruction and coupling detection. *Phys. Rev. E*, 82, (1 July 2010), 016207.

[318]   R. Vogelsang and B. Keding. Climate, culture, and change: From hunters to herders in northeastern and southwestern Africa. In: *Comparative Archaeology and Paleoclimatology: Socio-cultural responses to a changing world*. Archaeopress, 2013, pp. 43–62.

[319]   Ralf Vogelsang, Olaf Bubenzer, Martin Kehl, Svenja Meyer, Jürgen Richter, and Bahru Zinaye. When Hominins Conquered Highlands—an Acheulean Site at 3000 m a.s.l. on Mount Dendi/Ethiopia. *Journal of Paleolithic Archaeology*, 1 (4), (2018), 302–313.

[320]   E Vrba. The fossil record of African antelopes (Mammalia, bovidae) in relation to human evolution and paleoclimate. *Paleoclimate and evolution, with emphasis on human origins*, (1996), 385–424.

[321]   C. L. Webber, Jr. and N. Marwan. *Recurrence Quantification Analysis – Theory and Best Practices*. Cham: Springer, 2015, p. 421.

[322]   C. L. Webber, Jr., N. Marwan, A. Facchini, and A. Giuliani. Simpler methods do it better: Success of Recurrence Quantification Analysis as a general purpose data analysis tool. *Physics Letters A*, 373, (2009), 3753–3756.

[323]   C. L. Webber, Jr. and J. P. Zbilut. Dynamical assessment of physiological systems and states using recurrence plot strategies. *Journal of Applied Physiology*, 76 (2), (1994), 965–973.

[324]   C. L. Webber, Jr. and J. P. Zbilut. Recurrence quantification analysis of nonlinear dynamical systems. In: *Tutorials in Contemporary Nonlinear Methods for the Behavioral Sciences Web Book*. Ed. by M. A. Riley and G. C. Van Orden. National Science Foundation (U.S.), 2005, pp. 26–94.

[325]   D. Wendi, N. Marwan, and B. Merz. In search of determinism-sensitive region to avoid artefacts in recurrence plots. *International Journal of Bifurcation and Chaos*, 28 (1), (2018), 1850007.

[326]   Dadiyorto Wendi and Norbert Marwan. Extended recurrence plot and quantification for noisy continuous dynamical systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28 (8), (2018), 085722.

[327]   Pål Wessel and Walter H. F. Smith. A global, self-consistent, hierarchical, high-resolution shoreline database. *Journal of Geophysical Research: Solid Earth*, 101 (B4), (1996), 8741–8743.

Bibliography

[328]   Thomas Westerhold et al. An astronomically dated record of Earth's climate and its predictability over the last 66 million years. *Science*, 369 (6509), (2020), 1383–1387.

[329]   Hassler Whitney. Differentiable Manifolds. *Annals of Mathematics*, 37 (3), (1936), 645–680.

[330]   Mahesh Wickramasinghe and István Z. Kiss. Effect of temperature on precision of chaotic oscillations in nickel electrodissolution. *Chaos*, 20 (2), (June 2010), 023125.

[331]   M. Wiedermann, J. F. Donges, J. Kurths, and R. V. Donner. Mapping and discrimination of networks in the complexity-entropy plane. *Physical Review E*, 96, (2017), 042304.

[332]   Jonathan R Wolpaw, Niels Birbaumer, Dennis J McFarland, Gert Pfurtscheller, and Theresa M Vaughan. Brain–computer interfaces for communication and control. *Clinical neurophysiology*, 113 (6), (2002), 767–791.

[333]   Hao Ye, Ethan R. Deyle, Luis J. Gilarranz, and George Sugihara. Distinguishing time-delayed causal interactions using convergent cross mapping. *Scientific Reports*, 5 (1), (2015), 14750.

[334]   Han Yuan and Bin He. Brain–computer interfaces using sensorimotor rhythms: current state and future perspectives. *IEEE Transactions on Biomedical Engineering*, 61 (5), (2014), 1425–1435.

[335]   J. P. Zbilut, A. Giuliani, and C. L. Webber, Jr. Recurrence quantification analysis and principal components in the detection of short complex signals. *Physics Letters A*, 237 (3), (1998), 131–135.

[336]   J. P. Zbilut and C. L. Webber, Jr. Embeddings and delays as derived from quantification of recurrence plots. *Physics Letters A*, 171 (3–4), (1992), 199–203.

[337]   J. P. Zbilut, J.-M. Zaldívar-Comenges, and F. Strozzi. Recurrence quantification based Liapunov exponents for monitoring divergence in experimental data. *Physics Letters A*, 297 (3–4), (2002), 173–181.

[338]   Joseph P. Zbilut and Norbert Marwan. The Wiener–Khinchin theorem and recurrence quantification. *Physics Letters A*, 372 (44), (2008), 6622–6626.

[339]   A. Zimek, E. Schubert, and H. P. Kriegel. A survey on unsupervised outlier detection in high-dimensional numerical data. *Statistical Analysis and Data Mining*, 5 (5), (2012), 363–387.

[340]   Y. Zou, R. V. Donner, N. Marwan, J. F. Donges, and J. Kurths. Complex network approaches to nonlinear time series analysis. *Physics Reports*, 787, (2019), 1–97.

Selbständigkeitserklärung

Über die Bedeutung einer zur Vorlage bei Gericht bestimmten eidesstattlichen Versicherung und über die Strafbarkeit einer falschen Versicherung belehrt, erkläre ich, Kai Hauke Krämer, geb. am 22.11.1985 in Göttingen, folgendes an Eides statt:

Ich erkläre, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung derangegebenen Literatur und Hilfsmittel angefertigt habe. Ich habe mich nicht anderwärts um einem Doktorgrad im Promotionsfach Physik beworben und besitze keinen Doktorgrad im Promotionsfach Physik.

Die Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät, veröffentlicht in den Amtlichen Bekanntmachungen der Universität Potsdam Nr. 19/2013 am 19.12.2013 habe ich zur Kenntnis genommen.

Potsdam, den 11. Februar 2022, Kai Hauke Krämer