Springer Proceedings in Physics 180

Charles L. Webber, Jr.
Cornel Ioana
Norbert Marwan *Editors*

Recurrence Plots and Their Quantifications: Expanding Horizons

Proceedings of the 6th International Symposium on Recurrence Plots, Grenoble, France, 17–19 June 2015

Springer

Springer Proceedings in Physics

Volume 180

More information about this series at http://www.springer.com/series/361

Charles L. Webber, Jr. · Cornel Ioana
Norbert Marwan
Editors

Recurrence Plots and Their Quantifications: Expanding Horizons

Proceedings of the 6th International Symposium on Recurrence Plots, Grenoble, France, 17–19 June 2015

Springer

Editors
Charles L. Webber, Jr.
Department of Cell and Molecular Physiology
Loyola University Chicago
Maywood, IL
USA

Cornel Ioana
GIPSA-Lab
University Grenoble Alpes
Grenoble
France

Norbert Marwan
Potsdam Institute for Climate Impact Research
Potsdam
Germany

ISSN 0930-8989 ISSN 1867-4941 (electronic)
Springer Proceedings in Physics
ISBN 978-3-319-29921-1 ISBN 978-3-319-29922-8 (eBook)
DOI 10.1007/978-3-319-29922-8

Library of Congress Control Number: 2016936417

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

This book is dedicated to the next generation of scientists who will push the expanding horizon of nonlinear dynamics to worlds unknown.

Foreword

Zeitgeist is an evocative German word which can be translated as "Spirit of the Age" or "Spirit of the Time." The basic idea is that, besides the obvious differences discriminating various fields of inquiry (science, art, religion, politics, social organization, etc.), there are some shared and recognizable invariant "structures" shaping the way of thinking for every age (historical epoch). Until the eighteenth century the privileged zeitgeist of that time resided within the figurative arts. For an example, one needs only to recall the rationalist anthropocentrism of the Renaissance marked by the development of perspective laws in painting. Then one can also consider the shift from realism to include human emotions, personal drama and introspection as emoted by the Italian painter Caravaggio at the dawn of seventeenth century. Many shifts later, however, the modern zeitgeist is to be found in the natural sciences.

Besides some early visionary ideas from thinkers like Weaver [1], the turning point toward the contemporary zeitgeist occurred within the last two decades of the previous century. One of the most clear and synthetic papers overtly initiating the new era was the work of Laughlin et al. [2] who spoke of "The Middle Way." These authors stated that the "really interesting things" are no longer located at the very bottom of the scale (sub-atomic particles, genes, molecules, neurons, etc.), but on the mesoscale. It is within this "middle way" space where the highest number of correlations between different organization layers can be found. This meso-perspective allows the organization of the studied system to be understood in terms of how its various parts are interconnected. This "Third Way" is proving to be much more fruitful than merely acquiring more intricate details at the deepest level (bottom-up strategies) or accumulating mere description of general laws governing the studied phenomena with no utility (top-down strategies).

Laughlin at al. [2] sketched some "mesoscopic principles" that were mostly dependent on the organization of the system, but largely independent of the material instantiation of the system itself. This critical idea was corroborated and expanded upon by others like Mikulecky [3] and Giuliani et al. [4]. Giuliani et al. [4] even called this new perspective the "Middle-Out" strategy in which the emphasis is

shifted away from individual parts or system components to the relationships within and between systems. In this paradigm-shift "relation" is the new operative starting point from where investigators can try to understand the world. The World Wide Web is one material counterpart of this operative because it uniquely focuses on the networked relations among multiplicative fields of knowledge. In other words, the WWW embodies the present zeitgeist which reaches beyond the realm of science with its long "sticky" fingers!

With this new emphasis on relations over components, the "perfect science icon" is no longer an equation-based (think of Einstein's famous equation: $e = mc^2$), but a graphically centered. Case in point, the recurrence plot [5] is actually an adjacency matrix having as rows and columns the elements of the system (nodes) and at the intersection a label indicating the presence (or absence) of a direct relations (edges) between the corresponding elements. It does not matter what the elements of the system are exactly, but merely their relationships. It is for this very reason that recurrence plots are so ubiquitous across different fields of inquiry. Recurrence quantifications help evolve general portraits of the system in different states by means of deriving descriptors from the Recurrence Plot. Indeed, the Recurrence Plot graphs the "wiring architecture" of the system whose nodes are different (short) epochs and whose edges correspond to the identity (approximated by a very high similarity) between them. These descriptors are completely superimposable to graph invariants and allow one to go from the time-series realm to other systems from any field as long as it can be segmented into constituent parts.

This present book is continuing testimony as to just how far recurrence strategies have developed and matured since they were originally introduced as Recurrence Plots in 1987 by Eckmann et al. [6] and as Recurrence Quantification Analysis in 1992 and 1994 by Zbilut and Webber [7, 8]. These early investigators had no idea of how far and wide-reaching such simple and fundamental rules of dynamical systems would spread and invade ever-new fields of science (and non-science). Indeed, the recurrence horizon is ever-expanding and one wonders what the next decades will bring? One thing is for sure, recurrence analyses have the unparalleled power of integrating systems and dissuading any disruptive fragmentation of discovery between fields that is a serious menace to the furtherance of true knowledge and deep understanding.

Alessandro Giuliani
Istituto Superiore di Sanità
Rome, Italy
e-mail: alessandro.giuliani@iss.it

References

1. W. Weaver, Science and complexity. Am. Sci. **36**, 536–539 (1948)
2. R.B. Laughlin, D. Pines, J. Schmalian, B.P. Stojković, P. Wolynes, The middle way. Proc. Natl. Acad. Sci. **97**, 32–37 (2000)
3. D.C. Mikulecky, Network thermodynamics and complexity: a transition to relational systems theory. Comp. Chem. **25**, 369–391 (2001)
4. A. Giuliani, S. Filippi, M. Bertolaso, Why network approach can promote a new way of thinking in biology. Front. Genet. **5**, 83 (2014). doi:10.3389/fgene.2014.0008
5. N. Marwan, J.F. Donges, Y. Zou, R.V. Donner, J. Kurths, Complex network approach for recurrence analysis of time series. Phys. Lett. A **373**, 4246–4254 (2009)
6. J.-P. Eckmann, S.O. Kamphorst, D. Ruelle, Recurrence plots of dynamical systems. Europhys. Lett. **4**, 973–977 (1987)
7. J.P. Zbilut, C.L. Webber, Jr, Embeddings and delays as derived from quantification of recurrence plots. Phys. Lett. A **171**, 199–203 (2002)
8. C.L. Webber, Jr., J.P. Zbilut, Dynamical assessment of physiological systems and states using recurrence plot strategies. J. Appl. Physiol. **76**, 965–973 (1994)

Preface

The title of this book—"Recurrence Plots and Their Quantifications: Expanding Horizons"—suggests humble beginnings, leaps in understanding, and continued spreading of recurrence ideas into multiple fields of inquiry. The world and its uncountable dynamical systems seen to function on fundamental principles of recurrence in linear and as well as nonlinear domains, operating on different temporal and spatial scales and over various dimensions. When one thinks about it, the sun recurs every day, but at slightly different positions along the horizon. The beating wings of birds and insects recur, producing up and down motions for propulsion. Movements in humans recur, from the scratching of an itch, to the many gaits of locomotion and stages of sleep/wake patterns. Words in a book recur, making possible the writing very long texts larger than the vocabulary size selected. And DNA codes recur, giving instructions for the production of specific proteins that make life itself possible. Indeed, life itself recurs through reproduction!

Yes, we live in a recurrent world, nay universe, which explains why symposia after recurrence symposia, new and useful research applications keep popping up for recurrence plots (RPs) and recurrence quantifications (RQs) in virgin systems as it were. There is literally no end in sight as the collection of papers herein bears witness. Specifically, this present volume represents 19 selected papers presented, discussed, and debated over at the Sixth International Recurrence Plots Symposium held in Grenoble France (June 17–19, 2015). More than 50 participants from around the world interacted. Some scientists were new to the meeting, but others "recurred" having attended one or more of the five previous symposia (Potsdam 2005, Siena 2007, Montreal 2009, Hong Kong 2011, Chicago 2013).

The structure of this book is divided into two parts: methodological and practical. Part one addresses theoretical topics with examples such as recurrences in large data sets, transient and non-stationary signals, complexity testing, approximate recurrences and the new splayed recurrence analysis (Chaps. 1–7). Part two focuses on specific dynamical systems that employ recurrence strategies and are grouped into four categories consisting of three chapters each. The first group readdresses now familiar recurrent systems in human physiology including heart

and brain rhythms (Chaps. 8–10). Second coordinated systems are studied including those related to human interpersonal behaviors, social-motor coordination, and even monetary systems which are heavily influence by human activity (Chaps. 11–13). Papers in the third group of center on hydraulics and hydrology including the complex dynamics of underwater acoustics, water temperature fluctuations, and eddy current fluxes (Chaps. 14–16). Finally, group four introduces uses of recurrence analyses in combustion dynamics and flashbacks, turbulence in plasmas, and ultrasonic testing in polymers (Chaps. 17–19). Indeed, creative recurrence ideas as strategies are expanding the horizon into numerous scientific disciplines. That is precisely why "Expanding Horizons" was chosen as part of this book's title.

The editors and authors welcome you to the fascinating, even beautiful world of recurrences. By reading this book may you become excited about how recurrence analysis might be useful to your own field of study no matter what interests you scientifically. The invitation is extended to all to feel free to contact any of the authors regarding their contributions.

Chicago — Charles L. Webber, Jr.
Grenoble — Cornel Ioana
Potsdam — Norbert Marwan

Acknowledgments

The editors are indebted to the Springer Publisher who have faithfully followed and supported this relatively new field of recurrence analysis. This is the third book on the topic and Springer has proven to be as professional to work with as they are editor-friendly! Sincere thanks are also extended to all the support staff at Grenoble-INP, Gipsa-lab who worked out the practical details involved in hosting an international symposium in such a beautiful venue at the foot of the European Alps. Finally, to the many expert reviewers who provided critical feedback to all authors contributing to this volume, we thank you for your fruitful labors which have helped shape, not only this book, but the furtherance of the field of nonlinear dynamics.

Contents

Part II Practical/Utilitarian Recurrences

Contributors

Humberto Arce Laboratorio de Biofísica de Sistemas Excitables, Facultad de Ciencias, Universidad Nacional Autónoma de México, Mexico, Mexico

Cindy Bernard Gipsa-lab, Universités Grenoble Alpes, Grenoble-INP, Grenoble, France

Olivier Le Bot GIPSA-Lab, Univ. Grenoble Alpes, Grenoble, France; GIPSA-Lab, CNRS, Grenoble, France

Carsten Brandt Centre for Industrial Mathematics, Bremen, Germany

I.L. Caldas Institute of Physics, University of São Paulo, São Paulo, Brazil

Ion Candel GIPSA-lab, Grenoble Institute of Technology, Grenoble, France

A. Carrión Instituto de Telecomunicación y Aplicaciones Multimedia (iTEAM), Universitat Politècnica de València, Valencia, Spain

Ralf F.A. Cox Department of Psychology, Faculty of Behavioural and Social Sciences, University of Groningen, Groningen, The Netherlands

Angela Digulescu GIPSA-lab, Grenoble Institute of Technology, Grenoble, France; Military Technical Academy, Bucharest, Romania

Marijn van Dijk Department of Psychology, Faculty of Behavioural and Social Sciences, University of Groningen, Groningen, The Netherlands

Thomas Dionysopoulos ADVENIS Investment Managers, Paris, France; AXIANTA Research, Nicosia, Cyprus

Paula Fitzpatrick Assumption College, Worcester, USA

Milan Flach Max-Planck-Institute for Biogeochemistry, Jena, Germany

Thomas Foken Bayreuth Center of Ecology and Environmental Research (BayCEER), University of Bayreuth, Bayreuth, Germany

Federica Frigerio Department of Neuroscience, IRCCS—Istituto di Ricerche Farmacologiche Mario Negri, Milan, Italy

Ayari Fuentes Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Mexico, Mexico

K.W. Gentle Department of Physics and Institute for Fusion Studies, The University of Texas at Austin, Austin, TX, USA

Cédric Gervaise Chaire Chorus, Foundation of Grenoble INP, Grenoble Cedex 1, France

Alexandre Girard EDF R&D, Chatou, France

Moacir Fernandes Godoy NUTECC-Transdisciplinary Nucleus of Studies on Complexity and Chaos, Sao Jose do Rio Preto Medicine School (FAMERP), Sao Jose do Rio Preto, São Paulo, Brazil

G. Hortensia González Laboratorio de Biofísica de Sistemas Excitables, Facultad de Ciencias, Universidad Nacional Autónoma de México, Mexico, Mexico

Marlenny Guevara Institute of Psychology, Universidad del Valle, Cali, Colombia

Z.O. Guimarães-Filho Institute of Physics, University of São Paulo, São Paulo, Brazil

Michael Hauhs University of Bayreuth, Bayreuth Center of Ecology and Environmental Research (BayCEER), Bayreuth, Germany

Rosangela Akemi Hoshi NUTECC-Transdisciplinary Nucleus of Studies on Complexity and Chaos, Sao Jose do Rio Preto Medicine School (FAMERP), Sao Jose do Rio Preto, São Paulo, Brazil

Cornel Ioana Gipsa-lab, Universités Grenoble Alpes, Grenoble-INP, Grenoble, France

Valentina Iori Department of Neuroscience, IRCCS—Istituto di Ricerche Farmacologiche Mario Negri, Milan, Italy

Lisette de Jonge-Hoekstra Department of Psychology, Faculty of Behavioural and Social Sciences, University of Groningen, Groningen, The Netherlands

Lipika Kabiraj Chair of Fluid Dynamics, Hermann Föttinger Institute, Technische Universität Berlin, Berlin, Germany

Nader Karimi School of Engineering, University of Glasgow, Glasgow, UK

Mariola Kędra Department of Environmental Engineering, Institute of Engineering and Water Management, Cracow University of Technology, Cracow, Poland

Holger Lange Norwegian Institute of Bioeconomy Research, Ås, Norway

Grzegorz Litak Lublin University of Technology, Lublin, Poland

Jérôme I. Mars GIPSA-Lab, Univ. Grenoble Alpes, Grenoble, France

Norbert Marwan Potsdam Institute for Climate Impact Research, Potsdam, Germany

R. Miralles Instituto de Telecomunicación y Aplicaciones Multimedia (iTEAM), Universitat Politècnica de València, Valencia, Spain

Romuald Mosdorf Bialystok University of Technology, Bialystok, Poland

Irina Murgan GIPSA-lab, Grenoble Institute of Technology, Grenoble, France

I.C. Nascimento Institute of Physics, University of São Paulo, São Paulo, Brazil

Holger Nawroth Chair of Fluid Dynamics, Hermann Föttinger Institute, Technische Universität Berlin, Berlin, Germany

C.O. Paschereit Chair of Fluid Dynamics, Hermann Föttinger Institute, Technische Universität Berlin, Berlin, Germany

Carlos Marcelo Pastre UNESP-State of Sao Paulo University, Campus of Presidente Prudente, Physical Therapy Department, Presidente Prudente, Brazil

Tobias Rawald German Research Center for GeoSciences GFZ, Potsdam, Germany; Humboldt-Universität zu Berlin, Berlin, Germany

Michael J. Richardson University of Cincinnati, Cincinnati, USA

Massimo Rizzi Department of Neuroscience, IRCCS—Istituto di Ricerche Farmacologiche Mario Negri, Milan, Italy

Veronica Romero University of Cincinnati, Cincinnati, USA

Andrzej Rysak Lublin University of Technology, Lublin, Poland

Aditya Saurabh Chair of Fluid Dynamics, Hermann Föttinger Institute, Technische Universität Berlin, Berlin, Germany

R.C. Schmidt College of the Holy Cross, Worcester, USA

David Schultz DAI-Lab, Technische Universität Berlin, Berlin, Germany

Alexandru Serbanescu Military Technical Academy, Bucharest, Romania

Mike Sips German Research Center for GeoSciences GFZ, Potsdam, Germany; Humboldt-Universität zu Berlin, Berlin, Germany

Stephan Spiegel DAI-Lab, Technische Universität Berlin, Berlin, Germany

Steffie van der Steen Department of Special Needs Education and Youth Care, Faculty of Behavioural and Social Sciences, University of Groningen, Groningen, The Netherlands

R.I. Sujith Department of Aerospace Engineering, Indian Institute of Technology Madras, Chennai, India

Dennis L. Toufen Instituto Federal de Educação, Ciência e Tecnologia. Campus de Guarulhos, Guarulhos, São Paulo, Brazil

George Tzagkarakis EONOS Investment Technologies, Paris, France

Luiz Carlos Marques Vanderlei UNESP-State of Sao Paulo University, Campus of Presidente Prudente, Physical Therapy Department, Presidente Prudente, Brazil

R.L. Viana Department of Physics, Federal University of Parana, Curitiba, Parana, Brazil

Charles L. Webber, Jr. Health Sciences Division, Department of Cell and Molecular Physiology, Loyola University Chicago, Maywood, IL, USA

Carl Witt Humboldt-Universität zu Berlin, Berlin, Germany

Part I
Methodological/Theoretical Recurrences

Chapter 1
Torwards Visual Analytics for the Exploration of Large Sets of Time Series

Mike Sips, Carl Witt, Tobias Rawald and Norbert Marwan

Abstract In this chapter, we discuss the scientific question whether the clustering of time series based on **RQA** measures leads to an interpretable clustering structure when analyzed by human experts. We are not aware of studies answering this scientific question. Answering it is the crucial first step in the development of a **Visual Analytics** approach that support users to explore large sets of time series.

1.1 Introduction

The advancement of recent technology allows scientists to measure different variables of environmental systems at many positions; e.g., the National Oceanic and Atmospheric Administration (NOAA) [1] measures the surface temperature around the globe. The resulting time series encompass a sequence of data points over long time periods at high sampling rates. To study the temporal behavior of an environmental system, scientists need to detect positions with similar temporal dynamics in large sets of time series.

The number and duration of recurrent states in a system is an important aspect of its temporal behavior. A well-established approach to quantify the number and duration of recurrent states is **Recurrence Quantification Analysis (RQA)** [2]. **RQA**

M. Sips (✉) · T. Rawald
German Research Center for GeoSciences GFZ, Potsdam, Germany
e-mail: sips@gfz-potsdam.de

T. Rawald
e-mail: trawald@gfz-potsdam.de

M. Sips · C. Witt · T. Rawald
Humboldt-Universität zu Berlin, Berlin, Germany
e-mail: wittcarl@googlemail.com

N. Marwan
Potsdam Institute for Climate Impact Research, Potsdam, Germany
e-mail: marwan@pik-potsdam.de

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8_1

calculates quantitative measures that enable scientists to understand the temporal behavior of systems.

In this chapter, we address the scientific question of whether the clustering of time series based on their **RQA** measures produces a clustering structure that is interpretable by human experts. To study whether the grouping of time series based on their **RQA** measures into clusters produces an interpretable clustering structure, we conduct two independent experiments: one experiment with synthetic signals and one experiment with real-world signals. We utilize the **iVAT** method to analyze the clustering structure in these two experiments. We are not aware of similar experiments. Therefore, a positive answer to this question is a critical first step in the development of a **Visual Analytics** approach to support the exploration of large sets of time series. The main contributions of this chapter follow.

- We present an experiment that demonstrate how the clustering of synthetic time series based on their **RQA** results leads to an interpretable clustering structure (Sect. 1.3).
- We show that the clustering structure of the synthetic time series is robust against noise (Sect. 1.3).
- We compute the clustering structure of nine real-world time series and show that the resulting clustering structure is interpretable by human experts (Sect. 1.4).

1.2 Methodology

In this section, we briefly discuss important methods utilized in our study and define important terms of this chapter.

1.2.1 Recurrence Quantification Analysis

Recurrence plots (RPs) and recurrence quantification analysis (RQA) are powerful methods for analyzing recurrences in measured time series [2]. Their application in many fields have proven their potential for various kinds of analyses [3]. A recurrence plot is a two-dimensional representation of a time series when a m-dimensional phase space trajectory recurs to former (or later) states. Recurrence of a state at time i at a different time j is captured within a squared matrix $\mathbf{r}$ [2]:

$$r_{i,j} = \Theta\left(\varepsilon - \left\|\mathbf{x}_i - \mathbf{x}_j\right\|\right), \quad \mathbf{x}_i \in \mathbb{R}^m, \quad i,j = 1 \ldots N. \tag{1.1}$$

Both of its axes represent the time steps. N is the number of considered states x_i (length of phase space trajectory). ε is a threshold distance, $\|\cdot\|$ a norm, and $\Theta(\cdot)$ the Heaviside function. A pair of states that fulfills the threshold condition is assigned with the value 1 (*recurrence point*), whereas a pair that is considered to be dissimilar

is assigned with the value 0. Further details about the reconstruction of phase space vectors from a scalar time series, the recurrence parameters, as well as the typical visual characteristics of RPs can be found in [2].

Small scale structures in the RPs, like diagonal lines, are used to define measures of complexity called **Recurrence Quantification Analysis (RQA)** [2, 4, 5]. As an example, we present the RQA measure *percent determinism (DET)*:

$$DET = \frac{\sum_{l=d_{\min}}^{N} l\, H_{\mathrm{D}}(l)}{\sum_{i,j=1}^{N} r_{i,j}}. \tag{1.2}$$

It is the fraction of recurrence points that form diagonal lines; $H_{\mathrm{D}}(l)$ is the number of diagonal lines of exactly length l and $d_{\min}$ is a minimal length necessary to be a diagonal line. This measure characterizes the deterministic nature of a dynamical system from a heuristic point of view (further discussions can be found in [2, 6]). Further measures quantify average line lengths or the complexity of the line length frequency distributions $H_{\mathrm{D}}(l)$ (diagonal lines) and $H_{\mathrm{V}}(l)$ (vertical lines).

1.2.2 VAT and iVAT

Visual Assessment of Clustering Tendency [7] is a method to depict the cluster tendency of a set of data objects. The input to the VAT method is usually the pairwise dissimilarity matrix of the data points. The basic idea of the **VAT** method is to rearrange the rows and columns of the dissimilarity matrix in such a manner that it depicts the clustering structure of the data objects. **VAT** visualizes the reordered matrix as intensity image utilizing gray levels. Pure white depicts the largest dissimilarity value and pure black depicts zero dissimilarity in the intensity image. Users utilize the **VAT** method to determine whether clusters are present before applying particular clustering algorithms to their data.

To reorder the rows and columns, VAT computes a permutation of the original row and column indices of the dissimilarity matrix. The algorithm to compute this permutation is similar to Prim's algorithm [8]. The basic idea is to interpret the dissimilarity matrix as a weighted graph of the input data points; i.e., the nodes of the graph are the data points that are connected to each other. The dissimilarity value denotes the distance between the nodes. **VAT** starts with the node v that has the largest distance to another node. The row index of v is the first element of the permutation. It then searches for the node w that has the smallest distance to v. The column index of w is the second element of the permutation. Next, VAT searches for the node x that has the smallest distance to either v or w. The column index of x is the next element of the permutation. This algorithm terminates if all nodes has been visited.

This mechanism is responsible for the emergence of black blocks along the diagonal of the reordered matrix. Without this proper reordering of the rows and columns, it is essentially impossible to assess the clustering tendency in the intensity image (we refer to [7] for more detailed discussion).

Improved VAT (iVAT) [9, 10] method is an extension of the **VAT** algorithm, and transforms the dissimilarity matrix using a graph theoretic measure. It utilizes the VAT algorithm on the transformed matrix to improve the rearrangement of rows and columns.

1.2.3 Definitions

The following definitions are valid throughout the chapter.

Clustering Structure Given a set of time series. A clustering structure is a set of distinct groups where (a) each group consists of a subset of time series and (b) each time series is assigned to exactly one group.

Interpretable Clustering Structure We call a clustering structure interpretable by human experts if (a) an expert identifies a cluster as the time series or recurrence plots of a particular system, and (b) the expert considers the group members as similar to each other.

1.3 Experiment with Synthetic Signals

To study whether the grouping of synthetic time series based on their **RQA** measures into clusters produces an interpretable clustering structure, we do not introduce independent variables into the experimental setup. In the next step, we extend the experiment by introducing one independent variable. This independent variable is noise ratio. Note: the clustering structure is always the dependent variable in the experiments.

1.3.1 Experimental Setup

The set of time series is finite and well-known. We utilize nine systems to generate the set of time series and compute the data points for each time series according to the equation and parameters listed in Table 1.2. To generate data points of time series representing the periodic motion system, λ traverses the unit circle three times and each circulation generates at least 16 data points. Figure 1.1 presents representative RPs of the utilized systems.

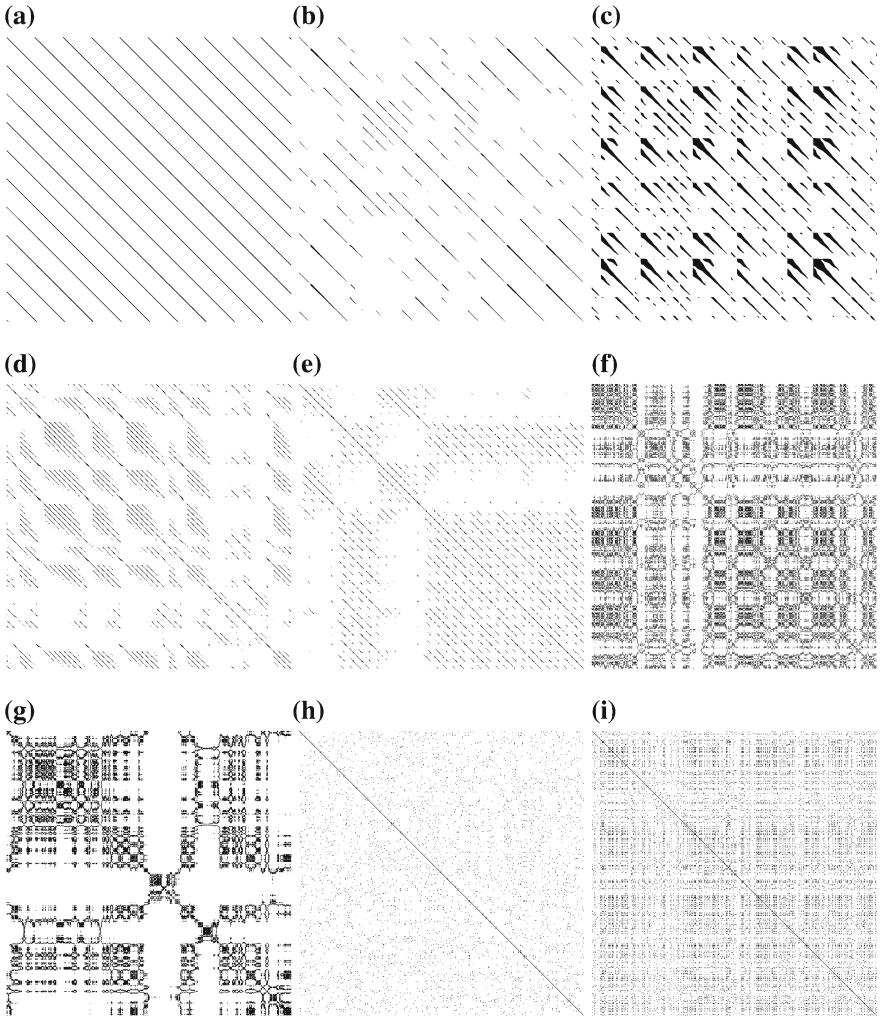

Fig. 1.1 Example RPs of the nine systems utilized in the experiments. **a** OSC. **b** R. **c** RF. **d** L. **e** LO. **f** AR−. **g** AR+. **h** N. **i** NG. Table 1.1 resolves the abbreviations

We generate 100 time series for each system. Hence, the number of time series we use in the experiment is 900. This set represents a broad range of well-known temporal behavior. Each time series has 25,000 data points. To generate the time series for the R, RF, L, LO, we generate $100 \times 25{,}000 + 1000$ data points. We discard the first 1000 data points and group the remaining data points into 100 time series. To compute time series for the other systems, we change the seed of the random number generator for each time series.

Table 1.1 Parameter assignments of the RQA method utilized in the experiments

RQA parameter	Assignment
Embedding dimension	1
Time delay	0
Recurrence threshold	5 % of the trajectory diameter
Minimal black diagonal and vertical line lengths	2
Minimal white diagonal and vertical line lengths	2
Theiler corrector	1

To minimize variations in the clustering structure introduced by the **RQA** method, we keep all **RQA** parameters at predefined values. Table 1.1 summarizes the parameter assignments of this experiment. An adaptive choice of the **RQA** parameters may improve the clustering structure. To study the impact of certain **RQA** parameters on the clustering structure, we extend this experiment in the next step and introduce noise ratio as independent variable.

Note that our experiment contains the two noise systems white noise (N) and gamma noise (NG). The reason for including these two noise systems is to see whether the clustering structure also separates the noise systems in two distinct clusters.

1.3.2 Experimental Procedure

First, we compute 16 **RQA** measures (see Table 1.3) for each time series, and store these measures in a 16-dimensional vector. To minimize variations introduced by the different dynamic ranges of the measures, we normalize each value of the 16-dimensional vectors. Let $\{v_i \in \mathbb{R}^{16} \quad i = 1 \ldots 900\}$ be the set of all 16-dimensional vectors. We decided to normalize the i-th value $v_i[i]$ with $\max_{j=1\ldots900}\{v_j[i]\} - \min_{j=1\ldots900}\{v_j[i]\}$.

We utilize the Euclidean distance to determine the dissimilarity between two vectors. This means that small distances denote low dissimilarities between time series, and large distances denote high dissimilarities between time series. Second, we compute all pair-wise dissimilarity values between the **RQA** vectors and store these values in a 900×900 matrix. This matrix is the input to the **iVAT** method. The **iVAT** method determines an ordering of the 900 time series that depicts the clustering structure of the time series.

To see whether the clustering structure is in line with expert opinion, in addition we conducted an experiment with two experts. In this experiment, the expert determines the dissimilarity between time series by comparing their RPs. To this end, we present two RPs to the expert who determined the dissimilarity between them at a scale between zero and 100; zero denotes a low dissimilarity and 100 denotes

a high dissimilarity. The expert determined the pair-wise dissimilarity between 45 RPs. We randomly selected 36 pairs of RPs from different systems, and nine pairs from the same system. These RPs were randomly selected from the set of 900 time series described above. We decided to ask the expert to make 45 comparisons of RPs because the attention of the human expert is the limiting factor in this experiment. We stored these pair-wise dissimilarities in a 45×45 matrix. This matrix is the input to the **iVAT** method.

Note: the pair-wise dissimilarity values are different in the experiments, and therefore, the arrangement of rows and columns may differ in the resulting intensity images.

1.3.3 Discussion

Figure 1.2a shows the result of the **iVAT** method computed from **RQA** measures. Dark gray colors denote high similarities between time series and light gray colors denote low similarities between time series.

We see nine well-separated black blocks along the diagonal of the **iVAT** matrix in Fig. 1.2a. Each black block contains 100 time series, and the black color indicates that the members of each block are almost indiscernible based on their **RQA** measures. Furthermore, each block contains time series from the same system only. The labels of the blocks are listed in Table 1.2. The cluster OSC that represents the time series of periodic motion, for example, is composed of the 100 time series of periodic motion only. We conclude from the nine well-separated black blocks in Fig. 1.2a that **RQA** measures consider the time series of one system as similar to each other; each

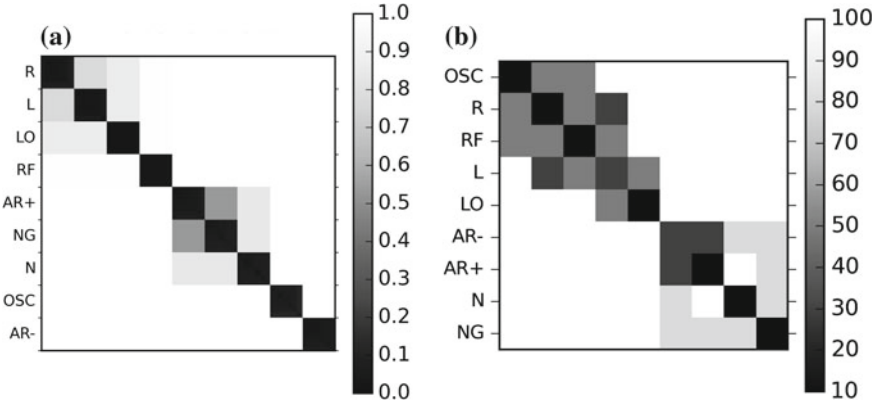

Fig. 1.2 **iVAT** matrix depicting the clustering structure of the experiment. Each *black block* in this matrix contains 100 time series of a particular system. *Note* the different orderings of the systems along the y-axis are due to the **iVAT** method and cannot be changed to maintain the visual cluster tendency. **a** RQA measures. **b** Expert opinion

Table 1.2 The nine systems utilized in the experiments and their initial conditions

System	Abbr.	Equations	Parameterization
Periodic motion	OSC	$x(t) = \cos(\lambda t)$, $y(t) = \sin(\lambda t)$	See text
Rössler system, standard parameters	R	Rössler equations	a = 0.2, b = 0.2, c = 5.7, Δt = 0.085
Rössler system, funnel regime	RF	Rössler equations	a = 0.2925, b = 0.1, c = 8.5, Δt = 0.085
Lorenz system, standard parameters	L	Lorenz equations	$\sigma = 10$, $\beta = 2.66$, $\rho = 28$, Δt = 0.025
Lorenz system, noisy oscillation	LO	Lorenz equations	$\sigma = 10$, $\beta = 2.66$, $\rho = 198$, Δt = 0.025,
Autoregressive, negative correlation	AR−	$x(t) = \phi x(t-1) + \mathcal{N}(0, 0.1)$	$\phi = -0.95$
Autoregressive, positive correlation	AR+	$x(t) = \phi x(t-1) + \mathcal{N}(0, 0.1)$	$\phi = 0.95$
Normal noise	N	$x(t) \sim N(\mu, \sigma^2)$	$\mu = 0, \sigma = 1$
Gamma noise	NG	$x(t) \sim \Gamma(k, \theta)$	$k = 0.5, \theta = 1$

block contains the 100 time series of one system only. Furthermore, **RQA** measures consider time series from different systems as dissimilar to each other.

Figure 1.2b shows the result of the expert assessment. We present the result of one expert here; the results among the two experts are similar to each other. Again, dark gray colors denote high similarities between time series and light gray colors denote low similarities between time series. Note, the dissimilarity values range from 0 to 100. We see seven well-separated black blocks along the diagonal of the **iVAT** matrix in Fig. 1.2b. Furthermore, we observe similar properties of these seven clusters in comparison to the nine clusters produced by **RQA** measures. We identify the seven blocks OSC, R, RF, LO, AR+, N, and NG from Fig. 1.2a also in Fig. 1.2b. Each of these blocks contain 100 time series of one system only.

The **iVAT** result suggests that time series of the L system are in the R block, and members of the AR− system are in the AR+ cluster (see Fig. 1.2b). L and R are chaotic systems, and thus, both belong to the same class of systems. AR+ and AR− are correlated noise systems, and thus, both belong to the same class. We argue that the mixing of L with R and AR+ and AR− does not constitute a problem. The interesting point here is that the expert can clearly see the similarity between these systems in the RPs but the quantitative properties of these examples are much more different. This leads to distinct clusters for each of these systems. Nevertheless, there is a high degree of consensus between the two clustering structures of Fig. 1.2a, b. We conclude from the comparison of Fig. 1.2a, b that the clustering structure of time series based on their **RQA** measures is interpretable.

1.3.4 Potential of Clustering Algorithms

In this section, we discuss the clustering structure that is likely to be detected by clustering algorithms. Figure 1.2a shows that clustering algorithms experience difficulties to derive exactly the nine clusters as from our visual inspection of the **iVAT** matrix. The reason for this are the light gray blocks around some of the black blocks. The gray box around the R, L, and LO clusters indicate that some time series of the R system are more similar to some time series of L and/or LO systems. This is a reasonable result since these systems belong to the same class, the chaotic systems. A similar observation holds for AR+, NG, and N systems. These systems belong to the same class, the stochastic systems.

To estimate the clustering structure, we assume that clustering algorithms cannot separate clusters in Fig. 1.2a that are close to each other. Hence, the time series of these clusters will be merged into a bigger cluster; together with time series from other systems. Based on this assumption, we conclude from Fig. 1.2a that clustering algorithms are likely to derive the following clusters: c1 = (R, L, LO), c2 = (RF), c3 = (AR+, NG, N), c4 = (OSC) and c5 = (AR−).

To see whether this clustering structure is interpretable, we compare our estimated clustering structure with the clustering structure based on expert opinion (see Fig. 1.2b). We see that experts judge time series of the R, and L systems to be similar to each other. We argue that the merging of these two systems into a common cluster does not constitute a problem since this grouping is in line with expert opinion. The LO system needs a careful discussion. Figure 1.2b shows that LO contains almost 100 time series of the LO system. The ordering along the diagonal of the matrix suggest that the LO cluster is close to the R and L clusters. We conclude from this observation that the cluster c1 is interpretable.

An interesting observation is that clustering algorithms are likely to experience difficulties to separate noise systems N, NG and autoregressive systems (positively correlated). According to expert opinion, the systems AR− and AR+ are similar to each other. In contrast, the system AR+ shares some similarity to the systems NG and N according to expert opinion. We argue that this does not constitute a problem since cluster c3 = (AR+, NG, N) is still in line with expert opinion. We conclude from these observations that the clustering structure computed by clustering algorithms are likely to be interpretable.

1.3.5 Effect of Noise

We extend the experiment to study the effect of noise on the clustering structure of the nine systems. We utilize noise ratio as the independent variable because we expect to see that time series represent certain noise signals. In this experiment, the experimental setup is the same as in the experiment reported above:

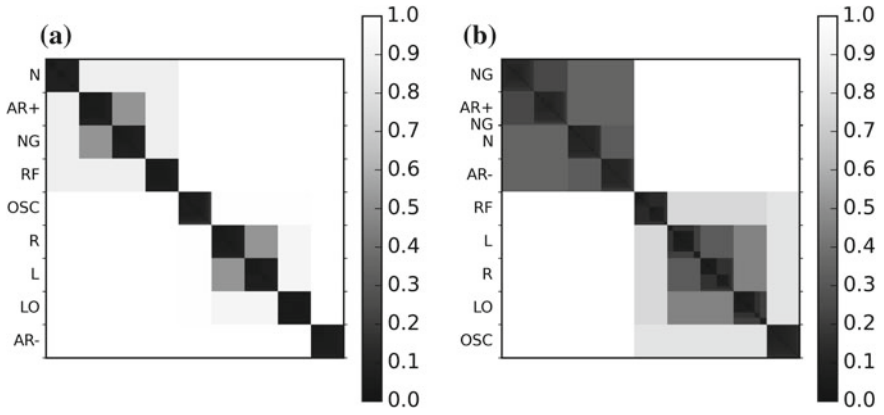

Fig. 1.3 **iVAT** matrix depicting the clustering structure of noise ratio 0.2 and 0.8. Each *black block* in this matrix contains 100 time series of a particular system. *Note* the different orderings of the systems along the y-axis are due to the **iVAT** method and cannot be changed to maintain the visual cluster tendency. The NG label appears two times in the determined ordered sequence in (**b**). It is difficult to see in the **iVAT** matrix that a few time series from the NG system are located between AR+ and N. **a** 0.2. **b** 0.8

- **RQA** parameters at predefined values (see Table 1.1).
- each time series has 25,000 data points. In this experiment, we use the time series of this experiment (see Table 1.2).
- 16 **RQA** measures from Table 1.3.

In contrast to our first experiment reported above, we add Gaussian white noise to each of the 900 time series to generate different noise ratios (our noise model is similar to [11]). We generate six sets of 900 time series representing noise ratios of 0.2, 0.4, 0.6, 0.8, 0.9, 1.0. Hence, we generate 5400 time series. The experimental procedure for each set is the same as in the experiment reported above. Again, the pair-wise dissimilarity values are different in our experiment, and therefore, the arrangement of rows and columns may differ in the resulting intensity images.

Figure 1.3a shows the **iVAT** matrix for noise ratio 0.2. This matrix is similar to the **iVAT** matrix in our initial study (see Fig. 1.2a). We observe similar results for noise ratios 0.4 and 0.6., and we conclude from this result that the clustering structure is robust against low noise ratios, i.e., 0.2, 0.4 and 0.6.

For high noise ratios, the situation becomes more complicated. Figure 1.3b shows the **iVAT** result for noise ratio 0.8. We see a noticeable difference compared to the result of the first experiment. Although we see the nine clusters along the diagonal, the black blocks are not homogeneous as with the first experiment (we see tiny block structures with each system). Furthermore, dark gray blocks capture some of the black blocks. These dark gray blocks make it difficult to decide whether NG, N, AR+, and AR− form distinct clusters. Figure 1.3b indicates that NG, N, AR+, and AR− are likely to be members of the same cluster. This is a reasonable result since

Table 1.3 RQA measures of the experiment

Description	Formula	Diagonal	Vertical	White
Filter ratio	$\frac{\sum_{l=l_{\min}}^{N} lP(l)}{\sum_{l=1}^{N} lP(l)}$	DET	LAM	–
Average length	$\frac{\sum_{l=l_{\min}}^{N} lP(l)}{\sum_{l=l_{min}}^{N} P(l)}$	L	TT	W_{mean}
Maximum length	$\max(\{l \mid P(l) > 0\})$	$L_{\max}$	$V_{\max}$	$W_{\max}$
1/Maximum length	See above	DIV	–	W_{div}
Entropy	$-\sum_{l=l_{\min}}^{N} p(l) \ln p(l)$	L_{entr}	V_{entr}	W_{entr}

In addition to these measures, we also utilize the recurrence rate (RR) as well as the ratios DET/RR and LAM/DET

these systems belong to the same class, the stochastic systems. We see a similar result for L, LO, and R systems. However, these results are still in line with expert judgment (see Fig. 1.2b). We conclude from these observations that the clustering structure is interpretable for the noise ratio 0.8 (Fig. 1.3b).

1.4 Experiment with Real-World Signals

In this section, we study whether the clustering of real-world time series based on their **RQA** measures produces an interpretable clustering structure, when compared to standard climate classification. In this experiment, we utilize climate time series ("Quality Controlled Local Climatological Data") from the National Oceanic and Atmospheric Administration (NOAA). NOAA provides a variety of time series for stations across the United States.

We select nine stations from three different zones; three stations per zone (see Table 1.4). According to the Köppen climate classification [12], the stations either

Table 1.4 Station names and their climate

Label	Location of station	Climate zone
1	McCarran International Airport, Nevada	BWh
2	Tucson International Airport, Arizona	BWh
3	Phoenix Sky Harbor International Airport, Arizona	BWh
4	Dickinson Municipal Airport, North Dakota	Dfb
5	Minot International Airport, North Dakota	Dfb
6	Bismarck Municipal Airport, North Dakota	Dfb
7	Winkler County Airport, Texas	BSk
8	Lubbock International Airport, Texas	BSk
9	Midland International Airport, Texas	BSk

The corresponding time series are depicted in Fig. 1.4. BWh = hot desert climate, Dfb = humid continental climate and BSk = semi-arid climate

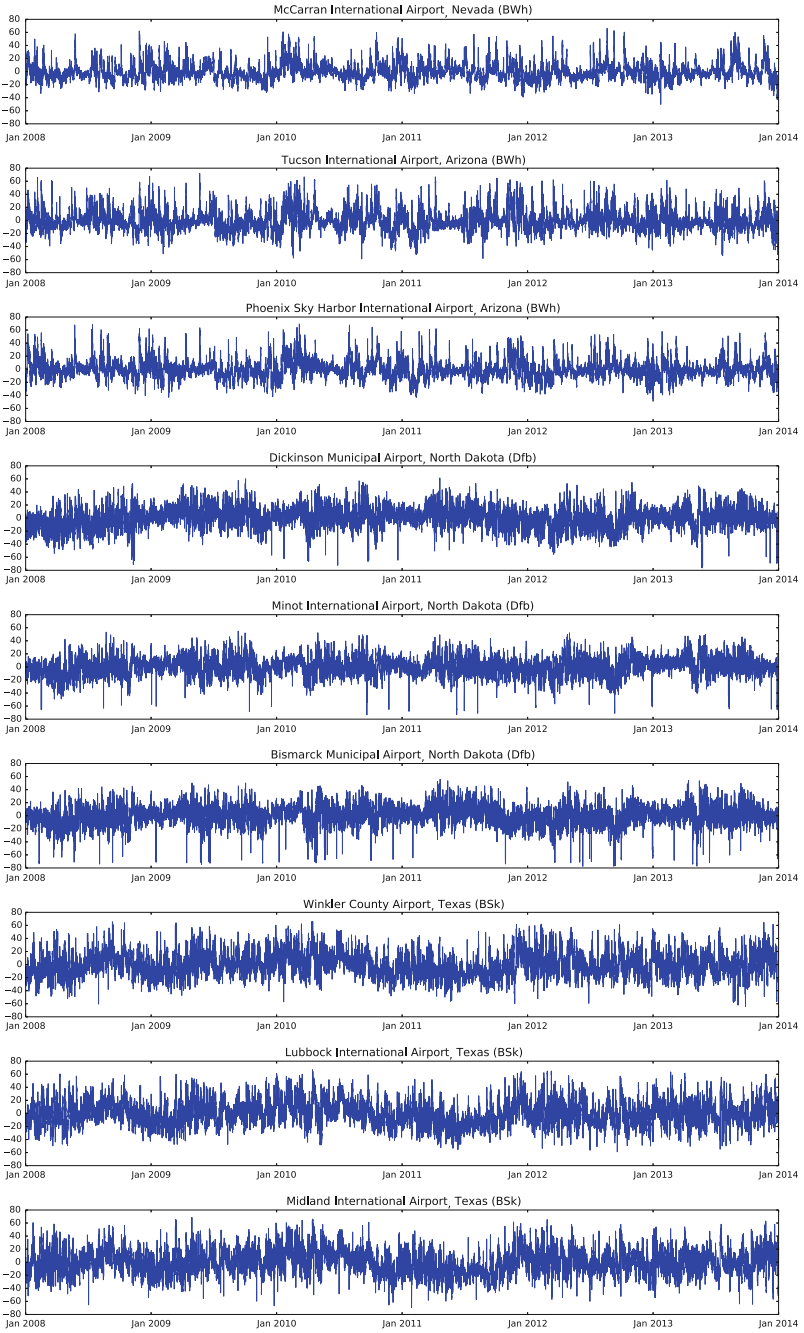

Fig. 1.4 Time series of the NOAA stations utilized in our experiment

belong to the humid continental climate (Dfb), semi-arid climate (BSk) or the hot desert climate (BWh). The aim of this experiment is to see whether the **RQA** measures will group stations belonging to the same climate into a common cluster.

In this experiment, we consider time series from January 2008 to December 2013 at an hourly resolution that capture the relative humidity measured at each station. We exclude the annual and the daily humidity cycle from the time series. Figure 1.4 presents the time series utilized in this experiment. To compute the **RQA** measures, we use the parameter assignments (see Table 1.1) and the same normalization procedure of the experiment with synthetic signals.

Figure 1.5 depicts the **iVAT** matrix for the nine stations from Table 1.4. We see three clusters along the diagonal of the **iVAT** matrix (highlighted with red bounding boxes). A close inspection of these clusters revealed that each cluster covers time series from the same climate only. Hence, the clustering structure groups stations belonging to the same climate into the same clusters.

Comparing the similarities of the stations belonging to the BWh climate, it becomes apparent that the temporal dynamics of the time series captured at the McCarran International Airport in Nevada are different to the remaining time series from the same climate. A reason for this behavior may be the differences in geographic location. Generally being a dry area, the airport is located in the Mojave desert. Nevertheless, its dynamics are more similar to the time series belonging the BWh climate than to those belonging to different climate zone, as depicted in the **iVat**.

Based on these findings, we conclude that the clustering structure of the nine stations is interpretable.

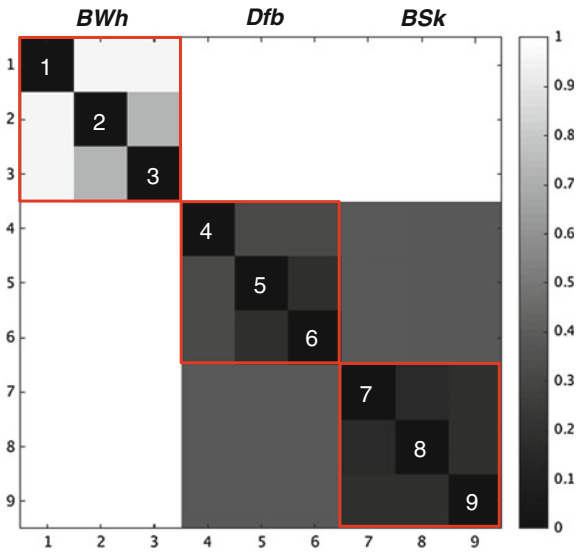

Fig. 1.5 **iVAT** matrix for the nine stations from NOAA. The *three red blocks* along the diagonal represent distinct clusters of the climate zones. Each *red block* contains three time series. Table 1.4 lists the station names for each diagonal item (labels are identical). BWh = hot desert climate, Dfb = humid continental climate and BSk = semi-arid climate

1.5 Future Lines of Research

We report these experiments for two reasons. First, answering the scientific question of this chapter is critical for our **Visual Analytics** approach. In close collaboration with experts in the field, we develop a **Visual Analytics** approach that supports users to explore large sets of a time series to identify regions with similar temporal dynamics. **Visual Analytics** has the potential to extent the users toolbox, since users often need to specify in advance what constitutes similar temporal dynamics when trying to study large sets of time series. The positive answer to the scientific question whether the clustering of time series based on their **RQA** measures facilitates the further development of our **Visual Analytics** approach.

We also plan to extend our study to multi-scale algorithms. The basic idea is to group the multi-scale components of time series into clusters of similar temporal dynamics based on their **RQA** measures and to explore this clustering structure. We argue that an exploration of this clustering structure supports users to gain a better understanding of the complex temporal behavior of systems. This study is crucial for the development of a **Visual Analytics** approach for Multi-Scale RQA.

1.6 Conclusion

The scientific question discussed in this chapter is whether the clustering of time series based on their **RQA** measures leads to an interpretable clustering structure when analyzed by human experts. To address this scientific question, we described a first experiment in which we do not introduce independent variables. In this experiment, we utilized nine well-known dynamic systems and 16 **RQA** measures. We generated 100 time series for each of the nine systems, and each time series had 25,000 data points. The dependent variable was the clustering structure.

Furthermore, we calculated the Euclidean distance between the **RQA** vectors created based on those time series and stored these pairwise distances in a matrix. We then utilized the **iVAT** method to uncover the clustering structure. The **iVAT** matrix in the first experiment shows nine distinct and well-separated clusters along the diagonal of the **iVAT** matrix. To see whether this clustering structure is interpretable, we compare our estimated clustering structure with the clustering structure based on expert opinion. We concluded from this comparison that **RQA** measures produce an interpretable clustering structure for a synthetic data set. To estimate the result of clustering algorithms, we assumed that close clusters are likely to be merged into a common cluster. The estimated clusters are in line with expert opinion.

We extended the experiment, and introduced one independent variable. We utilized noise ratio as independent variable. The experimental setup and procedure were similar to the first experiment. The **iVAT** results show that the clustering structure is robust against noise ratios up to 0.8. In addition, we conducted an experiment with the embedding dimension as the independent variable. Again, the experimental

setup and procedure were similar to the first experiment. Our result suggested that the embedding dimension has only a minor effect on the clustering structure of the synthetic time series.

In addition, we conducted an experiment with real signals. We select nine stations from three different climate zones; according to the Köppen climate classification. The **iVAT** matrix shows a well-organized clustering structure for these real-world time series. Time series from the same climate are grouped into a common cluster. We concluded from this observation that the clustering structure of the nine stations is interpretable.

Finally, our experiments to determine the similarity between RP's with Hamming distance and Spatiogram distance as alternative approaches to **RQA** vectors did not yield better results.

References

1. National centers for environmental information. Quality controlled local climatological data (2015). http://www.ncdc.noaa.gov/qclcd/QCLCD?prior=N
2. N. Marwan, M.C. Romano, M. Thiel, J. Kurths, Phys. Rep. **438**(5–6), 237 (2007)
3. N. Marwan, Eur. Phys. J. Spec. Top. **164**(1), 3 (2008)
4. J.P. Zbilut, C.L. Webber Jr., Phys. Lett. A **171**(3–4), 199 (1992)
5. C.L. Webber Jr., J.P. Zbilut, J. Appl. Physiol. **76**(2), 965 (1994)
6. N. Marwan, S. Schinkel, J. Kurths, Europhys. Lett. **101**, 1 (2013)
7. J. Bezdek, R. Hathaway, in *Proceedings of the 2002 International Joint Conference on Neural Networks IJCNN '02*, vol. 3, pp. 2225–2230 (2002)
8. R.C. Prim, Bell Syst. Tech. J. **36**(6), 1389 (1957)
9. L. Wang, U. Nguyen, J. Bezdek, C.A. Leckie, K. Ramamohanarao, in Advances, in *Knowledge Discovery and Data Mining*, vol. 6118, Lecture Notes in Computer Science, ed. by M. Zaki, J. Yu, B. Ravindran, V. Pudi (Springer, Berlin Heidelberg, 2010), pp. 16–27
10. T.C. Havens, J.C. Bezdek, IEEE Trans. Knowl. Data Eng. **24**(5), 813 (2012)
11. M. Thiel, M. Romano, J. Kurths, R. Meucci, E. Allaria, F. Arecchi, Physica D **171**(3), 138 (2002)
12. M.C. Peel, B.L. Finlayson, T.A. McMahon, Hydrol. Earth Syst. Sci. **11**(5), 1633 (2007)

Chapter 2
Applications of Transient Signal Analysis Using the Concept of Recurrence Plot Analysis

Angela Digulescu, Irina Murgan, Cornel Ioana, Ion Candel and Alexandru Serbanescu

Abstract Transient signals are universally characterized by a short duration and a broad spectrum which are often present in various phenomena such as sudden acoustic pressure changes, seismic waves, electrical discharges, etc. In order to efficiently monitor the systems where they happen, it is very important that the signals generated by transient phenomena be detected, located and characterized. This significantly helps to better understand their effects in the given application context. This chapter presents new tools derived from the concept of Recurrence Plot Analysis (RPA) and applied on three real applications. Two of the applications concern the detection, localization and characterization of the electrical partial discharges (measured from photovoltaic panels and on electrical cables, respectively). Another application refers to the quantification of the water hammer effect using two acoustic sensors placed on a pipe line.

2.1 Introduction

Complex systems are often met in real life and they usually present highly nonlinear (and sometimes linear) deterministic, stochastic and random characteristics [1]. These systems comprise different subparts which are strongly interconnected, hereby

A. Digulescu (✉) · I. Murgan · C. Ioana · I. Candel
GIPSA-lab, Grenoble Institute of Technology, Grenoble, France
e-mail: Angela.Digulescu@gipsa-lab.grenoble-inp.fr; adigulescu@mta.ro

I. Murgan
e-mail: Irina.Murgan@gipsa-lab.grenoble-inp.fr

C. Ioana
e-mail: Cornel.Ioana@gipsa-lab.grenoble-inp.fr

I. Candel
e-mail: Ion.Candel@gipsa-lab.grenoble-inp.fr

A. Digulescu · A. Serbanescu
Military Technical Academy, Bucharest, Romania
e-mail: alexe1serbanescu@yahoo.com

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8_2

the interdependence of their characteristics is difficult to depict, and, therefore, the system has to be characterized as a whole and not individually. This is the reason why for most real applications, the measurements are very well suited to reveal the resultant effect of the processes that describe a phenomenon under study [2].

Hereby, the choice of the RPA concept for the analysis of transient signals is based on the fact that it is a data-driven method which does not require a priori information about the system, knowing that such information is not always available [3].

Our applications under study concern the transient signals that appear in hydraulic and power systems. The major issue is that these signals reflect a sudden change of the dynamical system which can cause, in an unpredictable laps of time, a breakdown of the system.

The recurrence information is very important, offering us new insights in the analysis of transient signals which represent totally different states of the systems. In our work, we are interested in the system's state changes that are not determined by random causes, but they are the results of a nonlinear input that causes them to change their state suddenly, exposing the system to major collapse.

The first application relates to the electrical *partial discharge* (PD) detection and characterization [4]. The PDs indicate that some changes have occurred in the insulation due to chemical and/or mechanical transformations [5], which, in time, can lead to the failure of the equipment. Hereby, the PD measurement is a routine procedure for testing important components from the power system (high-voltage cables, transformers, etc.).

The second application concerns the detection, localization and characterization of *electrical arcs* generated in photovoltaic panels [6, 7]. The need of detection, localization and characterization of the electrical arcs is a growing demand as these systems continue to develop and the environmental conditions still unexpectedly change.

Next, our application refers to the *water hammer effect* which appears in pipelines when a valve is suddenly closed, so it forces the fluid to change its direction or to stop its flow. This translates to a pipe pressure sudden increase/decrease which causes from vibrations of the pipe to system collapse. Thus, this phenomenon must be supervised and characterized in order to control its damaging effects to the hydraulic system.

Through these specific applications, our chapter shows the interest of RPA approaches for the analysis of the transient signals in various applications of nowadays interest.

The chapter is organized as follows: the second section presents some relative new signal analysis tools based on the RPA concept. Next, each section presents the applications mentioned above and discusses on the subject. The last section illustrates the conclusions and perspectives of our work.

2.2 Signal Analysis Tools Based on the RPA Concept

In this section, starting from the concept of recurrence plot analysis, two new analysis tools derived from RPA concept are presented.

Firstly, there will be highlighted the measure used for the **detection of a transient signal**, namely the *time-distributed recurrence* (TDR) measure. Then, the *multi-lag phase-space analysis* will be introduced. This concept is very useful for the **characterization of transient signals**.

2.2.1 The Time-Distributed Recurrence Measure

The basis of this measure starts from the idea that a sudden change in a time series represents a new state of the dynamical system [3, 8–14], namely there is no recurrence with the previous states. Therefore, when an appropriate distance is used, the recurrence matrix presents a horizontal/vertical band with much fewer recurrences. When the sum of the lines/columns of the recurrence matrix is computed, we actually obtain the column average [3], which, in the case of transient signals, significantly changes.

Considering a measured signal as the following time series [15–21]:

$$\mathrm{s}[n] = \{s[1],\ s[2],\ \ldots,\ s[N]\} \tag{2.1}$$

where N is the length of the signal, then the phase-space points of the system are obtained from the available time series:

$$\overrightarrow{v_i} = \sum_{k=1}^{m} s[i + (k-1)d] \cdot \overrightarrow{e_k} \tag{2.2}$$

where m is the embedding dimension of the phase-space, d is the delay (lag) chosen between the samples of the time series and $\overrightarrow{e_k}$ are the axis unit vectors corresponding to each dimension of the phase-space.

Then, the distance/recurrence matrix is obtained:

$$R_{i,j} = \Theta\left(\varepsilon - \left\|\overrightarrow{v_i} - \overrightarrow{v_j}\right\|\right),\quad i, j = \{1, 2, \ldots, M\} \tag{2.3}$$

where $\|\cdot\|$ is a certain chosen distance (Euclidean distance [8, 21], angular distance [3, 8], L1 norm [22], etc.) and the $\Theta(\cdot)$ is the Heaviside step function. For our applications, the threshold ε is considered constant and $M = N - (m-1)d$.

Hereby, the time-distributed recurrence (TDR) measure is defines as [3, 8]:

$$TDR[n] = \frac{1}{M}\sum_{i=1}^{M} R_{i,n} \tag{2.4}$$

This measure can be interpreted as the column average recurrence of a given point i or the recurrence density heterogeneity in the point i. Hereby, a solitary position of a phase space vector changes significantly its average recurrence and can be highlighted through the use of the complementary version of the measure from (2.4):

$$TDR^{*}[n] = 1 - \frac{TDR[n]}{\max\{TDR[n]\}} \tag{2.5}$$

Moreover, in order to detect only the transient signal from the analyzed observation, the signal-to-noise ratio (SNR) must be computed.

The SNR is computed as follows: the last part of the acquired signal (when no phenomena is happening) is considered as noise, $z[i]$, (unwanted signal recording environmental noise, cable noise, noise caused by imperfect connections, etc.), whereas the part that has a different behavior is considered as the interest signal, $s[i]$. Both parts of the signal, $s[i]$ and $z[i]$ have the same length, N. The SNR is computed as:

$$SNR = 10\log_{10}\left(\frac{\sum_{i=1}^{N}(s[i])^2}{\sum_{i=1}^{N}(z[i])^2}\right) \tag{2.6}$$

Accordingly, the threshold ε of the recurrence matrix is chosen so that it includes the a percentage, α, of the maximum value of the equivalent noise which has the same power as the interest signal:

$$\varepsilon = \alpha \cdot \max\left(\left|SNR_{signal} \cdot noise\right|\right) \tag{2.7}$$

where α is a constant that is chosen to nonlinearly filter the noise. In our applications, α varies from 0.4 to 0.95.

Considering the (2.7), the components of the noise (undesired parts of the signal) are considered as recurrences, therefore the transient signal (useful part of the signal) is highlighted by the proposed measure.

Figure 2.1 emphasizes the advantage brought by the TDR measure, namely that the detection provided by the detection curve is more robust to noise the actual recorded signal (electrical arc acquisition). In the examples above, the detection curve exhibits a SNR improvement of 30 dB.

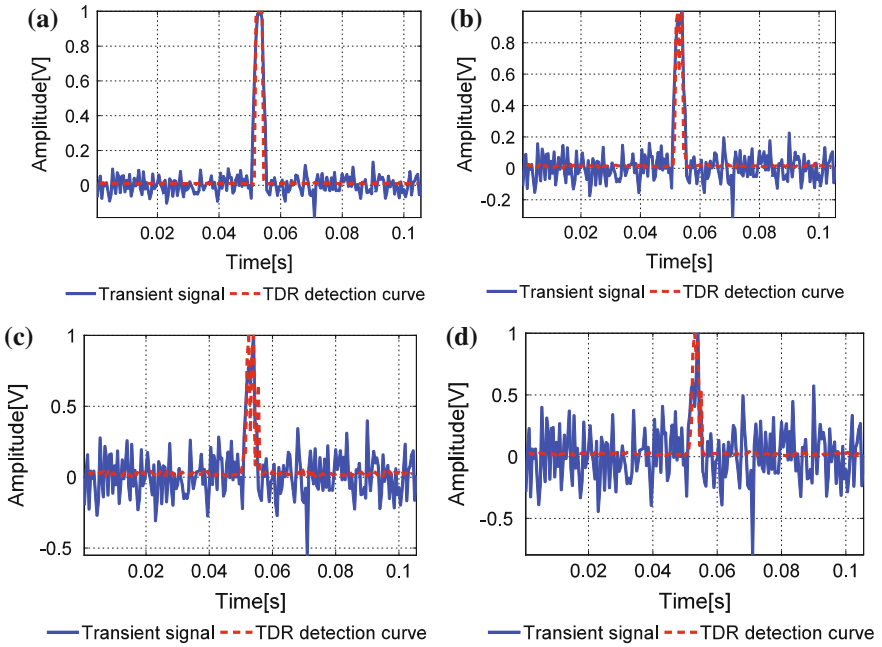

Fig. 2.1 Transient signal detection using the TDR measure for a transition of only 11 samples where $m = 3$, $d = 2$ and $\alpha = 0.8$. **a** Transient signal detection for signal with SNR = 20 dB. **b** Transient signal detection for signal with SNR = 15 dB. **c** Transient signal detection for signal with SNR = 9 dB. **d** Transient signal detection for signal with SNR = 5 dB

The major advantage of this method, is that, in applications where the SNR level of the acquired signal varies a lot, the TDR detection curve approach improves the SNR level after filtering the acquired signal with the detection curve. Through this approach, the signal's power after filtering is significantly increased (in our applications, at least 10 dB), which is very helpful for the part of signal classification (Sects. 2.3 and 2.4). Moreover, this provides the robustness to the TDOA (time difference of arrival) estimation. This comes from the measure's invariance to the group velocity effect (thanks to the concept of recurrence), whereas the peak detection or the correlation function is very sensitive to this effect (Sect. 2.5).

2.2.2 Multi-lag Phase-Space Analysis

The RPA method stands, as its name suggests it, on the concept of recurrences. Still, for our applications, a step backward has been made and a closer attention is given to the phase-space in order to achieve richer characterization of similar signals coming from the same source and having the same propagation and acquisition conditions.

The phase-space representation is very rich in information regarding the evolution of a dynamical system [22, 23], therefore the analysis can be made in any dimension, but our approach restrains, for the moment, to a bi-dimensional representation.

Therefore, considering the Cartesian coordinate system, the phase-space representation is given by:

$$\begin{cases} x_n = s[n] \\ y_n = s[n+d] \end{cases} \tag{2.8}$$

where n is the nth sample of the recorded time series. It goes that (2.2) can be rewritten as: $\overrightarrow{v_n} = s[n] \cdot \overrightarrow{i} + s[n+d] \cdot \overrightarrow{j}$.

Firstly, the main attributes of this representation are recalled. Considering three signals s_1, s_2 and s_3 defined as:

$$\begin{aligned} s_1[n] &= s[n+\delta] \\ s_2[n] &= s[\alpha n] \\ s_3[n] &= \beta \cdot s[n] \end{aligned} \tag{2.9}$$

where α, β and δ are constant which modify the signal s through translation in time, scale and or amplitude, the phase-space points present the following attributes [20, 24, 25]:

$$\begin{aligned} \overrightarrow{v_{1[i]}} &= \overrightarrow{v_{[i+\delta]}} \\ \overrightarrow{v_{2[i]}} &= \overrightarrow{v_{[\alpha i]}} \\ \overrightarrow{v_{3[i]}} &= \beta \cdot \overrightarrow{v_{[i]}} \end{aligned} \tag{2.10}$$

The phase-space trajectory is invariant to translation and it points out the scale and amplitude change.

In order to introduce the concept of multi-lag phase-space analysis, three transient signals with similar characteristics are considered. These signals are given by the generic signal:

$$s_{(a,f,b)}[n] = \begin{cases} a \cdot (\sin[2\pi \cdot f \cdot n] + b[n]), \ n = \{1, 2, \ldots, N\} \\ a \cdot b[n], \ otherwise \end{cases} \tag{2.11}$$

where $N = \left[\frac{f_s}{2f}\right]$ and $f_s = 10\,\text{MHz}$ is the sampling frequency.

The considered signals $s_{1(a_1, f_1, b_1)}$, $s_{2(a_2, f_2, b_2)}$ and $s_{3(a_3, f_3, b_3)}$ have the following relationships between their parameters: $a_1 / a_2 = 1 / 0.6 = 1.66$, $a_1 / a_3 = 1 / 0.3 = 3.33$, $f_1 / f_2 = (2 \times 10^5) / (1.9 \times 10^5) = 1.05$, $f_1 / f_3 = (2 \times 10^5) / (1.6 \times 10^5) = 1.25$, $b_1 = b_2 = b_3$ and $SNR_i = 20\,\text{dB}$, $i = \{1, 2, 3\}$.

The chosen signals from Fig. 2.2 seem to be quite similar, but, at a closer look, slight differences appear. Firstly, these signals are studied using the wavelet analysis [26–28].

From Fig. 2.3, it can be observed that the slight differences between the signals cannot be highlighted by the wavelet transform: their presence is detected, by their shape does not present any discriminating element.

In order to distinguish between these transient signals, the multi-lag phase space analysis is considered. The study of the representation of the trajectory for multiple lags is done by two approaches: the elliptic modeling of the trajectory for the area estimation in the phase-space, respectively, the polar coordinate representation.

The elliptic modeling supposes to determine the ellipse that circumscribes the phase-space trajectory and, therefore, to estimate the area of the trajectory through the ellipse's area computation. The area of the modeling ellipse is a new descriptor of the transient signal.

The trajectory is modeled by considering the solution that minimizes the following system [29]:

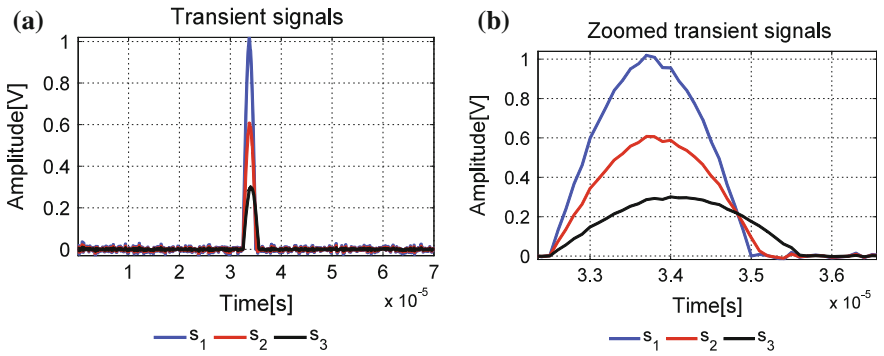

Fig. 2.2 Transient signals considered for the multi-lag phase-space analysis

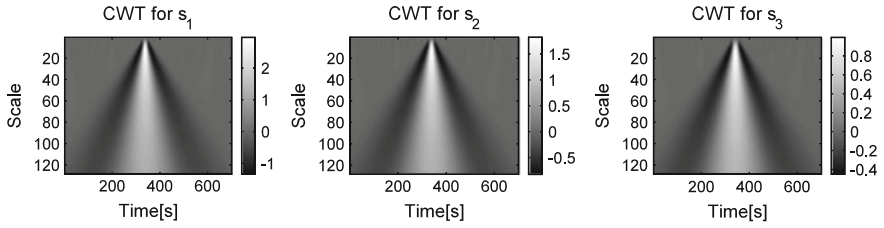

Fig. 2.3 The continuous wavelet transform (with the Mexican Hat mother-function) applied on the three transient signals

$$S = \sum_{i=1}^{M} [F(x_i, y_i)]^2 \tag{2.12}$$

where $F(x, y) = \Gamma x^2 + \Lambda y^2 + 1$. The least mean square estimation of (2.12) gives the couple (Γ, Λ). It goes that the major semi axes $a = 1/\Gamma$ and the minor semi axes $b = 1/\Lambda$.

The next step after the elliptic modeling is the area estimation and the estimation of the optimal delay. The delay is considered to be optimal for the value that provides an average value of the area and provides an adequate representation of the trajectory. This average value area assures a suitable phase-space representation where the trajectory does not evolve too close to the main diagonal (case of redundancy) or its evolution is too complicated (case of irrelevance) [30].

From Fig. 2.4, it can be noticed that the area of the signals has a similar trend with some differences for the normalized signals. Hereby, considering the average value of the area, $A = 3$, the optimal delay for s_1 is $\tau_1 = 5$, for s_2, $\tau_2 = 5$ and for s_3, $\tau_3 = 6$.

With the chosen lags, the next step is to plot the phase-space trajectory into polar coordinates:

$$\begin{aligned} \varphi[n] &= \sqrt{x_n^2 + y_n^2} \\ \theta[n] &= \arctan\frac{y_n}{x_n} \end{aligned} \tag{2.13}$$

This representation is very useful for signals with different amplitudes, because it shows the evolution of the position vector regardless of its length.

Figure 2.5 shows the evolution of the transient signals on the phase diagram. It can be observed that the noise is concentrated in a small region of the phase space which is translated in a small points conglomeration around the peak corresponding to the angle $-3\pi/4$.

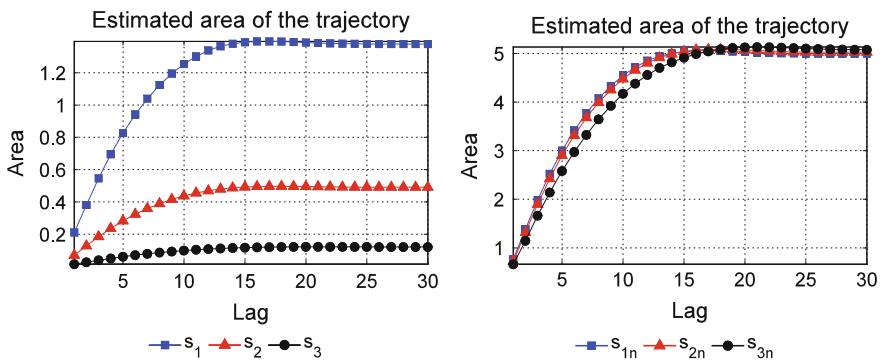

Fig. 2.4 The evolution of the area according to the lag (delay) for the signals presented in Fig. 2.2 (*left figure*) and for the normalized signals (*right figure*)

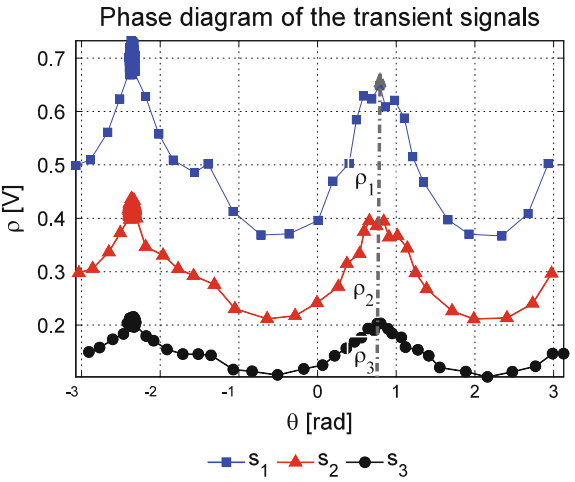

Fig. 2.5 The polar coordinates representation for the transient signals s_1: $\tau_1 = 5$, s_2: $\tau_2 = 5$ and s_3: $\tau_3 = 6$

In this way, through the isolation of this region, the noise can be eliminated in the process of signal characterization. Moreover, the evolution of the trajectories is different for the peaks corresponding to the angle $\pi/4$. Measuring the values of the maximum length of the vector corresponding to the $\pi/4$ angle, it goes: $\rho_1/\rho_2 = 1.67$ and $\rho_1/\rho_3 = 3.43$. So, the ratio of the amplitude of the transient signals is then conserved in the phase diagram.

Furthermore, the number of points between the two angles $-3\pi/4$ and $\pi/4$ (after excluding the points corresponding to the noise) is directly related to the fundamental frequency of the signal. Let n_0 be this number of points from the phase diagram. The fundamental period of the signal, T_0 is:

$$T_0 = \frac{2n_0}{f_s} \tag{2.14}$$

From (2.14), it goes that: $f_1/f_2 = 1.04$ and $f_1/f_3 = 1.24$.

This results allow us to discriminate between the proposed signals although their characteristics are very close.

2.3 Characterization of Partial Discharges in High Voltage Cables

In high voltage systems the presence of partial discharges (PD) are an indication of insulation weakness which, in time, may lead to total damage of the equipments [5]. Therefore, is it absolutely necessary to monitor such systems (power cables, transformers, etc.) in order to detect and localize the PD source, namely the position of the

insulation troubles. Moreover, the characterization of these signals provides extra information regarding the long terms effects that they have upon the system [5].

The experiment was made on the grounding connection of the cable in order to record signals of 20 ms using high current inductive sensors and high speed data acquisition. It goes that each recorded signal has 2 million samples and a preliminary detection of the potential harmful zones is achieved using the spectrogram (Fig. 2.6).

Then, the *TDR* measure is applied on partial discharges that have different *SNRs*.

The PDs presented in Fig. 2.7 are detected with the *TDR* measure so that the noise has no impact upon the detection curve. Hereby, after the filtering of these signals, the filtered PDs have a *SNR* improved with at least 20 dB.

Next, these PDs are characterized using the multi-lag phase-space analysis. Firstly, the signals were normalized in order to eliminate the drawbacks that the different PDs amplitudes would involve. Then, on these signals, the area estimation is performed after the elliptic modeling. Figure 2.8 presents the obtained results.

It can be observed that, even if the signals are normalized, their area evolution is different. The choice of the lag is done so that the area of the trajectory on the phase space has the same average value ($A = 1.5$): $d_{PD_1} = 3$, $d_{PD_2} = 4$ and $d_{PD_3} = 6$.

Using these lags, in Fig. 2.9 the phase diagram representation is presented.

The evolution of the trajectory is better pointed out in the $\pi/4$ angle region than in the $-3\pi/4$ angle region. Moreover, the noise in distributed all along the peak corresponding to the $-3\pi/4$ angle depending on the *SNR*.

Recalling the fact that the amplitude ratio is conserved in the ratio between the lengths of the vectors corresponding to the $\pi/4$ angle, means that: $a_{PD_1}/a_{PD_2} = 1.53$ and $a_{PD_1}/a_{PD_3} = 2.87$. Moreover, according to (2.14), the frequency ratio are $f_{PD_1}/f_{PD_2} = 1.15$ and $f_{PD_1}/f_{PD_3} = 1.37$.

Hereby, the results show that the signals PD_2 and PD_3 suffer not only an attenuation, but also a frequency shift with respect to PD_1 which helps to establish their source characteristics.

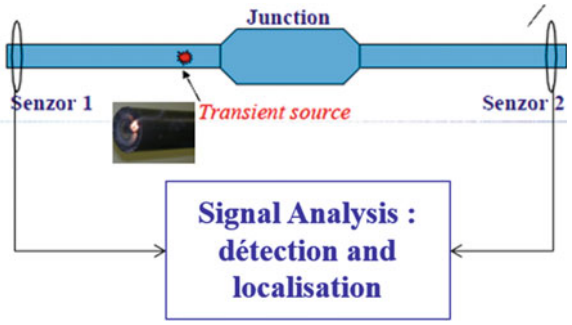

Fig. 2.6 The experimental configuration for the PD measurements

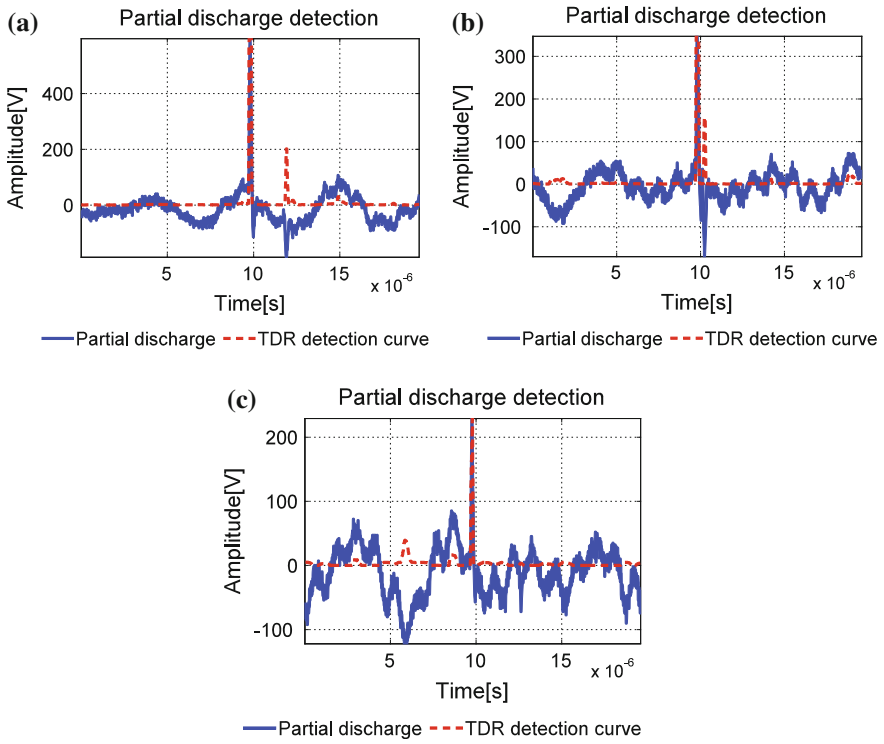

Fig. 2.7 Recorded partial discharges with different *SNRs* and ratios α for ε (2.7) **a** $SNR = 19.7$ dB, $m = 3$, $d = 3$, $\alpha = 0.4$; **b** $SNR = 14$ dB, $m = 3$, $d = 4$, $\alpha = 0.7$; **c** $SNR = 8$ dB, $m = 3$, $d = 6$, $\alpha = 0.95$

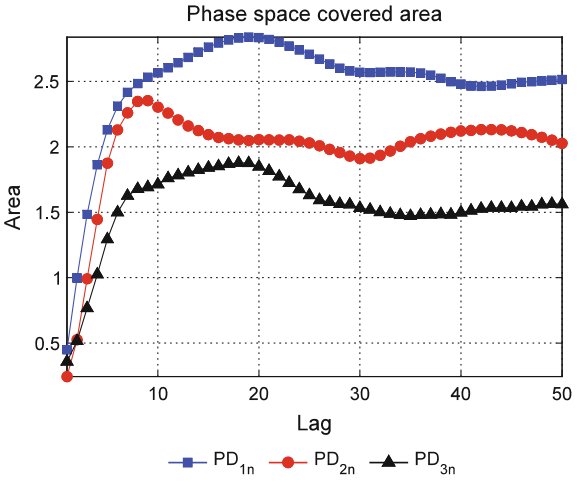

Fig. 2.8 The area estimation after the elliptic modeling in multi-lag representation

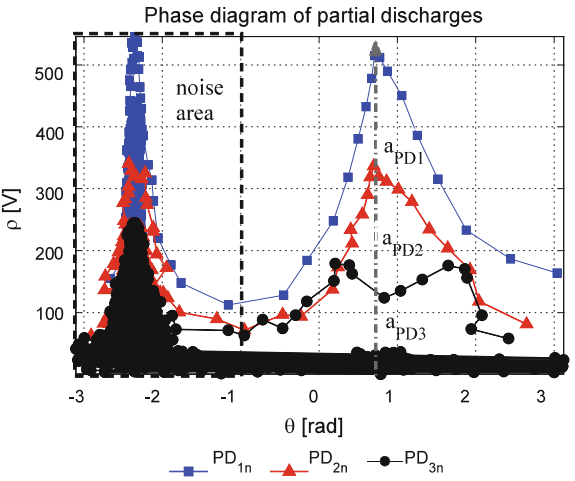

Fig. 2.9 The phase space representation in polar coordinates for $PD_1: d_{PD_1} = 3$, $PD_2: d_{PD_2} = 4$ and $PD_3: d_{PD_3} = 6$

2.4 Electrical Arcs in Photovoltaic Panels

Photovoltaic panels are very important in the landscape of renewal energy sources of strategic interest, for both ecological reasons and the worldwide growing energy demand. The electrical arcs (EA) that appear in these systems can be a major problem, so it is necessary to supervise such phenomena in the system in order to keep it safe [7].

The experiment was performed with an electrical arc locator system composed of three acoustical microphones place in a 3D configuration and an wide band antenna placed in the center of the system (Fig. 2.10).

The detection and estimation of times of arrival of the transient signals generated by electrical arc at each sensor has been done using the *TDR* measure, the spectrogram and the wavelet transform. Figure 2.11 presents the recorded signals and their detection curves obtained with the TDR measure.

The recording is done at a sampling frequency of $f_s = 5\,\text{MHz}$ for a period of 10 ms (50000 samples). In terms of location accuracy, the *TDR* measure is compared with the classical time-scale approaches. The spatial localization is achieved by solving the geometrical system (2.15) based on the time-of-arrival (TOA) of the electrical arc at each microphone. The TOA is obtained by imposing the same threshold (0.5) at the normalized detection curves based on each method.

$$\begin{cases} d_{PS_2} - d_{PS_1} = v \cdot t_{21} \\ d_{PS_3} - d_{PS_1} = v \cdot t_{31} \\ d_{PS_4} - d_{PS_1} = v \cdot t_{41} \end{cases} \tag{2.15}$$

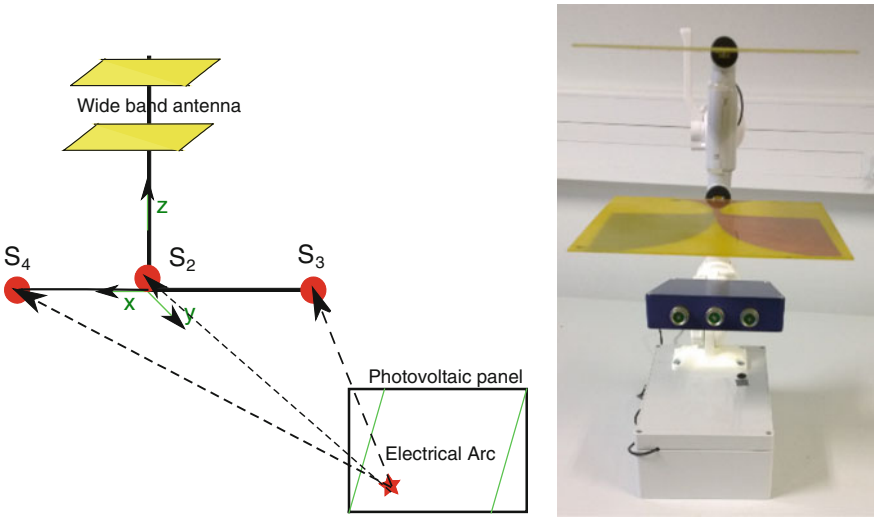

Fig. 2.10 The experimental configuration of the EA locator system

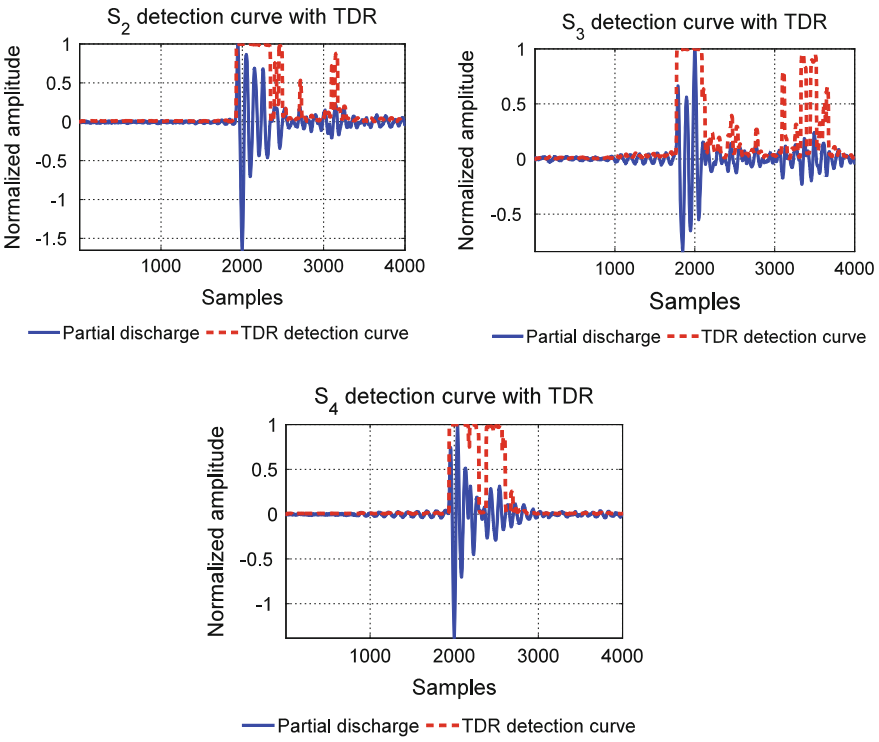

Fig. 2.11 The electrical arc locator system, the recorded AEs and the detection curve obtained with the TDR measure $m = 3$; $d = 8$; $\alpha = 0.7$

Table 2.1 The spatial localization accuracy for the electrical arcs

Method	Relative error (%)
TDR measure	6.2
Wavelet	11.2
Spectrogram	9.4

where $d_{PS_i} = \sqrt{(x_{s_i} - x_P)^2 + (y_{s_i} - y_P)^2 + (z_{s_i} - z_P)^2}, i = \overline{1,4}$, each position of the acoustic microphone and antenna (S_1) is known $S_i(x_{S_i}, y_{S_i}, z_{S_i})$, $t_{(i+1)i}$ is the time of arrival of the electrical arc at each microphone triggered by the S_1 wide band antenna.

The precision accuracy using each method is presented in Table 2.1.

In this application, the signal is propagated on multiple paths, because the experiment is performed in a closed facility: a 75 m^2 laboratory sustained by 6 columns and equipped with test tables, desktops and test boards.

Although, the classical techniques detect and localize the electrical arc source in an effective manner, when it comes to multi-path signal discrimination, these methods are limited.

But, the multi-lag phase space analysis provides better results. Using one of the signals arrived at S_2 microphone, the reflections s_2, s_3 of the electrical arc are compared with the direct path signal s_1.

Therefore, for the signals highlighted in Fig. 2.12, the elliptic modeling is applied and the area of the estimated ellipse is determined. Previously, the signals are normalized in order to bring the signals at the same amplitude level (Fig. 2.13).

The estimated ellipses evolve in a different manner, but they have the same trend. After an average value of the area is chosen, $A = 1$, an optimal lag is determined: $d_1 = 11$, $d_2 = 12$ and $d_3 = 8$. The results of the phase diagram representation are shown in Fig. 2.14.

In terms of attenuation, the phase diagram points out the same information as the time evolution of the signals. Moreover, the reflections present fewer curves that the direct path signal. This means that the reflections contain fewer oscillations than the direct signal, hereby, the reflections are dispersed.

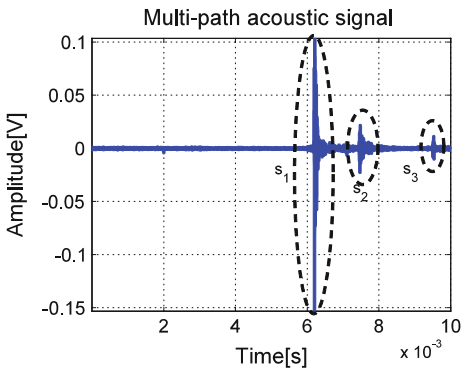

Fig. 2.12 The electrical arc recording s_1 and its reflections s_2, s_3

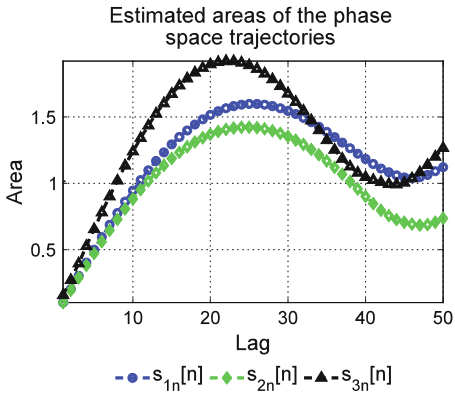

Fig. 2.13 The areas of the estimated ellipses that circumscribe the phase-space trajectory

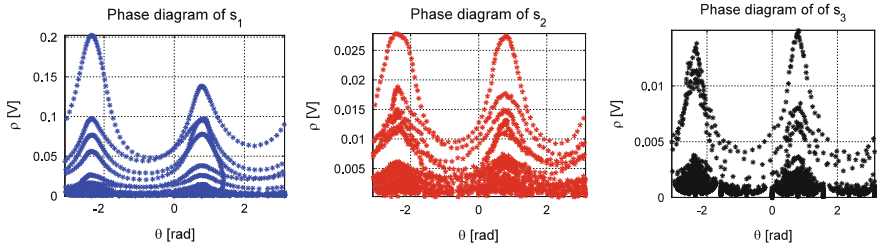

Fig. 2.14 The polar coordinates representation for the multi-path acoustic signals $s_1: d_1 = 11$, $s_2: d_2 = 12$ and $s_3: d_3 = 8$

Using (2.14), it goes that: $f_{s_1}/f_{s_2} = 0.94$ and $f_{s_1}/f_{s_3} = 0.92$, meaning that the reflections suffer a time dilatation. Because the frequency ratios are close to 1, it leads to the idea that the signals have the same source, but on their propagation path, the reflections are affected by multiple phenomena: diffusion, dispersion, attenuation, etc.

Concluding this part, the RPA approach provides a better localization accuracy than the time-scale methods and with the use of multi-lag phase-space analysis, it highlights new information regarding the characterization of transient signals and the changes that they suffer.

2.5 Water Hammer Effect Quantification

The water hammer is a phenomenon that take place in a closed pipeline when a vane is suddenly closed while the water is flowing. The effect is that the liquid is forced to change its direction or to stop its motion. The risks of this operation vary from pipeline vibration to pipe collapse.

Usually, in industrial applications, this phenomenon is quantified using either intrusive pressure sensors, either a hydraulic formula (2.16) with the condition that the characteristics of the system are well known [31–33].

$$c = \frac{1}{\sqrt{\rho\left(\frac{1}{\varepsilon_w} + \frac{D}{e \cdot E}\right)}} \tag{2.16}$$

where c is the pressure wave speed, ε_w is the bulk modulus of the fluid, D is the pipe diameter, e is the pipe wall thickness and E is the Young modulus of the pipe.

The direct relation between the speed of the pressure wave and the pressure variation Δp is given by Joukovski's equation [31]:

$$\Delta p = \rho \cdot c \cdot v_0 \tag{2.17}$$

where ρ is the fluid density and v_0 is the steady flow velocity.

Our approach consists in placing a pair of ultrasonic sensors on the pipe and to record the acoustical effect of the water hammer in order to compute the pressure wave speed. This approach has the advantage to supervise the system as it is and not to require any additional intrusive changes to the system as inserting a pressure sensor inside the pipe (2.17) or determining the exact characteristics of the hydraulic system (2.16).

The experiment is done on a horizontal pipe supplied by a tank (ST) of 200 l volume. The pipe has a length of $L = 10.11$ m, it is made from Plexiglas with an exterior section of $50 \times 50\,\text{mm}^2$ and a circular interior section of $D = 39$ mm. The water evacuation diaphragm has a diameter of $d = 20$ mm.

Our ultrasonic transducers are placed on the pipe at a distance of 8 cm (S_2), respectively 16 cm (S_1) from the closing vane (CV). Next to them, the pressure sensor P is already installed in the pipe (Fig. 2.15).

The acquired signals, by both ultrasonic and pressure sensors are presented in Fig. 2.16. The highlighted areas emphasize that the acoustical effect happens simultaneously with the pressure variations.

The acoustic signal arrived at sensor S_2 is more clearer than the one arrived at S_1, because the first one is closer than the second from the CV. This happens because the wave starts to diffuse. The results obtained with the *TDR* measure are shown in Fig. 2.17.

Although, the effect of the mechanical vibrations of the pipe are also recorded by the sensors (especially sensor S_1), the transient detection is successfully obtained.

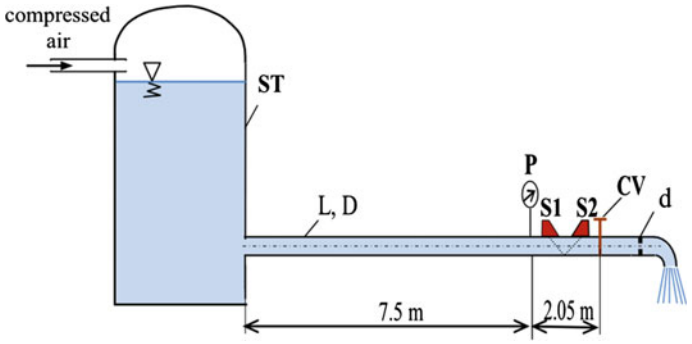

Fig. 2.15 The experimental configuration

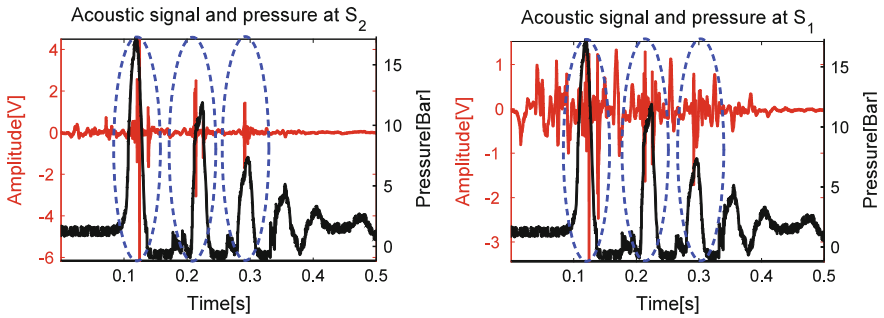

Fig. 2.16 The acoustic and pressure signals at sensor S_2 (*left figure*), respectively at sensor S_1 (*right figure*)

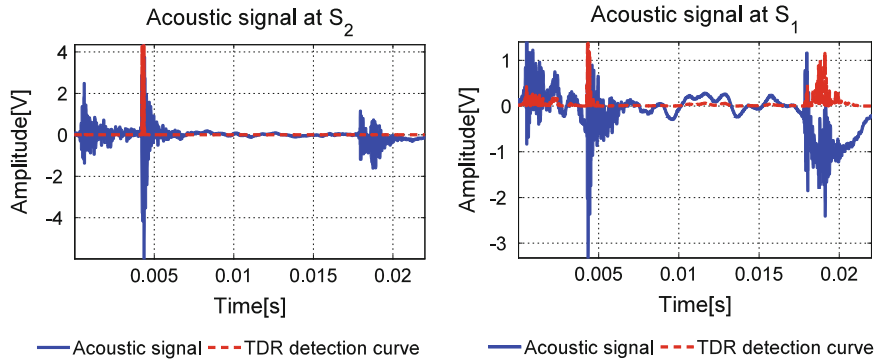

Fig. 2.17 The water hammer effect detection on the signals recorded by the two sensors $m = 3$, $d = 10$, $\varepsilon = 70$

Table 2.2 The relative error using different approaches

Method	Relative error
Experimental (c_e)	–
TDR measure (c_{TDR})	1.69 %
Maximum peak (c_{pk})	16.02 %
Theoretical (c_{th})	3.43 %

After imposing a threshold for the *TDR* measure of 80 % from its maximum value, two times of arrival are determined. Hereby, the time difference of arrival obtained is $TDOA = 82\,\mu s$. Therefore, the estimated wave speed is $c_{TDR} = 975\,m/s$.

Computing only the $TDOA = 99\,\mu s$ for the maxima of the recorded signals, the estimated wave speed is $c_{pk} = 805\,m/s$. The difference between these approaches is explained by the diffusion.

With the use of (2.16) and considering $\rho = 1000\,kg/m^3$, $\varepsilon_w = 2.1$ GPa, $e = 10$ mm and $E = 5.66$ GPa, the theoretical value of the wave speed is $c_{th} = 926\,m/s$.

Still, the pressure variation recorded by the pressure sensor P is $\Delta p = 15.9$ bar. Knowing that the flow rate of the water inside the pipe before the closure is $Q = 1.98\,l/s$, it goes that the speed of the water in steady state is $v_0 = 1.66\,m/s$. Making use of (2.17), the experimental pressure wave speed is $c_e = 959\,m/s$.

Relating the pressure wave speed the pressure variation using (2.17), the relative error for the estimation of the pressure variation is performed, namely the water hammer effect is quantified.

The results from Table 2.2 state that the *TDR* measure based on the RPA concept is very efficient in the water hammer effect quantification with an error below 2 %.

2.6 Conclusions

The concepts of RPA and phase space bring new insights which together with classical signal processing methods can help the analysis the transient signals.

The choice of the RPA concept has the advantage to be a *data-driven method*, therefore, it is an *alternative* to the classical transient signal processing techniques based on projection of analyzed signals on a given dictionary. Three applicative contexts have been addressed in our work.

The first one is the electrical partial discharge analysis. The detection method is based on the RPA method parameters using the *TDR* measure. Furthermore, the signals are characterized and discriminated using the concept of *multi-lag phase space analysis*. In addition, for the *electrical arcs* the detection and characterization is similarly obtained. The localization precision accuracy outperforms the classical non-stationary signals processing methods.

The third application, related to the *water hammer* phenomenon analysis, is pointed out through our ultrasonic non-intrusive approach that proves to be the closest to the reference method. The effect of the phenomenon is quantified using

the *TDR* measure which is more robust to the effect of pressure wave diffusion that takes place inside the pipe during the experiment.

Our future work foresees to correlate the information given by this new approach with the physical parameters of the system. For this purpose, our main research efforts will concentrate to the development of new descriptors of transients signals, derived from multi-lag phase diagram analysis. The characteristics of such descriptors that we look for are both the parsimony and the robustness to disturbing factors.

In parallel, new applications domains will be addressed aiming to provide new practically-oriented approaches for transient phenomena.

Acknowledgement This work has been supported in part by the Dema'Loc project funded by Institut Carnot "Energies du futur" and by the "Smart Hydro Monitoring" project (Rhone Alpes "Tenerrdis" and "Minalogic" research clusters).

References

1. J. Gao, Y. Cao, W.W. Tung, J. Hu, *Multiscale Analysis of Complex Time Series: Integra-tion of Chaos and Random Fractal Theory, and Beyond* (Wiley, New Jersey, 2007)
2. A. Serbanescu, L. Cernaianu, C. Ivan, New approaches in nonlinear dynamics analysis of complex systems and processes, in *International Symposium of Electronics Computers and Artificial Intelligence,* Pitesti, 2009
3. C. Ioana, A. Digulescu, A. Serbanescu, I. Candel, F.-M. Birleanu, in *Recent Advances in Non-Stationary Signal Processing Based on the Concept of Recurrence Plot Analysis, Springer Proceedings in Mathematics and Statistics*, ed. by M. Marwan et al. Translational Recurrences, pp. 75–93. Cham, Switzerland, 2014
4. I. Candel, A. Digulescu, A. Serbanescu, E. Sofron, Partial discharge detection in high voltage cables using polyspectra and recurrence plot analysis, in *9th International Conference on Communications (COMM).* Bucharest, Romania, 2012
5. S.A. Boggs, G.C. Stone, Fundamental limitations in the measurement of corona effect and partial discharges, in *Ontario Hydro Research, Toronto, Canada IEEE Transactions on Electrical Insulation,* vol. EI–17, no.2, (1982), p. 143
6. A. Digulescu, M. Paun, C. Vasile, T. Petrut, D. Deacu, C. Ioana, R. Tamas, in Electrical *Arc Surveillance and Localization System Based on Advanced Signal Processing Techniques IEEE International Energy Conference (ENERGYCON)* (Dubrovnik, Croatia, 2014)
7. C. Strobl, P. Meckler, Arc Faults in photovoltaic systems, in *Proceedings of 25th ICEC and 56th IEEE Holm Conference on Electrical Contacts,* Charleston, 2010, pp. 216–222
8. F.M. Birleanu, I. Candel, C. Ioana, C. Gervaise, A. Serbanescu, G. Serban, A vector approach to transient signal processing, in *The 11th Conference on Information Science, Signal Processing and their Applications,* Montreal, Canada, 3–5 July 2012
9. N. Marwan, J. Kurths, Nonlinear analysis of bivariate data with cross recurrence plots. Phys. Lett. A **302**, 299–307 (2002)
10. F. Popescu, F. Enache, I.C. Vizitiu, P. Ciotîrnae, Recurrence plot analysis for characterization of appliance load signature, in *10th International Conference on Communications*, Bucharest, Romania, 2014
11. H. Yang, Multiscale recurrence quantification analysis of spatial cardiac vectorcardiogram (VCG) signals. IEEE Trans. Biomed. Eng. **58**(2), 339–347 (2011)

12. Y. Chen, H. Yang, Multiscale recurrence analysis of long-term nonlinear and nonstationary time series. Chaos, Solitons Fractals **45**(7), 978–987 (2012)
13. G.M. Ramirez Avila, A. Gapelyuk, N. Marwan, H. Stepan, J. Kurths, T. Walther, N. Wessel, Classifying healthy women and preeclamptic patients from cardiovascular data using recurrence and complex network methods, Auton. Neurosci. Basic Clin. **178**(1–2), 103–110 (2013)
14. C.L. Webber Jr., J.P. Zbilut, in *Recurrence Quantification Analysis of Nonlinear Dynamical Systems*, ed. by M.A. Riley, G.C. Van Orden. Tutorials in Contemporary Nonlinear Methods for the Behavioral Sciences Web Book (National Science Foundation, U.S., 2005), pp. 26–94
15. J.P. Zbilut, C.L. Webber Jr., in *Recurrence Quantification Analysis*, ed. by M. Akay. Wiley Encyclopedia of Biomedical Engineering, Wiley, 2006
16. J. Fan, Q. Yao, *Nonlinear Time Series: Nonparametric and Parametric Methods* (Springer, New York, 2005)
17. J.C. Sprott, *Chaos and Time-Series Analysis* (Oxford University Press, New York, 2003)
18. J. Gao, Y. Cao, L. Gu, J. Harris, J. Principe, Detection of weak transitions in signal dynamics using recurrence time statistics. Phys. Lett. A **317**, 64–72 (2003)
19. J. Gao, Y. Cao, W.W. Tung, J. Hu, *Multiscale Analysis of Complex Time Series: Integration of Chaos and Random Fractal Theory, and Beyond* (Wiley, New Jersey, 2007)
20. N. Marwan, S. Schinkel, J. Kurts, Recurrence plots 25 years later—Gaining confidence in dynamic transitions. Europhys. Lett. **101**, 20007 (2013)
21. N. Marwan, M.C. Romano, M. Thiel, J. Kurths, Recurrence plots for the analysis of complex systems. Phys. Rep. **438**(5–6), 237–329 (2007)
22. H. Kantz, T. Schreiber, *Nonlinear Time Series Analysis* (University Press, Cambridge, 1997)
23. J.P. Eckmann, S.O. Kamphorst, D. Ruelle, Recurrence plots of dynamical systems. Europhys. Lett. **5**(9), 973–977 (1987)
24. F.-M. Birleanu, C. Ioana, C. Gervaise, J. Chanussot, A. Serbanescu, G. Serban, On the recurrence plot analysis method behaviour under scaling transform, in *2011 IEEE Workshop on Statistical Signal Processing (SSP2011)* (2011), pp. 789–792
25. S. Mallat, S. Zhong, Characterization of signals from multiscale edges. IEEE Trans. Pattern Anal. Mach. Intell. **14**(7), 710–732 (1992)
26. S. Mallat, *A Wavelet Tour of Signal Processing*, 2nd edn, Academic Press, 1999
27. S. Mallat, A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. IEEE Trans. Pattern Anal. Mach. Intell. **11**(7), 674–693 (1989)
28. M. Misiti, Y. Misiti, G. Oppenheim, J-M. Poggi, Wavelets Toolbox Users Guide, the MathWorks, Wavelet Toolbox, for use with MATLAB (2000)
29. C. Bernard, T. Petrut, G. Vasile, C. Ioana, Multi-lag phase space representations for transient signal characterization, in *Europeean Signal Processing Conference*, Portugal, 2014, pp. 2115–2119
30. M. Casdagli, S. Eubank, J. Doyne Farmer, J. Gibson, State space reconstruction in the presence of noise. Phys. D **51**(1–3), 52–98 (1991)
31. E.B. Wylie, L.V. Streeter, *Fluid Transients in Systems* (Prentice-Hall Inc., 1993)
32. F.E. Hachem, A.J. Schleiss, Effect of drop in pipe wall stiffness on water-hammer speed and attenuation. J. Hydraul. Res. **50**(2) (2012)
33. N.H.C. Hwang, R.J. Houghtalen, *Fundamentals of Hydraulic Engineering Systems*, 3rd edn (Prentice Hall Inc., 1996)

Chapter 3
Multi-lag Phase Diagram Analysis for Transient Signal Characterization

Cindy Bernard, Angela Digulescu, Alexandre Girard and Cornel Ioana

Abstract Phase diagram analysis is a potential technique that can offer interesting information regarding the signal shapes and eventually transient signal characterization. Indeed by choosing wisely the lag in phase diagram representations, it is possible to highlight mathematical properties such as time-shift and time-scale operators, as well as amplitude modifications. Therefore, this chapter develops the concept of multi-lag phase diagram analysis (MLPDA), as well as different methods aimed to extract parsimonious parameters from signal's phase diagrams calculated for different values of lags. By combining all of them, we are then able to explore new ways of transient signal characterization.

3.1 Introduction

In this chapter, we only consider transient signals that are characterized by sudden amplitude changes. They usually traduce mechanical, electrical, or electromagnetic phenomena that are very important to monitor. Such signal can be given by:

$$s(t) = A(t),\ t \in \left[t_0, t_0 + D\right] \tag{3.1}$$

C. Bernard (✉) · A. Digulescu · C. Ioana
Gipsa-lab, Universités Grenoble Alpes, Grenoble-INP, Grenoble, France
e-mail: cindy.bernard@gipsa-lab.fr

C. Ioana
e-mail: cornel.ioana@gipsa-lab.fr

A. Digulescu
Military Technical Academy, Bucharest, Romania
e-mail: angela.digulescu@gipsa-lab.grenoble-inp.fr

A. Girard
EDF R&D, Chatou, France
e-mail: Alexandre.girard@edf.fr

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8_3

where the duration D is assumed much shorter than the signal observation time, and t_0 the starting time of the signal. The amplitude function $A(t)$ of the signal is assumed to have high order derivatives.

We particularly focus on the case of transients observed at a given distance from the source and that consists of not only the transient at the origin (as it is generated by the phenomenon) but also of the propagation and the receiver processing effects. Thus, the deformations introduced by the propagation are of great interest to characterize the environment.

The methods proposed in our work will contribute to offer a potential interesting solution, attempting to provide a general analysis framework of such signals. To do so, we propose to investigate the signals similarities and dissimilarities through recurrence plot (RP) which has been introduced by Eckmann et al. [1] in 1987 in order to visualize recurrences of higher-dimensional phase space trajectories in nonlinear data time series. By recurrence, we refer to the return of a state of a system to a previously visited point. Recurrence Plot Analysis (RPA) has then been derived from RP and is based on three major steps: time-delay embedding (1980–1981) [2, 3], recurrence plots (1987) and recurrence quantification analysis (RQA) (1992–2002) [4, 5]. More details about the history of RPA can be found in [6] and an international website gathers all the advances and communications of the community [7].

Therefore, the main idea of this chapter is to investigate the different signals's properties in their phase diagrams in order to explore the similarities and dissimilarities between them. To do so, Sect. 3.2 first provides different mathematical properties such as time-scale transformations and amplitude modifications between transients that can be highlighted by the lag diversity in phase diagrams. Then Sect. 3.3 proposes five new descriptors for transient characterization that enables to extract the previously cited transformations. An application example is then presented in Sect. 3.4. Finally, Sect. 3.5 provides some conclusions and perspectives of work.

3.2 Mathematical Properties of Lag Diversity in Phase Diagram

In order to illustrate the concept of multi-lag PDA, a transient is modeled as a modulated cosine $s(n)$ defined for $n \in \{1, \ldots, N\}$ such as:

$$s(n) = \begin{cases} \cos\left(2\pi f_0 n\right) w(n) \text{ if } \quad n \in \left[n_0, n_0 + \Delta\right] \\ 0 \text{ otherwise} \end{cases} \tag{3.2}$$

with $w(n)$ a modulating window (such as a Hanning window), f_0 the central frequency of the modulation (we consider $f_0 = 3$ in our examples), n_0 the beginning of the transient and Δ the transient duration.

Phase space diagram analysis of signal $s(n)$, using the embedding dimension m and lag τ, corresponds to a representation T that is defined as follows:

$$\begin{array}{rcl} T : \mathbb{R}^N \mathrm{x} \mathbb{N} \mathrm{x} \mathbb{N} & \longrightarrow & \mathscr{M}_{N-(m-1)\tau,m}(\mathbb{R}) \\ (s,\tau,m) & \longmapsto & T_{m,\tau}(s) \end{array} \tag{3.3}$$

where:

$$T_{m,\tau}(s) = \begin{pmatrix} s(1) & \dots & s(1+(j-1)\tau) & \dots & s(1+(m-1)\tau) \\ s(2) & \dots & s(2+(j-1)\tau) & \dots & s(2+(m-1)\tau) \\ \vdots & & & & \\ s(i) & \dots & s(i+(j-1)\tau) & \dots & s(i+(m-1)\tau) \\ \vdots & & & & \\ s(N-(m-1)\tau) & \dots & s(N-(m-1)\tau+(j-1)\tau) & \dots & s(N) \end{pmatrix} \tag{3.4}$$

which can be summarized by:

$$T_{m,\tau}(s) = \left\{ t_{i,j} \right\}_{i=\{1,2\dots,N-(m-1)\tau\}, j=\{1,2,\dots,m\}} \tag{3.5}$$

with:

$$t_{i,j} = s(i+(j-1)\tau) \tag{3.6}$$

Each row of the T matrix corresponds to a phase space vector. Each column corresponds to the coordinates of one axis of the phase space.

Let us now investigate the properties of this representation to the time-shift operator, the scale operator and amplitudes changes. We define three signals $s_1(n)$, $s_2(n)$ and $s_3(n)$ such as:

$$s_1(n) = s(n+\delta) \tag{3.7}$$

$$s_2(n) = s(\alpha n) \tag{3.8}$$

$$s_3(n) = \beta s(n) \tag{3.9}$$

with δ the time-shift delay, $\alpha \in \mathbb{R}^+$ the dilation coefficient and β an amplitude modification coefficient. We propose to illustrate the different properties with equations and numerical examples that are shown in Fig. 3.1. The numerical examples are computed with the following parameters: $\delta = 55$, $\alpha = 2$ and $\beta = 1.5$.

Investigation of the Time-Shift Invariance Property

We can start by considering the phase space vector of $s_1(n)$ at instant n and the relation given by (3.7). We have:

$$\begin{array}{rl} \left[s_1(n), s_1(n+\tau), \dots, s_1(n+(m-1)\tau) \right] & = \left[s(n+\delta), s(n+\delta+\tau), \dots, s(n+\delta+(m-1)\tau) \right] \\ & = \left[s(n_0)\, s(n_0+\tau), \dots, s(n_0+(m-1)\tau) \right] \end{array} \tag{3.10}$$

This result states that even **if two identical signals are time-shifted, their phase space diagrams are invariant**. Mathematically speaking, it means that the rows of $T_{m,\tau}(s_1)$ are the same as $T_{m,\tau}(s)$'s but they suffered the circular permutation and are

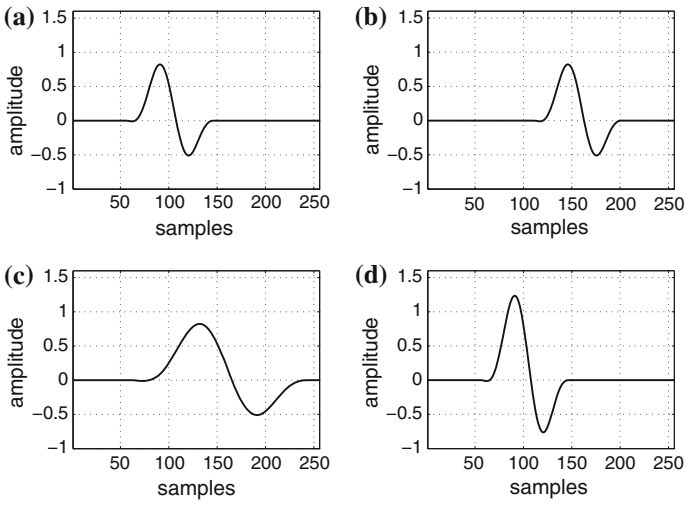

Fig. 3.1 Temporal data of: **a** $s(n)$, **b** $s_1(n)$, **c** $s_2(n)$ and **d** $s_3(n)$

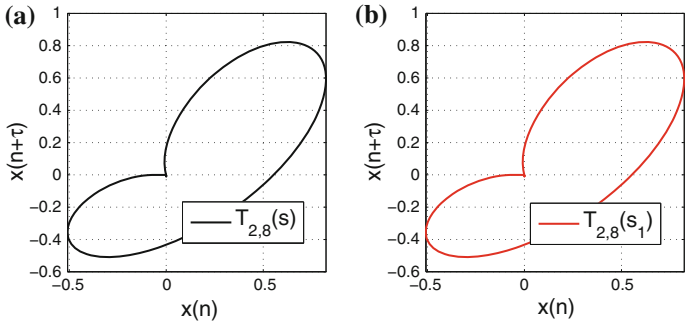

Fig. 3.2 **a** $T_{2,4}(s)$ and **b** $T_{2,4}(s_1)$ phase space diagrams. The trajectories are the same despite the time-shift

shifted compared to $T_{m,\tau}(s)$. This property is illustrated by Fig. 3.2 where we can see that for $m = 2$ and $\tau = 8$ the phase space diagrams of the two signals are exactly the same despite the time-shift between the signals.

Investigation of the Time-Scaling Property

We now consider $s_1(n)$'s phase space vector at instant n and the relation given by (3.8). We have:

$$\begin{aligned}\left[s_2(n), s_2\left(n+\tau\right), \ldots, s_2\left(n+(m-1)\,\tau\right)\right] &= \left[s\left(\alpha n+n_0\right), \ldots, s\left(\alpha\left(n+(m-1)\,\tau\right)+n_0\right)\right] \\ &= \left[s\left(n_1\right), s\left(n_1+\alpha\tau\right), \ldots, s\left(n_1+(m-1)\,\alpha\tau\right)\right] \\ &= \left[s\left(n_1\right), s\left(n_1+\tau_0\right), \ldots, s\left(n_1+(m-1)\,\tau_0\right)\right]\end{aligned} \tag{3.11}$$

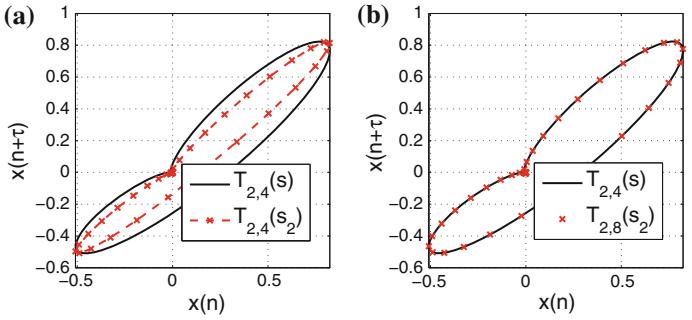

Fig. 3.3 Phase space diagrams of $s(n)$ and $s_2(n)$ for $m = 2$ and different sets of lags: **a** [4, 4] and **b** [4, 8]. The dilation connection between the two signals can be enlightened by using the appropriate set of lags

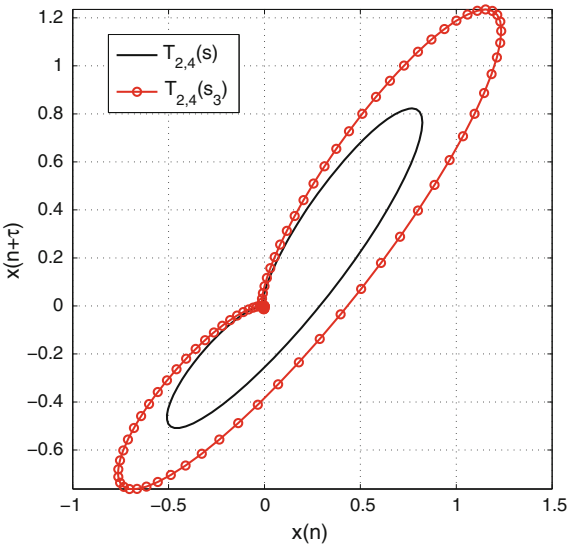

Fig. 3.4 $T_{2,4}(s_3)$ and $T_{2,4}(s)$ phase space diagrams. They are related by a shape-invariant scale transformation

This result states that **for a given** m **there exists many sets of lags** $[\tau, \alpha\tau]$ **that enable an invariance of phase space diagrams.** It means that it is possible to identify two signals that are related by a dilation. Figure 3.3 presents the phase space diagrams of $s(n)$ and $s_2(n)$ for two differents sets of lags: [4, 4] and [4, 8]. For the first set, the two diagrams do not overlap, whereas the second set enables a perfect superposition of the two diagrams.

Investigation of the Amplitude Coefficient Modification

Finally, we consider $s_3(n)$'s phase space vector at instant n and the relation given by (3.9). We have:

$$\begin{aligned}\left[s_3(n), s_3(n+\tau), \ldots, s_3(n+(m-1)\tau)\right] &= [\beta s(n), \beta s(n+\tau), \ldots, \beta s(n+(m-1)\tau)] \\ &= \beta\,[s(n), s(n+\tau), \ldots, s(n+(m-1)\tau)]\end{aligned} \tag{3.12}$$

which can be summarized by:

$$T_{m,\tau}(s_3) = \beta T_{m,\tau}(s) \tag{3.13}$$

This result shows that **the amplitude change is equivalent, in the phase space diagram representation, with a shape-invariant scale transformation.** This is illustrated by Fig. 3.4.

3.3 Multi-lag Phase Diagram Analysis

The previous section has shown that phase diagram representations can potentially be invariant to the main transforms of signals such as: time-shift, time-scale changes, amplitude modification, etc. The key point of the multi-lag based representation is that the invariance can be controlled by the lag choice, which makes possible a better exploration of the analyzed signals.

In this section, we propose different descriptors to extract parsimonious parameters from each diagram acquired for a given lag. The evolution of these descriptors with respect to the lags is then explored conducting to new representation tools for transients.

In this Chapter, we restrained ourselves to $m = 2$ in order to visualize the results but the work can be extended to higher embedding dimension. We also note $y(n)$ for $x(n+\tau)$ in order to simplify the notations.

3.3.1 Ellipse Modeling

Generally, signals can have various trajectories with different shapes, but in this subsection, we propose a general model for **approximation of trajectories based on ellipse shape** (Fig. 3.5). This choice is quite natural since the ellipse model is specific to harmonic signals. Therefore, the model is simple and enables to extract three parameters per phase space diagrams:

- the polar angle θ of the ellipse assuming that θ is the angle between the first axis and the major semi-axis
- the major semi-axis a
- the minor semi-axis b

We assume the center of the ellipse being the center of phase space diagrams as transients can be considered as zero-mean signals.

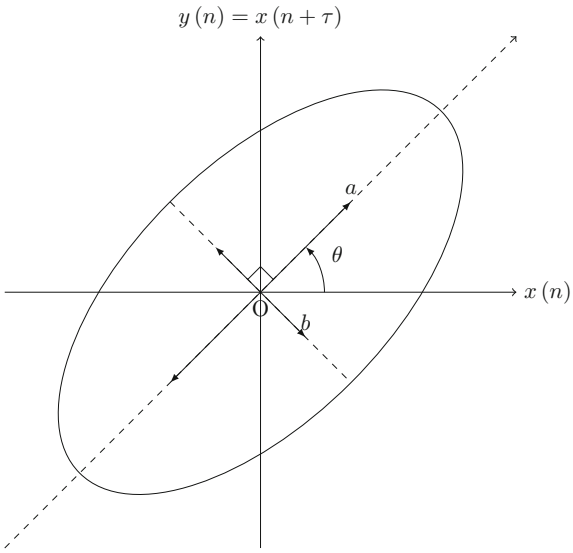

Fig. 3.5 Ellipse modeling: major and minor axis are determined by an SVD on phase space diagram, and a least square fitting method is used to calculate a and b

The first step consists in performing a singular value decomposition (SVD) of the phase space diagram to calculate its eigenvectors. Those are of great interest as they define a new basis that reflects the distribution of the data and also correspond to the major and minor axis of the ellipse model. Then, a least squares fitting method is performed to determine the ellipse that satisfies the following equation in the new basis defined by the eigenvectors:

$$Ax^2 + By^2 = 1 \tag{3.14}$$

where the semi-major and semi-minor axis are given by:

$$\begin{cases} a = 1/\sqrt{A} \\ b = 1/\sqrt{B} \end{cases} \tag{3.15}$$

At this point of the study, each diagram is modeled by an ellipse with these three parameters $[a, b, \boldsymbol{\theta}]$. The evolution of $\boldsymbol{\theta}$ enables to estimate an apparent periodicity of the transient, while the evolution of a and b enables to know how the data is distributed in the phase space. If they are scattered over a large area, it means that the lag used to construct the representation is not representative of the transient construction. On the other hand, if they are rather confined into a smaller area, it means that the lag is representative and is well adapted to the study of this transient.

In order to illustrate the method, let consider three signals $s_1(n)$, $s_2(n)$ and $s_3(n)$ that are modulated cosines. $s_1(n)$ and $s_2(n)$ are related by a time-scale transformation while $s_3(n)$ is the result of a 10-th low-pass FIR digital filtering applied on $s_1(n)$ with the normalized cutoff frequency of 0.1. Temporal signals are shown in Fig. 3.6. At

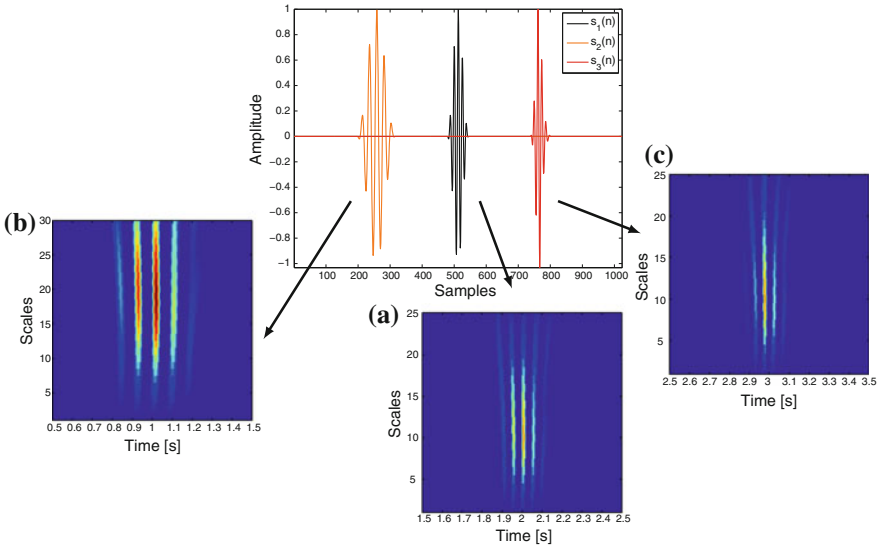

Fig. 3.6 Temporal signals studied in this section and its wavelet transforms using the Daubechies mother wavelet

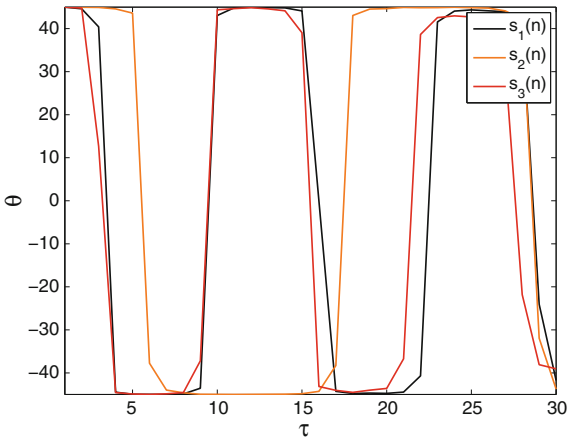

Fig. 3.7 Evolution of θ with respect to τ for the three studied signals

first look, it is quite difficult to tell appart $s_1(n)$ from $s_3(n)$ and a time-scale analysis would not do better as their frequency contents are really close (Fig. 3.6).

For the three signals and $\tau \in \{1, \ldots, 30\}$, we model each phase diagram by an ellipse and record the variation of a, b and θ with respect to the lags. As we can see in Figs. 3.7 and 3.8, the 3 parameters present an apparent periodicity that are related to the apparent periodicities of the signals. We call *apparent periodicity the number of samples between two successive zero-crossing*. As an example, the number of lags between two successive maxima of θ is equal to 13 for $s_1(n)$, 23 for $s_2(n)$ and 12 for

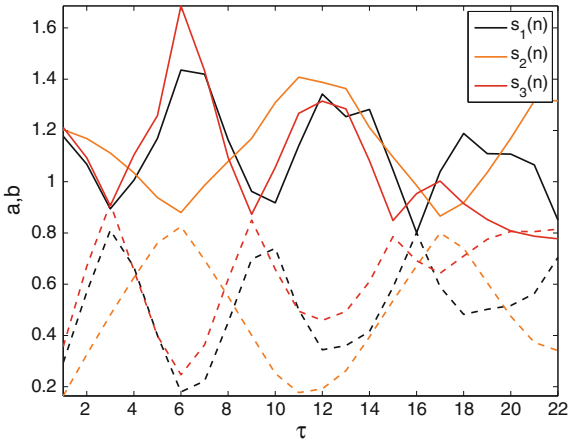

Fig. 3.8 Evolution of a (*continuous line*) and b (*dashed line*) with respect to τ for the 3 studied signals

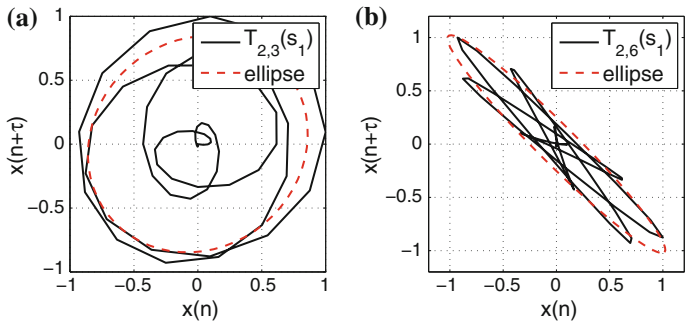

Fig. 3.9 **a** $T_{2,3}\left(s_1\right)$ phase space diagram and its associated ellipse modeling. **b** $T_{2,6}\left(s_1\right)$ phase space diagram and its associated ellipse modeling

$s_3(n)$, while the apparent periodicity is of 12.5 samples for $s_1(n)$, 22.25 for $s_2(n)$ and 11.75 for $s_3(n)$ which is coherent with the previous values.

The periodicity of a and b corresponds to the half period of the signals (Fig. 3.8). Their maximal values correspond to phase diagrams that can be modeled by circles (same value for both parameters), meaning that the phase diagrams are more scattered in the phase space. This can be seen in Fig. 3.9 where we plotted $s_1(n)$'s phase diagrams for $m = 2$ and $\tau = \{3, 6\}$. For $\tau = 3$ that enables to obtain a maximal value for b, the phase diagram can be modeled by a circle and is well distributed in the space, while for $\tau = 6$, the phase diagram is concentrated into a smaller area.

We have seen that with the ellipse modeling it was possible to summarize each representation by only three parameters and estimate the apparent frequency of a transient. We can also explore the distribution of the diagram in the phase space which is of great interest as it enables to highlight the lags that provide a great concentration of the data or on the contrary a dispersal.

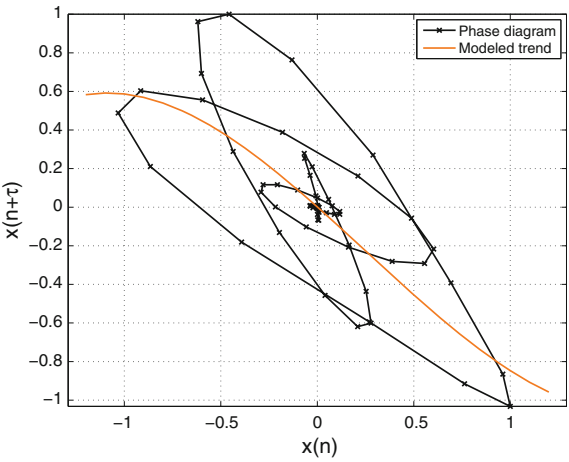

Fig. 3.10 Phase diagram's trend is modeled by a third degree polynomial

3.3.2 Trend Modeling

In the previous subsection, we have seen that phase diagrams were more or less concentrated around a line that tends to rotate around the origin of the phase space (due to the zero-mean of the modeled signals). This is why, we want to model this line as a third degree polynomial (Fig. 3.10) in order to quantify the rotation and the natural trend of the diagram [8]. The model is defined as follows:

$$y = \hat{a}x^3 + \hat{b}x^2 + \hat{c}x + \hat{d} \tag{3.16}$$

To do so, we consider the diagram as a scatterplot and perform a least square fitting estimation by minimizing the following sum:

$$\underset{\hat{a},\hat{b},\hat{c},\hat{d}}{Argmin} \sum_{i=1}^{N} \left(s\left(i+\tau\right) - \left(\hat{a}s^3\left(i\right) + \hat{b}s^2\left(i\right) + \hat{c}s\left(i\right) + \hat{d} \right) \right) \tag{3.17}$$

where $s\left(i\right)$ corresponds to the analyzed signal for $i \in \{1, \ldots, N\}$; N being the number of samples.

Therefore, for each representation, phase diagrams are summarized by four parameters $\left\{ \hat{a}_\tau, \hat{b}_\tau, \hat{c}_\tau, \hat{d}_\tau \right\}$ that vary with respect to τ. As studied signals can always be considered as zeros mean, we can remove from consideration $\hat{d}$: this parameter will always be equal to zero. Thus, three parameters remain: $\hat{a}$, $\hat{b}$ and $\hat{c}$. They enable to discriminate transients by looking at their evolutions with respect to τ and they also permit to highlight similitudes by looking at one parameter with respect to another. This last representation allows to get rid of the evolution of τ and investigate if whether or not the diagrams have similar trends for different lags: this can reflect a time-scaling operation.

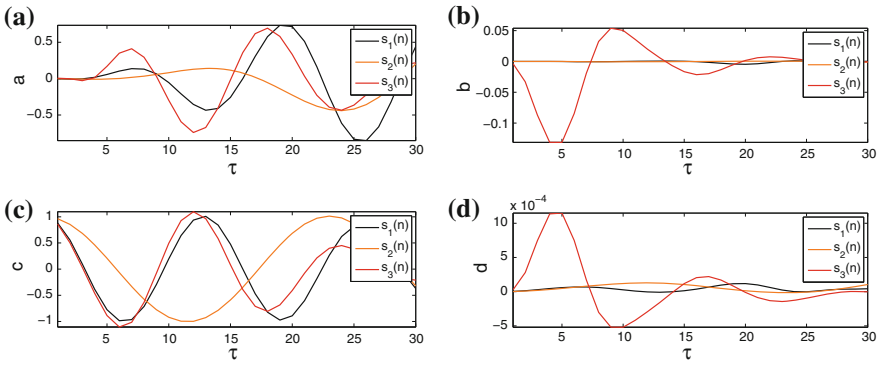

Fig. 3.11 Evolution of **a** $\widehat{a}$, **b** $\widehat{b}$, **c** $\widehat{c}$ and **d** $\widehat{d}$ with respect to τ

To illustrate the concept of trend modeling, we consider the three signals that were introduced in the previous subsection. For all of them and $\tau = \{1, \ldots, 30\}$, we model each phase diagram's trend by a third degree polynomial and record the variations of the four parameters with respect to the lag. As we can see in Fig. 3.11, it is quite easy to discriminate between $s_1(n)$ and $s_3(n)$ by looking at the evolution of the parameters of interest. The apparent periodicity of the parameters corresponds to the apparent periodicity of the signals. As an example, the number of lags between two successive maxima of $\widehat{c}$ is equal to 12 for $s_1(n)$, 22 for $s_2(n)$ and 11 for $s_3(n)$, while the apparent periodicity is of 12.5 samples for $s_1(n)$, 22.25 for $s_2(n)$ and 11.75 for $s_3(n)$ which is coherent with the theoretical values. We can also notice that the evolution of $\widehat{d}$'s is very small (below 0.0001), as well as for the evolution of parameter $\widehat{b}$. This is why we remove this last parameter from consideration as well.

Figure 3.12 presents the evolution of $\widehat{c}$ with respect to the evolution of $\widehat{a}$. This representation is interesting as we get rid of the evolution of τ. It enables to highlight signals that would have the same phase diagram's trends for different values of lags.

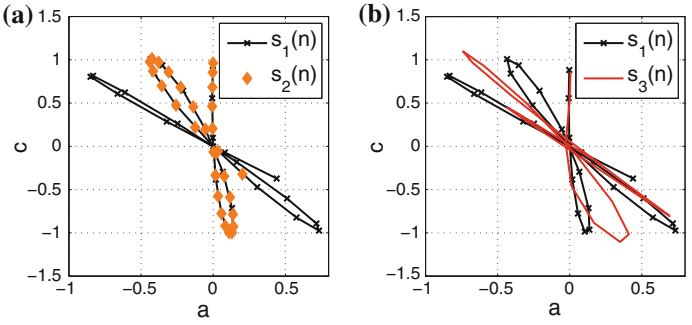

Fig. 3.12 Evolution of $\widehat{c}$ with respect to $\widehat{a}$ for the 3 studied signals

This is the case for $s_1(n)$ and $s_2(n)$ where the evolution of $[\widehat{a}, \widehat{c}]$ for both signals overlap in the representation. This is coherent as the two signals are related by a time-scale operator. As $s_1(n)$ and $s_3(n)$ are related by a low-pass filtering, we can see that even if their representations are similar, they do not overlap. This shows the complexity introduced by the filtering.

The modeling of phase diagrams's trend by a third degree polynomial is of great interest as it enables to detect if transients are related by a time-scale transformation. However, at this point of the study, there is no method that tells us what couples of lags we should use to highlight this transformation. This is why we moved toward **a matching phase diagram technique** that is presented in the next subsection.

3.3.3 Extremum Points/Bounding Box

Generally, when we talk about SNR for transient analysis, we only consider signal and noise over the duration of the transient. Thus, the SNR is defined as follows:

$$\mathbf{SNR}_{dB} = 10 \log_{10} \frac{\sum_n s(n)}{\sum_n b(n)} \tag{3.18}$$

where $s(n)$ corresponds to the noise-free transient and $b(n)$ to the noise, defined for $n \in \Delta$ with Δ being the time interval where is defined the transient of interest.

When it comes to noise in phase diagram, we can see in Fig. 3.13 that the external contour remains more or less the same depending on the level of noise considered. This is the reason why we focus on the bounding box of the trajectory that is delimited by the maximal and minimal values of the studied signal.

We thus define 4 remarkables coordinates defined as follows and illustrated in Fig. 3.14:

$$A : \begin{cases} \forall n, \quad x(n_1) = \max\limits_n (x(n)) \\ y(n_1) \end{cases} \tag{3.19}$$

$$B : \begin{cases} x(n_2) \\ \forall n, \quad y(n_2) = \max\limits_n (y(n)) \end{cases} \tag{3.20}$$

$$C : \begin{cases} \forall n, \quad x(n_3) = \min\limits_n (x(n)) \\ y(n_3) \end{cases} \tag{3.21}$$

$$D : \begin{cases} x(n_4) \\ \forall n, \quad y(n_4) = \min\limits_n (y(n)) \end{cases} \tag{3.22}$$

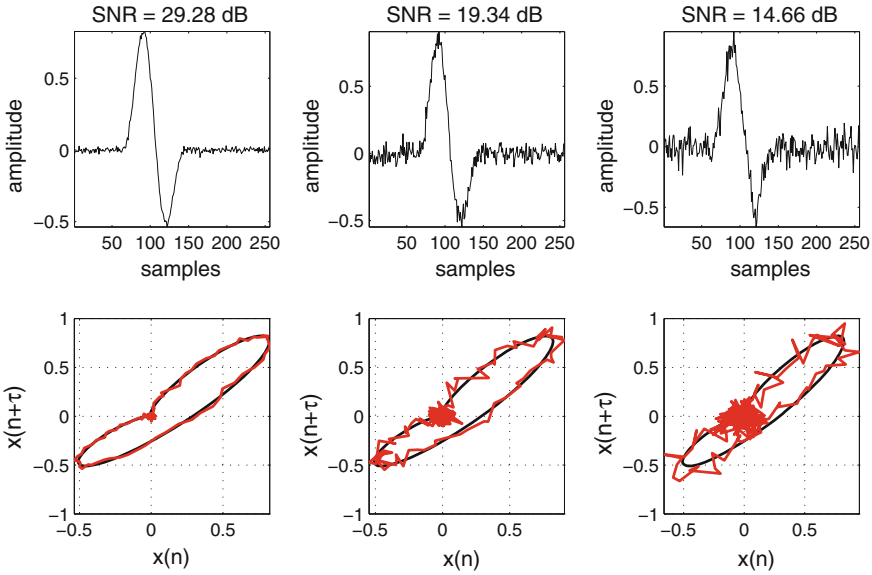

Fig. 3.13 Noisy transients and their respective phase diagrams for different level of noise

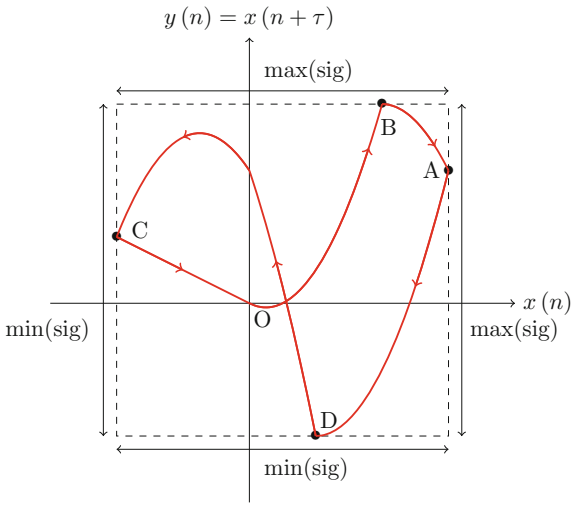

Fig. 3.14 Bounding box that confines transient trajectories are delimited by the maximal and minimal values of transients

In order to compare two transient signals analysis, we consider the signals' phase space diagrams for different values of lag $\left[\tau_{1,i}, \tau_{2,i}\right]_{i\in[1,\ldots,\tau_{max}]}$ and we look for their extremum points $\left[A_{k,i}, B_{k,i}, C_{k,i}, D_{k,i}\right]_{k=1,2}$. Then, we compute 4 matrices H_A, H_B, H_C and H_D defined as follows:

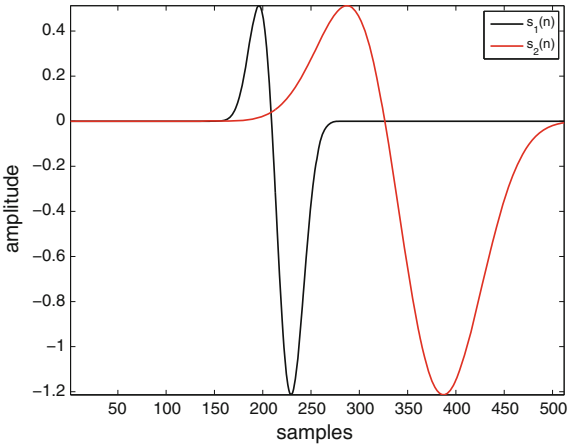

Fig. 3.15 Temporal data of the studied signals

$$H_X = \left\{h_{X,i,j}\right\}_{i,j\in[1,\ldots,\tau_{max}]} \tag{3.23}$$

$$= \left\{\mathscr{D}\left(X_{1,i}, X_{2,j}\right)\right\} \tag{3.24}$$

where $\mathscr{D}$ describes a given metric and X the extremum point taken into consideration.

Each matrix provides a "map" of the distances between the extremum points of each phase space diagrams calculated for different values of lags. Therefore, it enables to discover which couples of lags $\left[\tau_1, \tau_2\right]$ need to be used to provide a match between two extremum points.

In order to illustrate this concept, let consider two transients related by a time-scale transformation with the dilation coefficient $\alpha = 3$. Temporal data presented in Fig. 3.15 shows that $s_2(n)$ (in red) is only dilated compared to $s_1(n)$ (in black) and does not suffer amplitude changes. In this part, the purpose is to highlight the time-scale relation between these signals.

We first calculate the 4 matrices described previously using the Euclidean norm $\mathscr{D}_2$ defined as follows:

$$\mathscr{D}_2\left(x\right) = \left\|X_{1,i} - X_{2,j}\right\|_2 \tag{3.25}$$

Those 4 matrices are displayed in Fig. 3.16 with the same colorbar. In this example, we can see that the error is always really small for the extremum point B contrary to the other points. Nevertheless, they all have in common a 'line' representing the set of lags where the error is minimum. Those 'lines' are shown in Fig. 3.17 and we can notice that they overlap. They describe the sets of lags $\left[\tau_1, \tau_2\right]$ to use that would guaranty a match of the 4 extremum points for the two signals. Thus, we can deduce a relationship between the sets of lags by performing a linear regression. We obtain that:

$$\tau_2 = 3\tau_1 \tag{3.26}$$

which is consistent with the dilation coefficient.

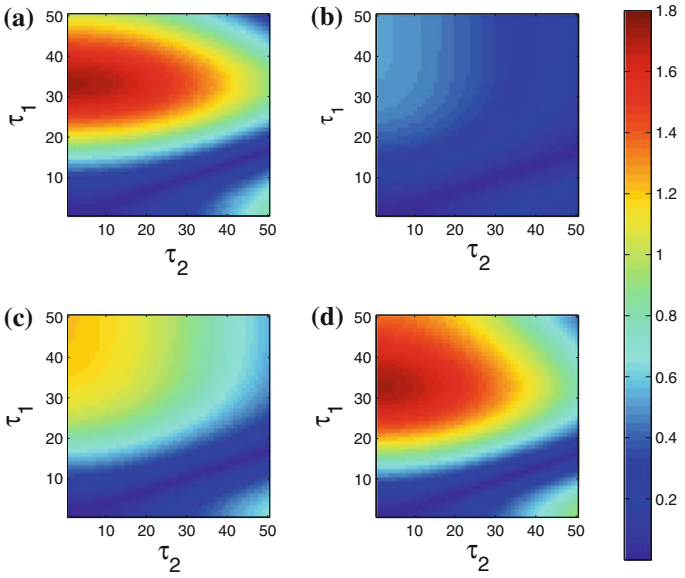

Fig. 3.16 The four matrices displaying the distance between extremum points of $s_1(n)$ and $s_2(n)$'s phase space diagrams for different values of lags τ_1 and τ_2: **a** H_A, **b** H_B, **c** H_C and **d** H_D

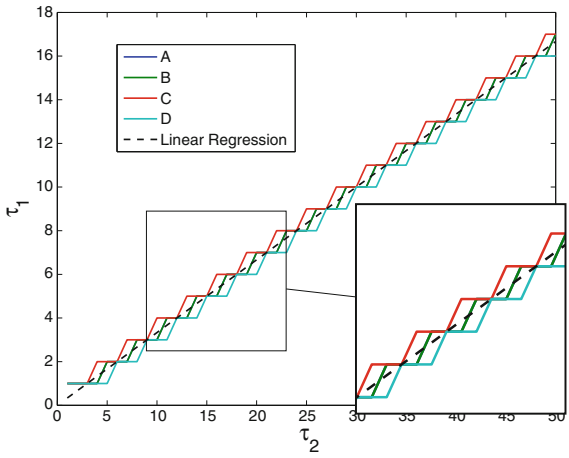

Fig. 3.17 Sets of lags $[\tau_1, \tau_2]$ that provide matches for the extremum points

To validate this result, we draw $s_1(n)$ and $s_2(n)$'s phase space diagrams by using the set of lags $[\tau_1 = 3, \tau_2 = 9]$ (Fig. 3.18). As a matter of fact both phase space diagrams overlap perfectly.

This technique allows to highlight time-scale transformations. However, even if this technique enables to match extremum points having the same coordinates for certain couples of lags, we need to keep in mind that trajectories can still be different

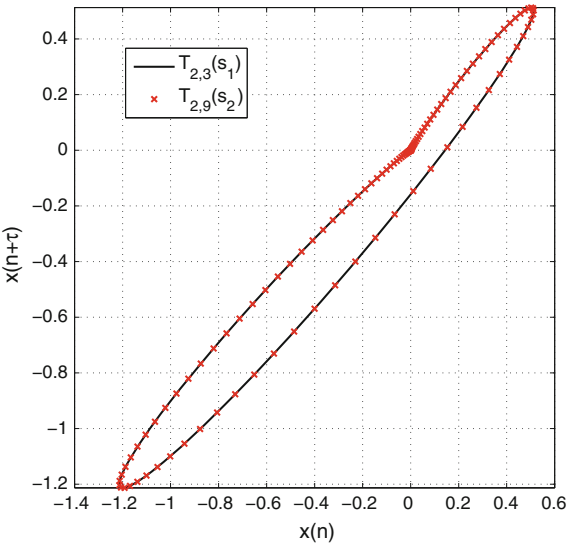

Fig. 3.18 Phase space diagrams of $T_{2,3}(s_1)$ and $T_{2,9}\left(s_2\right)$ respectively in *black* and *red*

as the other coordinates are not considered. This is why it is interesting to monitor *the area covered by the trajectory in the phase space* which is the subject of the next subsection.

3.3.4 Area Calculation

Previous subsection has shown that it was possible to define four reference coordinates for each phase space diagrams that allow to find matching correspondances, however, it is not enough to guaranty the invariance. The next idea consists in calculating the area of the diagram to quantify the surface. To do so, the diagram is considered as a curve having at each instant n polar coordinates $\left[\rho_{\tau_k}(n), \theta_{\tau_k}(n)\right]$. We then calculate the area $\mathscr{A}\left[s, \tau_k\right]$ covered by the phase space diagram calculated for $\tau = \tau_k$:

$$\mathscr{A}\left[s, \tau_k\right] = \int_{\theta_{\tau_k}(1)}^{\theta_{\tau_k}(N)} \rho_{\tau_k}^2(n) \left| d\theta_{\tau_k}(n) \right| \tag{3.27}$$

We also define the matrix $A\left[s_1, s_2\right]$ defined as follows that enables to calculate the distance between two signals s_1 and s_2's phase space diagram areas computed for different sets of lags $\left[\tau_1, \tau_2\right]$:

$$A\left[s_1, s_2\right] = \left\{a_{i,j}\right\}_{i,j \in [1,\ldots,\tau_{max}]} \tag{3.28}$$

$$= \left\{ \left| \mathscr{A}\left[s_1, \tau_i\right] - \mathscr{A}\left[s_2, \tau_j\right] \right| \right\} \tag{3.29}$$

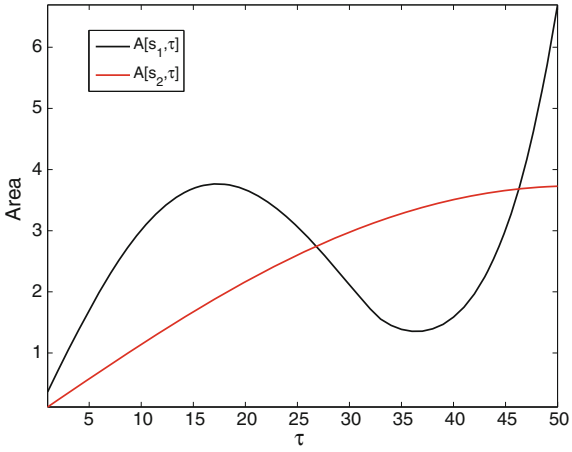

Fig. 3.19 Evolution of the two signals area according to the lag used to compute their phase space trajectories

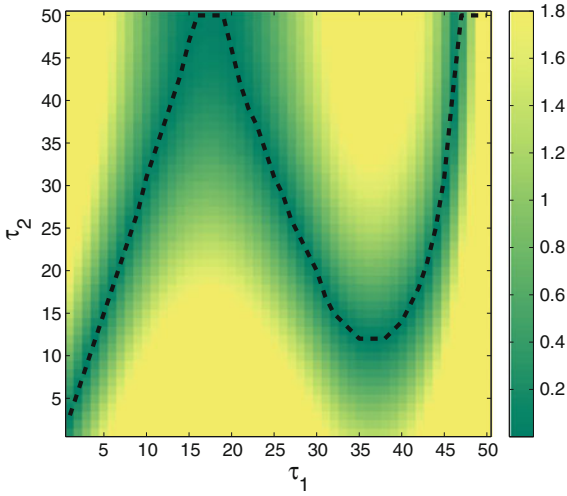

Fig. 3.20 The matrix $A\left[s_1, s_2\right]$ enables to calculate the distance between two signals s_1 and s_2's phase space diagram areas computed for different sets of lags $\left[\tau_1, \tau_2\right]$

It permits to find the sets of lags that offer similar areas for two phase space diagrams. This analysis can be complementary to the bounding box method to highlight time-scale transformation for instance.

In order to illustrate this concept, let consider the example presented previously that only presents a time-scale transformation. We first compute the area for $s_1(n)$ and $s_2(n)$ using $\tau = 1, \ldots, 50$. The evolution of the area according to the lag is displayed for both signals in Fig. 3.19. As we can see, $\mathscr{A}\left[s_1, \tau\right]$ presents a clear maxima for $\tau = 17$, while $\mathscr{A}\left[s_2, \tau\right]$ does not have one. A maxima notifies a change on phase space diagrams that can as an example change its first eigenvector.

The computation of matrix $A\left[s_1, s_2\right]$ provides the sets of lags that offer matching areas. They are highlighted by a black dashed line in Fig. 3.20. For $\tau_1 = 1, \ldots, 17$, we can see that this line is also a straight line that verifies the following equation:

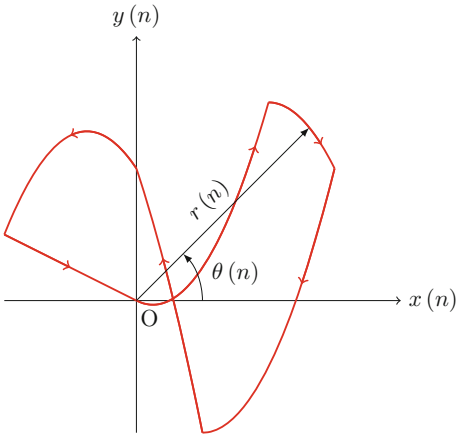

Fig. 3.21 Trajectory expressed in polar coordinate system. Each coordinates is represented by a radius and a polar angle

$$\tau_2 = 3\tau_1 \tag{3.30}$$

This relationship confirms the dilation coefficient that has been used for the signals simulation.

This method is really useful when coupled with the bounding box method. They enables to highlight time-scale transformations between signals.

3.3.5 Polar Coordinates Analysis

We have shown in Sect. 3.2 that amplitude changes can be enlightened from phase space diagrams. Indeed there exists a scale factor between two signals having an amplitude relationship. To highlight it, phase diagrams are turned into polar coordinates as shown in Fig. 3.21. The assessment is that for a given polar angle, radius are directly connected by the amplitude coefficient.

Let consider $s_1(n)$ defined as in (3.2) and $s_2(n)$ defined as follows:

$$s_2(n) = \frac{1}{\beta} s_1(\alpha n) \tag{3.31}$$

with $\alpha = 2$ and $\beta = 1.4$, as an example. That is, s_2 is derived from s_1 by a double scale and amplitude modification.

Figures 3.15 and 3.22 present temporal data and phase diagrams obtained for $m = 2$ and respectively $\tau = 8$ and $\tau = 16$. Using this set of parameters, we know that the phase diagrams would be superposed if the signals's amplitudes were identical. However due to this difference they present a shape-invariant scale transformation.

We turn the cartesian coordinates into polar coordinates and draw the functions $\rho(n) = f(\theta(n))$ for the two signals that are shown in Fig. 3.23. As we can see, both

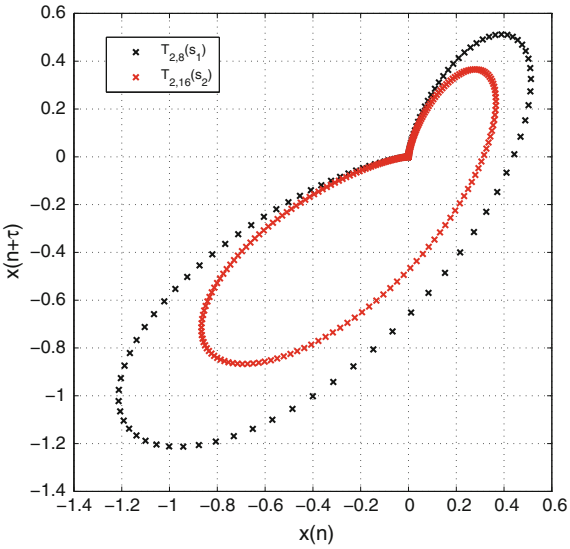

Fig. 3.22 Phase space diagrams of $T_{2,8}(s_1)$ and $T_{2,16}(s_2)$ respectively in *black* and *red*

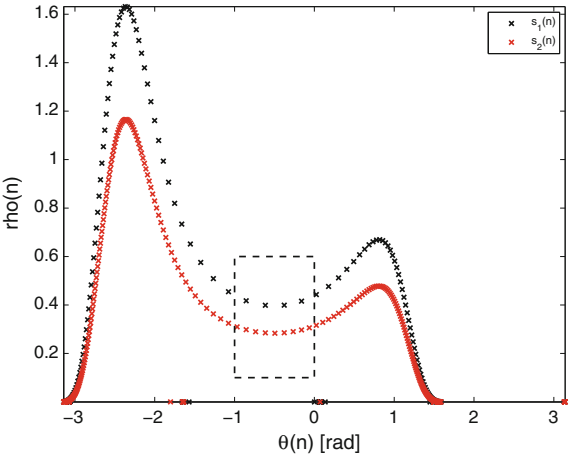

Fig. 3.23 $\rho(n) = f(\boldsymbol{\theta}(n))$ for $T_{2,8}(s_1)$ and $T_{2,16}(s_2)$ respectively in *black* and *red*

curves are similar and it is easy to imagine a linear relationship between them. However, for a given polar angle, there is not always a correspondance between the two curves (Fig. 3.24). This is the reason why we propose an algorithm to match corresponding coordinates with respect to polar angles.

Algorithm description:

For each polar angle $\boldsymbol{\theta}_1(n)$ of the first curve, the algorithm searches for the corresponding polar angle on the second curve $\boldsymbol{\theta}_2(n)$. A match is enlightened if the two conditions are met:

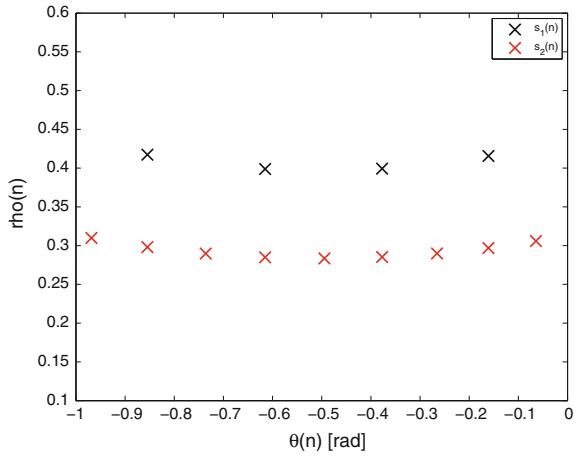

Fig. 3.24 Zoom in of Fig. 3.23

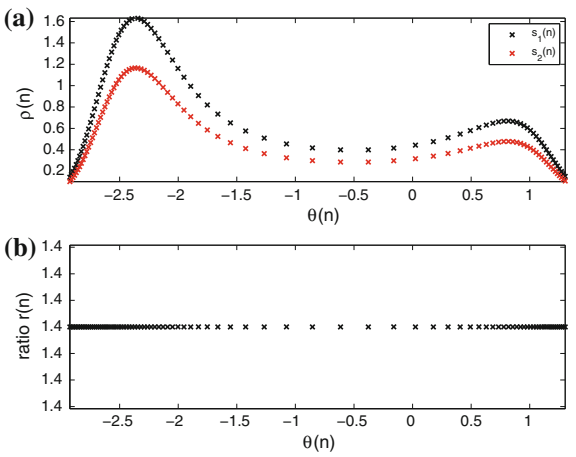

Fig. 3.25 **a** Pairs of coordinates that have been selected by the algorithm to calculate the amplitude modification known as $r(n)$ (**b**)

$$\begin{cases} \theta_2\left(n_1\right) = argmin\left|\theta_1\left(n_0\right) - \theta_2\left(n\right)\right| \\ \theta_2\left(n_1\right) < \epsilon \end{cases} \tag{3.32}$$

where n_0 is the index of $\theta_1\left(n\right)$ we are looking for, n_1 is the corresponding index for $\theta_2\left(n\right)$ and ϵ a threshold error that enables to discard certain associations that are not consistent.

The algorithm selects M pairs of coordinates and calculates a ratio $r(i)$ defined for $i \in \{1, \ldots, M\}$ such as:

$$r(i) = \frac{\theta_1\left(n_{0_i}\right)}{\theta_2\left(n_{1_i}\right)} \tag{3.33}$$

Figure 3.25 presents the pairs of coordinates that have been selected by the algorithm and the associated ratio. In this example, we can see that the ratio is consistent along the entire polar angles interval which corroborates the idea of a shape-invariant scale transformation. Moreover, the ratio is here equal to 1.4 which is the value of β.

We have seen in this example that it was possible to estimate a shape-invariant scale transformation. The same work can be done along an interval of polar angles that would highlight amplitudes changes on this interval.

3.4 Application Example

In order to characterize transient modifications due to its propagation through a cable, we now conduct an experiment using the facilities existing in our lab. We generate a partial discharge (PD) on an electric cable whose ends T_1 and T_2 are connected to a data acquisition system. The PD source is respectively localized at L_1 and L_2 distances from the recording devices, as presented in Fig. 3.26. The main idea is to estimate the relative propagation distances by comparing the relative deformations between the recorded signals. We note $s_0(n)$ the generated PD, $s_1(n)$ the signal recorded at T_1 and $s_2(n)$ the signal recorded at T_2.

Time representation of emitted and recorded signals are presented in Fig. 3.27. As we can notice, signals recorded after propagation through the cable reels differ from the emitted one due to the propagation through the medium. They present non-linear dilation and amplitude changes that are more or less stronger depending on the propagation range. Temporal changes are shown in Fig. 3.28 where studied signals have been superposed and normalized.

What is interesting with this kind of signals is that we can easily visualize that a time-scale analysis would not be enough to characterize the modifications caused by the propagation. There are dilation and also a modification of the envelop of the signal that would be difficult to analyze with a single type of mother wavelet function. This is shown by Fig. 3.29 where a time-scale study is performed for the three signals using the Symlet mother wavelet. The results highly depend on the choice of mother wavelet used and for the three cases, even if it is easy to detect them, it is quite difficult to claim that they come from the same source as waveforms changes with the propagation. This is the reason why it is useful to move forward MLPDA which is presented next.

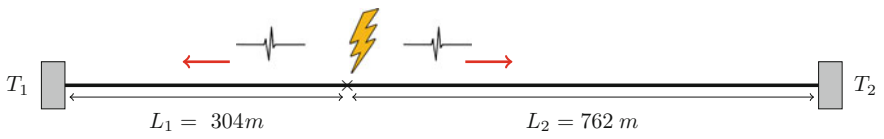

Fig. 3.26 Experimental outline

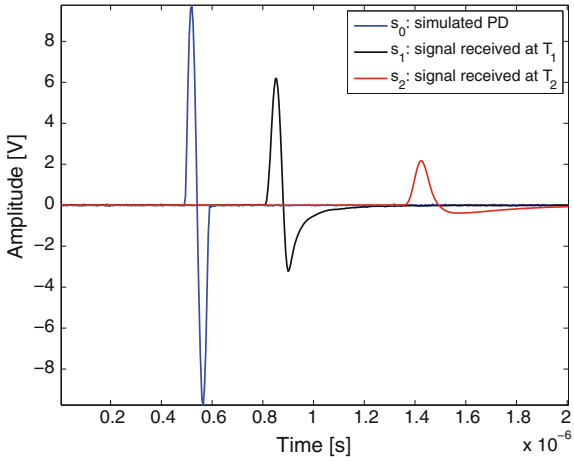

Fig. 3.27 Time representation of emitted signal and recorded signals after propagation through the cable reels

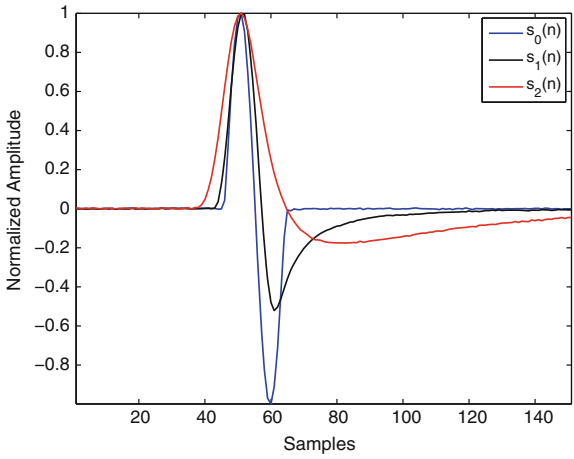

Fig. 3.28 Time representation of the superposition of normalized signals that highlights dilation and amplitude changes

To begin with, we need to normalize the signals by their maximal values, then compute their phase diagrams for $\tau = 1, \ldots, 20$, search for the 4 extremum points A, B, C and D, and finally compute the distance matrices H_A, H_B, H_C and H_D as presented in Sect. 3.3.3. We then look for the sets of lags $[\tau_1, \tau_2]$ that minimize the distances between each extremum points.

Thus, after the computation, we can see that it is not possible to superpose the different C and D coordinates, which is coherent with the normalization of the signals by their maximum values. On the contrary, we can superpose the A and B coordinates. To illustrate this, we compute $T_{2,4}(s_1)$ and $T_{2,6}(s_2)$ shown in Fig. 3.30, and $T_{2,5}(s_1)$ and $T_{2,8}(s_1)$ shown in Fig. 3.31. We can see that the top right parts of the two trajectories superpose well in the first case compared to the other one, even if the distances between A_1 and A_2, and B_1 and B_2 are smaller in the second case. This result confirms that the bounding box method cannot be used alone to determine dila-

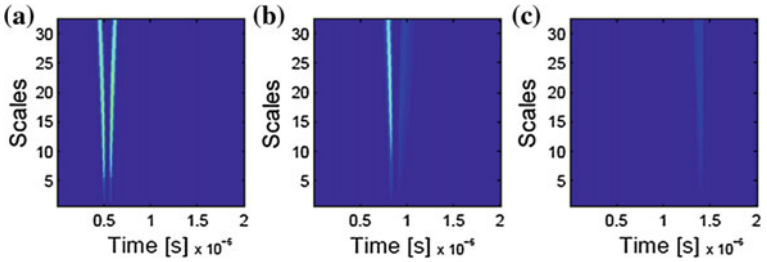

Fig. 3.29 Time-scale study using the symlet mother wavelet for **a** $s_0(n)$, **b** $s_1(n)$ and **c** $s_2(n)$

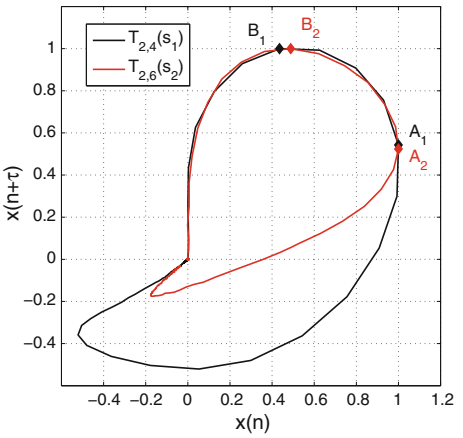

Fig. 3.30 Superposition of $T_{2,4}(s_1)$ and $T_{2,6}(s_2)$ phase space diagrams. Even if the distances $d(A_1, A_2)$ and $d(B_1, B_2)$ are not minimal, the top right parts of the diagrams superpose well

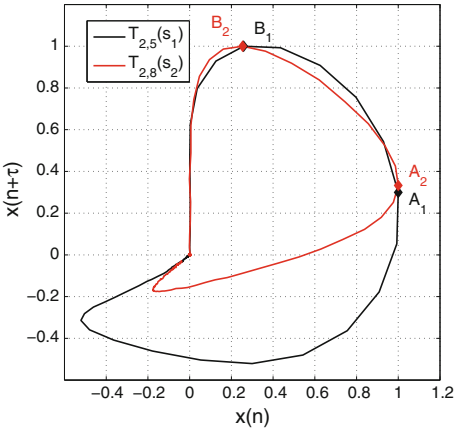

Fig. 3.31 Superposition of $T_{2,5}(s_1)$ and $T_{2,8}(s_2)$ phase space diagrams. The distances $d(A_1, A_2)$ and $d(B_1, B_2)$ are smaller than in the first case but the diagrams do not superpose well

tion relationship. It needs to be coupled with a verification of the trajectories and/or other methods. Nevertheless, according to Fig. 3.30, we can conclude that there is a dilation coefficient equal to $\frac{\tau_2}{\tau_1} = 3/2$ between the positive parts of $s_1(n)$ and $s_2(n)$.

In order to quantify the time-scale coefficients obtained with MLPDA tool, we propose to define an estimated dilation coefficient obtained as follows. For the normalized signal's parts of interest, we calculate the number of samples where the signal's amplitude is higher than half the maximal value, i.e. 0.5 and then compute the ratio between the two numbers to obtain the estimated dilation coefficient. For positive part, we obtain a ratio τ_1/τ_2 equal to $8/12$, i.e. $2/3$ which corresponds to the coefficient we previously estimated.

As a conclusion, we can say that multi-lag tools have been used on partial discharge in order to characterize two signals received at two ends of cable reels. We have highlighted a time-scale transformation of the positive parts of the recorded data after propagation. Therefore, MLPDA tools have shown that they are promising to extract information from signals that suffer nonlinear modification.

3.5 Conclusions and Perspectives

In this article, we have shown in a first part that transients that were connected by time-shifted operator, time-scaling operator and amplitude modification, have invariant (or shape-invariant) phase diagrams if the lags were chosen wisely. We then proposed different methods to extract parsimonious parameters from each representation and combine all of them to highlight the named properties. So far, the developed methods provided good results for numerical examples as we showed in the last section.

Future axis of research can propose to extend this work to higher embedding dimensions. One possible idea could be to apply the same methods to the projections of phase diagrams onto the different planes of the phase space. More work should also be done to explore linear and nonlinear amplitude modifications.

References

1. J.-P. Eckmann, S.O. Kamphorst, D. Ruelle, Recurrence plots of dynamical systems. Europhys. Lett. (EPL) **4**(9), 973–977 (1987)
2. F. Takens, *Dynamical Systems and Turbulence, Warwick 1980*, Lecture Notes in Mathematics, vol. 898 (Springer, Berlin, 1981)
3. N. Packard, J. Crutchfield, J. Farmer, R. Shaw, Geometry from a time series. Phys. Rev. Lett. **45**(9), 712–716 (1980)
4. J.P. Zbilut, C.L. Webber, Embeddings and delays as derived from quantification of recurrence plots. Phys. Lett. A **171**(3–4), 199–203 (1992)
5. N. Marwan, J. Kurths, Nonlinear analysis of bivariate data with cross recurrence plots. Phys. Lett. A **302**(5–6), 299–307 (2002)

6. N. Marwan, A historical review of recurrence plots. Eur. Phys. J. Special Topics **164**(1), 3–12 (2008)
7. Recurrence Plots and Cross Recurrence Plots. http://www.recurrence-plot.tk/. Accessed: 2015-07-05
8. C. Bernard, T. Petrut, G. Vasile, C. Ioana, Multi-lag phase space representations for transient signals characterization, in *22nd IEEE European Signal Processing Conference (EUSIPCO 2014)* (2014), pp. 2115–2119

Chapter 4
Analysis of Non-stationary Signals by Recurrence Dissimilarity

Andrzej Rysak, Grzegorz Litak and Romuald Mosdorf

Abstract We propose a new method for testing non-stationary and intermittent signals. Using a basic recurrence plot quantifier, i.e. a recurrence rate, we create a relative measure which is sensitive to changes in the nature of signal. Given the specificity of this measure, we call it recurrence dissimilarity (RD). First, we test it using well-known non-linear systems for which we generate signals with different intermittent characteristics. In addition, the generated signals are disturbed by noise. The effectiveness of our measure is verified by applying different variables and noise levels. The results allow us to draw a number of conclusions concerning the proposed method. Finally, we give examples of using this method for experimental data analysis. We report the results of detecting changes in flowing patterns in two-phase flows and of switching heart modes in the signals recorded in ECG Holter tests.

4.1 Introduction

According to basic statistical moments, nonlinear signals are stationary at longer intervals without showing any trends. However, due to the nonlinear character of the system, multiple solutions of periodic and non-periodic attractors could appear together with short time reversible switching in time evolution. This effect is referred to intermittency and local non-stationarities and can be measured by recurrence quantification including *RR* [1]. Intermittent signals were studied systematically by Pomeau and Mannevielle [2]. Their studies were motivated by observations of the intermittent route to turbulence in convective fluids. More generally, intermittency

A. Rysak (✉) · G. Litak
Lublin University of Technology, Nadbystrzycka 36, 20-618 Lublin, Poland
e-mail: a.rysak@pollub.pl

G. Litak
e-mail: g.litak@pollub.pl

R. Mosdorf
Bialystok University of Technology, Wiejska 45C, 15-351 Bialystok, Poland
e-mail: r.mosdorf@pb.edu.pl

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8_4

in dynamical systems consists in irregular time alternation of phases with periodic and chaotic dynamics or of different forms characterized by chaotic dynamics [3]. Pomeau and Manneville described three different routes to intermittency where a nearly periodic system showed irregular bursts of chaos [2]. They classified intermittences into type I, II and III, these types corresponding to a saddle-node bifurcation, a subcritical Hopf bifurcation and an inverse period-doubling bifurcation, respectively. In periodic phases, the behavior was only nearly periodic, slowly drifting away from an unstable periodic orbit. Evolving with time, the system became closer to the unstable orbit again and returned to a nearly periodic behavior. Lengths of individual phases were unpredictable and characterized by different statistics depending on an intermittency type. In this chapter we examine the stationarity of time series using criteria that are based on their recursive properties. The main assumption is that non-stationary signals are those signals for which the $RR(\varepsilon)$ curve changes over time. The main aim was to convert this concept into a stationarity measure and test its potential with respect to detecting periodicity, non-periodicity, and intermittency in several exemplary mathematical and real signals. Signal detection and extraction as well as estimation of its periodicity and noise level are intensively investigated problems, particularly with regard to the physiological time series. In one of the first works in developing recurrence quantification analysis (RQA), Zbilut et al. [4] proposed a new technique, based on cross-recurrence quantification analysis, that allows to extract signals with very low signal-to-noise ratio and immediately assesses their degree of periodicity. Another interesting method for the detection of recurring patterns in the ECG signals was described by Sternickel [5]. This technique uses neural networks, which are trained with wavelet transformed templates. It allows to detect even temporally varying patterns within the ECG time series and works stable for signal-to-noise ratios larger then one. More recently, Urbanowicz et al. [6] used a general method estimating the noise level in heart rate variability by using a coarse-grained entropy and discovered that usually the level of such noise is within 5–15 %. The methods of nonlinear dynamics are often used in studies of ECG signals as a very effective tool in cardiac research [7, 8]. A review of some methods implementing recurrence quantification analysis in diagnosing changes of non-stationary cardiac signals was presented by Zbilut et al. [9]. They shown that this approach can be successful in regarding cardiac system state changes as well as its degrees of complexity and/or randomness. In the approach of Zhu et al. [10] heart rate variability and QT variability were characterized by short-term nonlinear RQA indexes. They tested Holter ECG signals of two groups: patients with dysfunction of the autonomic nervous system and healthy people. Significant differences were discovered between mutual information of paired RQA indexes (DET and LAM) in comparison of both groups. Application of nonlinear methods to cardiac study becomes more common. Therefore, the ECG time series seems to be a most attractive data to perform the first tests of the new recurrence measure.

An intermittent flow can also occur in two-phase (air-water) flows in mini-channel. This phenomenon is reflected in the flow patterns consisting with the water content and various sizes bubbles including small and elongated bubbless (slugs). Various gas-liquid two-phase flow patterns were visualized and identified by Triplett

et al. [11] and Serizawa et al. [12]. Recently, the formation of periodic patterns was investigated by Gorski et al. [13, 14]. It is worth noticing that two-phase flow typical of technical systems where, apart from a liquid phase, a gas phase appears. Such systems include boilers producing steam for turbines or in pumps, which operate close to a vapor pressure of a fluid being pumped. Intermittencies and local non-stationarity were studied by various tools including wavelets, multifractals, multiscale entropy, and recurrence plots [1, 15–17]. In particular, the recurrence plots quantification technique was applied to successive windows in the corresponding time series [1].

In this study, we test a new recurrence measure designed to distinguishing states of non-stationary signals. The main focus is to assess effectiveness of this measure. Therefore, we first generate a series of numerical values which are then tested using this measure. We employ well-known nonlinear systems as source of the non-stationary signals. To simulate real conditions, a noise component is added to the numerical series. We test all simulated results for different noise levels. Given the limited scope of this work, only selected results are presented and discussed in Sect. 4.4. Section 4.3 describes the measure of signal state applied in this work. In Sect. 4.5 we present the results of ECG signals analysis produced with this method. In the subsequent section, we report the results of applying this method to study changes in the flowing patterns of the two-phase flows measured in circular channels. A comparison with the traditional method of periodogram is presented in Sect. 4.7, where the selected mathematical time series have been tested by both methods. The results show good agreement between both test methods, pointing to the advantage of the RD measure in some cases.

4.2 Investigated Nonlinear Systems

Prior to using the new measure of the signal status defined in Sect. 4.3 for analysis of real waveforms, we first tested this measure using numerically generated time series. These time series were obtained as a solutions of well known nonlinear systems:

1. Burke–Shaw system
2. Duffing–van der Pol oscillator
3. Windmi system.

In order to unify the description of all investigated waveforms in both the time and frequency domain, we denote the independent variable of the studied nonlinear systems as time. By applying the proposed measure to above systems, we examine how this measure can be used to distinguish between the intermittent signal states. Artificial data were prepared in the following way. For each system were found three solutions with different intermittent characteristics:

- Solution 'A'—a periodic signal with rare and short-term switching to a chaotic state
- Solution 'B'—a signal which switches intermittently between periodic and chaotic states
- Solution 'C'—a chaotic signal with rare and short-term switching to a periodic state.

The above characteristics of signals A, B and C were selected to check effectiveness of the RD measure regarding each of the above-mentioned nonlinear systems. The A-type signal enables us to test the measure with respect to finding chaotic windows or transient disturbances. The time series of type B typically have intermittent characteristics. The researches on these series focuses on detecting intermittent switching between different signal states. In the case of signal C, the task of the RD measure is to find a short periodic windows immersed in the chaotic background. Tables with the applied parameter values and diagrams illustrating the generated time series are given in Sects. 4.2.1–4.2.3. To give an exhaustive presentation of the investigated mathematical data, the time series are shown together with their time-frequency representations in the form of spectrograms. In addition, different levels of noise were added to the numerically generated time series. The use of the proposed measure to investigate the numerical waveforms with a noise enables us to assess its suitability for analysis of real signals. The procedure for interfering waveforms by adding noise is described in Sect. 4.2.4.

4.2.1 Burke–Shaw

The first set of a basic time series is obtained as a solution of the Burke–Shaw system. This system was derived by B. Burke and R. Shaw from the Lorenz equations [18]. The set of ordinary differential equations is

$$\begin{aligned} \dot{y}_1 &= -U(y_1 + y_2) \\ \dot{y}_2 &= -Uy_1y_3 - y_2 \\ \dot{y}_3 &= Uy_1y_2 + V, \end{aligned} \tag{4.1}$$

where y_1, y_2 and y_3 are variables while U and V are constants. The superscript dot denotes the time derivative. The corresponding time-series for y_1 solutions: A, B and C (Fig. 4.1) were obtained by using selected values of the U and V coefficients (Table 4.1). The characteristics of signal A are described in the introduction to this section. This time series reveals the presence of two chaotic windows and a few short-term disturbances. Signals B and C are similar. In the chaotic time-series we can observe visible periodic windows. The system coefficients are selected such that periodic windows occur more frequently in time series B. In Fig. 4.2, the spectrograms of the A, B, and C signals present their time–frequency characteristics.

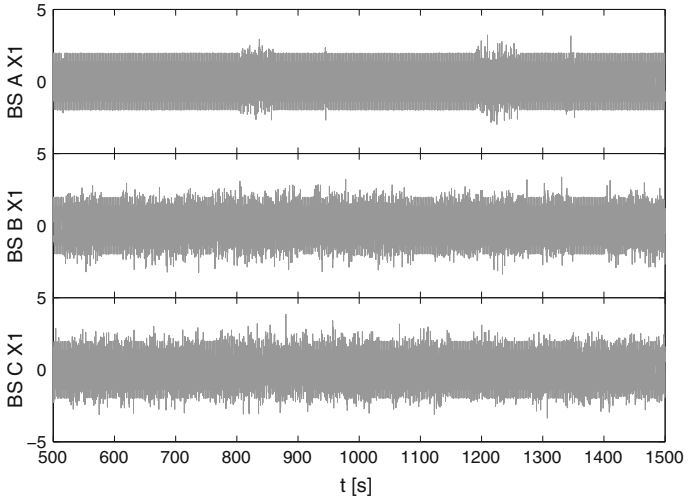

Fig. 4.1 Time series of the first coordinate of the BS system's solution. The diagrams of solutions A, B, C are shown in the same sequence from *top* to *bottom*. Diagrams B and C show time lengths with periodic fluctuations. They occur less frequently in C

Table 4.1 Coefficients of the Burke–Shaw system

Type of signal	*U*	*V*
A	13.979645	13
B	13.97755	13
C	13.97	13

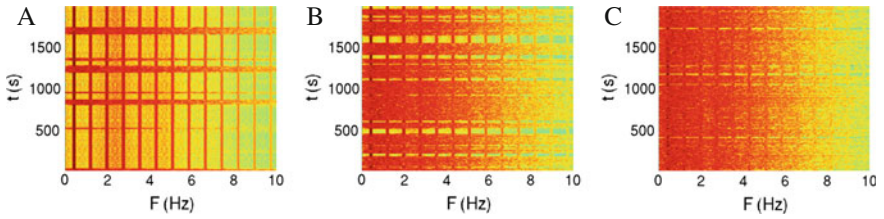

Fig. 4.2 Time–frequency characteristics of X1 solution of the Burke–Shaw system obtained for signals A, B, and C

4.2.2 Duffing–van der Pol Oscillator

Another system we employed to generate non-stationary signals was the extensively studied periodically forced Duffing–van der Pol oscillator [19, 20]. This system is defined by the equation

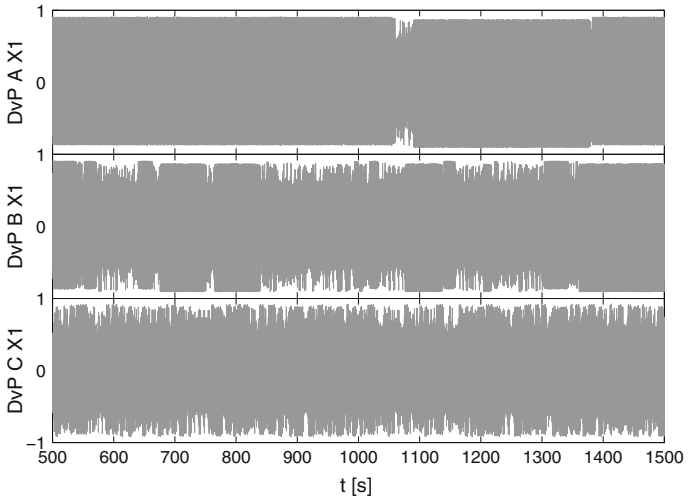

Fig. 4.3 Time series A, B and C as were generated by the Duffing–van der Pol system. The first coordinate of the solution obtained for the coefficient values given in Table 4.2

Table 4.2 Coefficients of the Duffing–van der Pol system

Type of signal	μ	γ	A	ω
A	0.20	9	0.358	1.06239
B	0.20	9	0.35	1.06239
C	0.18	9	0.35	1.06239

$$\ddot{x} - \mu(1 - \gamma x^2)\dot{x} + \eta x + x^3 = A\, sin(\omega t), \tag{4.2}$$

where x is a variable and μ, γ are constants parametrizing the self excitation effect, η is a spring coefficient and $A\sin(\omega t)$ is an external harmonic force with the amplitude A and frequency ω. To obtain the corresponding time series for x: A, B and C (Fig. 4.3), we have assumed that $\eta = 0$ and that the values of other coefficients are set as listed in Table 4.2. Signal A has a very regular amplitude and only two visible periodicity breaks. Signal B is typically intermittent. Periodic areas are difficult to see in the diagram for signal C. This being the case, brief periodic windows could only be detected by detailed examination of the magnified fragments of the series. The spectrograms of the Duffing–van der Pol signals, A, B, and C are presented in Fig. 4.4.

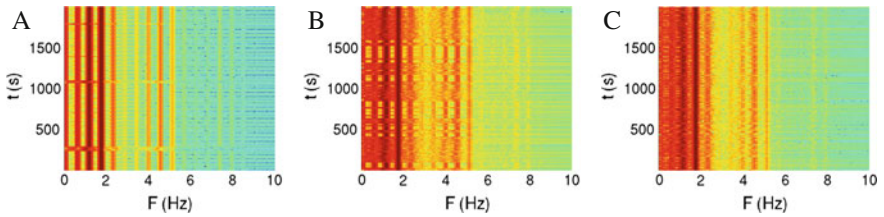

Fig. 4.4 Time–frequency characteristics of the Duffing–van der Pol X1 solution of signals A, B, and C

4.2.3 Windmi

The third analyzed nonlinear system was the WINDMI attractor introduced by Horton and Doxas [21] to describe the solar-wind-driven magnetosphere-ionosphere system. In this chapter we analyze a simplified form of the system which was developed by Sprott [22]:

$$\begin{aligned} \dot{y}_1 &= y_2 \\ \dot{y}_2 &= y_3 \\ \dot{y}_3 &= -ay_3 - y_2 + b - exp(y_1), \end{aligned} \tag{4.3}$$

where y_1, y_2 and y_3 are variables while a, b are constants. The corresponding time series for y_3 are shown in Fig. 4.5. They were obtained for the coefficients listed in Table 4.3. For this system we generate an additional time series (P), characterized by switching between two periodic signal states. As in the case of the Duffing–van der Pol system, signal C appears to be completely chaotic. Therefore, it seems to be a good experimental field to assess effectiveness of the applied measure in detecting periodicity. Spectrograms of the Windmi system are shown in Fig. 4.6.

4.2.4 Noise Generation

The aim of this work is to detect intermittent states of non-stationary signals using a new recurrence measure. To test the measure's usefulness under conditions similar to real ones, different levels of noise are added to the numerically obtained time series. Values of the noise signal are drawn from a standard normal distribution (Gaussian noise). For each coordinate k of a non-linear system the noise vector is expressed as follows:

$$\mathbf{n_k} = \{n_{ki}\}, i = 1, 2, \ldots, n \tag{4.4}$$

The noise level is defined with respect to a source signal level. Whenever noise is added to the numerical signal, its amplitude is multiplied by a coefficient calculated as part of the standard deviation σ_k of the source signal $\mathbf{y_k}$

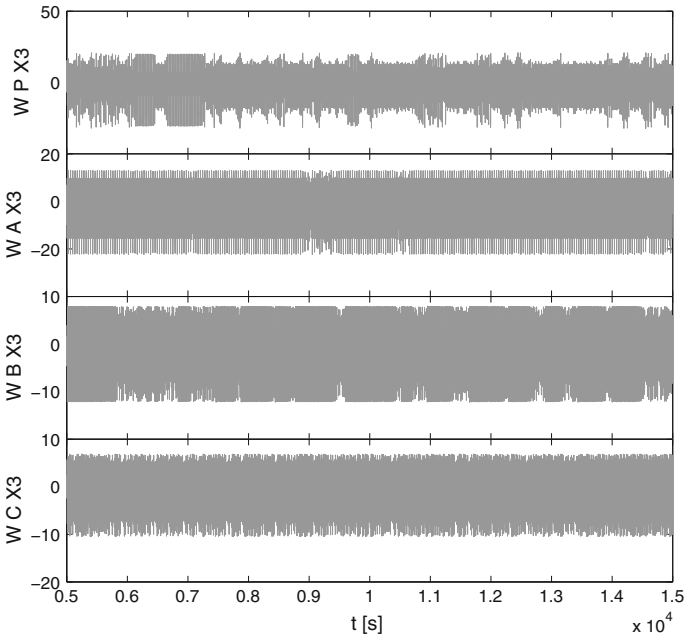

Fig. 4.5 Time series obtained for the Windmi system, i.e. third coordinate of the solution. The *top* trace is a P-type waveform in which two different periodic-like states were observed

Table 4.3 Coefficients of the Windmi system

Type of signal	*a*	*b*
P	0.5061541215	2.5
A	0.704695	2.684
B	0.678795	1.99816
C	0.715	1.95

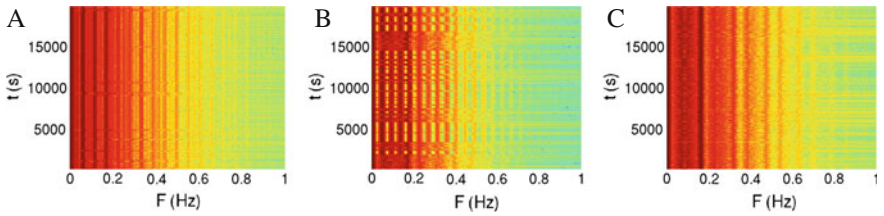

Fig. 4.6 Time–frequency characteristics of the Windmi X1 solution for signals A, B, and C

Table 4.4 Signal to noise ratio for different noise levels

Noise level (%)	SNR	SNR (dB)
5	400	26
10	100	20
20	25	14
40	6.25	8
60	4.44	2.8
80	1.94	1.56

$$\mathbf{y_k} = \{y_{ki}\}, i = 1, 2, \ldots, n$$

$$\sigma_k = \left[\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \tilde{y})^2\right]^{\frac{1}{2}}$$

The value of this coefficient is marked in the figures as the noise level and expressed as a percentage. It is denoted as NP. Then the signal disturbed by noise about the NP level is calculated as follows:

$$\mathbf{y_k}^{NP} = \mathbf{y_k} + \frac{NP}{100}\sigma_k \mathbf{n_k} \tag{4.5}$$

All non-stationary signals listed in Sect. 4.2 were tested in terms of adding different noise levels. By convention, the noise was added to the investigated signals at the following levels: 5, 10, 20, 40, 60 and 80 %. The values of SNR for different noise levels are listed in Table 4.4. The disturbance of the signal due to different noise levels is clearly visible in the case of the BS-A (Burke–Shaw A-type) signal (Fig. 4.7).

4.3 Recurrence Dissimilarity

In this chapter, we propose a new measure to distinguish between non-stationary signal based on differences in recurrence plot patterns. The Recurrence Plot method was developed by Eckmann [23] and extended by Webber and Zbilut [24], Casdagli [25], later by Marwan et al. [26, 27] and others. This approach is very useful for analysis of short and noise-affected experimental data. Prior to the application of the RP method, it is necessary to move a considered time series to an m-dimensional embedding space. After Takens [28] and based on our initial time series $x(t_i), i = 1, 2, 3, \ldots, n$, we define the vector $\mathbf{Y}$ in the m-dimensional embedding space as:

$$\mathbf{Y}(t_i) = [x(t_i), \quad x(t_i - \tau), \quad x(t_i - 2\tau), \quad \ldots, \quad x(t_i - (m-1)\tau)] \tag{4.6}$$

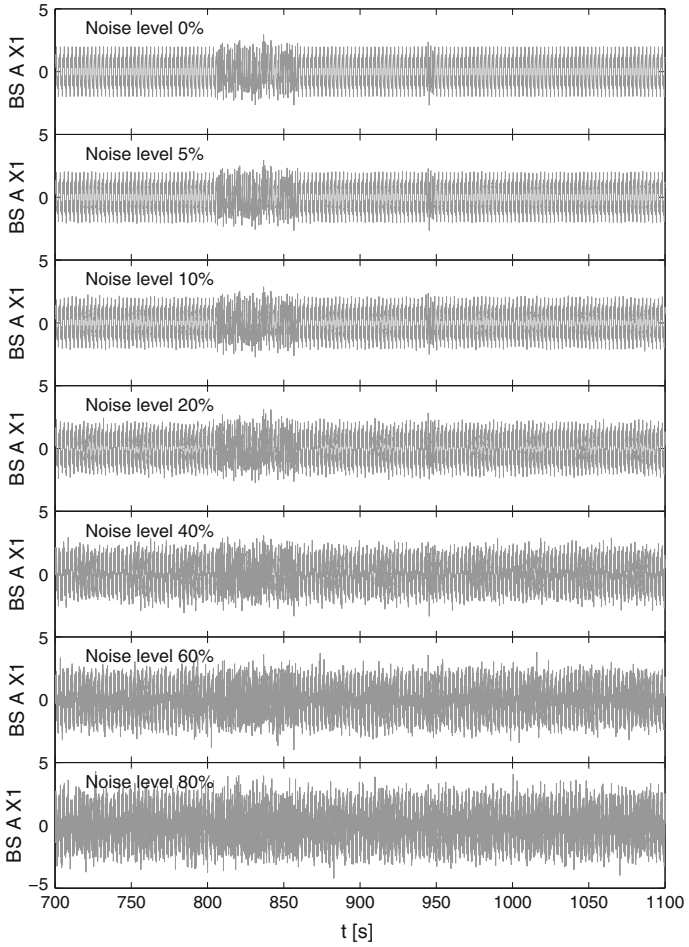

Fig. 4.7 Burke–Shaw A-type signal (time series) with different noise levels

In the embedding space, for the vector $\mathbf{Y}$ we define the distance matrix R which determines neighbors for each element of the vectorial time series. Two vector elements are called neighbors if the distance between them in the embedding space is lower than the threshold value ε. The elements R_{ij}^{ε} of the distance matrix R are defined as:

$$R_{ij}^{\varepsilon} = \Theta\left(\varepsilon - ||\mathbf{Y}_i - \mathbf{Y}_j||\right), \tag{4.7}$$

and they are equal to 1 for the components that are neighbors, and 0 otherwise. A recurrence rate (RR) is a recurrence plot quantifier, the value of which is equal to the number of elements of the R matrix equal to 1. By calculating the RR values for different values of ε, we obtain an $RR(\varepsilon)$ curve.

The proposed measure determines a relative difference between the state of the signal on a reference section at about $t = t_1$ (reference window) and the state of the signal on a section of the same length about the point t. For each window, the RP patterns are calculated for a few threshold values ε, and for each pattern (R matrix) the RR quantifier is determined. In this manner, a $RR(\varepsilon)$ curve is created for every window. This curve is the core of the measure. In our preliminary studies, we examined the relationship between the RR curves obtained for windows set in various states of non-stationary signals. Both the tested time series and the windows for which the $RR(\varepsilon)$ curves were computed are shown on the left in Fig. 4.8.

The first calculations demonstrated that the $RR(\varepsilon)$ curve obtained for both windows (colors agreement) differ only when both windows are placed in the areas described by a different signal state (Fig. 4.8b). The black RR curve was obtained for a cross-recurrence plot pattern of both windows. This observation inspired us to define the measure of non-stationary signal state as a distance between the $RR(\varepsilon)$ lines of both windows (at two time points). We call this measure 'recurrence dissimilarity' because it reflects the difference between recurrence plot patterns of both windows. In this work, the result of the RD method, i.e. the distance between lines of $RR(\varepsilon)$, is denoted as D_{RR}. The D_{RR} distance is calculated with a simple formula

$$D_{RR} = \left\{ \sum [RR_1(e_i) - RR_2(e_i)]^2 \right\}^{\frac{1}{2}}. \tag{4.8}$$

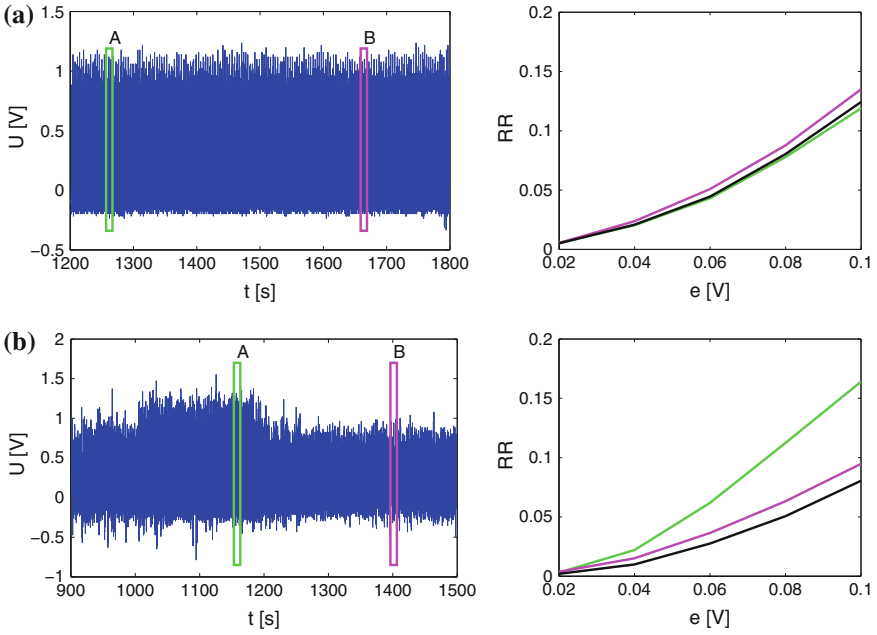

Fig. 4.8 Two windows located in places with similar (**a**) and different (**b**) signal characteristics. The corresponding $RR(\varepsilon)$ curves do not overlap in (**b**). The *black curve* is the $RR(\varepsilon)$ function calculated for cross-RP

Below are listed the most important parameters for this method.

- W—window width
- S—sliding window step
- t1—starting point (reference window position)
- m—space dimension
- τ—delay
- Δ—ε growth rate (in m-dimensional space), $\varepsilon(i) = i\Delta$, $i = 1 \ldots 5$.

Without proper selection of values of the embedding coefficients: m, τ, we do not get the results determining recurrence similarity. We found experimentally that the method is effective for m set to 3 (even for the 2-dimensional Duffing–van der Pol system) and τ made equal to at least 10. With a correctly specified range of values of the ε parameter, we can assess the sensitivity of the method. When Δ is too high or small, $D_{RR} = 0$. Optimal values of Δ were determined experimentally for each system depending on the level of noise. The step S is related to the time resolution of the result, whereas the width W of the sliding window determines its precision. By definition, the relationship $RR(\varepsilon)$ is always on the increase, but the variability of the RR curve can be different in both windows, i.e. the reference and current window. As a result, the difference between relative $RR(\varepsilon_i)$ values can be either positive or negative. The measure adopted in this work (4.8) is rough. The use of more precise measure, the D_{RR}, which takes into account the sign when calculating the distance between the two curves should increase the functionality and sensitivity of this technique. Further improvement of this method can be achieved by reducing the step shift of the sliding window. This should significantly improve the resolution of the method. Good quality time series $D_{RR}(t)$ can then be used to determine the recurrence plot of the RD measure (RD-RP). Diagrams obtained in this way can be an interesting way of presenting characteristic properties of the examined time series.

4.4 Examination of the Signal State

This study involved testing all generated signals with an addition of different noise levels. An analysis of signals with numerically generated noise under controlled conditions enables formation of signals with known characteristics. By analyzing these signals, it is relatively easy to interpret the results and assess the proper operation of the employed method. In this section, we present only some examples to demonstrate what kind of results can be obtained for different types of signals and to show the impact of the noise on recurrence dissimilarity measurements. In every case, the reference window is set such that it is compatible with the purpose of the test in terms of recurrence dissimilarity. That is, for example, searching for chaotic inserts, we must set the reference window in location with a periodic signal, and vice versa. The reference window is marked in the figures by a rectangular green frame. If it is not visible, this means that the image is magnified and the window is beyond the scope of the figure.

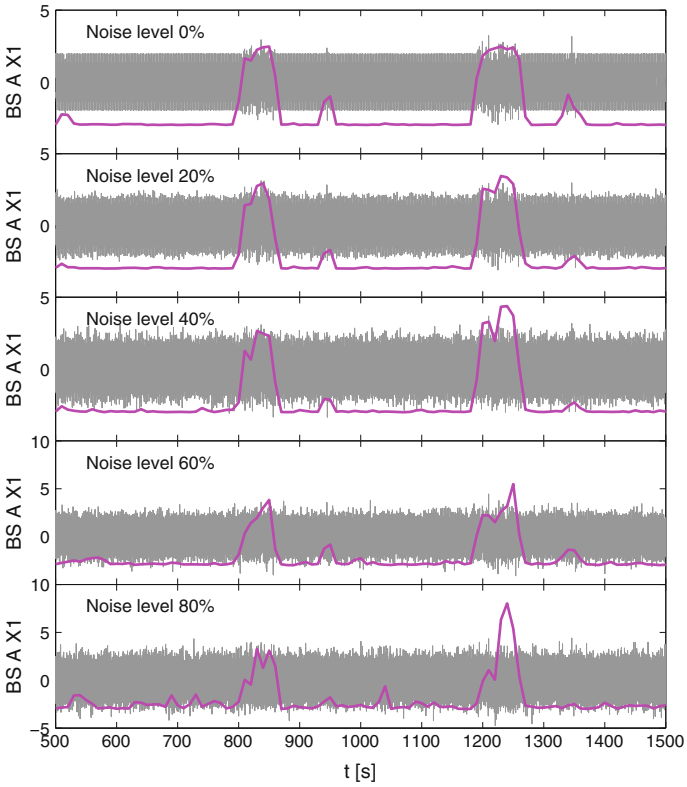

Fig. 4.9 The function $D_{RR}(t)$ (*magenta*) calculated for the time series A-X1 of the Burke–Shaw system (corresponding time series in the background). Detection of chaotic windows in the signal with different noise levels. In each case, the reference window was set outside the scope of the figure

4.4.1 Detection of Chaotic Windows in the Burke–Shaw System

The D_{RR} curves detected in Burke–Shaw system with different noise levels are presented in Fig. 4.9. The RD method identifies well chaotic episodes in the A-type signal which is dominated by a periodic-like behavior of the system. As can be seen in the figure, the applied technique detects the areas of the chaotic windows correctly, even after the addition of 80 % noise to the original signal.

4.4.2 Recurrence Dissimilarity in the Duffing–van der Pol System

The recurrence dissimilarity for the B signal in the Duffing–van der Pol system is presented in Fig. 4.10. For the B-type signal, the nonlinear system often and

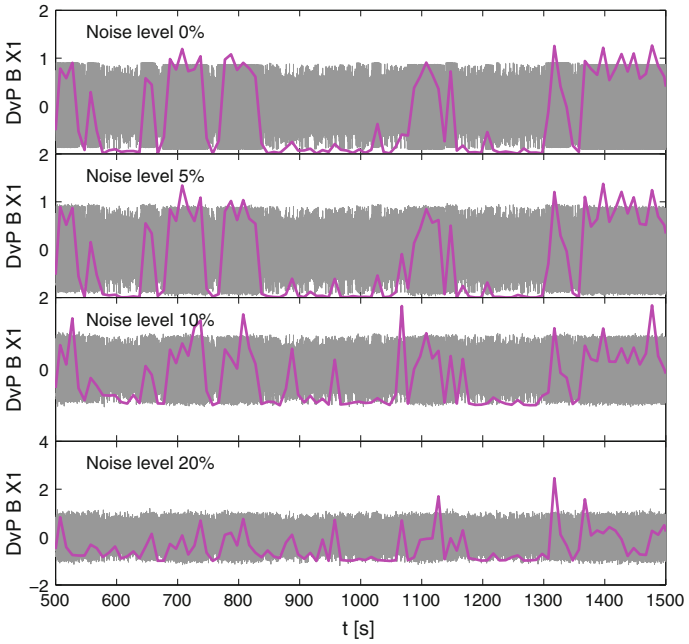

Fig. 4.10 D_{RR} function (*magenta*) calculated for the B-X1 signal of Duffing–van der Pol system. Investigation of the intermittent switching for the signal with different noise level. In each case, the reference window was set outside the scope of the figure

intermittently switches between periodic-like and chaotic states of the signal. It can be observed that periodic states are still detectable up to a noise level of 20 %.

In Fig. 4.11 we present the results obtained for the C-type signal dominated by chaotic changes where short and infrequent periodic excitations mingle. In addition, the signal was disturbed by a 10 % noise level. Even under these difficult conditions, weak and narrow periodical areas can still be detected with the RD method. Figure 4.12 shows a magnified image of one of D_{RR} peak and thereby detected periodic-like fluctuations.

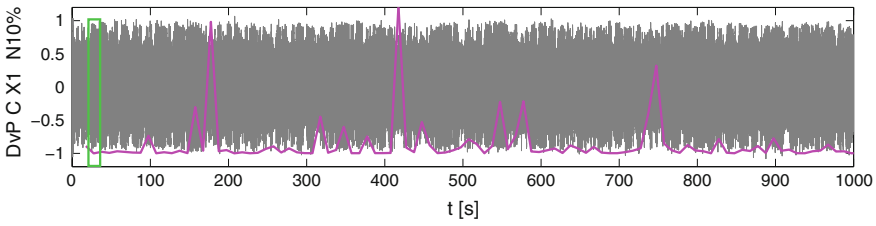

Fig. 4.11 The first coordinate of the C-type signal generated by the Duffing–van der Pow system with a 10 % noise level. The D_{RR} line is plotted in *magenta*. The reference window is marked in *green*

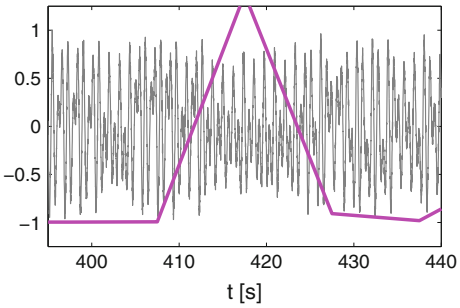

Fig. 4.12 Magnified fragment of the Fig. 4.11 with a D_{RR} peak, marking a short periodic area

4.4.3 Detection of Chaotic and Periodic Windows in the Windmi System

The C-type signal of the Windmi system seems to be completely chaotic (Fig. 4.13). But the RD analysis reveals two high peaks of the D_{RR} function, which indicates some differences of the signal in these places. In Fig. 4.14 we show a magnified image of these two peaks. The area identified by the left peak (left figure) is clearly periodic and stands out from the chaotic signal. But the regularity indicated by the right peak (right figure) is difficult to distinguish from the surrounding chaotic background. After adding 10 % noise to the W-C-X3 solution (Fig. 4.15), the RD measure produces a result which is partially consistent with that which was obtained for the undisturbed signal. Both areas identified in Fig. 4.13 can still be detected. However, a comparison of the results presented in Figs. 4.13 and 4.15 shows some distinctive phenomena caused by noise. Following the addition of noise, some vertices of the D_{RR} function—which in the noise-free case had a small value—significantly increase. This effect occurs for the peaks located in the vicinity of $t = 2100$ s and $t = 9000$ s. The noise also induces an entirely new peaks which are invisible in the undisturbed signal (for $t = 5500$ s).

The recurrence dissimilarity obtained for the A-type signal of the Windmi system for different noise levels is shown in Fig. 4.16. As can be seen in the figure, the noise strongly alters the results obtained with the RD method. The increase in the noise

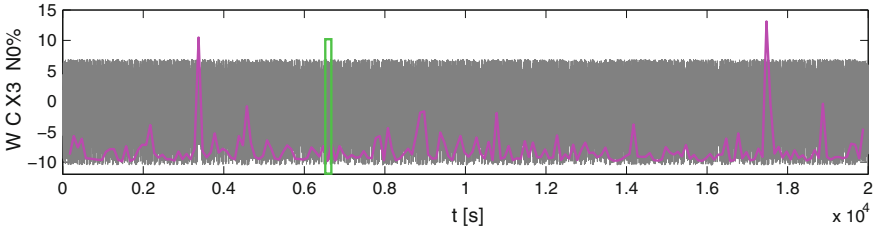

Fig. 4.13 D_{RR} function (*magenta*) of C-X3 solution of the Windmi system with a zero noise level. Short-term periodic areas are effectively detected. The corresponding time series is plotted in the background

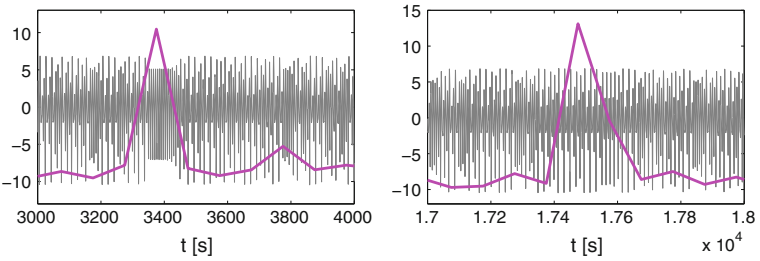

Fig. 4.14 Fragments of the Fig. 4.13 which show areas with the D_{RR} peaks detecting narrow periodic windows

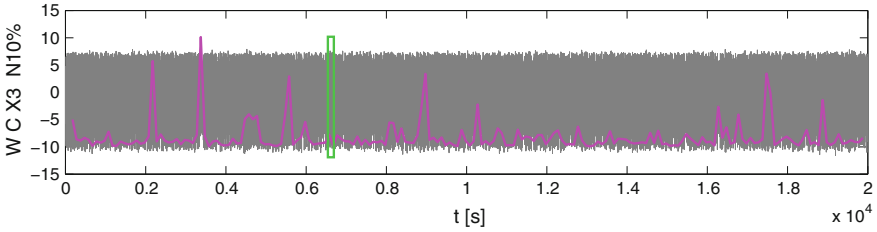

Fig. 4.15 D_{RR} function (*magenta*) of C-X3 of the Windmi system time series with 10 % noise level

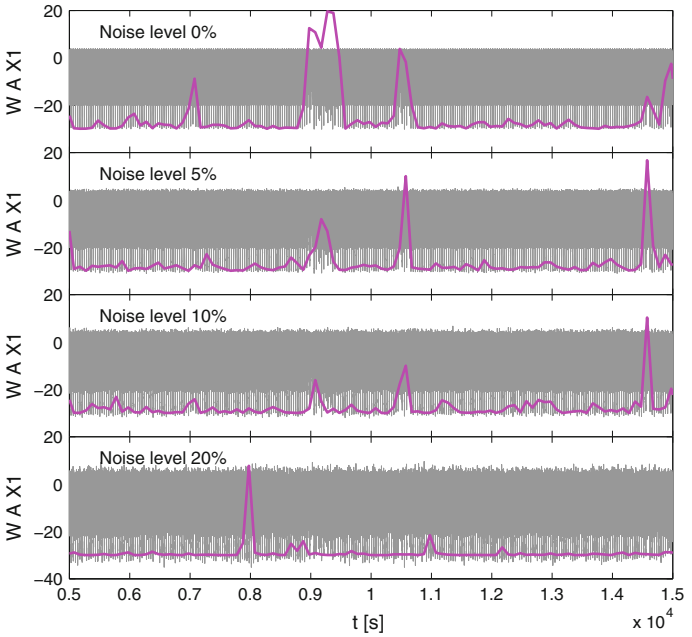

Fig. 4.16 D_{RR} function (*magenta*) of A-X1 solution of the Windmi system time series with different noise levels. In each case, the reference window was set outside the scope of the figure

level significantly reduces the detection of the chaotic windows. For the signal with added 20 % of noise, the chaotic windows are no longer detectable. Comparing these results with the analogous results obtained for the Burke–Shaw system (Fig. 4.9), we can notice that the Windmi system is much less resistant to the effect of noise. Therefore, we can conclude that in terms of recurrence similarity, the difference between chaotic and periodic signals is significantly higher for the Burke–Shaw system.

4.5 Recurrence Dissimilarity in the ECG Signals

To test the method for real signals, we used the RD measure to study the variability of ECG Holter signals. Analyzed were the results derived from active patients with possible heart defects. Under such conditions it can be expected that in the results of the study will reveal a visible impact of such phenomena as:

- changes in heart rate
- changes in the amplitude of the ECG signal (the effect of apparatus)
- electrical interference
- error in electrodes connections
- changes in the ECG signal resulting from the revealing of a disease state.

Below we provide a few examples which illustrate how the RD measure works under such circumstances. Wherever it is drawn, the red horizontal line marks the zero of the D_{RR} signal level. Figure 4.17a shows the ECG signal, a portion of which (at around t = 1200 s) has an amplitude which is significantly higher than the rest. The D_{RR} line takes high values over this length, which indicates that its nature differs from that of the signal in the selected reference window. In the analysis of the next ECG waveform Fig. 4.17b only one short disturbance is identified.

The ECG signal, shown in Fig. 4.18 (top) has in its central part evident excitations. The D_{RR} function detects strong but short-term disturbances in the region of these excitations. It can be noted, that the D_{RR} function increases once these excitations are over. The lower part of Fig. 4.18 shows magnified fragments of the ECG signal before (left) and after (right) these excitations. Comparing them, we can detect a difference between both signals responsible for the increase in the D_{RR} function. Further on, we analyze the ECG waveform which has a long region with a higher amplitude in the centre (Fig. 4.19). The D_{RR} function gives high peaks only in the boundaries between the time regions with different signal levels. A similar signal is presented in Fig. 4.20. In this case, intermittent changes are visible in the ECG signal level. The D_{RR} function does not detect either changes in the nature of the signal or transitions between the areas with different signal levels. Thus, the RD detects only disturbance points and is insensitive to signal levels. One visible D_{RR} peak appears in the region of noticeable time series disturbance at about t = 3310 s. The next two diagrams in Fig. 4.20 show the magnified details of the ECG results for lower (left) and higher (right) signal levels. Both series are qualitatively identical and differ only with respect to signal level.

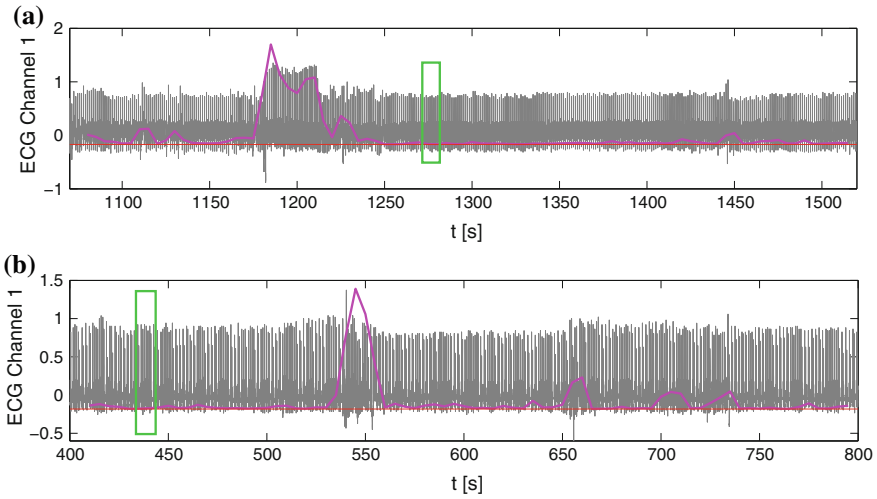

Fig. 4.17 Recurrence dissimilarity (magenta) detects a fragment of signal with different characteristics (**a**) and relatively short signal disturbance (**b**). In both cases, the location of the reference window is marked by *green*. The corresponding time series is plotted in the background

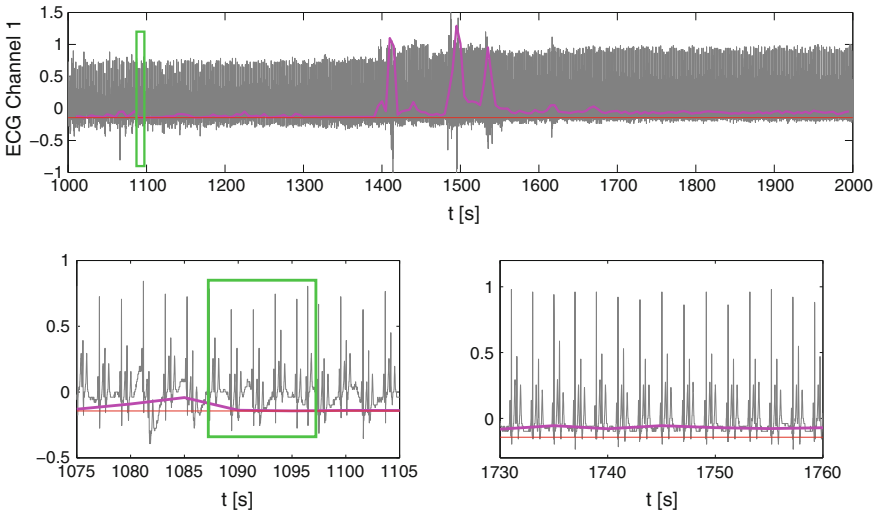

Fig. 4.18 Recurrence dissimilarity (magenta) detects a signal disturbance after which the level of dissimilarity slightly increases. Below are shown magnified images of fragments from both sides of the signal disturbance

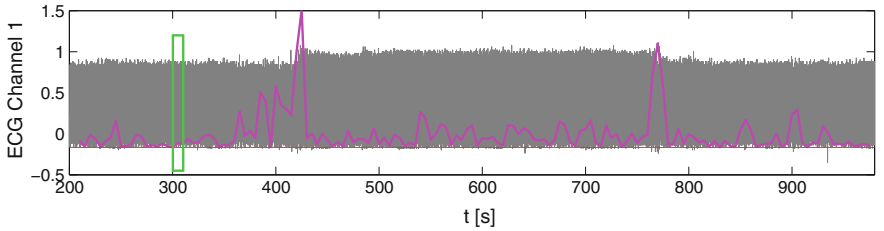

Fig. 4.19 RD (magenta) detects signal disturbance in the regions where the signal level changes

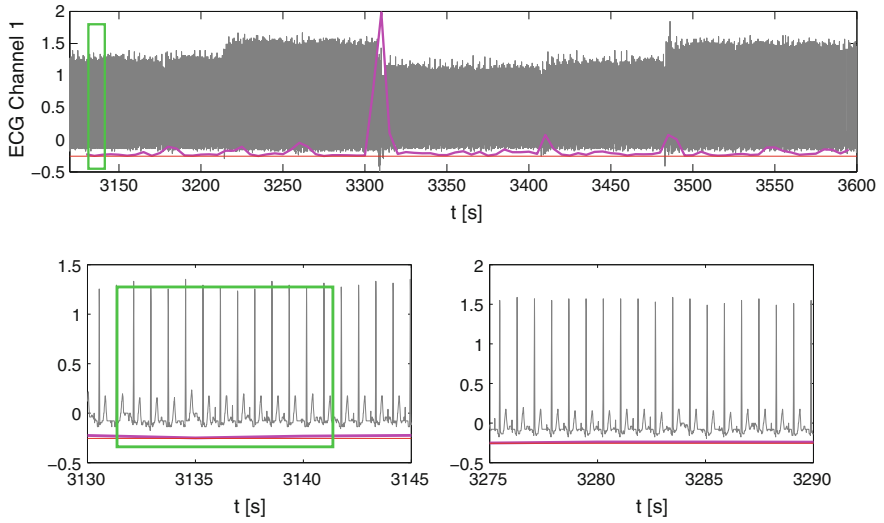

Fig. 4.20 Signal level switching. RD (magenta) detects only disturbance points and is insensitive to the signal levels. Below are presented magnified fragments

4.6 Evolution of the Two-Phase Flowing Systems

In this section, we test changes of flowing structures in two-phase air-water flow systems by the RD method. In the experiment, the gas flow rate is kept constant whereas the water flow rate is systematically decreased with time, thus the flow velocity of the mixture decreases as well. We report here the results obtained for the gas flow rate set equal to 0.1, 0.2, and 0.3 l/min. In each case, the water flow rate decreases from 0.085 to 0.01 l/min. A schematic design of the experimental stand is shown in Fig. 4.21, where the water loop (blue) and air branch (yellow) supply a bubble generator (BG). In both subsystems, the most important elements such as pumps (WP, AP), tanks, valves (WV, AV), and flow meters (WM, AM) are marked. Flow patterns were recorded with a Phantom digital camera at 5000 fps. The content of the minichannel (bubbles or liquid) was qualitatively assessed using a laser-phototransistor sensor. The sensor consisted of a laser which generated the laser beam with a

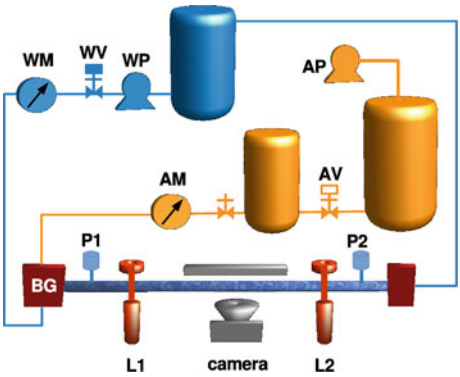

Fig. 4.21 Two-phase system flows in a channel. Air is provided to the water loop (*blue*) from a side branch (*yellow*). Pressure is measured by two sensors (P1, P2). Two lasers measure the perturbations of the system (L1, L2). The camera collects video frames

diameter of 3 mm, the laser and silicon sensor placed in focal point of lens. Data from the laser-phototransistor sensors was acquired by the acquisition system at a sampling rate of 5 kHz. If both, gas and water flow rates are high the flow nature is turbulent and chaotic. This occurs at the beginning of measurement for flows with a high value of the gas flow rate. The data are collected by a laser probe. The laser beam has a diameter of 3 mm. Therefore, only the structures with comparable dimensions have a significant impact on the results.

The non-stationarity of the two-phase flow increases with changing the ratio of both phases. These processes lead to reducing flow instabilities; they also change the flowing structures and their organization. As the water flow rate decreases, the frames collected with a high-speed camera reveal changes in the shape of flowing gas structures from chaotic small bubbles, and mixed flows of bubbles and slugs, to more stable flows of long slugs. Along with the changes in shape, the distances between the structures and their configurations change, too. During these changes, the flowing sequences of structures often become periodically arranged. At a relatively high gas content, the flow system tends to an annular or long slugs flow. Figures 4.22, 4.23, 4.24 present the RD results of the laser-collected data for the gas flow rate set to 0.1, 0.2, and 0.3 l/min, respectively. In the first case, for the water flow rate q = 0.1 l/min (Fig. 4.22) flow begins as a chaotic motion of bubbles, later becoming a mixture of bubbles and unstable slugs. During this period of time, the D_{RR} function value remains unchanged. Starting from t = 55 s the number of chaotic bubbles decreases and the movement of slugs becomes more stable. The value of the

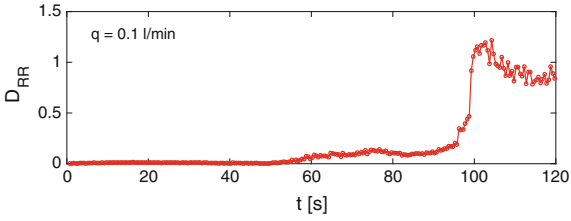

Fig. 4.22 Changes in recurrence similarity in the two-phase flow when the gas flow rate is set to to 0.1 l/min

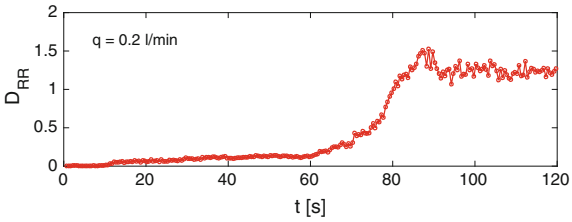

Fig. 4.23 Changes in recurrence similarity in the two-phase flow when the gas flow rate is set to 0.2 l/min. The flow rate of the mixture increases with increasing the gas flow rate

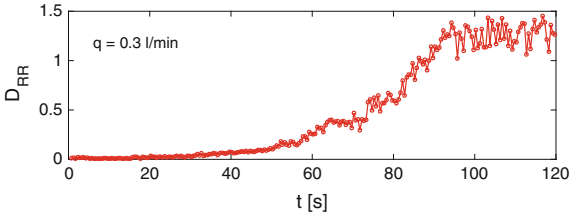

Fig. 4.24 Changes in recurrence similarity in the two-phase flow when the gas flow rate is set to 0.3 l/min. The highest gas flow rate

function increases slightly. Until about t = 100 s, the flowing slugs change their configuration and size only to a little extent. After this time, their size increases rapidly, which causes a sudden increase in the D_{RR} function. A comparison of changes in the D_{RR} function with video images enables better interpretation and verification of the results produced with the RD method. The selected video frames collected during measurement with the gas flow rate set equal to 0.1 l/min are presented in Fig. 4.25.

In the next experiment, for q = 0.2 l/min (Fig. 4.23), the flow begins in a similar vein, but then slugs quickly emerge. Because of slugs instability and the presence of chaotic bubbles, this change causes, however, only a small increase in D_{RR}. After the t = 60 s, the slugs are more stable and the number of bubbles is reduced. Starting from this point, they length increases, which makes the D_{RR} function increase. In the last example (q = 0.3 l/min, Fig. 4.24), the flow starts as a chaotic movement of a bubble and slug mixture. With a lapse of time, both the flow rate and the number of chaotic bubbles decrease. At the same time the slugs become stable. Looking at the

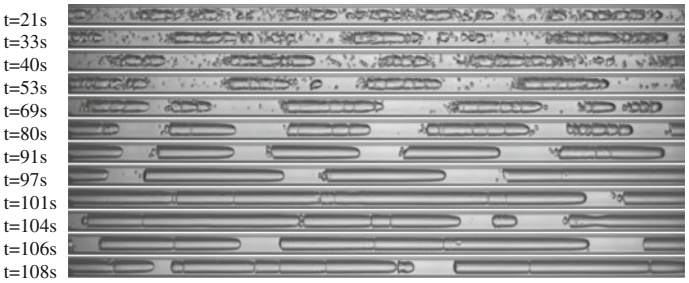

Fig. 4.25 Photo-frames of the structures flowing in the channel taken for the flow where the gas flow rate equals to 0.1 l/min

function D_{RR}, we can observe a slow increase in their value in a wide range of time (decreasing the water flow rate), corresponding to the expansion of slugs length and increased stabilization of the flow.

The RD results reflect well the evolution of gas structures in two-phase flows, even when analyzing signals with a low spatial resolution (laser). The method detects chaotic flow and slug formation, and produces a high value of the D_{RR} function for a steady flow of long structures. In the final stage of the measurement of each flows, the D_{RR} function's value saturates. This is due to the stabilization of the water flow rate at minimum and the reconfigurations of gas structures under these conditions. The changes in gas structure configurations occur together with changes in the relative flow velocity of water and gas. In the RD measurements, such reconfigurations of gas structures lead to variations in the D_{RR} function.

4.7 Recurrence Dissimilarity Versus Periodogram Map

As has been shown in the above sections, the RD effectively detects changes in signal characteristics. Because the RD measure analyzes the recurrence similarity, it is not sensitive to changes in signal amplitude but successfully detects many other signal exceptions. To assess the effectiveness of the RD method, we compare its results with the results produced with one of the standard time series analysis methods—the periodogram technique. The $D_{RR}(t)$ function is determined by the sliding window technique. For each analyzed waveform the appropriate window width and time step are used. Comparing both techniques, the periodograms are calculated for the same values of the window widths and step shifts that have been applied in determining the function $D_{RR}(t)$. Figure 4.26 shows exemplary periodograms obtained for the studied mathematical systems. For each system we report PSD (power spectrum density) results obtained for one selected time window. The frequency axis is scaled to a 0–1 value range. By moving the time window, we get a series of periodograms—one periodogram for each point of the $D_{RR}(t)$ curve. To compare the results obtained by both methods, each series of periodograms is presented on one 3D time-frequency chart. To simplify the graph, we project the 3D chart to a 2D color map and convert the colors to gray scale so that the drawing's background is white. For the obtained periodogram maps, the gray color marks the frequencies for which the signal amplitude at any given time is relatively high. On such a map, a periodic component stable in time is indicated as a dark horizontal line. The darkness of the line is proportional to the frequency peak value. In the all presented periodogram maps we narrow the frequency range to a scope where the PSD have significant values.

In Fig. 4.27 the results of the RD measure and periodogram map (RD-P) are compared for the time series type-A of the BS system without added noise. The figure shows long gray lines corresponding to the frequency peaks of the PSD function, which do not change theirs position over time. The breaks in these lines correspond to the peaks of the $D_{RR}(t)$ function and show time windows in which the system switches to non-periodic solutions. As once can observe, both methods work simi-

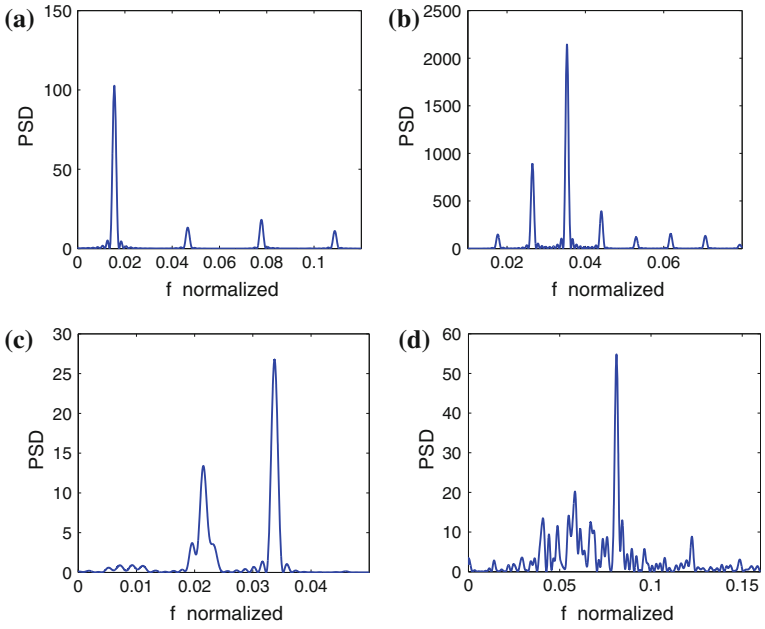

Fig. 4.26 Examples of periodograms calculated for the mathematical systems: **a** BS A X1 N0, **b** W B X3 N10, **c** DvP B X1 N10, and **d** BS C X3 N10

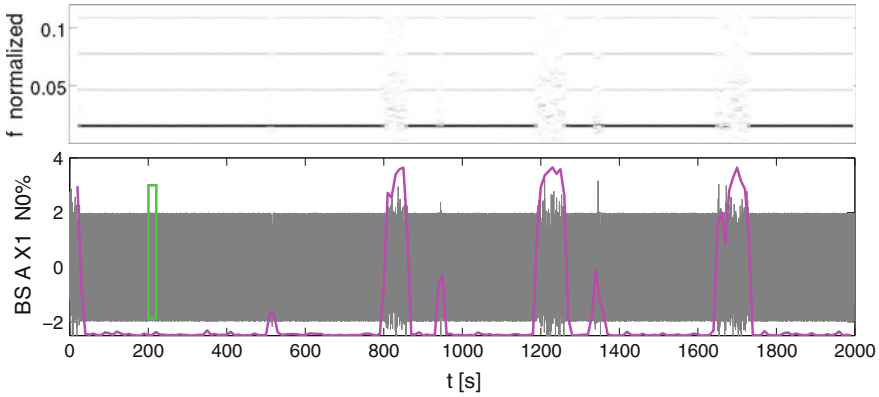

Fig. 4.27 The RD-P (RD in Magenta) comparison of the time series type-A of the Burke–Shaw system without added noise

larly, but some D_{RR} peaks indicate signal changes that are difficult to notice on the periodogram map (about 500, 950, and 1350 s). In Fig. 4.28 both methods are applied to the W B X3 signal with a 10 % noise level. The signal is characterized by frequent intermittent hops between different solutions. The results of both methods are consistent. In the subsequently analyzed case we compare the results of both methods for

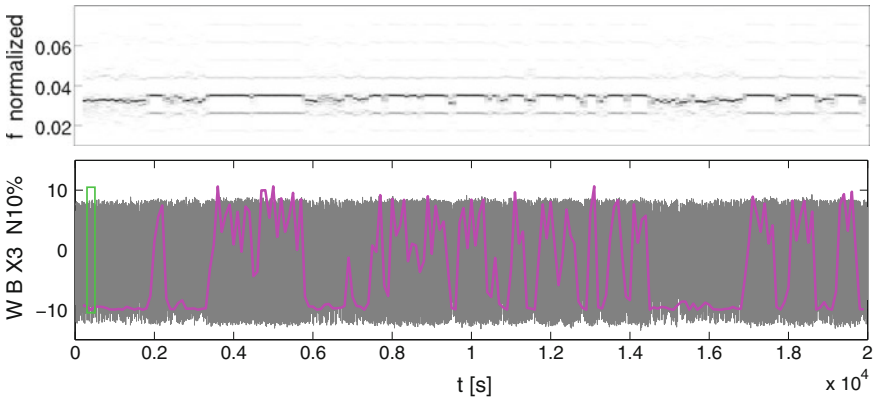

Fig. 4.28 The RD-P (RD in Magenta) comparison of the time series type-B of the Windmi system with a 10 % noise level

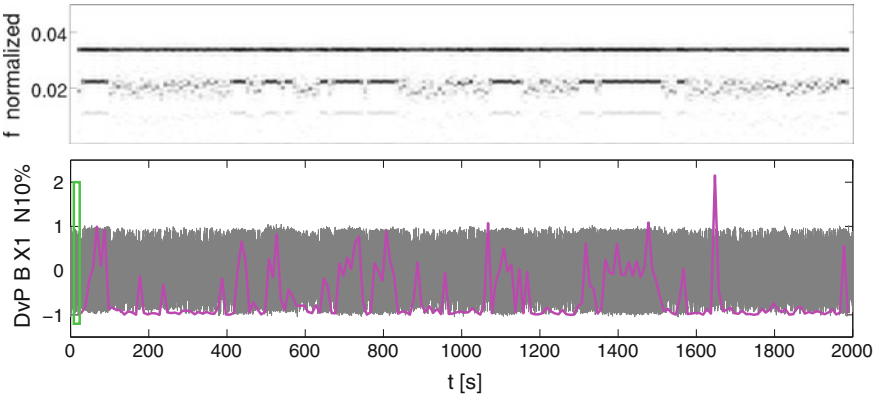

Fig. 4.29 The RD-P (RD in Magenta) comparison of the time series type-B of the Duffing–van der Pol system with a 10 % noise level

the signal DvP B X1 with a 10 % noise level (Fig. 4.29). As in the previous case, the signal is characterized by intermittent hopping between periodic and non-periodic solutions. Both methods give similar results. However, we can observe the presence several of the $D_{RR}(t)$ function peaks detecting windows with signal changes which are invisible on the periodogram map. For example, there are two peaks in the range (100–300 s), two peaks in the range (850–1000 s) and the highest peak of the DR function occurring around t = 1650 s. In this case, the RD analysis detects more changes in the signal state than is shown by the periodogram map. The last discussed comparison of the two methods pertains to the results of analysis for the signal BS C X3 with a 10 % noise level. For the type-C non-periodic signal, rare and narrow periodic windows are practically unnoticeable on the periodogram map, while the DR function identifies them pointing to clear peaks at these moments of

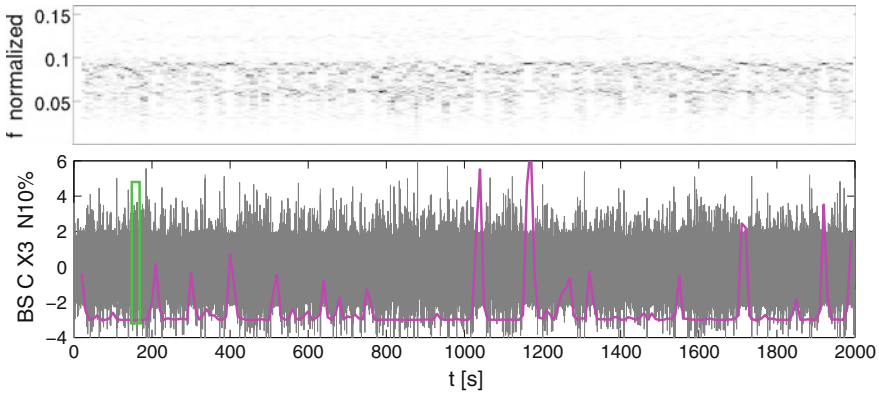

Fig. 4.30 The RD-P (RD in Magenta) comparison of the time series type-C of the Burke–Shaw system with a 10 % noise level

time (Fig. 4.30). Concluding the above, it can be said that both the RD and periodogram method can be applied to test stationarity of the time series. However, on the basis of this comparison, we can conclude that the results of the RD are more legible and that this method detects more signal changes. This is primarily due to the fact that the concept of recurrensivity includes but is not limited to periodicity.

4.8 Conclusions

The results demonstrates effectiveness of the proposed method in analysis of chaotic non-stationary signals. The method successfully detects different events, including chaotic and periodic windows, changes in the nature of the signal, signal modifications and short-term disturbances. The application of this method to study signals disturbed by environmental noise produces different results for different chaotic systems. This study showed that the RD method detects chaotic windows in signals with even 80 % noise level in the case of the Burke–Shaw system, while ceases to be effective for the Duffing–van der Pol system where the noise level is 20 %. On this basis it can be concluded that the recurrence dissimilarity between chaotic and periodic fragments has a completely different level in the two nonlinear systems. The usefulness of the RD measure has been demonstrated in Sect. 4.7 by comparing its results with periodoram maps. We can assume that following some improvements the RD measure will be useful for studying basic problems related to nonlinear signals such as chaos, order, periodicity and noise level. The RD results greatly depend on the reference window choice and the characteristics of a signal in a selected period of time. Other parameters also affect the quality of the collected data. To improve the method such that it could be used to study general characteristics of nonlinear signals, we must first solve the problem pertaining to selection of a reference window.

We claim that the most important improvement is to development a technique for artificial creation of a reference signal based on a preliminary analysis of the examined time series. This should render RD measurement results more universal and comparable.

The software for determining the RD measure will be prepared and made available to interested researchers on request.

Acknowledgments The research was funded by the National Science Centre, Poland—the number of decision: DEC-2013/09/B/ST8/02850. The authors express their gratitude to the Department of Cardiology, John Paul II Hospital in Zamość and to the Clinic of Cardiology, Hospital SPSK 4 in Lublin for providing the ECG data.

References

1. G. Litak, A. Syta, J. Gajewski, J. Jonak, Meccanica **45**, 603 (2009)
2. Y. Pomeau, P. Mannevielle, Commun. Math. Phys. **74**, 189 (1981)
3. E. Ott, *Chaos in dynamical systems* (Cambridge University Press, Cambridge, 2002)
4. J.P. Zbilut, A. Giuliani, C.L. Webber Jr., Phys. Lett. A **246**, 122 (1998)
5. K. Sternickel, Comput. Methods Progr. Biomed. **68**, 109 (2002)
6. K. Urbanowicz, J.J. Zebrowski, R. Baranowski, J.A. Holyst, Physica A **384**, 439 (2007)
7. J.P. Zbilut, C.L. Webber Jr, M. Zak, Quantification of heart rate variability using methods derived from nonlinear dynamics Analysis and Assessment of Cardiovascular Function, Chapter 19, ed. by G. Drzewiecki, J.K.-J. Li (Springer, New York, 1998), pp. 324–334
8. O. Durin, C. Pedrinazzi, G. Donato, R. Pizzi, G. Inama, Ann. Noninvasive Electrocardiol. **13**, 219 (2008)
9. J.P. Zbilut, N. Thomasson, C.L. Webber Jr., Med. Eng. Phys. **24**, 53 (2002)
10. Y. Zhu, X. Yang, Z. Wang, Y. Peng, IEEE Trans. Biomed. Eng. **61**(3), 938 (2014)
11. K.A. Triplett, S.M. Ghiaasiaan, S.I. Abdel-Khalik, D.L. Sadowski, Int. J. Multiphase Flow **25**, 377 (1999)
12. A. Serizawa, Z. Feng, Z. Kawara, Exp. Thermal Fluid Sci. **26**, 703 (2002)
13. G. Gorski, G. Litak, R. Mosdorf, A. Rysak, Eur. Phys. J. B **88**, 239 (2015)
14. G. Gorski, G. Litak, R. Mosdorf, A. Rysak, Zeitschrift fur Naturforschung A **70**, 843 (2015)
15. G. Litak, A. Syta, R. Rusinek, Int. J. Adv. Manuf. Technol. **56**, 445 (2011)
16. A.K. Sen, G. Litak, K.D. Edwards, C.E.A. Finney, C.S. Daw, R.M. Wagner, Appl. Energy **88**, 1649 (2011)
17. A.K. Sen, G. Litak, A. Syta, R. Rusinek, Meccanica **48**, 738 (2013)
18. R. Shaw, Zeitschrift fur Naturforsch A **36**, 80 (1981)
19. G. Duffing, Erzwungene Schwingungen bei Veränderlicher Eigenfrequenz (F. Vieweg u. Sohn, Braunschweig, 1918)
20. B. van der Pol, Radio Rev. **1**, 701 (1920), **1**, 754 (1920)
21. W. Horton, I. Doxas, J. Geophys. Res. A **103**(A3), 4561 (1998)
22. J.C. Sprott, *Chaos and Time-Series Analysis* (Oxford University Press, New York, NY, USA, 2003)
23. J.-P. Eckmann, S.O. Kamphorst, D. Ruelle, Europhys. Lett. **5**, 973 (1987)
24. C.L. Webber Jr., J.P. Zbilut, J. Appl. Physiol. **76**, 965 (1994)
25. M. Casdagli, J. Royal Stat. Soc. **54**, 303 (1992)
26. N. Marwan, Encounters with Neighbours: Current Development of Concepts Based on Recurrence Plots and their Applications, Ph.D. Thesis, Universitaet Potsdam (2003)
27. N. Marwan, M.C. Romano, M. Thiel, J. Kurths, Phys. Rep. **438**, 237 (2007)
28. F. Takens, Detecting Strange Attractors in Turbulence Dynamical Systems and Turbulence, Lecture Notes in Mathematics (1981)

Chapter 5
New Insights for Testing Linearity and Complexity with Surrogates: A Recurrence Plot Approach

A. Carrión and R. Miralles

Abstract The detection and characterization of non-linearities in temporal series is a hot topic in some disciplines such as nondestructive testing of materials, bioacoustics and biomedical research domains. This is a complex interdisciplinary field where many different researchers are striving to achieve better and more sophisticated techniques. In this scenario, the search for new perspectives that can explain and unify some of the theories is of key importance. Recurrence Plots (RPs) and Recurrence Quantification Analysis (RQA) can play such a role. In this work, we show how RPs can be used to design tests for non-linear detection and characterization of complexity. The proposed tests are less parameter dependent and more robust than some of the traditional discriminating measures. We also illustrate the applicability of the proposed algorithms in simulations and real-world signals such as the analysis of anomalies in the voice production of mammals.

5.1 Introduction

Detecting non-linearities and complexity in time signals can be used in many situations as an indicator of changes in the underlying dynamical system that is responsible for the generation of those signals. In some disciplines, the study of these phenomena has not been addressed and it is a common practice to model these processes using suboptimal, but mathematically, tractable models. However, an appropriate detection and characterization of the non-linear and deterministic nature of the signal can convey important information in a large number of situations such as: early detection of epileptic symptoms with EEG signals [1], non-linear phenomena in

A. Carrión · R. Miralles (✉)
Instituto de Telecomunicación y Aplicaciones Multimedia (iTEAM),
Universitat Politècnica de València, Camino de Vera S/N, 46022 Valencia, Spain
e-mail: rmiralle@dcom.upv.es

A. Carrión
e-mail: alcarga4@upv.es

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8_5

mammalian voice production [2], stock market predictability [3], river flow discharge rates [4], etc.

Many authors have worked on different techniques to detect and characterize non-linearities in time series. One of the most commonly used methods is the Monte-Carlo approach, namely the surrogate data bootstrapping method. This approach is based on the computation of an ensemble of surrogates which are representative realizations of the null hypothesis under study. A statistical measure is computed for the original time series and the surrogates. If the statistic is significantly different from the values obtained for the surrogate set, the null hypothesis can be rejected. Therefore, there are three major aspects of the surrogate data method that need to be considered: (1) the exact definition of the null hypothesis, (2) the realization of the null hypothesis, and (3) the test statistic.

In the literature, different kinds of surrogate generation algorithms can be found based on the null hypothesis under study: stationarity/non-stationarity [5], determinism/randomness [6], linearity/non-linearity [7, 8], chaos [9], etc. A surrogate data generation algorithm for testing fluctuations and trends in data is the small-shuffle surrogate algorithm (SSS) [10], surrogate data generation algorithms for testing pseudo-periodic or oscillating time series are the pseudoperiodic surrogates (PPS) [11] and the twin surrogates (TS) [12, 13], and surrogate data generation algorithms for testing linearity include the well-known Amplitude Adjusted Fourier Transform (AAFT) [7] and its improved version, the iterative Amplitude Adjusted Fourier Transform (iAAFT) [14]. In each analysis, different statistical measures are applied to quantify the differences between the original data and the surrogates. Some of the methods applied in linear analysis are Kaplans δ-ε [15], Deterministic Versus Stochastic plots [16], or Delay Vector Variance (DVV) [17].

The use of the RP may have an important role in the signal modality characterization framework. In this work, we analyze how this visualization tool not only can be applied for the generation of surrogate data but also for the quantification of statistical differences between the original data and the surrogate data. We have mainly focused on two main features of signal modality: the characterization of the linear/non-linear nature of a signal and its complexity.

To explain these ideas, the following definitions are used throughout this work. We can compute the phase-space representation of a given N-point signal x_t for an embedding dimension (m) and a time lag (τ) by computing the Delay Vectors (DVs) using (5.1) [18]:

$$\mathbf{x}(i) = [x_{i-m\cdot\tau}, x_{i-(m-1)\cdot\tau}, \ldots, x_{i-\tau}] \tag{5.1}$$

Every delay vector has a corresponding target which is basically the next sample x_i [19]. The proper selection of τ and m is crucial in the following analysis since it

affects the correct representation of the data evolution in time. A common approach to determine the time lag is the one proposed by [20], which uses the first local minimum of the time-delayed mutual information as a reasonable value for τ. The selection of the minimum embedding dimension m is based on the false nearest neighbor algorithm proposed in [21, 22].

RPs are calculated using (5.2) [23]:

$$RP_{i,j} = \begin{cases} 1 & \text{if } ||\mathbf{x}(i) - \mathbf{x}(j)|| \leq \varepsilon \\ 0 & \text{otherwise} \end{cases} \tag{5.2}$$

where $||\cdot||$ is the Euclidean norm and the parameter ε is a threshold distance (or recurrence threshold). The quantification of RPs can be done by means of RQA.

The remainder of this work is structured as follows. In Sect. 5.2, we present a selection of surrogate techniques that can be applied to test the linear/non-linear nature of real signals that have a non-stationary behaviour. In this section, we also present a new surrogate technique to test the presence of high complexity in oscillatory signals. In Sect. 5.3, we illustrate how new discriminating measures for non-linear statistical tests can be designed using RQA. Section 5.4 presents some examples of signal modality characterization with both simulated and real signals. Finally, we present our conclusions in Sect. 5.5.

5.2 Surrogate Techniques

Theiler et al. [7] introduced the concept of 'surrogate data', which has been extensively used in the context of statistical non-linearity testing. A surrogate time series (or surrogate for short) is generated as a realization of the null hypothesis under study. Thus, given an original signal, realizations of this data must be generated by modifying only the desired characteristic of the signal that is being tested, while the rest of the characteristics remain the same. Care must be taken when using surrogate data to ensure that the statistical differences come from the desired characteristic and not from an undesired one, such as a failure of the surrogate algorithm to mimic non-stationary data. Choosing a surrogate technique that does not mimic these fluctuations or that changes the statistical distribution may lead to false positives.

Below we next present two techniques for the generation of surrogates that will be used later in a statistical analysis based on RQA. The first technique is used to detect the presence of non-linearities; whereas the second technique uses RP concepts to generate surrogates that are valid for testing high complexity in short oscillatory signals.

5.2.1 Surrogates for Testing Non-linearity

One of the key issues in signal processing is the definition of a linear signal. The standard definition is that it is a signal that is generated by a Gaussian linear stochastic process. Based on this definition, the null hypothesis refers to the properties of any AR-model that is driven by white noise. Since any AR-model can be described by its amplitude spectrum (and thus the phase spectrum is irrelevant), most surrogate data generation algorithms for testing linearity are based on a phase spectrum randomization. The most common established method for generating constrained surrogates is the Iterative Amplitude Adjusted Fourier Transform (iAAFT) [14].

Let x_t be the original time series, s_k the sorted version of x_t, and X_k the Fourier transform of the original data series. The original iAAFT algorithm is based on the following steps:

1. Make a random permutation of the time samples of the original time series x_t, namely r_t.
1. Compute the phase spectrum of r_t, namely ϕ_k.
3. Compute the Inverse Fourier Transform of $\{|X_k| \cdot exp(j\phi_k)\}$, namely c_t.
4. Obtain a new version of the time series r_t by rank-ordering (sort in increasing order) c_t so as to match s_k.
5. Repeat steps 2–4 until the discrepancy between $|X_k|$ and the amplitude spectrum of r_t is lower than a chosen tolerance.

This iterative algorithm has been shown to converge after a finite number of steps [14]. Each initial random permutation gives a different output surrogate data r_t with identical signal distribution and approximately identical amplitude spectra as the original signal.

The original iAAFT method has recently been refined to not only retain the signal distribution and amplitude spectrum of the original time series but also to retain the local mean and variance of the original time series [24]. This new approach uses a wavelet transform to preserve the behavior in the time-frequency plane. It makes this new technique very suitable for non-stationary signals whose time-changing properties would be destroyed using the original iAAFT algorithm.

Let x_t be the N-point original time series and s_k the sorted version; x_t is decomposed in J scales, where $N = 2^J$, $j = 1, 2, \ldots, J$. The Wavelet Iterative Amplitude Adjusted Fourier Transform (WiAAFT) algorithm can be summarized as follows:

1. Compute the Maximal Overlap Discrete Wavelet Transform (MODWT) of the original time series x_t using a high number of vanishing moments to deal with any potential non-stationarity in the series [25]. The result is a vector of wavelet detail coefficients, c_j, at each scale j.
2. Apply the iAAFT algorithm to each c_j to generate a constrained realization of the original detail coefficients, c'_j, preserving the original values and their periodicity.
3. Transpose c'_j so that the first detail coefficient in the transposed case is the last one in the new variant, c''_j.
4. Find the best match in each scale between c_j and the two variants, c'_j and c''_j, by circularly rotating until an error function is minimized. In this work, least-squares algorithm is used. This means that the positions with high energies in the original data are mimicked in the surrogates.
5. Invert the MODWT (using the original approximation coefficients) to yield a surrogate dataset, w_t.
6. Perform rank-ordering (sort in increasing order) to w_t so as to match s_t.
7. Use the new time series w_t as the initialization of the described iAAFT algorithm.

Figure 5.1 compares the surrogate data computed with the original iAAFT algorithm and the new WiAAFT for a heart-rate variability signal (HRV) that is recorded during a meditation session. HRV signals are widely used to analyze human health; however, from the signal processing point of view, one of the main drawbacks is the presence of many artifacts or temporal changes in the recorded time series due to different factors such as patient motion, eye blinking, etc. that may lead to non-stationarities. The blue line represents the original time series. The green line corresponds to the surrogate data that is computed using the iAAFT. Note that the iAAFT has almost destroyed the temporal structure of the original signal. The red line plots the surrogate data that is computed with the WiAAFT algorithm. Even though, this new surrogate data is still a random realization compared with the original data, it greatly preserves the time evolution of the original data. The WiAAFT algorithm has been observed to yield superior results when testing for linearity (Fig. 5.1).

5.2.2 *Pseudo-Periodic Twin Surrogates (PPTS)*

Oscillatory signals can often be found in the solution of real-world problems. Sometimes these signals have a high complexity in the oscillations, which may indicate the presence of non-linearities. Some examples are: the study of biphonation, subharmonics or other pathologies in animal sounds [26], the study of non-linearities in ultrasonic signals [27], etc. In most of these problems, conventional surrogates have

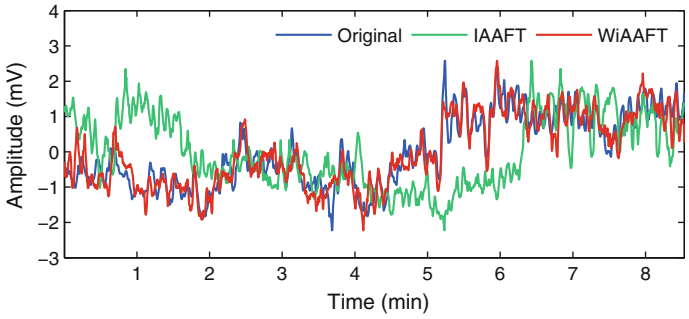

Fig. 5.1 Comparison between two surrogate data generation algorithms (iAAFT and WiAAFT) for a heart-rate variability signal (HRV) recorded during a meditation session. The *blue line* represents the original non-stationary signal; the *green line* represents the surrogate data computed with the iAAFT; and the *red line* corresponds to the surrogate data computed with the WiAAFT

a limited use, mainly due to the finite duration of the events that need to be analyzed and due to the sensitivity that needs to be achieved. Some of the techniques that rely on the Fourier Transform require long data series, and therefore, specific algorithms for the generation of surrogates in short-length oscillating data have to be designed.

In the literature, there are several methods that preserve the higher order moments for the surrogate generation of oscillatory signals. The most representative methods for this work are the pseudoperiodic surrogates (PPS) method [11], the twin surrogates (TS) method [12, 13], and the pseudoperiodic twin surrogate (PPTS) method [28]. The PPTS method is a combination of the PPS and the TS which allows tests to be designed in order to check the null hypothesis that the observed time series is a quasi-periodic signal obtained as the sum of sinusoids with incommensurate frequencies.

The PPTS algorithm uses the phase space and RP concepts to obtain the surrogates. This is achieved by employing the idea of jumping among twin points in the same way that the TS algorithm does. Twin points are points that are neighbors, $||\mathbf{x}(i) - \mathbf{x}(j)|| < \varepsilon_{TP}$, and they also share the same neighborhood $R_{i,l} = R_{j,l}\ \forall l$. Twin points are indistinguishable with respect to their neighborhoods, but they generally have different pasts and, more importantly, different futures. We can generate surrogates by changing the structures in the RP consistently with those produced by the underlying dynamic system. Jumping among twin points produces surrogates with RP representations that are very similar to the RP representation of the original signal if the time signal is periodic or quasi-periodic (similar futures). In contrast, jumping among twin points produces surrogates with RP representations that are quite different if the time signal is chaotic.

Unfortunately, jumping among twin points is not always enough to generate surrogates that allow the confidence level for null test rejection to be established. Additionally, there are some practical implementation problems with the TS algorithm for short time series (such us the limited number of twins [13]). Thus, the PPTS algorithm uses a second randomization technique. This technique consists of mov-

ing from point to point in the phase space in accordance with a probability that is inversely proportional to the distance between the two points (using (5.3) as is done in the PPS algorithm [11]). The proposed PPTS algorithm is summarized as follows:

1. Compute the RP of the original signal using (5.2), with an appropriate choice of ε, denoted by ε_{TP}, and identify the twin points ($R_{i,l} = R_{j,l}\ \forall l$). The choice of ε_{TP} is not crucial for the PPTS; it has been shown in [12] that a choice of ε_{TP} corresponding to 5–10 % of black points in the RP is appropriate.
2. Randomly choose an initial condition i_0 and make $i = i_0$. Initialize $n = 1$.
3. If there is a twin point for $\mathbf{x}(i)$, make the next point of the surrogate $\mathbf{x_s}(n) = \mathbf{x}(j)$, where j is randomly chosen among the twin points with the probability $1/T$ (T is the number of twin points for the state $\mathbf{x}(i)$). Let $i = j$ and $n = n + 1$.
4. For $\mathbf{x}(i)$, choose a neighbour $\mathbf{x}(j)$ from all of the elements of the phase space representation ($j = m \cdot \tau, \ldots, N - 1$) with probability

$$Prob(\mathbf{x}(j)) \propto exp\frac{-||\mathbf{x}(i) - \mathbf{x}(j)||}{\rho} \tag{5.3}$$

   where ρ is the noise radius studied in [11]. Make the next point of the surrogate $\mathbf{x_s}(n) = \mathbf{x}(j)$. Let $i = j$ and $n = n + 1$.
5. Repeat from Step 3 until $n = N$.

The surrogate is formed from the first scalar component of $\mathbf{x_s}(n)$.

The proposed PPTS algorithm generates surrogates that are very similar to the original signal as long as the original signal is periodic or quasi-periodic. When the original signal deviates from a periodic or quasi-periodic oscillation, the PPTS generates surrogates that have a RP matrix that is quite different while still preserving the approximate phase-space shape of the original signal. Thus, these surrogates are appropriate for testing the null hypothesis that the observed time series is consistent with a quasi-periodic orbit. This is illustrated in the following example.

5.2.2.1 Example

Consider the following signals: a quasi-periodic time series composed of the sum of two sinusoids with incommensurate frequencies and a Rössler chaotic time series [29].

The quasi-periodic signal was generated as given by:

$$x(t) = 8 \cdot sin(2 \cdot \pi \cdot f_1 \cdot t) + 4 \cdot sin(2 \cdot \pi \cdot f_2 \cdot t) \tag{5.4}$$

where $f_1 = \sqrt{3}$ Hz, $f_2 = \sqrt{5}$ Hz. The discrete time series version x_n^{quasi} is obtained using a sample period of $\Delta t = 0.02$ s. We only generate 400 samples of this signal.

The Rössler signal is generated as given by:

$$\begin{cases} \dot{x} = -y - z \\ \dot{y} = x + ay \\ \dot{z} = b + z(x - c) \end{cases} \tag{5.5}$$

with the initial conditions $x(0) = y(0) = z(0) = 0.1, a = 0.2, b = 0.2, c = 5.7$ (chaotic state) and a sampling time of $\Delta t_s = 0.2$. The system was integrated 1400 times using the Matlab ODE solver ODE45, and the time series $x_n^{Rossler}$ was obtained from the x component after discarding the first 1000 data points to avoid transient states (again, only 400 data points were used).

Figure 5.2 illustrates the phase-space reconstruction (with $m = 2$ and $\tau = 6$) of the following: (a) the original Rössler time series, and (b) the PPTS. From the comparison of (a) and (b), it can be observed how the proposed PPTS algorithm achieves a phase-space representation that removes most of the details that indicate chaos (the large number of trajectories that run arbitrarily close together) while at the same time preserving the shape of the original time series. In the same way, the panels (c) and (d) show the phase-space reconstruction of the original quasi-periodic time series and the phase-space reconstruction of the the PPTS of the quasi-periodic signal. In this situation both the phase-space diagram of the quasi-periodic and its PPTS look very similar.

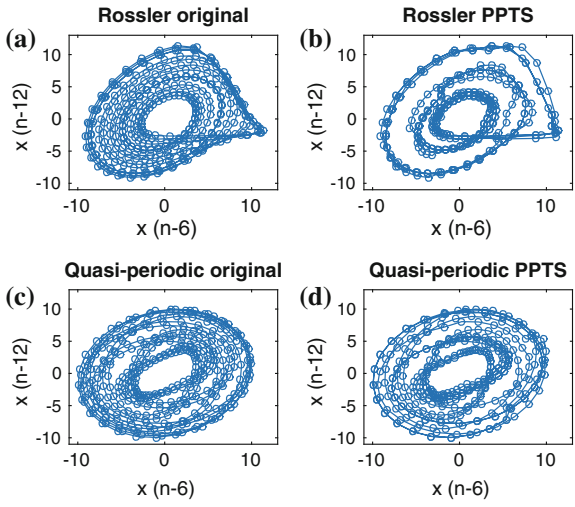

Fig. 5.2 **a** Phase-space reconstruction of the Rössler attractor; **b** phase space reconstruction of the PPTS for the Rössler attractor; **c** phase space reconstruction of the quasi-periodic time series; **d** phase space reconstruction of the PPTS for the quasi-periodic time series. The PPTS were computed with ε_{TP} corresponding to 10 % of *black* points in the RP and a $\rho = 0.25$

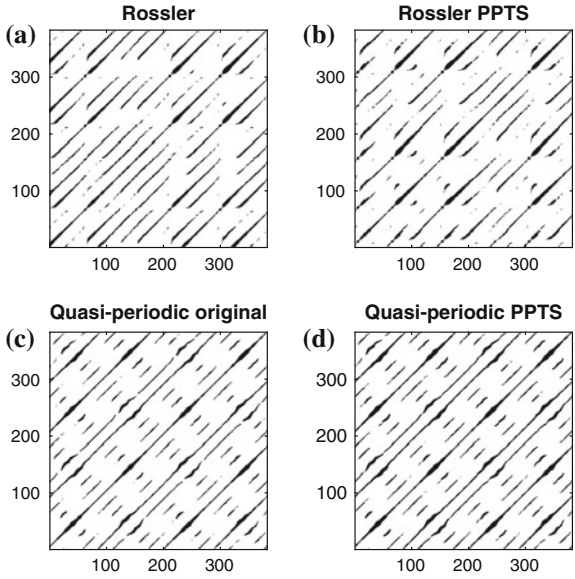

Fig. 5.3 **a** RP representation of the Rössler attractor; **b** RP representation of the PPTS for the Rössler attractor; **c** RP representation of the quasi-periodic time series; **d** RP representation of the PPTS for the quasi-periodic time series. The PPTS were computed with ε_{TP} corresponding to 10 % of *black* points in the RP and a $\rho = 0.25$

All of these details are captured and represented in the structures of the RPs (see Fig. 5.3). The appropriate metrics for testing using the PPTS (or some of the other surrogates presented) can then be obtained by means of the RQA. This idea, which is detailed in the following section, allows us to create tests to distinguish between oscillatory signals with different complex patterns of amplitude modulation.

5.3 Discriminating Non-linearity Tests and Statistics Based on Recurrence Plots

In order to test the null hypothesis of linearity, a statistical discriminating measure has to be performed on both the original data and the surrogates. Many different techniques and statistical tests have been suggested for this purpose. Some of the discriminating statistics are based on computing Lyapunov exponents, return maps, or some other graphs or functions that are representative of the topology of the underlying dynamics. The computation of these functions can be quite complex in real-world signals. As a result of this difficulty, a large number of techniques that compute a much more simple graphical representation of the underlying system have been developed. Some of the most cited methods are Deterministic versus Stochastic plots [16], Kaplans δ-ε [15], correlation exponent [30], or Delay Vector Variance (DVV) [17].

All of these methods introduce the DV $\mathbf{x}_i$ presented above (5.1), as well as the concept of the target of the DV x_i (basically, the next sample). The idea underpin-

ning these methods is that the study of the locality of the unknown models (which maps the DVs onto their corresponding targets) when combined with the method of the surrogates provides information of the non-linear behaviour of the underlying process. The degree of locality of a time series is closely related to the distribution of the nearest neighbour points; however, each method analyzes the degree of locality using different approaches: the mean of the targets, the variance of the targets, the prediction error, etc. Even though these studies are very thorough, none of them exploit the advantages of RP, and thus, the advantages of RQA.

RPs can help greatly in the design of robust and less parameter dependent tests for non-linearity detection. In the following, a detailed analysis of one of the above-mentioned methods is done in order to further understand the motivation of this new approach and its potential.

5.3.1 Reformulation of the DVV Using RPs

The delay vector variance (DVV) method is a phase-space based technique that examines the deterministic nature of a time series and that provides information of the non-linear behavior of the underlying process when it is combined with the method of surrogate data.

The DVV method can be summarized as follows [17].

1. Compute the optimal embedding parameters, m and τ, and generate the delay vectors (DVs) using (5.1). Every DV, $\mathbf{x}(i)$, has a corresponding target, namely the next sample x_i.
2. Compute the mean, μ_d, and the standard deviation, σ_d, over all pairwise Euclidean distances between DVs, $||\mathbf{x}(i) - \mathbf{x}(j)||(i \neq j)$.
3. Generate the sets $\Omega_k(r_d)$ such that $\Omega_k(r_d) = \left\{\mathbf{x}(i) \mid ||\mathbf{x}(k) - \mathbf{x}(i)|| \leq r_d\right\}$ (sets that consist of all DVs that lie closer to $\mathbf{x}(k)$ than a certain distance r_d, taken from the interval $\left[max\left\{0, \mu_d - n_d\sigma_d\right\}\right]$). For example, generate N_{tv} uniformly spaced distances, where n_d is a parameter controlling the span over which to perform the DVV analysis.
4. Compute the variance of the corresponding targets, $\sigma_k^2(r_d)$ for every set $\Omega_k(r_d)$. The average over all sets $\Omega_k(r_d)$, normalized by the variance of the time series, σ_k, yields the target variance, $\sigma^{*2}(r_d)$:

$$\sigma^{*2}(r_d) = \frac{\frac{1}{N_0}\sum_{k=1}^{N_0}\sigma_k^2(r_d)}{\sigma_x^2}, \tag{5.6}$$

   where N_0 denotes the total number of sets $\Omega_k(r_d)$.
5. Repeat steps 1–4 for the N_s surrogates.

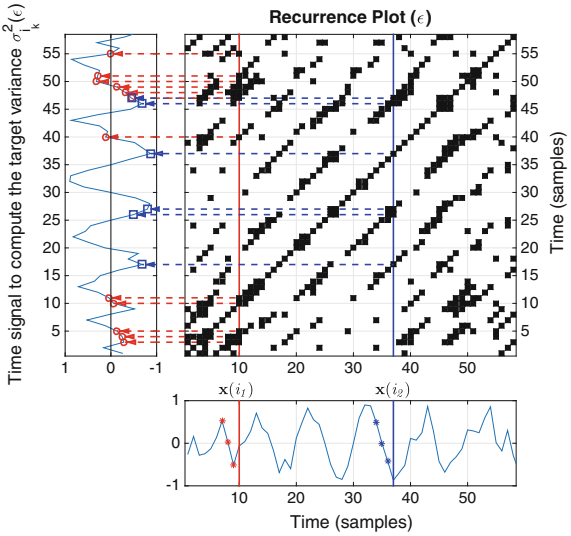

Fig. 5.4 Example of how to find the nearest DVs and the normalized variance of its targets using RP(ϵ). The *red* and *blue vertical lines* are two randomly chosen DVs ($\mathbf{x}(i_1)$ and $\mathbf{x}(i_2)$ using $m = 3$, $\tau = 1$) represented in the *bottom panel* using *asterisks*. The *red* and *blue horizontal dashed lines* are their corresponding neighbours that lie within the distance ε. The *red circles* in the *left panel* are the corresponding targets of $\mathbf{x}(i_1)$; the *blue squares* are the corresponding targets of $\mathbf{x}(i_2)$. The RP was computed for an ε corresponding to 14 % of *black* points

Due to the standardization of the intervals of r_d, the DVV analysis can be conveniently illustrated in the resulting DVV plots, where the x-axis corresponds to the standardized distance r_d and the y-axis corresponds to the target variance. If the surrogate data yields similar results to that of the original signal, the target variance of the original signal falls within the confidence interval and the null hypothesis can not be rejected.

Those readers familiar with RPs can readily understand that the computation of the DVV plot can be explained by using RP concepts. The RP is a 2D plot that for a given moment in time shows the times at which a phase-space trajectory visits roughly the same area in the phase space.

Thus, for a given DV, using the RP, we can easily find the DVs that are closer than a given distance r_d just by looking in the corresponding column of the RP(ε). The distance r_d is completely equivalent to the so-called ε in the RP (5.2). Figure 5.4 illustrates this idea and shows how we can obtain the corresponding set of targets (horizontal red and blue lines), $\Omega_k(r_d)$ in the DVV algorithm, for the later computation of its variance. Thus, for a given i_k and ε, we obtain $l_\varepsilon \subseteq \{j\}$ s.t. $R_{i_k j} = 1$, and $\sigma^2_{i_k}(\varepsilon) = VAR[x_{l_\varepsilon}]$, where $VAR[\cdot]$ is a variance estimator and x_{l_ε} is the target of the DV $\mathbf{x}(l_\varepsilon)$. Averaging this variance for N_0 different DVs and dividing by the variance of the time series (σ^2_x), we obtain (5.7) which is clearly equivalent to (5.6). Note that

the plot of this normalized variance $\sigma^{*2}(\varepsilon)$ as a function of the standardized distance is the DVV plot.

$$\sigma^{*2}(\varepsilon) = \frac{\frac{1}{N_0}\sum_{k=1}^{N_0}\sigma_{i_k}^2(\varepsilon)}{\sigma_x^2} \tag{5.7}$$

The analysis of target variance based on a distance r_d in the DVV algorithm, or a distance ε in the RP, is an indicator of the recurrence behavior of the studied phase space. This is not the only way to quantify the number and duration of recurrences of a dynamic system represented by its phase-space trajectory. A proper selection of a RQA measure allows an equivalent analysis that is less parameter dependent and that has greater robustness.

5.3.2 RQA Measures as Discriminating Statistics

The recurrence plots exhibit characteristic large-scale and small-scale patterns that are caused by typical dynamic behaviour. The appearance of diagonal lines is related to a similar local evolution of different parts of the trajectory, whereas the appearance of vertical and horizontal black lines means that the trajectory does not change its state for some time [31]. Zbilut and Webber developed Recurrence Quantification Analysis (RQA) to quantify a RP [32]. They define measures using the recurrence point density and the diagonal structures in the recurrence plot: the recurrence rate (RR), the determinism (DET), the maximal length of the diagonal structures, the entropy and the trend (see Table 5.1). The recurrent points delineate the number of embedded vector pairs that are near each other in the m-dimensional; however, the measures related to the diagonal patterns distinguish between points individually dispersed and those that represent parts of the signal where the signal similarly evolves. Gao and Cai [33, 34] carried out a detailed analysis about the reasons for the appearance of vertical lines (and horizontal lines for fixed values of ε) and its relation with the appearance of square-like textures in RPs. The underlying reasons for the abundance of vertical lines (and therefore square-like structures) are: (i) the usage of a high sampling frequency (i.e., a small sampling time); and (ii) the usage of a fairly large recurrence radius ε for constructing a RP. Based on this analysis, some recurrence time statistics corresponding to vertical structures were introduced. Marwan et al. [31, 35] extended this view on the vertical structures and defined measures of complexity based on the distribution of the vertical line length: laminarity (LAM), trapping time (TT), and the maximal length of the vertical structures (see Table 5.1). Those variables have widely been applied to identify laminar states and their transitions between regular and chaotic regimes [36], as well as to detect the presence of unstable singularities which are often found in biological dynamics [31, 37].

The main purpose of the statistical test within the surrogate data method is to be sensitive to any changes that are exclusively related to the null hypothesis under

study. For that reason, it is necessary to point out the need to understand the signal modality characteristic being studied and its relationship with the applied surrogate data generation algorithm.

As discussed in Sect. 5.2.1, the hypothesis of non-linearity is linked to the presence of information in the phase spectrum. The randomization of the phase spectrum leads to significant changes in the structure of the signal. Even though the surrogates of a non-linear signal (using the WiAAFT algorithm) preserve the probability distribution and the spectrum amplitude (and therefore the autocorrelation function) of the original signal, they have a new arrangement of closest points. The DVV algorithm detects these changes through the variance of the targets of the closest points. An equivalent measure of RQA that is susceptible to the distribution of the closest points is the Trapping Time (TT), which is a statistic that is linked to the laminarity of a dynamic system. The study of the TT measure as a function of the recurrence threshold ε, $TT(\varepsilon)$, allows the evolution of recurrent states to be studied as the percentage of nearby points increases. In the case of linear signals, the statistic vector computed for the surrogate data will coincide with the vector TT (ε) of the original signal. In the case of the non-linear data will not coincide, and this permits the non linear signals to be identified.

A similar argument can be made regarding the hypothesis of complexity using the PPTS algorithm (Sect. 5.2.2) for oscillatory signals. When generating the surrogates, jumping among twin points breaks the diagonals when the signals deviate from a quasi-periodic oscillation. This produces surrogates with different diagonal line lengths in the RP for temporal series having high complexity while maintaining approximately the same diagonal line lengths in quasi-periodic ones. An equivalent measure of RQA that is susceptible to the diagonal line length distribution is the average diagonal line length L, or any other measure related to the diagonal line length (see Table 5.1). The study of L as a function of the recurrence threshold ε, $L(\varepsilon)$, and the comparison with that of its surrogates allows the complexity to be identified.

Note that in both analyses, the phase-space reconstruction for the original and the surrogates data must be done with the same embedding parameters, τ and m. Moreover, the statistics compared (the original signal and the surrogates) need to be computed with the same percentage of black points in the RP matrix. In this work, a sweep between 10 and 80 % of the black points has been performed. This new representation (herein called DVRQA, Delay Vector Recurrence Quantification Analysis) eliminates the problem of choosing the most appropriate threshold for computing the RP in an unknown situation. Unlike the technique described above, which only takes into account some randomly chosen points, one of the advantages of statistical analysis based on RQA is that it uses the entire signal. This is supposed to be a more robust technique that is less parameter dependent and has smaller resultant confidence intervals. In addition, this technique eliminates the need for an estimator, such as the variance, that needs a high minimum number of points for a correct estimation.

The DVRQA analysis can be summarized as follows:

Table 5.1 Some of the possible measures for the recurrence quantification analysis (RQA) [31, 32, 35]

Symbol	Description	Equation	
RR	Recurrence rate: density of recurrence points	$\frac{1}{N^2}\sum_{i,j=1}^{N} R_{i,j}$	
TT	Averaged vertical line length[a]	$\frac{\sum_{v=v_{min}}^{N} v \cdot P(v)}{\sum_{v=v_{min}}^{N} P(v)}$	(5.8)
LAM	Laminarity: percentage of recurrence points that form vertical lines[a]	$\frac{\sum_{v=v_{min}}^{N} v \cdot P(v)}{\sum_{v=1}^{N} v \cdot P(v)}$	
L	Averaged diagonal line length[b]	$\frac{\sum_{l=l_{min}}^{N} l \cdot P(l)}{\sum_{l=l_{min}}^{N} P(l)}$	(5.9)
DET	Determinism: percentage of recurrence points that form diagonal lines[b]	$\frac{\sum_{l=l_{min}}^{N} l \cdot P(l)}{\sum_{i,j=1}^{N} R_{i,j}}$	

[a] $P(v)$ is the histogram of the lengths v of the black diagonal lines, and v_{min} is the minimal length of what should be considered to be a vertical line (typically, $v_{min} = 2$)
[b] $P(l)$ is the histogram of the lengths l of the black diagonal lines, and l_{min} is the minimal length of what should be considered to be a diagonal line (typically, $l_{min} = 2$)

1. Compute the distance vector ε that produces a RP with a percentage of black points between 10 and 80 % of the total number of points.
2. For each one of the elements in the vector ε, compute the RP (5.2) and obtain the selected RQA measure based on the null-hypothesis under study. We refer to the RQA measure as $DVRQA_{orig}(\varepsilon)$.
3. Repeat steps 1–2 for the N_s surrogates to establish the confidence intervals. The RQA measure for the ith surrogate is now referred to as $DVRQA_{surr,i}(\varepsilon)$.
4. Plot both variables on the same normalized graph. This will be referred as the DVRQA plot.

The computational complexity of the DVRQA analysis depends on the RQA measure selected. For simple RQA measures, the computational complexity can be similar to that of the DVV method. In more sophisticated measures, fast algorithms such as those described in [38] can be used.

5.3.3 Hypothesis Test

The modality tests done in this work compare different kinds of Recurrence Quantification measures that are computed for the original signal to those obtained for an ensemble of surrogates. So far in this work, only a visual comparison of the sta-

tistical measures studied for the original signal, $DVRQA_{orig}(\varepsilon)$, and the surrogates, $DVRQA_{surr,i}(\varepsilon)$ has been done. However, it is easy to implement a statistical test to quantify the level of similarity or difference that allow the hypothesis under study to be accepted or rejected.

The metric used in the analysis is the same as in the original DVV algorithm but substituting the variance estimator for each of the applied RQA measures, ξ, in each case:

$$t_{\xi} = \sqrt{\left\langle \left(\xi^2(\varepsilon) - \frac{\sum_{i=1}^{N_s} DVRQA_{surr,i}^2(\varepsilon)}{N_s} \right) \right\rangle_{\varepsilon}} \tag{5.10}$$

where $DVRQA_{surr,i}(\varepsilon)$ is the corresponding RQA metric at recurrence threshold ε for the ith surrogate, and the average $(\langle \cdot \rangle_{\varepsilon})$ is taken for each component of the threshold vector, ε. The variable ξ is replaced by $DVRQA_{orig}(\varepsilon)$ to compute the metric corresponding to the original signal, t_o, and replaced by $DVRQA_{surr,i}(\varepsilon)(i = 1, \ldots, N_s)$ to compute the metric corresponding to each surrogate, $t_{s,i}$.

Since the analytical form of the probability distribution of the applied metric is not known, a non-parametric, rank-based test is used [39]. In this work, for every original time series, we generate $N_s = 99$ surrogates. The metric for the original signal, t_o, and for the surrogates, $t_{s,i}(i = 1, \ldots, N_s)$, are computed, and the series $\{t_o, t_{s,i}\}$ is sorted in increasing order, after which the position index (rank) r of t_o is determined. A right-tailed test is rejected (and therefore the null hypothesis under study) if the rank r of the original time series exceeds 90. In this way, a single test statistic is obtained, and the above-mentioned right-tailed surrogate testing quantifies the level of rejection of the null hypothesis. In Sect. 5.4, the results of the corresponding rank tests for the analyzed examples are presented.

5.4 Applications

This section presents some results of signal modality characterization for the detection of non-linearities and complexity using both the surrogate data generation algorithms presented in Sect. 5.2 and the test statistics based on RQA measures presented in Sect. 5.3.

The linear and non-linear nature of time series is examined by performing the DVRQA analysis on both the original and 99 surrogates time series computed with the WiAAFT algorithm. The statistical measure applied was the *TT* measure. To verify the proposed technique, a number of time series with different linear and non-linear natures were generated. Each of the generated signals consisted of 1000 samples. The first two signals correspond to the examples in the text seen so far: the quasiperiodic signal (5.4) and the Rössler signal (5.5). Two random signals were also studied: Model 1 (5.11) and Model 2 (5.12).

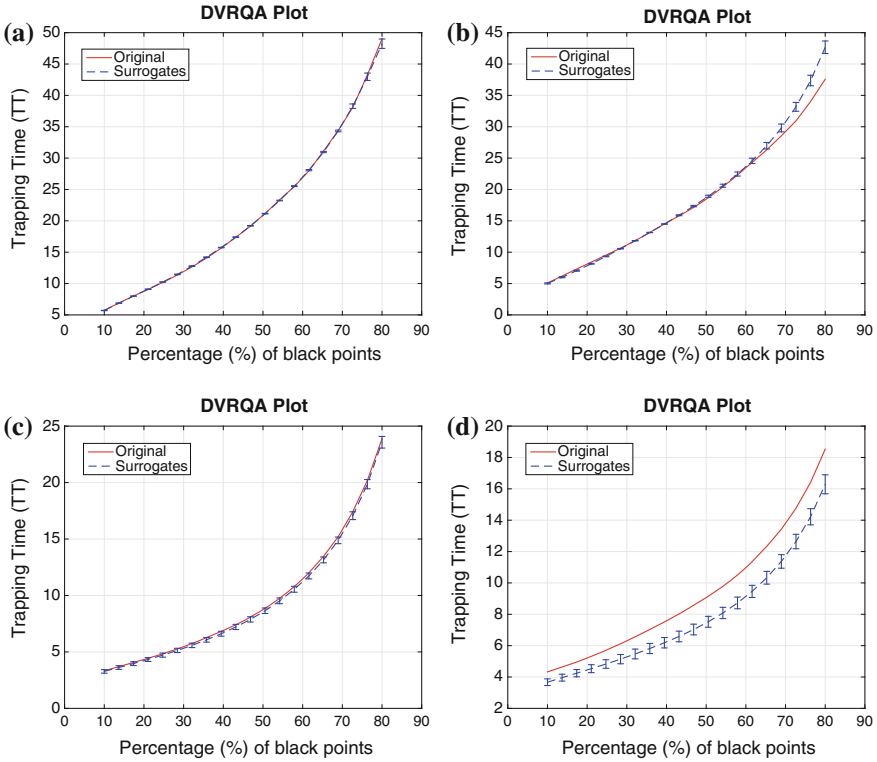

Fig. 5.5 DVRQA analysis with 99 WiAAFT surrogates performed on four simulated time series using the embedding parameters obtained via the mutual information algorithm (τ) and the Cao method (m): **a** Quasiperiodic signal; **b** Rossler signal; **c** Model 1; and **d** Model 2

$$y_k = 0.3 + 0.7 \cdot y_{k-1} + v_k + 0.4 \cdot v_{k-1} \tag{5.11}$$

where $y_0 = 0$ and v_k is a standard normal distribution, $N(0, 1)$.

$$y_k = v_k + 0.8 \cdot v_{k-1} \cdot v_{k-2} \tag{5.12}$$

where v_k is a standard normal distribution, $N(0, 1)$. For each signal, the optimal embedding parameters were determined using the mutual algorithm and the Cao method.

The results of the DVRQA analysis are illustrated in Fig. 5.5. The first signal analyzed is the quasiperiodic signal (Fig. 5.5a). The computed TT vector for the original signal coincides with the results obtained for the surrogates. The resulting rank-test is equal to 34, and therefore the linear null hypothesis can not be rejected. Figure 5.5b corresponds to the simulation of the Rössler signal, a well-known example of a chaotic (non-linear and deterministic) signal. In this case, the resulting TT values for the original signal significantly differ from the ones for the surrogates at

both ends of the graph. The resulting value of the rank-order test, r, is 100, and the signal can be identified as non-linear. For the stochastic Model 1, the values of TT fall into the confidence interval defined by the surrogates (the rank-test equals to 65). However, in the case of the non-linear Model 2 (Fig. 5.5d), the original and surrogate statistics are clearly different. Undoubtedly, the variable r is equal to 100 and the null-hypothesis related to linearity can be rejected.

Note that for the continuous time systems (the quasiperiodic and the Rössler signals), the maximum value of the TT measure is higher than the one computed for the stochastic signals, Model 1 and Model 2. As stated above, the trapping time is closely related to the sampling time and the recurrence radius. High values of TT appear for continuous time systems with a high time resolution and with a recurrence threshold ϵ that is not too small. However, the vertical structures are less prominent for maps and continuous systems that are reconstructed with small values of ϵ, unless the signals present any kind of singularities or intermittences [31, 34]. The DVRQA analysis based on the TT can also provide information about the laminarity of the phase space under study.

We are now going to illustrate how the proposed algorithms can be used in a real application for the characterization and detection of irregular animal vocalizations. We analyzed a recording containing dog barking sounds with different Harmonic to Noise Ratio (HNR). This parameter can be used in animal bioacoustics to quantify dysphonia. Normal sounding dogs occupy a middle HNR range, while dysphonic dogs exceed this range with higher as well as lower HNR values [40, 41]. To demonstrate this idea, we used a database from [42] which contains dysphonic sounds from animals that were recorded with a sample frequency of 22050 Hz.

Due to the oscillatory nature of the animal vocalizations and the fact that the appearance of this non-linear phenomenon results in the generation of an immeasurable number of new frequencies, it seems appropriate for the null hypothesis to be the measure of the complexity of each of the signals. The increase in complexity may indicate some kind of pathology [43]. Therefore, we propose the use of the PPTS algorithm and the average diagonal line length (L) to detect and characterize dysphonia phenomena.

Figure 5.6 (top) shows the time representation of a medium HNR dog bark together with the phase-space representation (only the first two components of the DVs). Figure 5.6 (bottom) shows the time representation of one of the computed PPTS together with its phase-space representation. It is important to note that the phase-space representation of both the original and one of the surrogates are quite similar.

A DVRQA test was done to analyze the dog barking sounds. We chose the average diagonal line length (L) as the discriminating measure. This measure (and any other RQA measures related to diagonal line length) captures determinism or predictability and is therefore a good candidate for the null test hypothesis of deviation of a periodic or quasi-periodic oscillation. We analyzed three different dog barks with three different HNR values: low HNR, medium HNR, and high HNR. The panels on the left in Fig. 5.7 show the temporal series together with their corresponding time frequency representation. Only fragments of 400 samples were analyzed. The start-

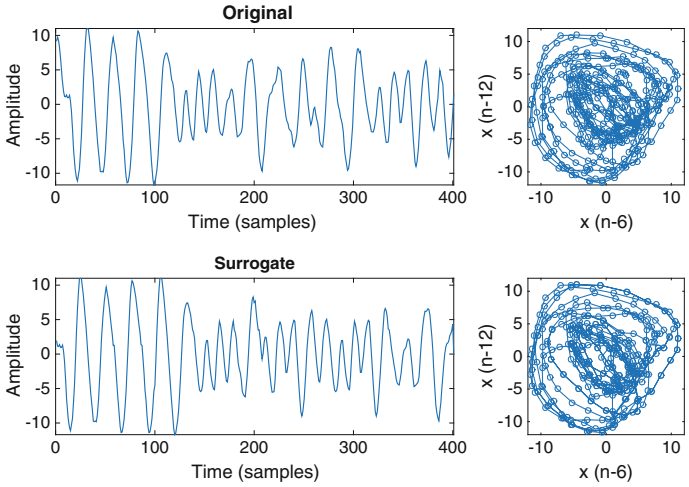

Fig. 5.6 *Top* Time representation of the medium HNR dog sound alongside with its phase-space representation. *Bottom* Time representation of one of the surrogates obtained with the PPTS together with its phase-space representation

ing point for the selection of the DVRQA analysis is not crucial (we obtained very similar results independently of the initial position of the 400 samples window). The red vertical line indicates the exact point where the DVRQA analysis was done. The panels on the right of the Fig. 5.7 show the DVRQA tests for each one of the dog barks. The confidence intervals were obtained computing 99 PPTS for each example of dog barks. The null hypothesis is that the signal can be obtained as the sum of sinusoids with incommensurate frequencies.

The low and medium HNR dog barks are very difficult to distinguish by looking at their corresponding time frequency representations. However, the DVRQA plot allows a clear differentiation between these two sounds. The null hypothesis can not be rejected for the low HNR bog bark since the rank-test results in r equals to 1, meaning that the $DVRQA_{orig}$ almost exactly match with the average of its surrogates, $DVRQA_{surr,i}$ (the red line in Fig. 5.7b falls in the middle of the confidence interval). On the other hand, the hypothesis test for the medium HNR dog bark can be rejected with a value of the rank-order test, r, equals to 90. This can be used to help distinguish between normal and dysphonic dogs.

In the time frequency representation of the high HNR dog barks, it is easy to see the presence of frequencies that lie between or below the main harmonic frequencies. This is a non-linear phenomenon known as subharmonic generation that also appears in pathological voices. For the high HNR dog bark, the DVRQA plot clearly shows a deviation from a periodic or a quasiperiodic oscillation, and the null test can be rejected with the variable r equals to 100 in the confidence interval. It must be highlighted that RQA measures combined with the surrogate approach can characterize phenomena that are usually analyzed by means of time-frequency techniques.

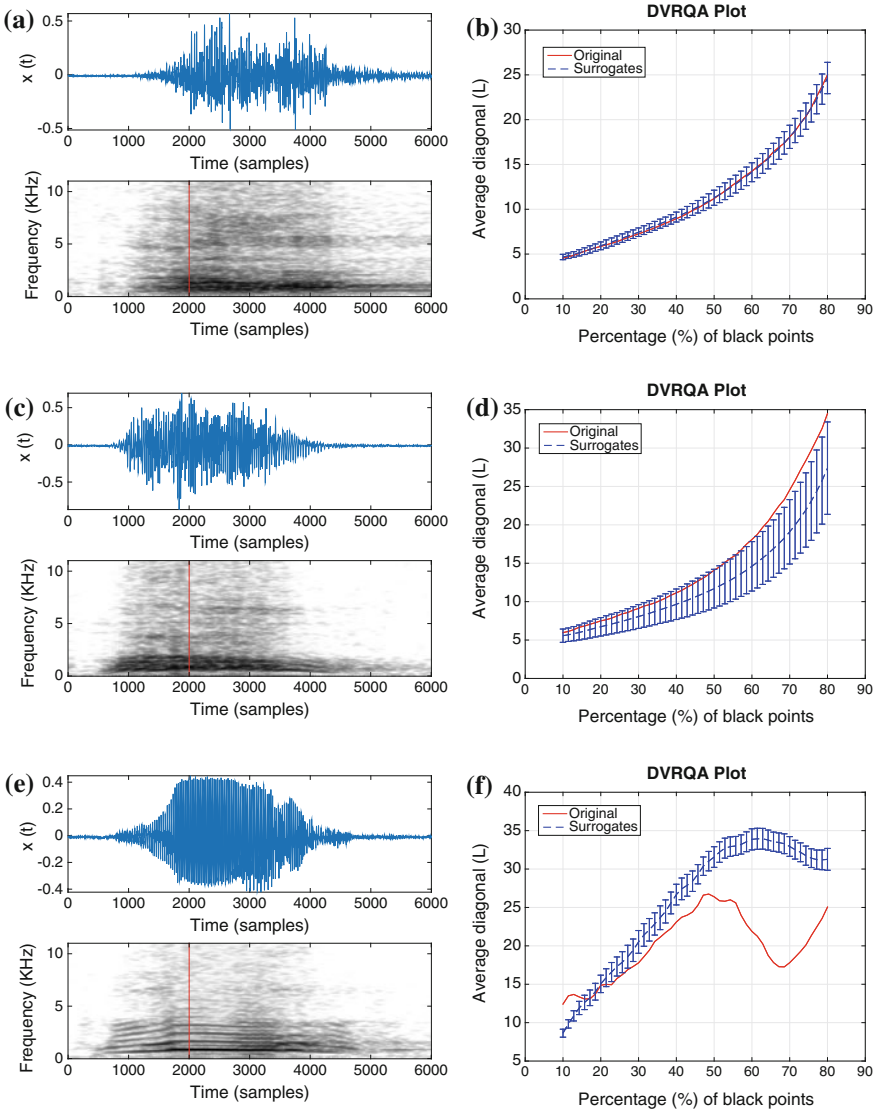

Fig. 5.7 **a** TFR of a low HNR dog bark, **b** DVRQA analysis of a low HNR dog bark, **c** TFR of a medium HNR dog bark, **d** DVRQA analysis of a medium HNR dog bark, **e** TFR of a high HNR dog bark, **f** DVRQA analysis of a high HNR dog bark. The *red vertical line* indicates where the DVRQA analysis was performed (99 PPTS were used for the analyses)

5.5 Conclusions

In this work, we have analyzed the problem of using hypothesis testing for the non-linear detection and characterization of complexity in temporal series from the RPs point of view. In order to do this, we have focused on both the generation of surrogates and the design of new statistical tests. To generate surrogates, we worked with the WiAAFT algorithm, which is a technique that maintains the temporal structure of the original signal and at the same time provides a sufficient degree of randomness to obtain valid surrogates for testing non-linearity (which are suitable for non-stationary signals). We have also created surrogates of a different kind, the PPTS, which are valid for detecting non-linear determinism and complexity in short oscillating signals. RPs play an important role both in the definition of twin points and the understanding of the PPTS algorithm. We have also demonstrated that RPs are not only a valid tool for the generation of surrogates but also for the design of the statistical tests. This has been illustrated by reformulating the DVV method using RPs.

This analysis has led to the creation of new discriminating tests that are based on RQA oriented to hypothesis testing (i.e., non-linear and complexity detection). The proposed test (DVRQA) has the advantage of analyzing the recurrent structures without having to choose the most appropriate recurrence threshold in an unknown situation. The selection of the RQA metrics must be carefully chosen based on the surrogates. The trapping time measure captures significant differences in the degree of locality and is therefore appropriate for testing non-linearity with the WiAAFT algorithm. The PPTS has proven to be useful for testing complexity in short oscillating signals using the averaged diagonal line length.

We have demonstrated through simulations that we can detect non-linearities in signals of different nature (stochastic and deterministic) using the proposed algorithms. Moreover, we have analyzed the problem of anomalies in the voice production of mammals and the use of algorithms for their detection and characterization. The analysis of a database containing real-world sounds from dysphonic dogs has shown that the L measure, together with the PPTS algorithm, allows the relationship among HNR, complexity and dysphonic dog barking sounds to be established. This technique has the potential to be applied in many other domains.

Acknowledgments This work has been supported by the Spanish Administration under grant TEC2011-23403.

References

1. K. Lehnertz, R.G. Andrzejak, J. Arnhold, T. Kreuz, F. Mormann, C. Rieke, A. Widman, C.E. Elger, Nonlinear EEG analysis in epilepsy: its possible use for interictal focus localization, seizure anticipation, and prevention. J. Clin. Neurophysiol.: Off. Publ. Am. Electroencephalogr. Soc. **18**(3), 209–222 (2001)

2. W.T. Fitch, J. Neubauer, H. Herzel, Calls out of chaos: the adaptive significance of nonlinear phenomena in mammalian vocal production. Animal Behav. **63**(3), 407–418 (2002)
3. D.G. McMillan, Non-linear predictability of UK stock market returns. Oxford Bul. Econ. Stat. **65**(5), 557–573 (2003)
4. F. Laio, A. Porporato, L. Ridolfi, S. Tamea, Detecting nonlinearity in time series driven by non-Gaussian noise: the case of river flows. Nonlinear Process. Geophys. **11**(4), 463–470 (2004)
5. P. Borgnat, P. Flandrin, P. Honeine, C. Richard, J. Xiao, Testing stationarity with surrogates: a time-frequency approach. IEEE Trans. Signal Process. **58**(7), 3459–3470 (2010)
6. M. Small, C.K. Tse, Detecting determinism in time series: the method of surrogate data. IEEE Tran. Circuits Syst. I: Fundam. Theory Appl. **50**(5), 663–672 (2003)
7. J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, J.D. Farmer, Testing for nonlinearity in time series: the method of surrogate data. Physi. D: Nonlinear Phenom. **58**(1), 77–94 (1992)
8. T. Schreiber, A. Schmitz, Improved surrogate data for nonlinearity tests. Phys. Rev. Lett. **77**(4), 635–638 (1996)
9. X. Luo, T. Nakamura, M. Small, Surrogate test to distinguish between chaotic and pseudoperiodic time series. Phys. Rev. E **71**(2) (2005)
10. T. Nakamura, M. Small, Small-shuffle surrogate data: testing for dynamics in fluctuating data with trends. Phys. Rev. E **72**(5), 056216 (2005)
11. M. Small, D. Yu, R. Harrison. Surrogate test for pseudoperiodic time series data. Phys. Rev. Lett. **87**(18) (2001)
12. M. Thiel, M.C. Romano, J. Kurths, M. Rolfs, R. Kliegl, Twin surrogates to test for complex synchronisation. Europhys. Lett. **75**(4), 535–541 (2006)
13. M.C. Romano, M. Thiel, J. Kurths, K. Mergenthaler, R. Engbert, Hypothesis test for synchronization: twin surrogates revisited. Chaos **19**(1), 015108 (2009)
14. T. Schreiber, A. Schmitz, Improved surrogate data for nonlinearity tests. Phys. Rev. Lett. **77**, 635–638 (1996)
15. D. Kaplan, Nonlinearity and Nonstationarity: The Use of Surrogate Data in Interpreting Fluctuations, in *Proceedings of the 3rd Annual Workshop on Computer Applications of Blood Pressure and Heart Rate Signals* (1997)
16. M.C. Casdagli, A.S. Weigend, Exploring the continuum between deterministic and stochastic modeling, in *Time Series Prediction: Forecasting the Future and Understanding the Past* (Addison-Wesley, Reading, MA, 1994)
17. T. Gautama, D.P. Mandic, M.M. Van Hulle, The delay vector variance method for detecting determinism and nonlinearity in time series. Phys. D: Nonlinear Phenom. **190**(3–4), 167–176 (2004)
18. N.H. Packard, J.P. Crutchfield, J.D. Farmer, R.S. Shaw, Geometry from a time series. Phys. Rev. Lett. **45**(9), 712 (1980)
19. D.P. Mandic, M. Chen, T. Gautama, M.M. Van Hulle, A. Constantinides, On the characterization of the deterministic/stochastic and linear/nonlinear nature of time series. Proc. Royal Soc. A: Math. Phys. Eng. Sci. **464**(2093), 1141–1160 (2008)
20. A. Fraser, H. Swinney, Independent coordinates for strange attractors from mutual information. Phys. Rev. A **33**, 1134–1140 (1986)
21. M.B. Kennel, R. Brown, H.D.I. Abarbanel, Determining embedding dimension for phase-space reconstruction using a geometrical construction. Phys. Rev. A **45**, 3403–3411 (1992)
22. L. Cao, Practical method for determining the minimum embedding dimension of a scalar time series. Phys. D: Nonlinear Phenom. **110**(12), 43–50 (1997)
23. J.P. Eckmann, S. Kamphorst, S. Oliffson, D. Ruelle, Recurrence plots of dynamical systems. Europhys. Lett. **4**(9), 973–977 (1987)
24. C.J. Keylock, Improved preservation of autocorrelative structure in surrogate data using an initial wavelet step. Nonlinear Proces. Geophys. **15**(3), 435–444 (2008)
25. D.B. Percival, A.T. Walden, *Wavelet Methods for Time Series Analysis (Cambridge Series in Statistical and Probabilistic Mathematics)* (Cambridge University Press, 2000)
26. C. Tao, J.J. Jiang, Chaotic component obscured by strong periodicity in voice production system. Phys. Rev. E **77**, 061922 (2008)

27. R. Miralles, L. Vergara, A. Salazar, J. Igual, Blind detection of nonlinearities in multiple-echo ultrasonic signals. IEEE Trans. Ultrason. Ferroelectr. Freq. Control **55**(3), 1–11 (2008)
28. R. Miralles, A. Carrion, D. Looney, G. Lara, D. Mandic, Characterization of the complexity in short oscillating time series: an application to seismic airgun detonations. J. Acoust. Soc. Am. **138**(3), 1595–1603 (2015)
29. M.T. Rosenstein, J.J. Collins, C.J. De Luca, A practical method for calculating largest Lyapunov exponents from small data sets. Phys. D: Nonlinear Phenom. **65**(1–2), 117–134 (1993)
30. P. Grassberger, I. Procaccia, Measuring the strangeness of strange attractors. Phys. D: Nonlinear Phenom. **9**(1–2), 189–208 (1983)
31. N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, J. Kurths, Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. Phys. Rev. E **66**(2), 026702 (2002)
32. C.L. Webber, J.P. Zbilut, Dynamical assessment of physiological systems and states using recurrence plot strategies. J. Appl. Physiol. **76**(2), 965–973 (1994)
33. J.B. Gao, Recurrence time statistics for chaotic systems and their applications. Phys. Rev. Lett. **83**, 3178–3181 (1999)
34. J. Gao, H. Cai, On the structures and quantification of recurrence plots. Phys. Lett. A **270**(12), 75–87 (2000)
35. N. Marwan, M.C. Romano, M. Thiel, J. Kurths, Recurrence plots for the analysis of complex systems. Phys. Rep. **438**(5), 237–329 (2007)
36. K. Klimaszewska, J.J. Żebrowski, Detection of the type of intermittency using characteristic patterns in recurrence plots. Phys. Rev. E **80**, 026214 (2009)
37. J.P. Zbilut, C.L. Webber Jr., Laminar recurrences, maxline, unstable singularities and biological dynamics. Eur. Phys. J. Spec. Topics **164**(1), 55–65 (2008)
38. T. Rawald, M. Sips, N. Marwan, D. Dransch, Fast Computation of Recurrences in Long Time Series. *Translational Recurrences, Springer Proceedings in Mathematics and Statistics*, vol. 103, pp. 17–29 (2014)
39. J. Theiler, D. Prichard, Constrained-realization Monte-Carlo method for hypothesis testing. Phys. D: Nonlinear Phenom. **94**(4), 221–235 (1996)
40. T. Riede, H. Herzel, K. Hammerschmidt, L. Brunnberg, G. Tembrock, The harmonic-to-noise ratio applied to dog barks. J. Acoust. Soc. Am. **110**(4) (2001)
41. T. Riede, B.R. Mitchell, I. Tokuda, M.J. Owren, Characterizing noise in nonhuman vocalizations: Acoustic analysis and human perception of barks by coyotes and dogs. J. Acoust. Soc. Am. **118**(1) (2005)
42. T. Riede, *Vocal changes in animals during disorders*. Ph.D. Report. Mathematisch-Naturwissenschaftlichen Fakultat I der Humboldt-Universitat zu Berlin (2000), http://edoc.hu-berlin.de/dissertationen/riede-tobias-2000-06-26/HTML/
43. I. Wilden, H. Herzel, G. Peters, G. Tembrock, Subharmonics, biphonation, and deterministic chaos in mammal vocalization. Bioacoustics **9**(3), 171–196 (1998)

Chapter 6
Approximate Recurrence Quantification Analysis (aRQA) in Code of Best Practice

Stephan Spiegel, David Schultz and Norbert Marwan

Abstract Recurrence quantification analysis (RQA) is a well-known tool for studying nonlinear behavior of dynamical systems, e.g. for finding transitions in climate data or classifying reading abilities. But the construction of a recurrence plot and the subsequent quantification of its small and large scale structures is computational demanding, especially for long time series or data streams with high sample rate. One way to reduce the time and space complexity of RQA are approximations, which are sufficient for many data analysis tasks, although they do not guarantee exact solutions. In earlier work, we proposed how to approximate diagonal line based RQA measures and showed how these approximations perform in finding transitions for difference equations. The present work aims at extending these approximations to vertical line based RQA measures and investigating the runtime/accuracy of our approximate RQA measures on real-life climate data. Our empirical evaluation shows that the proposed approximate RQA measures achieve tremendous speedups without losing much of the accuracy.

Keywords Recurrence plots · Recurrence quantification analysis · Time series · (Dis)similarity measures · Approximate solutions

S. Spiegel (✉) · D. Schultz
DAI-Lab, Technische Universität Berlin, Ernst-Reuter-Platz 7,
10587 Berlin, Germany
e-mail: spiegel@dai-lab.de

D. Schultz
e-mail: schultz@dai-lab.de

N. Marwan
Potsdam Institute for Climate Impact Research, Pappelallee 20,
14412 Potsdam, Germany
e-mail: marwan@pik-potsdam.de

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8_6

6.1 Introduction

In recent years, recurrence quantification analysis (RQA) has gained popularity in the time series community [1–3], where recurrence plot-based tools have been developed (i) to measure the pairwise (dis)similarity between temporal measurements based on co-occurring patterns [3, 4], (ii) for classification purposes in different scientific disciplines [5–9], (iii) to detect regime transitions [10–12], or even (iv) to study interrelationships and synchronization between different dynamical systems [13–15].

Since the quantification of recurring patterns is computational expensive, speedup techniques [16] and approximations [17] have been proposed. Speedup techniques commonly use distributed computing that ensures exact RQA results, e.g. by performing parallel processes on multiple Graphic Processing Units (GPUs), whereas approximation techniques estimate the RQA measures by means of less computational expensive algorithms. Given a time series with about one million data points, distributed computing with two GPUs has been shown to reduce the RQA calculation time by 1–2 orders of magnitude [16]. However, this work demonstrates that the proposed approximations [17] are able to reduce the RQA calculation time (for the same one million measurements) by 4 orders of magnitude. This tremendous speedup makes the approximation approach extremely valuable for many real-life data analysis tasks, although it does not yield exact results.

In this work we extend the approximation approach [17] to vertical line based measures, assess the runtime of our approximate RQA measures for relatively long time series (from climate impact research), and investigate the use our approximate RQA measures for transition detection.

6.2 Background and Notation

6.2.1 Recurrence Plots (RPs)

Recurrence plots (RPs) have been introduced to study the dynamics of complex systems that is represented in an m-dimensional phase space by its phase space trajectory $\mathbf{x}_i \in \mathbb{R}^m$ (assuming discrete sampling, $i = 1, \ldots, N$) [18]. A phase space trajectory can be reconstructed from a time series u_i ($t = i\Delta t$, where Δt is the sampling time) by different embedding schemes. The most frequently used scheme is the time delay embedding [19],

$$\mathbf{x}_i = \left(u_i, u_{i+1}, \ldots, u_{i+(m-1)\tau}\right), \tag{6.1}$$

with m the embedding dimension and τ the embedding delay. Both parameters can be estimated from the original data using false nearest neighbors and mutual information [20]. In the following we only consider the trajectory $\mathbf{x}$ and no longer the underlying time series u. That means the process of creating $\boldsymbol{x}$ from u by time delay

embedding is considered to be completed. Later on, we will apply time delay embedding to the trajectory $\boldsymbol{x}$ again. It is important to distinguish between both embedding procedures. The first is used for reconstruction purposes and the latter is used to express RQA-measures in a way that allows fast computation.

A RP is a 2-dimensional representation of those times when the phase space trajectory $\mathbf{x}_i$ recurs. As soon as a dynamical state at time j comes close to a previous (or future) state at time i, the recurrence matrix $\mathbf{R}$ at (i,j) has an entry one [20]:

$$R_{i,j} := \Theta(\varepsilon - \|\mathbf{x}_i - \mathbf{x}_j\|), \qquad i,j = 1,\ldots,N, \tag{6.2}$$

where $\|\cdot\|$ is a norm (representing the spatial distance between the states at times i and j), ε is a predefined recurrence threshold, and Θ is the Heaviside function (ensuring a binary $\mathbf{R}$). The RP has a square form and usually the identity $R_{i,i} \equiv 1$ is included in the graphical representation, although for calculations it might be useful to remove it [20]. The graphical representation of the RP allows to derive qualitative characterizations of the dynamical systems. For the quantitative description of the dynamics, the small-scale patterns in the RP can be used, such as diagonal and vertical lines. The histograms of the lengths of these lines are the base of the recurrence quantification analysis (RQA) developed by Webber and Zbilut and later by Marwan et al. [7, 21, 22].

6.2.2 Recurrence Rate (RR)

The simplest measure of RQA is the density of recurrence points in the RP, the recurrence rate:

$$RR := \frac{1}{N^2} \sum_{i,j=1}^{N} R_{i,j}, \tag{6.3}$$

that can be interpreted as the probability that any state of the system will recur.

6.2.3 Determinism (DET)

The fraction of recurrence points that form diagonal lines of minimal length μ is the determinism measure:

$$DET^{(\mu)} := \frac{\sum_{l=\mu}^{N} l \cdot D(l)}{\sum_{i,j=1}^{N} R_{i,j}} = \frac{\sum_{l=\mu}^{N} l \cdot D(l)}{\sum_{l=1}^{N} l \cdot D(l)} \tag{6.4}$$

where

$$D(l) := \sum_{i,j=1}^{N} \left\{ \left(1 - R_{i-1,j-1}\right) \cdot \left(1 - R_{i+l,j+l}\right) \cdot \prod_{k=0}^{l-1} R_{i+k,j+k} \right\}$$

is the histogram of the lengths of the diagonal lines. The understanding of 'determinism' in this sense is of heuristic nature.

6.2.4 Average Diagonal Line Length (L)

The average length of all diagonal lines (of at least length μ) in the RP is

$$L^{(\mu)} := \frac{\sum_{l=\mu}^{N} l \cdot D(l)}{\sum_{l=\mu}^{N} D(l)}, \tag{6.5}$$

and can be interpreted as the mean prediction time. As the diagonal lines in the RP are related to the divergence behavior of the phase space trajectory, its relationship with the Lyapunov exponents are obvious. Indeed, there is clear link between the distribution of the diagonal line lengths and the K_2 entropy of the system [23].

6.2.5 Laminarity (LAM)

Similar to the measure *DET*, the fraction of recurrence points that form vertical lines of a certain minimum length μ can be calculated. The corresponding measure is called laminarity:

$$LAM^{(\mu)} := \frac{\sum_{l=\mu}^{N} l \cdot V(l)}{\sum_{l=1}^{N} l \cdot V(l)}, \tag{6.6}$$

with

$$V(l) := \sum_{i,j=1}^{N} \left\{ \left(1 - R_{i,j-1}\right) \cdot \left(1 - R_{i,j+l}\right) \cdot \prod_{k=0}^{l-1} R_{i,j+k} \right\},$$

the histogram of the lengths of the vertical lines in the RP. Vertical (as well as horizontal) lines appear when states do not change or change only very slowly, as it is typical for intermittence and laminar regimes [7].

Further measures have been introduced that incorporate such line length distributions and also network properties [20, 24]. All these measures can be used to classify different dynamical regimes and to detect their transitions.

6.3 Approximate Recurrence Quantification Analysis

In this section we propose an alternative way of computing the RQA measures introduced previously. Our idea can be expressed as follows: (1) We propose two novel quantification techniques, namely pairwise proximities *PP* and stationary states *SS*, which account for diagonal and vertical lines by means of an embedded trajectory. (2) Based on *PP* and *SS* we introduce alternative formulations of the traditional RQA measures that are equivalent to the original formulations, provided that the phase space norm that measures the spatial distances (e.g. in (6.2)) is the maximum-norm, defined by $\|\mathbf{y}\|_\infty = \max_i |\mathbf{y}_i|$. (3) Using these new formulations, the RQA measures can be computed quickly if the similarity threshold is zero ($\varepsilon = 0$). (4) If the similarity threshold is greater than zero, we first discretize the data and then set the threshold to zero in order to make use of fast algorithms that are facilitated by our alternative formulations. In this case—due to discretization—we only get an approximation of the exact RQA measures.

In earlier work [17] we have proven the equivalence between our alternative and the original formulation of the RQA measures and, furthermore, analyzed the approximation error theoretically. Moreover, we have provided detailed information on the discretization and employed algorithms [17], which have complexity of $(N \log(N))$. Our implementation of the discretization and employed algorithms can be found in Sect. 6.4.

Important Note. In this section we assume that the similarity threshold is zero. That means the recurrent states we aim at quantifying are only states that are equal. This case is relevant if the trajectory $\mathbf{x}$ is discrete-valued or has been discretized beforehand in order to compute the approximate RQA-measures. To be more clear on the role of the threshold, we define *PP* and *SS* for general $\varepsilon \geq 0$, but the reader may imagine that in application of the fast (approximate) RQA algorithms we have $\varepsilon = 0$.

Given a phase space trajectory $\mathbf{x}$, the number of pairwise proximities *PP* can be defined as follows:

$$PP^{(\nu)} := \sum_{i,j=1}^{N-\nu+1} \Theta(\varepsilon - \|\mathbf{x}_i^{(\nu)} - \mathbf{x}_j^{(\nu)}\|), \tag{6.7}$$

where $\mathbf{x}^{(\nu)}$ is a time-delay embedded version of the trajectory $\mathbf{x}$ with embedding dimension $\nu \in \mathbb{N}$ and time-delay 1, i.e.,

$$\mathbf{x}_i^{(\nu)} = (\mathbf{x}_i, \ldots, \mathbf{x}_{i+\nu-1}), \qquad i = 1, \ldots, N - \nu + 1. \tag{6.8}$$

We want to emphasize that the key idea of our quantification techniques (*PP* and *SS*) is to embed the trajectory, since recurrent states of embedded trajectories indicate recurrent sequences in the original trajectory $\mathbf{x}$ if the phase space norm is $\|\cdot\|_\infty$. To see this, assume that, for instance, the recurrence plot of the embedded trajectory $\mathbf{x}^{(2)}$ indicates a recurrence point at position (i, j), that means

$$\|\mathbf{x}_i^{(2)} - \mathbf{x}_j^{(2)}\|_\infty \leq \varepsilon.$$

By definition of $\|\cdot\|_\infty$ and the trajectory embedding, this is equivalent to

$$\|\mathbf{x}_i - \mathbf{x}_j\|_\infty \leq \varepsilon \quad \text{and} \quad \|\mathbf{x}_{i+1} - \mathbf{x}_{j+1}\|_\infty \leq \varepsilon,$$

which exactly means that the recurrence plot of the original trajectory $\mathbf{x}$ contains a diagonal line of length 2 starting at position (i,j). Note that this equivalence is not true for arbitrary norms.

Our implementation of the general time delay embedding, (6.1), can be found in Sect. 6.4.1.

As shown in [17], if $\varepsilon = 0$, the measure of pairwise proximities $PP^{(\nu)}$ can also be interpreted as the sum over the squared frequencies of recurring states, which can be determined using the histogram $h(\mathbf{x}^{(\nu)})$ of the embedded trajectory:

$$PP^{(\nu)} = h(\mathbf{x}^{(\nu)}) \cdot h(\mathbf{x}^{(\nu)}). \tag{6.9}$$

In (6.9) the histograms are represented as vectors containing the frequencies of the elements in $\mathbf{x}^{(\nu)}$ and the dot denotes the inner product, defined by $u \cdot w = \sum_i u_i w_i$. This relation is the key for the fast computation of the RQA-measures since the histograms can be obtained in $\mathcal{O}(N \log(N))$, where N is the length of the trajectory $\mathbf{x}$. It is important to note that (6.9) does only hold for $\varepsilon = 0$. This is the reason why the data has to be discretized if $\varepsilon > 0$ is required.

Based on our definition of pairwise proximities PP we can introduce alternative formulations for the original diagonal line based RQA measures introduced in Sect. 6.2. In the following we discuss an alternative formulation for recurrence rate RR, determinism DET, average diagonal line length L, and laminarity LAM.

Before moving on to more advanced recurrence quantification measures, we want to provide an image representation of pairwise proximities PP. Figure 6.1 shows the recurrence plot $\mathbf{R}$ of a discrete-valued sample trajectory $\mathbf{x}$ and its corresponding histogram $h(\mathbf{x}^{(\nu)})$ for trajectory embedding dimension $\nu = 1$ (note that $\mathbf{x}^{(1)} = \mathbf{x}$). The pairwise proximities $PP^{(1)}$ are equal to the total number of recurrence points in $\mathbf{R}$ and are given by the sum over the squared frequencies:

$$PP^{(1)} = h(\mathbf{x}^{(1)}) \cdot h(\mathbf{x}^{(1)}) = 3^2 + 5^2 + 2^2 = 38$$

6.3.1 Reformulation of Recurrence Rate (RR)

The pairwise proximities $PP^{(\nu)}$ for trajectory embedding dimension $\nu = 1$ can be interpreted as the number of recurrence points, which are traditionally expressed by the sum over all recurrence plot entries $\sum_{i,j=1}^{N} R_{i,j}$ (see 6.3). To compute the

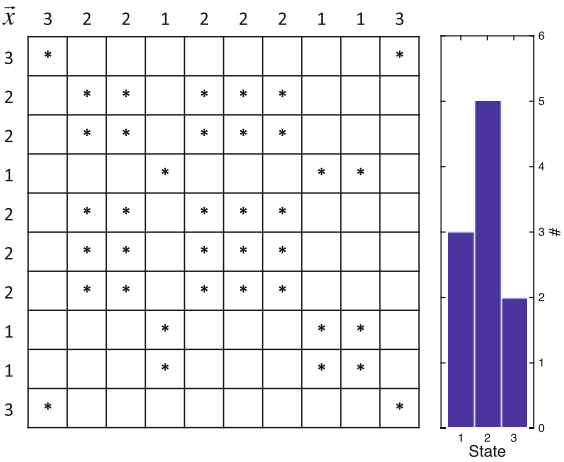

Fig. 6.1 Recurrence plot **R** of trajectory $\mathbf{x} = (3, 2, 2, 1, 2, 2, 2, 1, 1, 3)$ with similarity threshold $\varepsilon = 0$ (*left*) and its histogram $h(\mathbf{x}^v)$ for embedding dimension $v = 1$ (*right*), showing the frequencies of recurring states. The pairwise proximities $PP^{(1)}$ equal the total number of recurrence points in **R** and are given by the sum over the squared frequencies (see 6.9)

recurrence rate, the number of recurrence points is divided by the size of the recurrence plot, which is the squared length N^2 of the time series under study. Hence, the alternative way of computing the recurrence rate *RR* can be formalized as followed:

$$RR = PP^{(1)}/N^2. \tag{6.10}$$

For our sample trajectory **x** (shown in Fig. 6.1) with pairwise proximities $PP^{(1)} = 38$ and length $N = 10$ the recurrence rate is:

$$RR = 38/10^2 = 0.38$$

6.3.2 Reformulation of Determinism (DET)

The determinism *DET* can also be expressed in terms of pairwise proximities. Traditionally the determinism *DET* is described as the percentage of recurrence points which form diagonal lines (refer to 6.4). In the previous Sect. 6.3.1 we have already explained that the total number of recurrence points is equivalent to the pairwise proximities $PP^{(1)}$. Hence, the denominator of *DET* is known and it remains the question of how to compute the number of recurrence points that contribute to diagonal lines of minimum length μ. Our idea is to quantify the recurrence plot $\mathbf{R}^{(\mu)}$ of the embedded trajectory $\mathbf{x}^{(\mu)}$ in relation to the recurrence plot **R** of the original trajectory **x**. First note that each point in $\mathbf{R}^{(\mu)}$ indicates that there is a diagonal line of length $\geq \mu$ in **R**. Consequently, only lines we are interested in remain in $\mathbf{R}^{(\mu)}$. However, each diagonal line in $\mathbf{R}^{(\mu)}$ is $\mu - 1$ shorter than the corresponding line in **R**. Thus, we need to add the missing points. Evidently, the number of missing points is exactly given by "the number of diagonal lines of length $\geq \mu$" times $\mu - 1$, where "the

number of diagonal lines of length $\geq \mu$" in $\mathbf{R}$ is given by $(PP^{(\mu)} - PP^{(\mu+1)})$. To sum up, we have argued that

$$\sum_{l=\mu}^{N} l \cdot D(l) = PP^{(\mu)} + (PP^{(\mu)} - PP^{(\mu+1)}) \cdot (\mu - 1). \tag{6.11}$$

By simplifying (6.11) we achieve our alternative formulation for the determinism (proved in [17]), which is true for arbitrary similarity threshold $\varepsilon \geq 0$ and arbitrary minimum diagonal line length μ, provided that the phase space norm is the maximum-norm $\|\cdot\|_\infty$:

$$DET^{(\mu)} = \frac{\mu \cdot PP^{(\mu)} - (\mu - 1) \cdot PP^{(\mu+1)} \ldots - N}{PP^{(1)}}. \tag{6.12}$$

Depending on whether or not we want to include the recurrence points of the main diagonal in our calculation, we need to subtract N in the numerator (6.12).

Figure 6.2 illustrates how to employ the concept of pairwise proximities in order to compute the determinism for our sample time series $\mathbf{x}$ introduced in Fig. 6.1.

For example in Fig. 6.2, the determinism $DET^{(2)}$ for minimum diagonal line length $\mu = 2$ can be described as the number of recurrence points that rest on highlighted lines divided by the total number of recurrence points (i.e. 14/38). The total number of recurrence points is given by $PP^{(1)}$ (see Sect. 6.3.1) and the number of recurrence points that form diagonals of minimum length $\mu = 2$ can be expressed in terms of $PP^{(2)}$ and $PP^{(3)}$ (see 6.12). In Fig. 6.2, single circles illustrate the recurrences that are given by our formulation of pairwise proximities $PP^{(2)}$ for embedding dimension $\nu = 2$. By multiplying the length $\mu = 2$ and number $PP^{(2)}$ of the identified structures we quantify all recurrence points that rest on diagonal lines, including those of overlapping structures. To subtract recurrence points of overlapping struc-

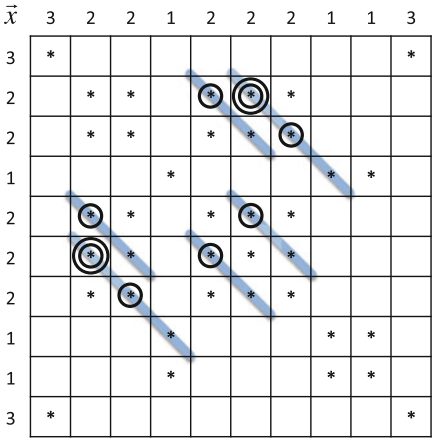

Fig. 6.2 Recurrence plot R of time series $\mathbf{x}$, where *highlighted lines* indicate diagonals that contribute the determinism DET, *single circles* illustrate recurrences that are given by our formulation of pairwise proximities $PP^{(2)}$ for embedding dimension $\nu = 2$, and *double circles* show recurrences that are quantified by $PP^{(3)}$ for $\nu = 3$ respectively

tures we compute the pairwise proximities $PP^{(3)}$ of higher embedding dimension $\nu = 2 + 1$, which are illustrated by double circles. Consequently, the determinism $DET^{(2)}$ for our sample time series $\mathbf{x}$ is computed in the following way:

$$DET^{(2)} = \frac{2 \cdot PP^{(2)} - 1 \cdot PP^{(3)}}{PP^{(1)}}$$
$$\left(= \frac{2 \cdot Single\ Circles - 1 \cdot Double\ Circles}{Total\ Number\ of\ Recurrences} \right)$$
$$= \frac{2 \cdot 8 - 1 \cdot 2}{38} = \frac{14}{38} \approx 0.37$$

6.3.3 Reformulation of Average Diagonal Line Length (L)

Given our new formulation for the determinism (see 6.12), the formalization of the average diagonal line length L in terms of pairwise proximities PP is straightforward. Informally speaking, L is defined as the number of recurrence points that form diagonals of minimum length μ divided by the number of diagonals of minimum length μ (see 6.5). We have already shown how to compute the first term or numerator in the previous Sect. 6.3.2. The second term or denominator can be computed by $PP^{(\mu)} - PP^{(\mu+1)}$, which is the number of diagonals with minimum length μ. Since we know that $PP^{(\mu)}$ accounts for all the diagonal line structures with minimum length μ including overlapping ones, we need to subtract the number of overlapping structures which are quantified by the term $PP^{(\mu+1)}$. Ultimately, under the same assumptions as for DET, our alternative formulation of L can be formalized as followed:

$$L^{(\mu)} = \frac{\mu \cdot PP^{(\mu)} - (\mu - 1) \cdot PP^{(\mu+1)} \cdots - N}{PP^{(\mu)} - PP^{(\mu+1)}}. \tag{6.13}$$

Same as for the determinism, we might not want to consider the main diagonal for our calculation and, thus, need to subtract N (the time series length) from the numerator.

For our sample time series $\mathbf{x}$ (shown in Fig. 6.2) we can compute the average diagonal line length for $\mu = 2$ as followed:

$$L^{(2)} = \frac{2 \cdot 8 - 1 \cdot 2}{8 - 2} = \frac{14}{6} \approx 2.33$$

6.3.4 Reformulation of Laminarity (LAM)

The laminarity is the percentage of recurrence points which form vertical lines (see 6.6) and **cannot** be computed by means of the PP measure, since it quantifies diagonal line structures. Therefore, we need to introduce a novel measure for stationary

states *SS*, which accounts for time intervals where the corresponding trajectory stays in the same phase space. Stationary states *SS* which stay stable for ν time points can be quantified in the following way.

$$SS^{(\nu)} := \sum_{i=1}^{N-\nu+1} \sum_{j=1}^{N} \Theta(\varepsilon - \|\mathbf{x}_i^{(\nu)} - \mathbf{1}^{(\nu)}\mathbf{x}_j\|), \tag{6.14}$$

where $\mathbf{1}^{(\nu)}\mathbf{x}_j = (\mathbf{x}_j, \ldots, \mathbf{x}_j)$ is the concatenation of ν copies of $\mathbf{x}_j$. Hence *SS* accounts for states where all elements in $\mathbf{x}_i^{(\nu)} = (\mathbf{x}_i, \ldots, \mathbf{x}_{i+\nu-1})$ are in a ε-neighborhood of $\mathbf{x}_j$, indicating that state $\mathbf{x}_i$ stays stationary for ν time points.

Analogously to (6.9) we can compute the stationary states efficiently in $\mathcal{O}(N \log(N))$ using histograms if $\varepsilon = 0$:

$$SS^{(\nu)} = \hbar(\mathbf{x}^{(\nu)}) \cdot \hbar(\mathbf{x}^{(1)}), \tag{6.15}$$

where $\hbar(\mathbf{x}^{(\nu)})$ is the *stationary state histogram* of the embedded trajectory, which—that is important—only accounts for stationary states of exact length ν (including overlapping structures in $\mathbf{x}$); and $\hbar(\mathbf{x}^{(1)})$ is the histogram of the original trajectory. Attention should be paid to the calculation of the inner product between stationary state histograms, since only frequencies of corresponding states are multiplied. For example the frequency of the stationary state $(1, 1)$ in $\mathbf{x}^{(2)}$ is multiplied with the frequency of state (1) in $\mathbf{x}$. Furthermore, it is important to mention that although **non**-stationary states may occur in an embedded trajectory (e.g. $(2, 1)$, in $\mathbf{x}^{(2)}$), their frequency in the corresponding stationary state histogram is always zero.

Given our new definition of stationary states *SS* and assuming that the phase space norm is the maximum-norm $\|\cdot\|_\infty$, we can compute the laminarity *LAM* for a given minimum vertical line length μ and an arbitrary threshold $\varepsilon \geq 0$ as follows:

$$LAM^{(\mu)} = \frac{\mu \cdot SS^{(\mu)} - (\mu - 1) \cdot SS^{(\mu+1)} \cdots - N}{SS^{(1)}}, \tag{6.16}$$

where the denominator $SS^{(1)}$ denotes the total number of recurrence points and the numerator $\mu \cdot SS^{(\mu)} - (\mu - 1) \cdot SS^{(\mu+1)}$ denotes the number of recurrence points that form vertical lines of minimum length μ. The thoughtful reader might have noticed that our new formulation of *LAM* and *DET* (6.16) and (6.12) resemble each other. The difference is that *DET* accounts for diagonal lines using *PP* and *LAM* quantifies vertical lines using *SS*. Therefore the proof of the *LAM* formula (6.16) is very similar to the proof of the *DET* formula presented in earlier work [17].

Figure 6.3 illustrates how to compute the laminarity $LAM^{(2)}$ for our sample trajectory $\mathbf{x}$. As defined in (6.16), the laminarity $LAM^{(2)}$ for minimum vertical (or rather horizontal) line length $\mu = 2$ can be computed in terms of $SS^{(1)}$, $SS^{(2)}$, and $SS^{(3)}$. For example in Fig. 6.3, the total number of recurrences points is described by the term $SS^{(1)}$, which can also be interpreted as the sum over the square frequencies of stationary states, given by the histogram $\hbar(\mathbf{x}^{(1)})$ (see Table 6.1).

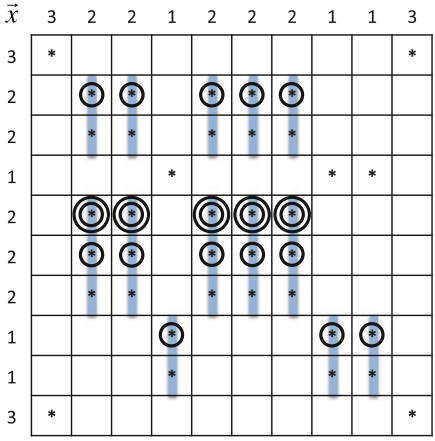

Fig. 6.3 Recurrence plot of trajectory $\mathbf{x}$ with similarity threshold $\varepsilon = 0$ and *highlighted vertical lines* of minimum length $\mu = 2$. The laminarity $LAM^{(2)} = 31/38$ is the percentage of recurrence points that form vertical lines of minimum length μ

Table 6.1 Stationary state histograms for our sample trajectory $\mathbf{x}$ in Fig. 6.3, showing the frequency of states that are stationary over μ time points

	State	#
$\hbar(\mathbf{x}^{(1)})$	[1]	3
	[2]	5
	[3]	2
$\hbar(\mathbf{x}^{(2)})$	[1,1]	1
	[2,2]	3
	[3,3]	0
$\hbar(\mathbf{x}^{(3)})$	[1,1,1]	0
	[2,2,2]	1
	[3,3,3]	0

Furthermore, the highlighted lines in Fig. 6.3 indicate all recurrence points that form vertical structures of minimum length $\mu = 2$, which can be quantified in terms of $SS^{(2)}$ and $SS^{(3)}$. For our sample trajectory $\mathbf{x}$ in Fig. 6.3, all stationary states $SS^{(2)}$ that are stable for 2 time points are illustrated by single circles. Moreover, double circles indicate stationary states $SS^{(3)}$ that are stable for 3 successive observations. The terms $SS^{(2)}$ and $SS^{(3)}$ can be computed by means of (6.14) and expressed in terms of the respective stationary state histograms $\hbar(\mathbf{x}^{(1)})$, $\hbar(\mathbf{x}^{(2)})$ and $\hbar(\mathbf{x}^{(3)})$:

$$
\begin{aligned}
SS^{(1)} &= \hbar(\mathbf{x}^{(1)}) \cdot \hbar(\mathbf{x}^{(1)}) = 38 \\
SS^{(2)} &= \hbar(\mathbf{x}^{(2)}) \cdot \hbar(\mathbf{x}^{(1)}) = 18 \\
SS^{(3)} &= \hbar(\mathbf{x}^{(3)}) \cdot \hbar(\mathbf{x}^{(1)}) = 5.
\end{aligned}
$$

The corresponding stationary state histograms for our plot in Fig. 6.3 are shown in Table 6.1. The multiplication of stationary state histograms is performed element-wise using the inner product, i.e. multiplying the frequencies of unique states with the corresponding stationary state counterparts and summing the results (see

Algorithm 4). For our sample trajectory $\mathbf{x}$, the inner product is $\hbar(\mathbf{x}^{(1)}) \cdot \hbar(\mathbf{x}^{(2)}) = 18$, since $3 \cdot 1 + 5 \cdot 3 + 2 \cdot 0 = 18$.

By multiplying the length $\mu = 2$ and the number $SS^{(2)}$ of identified stationary states we quantify all recurrence points that rest on highlighted lines, including those of overlapping vertical structures. To subtract the recurrence points of overlapping structures we compute the number of stationary states $SS^{(3)}$ with higher length $\mu = 2 + 1$, leading to the following formalization: $2 \cdot SS^{(2)} - 1 \cdot SS^{(3)}$. Eventually, we can compute the laminarity $LAM^{(2)}$ for the plot in Fig. 6.3 by dividing the amount of recurrence point that form vertical lines by the total number of recurrences $SS^{(1)}$:

$$\begin{aligned} LAM^2 &= \frac{2 \cdot SS^{(2)} - 1 \cdot SS^{(3)}}{SS^{(1)}} \\ &\left(= \frac{2 \cdot \textit{Single Circles} - 1 \cdot \textit{Double Circles}}{\textit{Total Number of Recurrences}} \right) \\ &= \frac{2 \cdot 18 - 1 \cdot 5}{38} = \frac{31}{38} \approx 0.82. \end{aligned}$$

Given our proposed definition of stationary states SS (6.14), we can also restate other vertical line based measures, such as trapping time TT or longest vertical line length V_{max}. However, this goes beyond the scope of this study.

6.4 Approximate Recurrence Quantification Analysis with MATLAB

Before we present our empirical results we want to discuss the implementation of our proposed approximate recurrence quantification analysis. The provided source code will help other researchers to reproduce our results and to continue with further ideas right where we left off. We decided to provide MATLAB code since it is very compact and often used in academia. However, our code snippets can also be executed in Octave, which is an open source alternative to MATLAB. Please note that our code is protected by copyright laws and is not provided for commercial use. If you plan to use our implementation for academic purpose (e.g. for reproduction of experimental results or further enhancements of the introduced concepts) we kindly remind you to cite this chapter.

In the following subsections we explain our implementation of: (i) time delay embedding (according to Takens' theorem), (ii) discretization (e.g. of a multivariate time series or an embedded phase space trajectory), (iii) pairwise proximities (for diagonal line based measures), (iv) stationary states (for vertical line based measures), and (v) our experimental protocol for the approximate recurrence quantification analysis of the Potsdam time series of hourly air temperature [16].

6.4.1 Time Delay Embedding

Our implementation of the time delay embedding is used for the following purposes: (i) given a sequence of temporal observations we aim at embedding the recorded time series into the phase space using predetermined parameters for the embedding dimension and time delay, (ii) given a (reconstructed) phase space trajectory we aim at embedding it (once again) in order to quantify recurring segments $PP^{(v)}$ and stationary intervals $SS^{(v)}$ with a certain number of time points v.

In general our time delay embedding function accepts an input time series of size $n \times d$, where n denotes the number of time points and d represents the dimensionality of the data. The function call furthermore requires us to specify the embedding dimension m and time delay τ. It is important to note that, in contrast to the time series embedding (i), the trajectory embedding (ii) always assumes unit time delay ($\tau = 1$) [17]. The output of our time delay embedding function is a time series of size $[n - (m - 1) * t] \times [d * m]$. What makes our *fTDE* implementation time efficient is the fact that the for loop in *Line* 11 does not run over the length n of the time series, but iterates over the embedding dimension m which is usually much smaller.

```
1  function X = fTDE(x,m,t)
2  %FTDE time delay embedding (C) Spiegel. et al.
3  %   x   ..  time series [n times d]
4  %   m   ..  embedding
5  %   t   ..  dalay
6  %   X   ..  time series [n-(m-1)*t times d*m]
7
8      [n,d] = size(x);
9
10     X = zeros(n-(m-1)*t,d*m);
11     for i = 1:m
12         a = i+(t-1)*(i-1);
13         b = a+n-1-(m-1)*t;
14         X(:,d*(i-1)+1:d*i) = x(a:b,:);
15     end
16 end
```

6.4.2 Discretization

In most real-life time series applications we aim at analyzing temporal data with continuous values. Since our concept of pairwise proximities *PP* and stationary states *SS* mainly relies on histograms, we need to apply some kind of binning to the continuous values beforehand. In our approach we first discretize the original time series and then create a histogram for the previously discretized data.

Our implemented discretization function requires an input time series of size $n \times d$ and the specification of a similarity threshold ε, which defines the size of the bins. In earlier work [17] we proposed to perform the discretization according to the size of the ε-neighborhood in the following manner: $\tilde{x} = \lfloor x/2\varepsilon \rfloor$ (see *Line* 23). The discretization is done element-wise and effects the approximation error of the subsequent recurrence quantification analysis [17].

```
17 function x = fDiscrete(x,eps)
18 %FDISCRETE Discretize Time Series (C) Spiegel et al.
19 %   x ..    time series [n-times-d]
20 %   eps ..  similarity thresholds [1-times-d]
21
22     if eps>0
23         x = floor(x*diag(1./(2*eps)));
24     end
25 end
```

6.4.3 Pairwise Proximity

Given a (reconstructed and subsequently embedded) phase space trajectory, we can use the concept of pairwise proximities $PP^{(\mu)}$ to quantify recurring segments of certain length μ that correspond to diagonal line structures in a recurrence plot. Having quantified the number of length of recurring segments, we can compute all diagonal line based RQA measures in a straightforward manner. For example in Sect. 6.3 we have explained how to use pairwise proximities *PP* to calculate the determinism *DET* and average diagonal line length *L*.

Our implementation of the pairwise proximities function takes an input time series of size $n \times d$ and returns the number recurring d-dimensional states, which (in our case) are the result of phase space reconstruction and subsequent trajectory embedding. The pairwise proximities function requires a time series with discrete values, since we aim at finding unique states (*Line* 31). Having identified unique states, we create a histogram that captures the frequency of the unique state in the next step (*Line* 32). Finally, we calculate the sum over the squared frequencies, which is equivalent to the inner product (dot-product) of the histogram with itself (*Line* 33).

```
26 function pp = fPProximities(x)
27 %FPPROXIMITIES pairwise proximities (C) Spiegel et al.
28 %   x ..    time series [nx-times-d]
29 %   pp ..   pairwise proximities
30
31     [~,~,ix] = unique(x,'rows');
32     hx = hist(ix,min(ix)-1:max(ix)+1);
33     pp = dot(hx,hx);
34 end
```

6.4.4 Stationary States

The idea of stationary states *SS* is an extension of earlier work [17] on approximate recurrence quantification analysis, but the concept is novel in that it enables us to quantify vertical line structures in an efficient way (without creating the recurrence plot). Although our definition of stationary states *SS* (6.14) resembles our definition of pairwise proximities *PP* (6.7), there is an important difference between the two concepts. In contrast to pairwise proximities *PP*, the computation of stationary states $SS^{(\nu)}$ with length ν is performed by comparing an embedded version of the reconstructed phase space trajectory $\mathbf{x}^{(\nu)}$ with the original trajectory $\mathbf{x}^{(1)}$ (see 6.14). This is due to the fact that we aim at identifying states that are stationary over a segment of ν time points (as explained at full length by our running example in Sect. 6.3.4).

Our implementation of the stationary state function requires as input the reconstructed phase space trajectory and its embedded version, regardless in which order. In *Line* 43–44 we identify stationary states by extracting those rows from the corresponding time series matrix, where all entries are the same. For this purpose we calculate the root mean squared value for each row vector, that is `sqrt(sum(x.^2,2)/d)`, and check for which row values the remainder after division by 1 equals 0, using the modulo operator ($\texttt{mod(value,1)}^{=} = 0$). In case that the remainder equals 0 we know that the corresponding row solely contains one and the same discrete entries.

Having identified the rows that contain only same entries, we create a stationary state histogram for the reconstructed phase trajectory as well as for its embedded version (see *Line* 46–47). By multiplying both histograms using the inner product (*Line* 48) we eventually get the number of stationary states that are steady for a certain time interval, whose length is given by the embedding dimension.

```
35 function ss = fSStates(x,y)
36 %FSSTATES stationary states (C) Spiegel et al.
37 %   x ..    time series [nx-times-d1]
38 %   y ..    time series [ny-times-d2]
39 %   ss ..   stationary states
40
41     [~,d1] = size(x);
42     [~,d2] = size(y);
43     x = x(mod(sqrt(sum(x.^2,2)/d1),1)==0,1);
44     y = y(mod(sqrt(sum(y.^2,2)/d2),1)==0,1);
45
46     hx = hist(x,min([x;y])-1:max([x;y])+1);
47     hy = hist(y,min([x;y])-1:max([x;y])+1);
48     ss = dot(hx,hy);
49 end
```

6.4.5 Approximate RQA

Having explained the implementation of time delay embedding, discretization, pairwise proximities, and stationary states, we are now in the position to introduce our experimental setup. First of all we load the Potsdam time series (*Line* 54), set the predefined parameters [16] (*Line* 57), and apply the time delay embedding (*Line* 59) to reconstruct the phase space trajectory. Afterwards we discretize the reconstructed trajectory (*Line* 62), which is a prerequisite for computing the approximate RQA measures. In the next step we embed the reconstructed and discretized trajectory (*Line* 63–64) in order to quantify the pairwise proximities (*Line* 66–68) and stationary states (*Line* 70–72). Given the number of pairwise proximities and stationary states for different embedding dimensions we are able to approximate the discussed diagonal and vertical line based RQA measures (*Line* 75–78). The results and runtimes of our experiments are presented in Sect. 6.5.

```
50 function fApproxRQA
51 %FAPPROXRQA approximate RQA (C) Spiegel et al.
52
53     % load Potsdam time series
54     x = load('../Data/temp_pdm_1893-2011.txt');
55
56     % set (predefined) parameters
57     eps = 1; m = 5; tau = 3; minL = 2;
58
59     x = fTDE(x,m,tau);          % time delay embedding
60     [n,~] = size(x);            % length of trajectory
61
62     x1 = fDiscrete(x,eps);      % discretized trajectory
63     x2 = fTDE(x1,minL,1);       % trajectory embedding
64     x3 = fTDE(x1,minL+1,1);     % trajectory embedding
65
66     pp1 = fPProximities(x1);    % pairwise proximities
67     pp2 = fPProximities(x2);    % pairwise proximities
68     pp3 = fPProximities(x3);    % pairwise proximities
69
70     ss1 = pp1;                  % stationary states
71     ss2 = fSStates(x1,x2);      % stationary states
72     ss3 = fSStates(x1,x3);      % stationary states
73
74     % compute approximate RQA measures
75     RR  = pp1/(n*n);
76     DET = (minL*pp2 - (minL-1)*pp3) / (pp1 + 10^-10);
77     L   = (minL*pp2 - (minL-1)*pp3) / (pp2 - pp3);
78     LAM = (minL*ss2 - (minL-1)*ss3) / (ss1 + 10^-10);
79 end
```

6.5 Empirical Results

The goal of our empirical evaluation is twofold: (i) we assess the runtime of original and approximate RQA measures for relatively long time series (with about a million data points); and (ii) investigate the correlation between original and approximate RQA measures for the purpose of finding transitions in time series streams using the sliding window technique. Both experiments are performed on the same real-life data set described in Sect. 6.5.1.

6.5.1 Data

For illustrating the approximation approach and to evaluate it we use the measured time series of hourly air temperature in Potsdam [16], which covers the period from 1893 until 2014 and contains 1,069,416 data points. This time series is one of the longest, non-interrupted, hourly climate records in the world. The Potsdam time series is divided into two intervals (1893–1974 and 1975–2014), because the warming trend of the annual mean temperature shows an abrupt change in 1975 [16]. However, recurrence quantification analysis has shown that, in contrast to longer time-scales, the short-term dynamics, and, thus, the short-term weather predictability, has not (yet) changed due to climate change [16]. Moreover, between 1975 and 1976 the measurement protocol has changed from manual to electronic recording. Such changes could be systematically influencing the measurements and should be visible by recurrence quantification. We, therefore, apply a windowing approach in order to investigate a potential shift in the recurrence properties after 1975. This approach can be further used to detect regime transitions in the local weather regime of Potsdam, but this is focus of a separate future study.

6.5.2 Experimental Protocol

We conducted experiments on (i) the runtime and (ii) the accuracy between original and approximate RQA measures.

Our experiments on (i) runtime were performed on different intervals of the Potsdam time series, namely 1893–1974, 1975–2014, and 1893–2014, as well as on yearly intervals (sliding window analysis). Before analysis the time series has been normalized to zero mean and standard deviation one. The original RQA measures have been calculated using three different implementations, basing on C++ for (i) single thread and (ii) multi thread CPU, and on Python (pyRQA) for (iii) GPU computing [16, 25]. The approximate RQA measures have been calculated by our own MATLAB implementation as described in Sect. 6.4.5, while the runtime has been determined for each measure individually (using one CPU). The runtime comparison

of both original and approximate RQA measures for all three time intervals can be found in Sect. 6.5.3.

Our evaluation on (ii) the accuracy between original and approximate RQA measures is presented in Sect. 6.5.4. For these experiments we considered the Potsdam time series in its entire length (1893–2014) and slide a '1-year' window with '1-year' step size from the beginning to the end. This approach is often referred to as sliding window technique and is commonly used to detect transitions in time series [20]. For each window we compute both original and approximate RQA measures (and using different recurrence thresholds to demonstrate the effect of the threshold), which gives us the temporal changes of the RQA measures under study. Given these temporal changes, we are able to compare the variations of the original and approximate RQA measures. For comparison we are using the Pearson correlation coefficient, the root mean square error, and the relative root mean square error. A high correlation, i.e., both measures vary in a similar way would confirm our hypothesis that the loss in accuracy is still a reasonable trade-off with the gain of speed and that the proposed approximations can be used to find transitions in time series streams.

6.5.3 Results on Runtime

Table 6.2 shows the runtimes of various RQA implementations. It is important to mention that the runtimes are merely an indicator for the performance of the examined implementations, since the experiments were performed under different conditions (using various hardware setups and programming languages).

The runtime experiments were conducted on a computer cluster (PIK HLRS2015—Lenovo/IBM NeXtScale nx360M5), consisting of compute nodes with Intel Xeon E5-2667v3 2×8 core CPUs at up to 3.2 GHz and 64 GB main memory. It furthermore includes NVIDIA Tesla K40 nodes that provide GPU processors running at up to 745 GHz, 2880 stream processors and each supplied with 12 GB of memory. The cluster runs on a 64-bit version of Suse Linux Enterprise Server 11 SP3 with version 7.0.28 of CUDA. The CUDA platform was utilized for the experiment that employ the OpenCL/Python implementation. One CPU node of the cluster was

Table 6.2 Runtime (in sec) for RQA calculation for Potsdam temperature time series

Data set	1893–1974	1975–2014	1893–2014
Data points	718,776	350,640	1,069,416
Single thread (CPU)	9,978.00	2,373.00	22,067.00
16 thread OpenMP (CPU)	782.00	188.00	1,743.00
OpenCL (1x GPU)	439.00	104.00	995.00
Approximation (CPU)	1.83	0.79	2.88

Table 6.3 RQA results for Potsdam time series for three different epochs calculated using maximum-norm, embedding dimension 5, embedding delay 3, threshold 0.75, and Theiler window 0

Data set	1893–1974	1975–2014	1893–2014
Data points	718,776	350,640	1,069,416
RR	0.15	0.15	0.15
RR approx.	0.12	0.12	0.12
DET	0.92	0.92	0.92
DET approx.	0.89	0.90	0,89
L	7.6	7.7	7.7
L approx.	6.9	7.2	7.0
LAM	0.96	0.95	0.96
LAM approx.	0.91	0.91	0.91

exploited for the single-thread and 16 core multi-thread (OpenMP) implementation, using C++ programming language. The approximation experiment was performed on the same hardware using MATLAB 2011b.

Although the runtime experiments were conducted with varying hardware setups and programming languages, the results give some indication of the speed-up factor achieved by our proposed approximation techniques. However, the approximation error is not to be neglected. Table 6.3 demonstrates the results for the original and approximate RQA measures, which also confirm the previous climatological findings and interpretations given in [16].

Although the results presented in Table 6.3 slightly differ for the original and approximate RQA measures, it is more important that the variation (tendency) of these measures is similar. This is considered in more detail in the next section.

6.5.4 Results on Correlation

In several applications, such as RQA based transition detection, the absolute values of the RQA measures are less important than the tendency of their variation with time. Here we compare the ability of the approximation approach to uncover variations that are similar to ones found by the exact measures. We apply both (exact and approximate) RQA measures to the Potsdam temperature series using a sliding windowing technique (non-overlapping windows with length of 1 year), which is the standard approach for detecting transitions or regime changes.

In Fig. 6.4 we find a similar variation between the RQA measures calculated using the exact as well as the approximation approaches. Figure 6.4a was derived with the same recurrence threshold $\varepsilon = 0.75$ that was taken for the analysis of the whole Potsdam time series (see Table 6.3). This threshold is considered as baseline here, meaning that the exact RQA measures show desired characteristics. Visually, the most

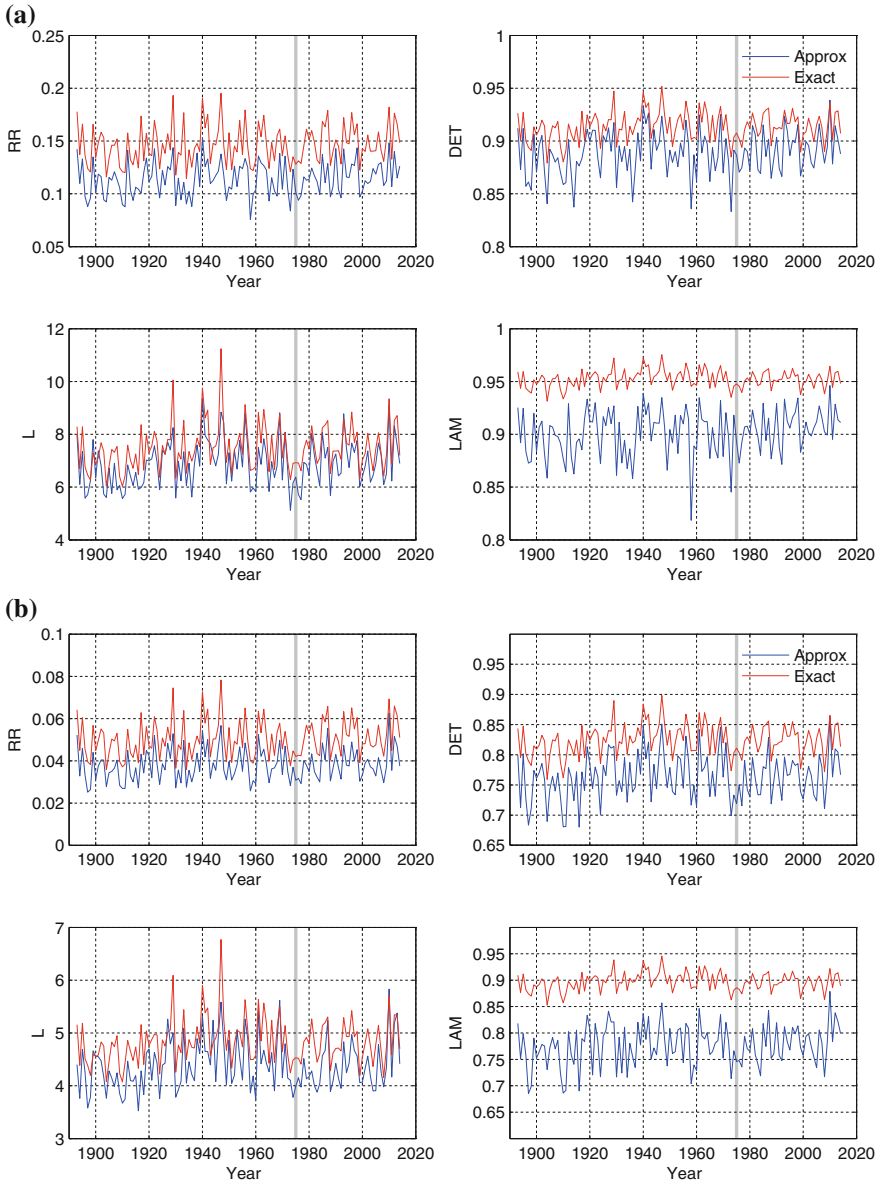

Fig. 6.4 Windowed RQA results for Potsdam time series calculated using a non-overlapping sliding window of length 1 year, maximum-norm, embedding dimension 5, embedding delay 3, Theiler window 0 and (**a**) threshold 0.75, (**b**) threshold 0.5. **a** Baseline similarity threshold: $\varepsilon = 0.75$, **b** Lower Similarity Threshold: $\varepsilon = 0.5$

Table 6.4 Correlation coefficient, root mean square and relative error between exact and approximate RQA measures for Potsdam temperature series as shown in Fig. 6.4

$\varepsilon = 0.75$	Correlation	RMSE	rel. RMSE (%)	$\varepsilon = 0.5$	Correlation	RMSE	rel. RMSE (%)
RR	0.79	0.034	17	RR	0.94	0.013	16
DET	0.81	0.028	3	DET	0.88	0.064	7
L	0.84	0.79	7	L	0.82	0.526	8
LAM	0.67	0.054	6	LAM	0.78	0.124	13

similar variation (and amplitude) is for the measure *L*, followed by *DET*, whereas the amplitude of the *LAM* measure has the largest deviation. This is quantitatively confirmed by the correlation coefficient and the error measures (Table 6.4). The highest correlations, thus the best coincidence of the variation, have the measures *L* and *DET*, whereas *LAM* has the lowest correlation. In contrast, the relative error is smallest for *DET* but largest for *RR*.

The values of the correlation and the errors can be controlled by changing the threshold ε (Fig. 6.5). For *L* and *LAM*, we find the best correlation for recurrence thresholds between 0.4 and 0.6 in units of the signal's standard deviation. The root mean square error decreases for *L* where it increases for *LAM* with decreasing threshold. The threshold that leads to the overall best correlation is $\varepsilon = 0.5$ (Fig. 6.4b), the corresponding windowed analysis is illustrated in Fig. 6.4b. Hence, if we choose this threshold for the sliding window analysis, then the mean correlation between the exact and the approximate RQA measures is greater than 0.85, which indicates a very strong linear relationship. But how does the choice of a lower threshold influence the sliding window analysis? Visually, the corresponding exact measures in Fig. 6.4 vary in a similar fashion, they only obey a different scale. This observation is confirmed by Table 6.5, all corresponding exact measures have a correlation of about 0.99 for the baseline threshold $\varepsilon = 0.75$ compared with the lower threshold $\varepsilon = 0.5$. In summary, the choice of a lower threshold can keep nearly the entire information on the variation of the exact RQA measures, and at the same time can significantly improve the approximate RQA in that the correlation between the exact and the approximate RQA measures increases.

From the climatological point of view, both approaches reveal some variation in the dynamics represented by the temperature time series. A general shift after the time point of the change in recording procedure 1975 is not visible. However, besides several short periods of decrease in the measures, around 1975 a clear drop can be identified and might be more an indication of a sudden change in the general climate regime by passing a tipping point [26] than of a change in the recording procedure. A future study will investigate these variations more systematically and should also consider significance tests [10].

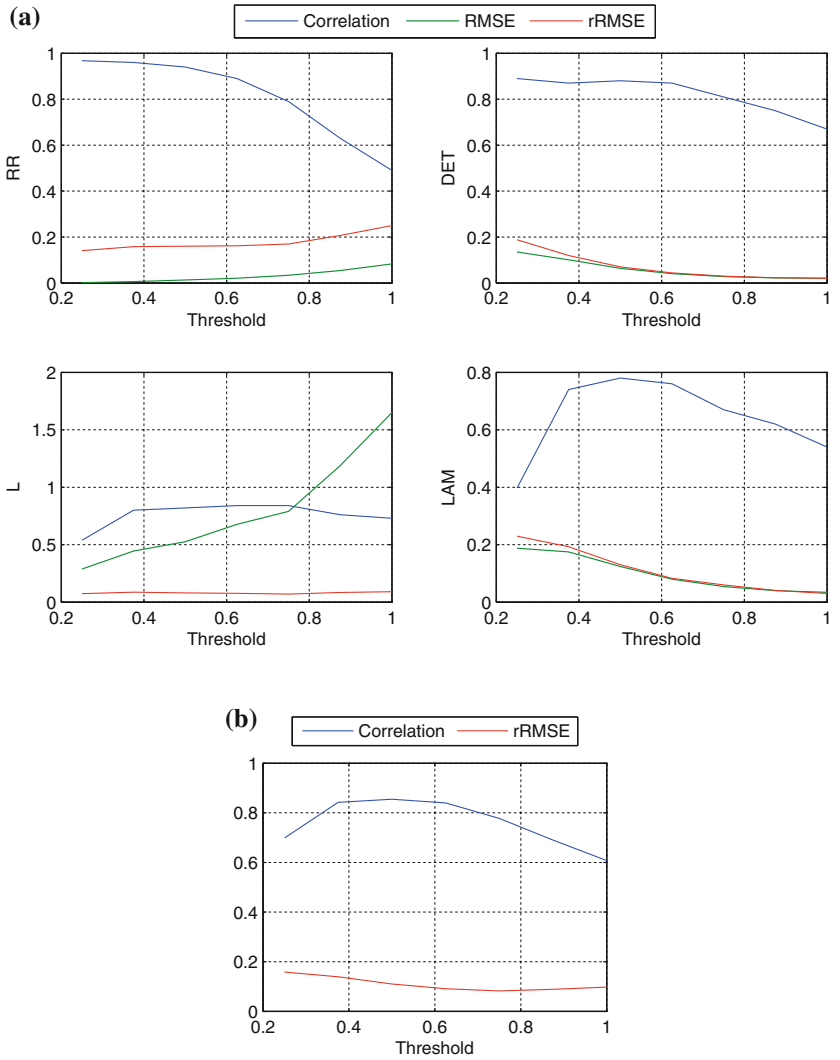

Fig. 6.5 Correlation coefficient, root mean square error (RMSE), and relative RMSE between exact and approximate RQA measures with varying thresholds ε, for our sliding window analysis of the Potsdam temperature series as shown in Fig. 6.4. **a** Individual RQA measures, **b** Mean over all RQA measures from (**a**)

6.6 Conclusion and Future Work

This chapter extends our theoretical work on approximate recurrence quantification analysis (aRQA) [17] and includes some practical considerations that occur when analyzing real-life data such as the Potsdam temperature time series [16]. We have

Table 6.5 Correlation between **exact** RQA measures for threshold $\varepsilon = 0.75$ and **exact** RQA measures for $\varepsilon = 0.5$ for Potsdam temperature series as shown in Fig. 6.4

	RR	DET	L	LAM
Correlation	0.9953	0.9915	0.9877	0.9942

not only discussed the formulation of diagonal line based measures by means of pairwise proximities (*PP*) [17], but also introduced our novel idea of stationary states (*SS*) that enables us to reformulate vertical line based RQA measures. In addition to our new formulation of the original RQA measures, we furthermore presented an efficient implementation that allows fast computation of the approximate RQA measures based on histograms.

Our experiments on relatively long time series (with about a million measurements) demonstrated that the proposed approximation is not only up to four orders of magnitude faster than single thread (exact) computations, but also gives results that are very close to the original measures. Furthermore, we were able to show that our approximate RQA measures strongly correlate with the corresponding exact RQA measures (when applying the sliding window technique) and, therefore, can be used for an efficient transition detection. The presented empirical results are also in agreement with our theoretical analysis [17] in that the error of the approximation is decided by the discretization or strictly speaking depends on the similarity threshold and the distribution of the data under study.

In future work we are going to investigate the discretization more deeply and develop time series representations that enable us to bound the approximation error. Moreover, we will transfer our idea of pairwise proximities (*PP*) and stationary states (*SS*) to cross recurrence plots (CRPs) and corresponding measures.

References

1. N. Marwan, A historical review of recurrence plots. Eur. Phys. J. Spec. Topics **164**(1), 3–12 (2008)
2. C.L. Webber Jr., N. Marwan, A. Facchini, A. Giuliani, Simpler methods do it better: success of recurrence quantification analysis as a general purpose data analysis tool. Phys. Lett. A **373**, 3753–3756 (2009)
3. S. Spiegel, J.B. Jain, S. Albayrak, A Recurrence Plot-Based Distance Measure, vol. 103 (Springer, Cham, 2014), pp. 1–15
4. S. Spiegel, D. Schultz, S. Albayrak, BestTime: Finding representatives in time series datasets. Lecture Notes in Computer Science: Artificial Intelligence, vol. 8726, pp. 477–480 (2014)
5. F. Hasselman, Classifying acoustic signals into phoneme categories: average and dyslexic readers make use of complex dynamical patterns and multifractal scaling properties of the speech signal. PeerJ **3**, e837 (2015)
6. M.A.F. Harrison, M.G. Frei, I. Osorio, Detection of seizure rhythmicity by recurrences. Chaos **18**(3), 033124 (2008)

7. N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, J. Kurths, Recurrence plot based measures of complexity and its application to heart rate variability data. Phys. Rev. E **66**(2), 026702 (2002)
8. A. Giuliani, M. Tomasi, Recurrence quantification analysis reveals interaction partners in paramyxoviridae envelope glycoproteins. Proteins: Struct. Func. Genetics **46**(2), 171–176 (2002)
9. S. Spiegel, *Discovery of Driving Behavior Patterns* (Springer, Cham, 2015), pp. 315–343
10. N. Marwan, S. Schinkel, J. Kurths, Recurrence plots 25 years later: gaining confidence in dynamical transitions. Europhys. Lett. **101**, 20007 (2013)
11. J.F. Donges, R.V. Donner, M.H. Trauth, N. Marwan, H.J. Schellnhuber, J. Kurths, Nonlinear detection of paleoclimate-variability transitions possibly related to human evolution. Proc. Natl. Acad. Sci. **108**(51), 20422–20427 (2011)
12. G. Litak, A.K. Sen, A. Syta, Intermittent and chaotic vibrations in a regenerative cutting process. Chaos Solitons Fractals **41**(4), 2115–2122 (2009)
13. M.C. Romano, M. Thiel, J. Kurths, I.Z. Kiss, J. Hudson, Detection of synchronization for non-phase-coherent and non-stationary data. Europhys. Lett. **71**(3), 466–472 (2005)
14. Y. Hirata, K. Aihara, Identifying hidden common causes from bivariate time series: a method using recurrence plots. Phys. Rev. E **81**(1), 016203 (2010)
15. J.H. Feldhoff, R.V. Donner, J.F. Donges, N. Marwan, J. Kurths, Geometric signature of complex synchronisation scenarios. Europhys. Lett. **102**(3), 30007 (2013)
16. T. Rawald, M. Sips, N. Marwan, D. Dransch, *Fast Computation of Recurrences in Long Time Series*, vol. 103 (Springer, Cham, 2014), pp. 17–29
17. D. Schultz, S. Spiegel, N. Marwan, S. Albayrak, Approximation of diagonal line based measures in recurrence quantification analysis. Phys. Lett. A **379**(14–15), 997–1011 (2015)
18. J.-P. Eckmann, S. Oliffson, Kamphorst, D. Ruelle, Recurrence plots of dynamical systems. Europhys. Lett. **4**(9), 973–977 (1987)
19. N.H. Packard, J.P. Crutchfield, J.D. Farmer, R.S. Shaw, Geometry from a time series. Phys. Rev. Lett. **45**(9), 712–716 (1980)
20. N. Marwan, M.C. Romano, M. Thiel, J. Kurths, Recurrence plots for the analysis of complex systems. Phys. Rep. **438**(5–6), 237–329 (2007)
21. J.P. Zbilut, C.L. Webber Jr., Embeddings and delays as derived from quantification of recurrence plots. Phys. Lett. A **171**(3–4), 199–203 (1992)
22. C.L. Webber Jr., J.P. Zbilut, Dynamical assessment of physiological systems and states using recurrence plot strategies. J. Appl. Physiol. **76**(2), 965–973 (1994)
23. M. Thiel, M.C. Romano, P.L. Read, J. Kurths, Estimation of dynamical invariants without embedding by recurrence plots. Chaos **14**(2), 234–243 (2004)
24. N. Marwan, J.F. Donges, Y. Zou, R.V. Donner, J. Kurths, Complex network approach for recurrence analysis of time series. Phys. Lett. A **373**(46), 4246–4254 (2009)
25. T. Rawald, pyRQA (2015)
26. M. Scheffer, J. Bascompte, W.A. Brock, V. Brovkin, S.R. Carpenter, V. Dakos, H. Held, E.H. van Nes, M. Rietkerk, G. Sugihara, Early-warning signals for critical transitions. Nature **461**(7260), 53–59 (2009)

Chapter 7
Splayed Recurrence Analysis of Iterated Dynamical Systems

Charles L. Webber, Jr.

Abstract Splayed Recurrence Analysis (SRA) is a new method for identifying and quantifying recurrent events in iterated systems. The technique is fully applicable to difference equations, Poincaré sections of continuous time series, and independent random events. Inspiration for SRA comes from American roulette wheel gaming. It has been postulated that non-random wheel determinism is introduced by unbalanced wheels (mechanical) and non-random repeated motions of house spinners (human). Primary data were taken from actual roulette outcomes in which ball landing slots were reported sequentially according to spin orders. These data were stored in a matrix [slot #, spin #] and lines were passed through all possible pairs of points in the matrix and extrapolated to the border. Centers of points falling exactly on these extended lines, including the initial pair, were scored as recurrent points. Necessarily, there were gaps between points which led to point-to-point intervals being splayed-out. Six variables were extracted from the recurrent points comprising lines: (1) number of recurrent points per line; (2) intervals between recurrent points; (3) lengths of lines; (4) slopes of lines; (5) entropy of line lengths; (6) density of recurrent points. Besides the American roulette data, these SRA strategies were also applied to natural random numbers, chaotic models, and natural phenomenon. No differences could be detected for roulette data and naturally occurring random processes. But SRA was able to detect non-random structures in mathematically chaotic systems as well as in eruption times of the Old Faithful geyser. Because the methodology does not depend upon embeddings and delays as required for nonlinear analyses, SRA is classified as fully linear.

C.L. Webber, Jr. (✉)
Health Sciences Division, Department of Cell and Molecular Physiology,
Loyola University Chicago, Maywood, IL 60153-3328, USA
e-mail: cwebber@luc.edu

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8_7

7.1 Introduction

This study was inspired by the challenge of a roulette wheel enthusiast who claimed the ability to predict future outcomes of roulette spins if given the results of a short series of runs. Casinos routinely post some 20 outcomes of previous rounds (spin #s), enticing gamers to detect patterns in the sequential numbers (slot #s). Indeed, there is a vast literature on beating the roulette wheel by finding patterns, using mathematical probabilities, ballistic physics and even applying chaos theory [1–4]. In the spirit of expert systems analysis, I sat down with the enthusiast and had him slowly describe and diagram his methodology for discovering patterns in sequences of actual roulette data. In return I challenged him with 10 data sets of slots from sequential roulette spins. Each set consisted of 30 spins, but I only gave him the first 10 slot outcomes. This enabled me to score the accuracy of his predictions among the following 20 hidden spins. The results are discussed later in this chapter.

At first glance, it appears that Roulette would produce sequences of totally random numbers, rendering future predictions impossible. The game is played by the croupier spinning the rotor in one direction (clockwise or counter clockwise) and rolling the ball in the opposite direction along the wheel track [5]. When the velocity of the ball slows, the ball falls onto the rotating apron, is jostled by deflectors, and finally lands into one of 38 slots discriminated by frets (American roulette). As shown in Fig. 7.1, the pockets are numbered from 1 through 36, half

Slot	Order
0	0
2	1
14	2
35	3
23	4
4	5
16	6
33	7
21	8
6	9
18	10
31	11
19	12
8	13
12	14
29	15
25	16
10	17
27	18

Slot	Order
00	19
1	20
13	21
36	22
24	23
3	24
15	25
34	26
22	27
5	28
17	29
32	30
20	31
7	32
11	33
30	34
26	35
9	36
28	37

Fig. 7.1 The American roulette wheel with 38 slots (18 *red*, 18 *black* and 2 *green*). The physical location of the slots are mapped (linearized) into 38 ascending integer values (0–37) in the counterclockwise direction (CCW) starting with *green* slot 0 and ending with *black* slot 28

red and half black, and with two green zeros (0 and 00), but are not positioned in any kind of numerical order. But two interfering factors may introduce bias (so-called determinism) into an otherwise random system [4]. First, wheel bias might cause certain slots to recur more frequently than others (like loaded dice). This is not unlike spinners in children's games often landing on the same (recurrent) number if the playing board is warped. Second, the spinning motions of the croupier are periodic, non-random and unique to the person. This may be one reason why croupiers are frequently rotated among different roulette wheels in the house. The gaming enthusiast defended these deterministic features using principles from nonlinear dynamics [6].

7.2 Graphical Theory

The roulette enthusiast showed me how his strategy for beating the house edge of 5.26 % for American roulette [2] was based on graphical theory. Figure 7.2 illustrates the methodology. Graphs (roulette matrices) are constructed by plotting points at the intersection of slot outcomes (vertical axis) and round spins (horizontal axis). Lines are defined as linear projections that pass exactly through the centers of three or more points. Lines passing through merely two points are rejected as being trivial. Using this simple rule, the data of Fig. 7.2 (22 spins with 38 possible slots) has a total of 16 recurrent points (circled) comprising 6 lines. For example, the first line, L1(3), is defined by 3 points [27, 1], [19, 3] and [0, 7] which are unevenly spaced. The second and third lines, L2(3) and L3(3), also pass through 3 points each as do the last two lines, L5(3) and L6(3). The fourth line, L4(4), however, passes through 4 points (rarer). Clearly, the six defined lines have different slopes, some lines share common points: L2 and L5 at [22, 11]; L3 and L4 at [3, 9]; L5 and L6 at [00, 16]. And six points (not circled) in the matrix are never recurrent: [26, 8]; [0, 10]; [34, 13]; [29, 14]; [25, 21]; and [11, 22].

Again, it is worth noting that the spacings of recurrent points are uneven and irregular. From the recurrence perspective, points (or hits) can be thought of as recurrent points (on lines) whereas missed points (off lines) define gaps between recurrent points. It is in this sense that the recurrent points are splayed out along linear trajectories defined by the slopes of lines each passing through a minimum of 3 points.

7.3 Algorithmic Implementation

Splayed recurrence analysis (SRA) consists of a suite of 9 programs (SRB, SRC, SRN, SRI, SRS, SRL, SRE, SRD and SRH, described below). Two key parameters control the size of the roulette matrix: (1) the number of possible outcomes (# slots)

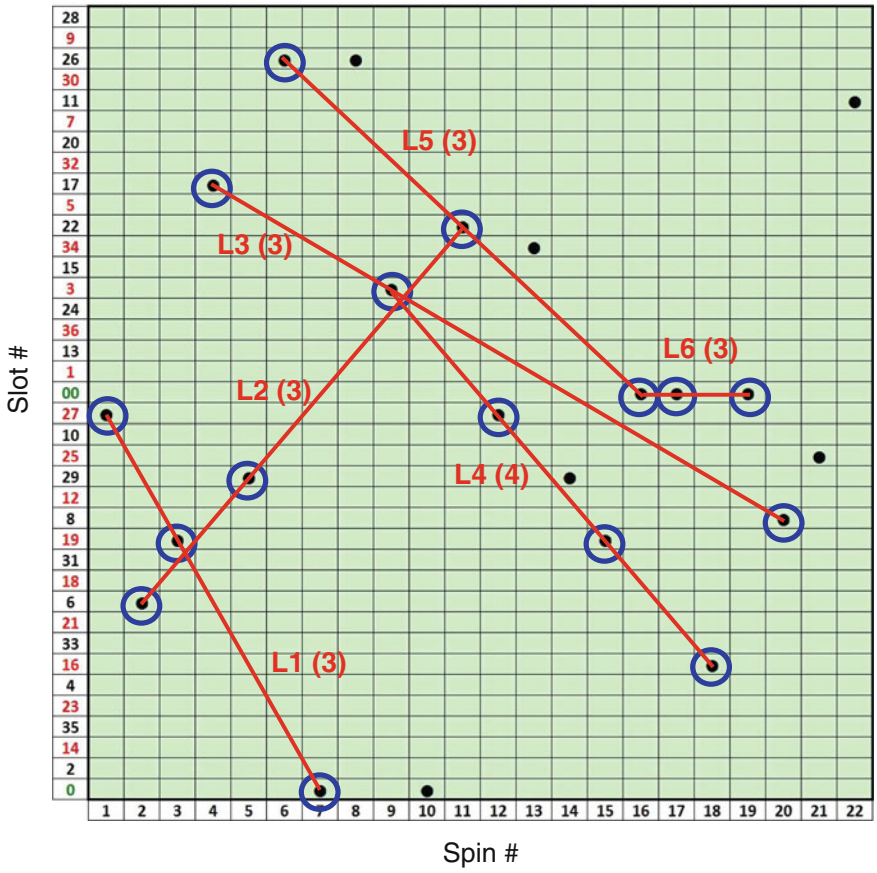

Fig. 7.2 Linear trajectories passing exactly through the centers of 3 or more points define splayed-out recurrent points (*circled*) with uneven spacing and variable slopes. Slot numbers (slot #) are given in their physical counterclockwise sequence order (spin #)

and (2) the length of the recurrence window (# spins). A third parameter adds statistical rigor to the analysis: (3) the number of sequential adjacent windows.

For roulette wheel data, the number of outcome slots or bins is fixed at 38 (American roulette). However, to generalize the methodology to any other dynamical system, program SRB (**B**in rescaling of input vector elements) is used to confine amplitudes of the floating point input into fixed integer bins of any selectable number (not just 38). Outputs of program SRB are then accepted as inputs to program SRC (**C**oordinates of 3 or more points falling alone linear trajectories) which reports the x, y coordinates of all recurrent points within each window of preselected window length.

Proof of concept was first explored by applying SRA to natural random numbers to check for chance recurrences in roulette matrices. A long series of 102,400

Table 7.1 Histograms of number of recurrent points per line based upon the number of outcomes (# slots) and the size of the window (# spins). Within most horizontal color bands, the number of bin counts decreases as the number of recurrent points per line increases

RANDOM HITS IN RECURRNECE MATRIX

Decreasing # Slots

Increasing # Spins

# Slots	# Spins	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
32	400	50869	4628	556	119	130	226	282	445	503	591	590	490	437	383	274	197	133	70	39	22	12	3	1	1	2
32	300	68126	6161	985	598	738	955	1073	1064	864	802	551	374	236	122	66	27	20	1	1	0	1				
32	200	103709	11050	3419	2603	2371	1886	1289	816	438	209	108	45	9	4	4	1	1								
32	100	107910	15266	4801	1959	774	296	76	31	3	0	1														
32	50	61360	7893	1485	281	50	7																			
32	25	30155	2653	265	25	1																				
64	400	62207	5734	1813	1485	1400	1181	854	605	297	155	87	34	10	6	3										
64	300	85734	9363	3668	2465	1707	1027	517	235	107	32	7	5	1												
64	200	137303	16574	5090	2063	858	306	99	23	9	2															
64	100	79928	8833	1645	322	55	9	1	1																	
64	50	40423	3143	299	36	1																				
64	25	18540	854	54	2	1																				
128	400	77813	8786	3248	1522	659	259	80	33	12	1															
128	300	109637	12177	3247	1036	297	89	25	5	1																
128	200	100047	9433	1726	365	49	14	3																		
128	100	51349	3528	336	43	1																				
128	50	23923	1033	49	2																					
128	25	12576	281	9																						
256	400	94460	8531	1593	338	59	16	1																		
256	300	92137	6757	967	159	22	3																			
256	200	62294	3686	381	40	6																				
256	100	29408	1128	53	4																					
256	50	13114	274	6																						
256	25	9911	121	1																						

Number of Points per Line

integers ranging from 0 to 255 (step 1) were obtained from genuine random numbers (HotBits) generated from natural beta radioactive decay of Cæsium-137 [7]. From this master set, four integer data sets were constructed ranging from 0 to different maxima (31, 63, 127, 257) for different number of slots (32, 64, 128, and 256 respectively). SRA was run on all four data sets were using six different window lengths (25, 50, 100, 200, 300, and 400 spins). The resultant 24 histograms are reported as bin counts in Table 7.1 and color coded according to number of outcomes (# slots) chosen. As anticipated, the number of chance recurrent points per line (horizontal axis) increases with increasing # spins and decreasing number of slots (vertical axes). Figure 7.3 captures this principle graphically. SRA results from these natural random numbers were used as controls for all dynamical systems studied (see below). In no cases were any of the data series scrambled, but admittedly this would be another way in future studies to shuffle or randomize the data sequences [8].

7.4 SRA Variables

Six recurrence variables are derived directly from the coordinates of recurrent points for each line. Each variable is computed by six separate programs using as inputs the outputs generated from program SRC (described in Sect. 7.3). Thus,

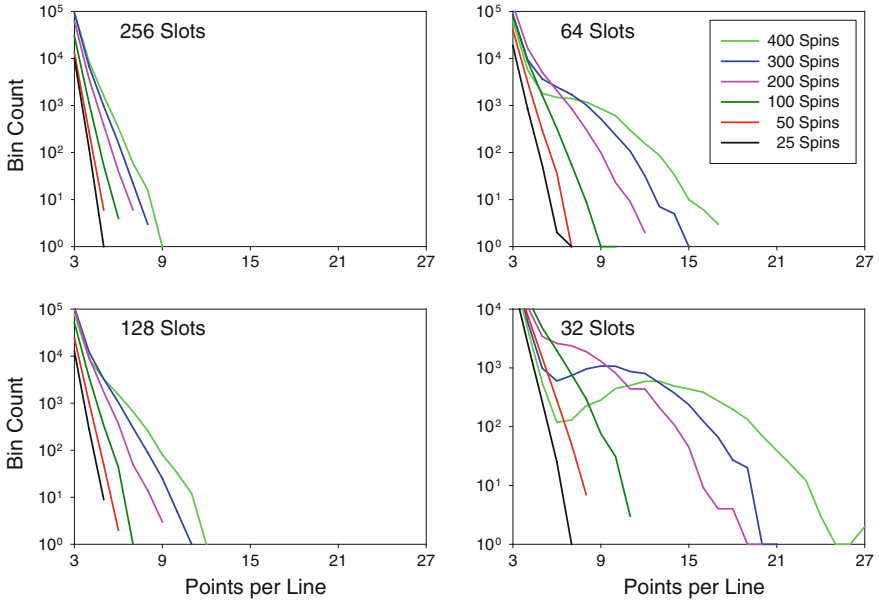

Fig. 7.3 Histograms of data from Table 7.1 color-coded by # spins and clustered by # slots selected

(1) program SRN computes the **N**umber of recurrent points per line (NUM). (2) Program SRI computes the **I**ntervals between recurrent points (INT). (3) Program SRL computes the **L**engths of lines defined by recurrent points (LEN). (4) Program SRS computes the **S**lopes of lines defined by recurrent points (SLP). (5) Program SRE computes the **E**ntropy of line-length distributions (ENT). And finally, (6) program SRD computes the **D**ensity of recurrent points over sequential epochs (DEN). Interval lengths and line lengths are computed as the diagonal distances between the x, y coordinates of interval points or line end-points, respectively, according to the following formula.

$$\text{INT or LEN} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{7.1}$$

A ninth and last program completes the SRA suite of programs. Program SRH computes the **H**istogram distributions of the first four SRA variables over sequential epochs. The only exceptions are the fifth and sixth variables, ENT and DEN, which are not a distributed variables, but rather compute as single values for each epoch (described below).

7.5 Test Data Sets

SRA strategies were applied first to the random and roulette data and next to six other systems (8 systems total). In order to compare all systems, the number of slots was held constant at 38 (Roulette defined) and the number of spins was held constant at 50. As demonstrated by Table 7.1 and Fig. 7.3, there is nothing special about numbers 38 and 50, of course, they were merely selected as proof of concept. Thus the roulette matrix was consistently sized as RM [38, 50] for the purposes at hand.

The eight systems were classified into one of two data sets. The data first set comprised four steady-state systems and the second set consisted of four transient signals. The four steady-state systems each possessed 14,979-point vectors and 299 epochs as dictated by the window length of 50 are shown in Fig. 7.4 (only the first 450 points are plotted for each system). These systems included HotBit natural random numbers (RANDOM), slot outcomes of roulette wheel spins (WHEEL) [9], the x-variable of the Hénon chaotic attractor (HENON, formulae 2 & 3), and geyser eruption intervals of Old Faithful in Yellowstone National Park (GEYSER) [10]. The vector elements of each system were rescaled into 38 discriminating amplitude bins (not shown) before processing by SRA.

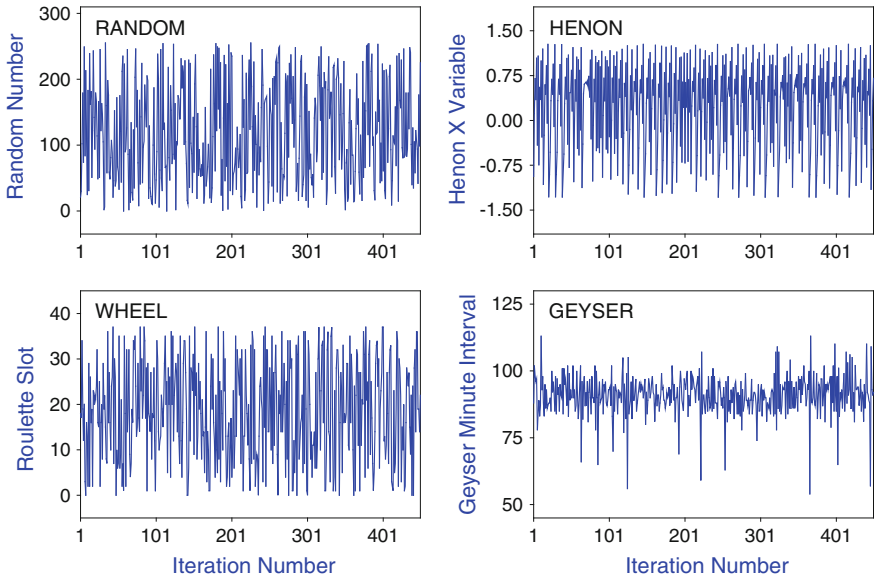

Fig. 7.4 Four steady-state systems each consisting of 14,979 points, only the first 450 which are graphed: RANDOM (HotBit natural random numbers); WHEEL (spin outcomes of American roulette wheel); HENON (the x variable of the Hénon chaotic attractor); GEYSER (geyser eruption intervals of Old Faithful in Yellowstone National Park)

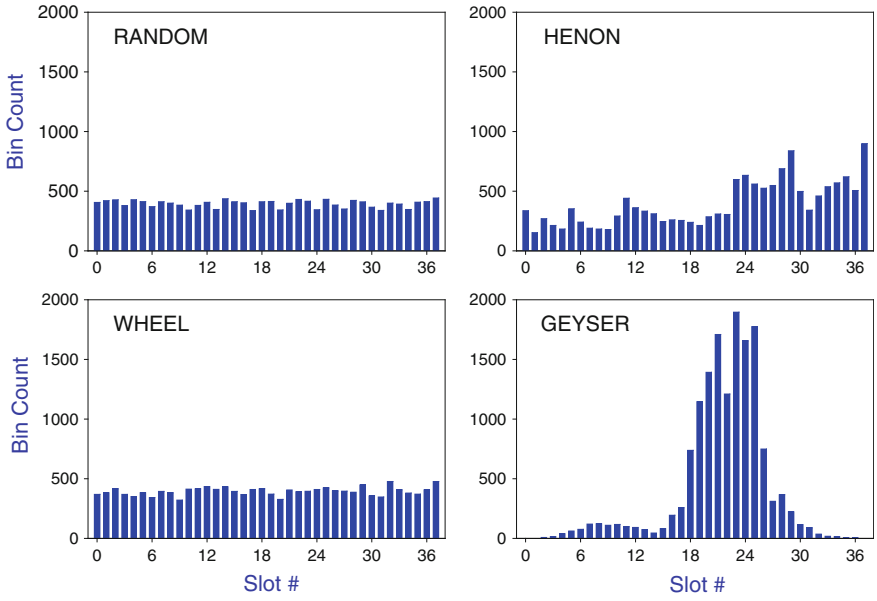

Fig. 7.5 Histograms of the 14,979 integer events comprising each of the four steady-state systems as identified in Fig. 7.4

$$x_{i+1} = y_i + 1 - 1.4\ x_i^2 \tag{7.2}$$

$$y_{i+1} = 0.3x_i \tag{7.3}$$

Prior to SRA, the integer data (14,979 points) of the four steady-state systems were distributed into histograms to check for bias as shown in Fig. 7.5. The distributions of the RANDOM and WHEEL integers are flat indicating lack of "wheel bias." However, the distributions of the HENON and GEYSER data are non-flat or biased with "wheel tilts" as it were. But of course, flat histograms do not rule out deterministic structures. And neither do biased histograms rule in deterministic structures. However, such differences can be discriminated by classical recurrence plots and their quantifications [11, 12].

The four transient systems each possessed 37,446-point vectors and 748 epochs as dictated by the window size of 50 are shown in Fig. 7.6. These systems included HotBit natural random numbers with added sine wave and ramp (RANDET, formula 7.4), vocabulary word in the English text of Genesis (TEXT), the x variable of the adiabatic logistic equation in its period doubling route to chaos (LOGISTIC, formula 7.5) [13], and biceps electromyogram *en route* to muscle fatigue (EMG) [14]. (Actually, the EMG was the single continuous system included in this study.) The vectors of each system were rescaled into 38 discriminating amplitude bins (not shown) before processing by SRA.

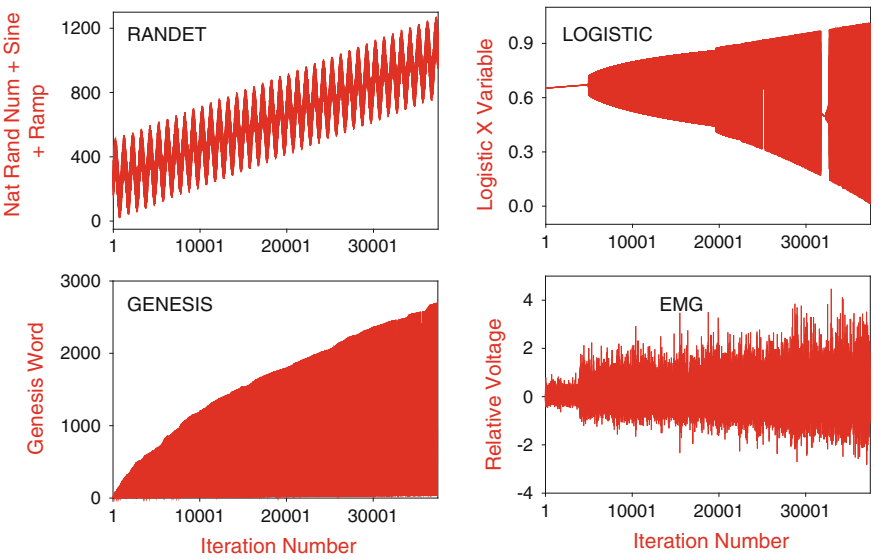

Fig. 7.6 Four transient systems each consisting of 37,446 points, all of which are graphed: RANDET (hot bit natural random numbers with added sine wave and ramp); GENESIS (vocabulary word in the English text of Genesis); LOGISTIC (the x variable of the adiabatic Logistic equation in its period doubling route to chaos); EMG (biceps electromyogram *en route* to muscle fatigue)

$$\mathrm{RANDET_i} = \mathrm{HB_i} + \mathrm{SINE(HB_i)} + \mathrm{RAMP_i} \tag{7.4}$$

where HB = 0 to 255; $\mathrm{RAMP_i} = \mathrm{INT(RAMP_{i-1} + 0.02)}$; $\mathrm{RAMP_0} = 0.0$.

$$\mathrm{x_{i+1}} = \mathrm{ax_i(1 - x_i)} \tag{7.5}$$

where a = 2.8 to 3.93 step 0.00003.

Prior to SRA, the integer data (37,446 points) of the four transient systems were distributed into histograms to check for bias as shown in Fig. 7.7. None of the distributions are flat indicating the presence of "wheel bias." These "wheel tilts" were unimodal (RANDT, GENESIS, EMG) or multi-modal (LOGISTIC) with uniform (RANDET, EMG) or skewed distributions (GENESIS, LOGISTIC).

7.6 Splayed Recurrence Analysis of Test Data Sets

SRA was applied to the two data sets, the first set with four steady-state systems and the second set with four transient signals. For simplicity and easy comparisons, for all eight examples the number of slot outcomes was consistently set to 38 (e.g. # slots in American roulette) and the window length was fixed at 50 (e.g. # spins).

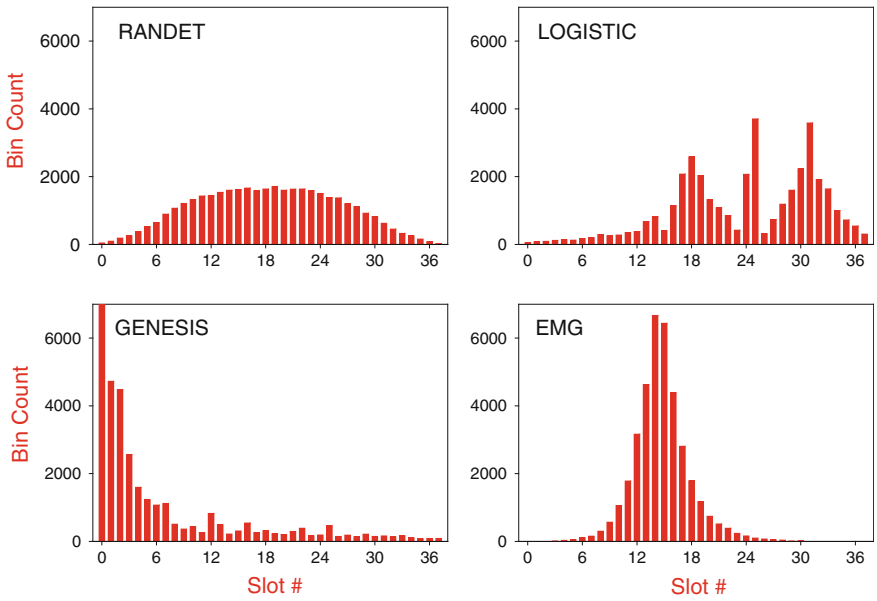

Fig. 7.7 Histograms of the 37,446 integer events comprising each of the four transient systems as identified in Fig. 7.6

Figure 7.8 shows the results displayed in histogram format (identical format as Fig. 7.3). For the steady-state systems (Fig. 7.8 left), no differences could be detected between the natural random numbers (RANDOM), roulette wheel outcomes (WHEEL) or the Hénon chaotic attractor (HENON). However, the Old Faithful eruption times (GEYSER) showed a clear skewing to the right of the logarithmic histogram counts indicative of non-random (deterministic?) structuring of this quasi-periodic and quasi-predictable event of nature. For the transient systems (Fig. 7.8 right) all four signals showed dramatic right skewing because of the

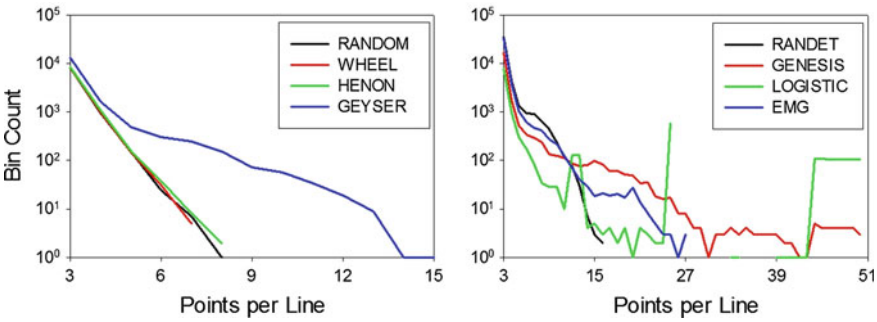

Fig. 7.8 Splayed recurrence analysis of steady-state systems (*left*) and transient systems (*right*). Abbreviations are identical to those used in Figs. 7.5 and 7.6

strong non-random (deterministic?) structuring in each of the signals. Because of the crossing over of the curves it is difficult to rank the degrees of structuring among these systems, but perhaps the Genesis text (GENESIS) possesses the strongest bias. The high bias in the RANDET is due to the addition of two linear (deterministic) signals to the natural random numbers.

7.7 Splayed Recurrence Intervals

The name Splayed Recurrence Analysis stems directly from the fact that recurrent points falling along linear trajectories in the roulette matrix are unevenly distributed. Gaps of all sizes are present depending upon the dynamic under study. Quantification of all the intervals over all combined epochs is best illustrated by histogram distributions of the SRA intervals. The four steady state systems (299 epochs) and four transient systems (748 epochs) can be compared directly (visually) because they have identical horizontal scales and vertical scales proportioned to the different number of epochs (max 4,000/max 10,000 = 299 steady state epochs/748 transient epochs).

Interval histograms for the four steady state systems are shown in Fig. 7.9. Non-paired student t-tests were performed (RANDOM as the reference) to provide statistical rigor ($P < 0.05$ considered significantly different). Each distribution is

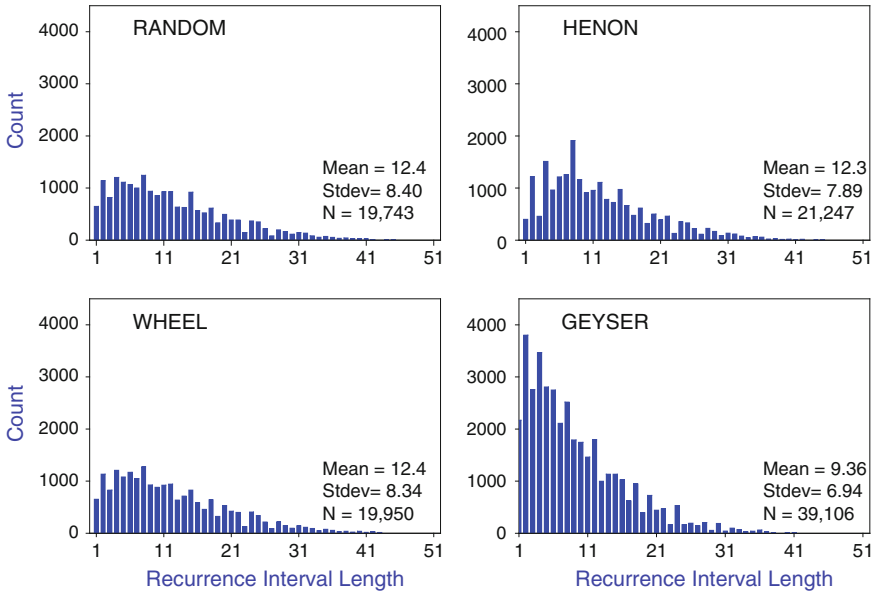

Fig. 7.9 SRA interval histograms of four steady-state systems

skewed to the left (toward the shorter intervals) with tails to the right (toward the longer intervals), but to different degrees.

Recurrence intervals and their distributions comprising the natural random numbers (RANDOM) and roulette wheel outcomes (WHEEL) are essentially identical (P = 0.600). Once again, this result indicates that the roulette wheel is an excellent random number generator (and therefore not predictable). The histogram intervals and distributions for the Hénon chaotic attractor (HENON) are barely significantly different from random (P = 0.050). This suggests that from the linear perspective of SRA, the Hénon chaotic attractor, although fully deterministic mathematically, is near-random statistically (poor sensitivity for chaoticity). Finally, the Old Faithful Geyser data (GEYSER) are strongly significant when compared to the random intervals (P = 0.010).

Interval histograms for the four transient systems are shown in Fig. 7.10. As seen for the steady state systems, each distribution is skewed to the left (toward the shorter intervals) with tails to the right (toward the longer intervals), but to different degrees. The modal intervals of these transient systems are all left (shorter periods) of those encountered with the steady state systems. This effect is due to the disallowance of long recurrences occurring across full transitions. For statistical comparisons, recurrent intervals from each of the four transient systems were compared against SRA random intervals (mean = 12.44; stdev = 8.39; n = 49,620) generated from natural random numbers (38 slots, 50 spins, 748 epochs) using the

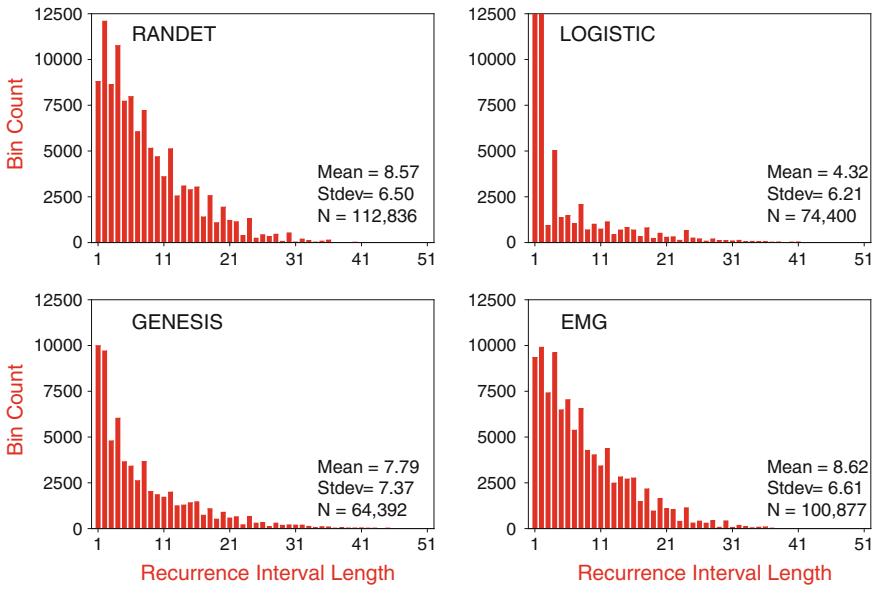

Fig. 7.10 SRA interval histograms of four transient systems. The first two bins of LOGISTIC are off-scale with counts of 34,945 and 16,787 respectively

non-paired Student's t test. In each case (RANDET, GENESIS, LOGISTIC and EMG) the P values were highly significant ($P < 0.00001$) implying they were non-random.

7.8 Line Entropy and Recurrence Density

Possibly the most instructive SRA variable for sorting out deterministic dynamics from stochastic system dynamics is the recurrence density variable (DEN). Alternately, line entropy (ENT) is probably the worst SRA variable for discriminating dynamics. Density is related to recurrence rate and is computed by scoring in a density matrix (identical in size to the roulette matrix) all x, y coordinates identified, epoch by epoch. The density matrix [# slots, # spins] was initially filled with zeros before hits were recorded as ones. As sequential epochs were computed, the density of hits increased as virgin cell hits were added. This analysis is much like watching rain fall on a dry sidewalk. At first the drops are sparsely spread out in a non-grid pattern, but over time the entire sidewalk is wetted leaving no spaces dry [15].

The cumulative line entropy and density results for the four steady-state systems and four transient systems are shown in Fig. 7.11. The line-length entropy data are not very instructive (Fig. 7.11, upper panels) for all maxed out at some 5

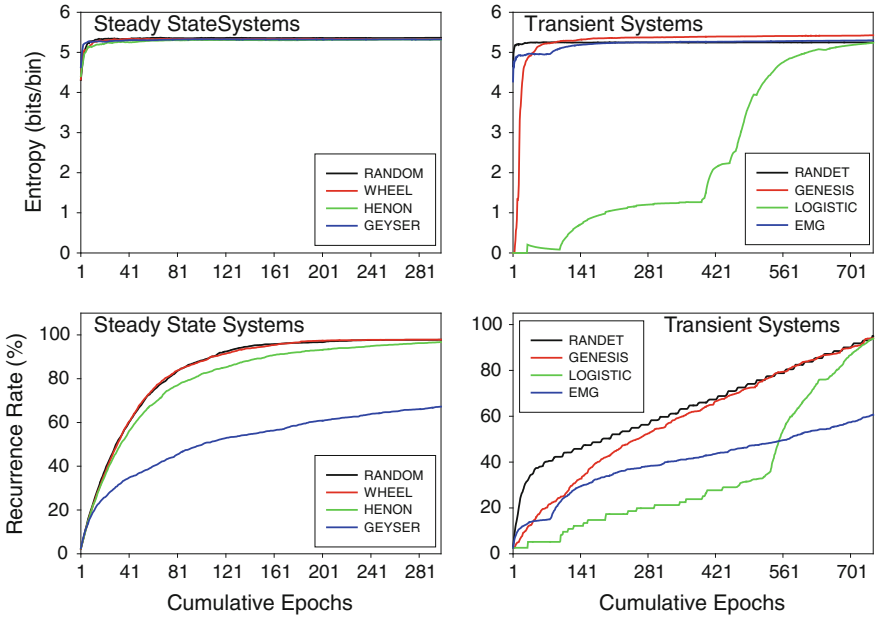

Fig. 7.11 Splayed recurrence analysis density curves over cumulative epochs for steady-state systems (299 epochs, *left*) and transient systems (748 epochs, *right*)

bits/symbol with similar profiles save the LOGISTIC system which was delayed in achieving the max entropy. Conversely, the density values reach or approach 100 % recurrence rate with different slopes for three of the steady-state systems (RANDOM, WHEEL, HENON) in 299 epochs and three of the transient systems (RANDET, GENESIS, LOGISTIC) in 748 epochs (Fig. 7.11, bottom panels). The Old Faithful Geyser (GEYSER) and the electromyographic (EMG) signals on the other hand both have very slow rise times, reaching only about 60 % recurrence rate. Assuming that the slower the increases in density the higher the degree of bias (determinism?), the seven of the eight systems can be ranked as follows (low to high bias): RANDOM/WHEEL (tied rank), HENON, RANDET, GENESIS, GEYSER and EMG. It is not clear how LOGISTIC system with its first low and then high recurrence rate should be ranked.

7.9 Line Lengths and Slopes

The remaining two variables, lengths of lines (LEN) and slopes of lines (SLP) were also studied. The results are presented in Table 7.2. For all systems, steady-state and transient, the distributions of slopes were all Gaussian-like with different standard deviations and ranges (compare with the statistics of recurrence intervals in Figs. 7.9 and 7.10). For the steady-state systems, the line length distributions were Gaussian-like (RANDOM, WHEEL, HENON) or left-skewed (GEYSER). For the transient systems, the distributions were left-skewed (RANDET, EMG) or non-Gaussian-like (GENESIS, LOGISTIC).

7.10 Roulette Wheel Challenge

Finally, we come to the results of the roulette-wheel enthusiast whom I challenged with 10 data sets from actual roulette spins. Each set consisted of 30 slot positions for 30 consecutive spins, but only the first 10 slot outcomes were provided (spins

Table 7.2 Histograms of number of recurrent points per line based upon the number of outcomes (# slots) and the epoch size of the roulette window (# spins)

Statistics on Line Lengths and Line Slopes

System =	RANDOM	SLOT	HENON	GEYSER	RANDET	TEXT	LOGISTIC	EMG
Mean LEN =	26.670	26.565	26.461	22.557	21.543	23.791	29.061	21.449
Stdev LEN =	10.7	10.6	10.7	11.7	12.3	12.6	12.7	12.2
N LEN =	9,190	9,290	9,842	16,219	44,870	21,071	11,066	40,536
Mean SLP =	0.095	0.080	0.066	0.010	0.016	0.060	0.044	0.034
Sdev SLP =	1.96	2.07	2.12	0.81	0.58	1.11	1.41	0.72
N SLP =	9,190	9,290	9,842	16,219	44,870	21,071	11,066	40,536

1–10). Using his own strategies the enthusiast returned his best guesses among the remaining 20 slot outcomes (spins 11–30). The results were scored for the 10 tests (T) as follows: T1 (12 lines, 73 points, 0 matches); T2 (8 lines, 31 points, 1 match); T3 (15 lines, 34 points, 0 matches); T4 (15 lines, 65 points, 2 matches); T5 (12 lines, 43 points, 1 match); T6 (14 lines, 60 points, 0 matches); T7 (13 lines, 40 points, 1 match); T8 (12 lines, 43 points, 2 matches); T9 (14 lines, 40 points, 1 match); and T10 (7 lines, 29 points, 0 matches). The probability of scoring a match by random chance in any one corresponding bin (ignoring the spin order) is 1/38. The probability rises to # points/38 when bin position is ignored. (The number of points guessed exceeds 20 because single points sometimes presented in multiple lines.) Clearly, the cumulative correct score of only 8 correct is below the predicted number of 13 matches by chance alone. These results demonstrate that either the roulette wheel data had no patterns or, if patterns were present, the enthusiast's logic was unable to detect them. All humans are experts at seeing patterns in otherwise random systems and are dogmatic about defending what is "seen"! As is written on Casino Advisor: "Like sportsmen gamblers are superstitious people. Some sportsmen carry lucky charms. They win on their ability but still find comfort in their lucky handkerchief or coin" [16].

7.11 Discussion

Splayed recurrence analysis (SRA) was inspired from the game of American roulette in which a rotating ball falls into one of 38 counter-rotating slots. The methodology can be generalized for any iterated dynamical system by selecting the number of slots and spins for as many epochs as desired. Thus SRA has only 3 parameters: # slots; # spins; and # epochs.

SRA methodology was encoded according to the graphical rules of a roulette-wheel enthusiast. Lines were defined as passing exactly through the centers of a minimum of 3 points within a roulette matrix [# slots, # spins]. The coordinates of all points on lines were collected from which 6 SPA variables were defined: (1) number of recurrent points per line (NUM); (2) intervals between recurrent points (INT); (3) lengths of lines (LEN); (4) slopes of lines (SLP); (5) entropy of line-length distributions (ENT); and (6) density of recurrent points over sequential epochs (DEN). The order of importance of these variables is distinguishing deterministic dynamics from stochastic dynamics or their mixtures is as follows (best to worst): DEN; NUM; INT; LEN; ENT; and SLP.

It must quickly be recognized that SRA is a linear methodology. In this context, one is reminded of the visibility graph (VG) technique of Lacasa et al. [17]. But the two systems are decidedly different. For example, SRA skips intervals to construct lines connecting distant points whereas VG draws lines between paired points at all interval spacings. Possibly SRA might map into VA, but this remains to be proven mathematically.

I am greatly indebted to the roulette-wheel enthusiast. Although his ideas did not pan out for roulette wheel predictions, his graphical algorithm for finding recurrent points within the roulette matrix has herein been shown effective in detecting biased (deterministic?) structures in other non-random and natural test systems. It is concluded that mechanical roulette wheels are carefully (but not perfectly) balanced, nullifying the formation of patterns. And any repeat (deterministic) motions of the croupier are non-repetitive and nullified by the starting positions of the ball and wheel, and ball deflections by the deflectors and frets on the rotating apron.

Future studies could be designed to compare linear SRA against nonlinear RQA (both originated by the author). The operating hypothesis (hunch) is that RQA is probably superior to SRA because it embodies more sophisticated embedding and delay theories [18]. Circular wheel wrap-around could be included in SRA extending lines to points arising beyond the first wheel rotation. Much more work is required to fine tune SRA strategies (theory) as well as apply the methodology to numerous other iterated systems (practice).

7.12 Conclusions

In summary, SPR failed to detect biased structures in both natural random numbers (RANDOM) and roulette slot outcomes (WHEEL). But SRA was good at detecting biases in random numbers to which linear signals were added (RANDET), mathematically chaotic systems (HENON, LOGISTIC); natural systems (GEYSER, EMG); and linguist texts (GENESIS). The link between bias detection and deterministic structures remains to be made. Lastly, SRA is another tool for studying dynamical systems, but it cannot be included in the non-linear tool box simply because SRA is linear. The methodology is integer-based (which constitutes a type of signal filtering) and is designed for iterated systems (e.g. Poincaré sections [19]). SRA may be applicable for continuous systems (e.g. EMG). The lesson learned is that recurrences and their quantifications come in many practical and useful forms.

References

1. M. Small, C.K. Tse, Predicting the outcome of roulette. Chaos **22**, 033150 (2012)
2. A. Snyder, *How to Win at Roulette—Part 1* (Blackjack Forum, 2012), http://www.blackjackforumonline.com/content/how-to-win-at-roulette-2.htm
3. E.O. Thorp, *The Mathematics of Gambling* (Secaucus, N.J., 1984), pp. 43–64
4. L. Scott, *Professional Roulette Prediction: Volume 1—Basic Methods* (RGE Publishing Ltd., 2006)
5. The Complete Roulette Guide of How to Play Roulette, http://www.fouroulette.com/roulette.htm
6. C.L. Webber, Jr, J.P. Zbilut, *Recurrence Quantification Analysis of Nonlinear Dynamical Systems* (Arlington, VA, National Science Foundation, 2005), pp. 26–94

7. HotBits: Genuine random numbers, generated by radioactive decay, https://www.fourmilab.ch/hotbits/
8. J.S. Eubank, A. Longtin, B. Galdrikian, J.D. Farmer, Testing for nonlinearity in time series: the method of surrogate data. Physica D **58**, 77–94 (1992)
9. Roulette data set consisting of 39,921 slot outcomes on American roulette wheels provided by Laurence Scott, professional roulette player and author
10. Old Faithful eruption time data from The Geyser Observation and Study Association, http://www.geyserstudy.org/geyser.aspx?pGeyserNo=OLDFAITHFUL
11. J.-P. Eckmann, S.O. Kamphorst, D. Ruelle, Recurrence plots of dynamical systems. Europhys. Lett. **4**, 973–977 (1987)
12. C.L. Webber, Jr, J.P. Zbilut, Dynamical assessment of physiological systems and states using recurrence plot strategies. J. Appl. Physiol. **76**, 965–973 (1994)
13. L.L. Trulla, A. Giuliani, J.P. Zbilut, C.L. Webber, Jr, Recurrence quantification analysis of the logistic equation with transients. Phys. Lett. A **223**, 255–260 (1996)
14. C.L. Webber, Jr, M.A. Schmidt, J.M. Walsh, Influence of isometric loading on biceps EMG dynamics as assessed by linear and nonlinear tools. J. Appl. Physiol. **78**, 814–822 (1995)
15. S.E. Brodie, Averaging Raindrops—an exercise in geometric probability (2010), http://www.cut-the-knot.org/Probability/RainDrops.shtml
16. Popular Roulette Betting Patterns, http://www.casinoadvisor.com/popular-roulette-betting-patterns-tip.html
17. L. Lacasa, B. Luque, F. Ballesteros, J. Luque, J.C. Nuño, From time series to complex networks: the visibility graph. Proc. Nat. Acad. Sci. **105**, 2972–2975 (2008)
18. C.L. Webber, Jr., N. Marwan (eds.), *Recurrence Quantification Analysis: Theory and Best Practices*, Springer series: Understanding Complex Systems. (Springer International Publishing, Cham Switzerland, 2015)
19. C.L. Webber, Jr., Recurrence Quantification of Fractal Structures, in *Fractal Analyses: Statistical and Methodological Innovations and Best Practices*, ed. by J. Holden, M. Riley, J. Gao, K. Torre. Frontiers in Fractal Physiology, vol. 3, Art. 382, 2012), pp. 1–11

Part II
Practical/Utilitarian Recurrences

Chapter 8
Assessment of Heart Rate Complexity Recovery from Maximal Exercise Using Recurrence Quantification Analysis

Rosangela Akemi Hoshi, Carlos Marcelo Pastre, Luiz Carlos Marques Vanderlei and Moacir Fernandes Godoy

Abstract Recurrence Plots have been described as suitable method for nonstationary and short data analysis, however, up to date, there is no registered study using this technique to evaluate exercise and recovery, which are conditions that provide unstable and variable heart rate time series. **Objectives**: To analyze heart rate complexity of healthy and young subjects during recovery from maximal exercise, using RQA measurements. **Methods**: Twenty participants were submitted to 120 min of passive recovery assessment after exercise at 100 % of maximal velocity until exhaustion. HR and R-R intervals (RRi) were recorded during the whole process and segments containing 300 RRi were selected from the baseline period and at each 10 min of recovery to be analyzed by the Kubios HRV Analysis software. **Results**: At baseline, the median values of Recurrence Rate, Determinism, Entropy and Lmean were 25.1 %, 97.2 %, 2.9 % and 8.6 %, respectively. After the exercise, all these variables showed significant increase until 80 min. Then, from 90 min and onwards, no significant difference was found between recovery moments and baseline. **Conclusions**: The regularity of heart rate rhythm increases expressively after maximal exercise and the complexity reduces. During recovery period, these conditions gradually return to rest levels and reach recovery after 80 min. As the variables showed pre-exercise levels at a similar moment, RQA seems to be a suitable method to evaluate exercise and recovery conditions.

R.A. Hoshi (✉) · M.F. Godoy
NUTECC-Transdisciplinary Nucleus of Studies on Complexity and Chaos, Sao Jose do Rio Preto Medicine School (FAMERP), 5416 Brigadeiro Faria Lima Ave., 15090-000 Sao Jose do Rio Preto, Sao Paulo, Brazil
e-mail: rosangela.hoshi@gmail.com

C.M. Pastre · L.C.M. Vanderlei
UNESP-State of Sao Paulo University, Campus of Presidente Prudente, Physical Therapy Department, 305 Roberto Simonsen Street, Presidente Prudente 19060-900, Brazil

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8_8

8.1 Introduction

Heart Rate Variability (HRV) is the measurement of periodic and non-periodic fluctuations between consecutive heartbeats employed as a tool for Autonomic Nervous System (ANS) assessment [1–5]. Since the 1960s, linear and nonlinear methods, based on mathematics, statistics and physics were developed and improved to provide reliable measurements [6]. Its application in clinical practice and research has been exploited by searching for its relationship with health and disease conditions, which helps to clarify their physiological meanings, yielding a more comprehensive description of indices and improving HRV analysis methods [4].

Linear methods based on time and frequency domains translate R-R intervals (RRi) oscillations into numeric values and are widely applied on research and clinical practice due to easy calculations and simple interpretation [1, 3, 7, 8]. According to the review article by Lombardi [9], time domain parameters should have a primary role as non-invasive stratifiers of patients with increased mortality, and frequency components are suitable indices of autonomic modulation of sinus node.

However, to ensure results reliability of linear indices, long and stationary time series are required, which may be impracticable in the real world, due to evaluation session time and heart rate natural oscillations [4, 10]. Sometimes, these characteristics may be considered a limitation and even meaningless, especially when the aim is to assess acute and unstable responses to certain conditions, such as during and after physical activities, when fast organic changes occur to provide the energy required by increased metabolism. Specifically regarding autonomic activity, there is a parasympathetic inhibition followed by sympathetic stimulation, which rise regularity and predictability of heart rate, reducing complexity. When the effort is interrupted, mechanical stimuli cease, metabolic stress reduces and autonomic functions are reset, then HR variations as well as complexity return to baseline levels [11, 12].

In these situations of rapid state changes, short time series are more appropriate, because organic functions are inherently time-varying processes that exhibit complex dynamics and state-dependent behaviors. Therefore, nonlinear analysis is suggested as suitable because it is less dependent of signal preprocessing, does not require stationarity and represents the complex nature of heart rate [13–15]. In this sense, Recurrence Quantification Analysis (RQA), based on Recurrence Plots (RP), is a nonlinear method described as ideal for physiological data. According to Webber et al. [16], RP does not impose rigid constraints on data set size, stationarity, or statistical distribution, thus it is suitable to assess conditions in which there are unstable behavior, fast oscillations and high levels of noise.

Several authors analyzed postexercise recovery using standard linear indices [12, 17–22], as well as nonlinear measures [23–26], concluding that the HRV is a reliable technique for ANS recovery assessment. However, as RQA application is a novel method in this field and so far there is no study, the objective of this research is to analyze heart rate complexity of healthy and young subjects during recovery

from maximal exercise, using RQA measurements, aiming to enhance the knowledge about this method and its physiological interpretations regarding heart rate time series.

8.2 Methods

Participants

Twenty healthy, young male university students (22.69 ± 2.98 years) were recruited and their anthropometric characteristics are presented in Table 8.1. All were classified as physically active, as they were regularly engaged in various intermittent activities (i.e., soccer, basketball, or handball three times per week), with each session lasting 1–2 h [12]. They were non-smokers and free of any known cardiovascular, metabolic or endocrine disease, as determined in interviews. None of the participants was taking any kind of medication. All participants signed a consent form after receiving verbal and written explanations about the procedures and the experimental protocol as approved by the Institutional Research Ethics Committee, protocol number 465/2011.

Study Design

All tests were performed in a quiet laboratory under standardized conditions (temperature: 21–23 °C; relative humidity: 40–60 %), between 2:00 and 6:00 pm. Each participant came to the laboratory two consecutive days and was asked to refrain from any exercise and not consume alcoholic or caffeinated beverages 24 h prior to each test. Additionally, participants were asked to have at least 2 h fasting prior to each test. On the first visit, participant height and weight were recorded, after which they underwent the Maximal Incremental Test in order to determine the peak of oxygen consumption (VO_{2peak}) and maximal aerobic speed (MAS).

Table 8.1 Anthropometric and cardiorespiratory parameters of participants (N = 20)

Variables	Mean ± SD
Age (year)	22 ± 1.2
Body mass (kg)	72.8 ± 8.4
Height (m)	1.76 ± 0.07
Body Mass Index (kg m^{-2})	23.49 ± 1.93
VO_{2max} (ml. kg^{-1} min^{-1})	55.85 ± 4.75
MAS (km h^{-1})	15.6 ± 0.88
HR_{max} (bpm)	180.7 ± 9.7

VO_{2max} maximum oxygen consumption; *MAS* maximal aerobic speed; *HR_{max}* maximal heart rate

On the second day, the participants first lay down for 30 min for the resting RRi recording, and then carried out constant exercise at the 100 % of MAS until volitional exhaustion, i.e., the Limit Time Test. Immediately after the exercise they kept a supine position for 120 min for the passive recovery period.

Maximal Incremental Test

After a standard warm-up, (which consisted in 5 min of 8 km h^{-1} jogging followed by 5 min in sitting position rest), VO_{2peak} and MAS were determined on a treadmill (Inbramed Super ATL, Inbrasport, Rio Grande do Sul, Brazil) using an incremental protocol. The slope was set at 1 % and the initial test speed was 10 km h^{-1}, increasing 1 km h^{-1} every minute until volitional exhaustion [27]. Breath-by-breath respiratory was measured and averaged every ten seconds by VO_{2000} Gas Analyzer (MedGraphics, Minnesota, USA). Expired oxygen (O_2) and carbon dioxide (CO_2) concentrations were determined using a galvanic fuel cell for the O_2 and a non-dispersive infrared analyzer for the CO_2. Flow was determined using a bi-directional differential pressure pneumotach (preVent™). Heart rate (HR) was recorded continuously with a Polar RS800CX monitor (Polar Electro, Kempele, Finland) in the RR mode [28].

At least two of the following criteria were required to ensure that VO_{2max} had been attained: (1) the occurrence of a plateau (i.e., <150 mL·min^{-1} increase between 2 successive stages); (2) a respiratory exchange ratio above 1.10; and/or (3) HR in excess of 90 % of age-predicted maximum. The MAS was recorded as the speed corresponding to the last complete stage on the incremental test [29].

Limit Time Test and Recovery

On the second day, after lying supine for 30 min for the recording of resting RRi, the participants performed the standard warm-up and then they ran until volitional exhaustion on a treadmill velocity set at 100 % of each individual's MAS. Additionally, strong verbal encouragement was provided. Exhaustion was determined by the same criteria as on the first day [12]. The sustained time average is shown in Table 8.1. Immediately following exercise, volunteers assumed a supine position for 120 continuous minutes for passive recovery.

Heart Rate Recordings and HRV Linear Analysis

The RRi were continuously recorded with Polar RS800 CX monitor (Polar Electro Oy, Kempele, Finland) in the RR mode at 1000 Hz of sampling rate. Time series were downloaded to the Polar Pro Trainer (v. 5, Polar Electro, Finland) software, which enables HR visualization. Only series with more than 95 % sinus rhythm were included in the study [30].

Suitable RRi were then selected for analysis and extracted into a *.txt* format to be analyzed. Then, a fixed number of 300 consecutive RRi were taken from the

baseline period, immediately after the end of the exercise, as well as right before each 10 min of the recovery period.

We analyzed linear indices SDNN (standard deviation of consecutive RR intervals) and RMSSD (root mean square successive difference between adjacent normal RR intervals), in time domain, and LF and HF (low and high frequency component of spectral analysis, respectively) in frequency domain [1].

Recurrence Quantification Analysis

The time series containing 300 RRi were analyzed by the Kubios HRV analysis software (v. 2.2, Kuopio, Finland), which considers vectors that represent the RRi time series as a trajectory in a m dimensional space ($m = 10$), with embedding lag ($\tau = 1$) and a fixed threshold 3.1623xSD.

We analyzed all the recurrence variables provided by this software: Recurrence Rate (%RR), defined as the percentage of recurring points, Determinism (%DET), the percentage of recurrence points forming diagonal lines, Shannon Entropy (ENT), which represents the complexity of line's distributions, Average Diagonal Line Length (L_{mean}) and Maximal Line Length (L_{max}) [23, 31, 32]. Since the last variable has a limited maximum value, we did not consider it for the statistical analysis because most of participants reached the peak, due to the intensity of the exercise performed.

Statistical Analysis

Kolmogorov-Smirnov test verified a nonadherence of variables to a Gaussian distribution. Therefore we used nonparametric Friedman's test, followed by Multiple Comparisons Dunn's test, to detect contrasts between observation moments and baseline data. StatsDirect software (v. 2.7.8b, StatsDirect Ltd., Cheshire, UK) was used for statistical analyses and the significance level was set at $P < 0.05$ for all tests.

8.3 Results

Figures 8.1, 8.2, 8.3 and 8.4 display the recovery course along time of %RR, % DET, ENT and Lmean, respectively, of Recurrence Quantification Analysis. We observe that in the first analyzed moment after exercise, all variables increased expressively in comparison to baseline. After 80 min and onwards, Friedman's test detected no differences between recovery moments and rest in none of these variables, suggesting recovery at the same moment for the four analyzed variables. Additionally, at 120 min we found values very close, but still higher, to pre-exercise levels.

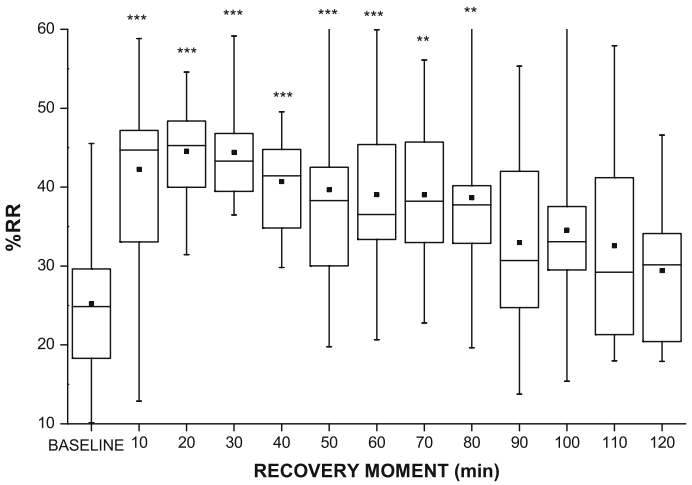

Fig. 8.1 Box Plots (Mean, Median, 1° and 3° quartiles, Minimum and Maximum) showing results of Recurrence Rate from RQA, according to the recovery time. **P < 0.01; ***P < 0.001

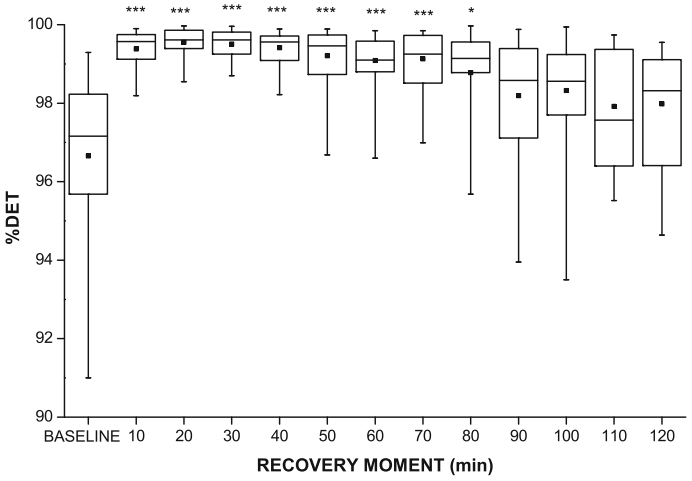

Fig. 8.2 Box Plots (Mean, Median, 1° and 3° quartiles, Minimum and Maximum) showing results of Determinism rate from RQA, according to the recovery time. *P < 0.05; ***P < 0.001

Table 8.2 shows linear indices variation during recovery time, compared to baseline values. We can observe that all variables presented an expressive decrease after exercise and SDNN as well as LF reach rest levels after 60 min, while RMSSD and HF, after 70 min.

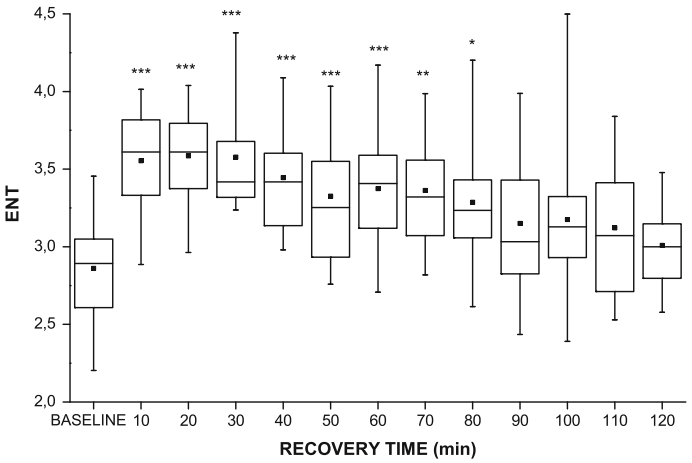

Fig. 8.3 Box Plots (Mean, Median, 1° and 3° quartiles, Minimum and Maximum) showing results of Shannon Entropy from RQA, according to the recovery time. *P < 0.05; **P < 0.01; ***P < 0.001

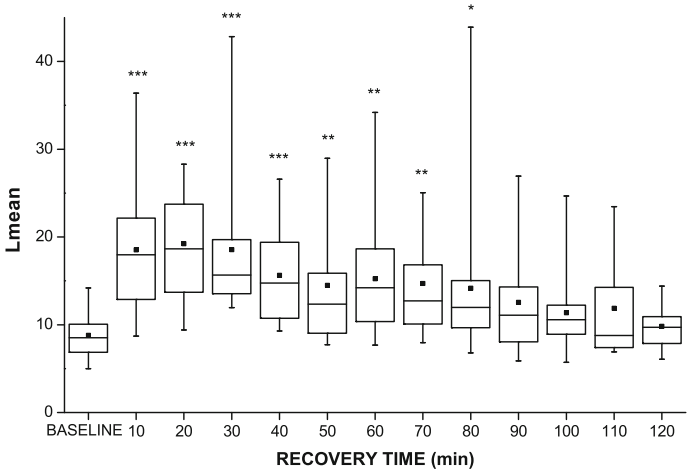

Fig. 8.4 Box Plots (Mean, Median, 1° and 3° quartiles, Minimum and Maximum) showing results of Mean Length of Diagonal Lines from RQA, according to the recovery time. *P < 0.05; **P < 0.01; ***P < 0.001

8.4 Discussion

In this study, we applied high-intensity exercise as a stressful element to the ANS to follow how RQA variables would behavior. Thus, we analyzed the moment immediately after the maximal effort, which presented high levels of recurrence

Table 8.2 Median (min-max) values of linear indices, according to the recovery time

LI	Baseline	Recovery time (minutes)											
		10	20	30	40	50	60	70	80	90	100	110	120
SDNN (ms)	60.5 (27–117)	7.5*** (5–20)	12*** (7–39)	17*** (8–63)	23*** (14–57)	26*** (11–60)	29** (12–80)	36 (22–89)	47.5 (22–80)	46.5 (20–79)	55.5 (17–80)	51 (21–93)	59.5 (26–113)
RMSSD (ms)	45.4 (19.4–97)	3.9*** (1.5–0.3)	4.9*** (2–31)	7.8*** (2.7–54)	10.1*** (6–40)	13.2*** (5.5–37.7)	17*** (4.9–71.2)	17.4*** (7–70.3)	30.3 (8.4–71.9)	32.9 (8.5–56)	37.9 (10.3–59.7)	39.1 (16.3–78.5)	41 (11.6–73.3)
LF (ms^2)	255 (24–058)	1*** (0–16)	4.5*** (0–33)	7*** (1–291)	15*** (5–111)	16.5*** (7–222)	33.5** (2–459)	59 (14–327)	94 (8–464)	129 (12–628)	227 (11–456)	206 (33–682)	204 (34–1067)
HF (ms^2)	128.5 (13–867)	0*** (0–5)	0.5*** (0–30)	2.5*** (0–185)	3.5*** (0–89)	6.5*** (0–109)	13.5*** (1–224)	9** (1–261)	44.5 (1–478)	40.5 (2–482)	77 (3–342)	75.5 (14–604)	120.5 (4–399)

LI Linear indices; *SDNN* standard deviation of consecutive RR intervals; *RMSSD* root mean square successive difference between adjacent normal RR intervals; *LF* low frequency component of spectral analysis; *HF* high frequency component of spectral analysis

**P < 0.01; ***P < 0.001

compared to baseline, due to parasympathetic inhibition and sympathetic stimulation during physical efforts. It is described that vagal modulation is the major effector over cardiac rhythm complexity, thus, this reciprocity that reduces RRi and increases regularity plays an important role decreasing complexity levels [33–35]. Sympathetic activity controls epinephrine and norepinephrine release into the blood stream, which will reach the heart, increasing HR and contraction strength to supply metabolic demand.

The following 120 min after exercise were also analyzed and we observed slow and gradual restoration of linear indices and RQA variables towards rest values. When the effort is interrupted, an efferent vagus nerve reactivation occurs followed by sympathetic activity inhibition [11, 12]. At the intensity applied in this study, we detected an apparent sequence of events: first, SDNN and LF, which reflect global variability, reached recovery levels; then RMSSD and HF, reflecting parasympathetic modulation and, lastly, RQA variables. This result converges to the fact that parasympathetic modulation is considered the major responsible for HR complexity [33, 35], since the periodic dynamics manifestations, as well as the interactions among them, the distribution of diagonal line lengths and the mean duration of a stable interaction, represented by %RR, %DET, ENT and Lmean [16, 36], respectively, showed recovery after autonomic branches reorganization, especially parasympathetic system.

RQA variables are mathematically calculated but they have physiological meanings, so that the results suggest some coherence, since they reached baseline levels at a similar moment, between 80 and 90 min. This parallelism of time courses was hypothesized since the studied variables are interdependent and result from the same features (recurrence points). In this sense, looking at the subjects individually, we detected that at 90 min, those participants who have reached pre-exercise levels in one of the four recurrence variables, also showed recovery on other two or three measurements.

Other studies, concerned about physiological recovery processes, mentioned at least one hour after intensive exercise as the time required for homeostasis restoration, including plasma catecholamines concentration removal [37], blood lactate clearance [38] and fast/slow components of oxygen uptake recovery [39]. These organic events contribute for autonomic activity normalization due to cessation of mechanical and metabolic stimuli [40].

Thus, our findings support the hypothesis that RQA is a consistent method for heart rate dynamics assessment after exercise. Although recurrence analysis is a nonlinear graphical method designed for nonstationary short time series analysis, we have not found other studies that applied this method to exercise, as mentioned before. Therefore, as "nonlinear techniques may offer advantages over linear techniques in identifying and quantifying the modulation of interactions among neurocardiac control mechanisms" [41], it is very worth the application of these measures over varied conditions.

Due to the importance of postexercise restoration processes, cardiac issues regarding performing physical exercises do not limit to the effort phase, but extend to recovery period, when the system is still unstable and vulnerable [42]. Therefore,

it is worthy to include RQA in the evaluation of physical conditions, likewise in performance tests or routine ergometric exams.

There is no doubt about nonlinear relevance for HRV analyses, but its application is not widespread due to several aspects that are still unclear, like results interpretation and practical relevance. According to the review article of Mansier et al. [43], "numerous methods of analysis, derived from classical signal processing or nonlinear dynamics, are available and no single measure is more appropriate than the others for physiological research or clinical practice" and other studies [44–47] describe that linear and nonlinear analysis are complementary in the evaluation processes, providing additional and non-redundant information. Therefore we highlight the importance of enhancing the knowledge about HRV, because it helps to understand various physiological aspects, such as cardiovascular function in health and disease.

As a limitation of this study we can point out the selected sample, because although all the participants had been classified as physically active, based on a questionnaire, and they reached similar levels of aerobic capacity, it was a marked variation in recurrence variables and linear indices. This may be due to e fact that it was not possible to ensure the activities specificity and, since the ANS is directly influenced by the frequency, load, and type of exercise [21, 48], it is reasonable to expect some variability among the participants. Moreover, fitness is not the only and not even the primary effector over autonomic modulation of cardiovascular system, but rather it is also the result of vasomotor and breathing centers activity, vascular autoregulation, and baroreflex and chemoreflex regulation [49, 50].

8.5 Conclusions

Based on the analysis performed in this study, we conclude that the regularity of heart rate rhythm increases expressively after maximal exercise and the complexity reduces. During recovery period, these conditions gradually return to rest levels and reach recovery after 80 min, in young healthy and trained men. As the variables showed pre-exercise levels at a similar moment and results are compatible to physiology, RQA seems to be a suitable method to evaluate exercise and recovery conditions.

References

1. Task Force of ESC and NASPE, Circulation **93**, 1043–1065 (1996)
2. A.E. Aubert, B. Seps, F. Beckers, Sport. Med. **33**, 889–919 (2003)
3. L.C.M. Vanderlei, C.M. Pastre, R.A. Hoshi, T.D. Carvalho, M.F. Godoy, Rev. Bras. Cir. Cardiovasc. **24**, 205–217 (2009)
4. G.E. Billman, Front. Physiol. **2**, 86 (2011). doi:10.3389/fphys.2011.00086
5. G.E. Billman, Front. Physiol. **4**, 222 (2013). doi:10.3389/fphys.2013.00222

6. F. Lombardi, Circulation **101**, 8–10 (2000)
7. L.J. Van Schelven, P.L. Oey, I.H. Klein, M.G. Barnas, P.J. Blankestijn, G.H. Wieneke, J. Auton. Nerv. Syst. **79**, 144–148 (2000)
8. D.S. Goldstein, O. Bentho, M.-Y. Park, Y. Sharabi, Exp. Physiol. **96**, 1255–1261 (2011). doi:10.1113/expphysiol.2010.056259
9. F. Lombardi, Card. Electrophysiol. Rev. **6**, 245–249 (2002)
10. A. Porta, T. Gnecchi-Ruscone, E. Tobaldini, S. Guzzetti, R. Furlan, N. Montano, J. Appl. Physiol. **103**, 1143–1149 (2007)
11. P. Kaikkonen, A. Nummela, H. Rusko, Eur. J. Appl. Physiol. **102**, 79–86 (2007)
12. F.N. Bastos, L.C.M. Vanderlei, F.Y. Nakamura, Int. J. Sports Med. **33**, 873–879 (2012)
13. C. Peng, S. Havlin, H. Stanley, A. Goldberguer, Chaos **5**, 82–87 (1995)
14. R.A. Hoshi, C.M. Pastre, L.C.M. Vanderlei, M.F. Godoy, Auton. Neurosci. **177**, 271–274 (2013). doi:10.1016/j.autneu.2013.05.004
15. M. Thiel, M.C. Romano, J. Kurths, Phys. Lett. A **330**, 343–349 (2004). doi:10.1016/j.physleta.2004.07.050
16. C.L. Webber, J.P. Zbilut, J. Appl. Physiol. **76**, 965–973 (1994)
17. L.R. Davrath, S. Akselrod, I. Pinhas, E. Toledo, A. Beck, D. Elian, M. Scheinowitz, Med. Sci. Sports Exerc. **38**, 2095–2101 (2006)
18. K. Martinmaki, K. Hakkinen, J. Mikkola, H. Rusko, Eur. J. Appl. Physiol. **104**, 541–548 (2008)
19. O.F. Barak, D.G. Jakovljevic, J.Z. Popadic Gacesa, Z.B. Ovcin, D.A. Brodie, N.G. Grujic, J. Sports Sci. Med. **9**, 176–182 (2010)
20. W. Niewiadomski, A. Gasiorowska, B. Krauss, A. Mróz, G. Cybulski, Clin. Physiol. Funct. Imaging **27**, 309–319 (2007)
21. S. Seiler, O. Haugen, E. Kuffel, Med. Sci. Sports Exerc. **39**, 1366–1373 (2007). doi:10.1249/mss.0b013e318060f17d
22. P. Terziotti, F. Schena, G. Gulli, A. Cevese, Eur. J. Appl. Physiol. **84**, 187–194 (2001)
23. M. Javorka, I. Zila, T. Balhárek, K. Javorka, Brazilian J. Med. Biol. Res. **35**, 991–1000 (2002)
24. J. Casties, D. Mottet, D. Le Gallais, Int. J. Sports Med. **27**, 780–785 (2006). doi:10.1055/s-2005-872968
25. R. Perini, A. Veicsteinas, Eur. J. Appl. Physiol. **90**, 317–325 (2003). doi:10.1007/s00421-003-0953-9
26. I. Hagerman, M. Berglund, M. Lorin, J. Nowak, C. Sylvén, Cardiovasc. Res. **31**, 410–418 (1996)
27. S.A. Harling, R.J. Tong, T.D. Mickleborough, Med. Sci. Sports Exerc. **35**, 663–668 (2003). doi:10.1249/01.MSS.0000058434.53664.EC
28. L.C.M. Vanderlei, R.A. Silva, C.M. Pastre, F.M. Azevedo, M.F. Godoy, Braz. J. Med. Biol. Res. **41**, 854–859 (2008)
29. T.D. Noakes, K. Myburgh, R. Schall, J. Sports Sci. **8**, 35–45 (1990)
30. M.F. Godoy, I.T. Takakura, P.R. Correa, Arq. Ciência da Saúde **12**, 167–171 (2005)
31. C.L. Webber, Front. Physiol. **1**, 382 (2012). doi:10.3389/fphys.2012.00382
32. N. Marwan, M. Carmen Romano, M. Thiel, J. Kurths, Phys. Rep. **438**, 237–329 (2007), doi:10.1016/j.physrep.2006.11.001
33. L. Mourot, M. Bouhaddi, E. Gandelin, Auton. Neurosci. **137**, 27–36 (2007). doi:10.1016/j.autneu.2007.06.284
34. J. Penttilä, A. Helminen, T. Jartti, T. Kuusela, H.V. Huikuri, M. Tulppo, H. Scheinin, Auton. Autacoid Pharmacol. **23**, 173–179 (2003)
35. K.S. Heffernan, C.A. Fahs, K.K. Shinsako, S.Y. Jae, B. Fernhall, Am. J. Physiol. Heart Circ. Physiol. **293**, H3180–H3186 (2007), doi:10.1152/ajpheart.00648.2007
36. J.E. Naschitz, I. Rosner, N. Shaviv, J. Hum. Hypertens. **17**, 111–118 (2003). doi:10.1038/sj.jhh.1001517
37. T.J. Horton, M.J. Pagliassotti, K. Hobbs, J.O. Hill, J. Appl. Physiol. **85**, 1823–1832 (1998)
38. F.B. Neric, W.C. Beam, L.E. Brown, L.D. Wiersma, J. Strength Cond. Res. **23**, 2560–2567 (2009)

39. M. Burnley, J.H. Doust, A.M. Jones, J. Appl. Physiol. **101**, 1320–1327 (2006)
40. J.V. Freeman, F.E. Dewey, D.M. Hadley, J. Myers, V.F. Froelicher, Prog. Cardiovasc. Dis. **48**, 342–362 (2006)
41. A. Schumacher, Biol. Res. Nurs. **5**, 211–221 (2004)
42. M. Tulppo, A.M. Kiviniemi, A.J. Hautala, M. Kallio, T. Seppänen, S. Tiinanen, T.H. Mäkikallio, H.V. Huikuri, Clin. Physiol. Funct. Imaging **31**, 272–281 (2011). doi:10.1111/j.1475-097X.2011.01012.x
43. P. Mansier, J. Clairambault, N. Charlotte, Cardiovasc. Res. **31**, 371–379 (1996)
44. M.J. Lewis, A.L. Short, J. Suckling, Comput. Methods Programs Biomed. **108**, 176–185 (2012)
45. G. Krstacic, A. Krstacic, A. Smalcelj, D. Milicic, M. Jembrek-Gostovic, Ann. Noninvasive Electrocardiol. **12**, 130–136 (2007)
46. M.M. Platisa, V. Gal, Physiol. Meas. **27**, 145–154 (2006). doi:10.1088/0967-3334/27/2/005
47. J.F. Valencia, I. Rivero, A. Voss, A.B. De, A. Porta, P. Caminal, J.F. Valencia, Phylos. Trans. A **373**, 20140092 (2015). doi:10.1098/rsta.2014.0092
48. K. Martinmäki, H. Rusko, Eur. J. Appl. Physiol. **102**, 353–360 (2008). doi:10.1007/s00421-007-0594-5
49. A. Voss, R. Schroeder, A. Heitmann, A. Peters, S. Perz, PLoS One **10**, e0118308 (2015). doi:10.1371/journal.pone.0118308
50. A. Porta, S. Guzzetti, N. Montano, R. Furlan, M. Pagani, A. Malliani, S. Cerutti, IEEE Trans. Biomed. Eng. **48**, 1282–1291 (2001)

Chapter 9
Recurrence Analysis of Cardiac Restitution in Human Ventricle

Humberto Arce, Ayari Fuentes and G. Hortensia González

Abstract The cardiac restitution curve describes functional relationships between diastolic intervals and their corresponding action potential durations. Although the simplest relationship is that restitution curves are monotonic, empirical studies have suggested that cardiac patients present a more complex dynamical process characterized, for instance, by a non-monotonic restitution curve. The purpose of this chapter is to analyze the dynamical properties of a non-monotonic cardiac restitution curve model derived from previously published clinical data. To achieve this goal, we use Recurrence Quantitative Analysis combined with Lyapunov exponents and Supertrack Functions in order to describe the complex dynamics underlying non-monotonic restitution curves. We conclude by highlighting that a consequence of the advanced complex dynamics that emerges from the aforementioned non-monotonicity, is the increasing risk of alternant rhythms.

9.1 Introduction

Cardiovascular accidents are a major cause of death worldwide. Among the diverse risk indicators, a significant one is the appearance of T-wave amplitude alternans on electrocardiograms consequence of a beat-to-beat oscillation on the intracellular *action potential duration* (APD) [1]. APD alternans have been studied using the *electrical restitution curve* or *cardiac restitution curve* (CRC) since the 70's of the

H. Arce (✉) · G.H. González
Laboratorio de Biofísica de Sistemas Excitables, Facultad de Ciencias, Universidad Nacional Autónoma de México, Mexico, Mexico
e-mail: harce@ciencias.unam.mx

G.H. González
e-mail: hortecgg@ciencias.unam.mx

A. Fuentes
Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Mexico, Mexico
e-mail: ayarifh@ccg.unam.mx

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8_9

last century, an approach based on studying the interaction between the action potential duration and the *diastolic interval* (DI) [2, 3]. The simplest interpretation is that the APD depends exclusively on the previous DI because cardiac cells only have a cycle length time to recover its ionic kinetics [4, 5]. However, theoretical and experimental results have suggested a memory effect that depends not just on the immediate previous cycle [6, 7].

The relevance of this approach is partly a consequence of results showing that the slope of the CRC can be modulated by pharmacological interventions [8–11] or by changing the propensity for ventricular fibrillation in cardiac preparations through pacing protocols [12–16]. In general, modeling studies consider the CRC as a monotonically-growing exponential curve [8, 13, 17–20] and, as a consequence, if the slope of the APD adaptation is higher than one, a reduction on the stimulation interval eventually leads to a period-doubling bifurcation [2, 21]. Indeed, a CRC slope >1 amplifies APD alternans and therefore it can lead to fibrillatory rhythms. For this reason, electrical restitution curve analysis has become an important tool for cardiac risk prediction [4, 6, 10, 13]. However, since 1975 multiple empirical studies have reported the occurrence of non-monotonic curves in different cardiac assays [11, 22, 23]. In particular, for the human ventricle, Franz and colleagues found in 1985 the existence of a shoulder or local maximum in the CRC [4, 23]. These results were exhaustively verified in 2005 by Yue et al. [24] using a S1–S2 stimulation protocol. Furthermore, their array of non-contact electrodes revealed CRC with diverse profiles in contiguous cardiac cells of the same patient; from purely monotonic to a non-monotonic profile with a very pronounced local maximum, a feature illustrated in Fig. 9.1a.

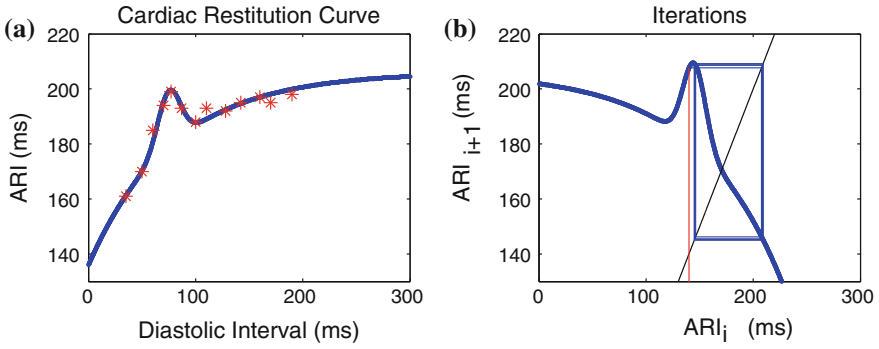

Fig. 9.1 Cardiac Restitution Curve with Local Maximum (CRC-LM). **a** the *red asterisks* correspond to 13 experimental points taken from Yue et al. [24] obtained using a S1-S2 stimulation protocol; the *blue* trace is the function that we used to fit the data based on the addition of an exponential and a Gaussian functions. Activation Recovery Interval (ARI) is an indirect measurement of the intracellular APD. DI denotes the Diastolic Interval and represents the resting time between activations. **b** iteration map used to obtain the time series (see methods for an extended explanation)

In this work, we show that the presence of a local maximum in a human data-driven CRC model can considerably advance the alternant behavior and, as a result, the *bifurcation diagrams* (BD) for these CRC are highly complex. Our analysis is based on a methodological extension of recurrence plots analysis known as *recurrence quantitative analysis* (RQA) based on the following indexes: *recurrence times of the second type* (T2), *recurrence rate* (RR), the *maximum diagonal length* (Lmax) and the *maximum vertical length* (Vmax). We then compare the predictions obtained using RQA with the results obtained using Lyapunov exponents. Using both approaches we demonstrate how an increment in the local maximum height has the effect of increasing the complex region size in the BD, as well as the magnitude of the alternans. Finally, we describe the different transitions observed.

9.2 Methods

9.2.1 Generation of Time-Series

In this chapter we propose a model based on data obtained by Yue et al. [24] who used an experimental protocol that introduces an extracellular electrode array into the right or left ventricle of a cardiac patient. This protocol allows the authors to measure a time interval equivalent to the intracellular APD known as the *activation recovery interval* (ARI). ARI was measured between the time of $(dV/dt)_{min}$ of the QRS and the $(dV/dt)_{max}$ of the T-wave on the unipolar electrograms of the ventricular endocardium employing a S1–S2 stimulation protocol based on a constant stimulation period maintained for 2 min, with an extra stimulus intercalated at a different time point [24].

For the purpose of this chapter we will define the CRC with local maximum as CRC-LM and estimate the experimental points in Fig. 9.2 panel C of Yue et al. [24]. We discarded the two shorter ones because they overlap with a third experimental point. Then, in order to fit the 13 remaining points, we added a Gaussian function to an exponential curve matching the local maximum position with the corresponding experimental maximum, as illustrated in Fig. 9.1a. We then obtain the following semi-quantitative function

$$\mathrm{ARI} = 206 - 70\exp(-\mathrm{DI}/78) + 20\exp(-((-75+\mathrm{DI})2/220)) \tag{9.1}$$

where the Gaussian coefficient (third term) is representative of its height, *h,* which in (9.1) corresponds to 20. Note that when this parameter equals zero the CRC is monotonic, and therefore here we are interested in changing its value to evaluate the CRC-LM deviation from the well-characterized behavior of monotonic curves.

We then consider a series of stimuli applied with a given period denoted by *P*. The effect of each stimulus will be to generate a voltage deflection with a given

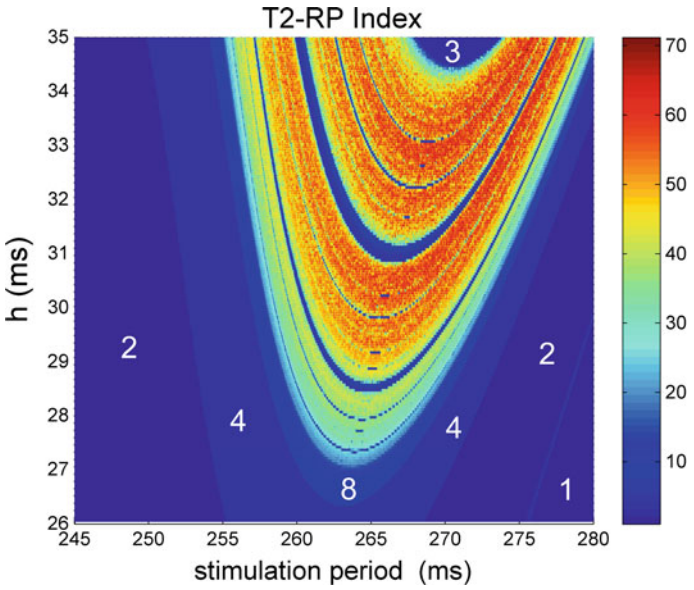

Fig. 9.2 Biparametric BD using T2. T2 was obtained from RP as index. N-Rhythms are longer as h is increased. *White numbers* indicate the basic rhythm observed in the corresponding band

ARI_i followed by a diastolic interval, DI_i and therefore we obtain the following constant value of P for each series of stimulus:

$$P = ARI_i + DI_i. \tag{9.2}$$

Now, (9.1) states that each ARI depends on the previous DI so, we can define a discrete map, denoted as F, such that

$$ARI_{i+1} = F(DI_i) \tag{9.3}$$

and thus from (9.2) we obtain

$$ARI_{i+1} = F(P - ARI_i). \tag{9.4}$$

So, in order to produce a time-series for a given pair P and h, we only need to propose an initial ARI and perform the necessary iterations. Figure 9.1 panel B shows the results obtained with $h = 30$, $P = 220$ and with an initial ARI of 140 ms. Note that after a short transient interval, a 2-period rhythm is achieved, that is, ARI values repeat every two other pulses. Indeed, N-rhythm is defined as repeated pattern every N elements, a pattern that depends on P. Longer periods move F to the right-hand side and the shorter ones to the left, displacing the intersection of F with the identity line. Then we have that each time-series depends on h for a given CRC-LM and P.

9.2.2 Dynamic Bifurcation Diagrams

We used three approaches to study the CRC-LM curves. Initially, we applied a method proposed by Trulla et al. [25] to obtain the BD and RQA indexes very quickly. This method is based on incrementing in small steps the bifurcation parameter before every new iteration of F. First, we obtained the BD for h with values ranging from 0 to 35 ms with steps every 0.05 ms. For each h, a modified ARI time-series was generated starting from $P = 100$ ms with increments of 0.001–300 ms. Once identified the region of interest inside the $P \times h$ space, we calculated the T2-RQA index and the maximum Lyapunov exponent in the region defined by h from 26 to 35 ms with 0.05 ms steps and P between 245 and 280 ms with 0.0001 ms steps. The time-series consists of 350,000 points that were sub-divided in 500-point blocks in order to calculate T2 and the higher Lyapunov exponent. Recurrence times of the second type (T2) were proposed by Gao [26] and correspond to the recurrence time remaining after the sojourn times in a specified epsilon are discarded. Here we obtained T2 using CRP-TOOLS (http://tocsy.pik-potsdam.de/CRPtoolbox/).

To calculate Lyapunov exponents, we used the logarithm of the derivatives product of F in each element of the 500 points orbit [27], moving along the time-series with 10-point steps. Based on previously published results [25, 26], we defined D = 2 and t = 1, as well as used the Euclidean norm and epsilon = 2^{-5}. Then, to compare our results with the standard stationary method, we obtained time-series for each P and h with 2,000 values in each case, but calculated T2 and the Lyapunov exponent based on the last 500 values of the time-series.

In the second part of this study, we followed the evolution of the system as h changes using the stationary procedure to estimate the *recurrence rate* (RR), the *maximum diagonal length* (Lmax) and the *maximum vertical length* (Vmax) from the time-series. These indexes are part of the RQA proposed as an extension of the Recurrence Plots [28, 29] where RR quantifies the percentage of recurrent points falling within the specified epsilon, Lmax is the length of the longest diagonal line segment in the plot (excluding the main diagonal line of the identity) and Vmax measures the length of the longest vertical line in the RP. This index was proposed later [30, 31] and reveals information about the time duration of the laminar states, allowing the investigation of intermittency possible.

Finally, in order to study the evolution of laminar zones, we obtained the *normalized recurrence frequency* (NRF) with a maximum that coincides with the *super track functions* (STF) for the CRC-LM under consideration. NRF is a histogram obtained for each P that measures how a time-series recurs to some particular value. The extreme ARI values in all the BD were identified; these intervals were divided in 0.001 ms bins to count the ARI values and normalized to the highest frequency found. The obtained values allowed us to analyze the most visited frequency bands in the BD. STF were obtained following Oblow [32]. In the map defined by F the invariant point corresponds to the maximum point, so we considered this value as the first s_1 function that results constant. The following s_n are then the n-th iteration value obtained from the map for each BD parameter, in our case P.

9.3 Results

9.3.1 Dynamic Method

The most relevant result we found is that, for a clinically and physiologically relevant context, the presence of a local maximum markedly advances the initiation of alternans. Furthermore, as alternans are cardiac risk predictors, we conclude that CRC-LM increases the risk for arrhythmia. We studied local maxima from 0 to 35 ms in height, with increments of 0.05 ms and, for each h studied, P was taken as a bifurcation parameter starting in 100 ms and going to 300 ms with 0.001 ms steps. When $h = 0$ the CRC corresponds to a monotonic curve and alternans appear when P is reduced to 121.3 ms. Moreover, increasing h leads to a new alternans zone. When $h = 10.5$ ms we see that the system is in a 2-rhythm alternant pattern between 244 and 246 ms. A larger value of h has the effect of increasing this region and therefore its behavior is more complex. For $h = 11$, the alternans region spans from 243 to 250 ms, while for $h = 35$, from 223.7 to 285.9 ms. To carefully analyze this alternans region, we defined the h interval between 26 and 35 ms and P from 245 to 280 ms and calculated T2 as well as the higher Lyapunov exponents using Trulla et al.'s method [25] with 0.0001 ms variations in the bifurcation parameter.

Figure 9.2 presents the results obtained for T2 index. The period P (in ms) is on the horizontal axis, and the height h (also in ms) on the vertical axis. T2 is indicated by the color bar at right. The general shape of the plot is of colored bands similar to upwards parabolas barely shifted to the right. As h grows, the rhythms (denoted with white numbers) move simultaneously inward from the left and right border, and new N-rhythms emerge. This pattern is interrupted by shorter periodicity bands (shown in blue).

Figure 9.3 shows the results obtained for the Lyapunov exponents. For clarity purposes, exponents <–0.2 have been truncated. Interestingly, this Figure has the same bands pattern as those shown in Fig. 9.2. The frontiers between 2, 4 and 8-rhythms are highlighted because Lyapunov exponents in these zones are very close to zero. We can see that zones with positive exponents (corresponding to chaotic regions) are related to zones with T2 higher than 20 in Fig. 9.2. In this Figure, we sketched a couple of white dotted horizontal lines; the first one in h ~ 26.9 ms is the highest limit to the periodic behaviors. It is important to notice that in the region between the two lines, chaotic behavior without laminar states can be present, while laminar states exist only above the higher line.

9.3.2 The Stationarity Method

For each time-series obtained with this procedure, the last 500 points were taken into account in order to obtain the RQA indexes (RR, Lmax and Vmax). Figure 9.4 illustrates the behavior of the system in the period-doubling bifurcation zone, from

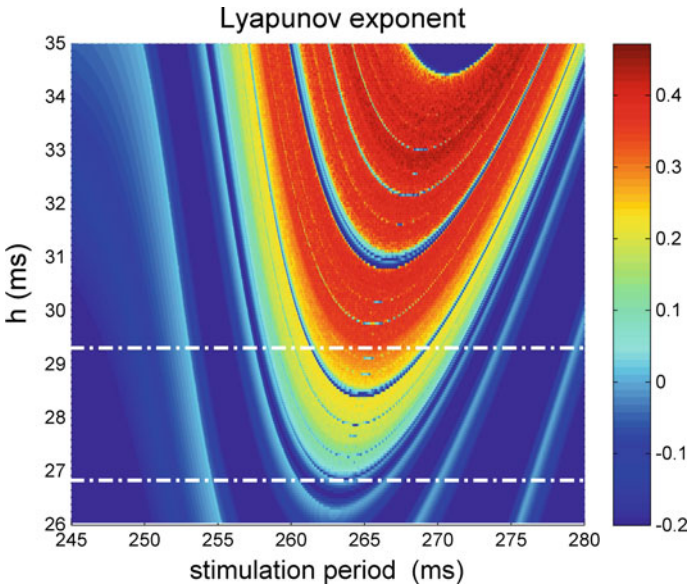

Fig. 9.3 Biparametric BD for Lyapunov exponent. Periodic states are found exclusively under the *bottom white line*, while chaos without laminar states can be found between both *dotted white lines*. Above the *top dotted white line* we can find chaos with laminar states. To recognize them we used the Lmax and Vmax evolution, explained afterwards in the text and shown in Figs. 9.4, 9.5 and 9.6

$h = 10.5$ up to approximately $h = 26.95$. For $h = 10.5$, the alternans zones is only of 2 ms. Interestingly, alternans magnitude calculated as the difference between highest and shortest ARI is 1.4 ms. In Fig. 9.4a shows the BD for $h = 20$ ms. Although alternans are still between 2-period rhythms, their magnitude has grown to 37.5 ms (and therefore arrhythmia risk is increased) and the width of the region where they exist is of 40 ms. We can see that alternans magnitude reduces as P grows.

Figure 9.4b displays the typical bubble shape of the BD associated with this system. Increments in P produce period-doubling bifurcations in a sequence 1, 2, 4, …$2n$ where n depends on h. As P increases, the inverse sequence $2n$, …4, 2, 1 develops. Figures 9.4c and 9.4d show RR (Fig. 9.4c), Lmax and Vmax (Fig. 9.4d) indexes. The RR inverse is the rhythm value of the corresponding orbits. An Lmax change in short steps is signaling transitions between periodic rhythms.

Figure 9.5 shows the behavior in the h region where chaos without laminar states can exist, an interval that spans from $h = 26.95$ to $h = 29.1$ ms. Figure 9.5a shows the BD for $h = 26.95$ ms, while Fig. 9.5b shows the evolution of Lmax and Vmax with P. It can be noted that there is a marked reduction in Lmax which is related to the appearance of chaotic activity, in this case the region has 2 ms in size. In Fig. 9.5c it is shown the BD for $h = 29.1$ ms and from Fig. 9.5d we can infer that

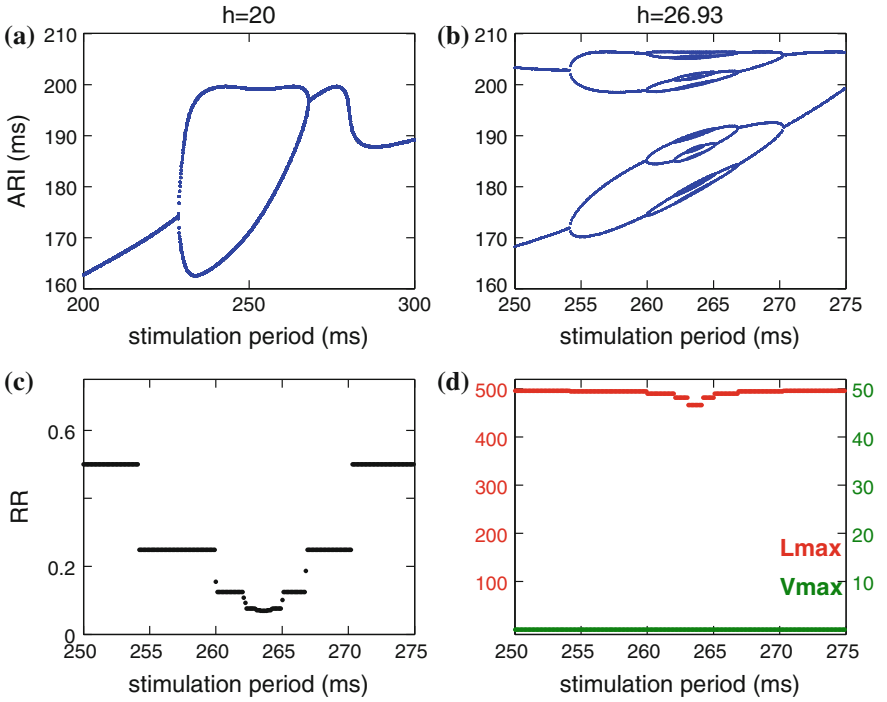

Fig. 9.4 Periodic Rhythms. In **a**, two consecutive period-doubling bifurcations are shown, occurring at h = 20 ms. **b** shows two bifurcation cascades from 2 to 32-rhythms and vice versa. **c** shows that recurrence density (RR) reflects the BD rhythms. **d** shows that Vmax = zero because there is not laminarity and Lmax has only small changes corresponding to high order periodic rhythms

chaotic region size increased more than 10 ms. In this case, in the interior of the chaotic region there are clearly distinguishable periodic windows. As Vmax is not different than zero (during the entire *h* interval) we know that there are no laminar states. The *h* interval for chaotic behavior and laminar states is from $h = 29.1$ to $h = 35$ ms, the latter being the highest *h* we considered in this study. Figure 9.6 shows the behavior found in this region.

Figure 9.6a displays the BD with the two chaotic bands that in Fig. 9.5c were unjointed and now partially merged, while Fig. 9.6b shows Lmax and Vmax for *P* values ranging from 262.2 to 269.3 ms, a region where there are windows with Vmax different from zero and therefore laminar states occur in these windows. To illustrate these states, we produced the corresponding RP for $P = 262.2$ ms that leads to chaotic behavior (only the last 200 points in the time-series were considered in this plot). It is important to mention that here, as in all the cases previously presented, we consider $D = 2$ and $t = 1$. Now, we can see that the observed pattern

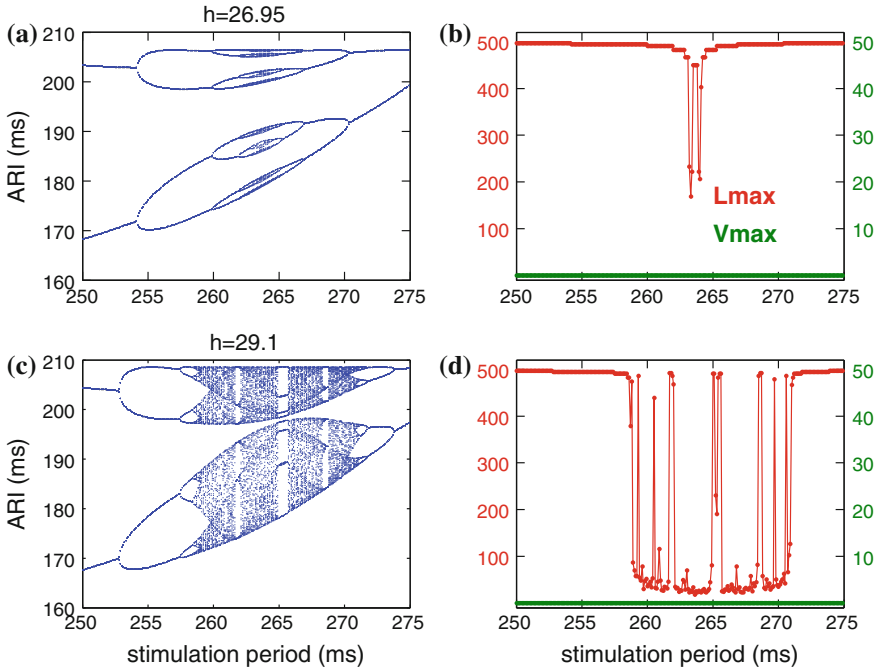

Fig. 9.5 Chaos without laminar states. **a** displays the BD for h = 26.95 ms. **b** shows that for this h there are already chaotic states since Lmax has sudden drops in respect to its value under periodic regimes. **c** shows BD for h = 29.1 ms and **d** indicates that the main effect due to h increase is the expansion of the region presenting chaos, revealed as Lmax index has more frequent drops

is characterized by small diagonals whose length is related with the exponential divergence of their phase space trajectories [31]. Figure 9.6d displays the RP for the last 200 points in the time-series generated when $P = 266.3$ ms. Note how small black blocks can be observed, corresponding to consecutive ARI with similar values.

To follow in detail the evolution of the region that presents laminarity, we calculated the NRF (see methods) for the case h = 30 ms. We see that the highest ARI value is slightly smaller than 210 ms and the lowest is slightly larger than 165 ms. We divided the interval in 45,000 equidistant windows and, for each P, counted the amount of elements in every window.

We also produced a histogram for the normalized time-series (with frequency values ranging from 0 to 1). Figure 9.7a shows the ARI when NRF = 1. This plot shows the surrounding lines that contain the two merged chaotic zones. In Fig. 9.7b, the first four STF are superposed to the plot in Fig. 9.7a. We see that the STF and the surrounding lines are in exact correspondence.

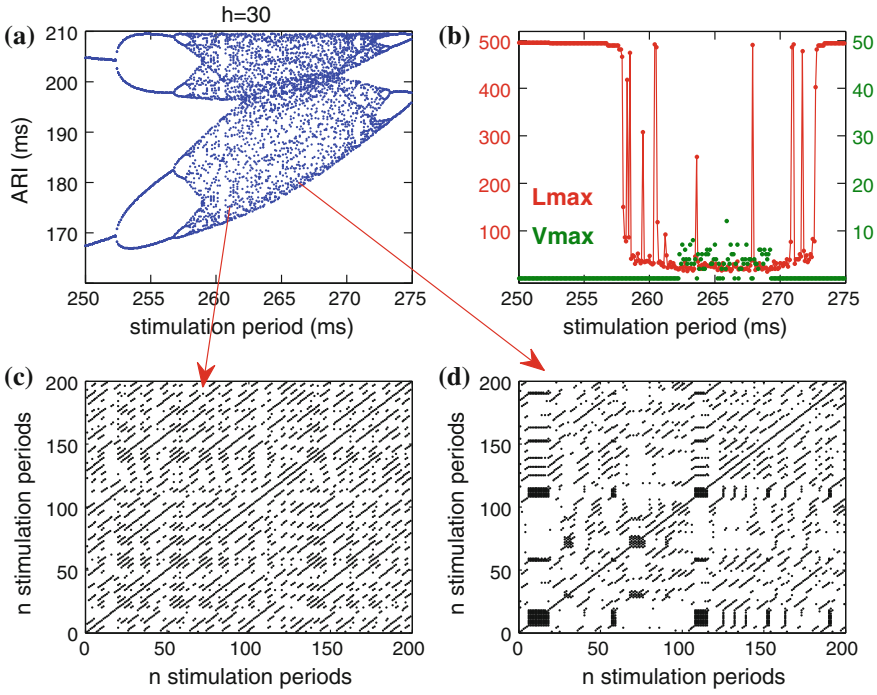

Fig. 9.6 Chaos with laminar states. **a** displays the BD for h = 30 ms. Note how in the middle of the plot, two chaotic zones are superposed. **b** shows that this superposition leads to a Vmax different from zero for some values of P. **c** and **d** show the RP built with 200 points, D = 2, delay = 1, epsilon = 0.2. **c** corresponds to P = 262.2 ms where there is chaotic behavior. **d** shows that, for P = 266.3 ms, there are laminar states. Recurrence is calculated after each stimulus

Figure 9.7c displays ARI values with STF > 0.5. Note that it is possible to identify new accumulation lines immersed in the BD; these lines coincide with the next four STF shown in Fig. 9.7d. The STF intersecting points are called star points [32] and we can observe two of them (indicated with black arrows) in Fig. 9.7d.

Finally, we observe that the laminar states occur inside the interval between these two star points. In Fig. 9.8 we show the growth of this inter-star zone with h variation. In Fig. 9.8a we have the case previous to the star point's emergence; in Fig. 9.8b there are already two star points with the intermittence zone between 264.22 and 266.58 ms. For $h = 30$ ms, shown in Fig. 9.8c, this interval ranges from 262.2 to 267.2 ms. In Fig. 9.8d, for $h = 35$ ms, the P interval spans from 257.9 to 278.8 ms, and shows a region where 3-period rhythm is present. This region is also observed in the right upper corner of Fig. 9.2. The P intervals for

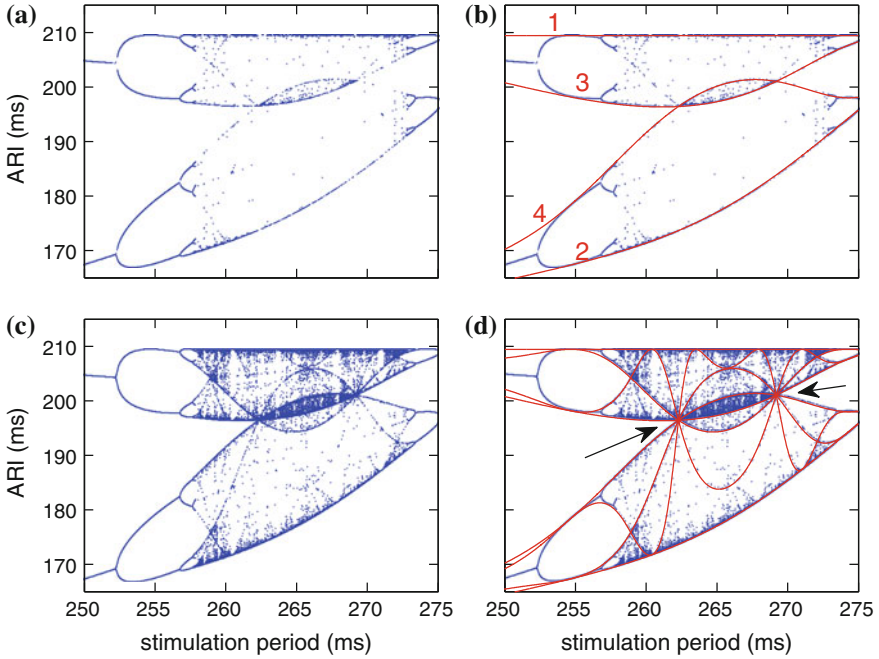

Fig. 9.7 Super Track Functions. **a** displays the plot for NRF = 1. **b** shows the superposition of points in **a** with the first four STF. In **c**, points with NRF > 0.5 are shown, while **d** shows the first eight STF superposed to points in **c**. *Arrows* indicate the star points where laminar states can be found in between

these reported laminar states were verified by calculating the corresponding Vmax. A final observation is that if $h = 29.3$ ms and there are not yet star points, there exists a very short laminarity zone ranging from 264.6 to 265.8 ms.

9.4 Discussion

CRC has been employed before as a theoretical and experimental tool to study the occurrence of alternans in portions of cardiac tissue or in isolated cardiac cells. The first authors to propose a graphic method to predict this phenomenon were Nolasco and Dahlen [3]. In 1984 Guevara et al. proposed an iteration map over the CRC obtained from the embryonic chick cardiomyocytes [2]. They showed that by reducing the stimulation period, a period-doubling bifurcation was achieved, moving from 1-period rhythm to a 2-period rhythm in the APD. This bifurcation occurred when the fixed-point in the map has a slope equal to 1. Later, it was shown

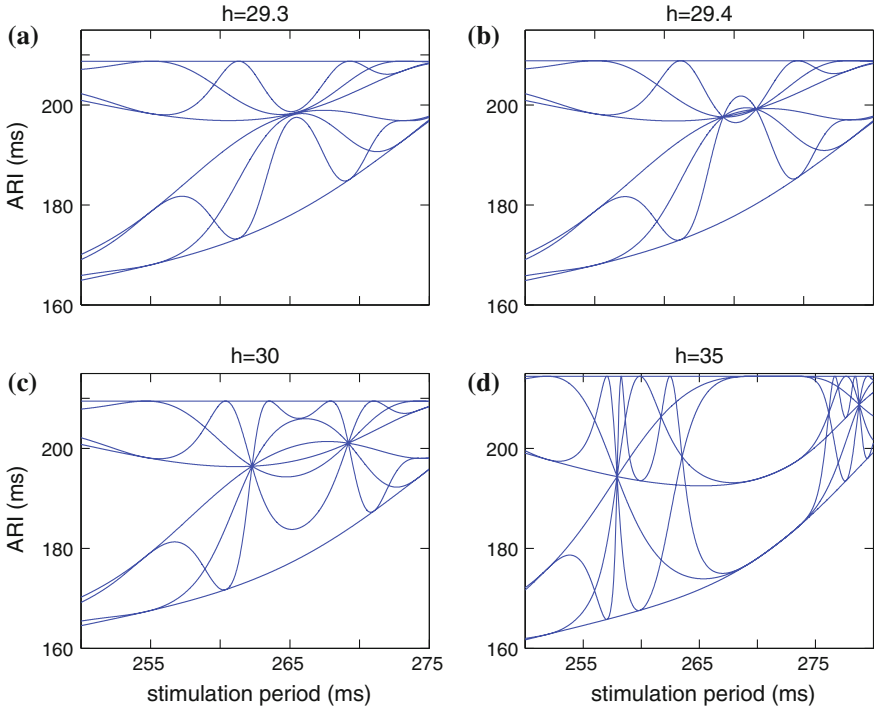

Fig. 9.8 Laminar states growth. **a** to **d** show how, as h increases, the region with laminar states (that is, between star points marked in Fig. 9.7) increases

that drugs like verapamil (a calcium channel blocker) modified the slope of the CRC and consequently the bifurcation behavior in cardiac assays [5, 10]. However, other studies have found that, in some cases, the alternant behavior is reached even with slopes less than 1, suggesting that theoretical models should consider further complexities, for instance including memory effects [6, 7, 12].

The experimental results for the human ventricle CRC of cardiac patients obtained by Franz and colleagues, showed the existence of a local maximum in some of the ventricular cells, contiguous to others with monotonic CRC. Furthermore, they showed that inhomogeneous conditions can prevail in the ventricle of a cardiac patient [24]. Although most of the experimental studies can be made in very controlled settings, this is, of course, faraway from clinical conditions [16, 20, 22]. Therefore, exploring different bifurcation scenarios for the onset of alternans in different CRC profiles is still a relevant task.

Here we have modeled the profile of the non-monotonic CRC as the addition of two functions, an exponential and a Gaussian function, adjusting the maximum of the Gaussian function to the local maximum of the experimental curve. We showed that this local maximum leads to an advance in the emergence of alternans of 122.7 ms respect to the observed P in a monotonic CRC. Using analytical tools

derived from RQA, we described how the alternans region is increased as the parameters are varied, as well as how the alternans magnitude and complexity of the system are modified. Also, RP was shown to be a very useful tool to study the complex behaviors nested in the irregular zones. In summary, we found that, as h grows, a period-doubling bifurcation in direct and inverse cascades appears, characterized by chaotic zones without laminar states and subsequently by chaotic zones with laminar states. We showed that the evolution of these regions can be followed-up using the STF built with the orbits obtained from time-series generated from the CRC-LM.

Although to our knowledge there is no experimental data in human ventricle to contrast with our theoretical results, there is a report by Watanabe et al. [11] that qualitatively reproduced our findings employing sheep epicardial strips with a biphasic CRC. Indeed, varying the stimulation period from 200 to 110 ms, the sequence of rhythms observed was 1:1, irregular, 8:8, 4:4, irregular, 2:2. This sequence starts with short couplings (1:1) goes through longer coupled rhythm to finish again with a short coupling (2:2) and with irregularity windows in between. In our analysis, we showed that, for h values higher than 26.95 ms, short stimulation periods produce a period-doubling bifurcation cascade that ends in chaotic behavior. Furthermore, subsequent stimulation period reduction leads to an inverse bifurcation cascade that ends in a 1-period rhythm, again with irregular behavior in between.

As we mentioned before, alternans are considered a cardiac risk predictor. In clinical examinations, alternant T-waves are generally not present without chronotropic challenge. Here we have shown, in a model for cardiac patient ventricle cells, that CRC with a local maximum advances the response to accelerated activation, as well as increasing the complexity of the system. However, it is not clear how this complex behavior propagates or influences the neighboring cells which may or may not have a monotonic CRC. A simulation considering an expanded range of CRC would be important for a better understanding of the evolution and dynamics of the CRC-LM and, hopefully, contribute in the development of a new type of indexes or cardiac risk metrics.

Acknowledgment We thank JoAnn Miller her valuable help in reviewing this manuscript.

References

1. J.M. Pastore, S.D. Girouard, K.R. Laurita, F.G. Akar, D.S. Rosenbaum, Mechanism linking T-wave alternans to the genesis of cardiac fibrillation. Circulation **99**, 1385–1394 (1999)
2. M.R. Guevara, G. Ward, A. Shrier, L. Glass, Electrical alternans and period doubling bifurcation. IEEE Comput. Cardiol. **167**, 18–24 (1984)
3. J.B. Nolasco, R.W. Dahlen, A graphic method for the study of alternation in cardiac action potentials. J. Appl. Physiol. **25**, 191–196 (1968)
4. M.R. Franz, The electrical restitution curve revisited: steep or flat slope—which is better? J. Cardiovasc. Electrophysiol. **14**, S140–S147 (2003)

5. J.N. Weiss, A. Garfinkel, H.S. Karagueuzian, T.P. Nguyen, R. Olcese, P.S. Chen, Z. Qu, Perspective: a dynamics-based classification of ventricular arrhythmias. J. Mol. Cell. Cardiol. **82**, 136–154 (2015)
6. P.N. Jordan, D.J. Christini, Determining the effects of memory and action potential duration alternans on cardiac restitution using a constant-memory restitution protocol. Physiol. Meas. **25**(4), 1013 (2004)
7. S.D. McIntyre, V. Kakade, Y. Mori, E.G. Tolkacheva, Heart rate variability and alternans formation in the heart: The role of feedback in cardiac dynamics. J. Theor. Biol. **350**, 90–97 (2014)
8. F. Fenton, E.M. Cherry, Cardiac Dynamics: Restitution (2012) http://www.scholarpedia.org/article/User:Flavio_H_Fenton/Proposed/Restitution
9. A. Garfinkel, Y.H. Kim, O. Voroshilovsky, Z. Qu, J.R. Kil, M.H. Lee, H.S. Karagueuzian, J. N. Weiss, P.S. Chen, Preventing ventricular fibrillation by flattening cardiac restitution. Proc. Natl. Acad. Sci. USA **97**, 6061–6066 (2000)
10. A. Karma, New paradigm for drug therapies of cardiac fibrillation. PNAS **97**(11) (2000)
11. M. Watanabe, N.F. Otani, R.F. Gilmour Jr, Biphasic restitution of action potential duration and complex dynamics in ventricular myocardium. Circ. Res. **76**, 915–921 (1995)
12. H. Dvir, S. Zlochiver, The Interrelations among stochastic pacing, stability, and memory in the Heart. Biophys. J. **107**, 1023–1034 (2014)
13. H. Jiang, D. Zhao, B. Cui, Z. Lu, J. Lü, F. Chen, M. Bao, Electrical restitution determined by epicardial contact mapping and surface electrocardiogram: its role in ventricular fibrillation inducibility in swine. J. Electrocardiol. **41**, 152–159 (2008)
14. B.C. Knollmann, T. Schober, A.O. Petersen, S.G. Sirenko, M.R. Franz, Action potential characterization in intact mouse heart: steady-state cycle length dependence and electrical restitution. Am. J. Physiol.- Heart Circul. Physiol. **292**(1), H614–H621 (2006). doi:10.1152/ajpheart.01085.2005
15. M. Watanabe, D.P. Zipes, R.F. Gilmour Jr, Oscillations of diastolic interval and refractory period following premature and postmature stimuli in canine cardiac Purkinje fibers. Pacing Clin. Electrophysiol. **12**, 1089–1103 (1989)
16. R. Wu, A. Patwardhan, Restitution of action potential duration during sequential changes in diastolic intervals shows multimodal behavior. Circ. Res. **96**, 634–641 (2012)
17. M.R. Guevara, L. Glass, A. Shrier, Phase locking, period-doubling bifurcations, and irregular dynamics in periodically stimulated cardiac cells. Science **214**, 1350–1353 (1981)
18. G.M. Hall, S. Bahar, D.J. Gauthier, Prevalence of rate-dependent behaviors in cardiac muscle. Phys. Rev. Lett. **82**, 2995 (1999)
19. A. Karma, Electrical alternans and spiral wave breakup in cardiac tissue. Chaos **4**, 461–72 (1994)
20. O.E. Osadchii, Effects of ventricular pacing protocol on electrical restitution assessments in guinea-pig heart. Exp. Physiol. **97**(7), 807–821 (2012)
21. H. Arce, A. Xu, H. González, M.R. Guevara, Alternans and higher-order rhythms in an ionic model of a sheet of ischemic ventricular muscle. Chaos **10**(2), 411 (2000)
22. B.J. Bas, Restitution of the action potential in cat papillary muscle. J. Physiol. **228**, 1717–1724 (1975)
23. M.R. Franz, C.D. Swerdlow, L.B. Liem, J. Schaefer, Cycle length dependence of human action potential duration in vivo. J. Clin. Investig. **82**(3), 972–979 (1988)
24. A.M. Yue, M.R. Franz, P.R. Roberts, J.M. Morgan, Global endocardial electrical restitution in human right and left ventricles determined by noncontact mapping. J. Am. Coll. Cardiol. **46** (6), 1067 (2005)
25. L.L. Trulla, A. Giuliani, J.P. Zbilut, C.L. Webber Jr, Recurrence quantification analysis of the logistic equation with transients. Phys. Lett. A **223**(4), 255–260 (1996)
26. J. Gao, Recurrence time statistics for chaotic systems and their applications. Phys. Rev. Lett. **83**(16), 3178–3181 (1999)
27. J.P. Eckmann, S.O. Kamphorst, D. Ruelle, S. Ciliberto, Liapunov exponents from time series. Phys. Rev. A **34**(6), 4971–4979 (1986)

28. C.L. Webber Jr, J.P. Zbilut, Dynamical assessment of physiological systems and states using recurrence plot strategies. J. Appl. Physiol. **76**, 965–973 (1994)
29. J.P. Zbilut, C.L. Webber Jr, Embeddings and delays as derived from quantification of recurrence plots. Phys. Lett. A **171**, 199–203 (1992). doi:10.1161/01.CIR.99.10.1385
30. N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, J. Kurths, Recurrence plot based measures of complexity and its application to heart rate variability data. Phys. Rev. E **66**(2), 026702 (2002)
31. N. Marwan, C.M. Romano, M. Thiel, J. Kurths, Recurrence plots for the analysis of complex systems. Phys. Rep. **438**(5–6), 237–329 (2007)
32. E.M. Oblow, Supertracks, supertrack functions and chaos in the quadratic map. Phys. Lett. A **128**(8), 406–412 (1988). doi:10.1016/0375-9601(88)90119-3

Chapter 10
The Early Phases of Epileptogenesis Induced by Status Epilepticus Are Characterized by Persistent Dynamical Regime of Intermittency Type

Massimo Rizzi, Federica Frigerio and Valentina Iori

Abstract Pre-clinical studies aimed to test potential anti-epileptogenic therapies by using the animal models of epileptogenesis induced by status epilepticus (SE), highlighted that the early days following the end of this primary insult represent a crucial temporal window for the subsequent development of epilepsy. In this study, we characterized the EEG dynamics during such crucial period of epileptogenesis, according to the conceptual framework of nonlinear dynamical systems. To this aim, we analyzed by recurrence quantification analysis (RQA) the EEG signals associated to the early days of epileptogenesis induced by SE in rodents according to two well-known experimental protocols, i.e., (i) SE induced by electrical stimulation of the hippocampus in rats (n = 7) and (ii) SE induced by the intra-amygdala administration of kainic-acid in mice (n = 6). We show that the EEG signals during the early 1–2 days post-SE are characterized by an enhanced and persistent rate of occurrence of dynamical regimes of intermittency type. This finding is common to both models of SE, hence it could represent the dynamical hallmark of pro-epileptogenic insults and could correlate with the efficacy of such insults to promote functional changes leading to the development of epilepsy. Future works aimed to deepen our findings could lead to the identification of a potential prognostic factor of the development of epilepsy as well as improve the portability of pre-clinical studies aimed to target new potential therapeutics designed to prevent the development of epilepsy.

10.1 Introduction

Brain insults as, for instance, traumatic brain injury, stroke or infectious diseases, are known as risk factors for the development of epilepsy and the prevention of this neurological disorder in individuals at risk still represents an important medical

M. Rizzi (✉) · F. Frigerio · V. Iori
Department of Neuroscience, IRCCS—Istituto di Ricerche Farmacologiche Mario Negri, Via La Masa 19, 20156 Milan, Italy
e-mail: massimo.rizzi@marionegri.it; massimo.rizzi@gmail.com

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8_10

challenge. In this context, the term epileptogenesis (or latent period) is referred to as the period which elapses between the occurrence of a pro-epileptogenic insult and the emergence of the first spontaneous seizure. Pre-clinical studies accomplished by using several animal models of epileptogenesis have highlighted the occurrence of many cellular and functional alterations in the brain tissue. Among these alterations, to name a few, the neurodegeneration of brain tissue and the reorganization of the molecular architecture of individual neurons, as well as alterations of the blood-brain barrier permeability and the alteration of patterns of expression of molecules related to the immune response [1]. Targeting of some of these pathological mechanisms has shown favorable effects on the development of epileptogenesis in clinical studies (for a comprehensive review see [2], and references there in), pointing out that all cases of potential antiepileptogenic therapies which show a beneficial effect are to be administered during the early days of the latent period, often soon after the end of the primary insult [2]. This bulk of evidence strongly suggests that the very early days following the induction of a pro-epileptogenic insult represent a crucial time window for the development of epilepsy.

Many efforts are currently aimed to investigate on the potentially epileptogenic cellular and functional alterations occurring during such crucial early phases of epileptogenesis. However, much less effort is aimed to characterize the brain electrical activity which subtend the EEG during the same period. It cannot be underestimated that the characterization of the nature of the dynamics of the EEG activity during the latent period may contribute to shed light on the mechanisms of epileptogenesis as well as suggest new strategies of therapeutic intervention aimed to prevent the development of epilepsy. Accordingly, in this study we characterized the EEG dynamics occurring during the early phases of epileptogenesis, according to the paradigm of nonlinear dynamical systems. To this aim, we analysed the EEG signals associated to two well-known experimental models of status epilepticus (SE) induced in rodents, i.e., SE induced by the electrical stimulation of the hippocampus in rats and SE induced by intra-amygdala administration of kainic-acid in mice. As analytical method we specifically choose the recurrence quantification analysis [3], a valuable mathematical tool, aimed to efficiently characterize the nonlinear dynamics embedded in short, noisy and nonstationary time-series, being the latter two features intrinsically associated to the nature of EEG signals [3–5]. In the context of epilepsy, several research groups have already applied the RQA to analyse EEG signals, gaining intriguing results (for a reliable and updated survey of the state-of-the art of the application of the RQA to epileptic EEGs, see http://www.recurrence-plot.tk/bibliography.php). Nonetheless, one notices that all the available studies were aimed to predict/detect the occurrence of seizures or classify epileptic EEG patterns in cases of overt epilepsy in humans or following acute seizures induced in experimental animals. At present, despite the growing importance of the experimental models of SE in the epilepsy research, the characterization of nonlinear EEG dynamics during the development of epileptogenesis is lacking. Therefore, the application of the RQA to investigate the EEG dynamics during the early phases of epileptogenesis can be reasonably considered not only as a novelty per se but also as a necessity to better characterize these models.

10.2 Methods

10.2.1 Animals

Male Sprague-Dawley rats (275–300 g; Charles-River, Calco, Italy; n = 7) and male C57BL6N mice (~20 g; Charles-River, Calco, Italy; n = 6) were maintained in SPF facilities at the Mario Negri Institute and housed at a constant room temperature (23 °C) and relative humidity (60 ± 5 %) with free access to food and water and a fixed 12 h light/dark cycle.

10.2.2 Models of Epileptogenesis

A major contribution to the understanding of mechanisms involved in epileptogenesis comes from animal models where the primary pro-epileptogenic insult is the status epilepticus (SE) [6, 7]. SE is a condition defined as a seizure that persists for at least 5–10 min, or it repeats itself frequently enough so that there is not a resolution of contiguity [8]. In humans, SE may occur in association with different forms of brain injury and conditions such as stroke, brain trauma, withdrawal of anti-epileptic drug treatment and exposure to toxic nerve agents. As an experimental model of epileptogenesis, SE is commonly induced in rodents by electrical stimulation of specific brain areas or by the systemic or local administration of convulsive drugs [2, 6, 7, 9, 10]. In this work, we exploited the EEG signals originated from two well-known experimental models of SE, i.e., SE induced by the electrical stimulation of the hippocampus in rats [11, 12] and SE induced by intra-amygdala administration of kainic-acid in mice [13, 14].

The following experimental procedures were conducted in conformity with institutional guidelines that are in compliance with national (D.L. n.26, G.U. March 4, 2014) and international guidelines and laws (EEC Council Directive 86/609, OJ L 358, 1, December 12, 1987, Guide for the Care and Use of Laboratory Animals, U.S. National Research Council, 1996), and were reviewed and approved by the intramural ethical committee.

10.2.2.1 Surgical Procedure and Induction of the SE for the Rat Model of Epileptogenesis

Rats were surgically implanted under 1.5 % isoflurane anesthesia with 2 bipolar Teflon-insulated stainless-steel depth electrodes placed bilaterally into the temporal pole of the hippocampus (from bregma, mm: AP − 4.7; L ± 5.0; −5.0 below dura, [15]). Two screw electrodes were positioned over the nasal sinus and the cerebellum, and used as ground and reference electrodes, respectively. Electrodes were connected to a multipin socket and secured to the skull by acrylic dental cement. After surgical

procedures, rats were treated locally with Cicatrene powder (Neomicyn; Bacitracin; Glicyne; L-Cistein; DL-Treonin) and injected with Ampicillin (100 mg/kg, s.c.) for 4 days to prevent infections. Rats were allowed to recover from surgery in their home cage for 7–10 days. Before electrical stimulation, EEG baseline hippocampal activity was recorded in freely-moving rats for 24 h. Then, rats were unilaterally stimulated (50 Hz, 400 μA peak-to-peak, 1 ms biphasic square waves in 10 s trains delivered every 11 s) in the CA3 region of the ventral hippocampus for 90 min to induce SE according to a well-established protocol [11, 12]. SE was defined as the presence of continuous spike activity with a frequency higher than 0.5 Hz intermixed with high amplitude and frequency discharges lasting for at least 5 s, with a frequency of ≥8 Hz and an amplitude twofold-higher than the baseline. Spikes were defined as sharp waves with amplitude of at least 2.5-fold higher than the baseline and duration lower than 100 ms, or as a spike-and-wave with duration lower than 200 ms [9]. The end of SE was determined by the occurrence of inter-spike interval >2 s. No pharmacological intervention was done to stop SE since no mortality is observed in this model. In order to ascertain the development of epilepsy according to this model of epileptogenesis, rats were continuously EEG recorded (24 h/day) from SE induction until 2 spontaneous EEG seizures occurred. EEG was recorded using the Twin EEG Recording System connected with a Comet AS-40 32/8 Amplifier (sampling rate 400 Hz, high-pass filter 0.3 Hz, low-pass filter 70 Hz, sensitivity 2000 mV/cm; Grass-Telefactor, West Warwick, R.I., U.S.A.)

For this model of epileptogenesis, we considered for nonlinear analysis, the EEG signal which originated from the hippocampal electrode used for the induction of the SE by electrical stimulation.

10.2.2.2 Surgical Procedure and Induction of the SE for the Mouse Model of Epileptogenesis

Mice were surgically implanted under general gas anesthesia (1–3 % isoflurane in O_2) and stereotaxic guidance. A 23-gauge cannula was unilaterally positioned on top of the *dura mater* for the intra-amygdala injection of kainic acid (from bregma, mm: nose bar 0; anteroposterior −1.06, lateral −2.75) [16]. A nichrome-insulated bipolar depth electrodes (60 μm OD) was implanted in the dorsal hippocampus ipsilateral to the injected amygdala (from bregma, mm: nose bar 0; anteroposterior –1.8, lateral 1.5 and 2.0 below dura mater) [16]. Additionally, a cortical electrode was implanted onto the somatosensory cortex in the contralateral hemisphere. Electrodes were connected to a multipin socket (PlasticOne Inc., USA). One week after surgery, mice were connected to the EEG set up and the SE was induced by the injection of kainic acid (0.3 μg in 0.2 μl) in the basolateral amygdala in freely moving mice using a needle protruding of 3.9 mm below the implanted cannula. SE developed after approximately 10 min from kainic acid injection, and was defined by the appearance of continuous spikes with a frequency >1.0 Hz. Spikes were defined as sharp waves with an amplitude at least 2.5-fold higher than the baseline and a duration of <20 ms, or as a spike-and-wave with a duration of <200 ms [17].

After 40 min from SE onset, mice received diazepam (10 mg/kg, ip) to improve their survival. SE lasted for 7.0 ± 1.0 h and its end was determined by the occurrence of inter-spike interval >2 s. In order to ascertain the development of epilepsy according to this model of epileptogenesis, mice were continuously EEG recorded (24 h/day) from SE induction until 2 spontaneous EEG seizures occurred. EEG was recorded using the Twin EEG Recording System connected with a Comet AS-40 32/8 Amplifier (sampling rate 400 Hz, high-pass filter 0.3 Hz, low-pass filter 70 Hz, sensitivity 2000 mV/cm; Grass-Telefactor, West Warwick, R.I., U.S.A.)

For this model of epileptogenesis, we considered for nonlinear analysis, the EEG signal which originated from the hippocampal electrode positioned ipsilaterally to the site of intra-amygdala administration of kainic-acid.

10.2.3 Selection Criteria of the Time Windows of Epileptogenesis and the Related EEG Epochs, for Nonlinear Analysis

The duration of the epileptogenesis was variable. For rats, which underwent the protocol of induction of SE by electrical stimulation of the hippocampus, the duration of the latent period ranged from 7 to 11 days (mean ± SD = 9.17 ± 1.60) whereas for mice, which underwent the protocol of induction of SE by the intra-amygdala injection of kainic-acid, ranged from 6 to 8 days (mean ± SD = 6.80 ± 0.84). However, we focused our attention on the early days of epileptogenesis, since they were shown as crucial for the development of epilepsy. Specifically, we evaluated EEG signals associated to the early 5 days of epileptogenesis for rats and the early 4 days of epileptogenesis for mice. This criterion allowed to evaluate comparable time windows of epileptogenesis between the two models of induction of SE. Indeed, for rats, the early 5 days of epileptogenesis corresponded to 55.99 ± 10.03 % (mean ± SD) of the total time of epileptogenesis, whereas for mice, the early 4 days of epileptogenesis corresponded to 59.52 ± 7.14 % (mean ± SD) of the total time of epileptogenesis. Additionally, in authors' opinion, these temporal windows were sufficiently far (at least 2 days) from the onset of the first seizure, thus reducing the risk of bias due to the possibility of the emergence of mechanisms leading to the onset of the first seizure.

The EEG signals associated to each day of the selected time windows, were sampled at 400 Hz and analog-to-digital converted with 12 bit precision. Respect to the original EEG recordings which were 0.3–70 Hz band-pass filtered, we further restricted the range of EEG frequencies according to evidence suggesting a more circumscribed frequency band (<50 Hz) as being more relevant to the aim of our investigation. Indeed, significant changes of the power spectral density of the EEG frequency bands, ranging from delta (1–4 Hz) to low-gamma frequency band (30–50 Hz), were shown to occur in cortical structures, as the hippocampus, during the early phases of epileptogenesis [18, 19]. These changes show a temporal profile

similar to that of cellular and functional alterations occurring in the brain during the latent period, and they correlate with the emergence of well-known comorbidities of epilepsy (e.g., spatial memory deficits and/or social behavioral dysfunctions), which were shown to develop during the early phases of epileptogenesis [18, 19]. Therefore, it is reasonable to expect that frequency bands below 50 Hz are those bearing the majority of information on EEG dynamics during the early phases of epileptogenesis. Accordingly, also considering the electrical noise due to the frequency of the power grid (50 Hz), we applied a low-pass 50 Hz filter to all EEG recordings. These recordings were also visually inspected for removal of evident gross artifacts, if any. Successively, in the time interval from 01:00 pm to 07:00 pm, we extracted 12-s epochs (i.e. 4800 data points) every 10 min of each EEG signal. For rats, the same procedure was applied also to basal EEG recordings obtained before the induction of the SE, whereas for mice there were no basal EEG recordings available. We chose the time interval from 01:00 pm to 07:00 pm since the ordinary tiding up of the animal housings is usually accomplished in the morning and the light/dark cycle switches from one state to the other at 07:00 am/pm, respectively. Therefore, we decided to extract EEG epochs during this period in order to minimize the potential effects of environmental stress and variation of circadian rhythms on the EEG dynamics of the animals. Each epoch was then analyzed by the recurrence quantification analysis.

10.2.4 The Recurrence Quantification Analysis

All the possible states of a nonlinear dynamical system are usually represented as trajectories in the phase (or state) space, being a trajectory the depiction of the time evolution of a set of states of the system. A fundamental property of nonlinear dynamical systems is the recurrence of states, which are expected to become arbitrarily close to each other after a sufficient time [20]. By exploiting this property of dynamical systems, Eckmann and colleagues [21] introduced the technique of the Recurrence Plot (RP), which is a graphical representation of a square matrix, in which the matrix elements correspond to those times at which a state of a dynamical system recurs. Technically, the RP reveals all the times when the phase space trajectory of the dynamical system visits roughly the same area in the phase space [22] (see also www.recurrence-plot.tk). The RPs allow to represent any *m*-dimensional phase space trajectory through a two-dimensional plot of its recurrences, since each recurrence of a state at two different times is marked within the two-dimensional squared matrix according to a binary decision, e.g., ones and zeros (black and white dots in the plot) to denote if these two states are recurrent (one/black) or not (zero/white). From the graphical perspective, the RPs exhibit characteristic large scale and small scale patterns denoted by Eckmann et al. [21] as typology and texture, respectively. Importantly, the small scale structures of texture consist of single dots, diagonal lines as well as vertical and horizontal lines. Single, isolated recurrence points can occur if states are rare and can be due to chance or

noise. A diagonal line occurs when a segment of the trajectory is parallel to another segment, i.e. the trajectory visits the same region of the phase space at different times. The length of a diagonal line is determined by the duration of such similar local evolution of the trajectory segments and was shown to be inversely related to the extent of the Lyapunov exponent. A vertical (horizontal) line marks a time length in which a state does not change or changes very slowly, as in laminar states (intermittency).

As intended since their introduction, RPs allow the visual inspection of high dimensional phase space trajectories by giving hints about their time evolution even in case of short, noisy and nonstationary dataset. However, although important behavioral properties of dynamical systems can be qualitatively inferred from a visual inspection of RPs by a well-trained investigator, being RPs intrinsically a visualizing tool, their usefulness is nevertheless limited when one considers to exploit RPs extensively as analytical technique. Such limitation was totally removed by Zbilut and Webber [3] who introduced the Recurrence Quantification Analysis (RQA), a powerful analytical technique based on the mathematical definition of variables introduced in order to measure some important properties of a nonlinear dynamical system by mean of an appropriate quantification of the small scale structures of the corresponding RP. At the present state of development of the RQA, several variables have been introduced [3, 23] (see also http://www.recurrence-plot.tk). In this work, we considered the variables which have been defined as follows:

- Recurrence Rate (REC), which represents the density of recurrence points in a RP. The recurrence rate corresponds with the probability that a specific state will recur and is expressed as

$$REC = \left(\frac{1}{N^2}\right) \sum_{i,j=1}^{N} R_{i,j} \tag{10.1}$$

with $R_{i,\,j} = \Theta\,(\mathrm{RAD_i} - \|\,\mathrm{x_i} - \mathrm{x_j}\|)$, $\mathrm{x_i} \in \mathfrak{R}^{\mathrm{m}}$, i, j = 1 … *N*, where *N* is the number of considered states, $\mathrm{x_i}$, $\mathrm{RAD_i}$ is a threshold distance (a.k.a. *radius*), $\|\cdot\|$ a norm and $\Theta(\cdot)$ the Heaviside function;

- Determinism (DET), which is the fraction of recurrence points forming diagonal lines. Diagonal lines represent epochs of similar time evolution of states of the system. Therefore, DET is related with the determinism of the system and is expressed as

$$DET = \frac{\sum_{l=l_{min}}^{N} lP(l)}{\sum_{l=1}^{N} lP(l)} \tag{10.2}$$

where l is the diagonal line length considered when its value is $\geq l_{min}$ and $P(l)$ is the probability distribution of line lengths;

- Longest Diagonal Line (DMAX), which measures the dynamical stability of the system, being inversely related to the largest Lyapunov exponent

$$DMAX = max(\{l_i, i = 1, \ldots N_i\}) \tag{10.3}$$

- Shannon Entropy (ENT) of the distribution of the line lengths, which is a measure of the complexity of the recurrence structure

$$ENT = -\sum_{l=l_{min}}^{N} p(l) \ln p(l) \tag{10.4}$$

where $p(l) = P(l)/N_l$ is the probability to find a diagonal line of exactly length l in the RP, being N_l the total number of diagonal lines.

- Laminarity (LAM), which is the fraction of recurrence points forming vertical lines. Vertical lines represent unchanged or slowly changing states of the system and are associated to laminar states. LAM is defined as

$$LAM = \frac{\sum_{v=v_{min}}^{N} vP(v)}{\sum_{v=1}^{N} vP(v)} \tag{10.5}$$

where v is the vertical line length considered when its value is $\geq v_{min}$ and $P(v)$ is the probability distribution of vertical line lengths;

- Trapping time (TT), which represents the average length of the vertical lines

$$TT = \frac{\sum_{v=v_{min}}^{N} vP(v)}{\sum_{v=v_{min}}^{N} vP(v)} \tag{10.6}$$

- Longest Vertical Line (VMAX), which represents the length of the longest vertical line

$$VMAX = max(\{v_i, i = 1, \ldots, N_v\}) \tag{10.7}$$

The computation of RQA variables is based on the preliminary reconstruction of the trajectories of the system in a phase space. In the specific case of a univariate time series, as the EEG, such reconstruction is based on the so-called time delay-embedding procedure as introduced by Takens [24]. The exceptional usefulness of the Takens' theorem can be particularly appreciated in the fields of biomedicine and life sciences in that the variables determining the underlining dynamics of systems are often unknown, and time series of just one single

observable can be the only data available. According to the Takens' theorem, the topological features of any high-dimensional system consisting of multiple coupled variables can be reconstructed by measuring just a single variable of that system [24, 25]. In other words, a single observable of the system is sufficient to allow the reconstruction of the trajectories in a high-dimensional phase space. This phase space is different from the 'true' phase space, and the reconstructed trajectories do not preserve the geometric shape as they would have had in the 'true' phase space. However, the reconstructed phase space preserves the properties of the dynamical system, since they are unaffected by smooth coordinate changes (topological invariants). A major advantage of the Takens' theorem is in that the 'true' phase space of the system can be totally unknown, as it is often the case in the context of biological systems. A single point of a trajectory in a phase space reconstructed according to the Takens' theorem, represents a vector made of m time points selected from the time series, where two consecutive time points are delayed by a predetermined time lag (a.k.a. *time delay*). The number of time points (m) chosen as components of vectors is referred to as the *embedding dimension* of the time-series. Therefore, the reconstructed phase space has m-axis, one for each of the m components of the vector. Besides the embedding dimension and the time delay, RQA requires to set a threshold value, the *radius* (RAD), by which a state point embedded in the reconstructed phase space can be considered as being recurrent, as defined above, hence included in the recurrence matrix.

The embedding dimension, the RAD and, to a lesser extent, the time delay are the most important parameters for an accurate RQA and their determination is a crucial step. To this aim, we accomplished a preliminary scaling analysis of subsets of epochs, according to the strategy suggested by Webber and Zbilut [25]. Scaling analysis allows to determine sets of values of embedding dimension and RAD for which the nonlinear deterministic patterns (if any) of the time series under investigation are sufficiently unveiled. A significant deterministic pattern to consider as a reference for an appropriate choice of parameters is the exponential scaling behavior of the variable REC versus the variable RAD. This scaling behavior is graphically manifested as a linear tract on a log-log plot of REC versus RAD and is usually expected to occur at low percentage values of REC, typically from 0.05 to 2 %. However, we noticed that the common approach of choosing a fixed RAD and then to determine the RQA variables might not be the most appropriate method for the analysis of our EEG epochs. Indeed, the log-log plots of REC versus RAD may show a linear tract of variable extension. Therefore, a fixed RAD does not ensure that the percentage of REC is within the linear range of the log-log plot of REC versus RAD, being this condition a stringent requirement for an appropriate estimation of RQA variables [25]. Therefore, in our study, it was advisable to consider the RAD as a variable and the REC as a parameter that we set to 1 %, a percentage which assured a sparse recurrence matrix and the preservation of the linear relationship of log-log plot of REC versus RAD in all cases considered. Additionally, scaling analysis suggested as a good choice for the embedding dimension, any value ranging from 10 to 15, hence we choose 12. We decided to not set a time delay common to all epochs and to determine the appropriate time delay for each epoch by considering the first

local minima of the mutual information function. The value of the time delay was used also for the determination of the Theiler's window [26], that we calculated according to Gao [27], i.e., [(m − 1) * (time delay)], thus reducing the influence of tangential motion in the estimate of the RQA variables [23].

The recurrence matrix was normalized to the maximum distance [25] and the variable RAD was expressed as percentage of the recurrence matrix, whereas the distance between vectors was computed as Euclidean distance. Additionally, the minimum diagonal line length was set at a conservative value of 5, since the observational noise was shown to significantly affect the value of DET by increasing the amount of diagonal lines with spurious diagonal segments with length ≤4 [28]. The minimum vertical line length was set at the default value of 2, since there are no indications available on the effect of noise on this parameter.

10.2.5 Surrogates Technique and Test of Significance

The significance of each RQA variable was tested against the null hypothesis of being the expression of a Gaussian linear stochastic process. According to the surrogate data technique as introduced by Theiler [26, 29], and similarly to Ouyang and colleagues [30], for each 12-s EEG epoch, we generated 50 surrogate epochs by the technique of the inverse Fourier transform with phase randomization [29, 31], thus creating surrogate epochs which preserved the Fourier amplitude as well as the same distribution of values of the original epoch, with the exception that nonlinear determinism, if any, was disrupted. For each of the 50 surrogate epochs, the RQA variables were calculated and then the so-called Theiler's sigma (TS) was determined [26, 30] as follows:

$$TS = \frac{\left| VAR_{orig} - \overline{VAR}_{surr} \right|}{SD\left(VAR^{i}_{surr} \right)} \tag{10.8}$$

where VAR_{orig} refers to any RQA variable (*RAD, DET, LAM*, etc.…) calculated for the original EEG epoch, and $\overline{VAR}_{surr}$ and $SD\left(VAR^{i}_{surr} \right)$ refer to the mean and the standard deviation of the same variable as calculated considering the 50 surrogates created from that original EEG epoch. For any RQA variable, when its corresponding value of TS > 1.96, the amplitude of that variable reflects actual nonlinear processes occurring in the original EEG epoch. In this work, only RQA variables which passed the TS test were considered.

10.2.6 Computational Resources and Software

The time for calculation of the RQA variables was considerably shortened by the High-Throughput-Computing Technology, a.k.a. Grid Computing Technology [32].

A Grid infrastructure represents a network of geographically-distributed computational resources which allow scientific communities to develop applications to be executed on distributed computational and storage resources across the Internet. The computational resources used in this study were those of the Italian e-Infrastructure of the INFN (National Institute of Nuclear Physics), which is part of the Italian Grid Infrastructure (IGI) that is fully integrated in the EGI project (European Grid Infrastructure), the Europe's leading grid computing infrastructure co-funded by the EU in the context of the 6th and 7th Framework Program. The middleware, i.e., the software specifically developed to manage the workload across the distributed computational and storage resources of a grid infrastructure, was gLite v. 3.1-3.2 and the interactive sessions were made by a Scientific-Linux command-line interface.

In this study the RQA was accomplished by the applications RQS and RQH, as developed by Webber [33] (freely available for Windows operating systems at http://homepages.luc.edu/~cwebber). However, in order to implement the Theiler's window and to make the applications executable in the Grid Computing environment, the applications RQS and RQH were adapted with minor modifications as ANSI C code. The applications for the calculation of the time delay as well as for construction of surrogates were *mutual, minima, extrema* and *surrogates*, available in the TISEAN software package [34] (freely available at http://www.mpipks-dresden.mpg.de/~tisean).

10.2.7 Statistics

Statistical tests were performed by the software package GraphPad Prism 6 (GraphPad Software Inc., San Diego, California, USA). The great majority of datasets did not passed the test for normality according to the D'Agostino and Pearson omnibus test for normality as well as the Shapiro-Wilk normality test in case of small datasets ($P < 0.05$). Therefore, we used the nonparametric Kruskal-Wallis test, followed by Dunn's multiple comparisons test, in case of significant variation of medians. For graphical purposes, datasets were represented by the descriptive statistics mean ± SEM.

10.3 Results

Results are organized in figures and tables so that for each figure representing the trends of specific RQA metrics the related statistics are reported in the table labeled with the same progressive number (e.g.: for metrics represented in Fig. 10.1 the related statistics are reported in Table 10.1, and so on).

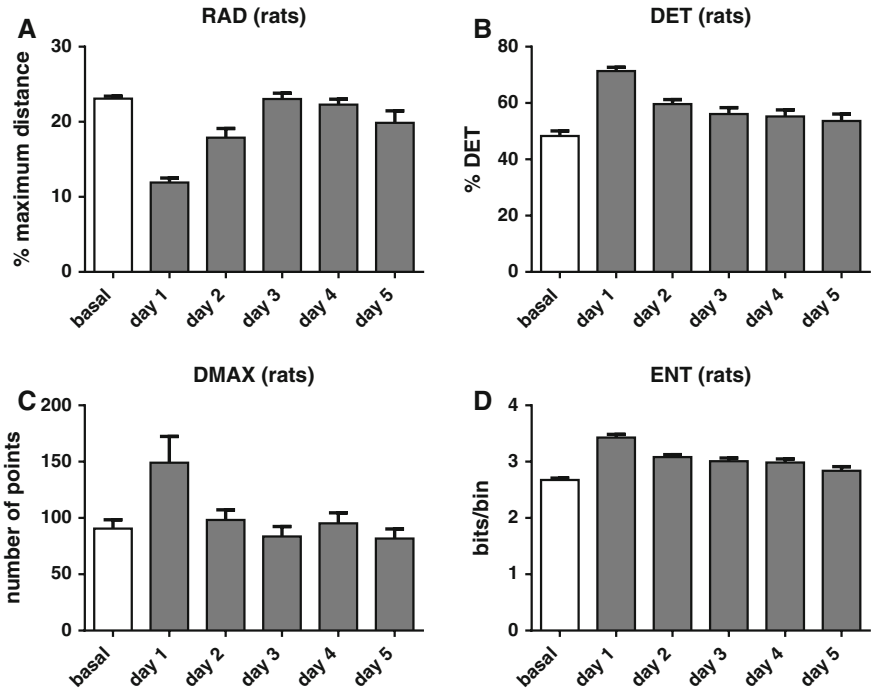

Fig. 10.1 Histograms (mean ± SEM) of datasets of RQA variables calculated in basal condition (*white bar*) and during the early 5 days of epileptogenesis (*grey bars*) for the model of SE induced in rats by the electrical stimulation of the hippocampus. *Panel A* depicts the variation of the variable RAD, whereas *panels B*, *C* and *D*, show how RQA metrics based on the *diagonal lines* of the RPs (DET, DMAX and ENT) are affected by the progression of epileptogenesis after the end of the primary insult

For both models of SE, the early 1–2 days of epileptogenesis are characterized by a significant reduction of the medians of the variable RAD (for rats, Fig. 10.1 panel A, $P < 0.0001$; for mice, Fig. 10.3 panel A, $P < 0.0001$—detailed statistics in Tables 10.1 and 10.3, respectively), thus suggesting an increased rate of recurrences during this time window, i.e., a higher density of the trajectories in the phase space of the system. Hence, a lower value of the variable RAD is sufficient to gain the percentage of recurrent points that we kept fixed at 1 % in order to assure the exponential scaling behavior of REC versus RAD, a basic requirement for a proper RQA [25].

The trends of the medians of the variables DET and LAM are reported in Figs. 10.1 (panel B) and 10.2 (panel A), for rats, whereas for mice are reported in Figs. 10.3 (panel B) and 10.4 (panel A). The detailed statistics for these metrics are reported in Tables 10.1 and 10.2 for rats and in Tables 10.3 and 10.4 for mice. The amount of DET significantly increases in the early 1–2 days post-SE for both models (for rats, Fig. 10.1 panel B, $P < 0.0001$; for mice, Fig. 10.3 panel B, $P = 0.0006$—Tables 10.1 and 10.3, respectively, for detailed statistics), whereas the amount of LAM shows opposite trends of variation (compare Fig. 10.2 panel A

Table 10.1 Nonparametric statistics (Kruskal-Wallis test, followed by Dunn's multiple comparisons test) of RQA metrics reported in Fig. 10.1

	RAD	DET	DMAX	ENT
Kruskal-Wallis test				
P value	<0.0001****	<0.0001****	0.0039**	<0.0001****
No of groups	6	6	6	6
No of values (total)	231	336	162	366
Kruskal-Wallis statistic	89.48	76.85	17.33	90.04
Dunn's multiple comparisons test				
Groups	*P value*	*P value*	*P value*	*P value*
Basal versus day 1	<0.0001****	<0.0001****	0.0074**	<0.0001****
Basal versus day 2	0.0054 **	0.0006 ***	ns	<0.0001****
Basal versus day 3	ns	ns	ns	0.0002***
Basal versus day 4	ns	ns	ns	0.0022**
Basal versus day 5	ns	ns	ns	ns
Day 1 versus day 2	0.0002***	<0.0001****	ns	0.0008***
Day 1 versus day 3	<0.0001 ****	<0.0001 ****	0.0043 **	0.0001***
Day 1 versus day 4	<0.0001 ****	<0.0001 ****	ns	<0.0001 ****
Day 1 versus day 5	<0.0001****	<0.0001****	0.0449*	<0.0001****
Day 2 versus day 3	ns	ns	ns	ns
Day 2 versus day 4	ns	ns	ns	ns
Day 2 versus day 5	ns	ns	ns	ns
Day 3 versus day 4	ns	ns	ns	ns
Day 3 versus day 5	ns	ns	ns	ns
Day 4 versus day 5	ns	ns	ns	ns

Asterisks summarize the level of significance
ns not significant

with Fig. 10.4 panel A—Tables 10.2 and 10.4, respectively, for detailed statistics). Indeed, LAM significantly increases the day after the end of the primary insult in rats (Fig. 10.2 panel A, $P < 0.0001$, as compared to basal and the following days of epileptogenesis), but significantly decreases in mice (Fig. 10.4 panel A, $P = 0.001$, as compared to the following days of epileptogenesis, no basal condition available). However, it is necessary to consider that when the RAD is a variable and the REC is a fixed parameter, the values of DET and LAM, per se, cannot correlate with the actual amount of determinism and laminarity of the time-series, because the RAD, although variable, can change only under the constrain to reach the imposed fixed amount of recurrences. This may underestimate/overestimate the actual amount of DET and LAM of the time-series. However, the values of DET and LAM calculated considering a small value of the RAD are necessarily the expression of a greater content of determinism and laminarity respect to the same values calculated for a large value of the RAD, since the lower the value of the RAD, the higher the density of trajectories. Therefore, when the RAD is a variable, the actual content of

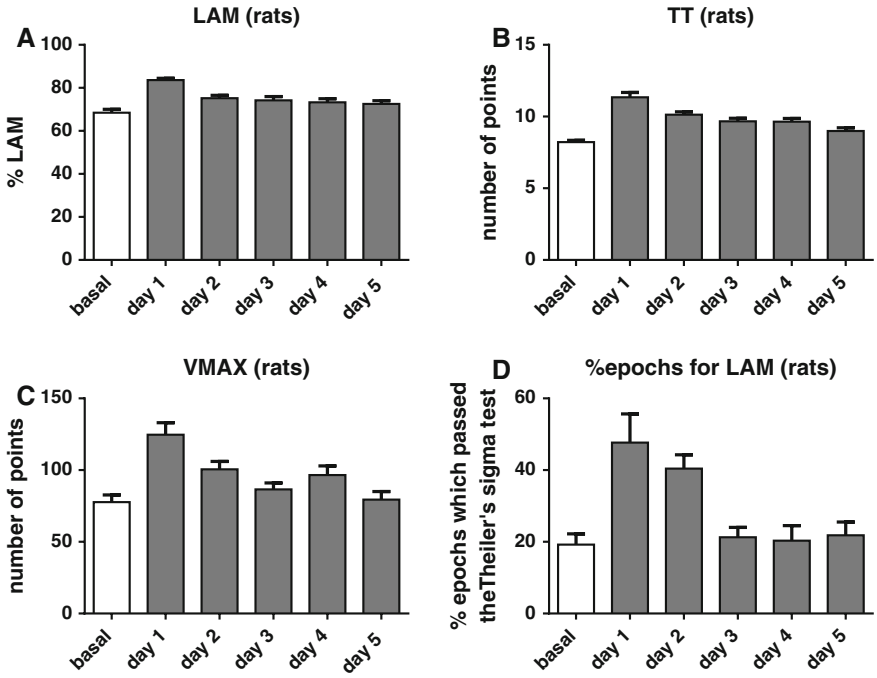

Fig. 10.2 Histograms (mean ± SEM) of datasets of RQA variables calculated in basal condition (*white bar*) and during the early 5 days of epileptogenesis (*grey bars*) for the model of SE induced in rats by the electrical stimulation of the hippocampus. *Panels A*, *B* and *C*, depict how RQA metrics based on the *vertical lines* of the RPs (LAM, TT and VMAX) are affected by the progression of epileptogenesis after the end of the primary insult. *Panel D* reports the percentage of epochs which passed the Theiler's sigma test for the RQA variable LAM during the early phases of epileptogenesis

determinism and laminarity of the time-series can be estimated by the ratios DET/RAD and LAM/RAD. Accordingly, the higher the ratios, the higher the amount of DET and LAM in the time-series. This normalization is equivalent to say that, in our context, if the RAD had been constant as it occurs in the classical RQA, the fraction of recurrent points which form diagonal/vertical lines (i.e., DET and LAM) would be greater for phase spaces with high densities of trajectories, because a fixed RAD would include a greater number of recurrent points.

The normalization of variables DET and LAM respect to the variable RAD immediately shows that the opposite trends of variation of LAM were only apparent. Indeed, the normalization of the variables DET and LAM, highlights how the early 1–2 days of epileptogenesis are characterized by a marked increase of determinism (for rats, Fig. 10.5 panel A, $P < 0.0001$; for mice, Fig. 10.5 panel C, $P < 0.0001$—the detailed statistics are reported in Table 10.5) and laminarity (for rats, Fig. 10.5 panel B, $P < 0.0001$; for mice, Fig. 10.5 panel D, $P < 0.0001$—see

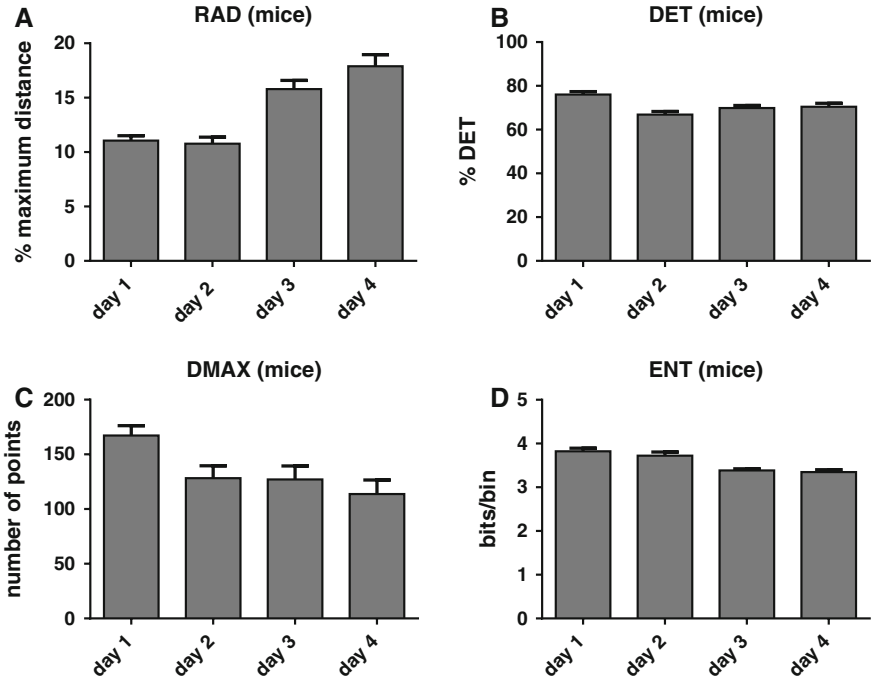

Fig. 10.3 Histograms (mean ± SEM) of datasets of RQA variables calculated during the early 4 days of epileptogenesis for the model of SE induced in mice by the intra-amygdala administration of kainic-acid. *Panel A* depicts the variation of the variable RAD, whereas *panels B*, *C* and *D* show how RQA metrics based on the *diagonal lines* of the RPs (DET, DMAX and ENT) are affected by the progression of epileptogenesis after the end of the primary insult

Table 10.5 for the detailed statistics) in the EEG dynamics which emerge after the end of the SE, a feature occurring in both models of epileptogenesis.

Not only the amount of determinism and laminarity are markedly increased following the end of the SE but, generally, all RQA variables show a significant increase in the early 1–2 days of the latent period. Indeed, this occurred for RQA metrics based on the diagonal line lengths (DMAX), and their related distributions (ENT) (for rats, Fig. 10.1 panels C and D, $P = 0.0039$ and $P < 0.0001$ respectively; for mice, Fig. 10.3 panels C and D, $P = 0.0101$ and $P < 0.0001$ respectively—the detailed statistics for these metrics are reported in Table 10.1 for rats and Table 10.3 for mice) as well as for metrics based on vertical line lengths, as the TT (for rats, Fig. 10.2 panel B, $P < 0.0001$—Table 10.2 for statistics; for mice, Fig. 10.4 panel B, $P = 0.0053$—Table 10.4 for statistics) and the VMAX (for rats, Fig. 10.2 panel C, $P < 0.0001$—Table 10.2 for statistics; for mice, Fig. 10.4 panel C, $P < 0.0001$—Table 10.4 for statistics).

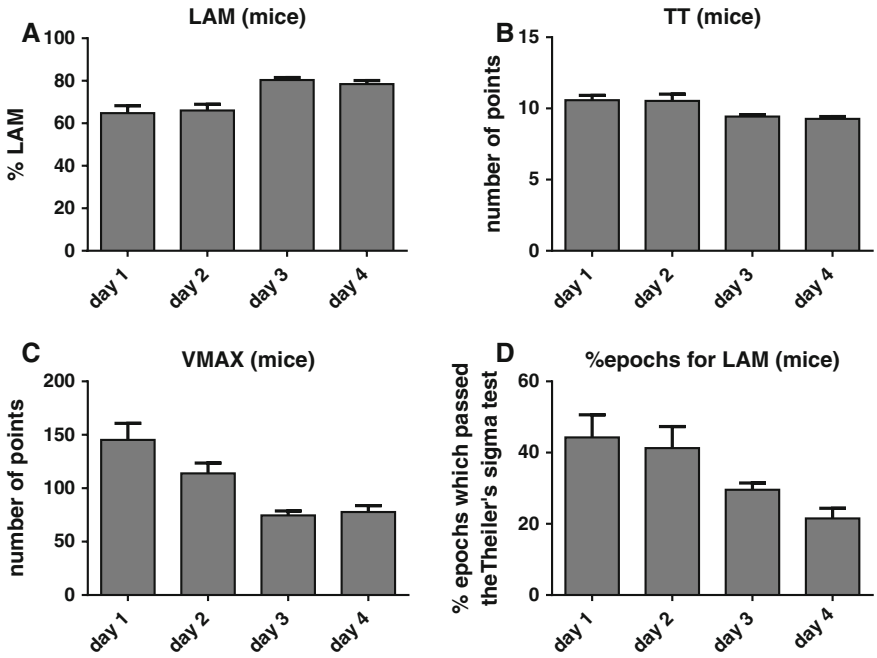

Fig. 10.4 Histograms (mean ± SEM) of datasets of RQA variables calculated during the early 4 days of epileptogenesis for the model of SE induced in mice by the intra-amygdala administration of kainic-acid. *Panels A*, *B* and *C*, depict how RQA metrics based on the *vertical lines* of the RPs (LAM, TT and VMAX) are affected by the progression of epileptogenesis after the end of the primary insult. *Panel D* reports the percentage of epochs which passed the Theiler's sigma test for the RQA variable LAM

Since the variation of the variable LAM gives a significant hint of the nature of the underlying dynamics, we also calculated the percentage of EEG epochs which passed the TS test for this variable. This metric depicts the probability of the occurrence of a laminar state in the time interval 1:00 pm–7:00 pm depending on the day of the epileptogenesis considered, i.e., how often this dynamic state emerges in function of the day of progression of the epileptogenesis. Not surprisingly, the early 1–2 days are those significantly affected by the highest percentage of occurrence of EEG epochs with significant values of the variable LAM (for rats, Fig. 10.2 panel D, $P = 0.0004$—statistics in Table 10.2; for mice, Fig. 10.4 panel D, $P = 0.0078$—statistics in Table 10.4), thus denoting a general persistent condition which is characterized by a high rate of laminarity states within the EEG epochs, i.e., high values of the variable LAM (Fig. 10.5 panels B and D, for rats and mice, respectively), with the additional evidence that this condition repeats frequently, i.e., high percentages of occurrence of EEG epochs which passed the TS test (Fig. 10.2 panel D, for rats; Fig. 10.4 panel D, for mice).

Table 10.2 Nonparametric statistics (Kruskal-Wallis test, followed by Dunn's multiple comparisons test) of RQA metrics reported in Fig. 10.2

	LAM	TT	VMAX	% epochs LAM
Kruskal-Wallis test				
P value	<0.0001****	<0.0001****	<0.0001****	0.0004***
No of groups	6	6	6	6
No of values (total)	307	352	186	41
Kruskal-Wallis statistic	66.94	82.44	35.35	22.67
Dunn's multiple comparisons test				
Groups	*P value*	*P value*	*P value*	*P value*
Basal versus day 1	<0.0001****	<0.0001****	<0.0001****	0.0168*
Basal versus day 2	0.0037**	<0.0001****	0.0159*	0.023*
Basal versus day 3	ns	<0.0001****	ns	ns
Basal versus day 4	ns	<0.0001****	ns	ns
Basal versus day 5	ns	ns	ns	ns
Day 1 versus day 2	0.0001***	ns	ns	ns
Day 1 versus day 3	<0.0001****	0.0116*	0.0017**	0.0434*
Day 1 versus day 4	<0.0001****	0.0014**	ns	ns
Day 1 versus day 5	<0.0001****	<0.0001****	0.0002***	ns
Day 2 versus day 3	ns	ns	ns	ns
Day 2 versus day 4	ns	ns	ns	ns
Day 2 versus day 5	ns	0.0451*	ns	ns
Day 3 versus day 4	ns	ns	ns	ns
Day 3 versus day 5	ns	ns	ns	ns
Day 4 versus day 5	ns	ns	ns	ns

Asterisks summarize the level of significance
ns not significant

Table 10.3 Nonparametric statistics (Kruskal-Wallis test, followed by Dunn's multiple comparisons test) of RQA metrics reported in Fig. 10.3

	RAD	DET	DMAX	ENT
Kruskal-Wallis test				
P value	<0.0001****	0.0006***	0.0101*	<0.0001****
No of groups	4	4	4	4
No of values (total)	184	208	76	259
Kruskal-Wallis statistic	44.23	17.36	11.33	32.62
Dunn's multiple comparisons test				
Groups	*P value*	*P value*	*P value*	*P value*
Day 1 versus day 2	ns	0.0002***	ns	0.024*
Day 1 versus day 3	0.0002***	0.0098**	ns	<0.0001****
Day 1 versus day 4	<0.0001****	ns	ns	<0.0001****
Day 2 versus day 3	<0.0001****	ns	ns	ns
Day 2 versus day 4	<0.0001****	ns	ns	ns
Day 3 versus day 4	ns	ns	ns	ns

Asterisks summarize the level of significance
ns not significant

Table 10.4 Nonparametric statistics (Kruskal-Wallis test, followed by Dunn's multiple comparisons test) of RQA metrics reported in Fig. 10.4

	LAM	TT	VMAX	% epochs LAM
Kruskal-Wallis test				
P value	0.001**	0.0053**	<0.0001****	0.0078**
No of groups	4	4	4	4
No of values (total)	192	269	92	23
Kruskal-Wallis statistic	16.17	12.72	22.84	11.88
Dunn's multiple comparisons test				
Groups	*P value*	*P value*	*P value*	*P value*
Day 1 versus day 2	ns	ns	ns	ns
Day 1 versus day 3	0.0166*	0.0192*	<0.0001****	ns
Day 1 versus day 4	ns	0.0105*	0.0049**	0.0177*
Day 2 versus day 3	0.0017**	ns	ns	ns
Day 2 versus day 4	ns	ns	ns	0.043*
Day 3 versus day 4	ns	ns	ns	ns

Asterisks summarize the level of significance
ns not significant

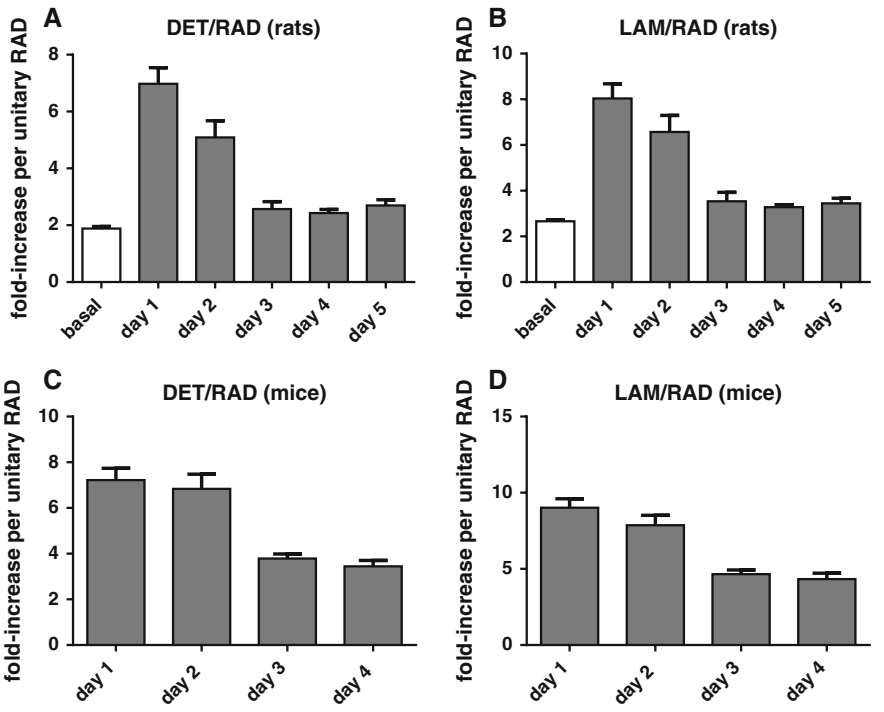

Fig. 10.5 Histograms (mean ± SEM) of datasets of the RQA variables DET and LAM after their normalization respect to the variable RAD, during the progression of the early phases of epileptogenesis in both models of SE

Table 10.5 Nonparametric statistics (Kruskal-Wallis test, followed by Dunn's multiple comparisons test) of RQA metrics reported in Fig. 10.5

	DET/RAD (rats)	LAM/RAD (rats)	DET/RAD (mice)	LAM/RAD (mice)
Kruskal-Wallis test				
P value	<0.0001****	<0.0001****	<0.0001****	<0.0001****
No of groups	6	6	4	4
No of values (total)	333	303	208	165
Kruskal-Wallis statistic	153.9	131.1	49.86	50.88
Dunn's multiple comparisons test				
Groups	*P value*	*P value*	*P value*	*P value*
Basal versus day 1	<0.0001****	<0.0001****	na	na
Basal versus day 2	<0.0001****	<0.0001****	na	na
Basal versus day 3	ns	0.0364*	na	na
basal versus day 4	ns	0.0166*	na	na
Basal versus day 5	0.0135*	ns	na	na
Day 1 versus day 2	<0.0001****	0.0003***	ns	ns
Day 1 versus day 3	<0.0001****	<0.0001****	<0.0001****	<0.0001****
Day 1 versus day 4	<0.0001****	<0.0001****	<0.0001****	<0.0001****
day 1 versus day 5	<0.0001****	<0.0001****	na	ns
Day 2 versus day 3	0.0002***	0.0108*	0.0003***	0.0007***
Day 2 versus day 4	0.0005***	0.0117*	<0.0001****	0.0002***
Day 2 versus day 5	ns	ns	na	ns
Day 3 versus day 4	ns	ns	ns	ns
Day 3 versus day 5	ns	ns	na	ns
Day 4 versus day 5	ns	ns	na	ns

Asterisks summarize the level of significance
ns not significant

10.4 Discussion

In this study we applied for the first time the RQA to investigate on the nature of EEG dynamics during the early phases of epileptogenesis, thus extending the application of the RQA to the characterization of important models of epileptogenesis in the context of the pre-clinical research in epilepsy. In particular, we deepened the nature of the dynamics involved in the early crucial phases of epileptogenesis elicited by the experimental models of SE.

In the following, we refer to the values of variables DET and LAM after their normalization versus the variable RAD.

For a meaningful evaluation of the nature of the dynamics subtending EEG activity following the end of a pro-epileptogenic primary insult as the SE, we considered the variations of all the RQA metrics based on diagonal and vertical

lines of RPs. At a first sight, the marked increases of ENT, DMAX and DET in the early 1–2 days post-SE may appear contradictory, since a greater degree of complexity (i.e. a wider distribution of diagonal line lengths, hence a higher ENT) usually implies an increase of the exponential instability, normally associated to a decrease of the variables DMAX and DET. Indeed, the DMAX is inversely related to the Lyapunov exponents, and the DET is a measure of the degree of predictability of the system. Nonetheless, the values of both these variables significantly increase.

However, when one takes into account also recurrence measures based on the vertical lines of RPs, it is immediate to notice that also the variables LAM, TT and VMAX significantly increase in the same post-SE temporal window. Altogether, the variations of these variables converge to point out the existence of bifurcation points which subtend a specific non-linear behavior known as intermittency, a dynamical state in which periods of apparent periodicity alternate irregularly with periods which show dynamics driven by the emergence of one or more different attractors, maybe of chaotic nature [23]. We did not determined the type of intermittency. However, this will be done in the near future, also considering an interesting application of the RQA in this context [35]. Intermittency may account for the broad variability of the DMAX, hence the reason for which the statistics related to the DMAX for both models of the SE, are those showing the highest P values (according to the Kruskal-Wallis test), as also the lack of statistical significances (according to the Dunn's post hoc test) for the same variable calculated from the EEG epochs of mice (Fig. 10.3 panel C).

The episodes of intermittency in the early latent period are not only more accentuated but, generally, they occur much more frequently, as evinced by the significantly high percentage of EEG epochs which passed the TS test for LAM. Therefore, the early post-SE days are affected by long-lasting temporal windows characterized by persistent instability of intermittency type. This condition appears as a general feature of the early phases of epileptogenesis triggered by SE, since we investigated two different models of SE, yielding similar results.

In the context of the epilepsy research, the occurrence of intermittency as dynamical regime is not a novelty since it was shown to emerge during ictal activities in epileptic patients or in rodents during the progression of the SE [36, 37]. However, our study shows that intermittency regimes persist beyond the end of the SE even without the occurrence of ictal events and emerge at high rates in the early days following the end of the primary insult.

It is of interest to notice that also without the emergence of typical widespread ictal activities, ictogenic brain areas are characterized by a high rate of occurrence of spatially distributed micro-seizures [41]. Since the intermittency may characterize the dynamics of ictal events [37], the emergence of intermittency regimes in the early phases of epileptogenesis could be the expression of such micro-seizures originating from an epileptogenic network under formation, due to the occurrence of neurodegenerative phenomena and functional alterations affecting the early phases of epileptogenesis following the end of the SE [38–40]. From this perspective, one also notices that our data show a general descending trend of the RQA

variables after day 2 post-SE, likely due to a progressive weakening of the contribution of nonlinear dynamics to the EEG signal detected by the measuring electrode. Interestingly, this descending trend could be compatible with the formation of topographically distributed ictogenic micro-domains, as a consequence of the progression of epileptogenesis [41, 42].

In authors' opinion, it is also of interest to notice that the percentage of epochs passing the TS test for the variable LAM in basal conditions (Fig. 10.2 panel D) is not negligible, thus suggesting that the intermittency could be a dynamical state which may occur also without an overt pathology. This finding is consistent with evidence of occurrence of micro-seizures also in basal conditions, as it was shown in healthy humans [41]. Altogether, our findings could reasonably support the hypothesis that the main aberrant feature of ictal events is not their nonlinear dynamical nature, but their spreading over an abnormally extended spatio-temporal scale, as proposed also by other investigators [41, 42].

Since it is reasonable to hypothesize that the emergence of intermittency could be the dynamic hallmark of pro-epileptogenic insults and could correlate with the efficacy of such insults to promote functional changes leading to the development of epilepsy, an important future work will be to establish the existence (if any) of a relationship between pro-epileptogenic insults and the emergence of such dynamics. To this aim, the same therapeutic interventions which were shown able to prevent/modify the development of epilepsy can be used, and this study could lead to (i) the identification of the emergence of intermittency as a potential prognostic factor of the development of epilepsy and (ii) a significant improvement of the portability of pre-clinical studies aimed to test new potential therapeutics able to prevent the development of epilepsy. From this perspective, it is worth considering that, nowadays, the direct comparison of results obtained in different models of SE is often impaired. Indeed, metrics which are commonly used to evaluate the efficacy of potential therapeutics in the experimental models (e.g., the number of spikes per hour) are based on the EEG patterns intrinsically associated to the progression of the prolonged ictal activity of the SE. However, these patterns of EEG activity are remarkably model-dependent. Therefore, it is often hard to compare experimental results from different models, when considering these metrics. However, despite the remarkable differences between the two models of SE that we investigated, we clearly show that the enhanced and persistent rate of occurrence of laminarity states in the EEG dynamics during the early days of epileptogenesis is a common feature of both models. This suggests that such dynamic behaviour could be model-independent and could be a better metric to validate results obtained from the application of potential therapeutics using different models of epileptogenesis, thus improving the portability of pre-clinical studies.

Acknowledgments We wish to thank Prof. Charles Webber, who generously provided the source codes of RQA applications used in this work. We also wish to thank Dr. Giuseppe La Rocca (Italian National Institute of Nuclear Physics, Division of Catania, Italy) and Prof. Giuseppe Barbera (Italian National Institute of Nuclear Physics, Division of Catania and Department of Physics and Astronomy of the University of Catania, Italy) for their technical assistance on the

usage of Grid Computing resources and services provided by the Italian Grid Infrastructure (IGI, http://www.italiangrid.it/).

Conflict of Interest Authors have no conflict of interest to declare.

References

1. A. Pitkänen, K. Lukasiuk, Molecular and cellular basis of epileptogenesis in symptomatic epilepsy. Epil. Behav. **14**(1), 16–25 (2009)
2. A. Pitkänen, K. Lukasiuk, Mechanisms of epileptogenesis and potential treatment targets. Lancet Neurol. **10**(2), 173–186 (2011)
3. C.L. Webber, J.P. Zbilut, Dynamical assessment of physiological systems and states using recurrence plot strategies. J. Appl. Physiol. **76**(2), 965–973 (1994)
4. J.P. Zbilut, A. Giuliani, C.L. Webber, Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification. Phys. Lett. A **246**(1), 122–128 (1998)
5. J.P. Zbilut, A. Giuliani, C.L. Webber, Recurrence quantification analysis as an empirical test to distinguish relatively short deterministic versus random number series. Phys. Lett. A **267**(2), 174–178 (2000)
6. W. Löscher, C. Brandt, Prevention or modification of epileptogenesis after brain insults: experimental approaches and translational research. Pharmacol. Rev. **62**(4), 668–700 (2010)
7. A. Pitkänen, Therapeutic approaches to epileptogenesis—hope on the horizon. Epilepsia **51**(s3), 2–17 (2010)
8. D.H. Lowenstein, B.K. Alldredge, Status epilepticus. N. Engl. J. Med. **338**(14), 970–976 (1998)
9. A. Pitkänen, I. Kharatishvili, S. Narkilahti, K. Lukasiuk, J. Nissinen, Administration of diazepam during status epilepticus reduces development and severity of epilepsy in rat. Epil. Res. **63**(1), 27–42 (2005)
10. M.J. Lehmkuhle, K.E. Thomson, P. Scheerlinck, W. Pouliot, B. Greger, F.E. Dudek, A simple quantitative method for analyzing electrographic status epilepticus in rats. J. Neurophysiol. **101**(3), 1660–1670 (2009)
11. M.G. De Simoni, C. Perego, T. Ravizza, D. Moneta, M. Conti, F. Marchesi, A. De Luigi, S. Garattini, A. Vezzani, Inflammatory cytokines and related genes are induced in the rat hippocampus by limbic status epilepticus. Eur. J. Neurosci. **12**(7), 2623–2633 (2000)
12. F. Noè, A.H. Pool, J. Nissinen, M. Gobbi, R. Bland, M. Rizzi, C. Balducci, F. Ferraguti, G. Sperk, M.J. During, A. Pitkänen, A. Vezzani, Neuropeptide Y gene therapy decreases chronic spontaneous seizures in a rat model of temporal lobe epilepsy. Brain **131**(6), 1506–1515 (2008)
13. G. Mouri, E. Jimenez-Mateos, T. Engel, M. Dunleavy, S. Hatazaki, A. Paucard, S. Matsushima, W. Taki, D.C. Henshall, Unilateral hippocampal CA3-predominant damage and short latency epileptogenesis after intra-amygdala microinjection of kainic acid in mice. Brain Res. **1213**, 140–151 (2008)
14. E.M. Jimenez-Mateos, T. Engel, P. Merino-Serrais, R.C. McKiernan, K. Tanaka, G. Mouri, T. Sano, C. O'Tuathaigh, J. Waddington, S. Prenter, N. Delanty, M.A. Farrell, D.F. O'Brien, M. R. Conroy, R.L. Stallings, J. deFelipe, D.C. Henshall, Silencing microRNA-134 produces neuroprotective and prolonged seizure-suppressive effects. Nat. Med. **18**(7), 1087–1094 (2012)
15. G. Paxinos, C. Watson, *The Rat Brain in Stereotaxic Coordinates* (Academic Press, New York, 2005)
16. K.B.J. Franklin, G. Paxinos, *The Mouse Brain in Stereotaxic Coordinates* (Academic Press, San Diego, 2008)

17. F. Frigerio, A. Frasca, I. Weissberg, S. Parrella, A. Friedman, A. Vezzani, F.M. Noe, Long-lasting pro-ictogenic effects induced in vivo by rat brain exposure to serum albumin in the absence of concomitant pathology. Epilepsia **53**(11), 1887–1897 (2012)
18. L. Chauvière, N. Rafrafi, C. Thinus-Blanc, F. Bartolomei, M. Esclapez, C. Bernard, Early deficits in spatial memory and theta rhythm in experimental temporal lobe epilepsy. J. Neurosci. **29**(17), 5402–5410 (2009)
19. J. Seo, S. Jung, S.Y. Lee, H. Yang, B.S. Kim, J. Choi, M. Bang, H.S. Shin, D. Jeon, Early deficits in social behavior and cortical rhythms in pilocarpine-induced mouse model of temporal lobe epilepsy. Exp. Neurol. **241**, 38–44 (2013)
20. H. Poincaré, Sur le problème des trois corps et les équations de la dynamique. Acta Math. **13**(1), A3–A270 (1890)
21. J.P. Eckmann, S.O. Kamphorst, D. Ruelle, Recurrence plots of dynamical systems. Europhys. Lett. **4**(9), 973 (1987)
22. N. Marwan, Encounters with neighbours: current developments of concepts based on recurrence plots and their applications. Ph.D. Thesis, University of Potsdam (2003)
23. N. Marwan, M. Carmen Romano, M. Thiel, J. Kurths, Recurrence plots for the analysis of complex systems. Phys. Rep. **438**(5), 237–329 (2007)
24. F. Takens, Detecting strange attractors in turbulence, in *Dynamical Systems and Turbulence*, Warwick 1980 (Springer, Berlin, 1981), pp. 366–381
25. C.L. Webber, J.P. Zbilut, Recurrence quantification analysis of nonlinear dynamical systems. in *Tutorials in Contemporary Nonlinear Methods for the Behavioral Sciences*, ed. by M.A. Riley, G.C. Van Orden, pp. 26–94 (2005). http://www.nsf.gov/sbe/bcs/pac/nmbs/chap2.pdf
26. J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, Farmer J. Doyne, Testing for nonlinearity in time series: the method of surrogate data. Physica D **58**(1), 77–94 (1992)
27. J. Gao, Z. Zheng, Direct dynamical test for deterministic chaos and optimal embedding of a chaotic time series. Phys. Rev. E **49**(5), 3807 (1994)
28. M. Thiel, M.C. Romano, J. Kurths, R. Meucci, E. Allaria, F.T. Arecchi, Influence of observational noise on the recurrence quantification analysis. Physica D **171**(3), 138–152 (2002)
29. T. Schreiber, A. Schmitz, Surrogate time series. Physica D **142**(3), 346–382 (2000)
30. G. Ouyang, X. Li, C. Dang, D.A. Richards, Using recurrence plot for determinism analysis of EEG recordings in genetic absence epilepsy rats. Clin. Neurophysiol. **119**(8), 1747–1755 (2008)
31. T. Schreiber, A. Schmitz, Improved surrogate data for nonlinearity tests. Phys. Rev. Lett. **77**(4), 635 (1996)
32. R. Barbera, G. La Rocca, M. Rizzi, Grid computing technology and the recurrence quantification analysis to predict seizure occurrence in patients affected by drug-resistant epilepsy, in *Data Driven e-Science* (Springer, New York, 2011), pp. 493–506
33. C.L. Webber, Introduction to recurrence quantification analysis RQA version 13.1 README. PDF. http://homepages.luc.edu/~cwebber/RQA131.EXE
34. R. Hegger, H. Kantz, T. Schreiber, Practical implementation of nonlinear time series methods: the TISEAN package. Chaos **9**(2), 413–435 (1999)
35. K. Klimaszewska, J.J. Żebrowski, Detection of the type of intermittency using characteristic patterns in recurrence plots. Phys. Rev. E **80**(2), 026214 (2009)
36. J.L. Velazquez, H. Khosravani, A. Lozano, B. Bardakjian, P.L. Carlen, R. Wennberg, Type III intermittency in human partial epilepsy. Eur. J. Neurosci. **11**(7), 2571–2576 (1999)
37. J.L.P. Velazquez, M.A. Cortez, O.C. Snead, R. Wennberg, Dynamical regimes underlying epileptiform events: role of instabilities and bifurcations in brain activity. Physica D **186**(3), 205–220 (2003)
38. J.A. Gorter, P.M.G. Pereira, E.A. Van Vliet, E. Aronica, F.H.L. Da Silva, P.J. Lucassen, Neuronal cell death in a rat model for mesial temporal lobe epilepsy is induced by the initial status epilepticus and not by later repeated spontaneous seizures. Epilepsia **44**(5), 647–658 (2003)

39. T. Araki, R.P. Simon, W. Taki, J.Q. Lan, D.C. Henshall, Characterization of neuronal death induced by focally evoked limbic seizures in the C57BL/6 mouse. J. Neurosci. Res. **69**(5), 614–621 (2002)
40. A. Vezzani, M. Conti, A. De Luigi, T. Ravizza, D. Moneta, F. Marchesi, M.G. De Simoni, Interleukin-1β immunoreactivity and microglia are enhanced in the rat hippocampus by focal kainic-acid application: functional evidence for enhancement of electrographic seizures. J. Neurosci. **19**(12), 5054–5065 (1999)
41. M. Stead, M. Bower, B.H. Brinkmann, K. Lee, W.R. Marsh, F.B. Meyer, B. Litt, J. Van Gompel, G.A. Worrell, Microseizures and the spatiotemporal scales of human partial epilepsy. Brain **133**, 2789–2797 (2010)
42. A. Bragin, C.L. Wilson, J. Engel, Chronic epileptogenesis requires development of a network of pathologically interconnected neuron clusters: a hypothesis. Epilepsia **41**(s6), S144–S152 (2000)

Chapter 11
Chromatic and Anisotropic Cross-Recurrence Quantification Analysis of Interpersonal Behavior

Ralf F.A. Cox, Steffie van der Steen, Marlenny Guevara, Lisette de Jonge-Hoekstra and Marijn van Dijk

Abstract Cross-recurrence quantification analysis (CRQA) is a powerful nonlinear time-series method to study coordination and cooperation between people. This chapter concentrates on two methodological issues related to CRQA on categorical data streams, which are commonly encountered in the behavioral sciences. Firstly, we introduce a more general definition of recurrence as 'behavioral matching', which can be applied to several kinds of matches simultaneously, visualized by a color coding. We will refer to this as cross-*matching*, and to the resulting quantification procedure as *Chromatic* CRQA. Secondly, cross-recurrence plots of categorical data often prominently consists of rectangular structures. This calls for a differential analysis of vertical and horizontal lines, rather than of diagonal lines. We introduce a simple procedure for this, referred to as *Anisotropic* CRQA. Both procedures are demonstrated with empirical studies on children's problem-solving behavior and by means of a model simulation. The authors hope that the ideas presented here increase the power and applicability of CRQA in the behavioral sciences, and that this chapter serves as a stepping stone for their mathematical and methodological development.

R.F.A. Cox (✉) · L. de Jonge-Hoekstra · M. van Dijk
Department of Psychology, Faculty of Behavioural and Social Sciences, University of Groningen, Groningen, The Netherlands
e-mail: r.f.a.cox@rug.nl

L. de Jonge-Hoekstra
e-mail: lisette.hoekstra@rug.nl

M. van Dijk
e-mail: m.w.g.van.dijk@rug.nl

S. van der Steen
Department of Special Needs Education and Youth Care, Faculty of Behavioural and Social Sciences, University of Groningen, Groningen, The Netherlands
e-mail: s.van.der.steen@rug.nl

M. Guevara
Institute of Psychology, Universidad del Valle, Cali, Colombia
e-mail: marlenny.guevara@correounivalle.edu.co

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8_11

11.1 Introduction

Behavioral scientists intend to study human behavior as it unfolds in time, aiming to capture the behavioral regularities at the different time scales that apply to a given task and context, and within the existing methodological constraints. These regularities are the key to understanding the processes underlying social behavior, learning and development. Typically the measurement scale of the behavioral dimension (eye movements, speech) is nominal in such studies, and signifies the specific set of categories (gaze here or there, say this or that kind of word) into which the researcher wishes to compress the behavior under study. Data collection hence results in an ordered sequence of observed behavioral categories. When the research context involves two people who are engaged in some form of interaction, the collected data set consists of two such categorical behavioral streams. The task at hand then is to analyze these streams in order to unravel the dynamic organization of the interactions.

To date, one of the best nonlinear techniques to analyze such data, which exploits the temporal structure of the behavioral streams, is recurrence quantification analysis (RQA) and its bivariate cousin cross-recurrence quantification analysis (CRQA). Both RQA and CRQA were originally designed for time series variables of a continuous measurement scale [1–4]. In the past decade, several methodological advances made these techniques more appropriate for the analysis of categorical data, for instance as collected in the behavioral sciences. As a result CRQA has found its way into the study of interpersonal behavior, and has now been applied successfully to explore, for instance, syntactic coordination between children and parents [5, 6], eye movement synchronization between speakers and listeners [7, 8], mother-infant gaze during a still-face procedure [9], infant-parent visual coordination in joint attention [10], nonverbal attunement in social interaction [11–13], and transitions in parent-child conflict conversation due to intervention [14]. In all these cases CRQA has provided valuable and unique knowledge about interpersonal coordination and its relation to other psychological constructs or its development.

Specific methodological issues which typically arise in CRQA on categorical time series (and which set it apart somewhat from CRQA on continuous time series) have been encountered in most if not all of the studies listed above. Although several of these issues have been noted, described and (sometimes even) solved in an ad hoc fashion (see also [15–18]), they have yet to be addressed more generally and more formally. Here we will focus on the following main issues, which are relevant to the behavioral sciences.

The first issue pertains to the definition of recurrence. Recurrence is essentially operationalized in a rather straightforward way akin to mimicking, that is, as ‘being of the same behavioral category’. However for interpersonal coordination to be successful, oftentimes, entirely different behaviors of the two interacting persons are

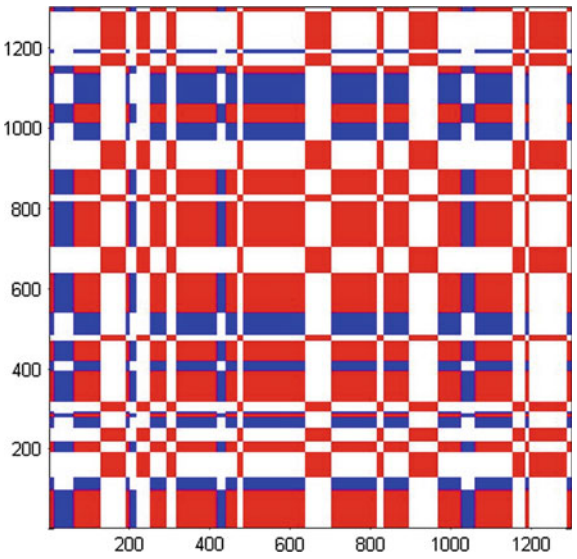

Fig. 11.1 A patchwork or multi-colored checkerboard CRP produced by the cross-matching procedure, forming the basis for Chromatic CRQA. In this case two different matches of behavioral categories were distinguished, depicted by a *red* and *blue* color code, together with the non-matching *white areas*. (For more details see Sect. 11.2.3)

called for (from the same functional behavioral class), for instance, being silent as the other person speaks. Recurrence procedures should be able to keep track of this, which necessitates a proper generalization of the concept of recurrence. As a corollary this creates the opportunity of distinguishing and analyzing different kinds of recurrences within one context.

The second issue has to do with the prominent rectangular structures which are common in cross-recurrence plots (CRPs) of categorical time series (see Fig. 11.1). These structures reflect some persistence in the interaction, partly due to the coarse-grained nature of the measurement. It makes sense to focus the analysis on the non-diagonal line measures (*LAM*, *TT*, *MaxL*), because for categorical time series these are often more reliable and informative. In addition, a differential analysis of the CRQA-measures along the horizontal and vertical directions in the (anisotropic) CRP might help to tear apart the respective contributions of each of the two subsystems to the dynamics.

Concretely, the goal of this methodological chapter is twofold: First, to introduce an alternate definition of cross-recurrence as cross-matching, for coupled dynamical systems measured on a nominal scale. Second, to initiate the development of recurrence procedures that enable the analysis of asymmetry in such systems, based on anisotropy in the CRP. To dub their existence we will call these methodological variations cross-*matching* and *Chromatic* CRQA and as *Anisotropic* CRQA, respectively. The merit of the proposed techniques will be demonstrated by two empirical studies and a model simulation. Note that although these issues will be presented somewhat more formally than has been done previously, no claims are made about mathematical rigour and methodological completeness.

11.2 Cross-Matching and Chromatic Quantification Analysis

11.2.1 Cross-Matching Procedure

There are numerous contexts and tasks in everyday behavior for which the atom of dyadic interaction is not defined on the basis of an equal or similar behavioral state for the two interacting people. More often than not effective and efficient interpersonal coordination requires individuals to behave quite differently from each other and sometimes even perform opposite actions in a coupled way (e.g. [11–21]). In addition, frequently behavioral scientists wish to keep track of, distinguish between, and relate several behavioral atoms, and wish to distil their temporal structure. The canonical concept of recurrence, that is, detecting similar behavioral categories at the same or any other moment in time, is not sufficiently powerful or accurate to capture the coordination dynamics of the interaction in those cases (cf. [16]). A variation of this basic procedure based on 'behavioral matching' should then be applied, so as to provide more flexibility in the analysis as well as a richer set of outcome measures.

The cross-matching (where 'cross' refers to the analysis of two time series) procedure on categorical data uses a fairly simple parameter setting for the analysis. It is largely similar to the canonical cross-recurrence procedure for categorical data which essentially reflects the typical arrangement of categorical data, but implements the research or context dependent choice of what recurrence should entail based on some desired matching of categories in the two time series. Taking an embedding dimension of 1 and a delay of 1 for the phase space reconstruction usually works well (e.g. [15, 22]). Setting the radius or threshold to 0 classifies 'perfect' recurrence within the reconstructed phase space, articulating a matching choice where behavioral categories in the two time series need to be equal to qualify as recurrent. Other matching choices can also be understood in terms of the radius parameter, when one realizes that the reconstructed phase space for categorical data has a discrete topology and the radius parameter can only take integer values. However, a more efficient and direct way to implement cross-matching and the subsequent chromatic quantification analysis is by means of the following procedure.[1]

The basic parameterization outlined above actually entails that the CRP can be drawn without any formal phase space reconstruction. In fact, the CRP is constructed by plotting all congruent appearances of some pre-specified matching values within a pair of time series in a plane, by placing one of the time series along the horizontal axis and the other along the vertical axis. In this way the CRP represents all those instances when the behavioral state of one subsystem at some moment in time is matched appropriately by the behavioral state of the other subsystem at the same or any other moment in time during the observation.

[1]Note that recoding the data oftentimes will also do the trick. However, complex cases with multiple matches and non-matches require a more general and powerful approach. Interestingly, this also gives rise to a set of additional measures, as will become clear further on.

This essentially generalizes the canonical 'recurrence' notion by a more appropriate 'behavioral matching' notion, specified by the researcher beforehand on theoretical or practical grounds in a specific context. And since this can readily be executed for more than one combination of categories simultaneously, cross-matching enables the distinction between several kinds of behavioral matches. The different matches can be tracked by means of a color coding. A straightforward way to implement this is by constructing multiple cross-recurrence matrices $\boldsymbol{CR}_{i,j}(color)$, one for each *color*. That is, one cross-recurrence matrix is constructed for each desired kind of match between the behavioral states of the dyadic partners on the two axes of the CRP. By applying this procedure, all recurrences of the different kinds of matching are represented by differently colored dots in the plane, resulting in a patchwork or multi-colored checkerboard CRP (see Fig. 11.1).

A straightforward application of the cross-matching procedure can be found for instance in the context of turn-taking in a conversation. For instance, in a study by Reuzel et al. [12, 13], speech rhythms of two conversing people were analyzed, by tracking in their nonverbal time series those instances where only one of the interlocutors was talking ('matches') and marking as 'non-matches' those instance where both were talking or where both were silent. So recurrence was defined as 'talking' (code 1) in one time series and 'silent' (code 0) in the other. The other combinations (0-0 and 1-1) were non-recurrent. This characterization of interlocutors nonverbal behavioral matching provided a meaningful partitioning of this particular interaction. In this way, it was demonstrated that people achieved interactional synchrony in their speech rhythm, which was related to the perceived quality of the conversation.

11.2.2 Chromatic CRQA

Fundamental to regular CRQA is that the spatial layout of dots in the CRP gives rise to several interesting recurrence measures [23]. Together the recurrence measures reveal the hidden structure concealed in the shared dynamics of the interacting subsystems across all possible time scales. The simplest measure of Chromatic CRQA, based on multiple kinds of behavioral matching (i.e. colors in the CRP), is the *recurrence rate* (*RR*) of a *color*:

$$RR(color) = \frac{1}{N^2} \sum_{i,j=1}^{N} CR_{i,j}(color), \tag{1}$$

where $\boldsymbol{CR}_{i,j}(color)$ are the cross-recurrence matrices underlying the CRP with respect to each individual *color*. *RR*(*color*) quantifies the density of one kind of match of behavioral states in the CRP. In other words, *RR*(*color*) reflects the extent to which the behaviors of one subsystem match those of the other subsystem in a specific way, across all possible time scales. As such, *RR*(*color*) is a basic measure

of the coordination between the two subsystems. Note that this definition equals that of the regular $RR(\varepsilon)$, for some radius ε, when the canonical concept of recurrence is applied.

A related measure is the *relative recurrence rate* (*rRR*) of a *color*:

$$rRR(color) = \frac{\sum_{i,j=1}^{N} CR_{i,j}(color)}{\sum_{color} \sum_{i,j=1}^{N} CR_{i,j}(color)}, \tag{2}$$

quantifying which proportion of the entire amount of behavioral matches (i.e. all colored dots) is of a particular kind (i.e. of a specific *color*). In other words, *rRR* (*color*) expresses the relative contribution of the different kinds of behavioral matches across all possible time scales to the overall coordination.

The other measures of Chromatic CRQA related to the diagonal and non-diagonal line structures in the CRP are straightforward generalizations of their regular definitions, which can be found, for instance, in [23]. Each of the measures can be calculated for the desired kinds of behavioral matches separately, that is, based on their respective cross-recurrence matrices $\boldsymbol{CR}_{i,j}(color)$. A number of the measures, viz. *rDET*(*color*) and *rLAM*(*color*) can also be calculated relative to all the colors included in the CRP, similar to *rRR*(*color*) above.

11.2.3 Application of Chromatic CRQA: Children's Dyadic Problem Solving

As an example application of Chromatic CRQA in the behavioral sciences, consider the following recent study on dyadic interaction between school-aged children in the context of science and technology education [24]. In six subsequent sessions over a period of 11 months, data was gathered about interpersonal behaviors of seven dyads of children with a mean age of 5.1 years, while they performed hands-on problem-solving tasks. For 20–25 min in each session, each child's individual interactive behaviors were coded at 1 Hz using five mutual exclusive behavioral categories: No work, Passive work, Copy work, Parallel work, and Collaborative work (for details see paper).

In this study the attractor dynamics of a small set of (predefined) quasi-stable states was analyzed, by applying Chromatic CRQA. To this end the five individual categories above, which potentially give rise to 25 a priori combinations, were conceptually coarse grained by means of cross-matching into three dyadic attractor states 'distributed dyadic interaction' (DDI) and 'unequal dyadic interaction' (UDI) and 'no dyadic interaction' (NDI), which were tracked separately. It should be noted that not all combinations were possible. Children could not copy each other at the same time, and collaboration always involved both individuals.

Figure 11.2 shows the discrete state space of the dyadic interaction. DDI indicates that both children were actively engaged with the task and contributed to the solution.

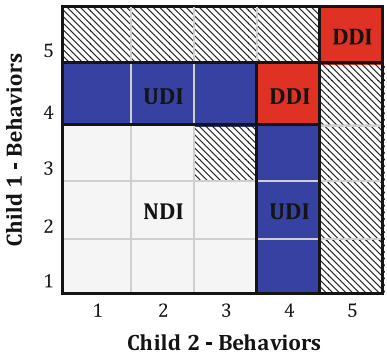

Fig. 11.2 State space or 'behavioral matrix' of dyadic problem-solving, presenting the potential combinations of dyadic interaction states. The colored areas define the behavioral matches used for Chromatic CRQA. The axis values represent the individual behaviors: No work *1*, Passive work *2*, Copy work *3*, Parallel work *4*, and Collaborative work *5*. DDI (*red*) corresponds to distributed dyadic interaction, UDI (*blue*) to unequal dyadic interaction, NDI to no dyadic interaction. The *dashed areas* indicate impossible combinations (i.e. non-matches), which entered the analysis in the same category as NDI

This dyadic interaction is the result of matching either collaborative-collaborative or parallel-parallel behaviors. UDI indicates that only one child was contributing to the solution while the other child was not active by the matching of three individual behaviors: parallel-no work, parallel-passive and parallel-copy. NDI refers to all other matches of behaviors that do not result in DDI or UDI. The result of the cross-matching procedure for one particular dyad is shown in Fig. 11.1.

The research questions of the study addressed how the dyadic interactions developed over six sessions in terms of the relative strength of the dyadic attractor states DDI, UDI and NDI, and how this attractor dynamics was related to task performance. In the confinement of this example we will only present the basic results with respect to coordination based on the relative recurrence rates of DDI, UDI, and NDI. The full article, however, also includes analyzes of other CRQA-measures concerning the non-diagonal line structures in the CRP [24].

As said, the recurrence rate provides a basic measure for the coordination between subsystem. To scale the degree of coordination and analyze the relative strength of the three quasi-stable states a baseline is needed. This baseline is readily provided by the discrete topology of the state space, with the color codes based on behavioral matching overlaying the 25 cells, as shown in Fig. 11.2. This generates a priori probabilities for the relative recurrence rates *rRR*(*color*) as follows: When *color* is *red* (DDI) the baseline is 8 % (i.e. 2 of 25 cells), for *blue* (UDI) it is 24 % (i.e. 6 of 25 cells), for *white* (NDI and other non-matches) it is 68 % (i.e. 17 of 25 cells). If the empirical *rRR*(*color*), especially *rRR*(*red*) and *rRR*(*blue*), significantly differ from these baselines this indicates coordination in the dyadic interaction. That is to say, children's behavioral systems involved in the problem-solving task are governed by a shared dynamics, with a non-random, non-uniform coupling between subsystems.

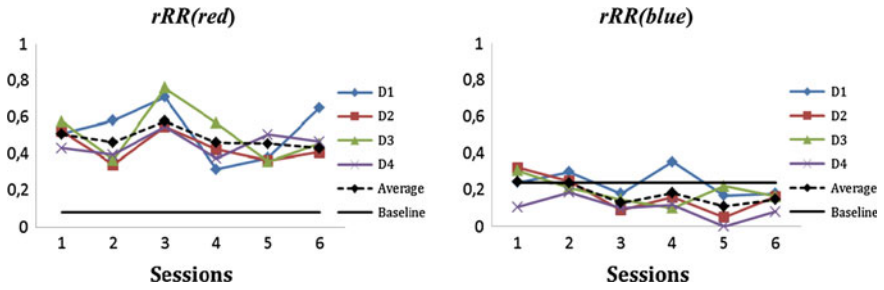

Fig. 11.3 Relative recurrence rates, *rRR*(*red*) and *rRR*(*blue*), for the two quasi-stable states DDI and UDI, respectively, for four dyads, over the six sessions

Figure 11.3 shows the *rRR*(*color*) of the two quasi-stable states DDI and UDI across all sessions, for four of the dyads in the study separately and averaged (dotted black line). It is clearly visible that *rRR*(*red*), referring to DDI, is above baseline (full black line) for the entire series of six sessions. In contrast, *rRR*(*blue*), referring to UDI, hovers around baseline (full black line). This makes DDI the predominant attractor state throughout the six observations. This basically implies that the distributed interactions remained more recurrent than the unequal interactions. This example briefly shows how Chromatic CRQA can reveal the coupling of both unequal and distributed 'matches' across all possible time scales and enables their comparison.

11.3 Anisotropic CRQA

11.3.1 Analysis of Anisotropic CRPs

As can be seen in Fig. 11.1, most colored dots in the CRP are not isolated but arrange to form rectangular and line structures. The "checkerboard texture" in Fig. 11.1 generally indicates recurrence domains which reveal metastability in the dynamics [25], an observation also made within the framework of symbolic dynamics [26–29]. In fact, the predominant formation of non-diagonal structures, as opposed to diagonal structures, is a typical feature of categorical time series, and is obviously related to the discrete topology of the phase space. (Note that this also appears when applying canonical recurrence, without matching and the color coding.) The rectangular and line structures provide information about the attractor dynamics of the interacting subsystems. They particularly inform about the strength and direction of the coupling between the two subsystems under investigation, as will be demonstrated later.

Thus far recurrence procedures have focused on diagonal lines with the measures *DET*, *MeanDiagLine*, *MaxDiagLine* and *ENT*, and vertical lines with measures *LAM*, *TT* and *MaxVertLine* ([23]; see below). However, as noted earlier, with

categorical time series dots mostly do not appear on diagonal lines but rather are part of non-diagonal structures. In addition, it is clearly visible in Fig. 11.1 that these non-diagonal line structures extend in two directions in the CRP, both horizontally as well as vertically. This directionality can potentially provide differential and complementary information about the coupling between the two subsystems represented by the time series along the two axes. Generally, line structures indicate instances where behavioral categories which are briefly expressed by one subsystem are accompanied by episodes of lingering in the matching behavioral category by the other subsystem. For instance, a vertical line structure in the CRP represents that a brief occurrence of some behavior in the time series along the horizontal axis is either followed (when above the Line-of-Synchrony) or preceded (when below the Line-of-Synchrony), with some delay, by a much longer occurrence of the matching behavior in the times series along the vertical axis.

Anisotropic CRQA, both in its cross-matching and cross-recurrence version, is founded on the general observation of anisotropy in the CRP, that is, on direction dependence along the two axes of the plot. Calculating the recurrence measures for the horizontal and vertical line structures separately, and allowing for their comparison, can easily be performed by first computing the transpose matrices, $\boldsymbol{CR}_{i,j}^{\mathrm{T}}(color)$, of the cross-recurrence matrices $\boldsymbol{CR}_{i,j}(color)$. Subsequently, the algorithm '*tt*' from Marwan's crp toolbox for Matlab (available at http://www.recurrence-plot.tk) can be applied to all $\boldsymbol{CR}_{i,j}(color)$ as well as all $\boldsymbol{CR}_{i,j}^{\mathrm{T}}(color)$. The '*tt*' function computes the distribution of the length of the vertical line structures in the recurrence plot. Based on these distributions of both the vertical and the horizontal line lengths, the corresponding direction-specific non-diagonal CRQA-measures can be calculated. This will be explained below for canonical cross-recurrence (or: cross-matching of a single kind) for the sake of simplicity.

The first measure derived from the non-diagonal line structures is *laminarity*, defined as the proportion of colored dots that are part of a vertical (LAM_V) or horizontal (LAM_H) line structure. Laminarity reflects the degree to which subsystems are 'trapped' into expressing matching behaviors for some period of time. LAM_V depicts how much the subsystem on the vertical axis constitutes larger structures, whereas LAM_H does so for the subsystem on the horizontal axis. Second, *trapping time* is the average length of either the vertical (TT_V) or horizontal (TT_H) line structures. *TT* is measured in units of time and estimates how long subsystems are, on average, lingering in a specific state. If TT_V is high, the system on the vertical axis tends to be trapped in relatively long periods of the same behaviors that are matched by the subsystem on the horizontal axis at some point, and for high TT_H the system on the horizontal axis tends to be trapped in relatively long periods of the same behaviors that are also expressed by the other subsystem at some point. Finally, *maximum line* also gives information about maximum duration of the non-diagonal line structures, with $MaxL_V$ being the length of the longest vertical line and $MaxL_H$ the length of the longest horizontal line. In other words, *MaxL* measures the duration of the longest behavioral matching for each subsystems. High $MaxL_V$ means that the vertical subsystem is lingering for a long period in a

single (matching) behavior, whereas $MaxL_H$ means that the horizontal subsystem lingers for a long period in a single (matching) behavior.

11.3.2 Application of Anisotropic CRQA: Asymmetric Gestures-Speech Attunement

In this example, the data do not concern dyadic interaction as in the previous examples, but two behavioral dimensions. More specifically we were concerned with the verbal and nonverbal problem-solving skills of children. This example study of Anisotropic CRQA addresses the coupled dynamics of children's gestures (i.e. manual kinesics and task manipulations) and speech as they construct cognitive understanding [30]. Contrary to the mainstream approach applied in most studies in this field, gestures were considered as in-the-moment actions, originating from an underlying complex dynamical system which is rooted in cognitive and perception-action subsystems. CRQA helps us to get a more detailed understanding of the role of gestures in learning and of their dynamic coordination with speech and cognitive processes.

The data of 12 children with a mean age of 5.4 years were analyzed. Children's expressed gestures and speech during a hands-on science task were scored in a common metric, which was based on their skill level as defined by dynamic skill theory [31]. Within this theory, a total of ten hierarchically ordered levels of cognitive skills (either actions or verbalizations) can be distinguished, divided over three tiers: sensorimotor, representational, and abstract. Anisotropic CRQA was performed on the (categorical) skill-level time series of children's gestures and speech. First the cross-matching as introduced above was applied: Skill levels of speech and gestures belonging to the same tier were considered behavioral matches of a specific kind. They received a designated color in the CRP, where blue was reserved for the sensorimotor tier, red for the representational tier, and green for the abstract tier. In Fig. 11.4 an example CRP is shown of the coupled gesture-speech subsystems. Note that in the following presentation of part of the results the color coding does not play a role. In other words, in the context of this example the data were not subjected to Chromatic CRQA for the sake of brevity (however see [30] for a more extensive account).

Results of the Anisotropic CRQA on the vertical and horizontal line structures in the CRP, for all 12 children in the study are shown in Table 11.1. Monte Carlo permutation tests[2] revealed that all vertical line measures were significantly higher

[2]These tests can readily be applied to small sample sizes or unequal groups of data, in this case to determine the statistical significance (p-value) of the differences between the line structures. The test repeatedly (e.g., 1000 times) compares the observed difference to the difference found within a range of simulated data drawn from the original sample. The p-value is computed by dividing the number of times the observed or a bigger difference occurs in these random samples by the number of drawn samples.

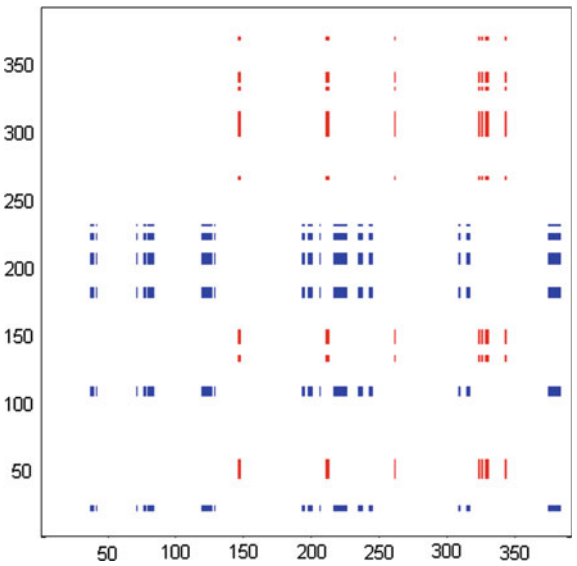

Fig. 11.4 Example CRP of the coupled gesture-speech subsystems for a single child, demonstrating the color coding for the three skill tiers used in the cross-matching procedure, with *blue* for sensorimotor, *red* for representational, and *green* for abstract (did not occur here). The child's speech is represented along the horizontal axis, gestures along the vertical axis

Table 11.1 Overview of *RR* and Anisotropic CRQA-measures (V = vertical lines; H = horizontal lines) of all 12 children in the study, not specified for color

Child	*RR*	LAM_V	LAM_H	TT_V	TT_H	$MaxL_V$	$MaxL_H$
1	0.013	0.99	0.91	5.2	3.4	21	7
2	0.019	1.00	0.89	6.4	3.8	19	10
3	0.004	0.97	0.69	4.3	2.6	12	3
4	0.011	1.00	0.89	7.4	5.1	26	11
5	0.002	0.89	0.90	3.2	3.1	5	6
6	0.010	0.96	0.70	6.6	2.6	16	5
7	0.009	0.98	0.92	5.8	4.0	18	12
8	0.006	0.97	0.62	4.8	2.8	12	5
9	0.025	0.99	0.92	6.3	5.1	15	15
10	0.016	1.00	0.79	6.0	5.5	25	27
11	0.002	0.96	0.63	5.4	2.7	18	3
12	0.018	1.00	0.79	8.3	3.6	24	6
Mean	0.011	0.98	0.81	5.8	3.7	17.6	9.2

than the horizontal line measures (all p-values < 0.01), with respectable effect sizes: $d(LAM_V > LAM_H) = 2.01$; $d(TT_V > TT_H) = 1.72$; $d(MaxL_V > MaxL_H) = 1.31$. This was not only true at the group level, but also for all children separately for *LAM* and *TT* and for 9 out of 12 children also for *MaxL*. This finding reveals an asymmetric dynamic attunement of gestures and speech, with gestures relatively more regularly and more rigidly displaying the same-tier skill level compared to speech. For the proposed meaning in terms of the difference in coupling between gestures and speech, we refer to the simulation study in Sect. 11.4.

11.4 Model Example of Anisotropic CRQA

In order to facilitate the interpretation of the Anisotropic CRQA-measures and to serve as a non-empirical illustration of Anisotropic CRQA on categorical time series, the following model of two time-discrete oscillators will briefly be explored:

$$\begin{aligned} y_1(n+1) &= \sin(\pi \cdot y_1(n)) + c_1 \sin(\pi \cdot y_2(n)), \\ y_2(n+1) &= \sin(\pi \cdot y_2(n)) + c_2 \sin(\pi \cdot y_1(n)). \end{aligned} \tag{3}$$

The model consists of a system of two coupled difference equations with coupling parameters c_1 and c_2, which represent the coupling strength between the oscillators y_1 and y_2.

The behavior of the model will be assessed in a coarse grained fashion to realistically simulate data as it is commonly collected in the behavioral sciences. Specifically, two dichotomous time series, x_1 and x_2, will be constructed by applying the following rule[3]:

$$x_i(n) = round\left[\frac{y_i(n) - \min(y_i(n))}{max(y_i(n)) - min(y_i(n))}\right], i \in \{1, 2\}. \tag{4}$$

Although there is some resemblance with a Kuramoto model [32, 33],[4] this is by no means the incentive of the present simulations. However, a potential connection to human behavior would be, for instance, and without any claim of validity, two people's gazes during a conversation. The model articulates the (dichotomized) dynamics of the weakly coupled oculomotor and cognitive systems of the two individuals, in terms of a pair of nonlinear coupled oscillators. Variable y_1 represents one person's eye movements over time on some continuous spatial dimension and variable y_2 that of the other person. When transcribed at the behavioral level in terms of the distinct categories (1) looking at the other person's face and (2) looking away from it, which is not uncommon in the behavioral sciences (e.g. [9, 12, 13]; see also [7, 8]), the resulting time series x_1 and x_2 are dichotomous.

Simulations will focus on the relation between relative coupling strength in the model and the Anisotropic CRQA-measures, that is, the measures of vertical and horizontal line structures in the CRP. The CRP will be based on straightforward cross-recurrence of the discrete-time dichotomized time series x_1 and x_2.

[3]This basically constructs a symbolic dynamics by introducing a static encoding of the time series (see e.g. [26, 27]).

[4]In fact written in this way the system models a system of three asymmetrically coupled Kuramoto oscillators under a specific set of parameter settings, among which, in particular equal intrinsic frequency.

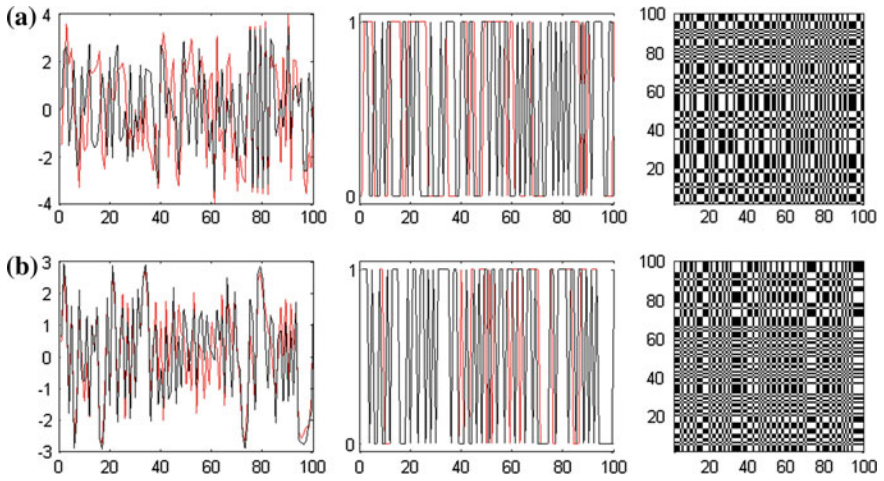

Fig. 11.5 Time series and CRPs for **a** $c_1 = 3.0$ and $c_2 = 2.5$, with resulting $RR = 0.51$, *vertical line* measures $LAM_V = 0.73$, $TT_V = 3.88$, $MaxL_V = 9$, *horizontal line* measures $LAM_H = 0.62$, $TT_H = 2.38$, $MaxL_H = 4$, and **b** $c_1 = 1.8$ and $c_2 = 2.0$, with resulting $RR = 0.50$, *vertical line* measures $LAM_V = 0.73$, $TT_V = 3.46$, $MaxL_V = 6$, *horizontal line* measures $LAM_H = 0.65$, $TT_H = 2.82$, $MaxL_H = 6$

Model simulations for both the continuous and coarse-grained variables and the related CRPs of the dichotomous system are given in Fig. 11.5, for two parameters settings, over 100 time steps, after transient. The parameter settings are drawn from a 'realistic' area of the parameters space (c_1, c_2). The impression of being realistic is based on a visual comparison of time series and CRPs of the model outcomes and the empirical data encountered in previous studies (esp. [12, 13]). We admit that this reflects no robust mathematical or other kind of formal criterion at this point. However, we would like to stress once more that the model is not intended as a realist model of some type of behavior. It merely serves as a demonstration of Anisotropic CRQA and to provide some directions for interpretation of anisotropy in CRPs.

Figure 11.6 shows how the relative difference in coupling strength, $(c1 - c2)/(c1 + c2)$, is related to the relative difference in vertical and horizontal line measures *LAM*, *TT* and *MaxL*, in two separate and realistic parts of the parameters space. These graphs reveal a strong association between coupling and anisotropy, for the Anisotropic CRQA-measures *LAM* and *TT* but not for *MaxL*, in both parameter ranges.

Results like these remains difficult to interpret however, since the relation between coupling and synchronization in complex dynamical systems is not straightforward and might rely heavily on the specifics of the chosen model and parameter range (cf. [23]).

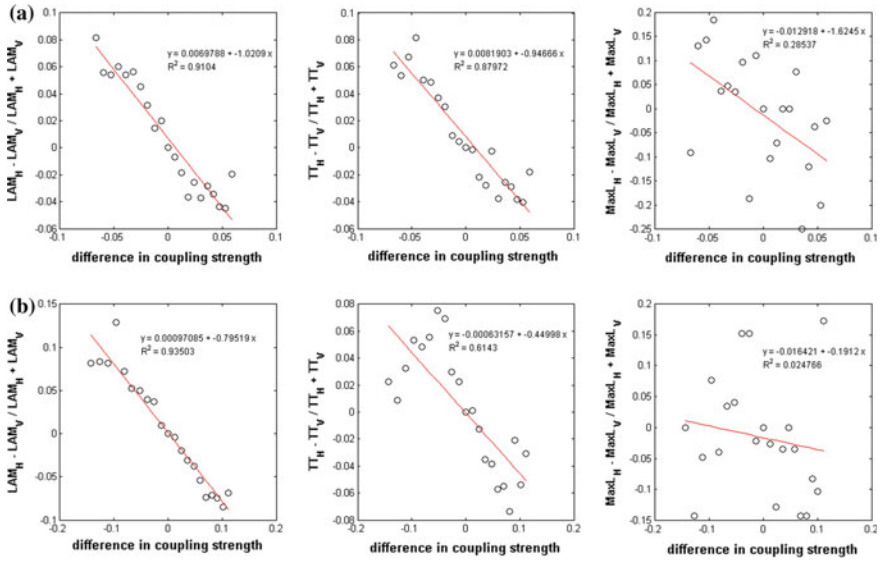

Fig. 11.6 Relative difference in coupling strength $(c_1 - c_2)/(c_1 + c_2)$, for **a** $c_1 = 2$ and $c_2 \in \{1.5 : 0.05 : 2.5\}$ and **b** $c_1 = 4$ and $c_2 \in \{3.5 : 0.05 : 4.5\}$, on the horizontal axis, and its relation to the relative difference in *vertical* and *horizontal line* measures *LAM*, *TT* and *MaxL*, on the vertical axis

11.5 Discussion

In the present chapter we tried to extend the existing recurrence procedures for the study of coupled dynamical systems which are measured on a nominal scale, by introducing two novel methodological variations: Chromatic CRQA, using cross-matching, and Anisotropic CRQA, based on anisotropy in the CRP. These advancements facilitate the study of a wider class of phenomena in the behavioral sciences and the detection of asymmetry in the dynamic organization of interpersonal behavior, respectively.

Several issues related to these procedures need to be developed further, in order for them to reach their full potential in the behavioral sciences. We will touch upon a few of these open-ends only briefly below. First, and foremost, the formal description as presented in this chapter should be extended and improved, with a focus on mathematical rigour and methodological completeness. This will allow the proposed methodological variations to be embedded in the larger framework of recurrence based methods (see [23]).

Second, applying Chromatic CRQA and Anisotropic CRQA, especially when done conjointly, results in substantial increase in the amount of measures to be considered, compared to regular CRQA. Specifically, there is a new set of measures for each added behavioral match (i.e. color) and the number of measures almost doubles with the anisotropic analysis. However, the possibility of directly comparing

the (relative) measures pertaining to the different kinds of matches in Chromatic CRQA reduces this problem to a large extent. Also, as already demonstrated, for Anisotropic CRQA, a focus on the (relative) differences of the anisotropic measures is probably the best way to express and detect asymmetries in the dynamics. The modeling approach initiated in Sect. 11.4 might be pursued further to find out which combinations of measures are most informative in specific contexts.

Third, subsequent efforts in the development of these procedures should look into the overlap between the anisotropic measures of rectangular structures. Since recurrent points within a rectangular structure are always part of both a horizontal line and vertical line at the same time, they 'count' for the measures in both directions. This is not the case for isolated recurrent points, or for those that form a line. Distinguishing between rectangular and (actual) line structures might provide a more reliable estimate of the anisotropy in the CRP and of the coupling between subsystems. Another promising way to deal with this is to quantify the rectangular structures directly, for instance in terms of their 'area' (i.e. horizontal length × vertical length), 'asymmetry' (i.e. horizontal length divided by vertical length), number, entropy etcetera (cf. [34]).

Fourth, and finally, the application of the behavioral matching procedure and chromatic analysis to a single system measured on a nominal scale might be interesting to pursue as well. Chromatic RQA on one categorical time series would give additional and unique information about the structural connections between different behavioral modes of a system. It might even be possible to (in a way) reverse the technique, and use recurrence measures to detect which behavioral matches are most dynamically stable in the behavior. Finally, it is not unlikely that even Anisotropic RQA and Anisotropic CRQA of continuous times series might find an application within a specific research context.

Acknowledgment The first author wishes to express his appreciation to Fred Hasselman, for the many fruitful discussions arising from shared encounters with the methodological issues regarding recurrence analysis on categorical behavioral time series.

References

1. N. Marwan, M. Thiel, N.R. Nowaczyk, Cross recurrence plot based synchronization of time series. Nonlinear Proc. Geophys. **9**, 325–331 (2002)
2. K. Shockley, M. Butwill, J.P. Zbilut, C.L. Webber, Cross recurrence quantification of coupled oscillators. Phys. Lett. **305**, 59–69 (2002)
3. J.P. Zbilut, A. Giuliani, C.L. Webber Jr, Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification. Phys. Lett. A **246**, 122–128 (1998)
4. J.P. Zbilut, C.L. Webber Jr, Embeddings and delays as derived from quantification of recurrence plots. Phys. Lett. A **171**, 199–203 (1992)
5. R.F.A. Cox, M. Van Dijk, Micro-development in parent-child conversations: from global changes to flexibility. Ecol. Psychol. **25**(3), 304–315 (2013)
6. R. Dale, M.J. Spivey, Unraveling the dyad: using recurrence analysis to explore patterns of syntactic coordination between children and caregivers in conversation. Lang. Learn. **56**, 391–430 (2006)

7. D.C. Richardson, R. Dale, Looking to understand: the coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. Cogn. Sci. **29**, 1045–1060 (2005)
8. D.C. Richardson, R. Dale, N.Z. Kirkham, The art of conversation is coordination. Psychol. Sci. **18**, 407–413 (2007)
9. J.A. De Graag, R.F.A. Cox, F. Hasselman, J. Jansen, C. De Weerth, Functioning within a relationship: mother-infant synchrony and infant sleep. Infant Behav. Dev. **35**, 252–263 (2012)
10. C. Yu, L.B. Smith, Joint attention without gaze following: human infants and their parents coordinate visual attention to objects through eye-hand coordination. PLoS ONE **8**(11), e79659 (2013)
11. M.M. Louwerse, R. Dale, E.G. Bard, P. Jeuniaux, Behavior matching in multimodal communication is synchronized. Cogn. Sci. **36**(8), 1404–1426 (2012)
12. E. Reuzel, P.J.C.M. Embregts, A.M.T. Bosman, R.F.A. Cox, M. Van Nieuwenhuijzen, A. Jahoda, Conversational synchronization in naturally occurring settings: a recurrence-based analysis of gaze directions and speech rhythms of staff and clients with intellectual disability. J. Nonverbal Behav. **37**, 281–305 (2013)
13. E. Reuzel, P.J.C.M. Embregts, A.M.T. Bosman, R.F.A. Cox, M. Van Nieuwenhuijzen, A. Jahoda, Verbal interactional dominance and coordinative structure of speech rhythms of staff and clients with an intellectual disability. Nonlinear Dyn. Psychol. Life Sci. **18**(4), 371–396 (2014)
14. A. Lichtwarck-Aschoff, F. Hasselman, R.F.A. Cox, A. Pepler, I. Granic, A characteristicdestabilization profile in parent–child interactions associated with treatment efficacy for aggressive children. Nonlinear Dyn. Psychol. Life Sci. **16**, 353–379 (2012)
15. R. Dale, A.S. Warlaumont, D.C. Richardson, Nominal cross recurrence as a generalized lag sequential analysis for behavioral streams. Int. J. Bifurcat. Chaos **21**, 1153–1161 (2011)
16. R. Fusaroli, I. Konvalinka, S. Wallot, *Analyzing Social Interactions: The Promises and Challenges of Using Cross Recurrence Quantification Analysis*, eds. by N. Marwan et al. Translational Recurrences, Springer Proceedings in Mathematics & Statistics, vol. 103 (2014)
17. B. Radstaak *Fighting AN. A nonlinear case-exploration of a client-therapist text message interaction.* Master thesis, Radboud University, Nijmegen, The Netherlands (2012). http://www.annabosman.eu/
18. E. Reuzel, P.J.C.M. Embregts, A.M.T. Bosman, M. Van Nieuwenhuijzen, *Verbal interactional dominance and linguistic coupling of staff and a client with an intellectual disability during WhatsApp conversations.* Manuscript submitted for publication (2015)
19. H.M. Endedijk, A.H.N. Cillessen, R.F.A. Cox, H. Bekkering, S. Hunnius, The role of child characteristics and peer experiences in the development of peer cooperation. Soc. Dev. **24**(3), 521–540 (2015)
20. K.L. Marsh, M.J. Richardson, R.C. Schmidt, Social connection through joint action and interpersonal coordination. Top. Cogn. Sci. **1**, 320–339 (2009)
21. N. Sebanz, H. Bekkering, G. Knoblich, Joint action: Bodies and minds moving together. Trends Cogn. Sci. 10, 70–76 (2006). K. Shockley, M. Butwill, J.P. Zbilut, C.L. Webber Jr. Cross recurrence quantification of coupled oscillators. Phys. Lett. A, 305, 59–69 (2002)
22. F. Orsucci, A. Giuliani, C.L. Webber Jr, J.P. Zbilut, P. Fonagy, M. Mazza, Combinatorics and synchronization in natural semiotics. Phys. A **361**, 665–676 (2006)
23. N. Marwan, M. Carmen Romano, M. Thiel, J. Kurths, Recurrence plots for the analysis of complex systems. Phys. Rep. **438**, 237–329 (2007)
24. M. Guevara, M. Van Dijk, R.F.A. Cox, P.L.C. Van Geert, *On the attractor dynamics of dyadic interaction: a recurrent based analysis.* Manuscript under review in Nonlinear Dynamics, Psychology, and Life Sciences (2015)
25. J.P. Eckmann, S.O. Kamphorst, D. Ruelle, Recurrence plots of dynamical systems. Europhys. Lett. **5**, 973–977 (1987)
26. P. Beim Graben, A. Hutt, Detecting recurrence domains of dynamical systems by symbolic dynamics. Phys. Rev. Lett. **110**, 154101 (2013)

27. P. Beim Graben, A. Hutt, Detecting event-related recurrences by symbolic analysis: applications to human language processing. Proc. Royal Soc. Lond. **A373**, 20140089 (2015)
28. R. Donner, U. Hinrichs, B. Scholz-Reiter, Symbolic recurrence plots: a new quantitative framework for performance analysis of manufacturing networks. Eur. Phys. J. Spec. Top. **164**, 85–104 (2008)
29. P. Faure, A. Lesne, Recurrence plots for symbolic sequences. Int. J. Bifurcat. Chaos **20**, 1731–1749 (2010)
30. L. De Jonge-Hoekstra, S. Van der Steen, P.L.C. Van Geert, R.F.A. Cox, *Asymmetric dynamic attunement of speech and gestures in the construction of children's understanding*. Manuscript under review in Frontiers in Psychology (2015)
31. K.W. Fischer, T.R. Bidell, *Dynamic Development of Action, Thought and Emotion*, eds. by W. Damon, R.M. Lerner. Theoretical Models of Human Development. Handbook of child psychology, 6th edn. (Wiley, New York, 2006)
32. J.A. Acebrón, L.L. Bonilla, C.J. Pérez Vicente, F. Ritort, R. Spigler, The kuramoto model: a simple paradigm for synchronization phenomena. Rev. Mod. Phys. **77**, 137–185 (2005)
33. M. Aguilera, X.E. Barandiaran, M.G. Bedia, F. Seron, Self-organized criticality, plasticity and sensorimotor coupling. explorations with a neurorobotic model in a behavioral preference task. PLoS ONE **10**(2), e0117465 (2015)
34. T. Xu, C. Yu, *Quantifying Coupled Behavior in Social Interaction using Cross-Recurrence Quantification Analysis.* Poster presented at the 5th Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics, Providence, USA (2015)

Chapter 12
Using Cross-Recurrence Quantification Analysis to Understand Social Motor Coordination in Children with Autism Spectrum Disorder

Veronica Romero, Paula Fitzpatrick, R.C. Schmidt and Michael J. Richardson

Abstract Interpersonal motor coordination is considered to be an integral part of maintaining successful social interactions. Research has shown that simply coordinating one's movements with another actor can influence rapport as well as feelings of social connection and social competence. Past research has also found that deficits in social motor coordination are associated with psychological dysfunction such as schizophrenia and borderline personality disorders. However, the potential association between interpersonal motor coordination and autism spectrum disorder (ASD) has only received a very limited amount of attention. In the current experiment, children who had been previously diagnosed with ASD and typically developing (TD) children were asked to synchronize with or imitate the movements of an experimenter in two different interpersonal motor tasks: object tapping and hand-clapping. Both the experimenter's and the participants' movements were captured and compared to each other using relative phase analysis and cross-recurrence quantification analysis (CRQA). The results reveal differences not only in the patterning of the coordination that occurred for ASD and TD children, but also in the stability and deterministic structure of the coordination. Of particular interest was the finding that children with ASD exhibited less stable (robust), but more deterministic patterns of interpersonal social motor coordination compared to TD children.

V. Romero (✉) · M.J. Richardson
University of Cincinnati, Cincinnati, USA
e-mail: romerovc@mail.uc.edu

M.J. Richardson
e-mail: richamo@ucmail.uc.edu

P. Fitzpatrick
Assumption College, Worcester, USA
e-mail: pfitzpat@assumption.edu

R.C. Schmidt
College of the Holy Cross, Worcester, USA
e-mail: rschmidt@holycross.edu

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8_12

12.1 Introduction

Children with autism spectrum disorder (ASD) exhibit numerous impairments that can severely impede learning and overall social functioning at home and in school, as well as make successful treatment difficult. Even though the dynamics of interpersonal motor coordination is considered to be an integral part of maintaining successful social interactions [1, 2], this dimension of behavioral order remains overlooked within the field of ASD research. This is despite the fact that there is now a well-established body of research demonstrating how intentionally or unintentionally coordinating one's movements with those of another individual can positively influence rapport, feelings of social connection, and feelings of social competence [2–5]. Previous research has also found that deficits in social motor coordination are associated with other social psychological disorders such as schizophrenia [6, 7] and borderline personality disorders [8]. It is therefore important to understand the degree to which social motor coordination (or the lack thereof) is related to social competency in children diagnosed with ASD.

With regard to the previous research examining social motor impairments in ASD, the general findings are often limited in scope and sometimes inconsistent and contradictory [9]. The limited scope and contradictory nature of this previous work is due, in part, to the fact that empirical, clinical, and observational research is often treated in isolation, without any attempt to formalize an integrative methodological approach for identifying and understanding the core ASD deficits in social motor coordination [9]. Thus, the social motor impairments that characterize children with ASD and the degree to which these impartments relate to ASD severity are largely unknown. Moreover, many researchers have simply focused their inquiries on whether children with ASD are or are not capable of carrying out certain social motor tasks, without examining *how* they perform these tasks over time—i.e., without examining the patterns or dynamics of the social motor behavior that occurs within a social context [9].

Of particular relevance for the current study, is that the few studies that have investigated the dynamics of social motor coordination in children with ASD have highlighted how these deficits might not be due to an absence of coordination, but rather to a difference in the patterning compared to typically developing (TD) children [1, 9, 10]. For example, Marsh and colleagues [1] investigated the interpersonal movement coordination that occurred between children and their parents sitting and rocking side-by-side in rocking chairs. The results revealed that both ASD and TD children could coordinate with the movements of their parents, but that the overall patterning (the relative phase relationship) that characterized the coordination for ASD and TD children was different, with TD children more spontaneously and reciprocally modifying their movements with respect to their parents than children with ASD. The goal of the current study was to further explore the dynamics of social motor coordination in children with ASD with the use of non-linear, as well as linear time-series methods.

Cross-Recurrence Quantification Analysis (CRQA) is a non-linear analysis method that quantifies the dynamic (time-evolving) similarity between two behavioral time- or event-series by identifying whether behavioral states/events in the two series reoccur over time. The benefits of this analysis method over other time-series methods (e.g., cross-correlation, cross-spectral analysis) is that it does not require a researcher to make any a priori assumptions about the structure or stationarity of the data analyzed. Moreover, this method can be employed to uncover recurrent structure between the behaviors of two individuals regardless of whether the behavior in question exists within a single state space or in a more abstract multi-dimensional (multivariate) state space. Defined at a more intuitive level, the method assesses whether the points (states) in a behavioral series visit the same states over time and then quantifies the dynamic patterns of these time-evolving recurrences using a range of statistics. Common statistics include: *Percent Recurrence* (%REC), which measures the percentage of recurrent states between the two time-series and is an index of the degree to which the two behavioral signals are confined to a similar region of state or phase space; *Percent Determinism* (%DET), which measures the percentage of recurrent states that form sequential lines of recurrent points (i.e., diagonal lines in a recurrence plot) and quantifies the degree to which the co-varying structure of two behavioral time-series is the result of deterministic vs. random processes; and *MaxLine,* which extracts the longest sequence of recurrent states (longest diagonal line in a recurrence plot), and provides a proportional measure of the stability of the coordination present between two behavioral time-series.

CRQA has previously been used to study the patterning and stability of coordination between people in social situations. For instance, Richardson and colleagues [11] demonstrated through the use of CRQA that the coordination that arises between the rhythmic limb movements of two people follows the same dynamics as a person coordinating her two arms intrapersonally. Furthermore, they were able to demonstrate that the increased variability in coordination observed in social situations was not due to an increase in movement noise (indexed with the use of %REC), but rather was due to the visual coupling of the interpersonal situation being significantly weaker than the biomechanical coupling that characterizes intrapersonal interlimb coordination (as indexed by MaxLine). In a more realistic, full-body conversational task, Shockley and colleagues [12] found that pairs of participants having a conversation in order to solve a puzzle showed higher amounts of postural coordination (greater %REC; greater recurrent postural activity) than when participants were conversing with a confederate. Additionally, they found that participants' postural trajectories were more stable (as indexed by MaxLine) when talking with each other, compared to when they were conversing with a confederate.

Accordingly, the aim of the current study was to employ CRQA to investigate the potential differences in the stability and patterning of the interpersonal coordination that occurred for children diagnosed with ASD compared to TD children. To achieve this aim we investigated two simple coordination tasks (i.e., an object tapping and hand clapping tasks) performed between an experimenter and ASD or

Table 12.1 Demographic information for all participants

Group	Age (in years)					Gender	
	6	7	8	9	10	Male	Female
ASD	10	7	7	10	7	34	7
TD	10	11	8	9	9	36	11

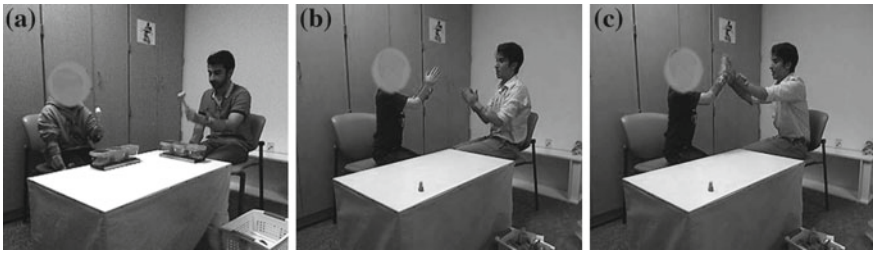

Fig. 12.1 Room set-up for the **a** tapping task and **b, c** interpersonal hand clapping game

TD children. Once this aim is achieved, it would be feasible to use our new-found understanding to extract the pertinent CRQA measures derived from an interpersonal task during the diagnostic process and include these findings as part of the criteria when deciding whether to apply the ASD diagnosis. Furthermore, these measures could instigate new intervention models focused on interpersonal coordination for children and adults already diagnosed with ASD.

12.2 Method

Forty three children previously diagnosed with ASD (and with a confirmed diagnosis made by a research certified clinician administering the Autism Diagnostic Observation Schedule second edition (ADOS-2)) who were classified as high functioning, and 47 typically developing children participated in the study.[1] Participants were between the ages of 6 and 10 years ($M = 7.92$, $SD = 1.45$ years, see Table 12.1 for age and gender distribution by group).

The study was conducted in a 10 by 12 foot laboratory room at Cincinnati Children's Hospital Medical Center (University of Cincinnati, Cincinnati, OH). Children came into the laboratory room and were asked to sit at a 2 foot wide × 4 foot long × 2 foot high table next to the seated experimenter (see Fig. 12.1). Four Polhemus Latus receptors were attached to the underside of the table top, one in each corner, to create a 10 × 12 × 8 foot capture volume around the table. As soon as the child was seated, the four Polhemus Liberty Latus wireless markers/sensors

[1]Due to measurement and experimenter error, data for two ASD participants in the object tapping task and four TD participants in the interpersonal hand clapping game was missing.

were placed in wristbands and slipped over the child's and experimenter's wrists (one marker on each wrist of the child and experimenter). The motion of the Polhemus sensors was recorded at 94 Hz on a Dell PC computer using a custom software application written by the authors using the Polhemus Latus C/C ++ SDK Library.

The data presented here was part of a bigger project, in which participants performed a large range of motor, social and cognitive tasks. Here, we selected two social motor coordination tasks that were performed by all of the children. The first coordination task was a sequence of *tapping movements*, which involved children using a stick to tap/hit three drum-like cylinders from left to right (see Fig. 12.1a). Children repeated this left-to-right drumming sequence six separate times with the experimenter showing the sequence previous to every child in a continuous manner. Participants were randomly assigned to either an imitation or synchrony condition for this task, in the ASD group 20 children were in the imitation condition, while 21 were in the synchrony, as for the TD group, 26 synchronized while 21 imitated. In the imitation condition, the child was asked to repeat the tapping pattern demonstrated by the experimenter. In the synchrony condition, children were asked to repeat the tapping pattern *in time* with the experimenter after every sequence. The second task was an *interpersonal hand clapping game* (pat-a-cake), in which children completed a simple repetitive sequence of clapping their hands together and then with the experimenter (see Fig. 12.1b, c). The hand clapping game was completed twice, with each sequence involving around 15 consecutive intrapersonal and interpersonal clapping movements.

12.2.1 Data Analysis

The *x*-plane (left-right), *y*-plane (forward-back) and *z*-plane (up-down) positional coordinates of the sensors placed on the wrists of the experimenter and child were recorded for each task. To best determine the stability and pattering of the behavioral coordination that occurred between the child and experimenter, we first isolated the primary plane of motion for each task. Since the primary plane of motion for the tapping task was in the left-right (sagittal) plane, the x-*plane* movement time series was used to assess the behavioral coordination that occurred for this task (for a sample time series see Fig. 12.2a). For the interpersonal hand clapping game, the largest amplitude of movement was in the up-down (transverse), *z-plane*, with the intrapersonal clapping events occurring at a lower height than the interpersonal clap events (see Fig. 12.2b). Accordingly, this plane of motion was employed to assess the behavioral coordination that occurred for this task.[2]

[2]An analysis of secondary planes of motion produced results that were consistent with those reported here.

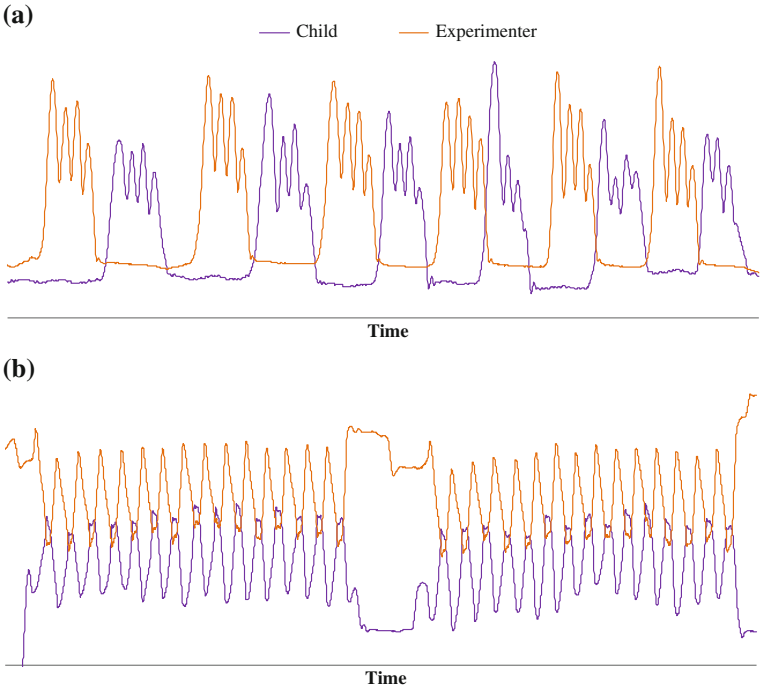

Fig. 12.2 Sample movement time series (raw unfiltered data) from a child in the **a** object tapping task and the **b** interpersonal hand clapping game

For the tapping task, we then performed a coordination analysis as well as a cross-recurrence quantification analysis using the primary plane of motion time series of the experimenter's right forearm (the experimenter always used his right hand/arm for all the tasks) and the primary plane of motion time series of the forearm used by the child for analysis. Note that for this task the child was free to use either left or right arm/hand. Although both arms/hands were employed by the experimenter and child for the hand clapping game, we only report the analysis of the right forearm movements of the experimenter and child because the coordination that occurred between the left forearm movements was completely redundant with the right.

Prior to analysis all of the pre- and post-non-task relevant movement transient periods were cropped from the different time series such as the period from the start of the recording to the start of the task and the periods in the synchrony condition where the experimenter repeated the sequence every time, otherwise all trials were included in the analysis. The time series were allowed to vary in length across participants. The resulting time series lengths in the tapping task were (ASD: Mean (*M*) number of data points = 3286.82, Standard Deviation (*SD*) = 761.30; TD:

$M = 3087.46$, $SD = 674.12$) and in the hand clapping game (ASD: $M = 4774.47$, $SD = 491.03$; TD $M = 4194.15$, $SD = 439.01$). These final motion time series were then low-passed filtered using 10 Hz 4th order Butterworth filter to remove measurement noise.

To determine the patterning and stability of the social motor coordination that occurred between the children and the experimenter for each task and condition, two measures of interpersonal coordination were employed: distribution of relative phase (see [13] for a review of studies that have employed this measure) and CRQA.

12.2.1.1 Distribution of Relative Phase Angles (DRP)

This measure evaluated the concentration of relative phase angles between the movements of the child and experimenter (i.e., the relative space-time angular location of the movements of the child and experimenter) across nine 20° regions of relative phase (0–20°, 21–40°, 41–60°, 61–80°, 81–100°, 101–120°, 121–140°, 141–160°, 161–180°). To determine these distributions, we computed the continuous relative phase of the two time series between −180° and 180° using the Hilbert transform [14]. We then computed the percentage of occurrence of the absolute value of the relative phase angles across the nine 20° relative phase regions from 0° to 180°. Previous research has demonstrated that stable social motor coordination is characterized by a concentration of relative phase angles in the portions of the distribution near 0° and 180° (e.g., [9, 15, 16]), which corresponds to inphase and antiphase patterns of coordination, respectively.

DRP was analyzed using a 9 (phase region: 0°, 30°, 50°, 70°, 90°, 110°, 130°, 150°, 180°) × 2 (diagnosis: ASD, TD) × 2 (condition: imitation, synchrony) mixed ANOVA for the object tapping task with phase region as the repeated measures factor (a Greenhouse-Geisser correction was employed where necessary). For the interpersonal hand clapping game, DRP was analyzed using a 9 (phase region) × 2 (diagnosis) mixed ANOVA. Of particular interest was the difference between ASD and TD children in the magnitude of inphase coordination (percentage of time in 0° bin) for synchronous object-tapping and hand-clapping tasks and in the magnitude of antiphase coordination (percentage of time in the 180° bin) for the imitation object-tapping task. Therefore, when a phase region by diagnosis interaction was found to be significant, planned *t*-tests were employed to compare these relative phase regions as a function of diagnosis.

12.2.1.2 Cross Recurrence Quantification Analysis (CRQA)

This measure evaluated the time-evolving phase-space structure of coordination that arose between the experimenter and the child when performing the two coordinative tasks. When analyzing the object tapping task data, an embedding dimension of 7, lag of 100, and radius of 15 % was used. For the interpersonal hand

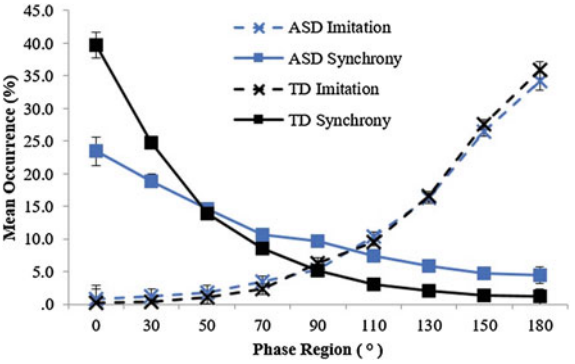

Fig. 12.3 Mean DRP observed in the object tapping task by phase region, diagnosis and condition. The error bars represent the standard error of the mean

clapping game, an embedding dimension of 5, lag of 45 and radius of 10 % was used.[3] These differences in parameter settings were due to the different pace and constraints of each experimental task. Since the goal of the study was to investigate the differences that might arise between children previously diagnosed with ASD and those who are typically developing, the parameters were kept constant for every child in each task. The CRQA measures %REC, %DET, and MaxLine were then analyzed using separate 2 (diagnosis: ASD, TD) × 2 (condition: imitation, synchrony) between subjects ANOVAs for the object tapping task. For the interpersonal hand clapping game, %REC, %DET and MaxLine were analyzed using separate independent samples *t*-tests.

12.3 Results

12.3.1 Object Tapping Task

For the object tapping task the DRP analysis revealed a main effect of phase region ($F(8,672) = 62.62$, $p < 0.01$, $\eta_p^2 = 0.43$), a two-way interaction between phase region and diagnosis ($F(8,672) = 8.07$, $p < 0.01$, $\eta_p^2 = 0.09$) a two-way interaction between phase region and condition ($F(8,672) = 316.52$, $p < 0.01$, $\eta_p^2 = 0.79$) and, most importantly, a three-way interaction between phase region, diagnosis and condition ($F(8,672) = 10.10$, $p < 0.01$, $\eta_p^2 = 0.11$; see Fig. 12.3). Planned t-tests revealed a lack of difference between the ASD and TD at the 180° phase region ($t(39) = -0.68$, $p = 0.50$) when the children were asked to imitate the experimenter. However, when asked to synchronize with the experimenter, the children in the TD group showed significantly larger occurrence of 0° phase relationship ($M = 39.71$, $SD = 15.53$) than those in the ASD group ($M = 23.44$, $SD = 9.71$; $t(45) = -4.18$, $p < 0.01$).

[3]The embedding dimension was chosen through the use of the false nearest neighbor (FNN) methodology while the lag was determined with the use of average mutual information (AMI).

Collectively, these DRP results would seem to indicate that children diagnosed with ASD were just as capable of imitating the experimenter's tapping movements as TD children. In contrast, however, the stability of the synchronous coordination for children with ASD was significantly weaker compared to TD children, which suggests that the synchronous motor coordination that occurred between the children with ASD and the experimenter was either less robust or more variable than the coordination that occurred between the experimenter and TD children.

The analyses of the CRQA measures revealed no significant main effects of diagnosis (all $ps > 0.19$). However, there was a significant main effect for condition in all measures (%REC: $F(1, 84) = 90.03$, $p < 0.01$, $\eta_p^2 = 0.52$; MaxLine: $F(1, 84) = 36.53$, $p < 0.01$, $\eta_p^2 = 0.30$; and %DET: $F(1, 84) = 22.08$, $p < 0.01$, $\eta_p^2 = 0.21$), such that children who were in the imitation group showed higher %REC ($M = 1.97$, $SD = 1.06$), %DET ($M = 98.85$, $SD = 1.17$) and MaxLine ($M = 112.68$, $SD = 97.71$) than those in the synchrony group (%REC: $M = 0.23$, $SD = 0.68$; %DET: $M = 92.20$, $SD = 8.82$; MaxLine: $M = 23.32$, $SD = 18.79$). Finally, there was a significant interaction between diagnosis and condition for %REC ($F(1, 84) = 8.36$, $p = 0.01$, $\eta_p^2 = 0.09$; for both %DET and MaxLine, $ps > 0.20$; see Fig. 12.4). Simple effects analyses

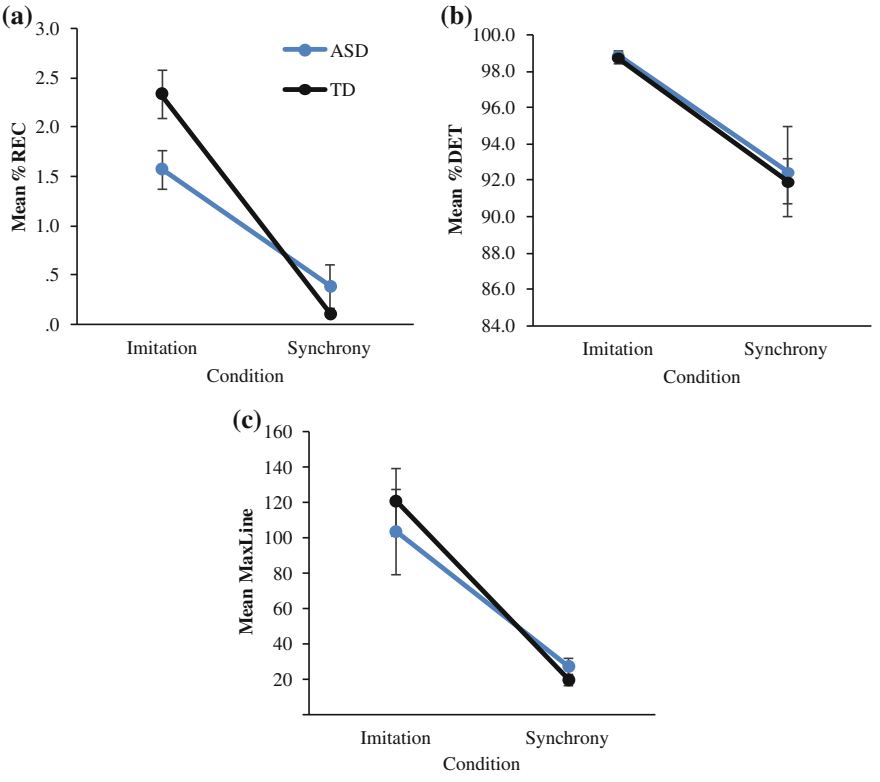

Fig. 12.4 Mean **a** %REC, **b** %DET and **c** MaxLine by diagnosis and condition for the object tapping task. The error bars show the standard error of the mean

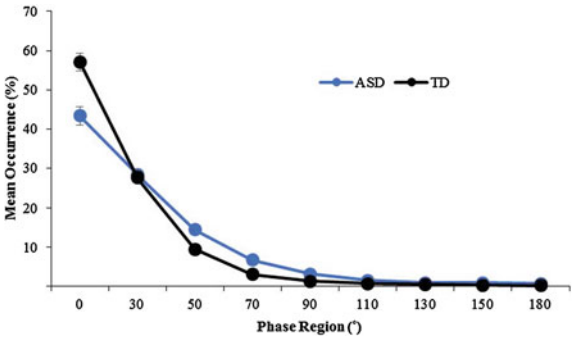

Fig. 12.5 Mean distribution of relative phase angles (DRP) by diagnosis. The error bars represent standard error of the mean

for %REC showed that in the imitation group, TD children had significantly more recurrent activity ($M = 2.34$, $SD = 1.10$) than the children with ASD ($M = 1.58$, $SD = 0.89$, $t(39) = -2.44$, $p = 0.02$), while there was no significant difference in the synchrony group ($t(45) = 1.42$, $p = 0.16$). Accordingly, the results for CRQA with respect to diagnosis are somewhat opposite the results observed for DRP above, in that they appear to indicate that the similarity between the child and experimenter phase-space trajectories is similar for ASD and TD children during synchronous performance, but that this is not the case during imitative performance. That is, for the imitation condition there was less recurrent activity for ASD compared to TD children.

12.3.2 Interpersonal Hand Clapping Game

Consistent with the synchrony results for the object tapping task, the DRP analysis for the synchronous hand-clapping task revealed a significant main effect of phase region ($F(8, 656) = 614.73$, $p < 0.01$, $\eta_p^2 = 0.88$) and a significant phase region by diagnosis interaction ($F(1, 656) = 15.12$, $p < 0.01$, $\eta_p^2 = 0.16$; see Fig. 12.5). Again, the planned t-test revealed that children with ASD spent significantly less time in the 0° phase region ($M = 43.37$, $SD = 15.28$) than those in the TD group ($M = 57.10$, $SD = 14.31$; $t(82) = -4.25$, $p < 0.01$), indicating that the coordination for the ASD children was less stable or more variable compared to TD children.

The analyses of the CRQA measures showed no significant differences in %REC between the ASD and TD children ($t(84) = 0.18$, $p = 0.86$). Surprisingly, however, there was a significant difference between the ASD and TD groups for MaxLine ($t(84) = 1.95$, $p = 0.05$), with children previously diagnosed with ASD having significantly longer MaxLine scores ($M = 141.53$, $SD = 120.25$) than the TD children ($M = 97.47$, $SD = 86.30$). Similarly, children previously diagnosed with

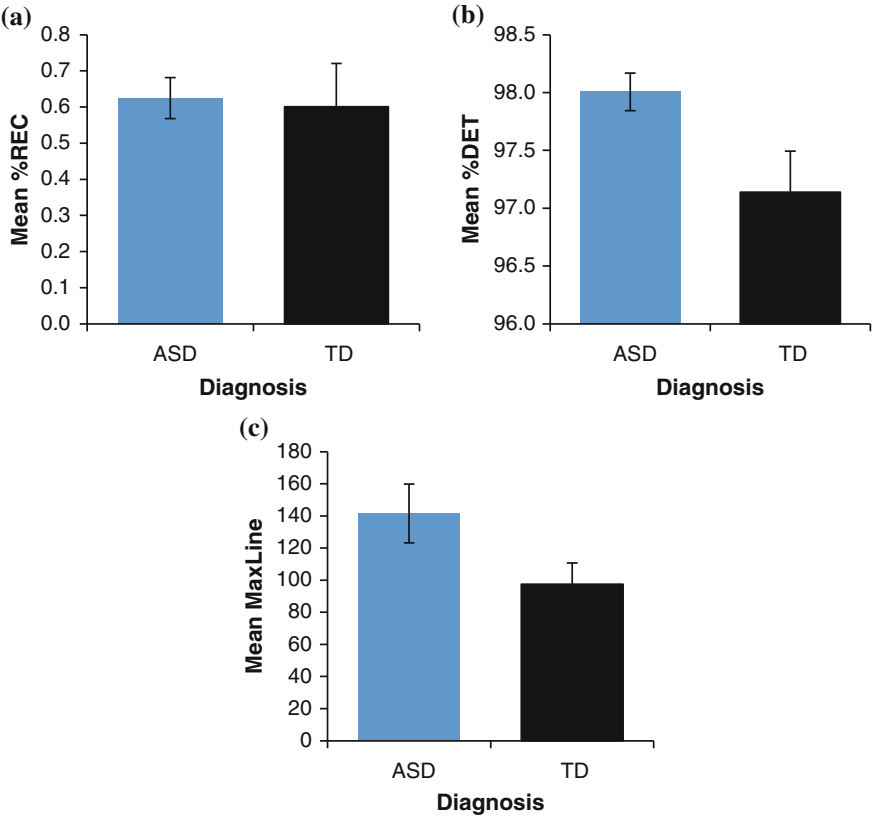

Fig. 12.6 Mean **a** %REC, **b** %DET and **c** MaxLine by diagnosis for the interpersonal hand clapping game. The error bars show the standard error of the mean

ASD showed significantly higher %DET[4] ($M = 98.01$, $SD = 1.06$) than the TD children ($M = 97.14$, $SD = 2.34$; $t(84) = 2.22$, $p = 0.03$; see Fig. 12.6). Together, these MaxLine and %DET results appear to suggest that the coordination observed for the ASD children was more regular and exhibited a greater level of local stability than the coordination observed for the TD children.

[4]Since %DET showed a ceiling effect in all of these analyses, the number of points required to be considered a line in the recurrence plot was systematically raised from 3 up to 7 points in order to investigate more fine-grained differences that might arise between the two groups in this particular measure. The results of this additional analysis however, matched the main results and did not add any new information.

12.4 Discussion

The goal of the current study was to better understand the differences in social motor coordination between typically developing children and those that have been diagnosed with autism spectrum disorder through the use of a nonlinear analysis technique (i.e. cross-recurrence quantification analysis, CRQA) as well as a linear analysis technique (i.e. distribution of the relative phase, DRP). These differences were explored using a simple imitative and synchronous interpersonal object tapping task and synchronous hand-clapping task. For the object tapping task, children's movements were in general more stable in the imitation condition than in the synchrony condition, as shown by the higher occurrence of antiphase relative phase angles (a relative phase of 180°) during imitation trials as opposed to the diminished occurrence of inphase relative phase angles in the synchrony trials. Additionally, the TD children were able to maintain the desired inphase relationship more often than those who had been previously diagnosed with autism in the synchrony condition. This finding was supported by the CRQA measures where MaxLine was significantly higher (i.e. showing a higher degree of stability) for the trials in the imitation condition than those in the synchrony condition. Also, the measure of %REC was able to capture the higher degree of coordination established between the children and the experimenter in the imitation trials compared to the synchrony trials. However, this measure was only able to differentiate the two diagnostic groups in the imitation trials, and not in the synchrony trials.

Regarding the seemingly contradicting results of the object tapping task in the context of diagnosis, where the DRP and CRQA measures differed, we would conclude that DRP and %REC in particular are isolating different aspects of the underlying coordination dynamics and, thus, that these measures provide a complementary index of the stability of coordination for the social object tapping task. Although conclusively defining what these different aspects are remains to be determined, one possibility is that DRP is providing a greater overall index of spatial-temporal coordination independent of movement amplitude or the degree of asymmetric movement kinematics, whereas %REC is providing a better measure of differences in the intrinsic dynamics (i.e., movement stiffness, variability, etc.) of the children's or experimenter's movements during social motor coordination [11, 17]. Thus, for the object tapping task, the current results may indicate that the local movement dynamics of children with ASD are more different than the experimenter's compared to TD children, and that this is more apparent in the imitation task due to the lack of local coupling (which is present in the synchronous condition).

For the interpersonal hand clapping game, the distribution of relative phase showed a significant difference by diagnosis, showing that TD children were able to maintain the desired inphase relationship more often than those previously diagnosed with ASD. Again, however, the CRQA analysis provides a slightly different picture compared to the DRP results. More specifically, although there was no difference in %REC, children with ASD exhibited more deterministic (i.e., higher %

DET) patterns of coordination compared to TD children, as well as a higher degree of local stability (i.e., greater MaxLine) than the TD children. As argued above, the DRP seems to capture an overall index of the spatio-temporal patterning of the coordination, but is less sensitive to the underlying dynamics of the coordination or the intrinsic dynamics of the movements themselves [11, 17]. Indeed, for the synchronous hand-clapping task the CRQA analysis appears to be indicating that the local stability and deterministic structure of the coordination for the ASD children is significantly greater than for TD children. Why then, is the overall coordination as measured by the DRP greater for TD children than for ASD children? Normally, when coordination is seen as less stable, it is assumed that this instability is created by increased random fluctuations, however, it seems that for the synchronous hand-clapping task employed here that the decrease in overall spatial-temporal coordination for ASD children compared to TD children is fueled by more deterministic movements of the child or experimenter when the coordination task involves a child with ASD. This is consistent with previous research examining the structure of movements in patients with Parkinson's disease [18] and Schizophrenia [6] and implies that the movements of the ASD children are less flexible than the movements of TD children [19]. Finally, the current results also highlight the need for future research more specifically aimed at examining the influence of the experimenter's movement modulation in cooperative tasks like the hand clapping game. Indeed, although we see differences and an effect of diagnosis in the task employed here, further work is required to tease apart how much of the difference is mediated by experimenter versus the child since the experimenter in this case was not blind to the child's group membership (due to the easily observable differences between the children's behaviors).

In conclusion, the current study provides new insights about the differences in social motor coordination in children with and without ASD and further exemplifies how the dynamics of social motor coordination can be used as an ASD related behavioral bio-marker in children, and eventually exploited as a diagnostic measure to improve this process. It is important to note that the differences observed between ASD and TD children were most pronounced during synchronized behavior, as opposed to imitative behavior. Accordingly, simple cooperative synchrony tasks seem to be better suited to uncovering ASD related social motor deficiencies [9]. The current results also demonstrate the importance of employing both linear (i.e., DRP) and nonlinear (i.e., CRQA) analysis techniques in order to fully understand the processes underlying social motor deficits in ASD (as well as in other disorders).

Acknowledgments The authors would like to thank Dr. Holly Barnard, Dr. Amie Duncan, Dr. Joseph Amaral, and Carrie Thomas for their help in data collection. Research reported in this publication was supported by the National Institutes of Health under Award Number R21MH094659. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. K.L. Marsh, R.W. Isenhower, M.J. Richardson, M. Helt, A.D. Verbalis, R.C. Schmidt, D. Fein, Autism and social disconnection in interpersonal rocking. Front. Integr. Neurosci. **7**(4), 1–8 (2013)
2. K.L. Marsh, M.J. Richardson, R.C. Schmidt, Social connection through joint action and interpersonal coordination. Top. Cogn. Sci. **1**(2), 320–339 (2009). doi:10.1111/j.1756-8765.2009.01022.x
3. F.J. Bernieri, Coordinated movement and rapport in teacher-student interactions. J. Nonverbal Behav. **12**(2), 120–138 (1988)
4. J.E. Grahe, F.J. Bernieri, The importance of nonverbal cues in judging rapport. J. Nonverbal Behav. **23**(4), 253–269 (1999)
5. L.K. Miles, L.K. Nind, C.N. Macrae, The rhythm of rapport: interpersonal synchrony and social perception. J. Exp. Soc. Psychol. **45**(3), 585–589 (2009)
6. M. Varlet, L. Marin, S. Raffard, R.C. Schmidt, D. Capdevielle, P. Boulenger, J. Del-Monte, B. G. Bardy, Impairments of social motor coordination in Schizophrenia. PLoS One **7**(1) (2012), doi:10.1371/journal.pone.0029772
7. F. Ramseyer, W. Tschachter, Nonverbal synchrony in psychotherapy: coordinated body movement reflects relationship quality and outcome. J. Consult. Clin. Psych. **79**, 284–295 (2011)
8. M. Gratier, G. Apter-Danon, The musicality of belonging: repetition and variation in mother-infant vocal interaction, in communicative musicality: narratives of expressive gesture and being human, in *Communicative Musicality: Exploring the Basis of Human Companionship*, ed. by S. Malloch, C. Trevarthen (Oxford University Press, Oxford, 2008), pp. 301–327
9. P. Fitzpatrick, R. Diorio, M.J. Richardson, R.C. Schmidt, Dynamical methods for evaluating the time-dependent unfolding of social coordination in children with autism. Front. Integr. Neurosci. **7**(21), 1–13 (2013)
10. C.M. Glazebrook, D. Elliot, J. Lyons, A kinematic analysis of how young adults with and without autism plan and control goal-directed movements. Mot. Control **10**(3), 244–264 (2006)
11. M.J. Richardson, R.C. Schmidt, B. Kay, Distinguishing the noise and attractor strength of coordinated limb movements using recurrence analysis. Biol. Cybern. **96**, 59–78 (2007)
12. K. Shockley, M. Santana, C.A. Fowler, Mutual interpersonal postural constraints are involved in cooperative conversation. J. Exp. Psychol. Hum. Percept. Perform. **29**(2), 326–332 (2003)
13. R.C. Schmidt, M.J. Richardson, Dynamics of interpersonal coordination, in *Coordination: Neural, Behavioral and Social Dynamics*, ed. by A. Fuchs, V. Jirsa (Springer, Heidelberg, 2008), pp. 281–308
14. A. Pikovsky, M. Rosenblum, J. Kurths, *Synchrony: A Universal Concept in Nonlinear Sciences* (Cambridge University Press, Cambridge, 2001)
15. M.J. Richardson, K.L. Marsh, R.W. Isenhower, J.R. Goodman, R.C. Schmidt, Rocking together: dynamics of intentional and unintentional interpersonal coordination. Hum. Mov. Sci. **26**(6), 867–891 (2007)
16. R.C. Schmidt, S. Morr, P.A. Fitzpatrick, M.J. Richardson, Measuring the dynamics of interactional synchrony. J. Nonverbal Behav. **36**, 263–279 (2012)
17. M.J. Richardson, S. Lopresti-Goodman, M. Mancini, B. Kay, R.C. Schmidt, Comparing the attractor strength of intra- and interpersonal interlimb coordination using cross-recurrence analysis. Neurosci. Lett. **438**, 340–345 (2008)
18. J.M. Schmit, M.A. Riley, A. Dalvi, A. Sahay, P.K. Shear, K.D. Shockley, R.Y. Pun, Deterministic center of pressure patterns characterize postural instability in Parkinson's disease. Exp. Brain Res. **168**(3), 357–367 (2006)
19. G.C. Van Orden, H. Kloos, S. Wallot, Living in the pink: intentionality, wellbeing, and complexity, in *Handbook of the Philosophy of Science, Philosophy of Complex Systems*, vol. 10, ed. by C. Hooker (Elsevier, Amsterdam, 2011), pp. 639–682

Chapter 13
Restoring Corrupted Cross-Recurrence Plots Using Matrix Completion: Application on the Time-Synchronization Between Market and Volatility Indexes

George Tzagkarakis and Thomas Dionysopoulos

Abstract The success of a trading strategy can be significantly enhanced by tracking accurately the implied volatility changes, which refers to the amount of uncertainty or risk about the degree of changes in a market index. This fosters the need for accurate estimation of the time-synchronization profile between a given market index and its associated volatility index. In this chapter, we advance existing solutions, which are based widely on the typical correlation, for identifying this temporal interdependence. To this end, cross-recurrence plot (CRP) analysis is exploited for extracting the underlying dynamics of a given market and volatility indexes pair, along with their time-synchronization profile. However, CRPs of degraded quality, for instance due to missing information, may yield a completely erroneous estimation of this profile. To overcome this drawback, a restoration stage based on the concept of matrix completion is applied on a corrupted CRP prior to the estimation of the time-synchronization relationship. A performance evaluation on the S&P 500 index and its associated VIX volatility index reveals the superior capability of our proposed approach in restoring accurately their CRP and subsequently estimating a temporal relation between the two indexes even when 80 % of CRP values are missing.

13.1 Introduction

In nowadays rapidly evolving financial world, design of modern data analysis techniques, as a means of understanding the underlying complex dynamics of economic systems, becomes an emerging necessity. Due to the complex behavior of financial

G. Tzagkarakis (✉)
EONOS Investment Technologies, Paris, France
e-mail: gtzag@eonos.com

T. Dionysopoulos
ADVENIS Investment Managers, Paris, France

T. Dionysopoulos
AXIANTA Research, Nicosia, Cyprus
e-mail: dplsthomas@gmail.com

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8_13

time series data, most of the existing methods, which either rely on linear stochastic models or require long data series, may often lead to pitfalls concerning the accurate description of the inherent dynamics, and subsequently the performance of further decision making based on the evolution of the data generating process.

A major feature of any financial system is that markets constitute a complex dynamically evolving universe, which consists of individual, yet interdependent, components. The extraction and precise description of such interrelations is critical towards designing efficient trading strategies. This can be done only indirectly by observing and analyzing the associated financial time series, which are considered to be realizations of the underlying dynamical process. The complex nature of financial data, which cannot be captured by linear stochastic models [1, 2], has been attributed to the fact that financial markets present a non-linear stochastic or chaotic behavior, or a combination of both. Knowledge about the properties of the underlying dynamics is crucial when investing in a market, since the incorporation of such information may improve the robustness and reduce the risk of a trading strategy. This makes non-linear time series analysis one of the most important tools, which can foster radically the identification of opportunities which appear for investors.

In the modern global financial scene, the underlying dynamics of markets is expected to vary across time. *Market indexes* are intended to represent an entire stock market and thus track its changes over time. Furthermore, prior knowledge regarding the level of confidence about the degree of changes in a market index is of high importance for investors, in order to design efficient and successful trading strategies. This confidence level is expressed via the so-called *implied volatility*, whose variations can cause losses to investors even when correct predictions of stock values included in a given market index are available. Implied volatility can have a negative impact on the performance of a trading portfolio, thus affecting adversely the overall investment returns. As such, traders and portfolio managers incorporate volatility expectations in their investment decisions, and adjust actively their positions to better manage the risk associated with volatility fluctuations [3]. The impact of implied volatility on market index returns necessitates the design of effective techniques for estimating the degree of interdependence between a market index and its associated volatility, in order to guarantee the performance of a trading strategy.

To this end, previous studies have given a strong empirical evidence for a negative relation between market index returns and volatility fluctuations [4]. More specifically, negative (respectively, positive) changes of index returns are correlated with positive (respectively, negative) changes of the associated volatility values. From another perspective, volatility can be considered as an indirect realization of the underlying market dynamics, as well as a descriptor of investors' sentiment. Existing methods for explaining the time-varying relation between market index returns and volatility values, primarily rely on the typical correlation coefficient [5–7], on simple regression analysis [8], or on GARCH models [9, 10]. However, these models often fail to capture the asymmetric effects of volatility on index returns, since the absolute effect of returns is considered, while ignoring the positivity or negativity of volatility changes [11]. To overcome some of those limitations, vector autoregres-

sion (VAR) approaches [12, 13] have been proposed recently to better understand the interrelation between index returns and volatility response. However, VAR-based approaches can be quite sensitive to the time-varying selection of the proper lag order for the VAR model.

Despite their drawbacks, all the previous works have identified contemporaneous and lead-lag relationships between index returns and changes in volatility. This fosters the need to design improved algorithmic tools for extracting and monitoring the time-synchronization profile between a given market index and the associated volatility index, whilst accounting for their distinct underlying dynamics, towards protecting the performance of a trading portfolio.

In late 1980's, the method of *recurrence plots* (RP) [14, 15] was proposed as an alternative to describe complex dynamics of a single system on the basis of recurrence theory, through phase-space reconstruction via time-delay embeddings. Then, a considerable amount of diverse information on the dynamics of the system can be extracted and quantified by means of appropriately defined quantitative measures [16–19], which constitute the framework of *recurrence quantification analysis* (RQA) [15, 16]. In order to analyze the dependencies and interrelations between the underlying dynamics of two different systems, the *cross recurrence plot* (CRP) was introduced as a bivariate extension of RPs [20]. An important property of CRPs, which will be exploited in our subsequent analysis of interdependencies between market index values and volatility fluctuations, is that it reveals the local difference of the dynamical evolution of close trajectory segments, represented in the form of bowed lines. Then, a *line of synchronization* (LOS) can be estimated from these curvilinear structures, which is capable of representing more complex lead-lag relationships when compared against the previous methods.

The advanced data analysis capabilities of RPs and CRPs yielded an increasing interest in financial research. In particular, RQA has been exploited to identify deterministic non-linearities in financial time series [21], to measure financial data volatility and detect correlations between currency time series [22], as well as for detecting the critical regime and estimate the bubble initial time in financial markets [23]. CRPs, in specific, have been already exploited in the financial industry to analyze convergence and synchronicity of business and growth cycles [24], to examine the interactive behavior between the hourly accepted weighted average price and the hourly required load in electricity markets [25], as well as for understanding the interrelation between commodity and stock indexes [26] or the coupling of the European banking and insurance sectors [27].

Although CRPs are very efficient in extracting simultaneously the individual and coupled dynamics of a pair of time series, however, there can be situations where information is missing in financial data, thus hindering a robust decision making. More specifically, we distinguish between *structural missing* and *observational missing* information. The former type is related to values that are not expected to be available, for instance, stock prices will be unavailable when stock markets are closed at weekends or holidays. On the other hand, the later type corresponds to the case where previously available values have been lost, for instance, due to malfunction of the storage media. Furthermore, notice that although structural missing information

refers mostly to the observed time series, however, this is also reflected to the associated CRP in the form of gaps in the time intervals where one of the two time series lacks observations. In both cases, appropriate meaningful values must be generated to fill the gaps, prior to further data processing.

Focusing again on the problem of examining potential interrelations between a pair of market and volatility indexes based on their CRP, any source of missing information may yield a completely erroneous interpretation of their time-synchronization profile. Subsequently, this can influence significantly the performance of a trading strategy, which is based on the accurate estimation of temporal associations between stock market and volatility indexes, in order to predict future investment returns.

To address the critical task of recovering missing information in CRPs, and subsequently of improving the estimation of the time-synchronization relation between a pair of stock market and volatility indexes, in this study we exploit the power of *matrix completion* (MC) [28], which enables the accurate recovery of unknown or missing entries of a matrix under certain conditions. More specifically, first we verify that the necessary conditions of the MC framework are satisfied for a CRP. Then, an appropriate MC step is applied on a corrupted CRP prior to the estimation of the time-synchronization relation between two time series, as expressed by their estimated LOS. The performance of our proposed approach is illustrated for the pair of S&P 500 and VIX indexes, which is of particular interest to investors and traders, as it will be described in the following section.

We emphasize, though, that our proposed approach is generic enough and can be applied for recovering missing information in any type of CRPs. The specific choice of financial time series is motivated by the fact that financial data are typically complex, whereas they are characterized by non-linear interdependencies, thus providing a challenging case study. Furthermore, to the best of our knowledge, this is the first attempt to bridge the fields of recurrence analysis and matrix completion, giving insights and paving the way for novel solutions in distinct application areas other than finance and economics. As such, this work should not be considered as an action plan for designing efficient trading strategies, but as a more general framework to remedy the case of missing data. A more detailed examination of the effects of our proposed framework on the performance of a trading portfolio is left as a separate thorough study.

The rest of the chapter is organized as follows: Sect. 13.2 introduces the main concepts and characteristics related to stock and volatility indexes, and specifically to the S&P 500–VIX pair. In Sect. 13.3, the process for estimating the time-synchronization relation between two time series from their CRP is reviewed, while in Sect. 13.4 our proposed approach for recovering the LOS of CRPs with missing entries, using matrix completion, is analyzed in detail. Section 13.5 evaluates the performance of our proposed approach applied on S&P 500 and VIX, in terms of estimating accurately the LOS from a restored CRP and then detecting the critical times, where the true S&P 500 values exceed the 30-day expected range, which is predicted by the current VIX level. To further validate the capabilities of our proposed approach, we compare against an interpolation-based technique, which is widely used for restor-

ing missing entries of image data. Finally, Sect. 13.6 highlights the major quantitative results and qualitative economic implications of our method, and provides directions for further improvements.

13.2 Market and Volatility Indexes: The S&P 500–VIX Case

In this section, the main concepts and properties of stock market and volatility indexes are briefly reviewed for completeness of presentation. Special focus is given to the S&P 500–VIX pair, which constitutes our selected case study for evaluating the performance of our proposed approach.

In finance, a volatility index is a mathematical measure of how much the market believes that a stock index will fluctuate in the future over a predetermined time interval. This belief is based on an analysis of the difference between current put and call option prices of the stock index. Although a volatility index is not expressed as a percentage, however, it should be understood like that. For instance, a volatility index value of 10 translates to an implied volatility of 10 % on the associated stock index. This means that the stock index has a probability of 68 % (that is, one standard deviation) of trading within a range, which is 10 % higher or lower than its current level over the next predetermined time interval [29]. Furthermore, a volatility index typically rises when the interest for buying *put options* increases, and falls when the buying activity of *call options* predominates[1] [30]. The importance for analyzing stock index–volatility index pairs stems from the fact that low values of a volatility index indicate a so-called *bearish market*, that is, a market whose economy is bad, its recession is looming, and stock prices are falling, thus making it tough for investors to select profitable stocks. On the contrary, high values of a volatility index indicate a so-called *bullish market*, that is, a market whose economy is robust and stocks are rising, thus enabling investors to identify profitable stocks [31].

A special case of a stock index–volatility index pair is the S&P 500–VIX pair. S&P 500 is a core index of equities in the United States, which is composed of 500 market value-weighted stocks from a broad range of industries, that are traded on the New York Stock Exchange (NYSE), the American Stock Exchange (AMEX), and the Nasdaq National Market System [32]. The weighting process controls the influence of each company on the overall performance of the index in such a way that is directly proportional to the market value of each company. It is exactly this weighting scheme, which ranks the S&P 500 index among the prevalent standards for measuring the performance of actual portfolios in the industry sector.

The volatility index which is associated with S&P 500, is the Chicago Board Options Exchange (CBOE) market volatility index. The CBOE volatility index, which is widely known under the trademarked ticker symbol “VIX”, is considered to

[1] A put option gives the purchaser the right, but not the obligation, to sell a security for a specified price at a certain time. A call option is a right to buy the same.

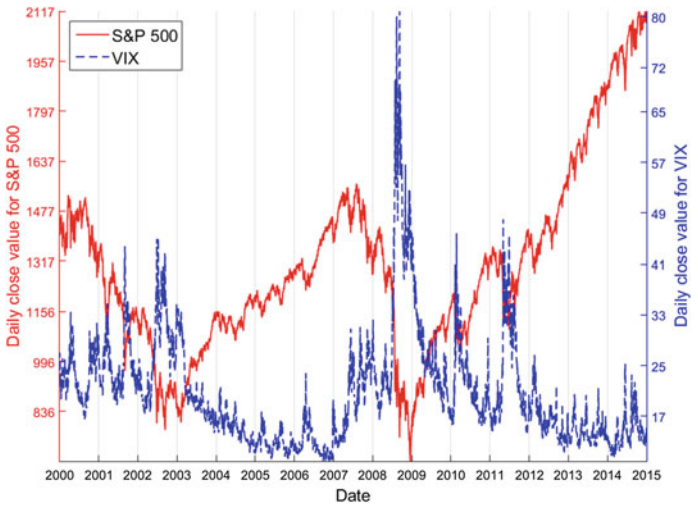

Fig. 13.1 Daily close values for S&P 500 and VIX for the period 03/01/2000–08/05/2015

be among the world's leading measures of equity market volatility. The calculation of VIX is based on real-time prices of options on the S&P 500 and is designed to reflect the consensus view of investors with respect to the future (30-day) expected stock market volatility [33]. A characteristic property of the S&P 500–VIX pair is that VIX has an inverse relationship with the market condition, as expressed by the values of S&P 500 (Fig. 13.1) [7]. More specifically, low VIX values (typically below 15) indicate a reduced interest of investors to take decisions under the current market conditions, which, in general, leads to a period of increased volatility. On the other hand, high VIX values (typically greater than 30) are associated with a high volatility as a result of investors' fear. In particular, the value of VIX increases as the investors become fearful and decreases when they feel confident about the future direction of the market [34].

The subsequent analysis and experimental evaluations are performed on S&P 500 and VIX values which cover the period between January 03, 2000, and May 08, 2015. More specifically, daily close prices were obtained for S&P 500 from the S&P Dow Jones Indices repository,[2] while the corresponding daily values for VIX were acquired from the CBOE website.[3]

The inverse relation between S&P 500 and VIX motivates the design of efficient methods for monitoring accurately the time-synchronization profile between the corresponding index values. In particular, by tracking the changes in VIX, while having estimated such a time-synchronization relation, we enable the prediction of critical events for the stock index, and subsequently the protection of an investment. To this end, the computation of the time-varying correlation between S&P 500 and VIX is

[2] http://us.spindices.com/indices/equity/sp-500.

[3] http://www.cboe.com/micro/vix/historical.aspx.

a commonly used method to monitor their lead-lag relation. However, correlation can be very sensitive to the presence of outliers, yielding erroneous results, whilst it does not account for the true underlying dynamics that generate the two indexes. Taking into consideration the dynamical processes, which drive the two indexes, may reveal a more complex interdependence among them. To address this problem, cross-recurrence analysis is employed as a powerful alternative in order to account for the underlying index dynamics and estimate an accurate lead-lag relationship between S&P 500 and VIX. The following section describes the process for extracting the time-synchronization mapping between S&P 500 and VIX by employing their cross recurrence plot.

13.3 Time-Synchronization Using CRPs

Recurrence is a fundamental feature of many non-linear dynamical systems, which is defined as a *time* when the trajectory returns to a location it has visited before. The study of recurrences, towards understanding the dynamics of non-linear systems, is facilitated by means of a powerful tool, the so-called *recurrence plot* (RP), and its quantification. The method of RPs was first introduced in [14] as an advanced technique of non-linear data analysis, to study recurrences and non-stationary behavior of dynamical systems. An RP is a visualization (graph) of a square matrix, in which the matrix elements correspond to those times at which a state of a dynamical system recurs (columns and rows correspond then to a certain pair of times). Such a recurrence of a state, occurring at time i, at a different time j is represented within the two-dimensional square matrix with ones (recurrence) and zeros (non-recurrence), where both axes are time axes. In fact, an RP reveals all the times when the *phase-space trajectory* of the dynamical system visits roughly the same area in the phase-space.

Given a time series with n samples, $\mathbf{s} = \{s_i\}_{i=1}^{n}$, a phase-space trajectory can be reconstructed via *time-delay embedding*,

$$\mathbf{x}_i = [s_i,\, s_{i+\tau}, \ldots, s_{i+(m-1)\tau}], \quad i = 1, \ldots, N\,, \tag{13.1}$$

where $\mathbf{x}_i$ denotes the i-th state, m is the embedding dimension, τ is the delay, and $N = n - (m-1)\tau$ is the number of states. The appropriate choice of the embedding parameters is crucial and needs special attention, as we will discuss in the following section. Having constructed a phase-space representation, an RP can be expressed mathematically as follows:

$$\mathbf{R}_{i,j} = \Theta\left(\varepsilon - \|\mathbf{x}_i - \mathbf{x}_j\|_p\right), \quad i, j = 1, \ldots, N\,, \tag{13.2}$$

where $\mathbf{x}_i, \mathbf{x}_j \in \mathbb{R}^m$, ε is a distance threshold, $\|\cdot\|_p$ denotes a general ℓ_p norm (Euclidean distance ($p = 2$) is commonly used), and $\Theta(\cdot)$ is the Heaviside step function, whose discrete form is defined as

$$\Theta(k) = \begin{cases} 1, & \text{if } k \geq 0 \\ 0, & \text{if } k < 0 \end{cases}, \quad k \in \mathbb{R} \,. \tag{13.3}$$

The resulting matrix $\mathbf{R}$ exhibits the main diagonal, $\mathbf{R}_{i,i} = 1$ $(i = 1, \ldots, N)$, also known as the *line of identity* (LOI). Typically, several linear (and/or curvilinear) structures appear in RPs, which give hints about the time evolution of the high-dimensional phase-space trajectories. Besides, a major advantage of RPs is that they can be also applied to rather short and even non-stationary data.

A special case of RP, which is useful in the study of phase-space trajectories, is the *unthresholded* RP [35] (also called *global* RP, or *distance plot*). Instead of plotting the recurrences, the unthresholded version is simply obtained by plotting the distances between states $\mathbf{x}_i$ and $\mathbf{x}_j$,

$$\mathbf{D}_{i,j} = \|\mathbf{x}_i - \mathbf{x}_j\|_p, \quad i, j = 1, \ldots, N \,. \tag{13.4}$$

The *cross recurrence plot* (CRP) is a bivariate extension of the RP, which was introduced to analyze the dependencies and interrelations between the underlying dynamics of two different systems by comparing their states [20]. A CRP is a graph showing all those times at which a state in one dynamical system occurs simultaneously in a second dynamical system. In other words, a CRP reveals all the times when the phase-space trajectory of the first system visits roughly the same area in the phase-space, where the trajectory of the second system is. Let $\mathbf{x}_i$ and $\mathbf{y}_j$ represent the trajectories of two dynamical systems. Then, the corresponding cross recurrence matrix is defined by

$$\mathbf{R}^{\mathbf{C}}_{i,j}(\varepsilon) = \Theta\left(\varepsilon - \|\mathbf{x}_i - \mathbf{y}_j\|_p\right), \quad i = 1, \ldots, N, \quad j = 1, \ldots, M \,, \tag{13.5}$$

where, in general, the two trajectories may not be of equal length ($N \neq M$), and hence the matrix $\mathbf{R}^{\mathbf{C}}$ is not necessarily square. However, we emphasize that both systems must be represented in the same phase-space, since a CRP looks for those times when a state of the first system recurs to one of the other system. If the embedding parameters estimated from the two time series are not equal, the higher embedding should be chosen.

Similarly to its RP counterpart, an *unthresholded* CRP is obtained by plotting the distances between states $\mathbf{x}_i$ and $\mathbf{y}_j$,

$$\mathbf{D}^{\mathbf{C}}_{i,j} = \|\mathbf{x}_i - \mathbf{y}_j\|_p, \quad i = 1, \ldots, N, \quad j = 1, \ldots, M \,. \tag{13.6}$$

Notice also that, since the values of the main diagonal $\mathbf{R}^{\mathbf{C}}_{i,i}$ (or $\mathbf{D}^{\mathbf{C}}_{i,i}$) $(i = 1, \ldots, N)$ are not necessarily equal to each other, usually there is not a (straight) main diagonal in the CRP (or the unthresholded CRP).

An important property of CRPs, as well as of unthresholded CRPs, which will be exploited in our subsequent analysis of interrelations between financial market indexes, is that they reveal the *local difference of the dynamical evolution* of close

trajectory segments, represented in the form of bowed lines. A time dilation or time compression of one of the trajectories causes a distortion of the diagonal lines. Then, the LOI will be disrupted, and in this case it will be called a *line of synchronization* (LOS) [36]. Specifically, if we stretch or compress the second trajectory slightly, then, a deformation of the LOS will be observed, and although still being continuous it will not be a straight diagonal line, whereas a time shift between the trajectories causes a displacement of the LOS. This enables the estimation of a *non-parametric rescaling function* between our financial time series.

We emphasize that, although our proposed approach is generic and can be applied on any type of CRPs, however, in the subsequent analysis we will mainly rely on unthresholded CRPs. For our selected financial application, this is justified by the fact that a *thresholded* CRP can reveal *convergence* properties between a stock market index and its volatility index, in the sense that the CRP values will indicate dynamics that occur to the two time series within some critical distance. On the other hand, since we are interested in estimating their synchronization profile, an *unthresholded* CRP can enhance the understanding of the phase-space trajectories and detect *phase synchronous* dynamics even when the two time series do not converge. This property of unthresholded CRPs, when compared with their thresholded counterparts, has been exploited for the analysis of synchronicity in growth and business cycles [37], as well as for the detection of financial crisis [38].

13.3.1 Estimation of Embedding Parameters

Without loss of generality, and for illustration purposes, we study the relationship of our selected S&P 500 and VIX indexes by fixing the delay parameter at $\tau = 1$. Concerning the embedding dimension m, a minimal sufficient value is estimated separately for the two time series, that is, S&P 500 prices and VIX values, using the method of false nearest neighbours (FNN) [39]. In practice, the minimal embedding dimension is defined as the dimension for which the fraction of false neighboring points is zero, or at least sufficiently small. Let $m_{\text{S\&P 500}}$ and m_{VIX} denote the embedding dimension of S&P 500 and VIX, respectively. As mentioned above, if $m_{\text{S\&P 500}} \neq m_{\text{VIX}}$, then, the higher embedding $m = \max\{m_{\text{S\&P 500}}, m_{\text{VIX}}\}$ should be chosen for the generation of the CRP or its unthresholded version.

Figure 13.2a shows the FNN ratio curves as a function of the embedding dimension for S&P 500 and VIX (Fig. 13.1). The minimal sufficient embedding dimensions, which are estimated separately, are equal to $m_{\text{S\&P 500}} = 5$ and $m_{\text{VIX}} = 6$, thus the corresponding CRP is constructed for an embedding dimension $m = 6$. Using this parameters setting, a detailed depiction of the relative distances between the corresponding states is given by the unthresholded CRP, which is shown in Fig. 13.2b.

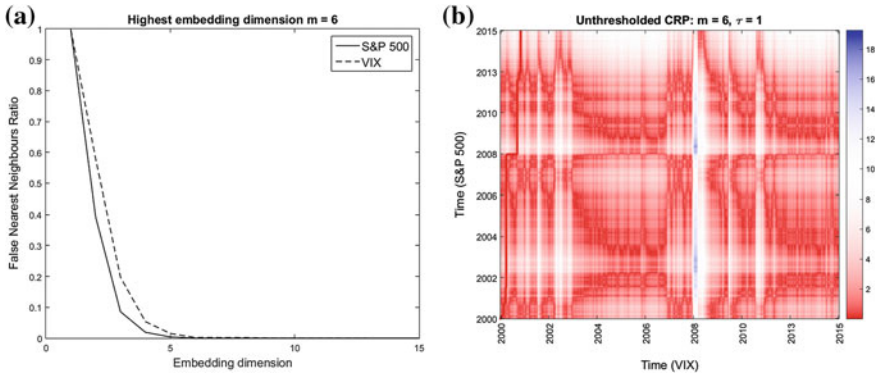

Fig. 13.2 **a** False nearest neighbours ratio as a function of the embedding dimension, **b** unthresholded CRP, for S&P 500 and VIX in the period 03/01/2000–08/05/2015 ($m = 6$, $\tau = 1$)

13.4 Restoring Missing Entries in Unthresholded CRPs Using Matrix Completion

Many of the sophisticated methods that have been developed for analyzing financial data can be particularly sensitive to data loss, thus affecting subsequent data processing and robust decision making. From the one hand, missing structural information typically exists when data is not expected to be available, for instance, when stock markets are closed at weekends or holidays. On the other hand, missing observational information is related to the case where previously available values have been lost, for instance, due to malfunction of the storage media. We emphasize again that, although structural missing information refers mostly to the observed time series, however, this is also reflected to the associated CRP in the form of gaps in the time intervals where at least one of the two time series lacks observations. As such, in both cases we address the problem of recovering missing information from a corrupted (unthresholded) CRP directly, prior to further data processing.

Several *imputation* methods [40] have been proposed in the past to address the problem of recovering missing information in financial data sets. To this end, there are two fundamental approaches: (i) fit a model and use it to predict the missing data; (ii) find the observation that is most similar to the one with the missing value given a predetermined "similarity" criterion. In the former case, the choice of a model can come from statistics or machine learning and includes techniques, such as, regression, neural networks, support vector machines, and decision trees. However, the main drawback of all those techniques is that complete data is needed to fit an accurate model. In the later case, typical methods include interpolation (e.g., linear, cubic, spline) or extrapolation methods to find the missing observation that is most "similar" to its neighboring values. However, the major limitation of such methods is that they are mostly based on geometric features of the data, while ignoring the inherent statistical or dynamical process generating the data.

Advanced signal processing techniques have been introduced recently for providing highly accurate restoration of missing data. Among them, *matrix completion* (MC) is a novel powerful framework, capable of recovering missing observations in a reliable fashion. In [28], it was shown that most low-rank matrices could be recovered from sufficiently large sets of entries by computing the matrix of *minimum nuclear norm* that agrees with the provided entries. The nuclear norm is equal to the sum of the singular values of a matrix and is the best convex lower bound of the rank function on the set of matrices whose singular values are all bounded by 1. Apart from the *low rank assumption* for our data, *no additional prior knowledge* is required to estimate the missing observations.

More specifically, given a $N \times M$ unthresholded cross-recurrence matrix $\mathbf{D}^{\mathbf{C}}$ (the same holds for the thresholded case), recovering the complete set of $N \cdot M$ entries of the matrix from a much smaller number of $K \ll N \cdot M$ entries is possible, provided that both the number of missing entries and the rank of the matrix are appropriately bounded [41]. Let $\mathscr{M} : \mathbb{R}^{N \times M} \rightarrow \mathbb{R}^{K}$ be a sampling operator (mask), which represents the process of selecting a subset of the entries of $\mathbf{D}^{\mathbf{C}}$, that is,

$$d_{i,j} = \mathscr{M}\left(\mathbf{D}^{\mathbf{C}}\right)_{i,j} = \begin{cases} 1, & \text{if } (i,j) \in \Omega \\ 0, & \text{if } (i,j) \notin \Omega \end{cases}, \tag{13.7}$$

where Ω denotes the set of indices of the *sampled entries*, or, equivalently, of the *available entries* of a given (unthresholded) CRP.

Let $\mathbf{X}$ be the target matrix obtained by recovering all the missing entries of $\mathbf{D}^{\mathbf{C}}$, whilst posing an equality constraint between the entries of $\mathbf{X}$ and $\mathbf{D}^{\mathbf{C}}$ whose indices belong to the set Ω. Then, the full matrix $\mathbf{X}$ can be estimated by solving the following nuclear norm optimization problem:

$$\min_{\mathbf{X}} \|\mathbf{X}\|_{*} \quad \text{subject to} \quad \mathscr{M}(\mathbf{X}) = \mathscr{M}(\mathbf{D}^{\mathbf{C}}) \,. \tag{13.8}$$

Recall that the nuclear norm of a matrix $\mathbf{X}$ is defined as the sum of its singular values $\|\mathbf{X}\|_{*} = \sum_{k} \sigma_{k}(\mathbf{X})$, where $\sigma_{k}(\mathbf{X})$ denotes the k-th largest singular value of $\mathbf{X}$. Furthermore, according to the spectral theorem associated with the singular value decomposition of $\mathbf{X}$, low-rank matrices, such as the ones produced by spatio-temporally correlated processes, as we expect to be the case with the (unthresholded) CRPs of stock and volatility indexes, are characterized by a small number of non-zero singular values. As such, before proceeding, the first requirement to be verified is whether the singular values of our unthresholded CRP decay very quickly. Indeed, as it is shown in Fig. 13.3, the singular values of the unthresholded CRP, which is depicted in Fig. 13.2b, decay very quickly to zero (notice the logarithmic scale of the x-axis), thus indicating that the MC could be applied successfully to restore its missing entries.

Various approaches have been proposed for the solution of the optimization problem (13.8). In this study, the Singular Value Thresholding (SVT) algorithm [42] is employed, due to the minimal storage space requirements and low computational

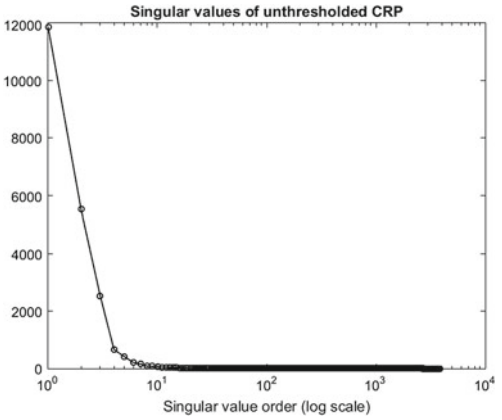

Fig. 13.3 Singular values of the unthresholded CRP of S&P 500 and VIX for the period 03/01/2000–08/05/2015

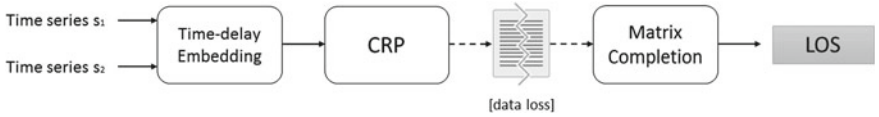

Fig. 13.4 Flow diagram for estimating the time-synchronization profile (LOS) of two time series from their corrupted CRP with missing information

cost per iteration, and also because it was shown to preserve the details (curvilinear structures) in an (unthresholded) CRP. Let $\hat{\mathbf{D}}^{\mathbf{C}}$ denote the restoration of the corrupted unthresholded CRP, obtained as the solution of (13.8). The LOS is then estimated for the corresponding stock and volatility indexes by applying the method described in [36] on the restored CRP $\hat{\mathbf{D}}^{\mathbf{C}}$. Our proposed approach for estimating the time-synchronization profile of stock and volatility indexes (or, in general, of any given pair of time series) from a corrupted (unthresholded) CRP with missing information is summarized in Fig. 13.4.

13.5 Performance Evaluation

In this section, the performance of our proposed approach is evaluated and compared with a well-established interpolation technique, in terms of estimating accurately the LOS from a corrupted unthresholded CRP between S&P 500 and VIX with missing entries. As described in Sect. 13.2, our data set consists of daily close prices for S&P 500, along with the corresponding VIX values, covering the period between January 03, 2000 and May 08, 2015.

First, an unthresholded CRP is calculated for the two indexes over the whole time period of 15 years, by setting the embedding parameters to $\tau = 1$ for the delay and $m = 6$ for the embedding dimension (Sect. 13.3.1). We emphasize that the subsequent performance analysis is not affected by the selection of the embedding parameters, in the sense that our proposed approach is valid as soon as the singular values of the calculated unthresholded CRP decay very quickly, as it was described in Sect. 13.4. Then, the scenario of missing information from the original unthresholded CRP is simulated by generating artificially corrupted versions of it. This is done by randomly selecting a percentage ρ of entries whose corresponding unthresholded CRP value is set equal to `NaN`, thus being equivalent to missing entries. In the subsequent evaluation, the percentage of missing entries varies in $\rho \in \{10\,\%, 20\,\%, \ldots, 80\,\%\}$.

The performance of both our proposed approach and the interpolation-based technique is measured in terms of the approximation accuracy between the original LOS, which is estimated from the original unthresholded CRP, and the LOS estimated from the restored unthresholded CRPs. In particular, let $\mathbf{l} \in \mathbb{R}^L$ and $\hat{\mathbf{l}} \in \mathbb{R}^L$ denote the original and restored LOS, respectively. In this study, we chose the root mean squared relative error (RMSRE) as a metric for measuring the reconstruction quality of the original LOS, which is defined as follows,

$$E(\mathbf{l}, \hat{\mathbf{l}}) = \sqrt{\frac{1}{L} \sum_{i=1}^{L} \left(\frac{l_i - \hat{l}_i}{l_i} \right)^2} \,. \tag{13.9}$$

Furthermore, in order to reduce the bias due to the random selection of missing entries, for each value of ρ the RMSRE is averaged over 100 Monte-Carlo runs, where in each run a different subset of missing entries is chosen at random. Algorithm 1 describes the steps for calculating the (unthresholded) CRP of two given time series, while Algorithm 2 summarizes the main steps of our proposed approach for restoring a corrupted CRP using MC and estimating the LOS.

Algorithm 1 Calculation of (unthresholded) CRP

1. Inputs: $\mathbf{s}_1 \in \mathbb{R}^{n_1}, \mathbf{s}_2 \in \mathbb{R}^{n_2}$ (input time series), τ (delay parameter), p (ℓ_p norm for the distance plot)

2. Estimate embedding dimension:

- Use False Nearest Neighbours (FNN) to estimate $m_{\mathbf{s}_1}$ for the given τ
- Use False Nearest Neighbours (FNN) to estimate $m_{\mathbf{s}_2}$ for the given τ
- $m = \max\{m_{\mathbf{s}_1}, m_{\mathbf{s}_2}\}$

3. Calculate CRP:

- Construct the states of $\mathbf{s}_1$, $\mathbf{x}_i = [s_{1,i}, s_{1,i+\tau}, \ldots, s_{1,i+(m-1)\tau}]$, $i = 1, \ldots, N$
- Construct the states of $\mathbf{s}_2$, $\mathbf{y}_j = [s_{2,j}, s_{2,j+\tau}, \ldots, s_{2,j+(m-1)\tau}]$, $j = 1, \ldots, M$
- Calculate unthresholded CRP, $\mathbf{D}^{\mathbf{C}}_{i,j} = \|\mathbf{x}_i - \mathbf{y}_j\|_p$, $i = 1, \ldots, N$, $j = 1, \ldots, M$

4. Output: $\mathbf{D}^{\mathbf{C}} \in \mathbb{R}^{N \times M}$ (unthresholded CRP)

Algorithm 2 Restoration of corrupted (unthresholded) CRPs using MC, and LOS estimation

1. Inputs: $\tilde{\mathbf{D}}^{\mathbf{C}} \in \mathbb{R}^{N \times M}$ (corrupted (unthresholded) CRP), Ω (set of indices of available CRP entries)

2. Restore missing entries using MC:

(*The SVT algorithm [42] is used to solve the following optimization problem*)

$$\hat{\mathbf{D}}^{\mathbf{C}} = \arg\min_{\mathbf{X} \in \mathbb{R}^{N \times M}} \|\mathbf{X}\|_* \text{ subject to } \mathscr{M}(\mathbf{X}) = \mathscr{M}(\tilde{\mathbf{D}}^{\mathbf{C}})$$

3. Estimate the LOS:

Estimate the LOS, $\hat{\mathbf{l}} \in \mathbb{R}^L$, by applying on $\hat{\mathbf{D}}^{\mathbf{C}}$ the method described in [36]

4. Outputs: $\hat{\mathbf{D}}^{\mathbf{C}} \in \mathbb{R}^{N \times M}$ (restored unthresholded CRP), $\hat{\mathbf{l}} \in \mathbb{R}^L$ (estimated LOS)

Concerning the interpolation-based technique against which we compare, among the several existing solutions, we employ a widely used data inpainting method that looks for `NaN` elements in an array and attempts to interpolate (or extrapolate) smoothly, in order to restore the missing values. To this end, a corrupted unthresholded CRP is treated as a monochromatic image, that is, an image with one color component per pixel, whereas the regions with the missing entries (`NaN` values) are considered to be the inpainting domains to be filled in. The core of this method consists of a partial differential equation (PDE), which is applied in the inpainting domains to fill the gaps [43]. Specifically, a fourth-order PDE model is formulated to allow for the transportation of available information from the exterior towards the interior of an inpainting domain and the simultaneous diffusion of the information inside the inpainting region. Due to the increased computational complexity of this method, in the subsequent evaluation we are restricted to a part of the original unthresholded CRP (Fig. 13.2b), by randomly selecting the time period 16/12/2005–01/12/2011 for both the S&P 500 and VIX. Then, the artificially corrupted CRPs, along with their restored versions, correspond to a square region of size 1500×1500, which is shown in Fig. 13.7a. In the rest of the text, whenever we refer to the original unthresholded CRP we mean this 1500×1500 region. We also emphasize that our objective is not to compare with the state-of-the-art data interpolation methods, but instead to verify the efficiency of the MC framework when dealing with missing CRP entries associated with financial data.

As a first evaluation, the reconstruction accuracy of the original LOS from the restored unthresholded CRPs is calculated as a function of the percentage of missing entries ρ. To this end, Fig. 13.5 shows the average RMSRE as a function of $\rho\,(\%)$ for our proposed approach based on MC (SVT) and the PDE-based method (Data Inpaint), along with the error bars corresponding to one standard deviation over the 100 Monte-Carlo runs. Clearly, the MC-based approach achieves a highly accurate restoration of the original unthresholded CRP, which yields an equally high accuracy in estimating the LOS. Furthermore, our proposed approach outperforms significantly the PDE-based method, with the difference in LOS estimation accuracy being more prominent for high percentages of missing entries. In addition, the small error bars of the MC-based approach, when compared with their much larger counterparts of the PDE-based method, indicate a more robust performance over the distinct Monte-Carlo runs.

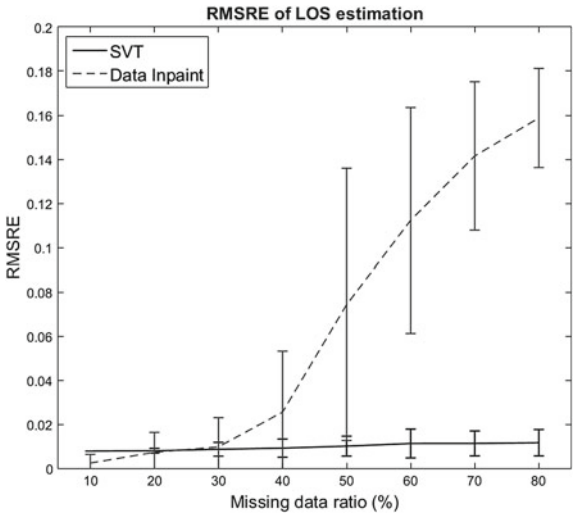

Fig. 13.5 Average RMSRE between the original LOS and its reconstructions from the restored unthresholded CRPs as a function of the percentage of missing entries, for our proposed MC-based approach (SVT) and the PDE-based method (Data Inpaint)

For a visual inspection of the restoration capabilities of both the MC and the PDE-based methods, we present the recovered unthresholded CRPs for two extreme values of ρ, that is, a low percentage of missing entries $\rho = 20\,\%$ and a very high percentage of missing entries $\rho = 80\,\%$. Specifically, Fig. 13.6 shows the original and corrupted unthresholded CRPs of S&P 500 and VIX in the time period 16/12/2005–01/12/2011, with 20 % of missing CRP values, along with its restored version using MC and data inpainting by employing the selected PDE-based method. At least visually, both of the restored CRPs look very similar to each other. This is also the case when 80 % of the original unthresholded CRP values are missing, as shown in Fig. 13.7, where the differences between the MC-based and the PDE-based recovered CRPs are hardly noticeable visually.

However, the difference in restoration accuracy between the proposed MC-based approach and its PDE-based counterpart becomes prominent in Fig. 13.8. This figure shows the LOS estimated from the original unthresholded CRP, along with instances of the LOS, which have been estimated from the restored unthresholded CRPs, for $\rho = 20\,\%$ and $\rho = 80\,\%$ of missing CRP values, by employing the two restoration methods. Once again, the superiority of the MC-based approach is highlighted, in terms of achieving a very accurate approximation of the original LOS, even for an extremely high percentage ($\rho = 80\,\%$) of missing CRP entries. This is not a surprising result, since most of the PDE-based techniques tend to smooth out the localized discontinuities of piecewise constant curves, or piecewise smooth regions, thus generating artificial edges in order to resemble the human visual perception, but which may confuse the algorithm used to estimate the LOS. On the other hand, this is not the case with the MC-based method, which accounts both for local and global correlations among the columns of the corrupted data matrix. Table 13.1 summarizes the main advantages and limitations of our proposed MC-based approach, as well as of the PDE-based method.

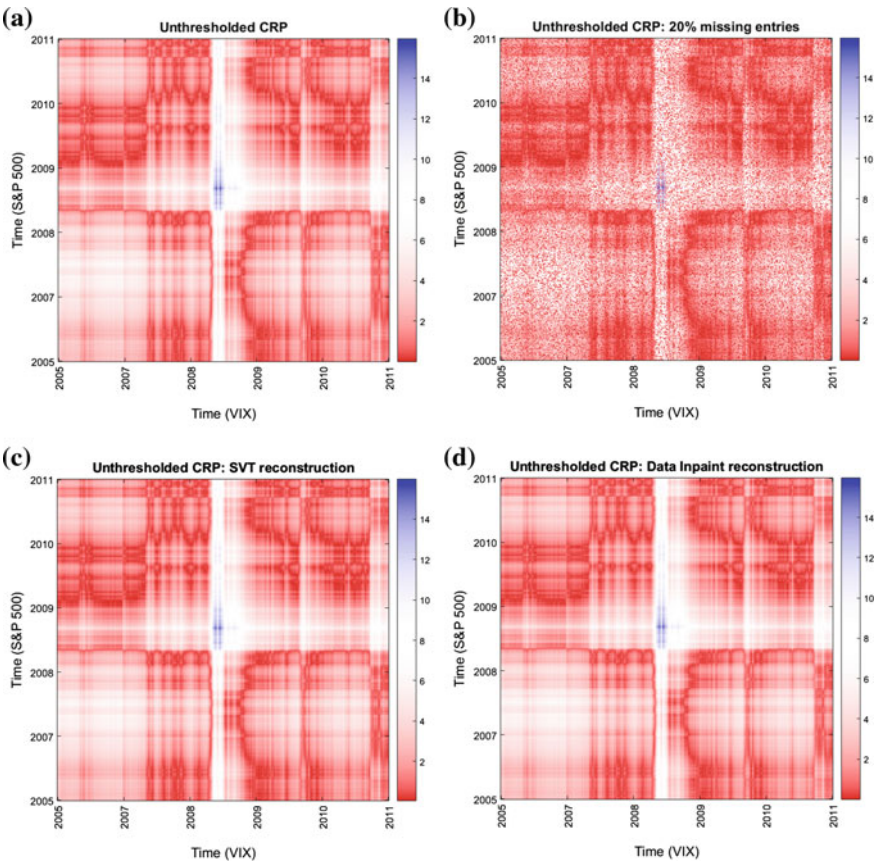

Fig. 13.6 Restored unthresholded CRP of S&P 500 and VIX in the time period 16/12/2005–01/12/2011 with 20 % of missing entries using the MC and the PDE-based methods. **a** Original unthresholded CRP. **b** 20 % missing entries. **c** Restoration using MC. **d** Restoration using data inpainting

13.6 Qualitative Economic Implications and Future Work

In this study, the efficiency of matrix completion (MC) was exploited to restore missing information in corrupted cross recurrence plots (CRP) and subsequently to examine the interrelation between a pair of time series based on their line of synchronization (LOS), which is estimated from the restored CRP. In particular, we focused on the analysis of interdependencies across time among a pair of stock market index and its associated volatility index, which is a challenging task due to the inherent complexity that typically characterizes financial data. To this end, the power of MC

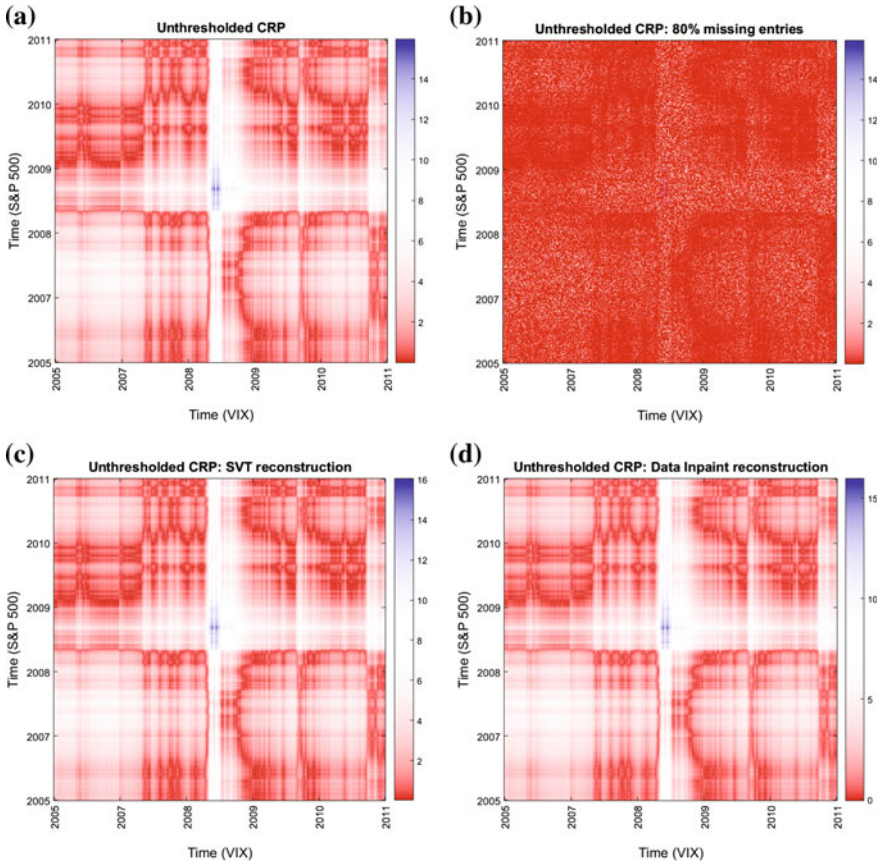

Fig. 13.7 Restored unthresholded CRP of S&P 500 and VIX in the time period 16/12/2005–01/12/2011 with 80 % of missing entries using the MC and the PDE-based methods. **a** Original unthresholded CRP. **b** 80 % missing entries. **c** Restoration using MC. **d** Restoration using data inpainting

in recovering accurately the missing entries of an unthresholded CRP, which is calculated for the market index–volatility index pair, was combined with the efficiency of a cross recurrence analysis, which extracts simultaneously the individual and joint underlying dynamics of the given time series.

For illustration purposes, we applied our proposed approach on a specific market index–volatility index pair, namely, the S&P 500–VIX pair. Our choice was based on the fact that S&P 500 is widely regarded as the best single gauge of large-cap U.S. equities, which captures approximately 80 % coverage of the available market capitalization, whilst VIX is considered to be a key measure for quantifying the sentiment of investors and market volatility. We emphasize, though, that the approach proposed in this study can be applied on any type of time series, as soon as the corresponding recurrence matrix satisfies a low-rank assumption, which is necessary

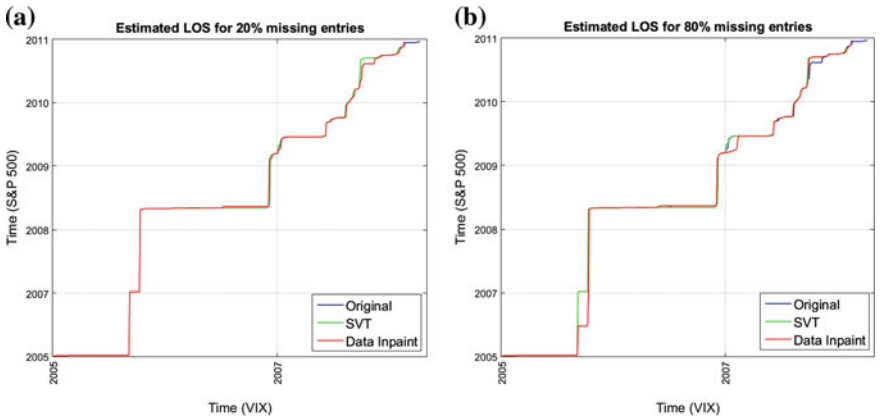

Fig. 13.8 Instances of reconstructed LOS from the restored unthresholded CRPs of S&P 500 and VIX, with 20 and 80 % of missing entries. **a** 20 % missing entries. **b** 80 % missing entries

Table 13.1 Major advantages (Pros) and limitations (Cons) of the MC and PDE methods for restoring corrupted unthresholded CRPs

	MC	PDE
Pros	✓ Adapts to complicated localized patterns	✓ Adapts to complicated localized patterns
	✓ No prior assumption is required for the statistical distribution of the data	✓ No prior assumption is required for the statistical distribution of the data
	✓ Accounts for both local and global correlations between distinct elements (columns) of the corrupted data matrix	✓ Preserves geometric structures
	✓ Computationally tractable	✓ Restored image resembles human visual perception
Cons	– Decreased performance if low-rank assumption is not satisfied	– Smooths discontinuities of piecewise constant curves, introducing artificial segments
	– Increased number of iterations may be required for high percentage of missing entries	– Only localized information is exploited to restore missing entries
		– High computational complexity
		– Diffusion process typically generates blurred patterns in restored image

for the MC framework to achieve an accurate recovery of missing entries, and subsequently, yield an accurate estimation of the time-synchronization relation (that is, the LOS) between the given time series.

The experimental evaluation revealed a superior performance of our proposed MC-based approach, when compared with a PDE-based interpolation technique, which is used for image inpainting. More specifically, from a quantitative point of view, our approach achieved an improved restoration quality of the original unthresh-

olded CRP, even for a high percentage of missing CRP values, when compared with the PDE-based counterpart. This was also verified by the significantly reduced approximation error between the LOS, which is estimated from the original CRP, and the LOS estimated from the CRPs restored using the MC-based method.

From a qualitative, financial, perspective, understanding the interrelation between S&P 500 and VIX is very useful towards building a successful investment strategy. More specifically, the actual VIX level represents the 30-day implied volatility of S&P 500, which conveys information about the likely range of possible index levels that the market *expects* in a month. Although this one-month-ahead index level is only an expectation, however, it can be exploited in order to adjust the positions of a trading portfolio.

In the following, let S_t and α_t denote the value of S&P 500 and the corresponding VIX level, respectively, at time t. Then, the VIX level can be easily converted into useful information about the future, 30-day, expected range of the market index values as follows:

1. The VIX level α_t is seen as a percentage, $\alpha_t\%$, which corresponds to the *annualized* implied 30-day volatility for the S&P 500;
2. The VIX percentage level is de-annualized to obtain its true monthly value. This is done simply by dividing with the square root of 12, that is, the number of months in a year, $\tilde{\alpha}_t\% = \frac{\alpha_t\%}{\sqrt{12}}$;
3. The current index value S_t is multipled by the de-annualized factor, $\tilde{\alpha}_t\%$, to obtain the bound b_{S_t} of the expected range, $b_{S_t} = S_t \cdot \tilde{\alpha}_t\%$;
4. Finally, the expected range of the market index value in a month is given by $[S_{t+30,lower}, S_{t+30,upper}] = [S_t - b_{S_t}, S_t + b_{S_t}]$. The range between $S_{t+30,lower}$ and $S_{t+30,upper}$ is where the market expects the S&P 500 to trade in 30 days from now, with a 68 % level of confidence (that is, one standard deviation of a Gaussian distribution).

Figure 13.9 shows the true S&P 500 values (blue curve), along with the 30-day expected range (red curves) as it is predicted by the VIX levels using the above four steps, in the period 01/02/2006–01/12/2011. As it can be seen, in general, the true S&P 500 values lie in the expected range. However, there are cases, such as the regions in the black rectangles, where the true S&P 500 values exceed the 30-day expected values. The early detection of those time instants, where the true S&P 500 values start declining from their 30-day predicted range, could protect the performance of a trading portfolio from becoming unstable due to positions that are taken based on an erroneous prediction of the stock index prices.

To address this problem, towards increasing the robustness of a trading portfolio, the use of the typical cross-correlation between S&P 500 and VIX does not suffice. Indeed, as it can be seen in Fig. 13.10, there is not an obvious relationship among the values of moving cross-correlation between the two indexes (using a 30-day window length with 1-day overlapping), and the time instants where the true S&P 500 values start declining from the 30-day expected range, indicated by the black vertical lines. These time instants have been estimated by means of the original LOS,

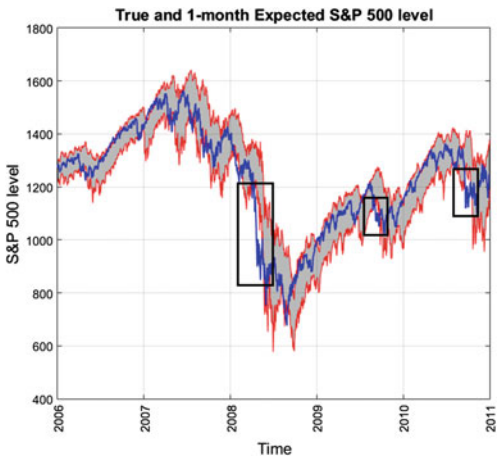

Fig. 13.9 True S&P 500 values (*blue curve*) and 30-day expected range (*red curves*) using the VIX levels, for the period 01/02/2006–01/12/2011

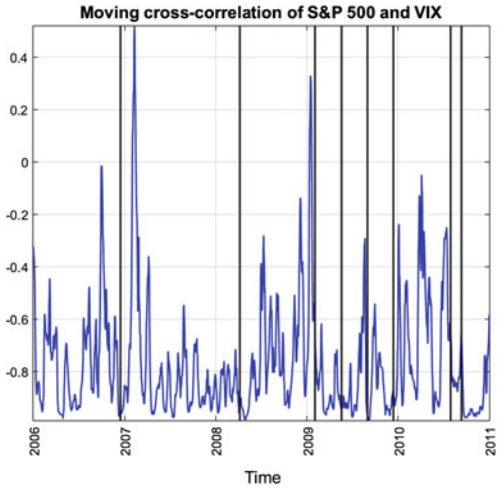

Fig. 13.10 Moving cross-correlation between S&P 500 and VIX (30-day window length. 1-day overlapping), along with the time instants (*black lines*) where the true S&P 500 values start declining from the 30-day expected range, for the period 01/02/2006–01/12/2011

that is, the LOS corresponding to the original unthresholded CRP, as follows: let $\mathbf{l} = (l_1, l_2, \ldots, l_L) \in \mathbb{R}^L$ be the estimated LOS vector. Then, a critical time, that is, a time when the expected range is exceeded, is defined as the index i for which the consecutive sample absolute difference $|l_i - l_{i-1}|$ exceeds a threshold δ, $|l_i - l_{i-1}| > \delta$ $(i = 1, \ldots, L)$. From a more intuitive perspective, the critical times are defined by those indices of the LOS vector, where "big steps" occur. In the following evaluation, and for illustration purposes only, we set $\delta = 30$ (one month). A detailed study of the optimal value of δ, which yields an optimal detection rate of the critical times, is beyond the scope of this work.

To illustrate the efficiency of our proposed approach for detecting such critical times, Fig. 13.11a shows the critical times obtained from the original LOS (black lines), along with the critical times obtained from the LOS of a restored CRP

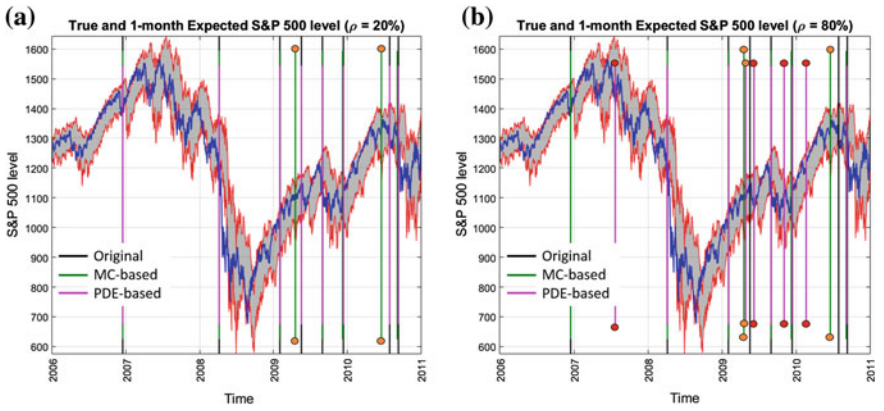

Fig. 13.11 True S&P 500 values (*blue curve*) and 30-day expected range (*red curves*), along with: (i) time instants estimated from the original LOS, where the true S&P 500 values start declining from the expected range; (ii) time instants obtained from estimated LOS using MC restoration; (iii) time instants obtained from estimated LOS using PDE restoration, for the period 01/02/2006–01/12/2011. The *orange circles* indicate detected but misaligned times, *red circles* indicate erroneous times. **a** $\rho = 20\,\%$, **b** $\rho = 80\,\%$ of missing CRP entries

using the MC-based method (green lines) and the PDE-based technique (magenta lines), when 20 % of the unthresholded CRP values are missing. For this restoration instance, we observe that both methods achieve to detect the original times, except for two misalignments (lines with orange circles) occurring for the MC-based method, which detect two critical times before their actual appearance. We note also that, for a better visualization, the three types (colors) of vertical lines are depicted with a different length.

The difference in performance between the MC-based and the PDE-based methods becomes more prominent for a higher percentage of missing CRP entries. In particular, Fig. 13.11b shows the critical times obtained from the original LOS (black lines), along with the critical times obtained from the LOS of a restored CRP using the MC-based method (green lines) and the PDE-based technique (magenta lines), when 80 % of the unthresholded CRP values are missing. Clearly, the MC-based approach achieves a high performance, in terms of detecting correctly all of the original critical times, except for two misalignments, before the mid 2009 and about the mid 2010. On the other hand, the performance of the PDE-based technique deteriorates dramatically for an increasing percentage of missing CRP values. More specifically, we can see that it fails completely to detect two of the critical times (2007, and mid 2010), whereas it generates four false critical times, as they are indicated by the magenta lines with the red circles.

We emphasize, though, that the above qualitative interpretation should be combined with a thorough study of the performance of a trading portfolio, whose future positions are adjusted according to the detected critical times. Furthermore, we also note that our present analysis is ex-post, in the sense that it is carried out by employ-

ing all the available data. To strengthen the validity of the approach proposed in this work, we have to examine the prediction capability of a (restored) LOS, in terms of detecting future critical times from the currently available data set. Both of these problems would be of great importance to investors, and are left as a separate thorough study.

As a future extension of this work, we are interested in exploiting the restored cross-recurrence profile to predict out-of-sample joint dynamics between a pair of stock market and volatility indexes. Besides, the performance of our proposed approach will be examined in terms of detecting jointly the co-movements between stock and volatility indexes based on alternative recurrence plots, such as the joint recurrence plot, in case of missing data. Furthermore, it would be also of great importance to investors a thorough study concerning the construction performance of S&P 500-based portfolios by employing the predicted interrelation between S&P 500 and VIX, expressed in terms of their estimated time-synchronization relation, as a risk premium indicator in case of missing information.

References

1. R. Mantegna, H. Stanley, *An Introduction to Econophysics* (Cambridge University Press, 2000)
2. N. Johnson, P. Jefferies, P. Ming Hui, *Financial Market Complexity* (Oxford University Press, 2003)
3. Y. Kim, C. Nelson, J. Fin. Econ. **12**(2), 307 (2014). doi:10.1093/jjfinec/nbt014
4. K. French, G. Schwert, R. Stambaugh, J. Fin. Econ. **19**(1), 3 (1987). doi:10.1016/0304-405X(87)90026-2
5. P. Giot, J. Portf. Manag. **31**(3), 92 (2005). doi:10.3905/jpm.2005.500363
6. B. Menachem, J. Shu, J. Zhang, J. Fut. Mark. **30**(9), 809 (2010). doi:10.1002/fut.20448
7. I. Vodenska, W.J. Chambers, Understanding the Relationship between VIX and the S&P 500 Index Volatility, in *Proceedings of the 26th Australasian Finance and Banking Conference* (2013)
8. R.E. Whaley, J. Portf. Manag. **35**(3), 98 (2009). doi:10.3905/JPM.2009.35.3.098
9. B. Awartani, V. Corradi, Int. J. Forecast. **21**(1), 167 (2005). doi:10.1016/j.ijforecast.2004.08.003
10. S.-M. Chiang, The Relationships between Implied Volatility Indexes and Spot Indexes, in *Proceedings of the International Conference on Asia Pacific Business Innovation and Technology Management*, Procedia—Social and Behavioral Sciences, vol. 57, p. 231 (2012)
11. S. Ercetin, S. Banerjee, *Chaos and Complexity Theory in World Politics* (IGI Global, 2014)
12. B.S. Lee, J. Bank. Finance **34**(6), 1257 (2010). doi:10.1016/j.jbankfin.2009.11.023
13. B.S. Lee, D. Ryu, Econ. E-journal **7**, (2013). doi:10.5018/economics-ejournal.ja.2013-3
14. J. Eckmann, K. Oliffson, D. Ruelle, Europhys. Lett. **4**(9), 973 (1987). doi:10.1209/0295-5075/4/9/004
15. J. Zbilut, C. Webber, J. Appl. Physiol. **76**(2), 965 (1994)
16. J. Zbilut, C. Webber, Phys. Lett. A **171**(3–4), 199 (1992). doi:10.1016/0375-9601(92)90426-M
17. L. Trulla, A. Giuliani, J. Zbilut, C. Webber, Phys. Lett. A **223**(4), 255 (1996). doi:10.1016/S0375-9601(96)00741-4
18. J. Gao, Phys. Rev. Lett. **83**(16), 3178 (1999). doi:10.1103/PhysRevLett.83.3178

19. N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, J. Kurths, Phys. Rev. E **66**(2), 026702 (2002). doi:10.1103/PhysRevE.66.026702
20. N. Marwan, J. Kurths, Phys. Lett. A **302**(5–6), 299 (2002). doi:10.1016/S0375-9601(02)01170-2
21. A. Antoniou, C. Vorlow, Neur. Net. World **10**, 131 (2000)
22. F. Strozzi, J.-M. Zaldivar, J. Zbilut, Phys. A **376**, 487 (2007). doi:10.1016/j.physa.2006.10.020
23. A. Fabretti, M. Ausloos, Int. J. Mod. Phys. C **16**(5), 671 (2005). doi:10.1142/S0129183105007492
24. P. Crowley, Eur. Phys. J. Sp. Top. **164**(1), 67 (2008). doi:10.1140/epjst/e2008-00835-3
25. N. Bigdeli, K. Afshar, Phys. A **388**, 1577 (2009). doi:10.1016/j.physa.2009.01.003
26. K. Guhathakurta, N. Marwan, B. Bhattacharya, A. Chowdhury, *Understanding the Interrelationship Between Commodity and Stock Indices Daily Movement Using ACE and Recurrence Analysis, Translational Recurrences* (Springer International Publishing, 2014). doi:10.1007/978-3-319-09531-8_13
27. P. Addo, Coupling Direction of the European Banking and Insurance Sectors using Inter-System Recurrence Networks, Documents de travail du Centre d'Economie de la Sorbonne (2015). ISSN: 1955-611X.2015
28. E. Candès, B. Recht, Found. Comput. Math. **9**(6), 717 (2009). doi:10.1007/s10208-009-9045-5
29. R. Whaley, J. Deriv. **1**(1), 71 (1993). doi:10.3905/jod.1993.407868
30. J. Noh, R. Engle, A. Kane, J. Deriv. **2**(1), 17 (1994). doi:10.3905/jod.1994.407901
31. W.-Y. Lee, C.-X. Jiang, D. Indro, J. Bank. Finance **26**(12), 2277 (2002). doi:10.1016/S0378-4266(01)00202-3
32. Standard & Poor's: S&P 500 Factsheet. Accessed 30 Nov 2015 (http://us.spindices.com/indices/equity/sp-500)
33. CBOE: The VIX White Paper, Chicago Board Options Exchange. http://www.cboe.com/micro/vix/. Accessed 30 Nov 2015
34. R. Whaley, J. Portf. Manag. **26**(3), 12 (2000). doi:10.3905/jpm.2000.319728
35. J. Iwanski, E. Bradley, Chaos **8**(4), 861 (1998). doi:10.1063/1.166372
36. N. Marwan, M. Thiel, N.R. Nowaczyk, Nonlinear Process. Geophys. **9**, 325 (2002). doi:10.5194/npg-9-325-2002
37. P. Crowley, A. Schultz, A New Approach to Analyzing Convergence and Synchronicity in Growth and Business Cycles: Cross Recurrence Plots and Quantification Analysis, Bank of Finland Research Discussion Papers 16 (2010)
38. P.M. Addo, M. Billio, D. Guégan, North Am J. Econ. Fin. **26**, 416 (2013). doi:10.1016/j.najef.2013.02.014
39. M.B. Kennel, R. Brown, H.D. Abarbanel, Phys. Rev. A **45**, 3403 (1992). doi:10.1103/PhysRevA.45.3403
40. R. Little, D. Rubin, *Statistical Analysis with Missing Data*, 2nd edn. (Wiley, 2002)
41. E. Candès, T. Tao, IEEE Trans. Inform. Theory **56**(5), 2053 (2010). doi:10.1109/TIT.2010.2044061
42. J.-F. Cai, E. Candès, Z. Shen, SIAM J. Optim. **20**(4), 1956 (2010). doi:10.1137/080738970
43. C. Peiying, W. Yuandi, A New Fourth-Order Equation Model for Image Inpainting, in *Proceedings of 6th International Conference on Fuzzy Systems and Knowledge Discovery*, Shanghai, China, 14–16 Aug 2009

Chapter 14
Time-Difference-of-Arrival Estimation Based on Cross Recurrence Plots, with Application to Underwater Acoustic Signals

Olivier Le Bot, Cédric Gervaise and Jérôme I. Mars

Abstract The estimation of the time difference of arrival (TDOA) consists of the determination of the travel-time of a wavefront between two spatially separated receivers, and it is the first step of processing systems dedicated to the identification, localization and tracking of radiating sources. This article presents a TDOA estimator based on cross recurrence plots and on recurrence quantification analysis. Six recurrence quantification analyses measures are considered for this purpose, including two new ones that we propose in this article. Simulated signals are used to study the influence of the parameters of the cross recurrence plot, such as the embedding dimension, the similarity function, and the recurrence threshold, on the reliability and effectiveness of the estimator. Finally, the proposed method is validated on real underwater acoustic data, for which the cross recurrence plot estimates correctly 77.6 % of the TDOAs, whereas the classical cross-correlation estimates correctly only 70.2 % of the TDOAs.

14.1 Introduction

Various applications make use of arrays of hydrophones to estimate the direction of arrival (DOA) of underwater acoustic sources, including military surveillance, bioacoustics, and environmental monitoring, to name but a few. Such estimations can be made through two broad classes of methods:

O. Le Bot (✉) · J.I. Mars
GIPSA-Lab, Univ. Grenoble Alpes, 38000 Grenoble, France
e-mail: lebotol@gmail.com

O. Le Bot · J.I. Mars
GIPSA-Lab, CNRS, 38000 Grenoble, France
e-mail: jerome.mars@gipsa-lab.grenoble-inp.fr

C. Gervaise
Chaire Chorus, Foundation of Grenoble INP, 46 Avenue Félix Viallet,
38031 Grenoble Cedex 1, France
e-mail: cedric.gervaise@gipsa-lab.grenoble-inp.fr

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8_14

1. Time-difference-of-arrival (TDOA) estimation.
2. Space-time processing for high-resolution estimation of the DOA, like beamforming [1] or spatial spectral estimation (Capon, MUSIC [2], ESPRIT [3, 4], and others).

The first class of methods is commonly used for bio-acoustic and environmental applications, such as the localization of cetaceans [5–8], and it generally uses a small number of hydrophones (i.e., from 2 to 10). The estimation of the TDOA, which consists of the determination of the travel-time of the wavefront between two spatially separated receivers, relies generally on the cross-correlation, which is classical and easy-to-use, even for non-specialists in signal processing. Moreover, the cross-correlation gives good performances when the signal-to-noise ratio (SNR) is high and the useful signal is slightly distorted. Thorough state-of-the-art reviews of cross-correlation-based TDOA estimators can be found in [8–10].

The second class of methods is mostly used when the array has a much greater number of receivers, and it is mostly used in applications such as SONAR [11], acoustical tomography [12], seismic-wave analysis [13], and digital communication. Methods from this class consist of the construction of a spatial spectrum by virtually steering the array in various directions and estimating the received power. When the array is steered in the direction of a source, the power received by the hydrophone array is maximized, whereas in the directions where no sources radiate, the received power is weak [1, 3, 4, 14].

In this article, we focus on the first class of methods, and more particularly, on estimation of the TDOA of acoustic signals with low SNR and strong distortions (when classical methods fail). Estimating the TDOA correctly is usually a challenging task, because underwater acoustic sources are immersed in a noisy environment and the propagation through the oceanic canal significantly modifies the signal by introducing amplitude and phase modulations. Thus, the signal received can be very different from one hydrophone to another, and only a few series of samples will remain similar on all of the hydrophones. After propagation, the received signals can be expressed as follows:

$$\begin{cases} s_1(t) = A_1(t)x(t) + n_1(t) \\ s_2(t) = A_2(t)x(t + TDOA) + n_2(t) \end{cases} \tag{14.1}$$

where $x(t)$ is the unknown source signal, $A_1(t)$ and $A_2(t)$ are random amplitude modulations introduced during the propagation between the source and each hydrophone, and $n_1(t)$ and $n_2(t)$ are uncorrelated additive white Gaussian noise (WGN).

Based on these observations, we propose a new approach to estimate the TDOA that is derived from the field of nonlinear analysis of dynamical systems. Our estimator relies on cross-recurrence plots (CRPs) [15, 16], which are used to identify series of samples that are similar within the signals received by two hydrophones. The TDOA is then estimated from the CRP using some specific measures, called recurrence quantification analysis (RQA), which are especially dedicated to quantification of the information contained in CRPs. In this article, we consider six different

RQA measures to estimate the TDOA. Four of these RQA measures come from the existing CRP literature [17, 18]. However, we will show that classical RQA measures are not appropriate for TDOA estimation. Therefore, in Sect. 14.2.2, we propose two new RQA measures that combine both the information contained in the CRPs and in the distance plots (or unthresholded CRPs). These latter RQA measures appear to be more appropriate for our purpose.

Cross recurrence plots have already been used for time-scale alignment of data series [16], to find nonlinear interrelations from bivariate time series with applications in climatology [17, 19], and to study synchronization of dynamic systems [18]. To our knowledge, this is the first time that CRPs are used with acoustic signals for TDOA estimation.

Section 14.2 recalls the basics of CRPs. Then some quantification measures are proposed to estimate the TDOA from the CRP. Section 14.3 tests the proposed method on simulated signals, and discusses the influence of the parameters involved in the CRP on the performance of the estimator. The proposed method is finally validated in Sect. 14.4 using real bio-acoustic signals recorded at sea by a hydrophone array. This validation set shows that our method gives an increase of the correctly estimated TDOAs of 7 % with respect to the classical cross-correlation estimator.

14.2 Method

14.2.1 About Cross-Recurrence Plots

Cross-recurrence plots are an extension to the bi-variate case of the recurrence plot analysis (RPA) from Eckman [20], which was later introduced by Zbilut et al. [15], and then described with a mathematical formalism by Marwan et al. [16]. CRPs are used to study the similarities and dependencies between two different complex systems, by comparisons of their respective states [16]. Three steps are involved to transform a data series from two systems to a CRP representation.

Considering two measured signals (denoted x and y) recorded on two different sensors, the first step aims at the simultaneous reconstruction of the trajectory of each signal in the same phase space using the time-delay embedding method [21, 22]. The same embedding dimension m and delay τ are used for both signals. The trajectory of each signal in the phase space domain is described by a series of phase space vectors given by:

$$\overrightarrow{x_m(i)} = [x(i),\ x(i+\tau), \ldots,\ x(i+(m-1)\tau)] \tag{14.2}$$

$$\overrightarrow{y_m(j)} = [y(j),\ y(j+\tau), \ldots,\ y(j+(m-1)\tau)] \tag{14.3}$$

where $i = \{1, 2, \ldots, N_x - (m-1)\tau\}$ and N_x are the sample index and the total number of samples of the measured signal x, respectively, and $j = \{1, 2, \ldots, N_y - (m-1)\tau\}$

and N_y are the sample index and the total number of samples of the measured signal y, respectively. To simplify the notations, we will consider that x and y have the same length $N = N_x = N_y$.

The second step consists of the measurement of the degree of similarity between the phase space vectors $\overrightarrow{x_m(i)}$ and $\overrightarrow{y_m(j)}$ associated to the two measured signals. The calculation of the similarities between all of the possible pairs of phase space vectors gives the similarity matrix with dimension $N \times N$ defined by:

$$d(i,j) = Sim\left(\overrightarrow{x_m(i)} , \overrightarrow{y_m(j)} \right) \tag{14.4}$$

where $Sim(.,.)$ is the function that is chosen to study the likeness of the phase space vectors. A lot of different mathematical functions can be used for this step [23]. The Euclidean norm is mostly used for this purpose by the recurrence plot community [18]. However, the Euclidean norm is not suitable for our application, because it is not adapted to the strong amplitude differences that exist between the noise samples and the useful signal samples. Using the Euclidean norm leads to finding similarities between vectors associated to the noise, because the noise samples generally have low amplitude, which leads to low Euclidean norm, although this also leads to finding similarities between vectors associated to the vectors of useful signals that look alike and where their Euclidean norms are also close to zero.

Instead, we prefer to use the dot product and Pearson's correlation coefficient (PCC), which are both particular cases of the cross-correlation classically used for TDOA estimation. These last two similarity functions are more appropriate to our signals because they both give high values when phase space vectors of the useful signal look alike, while they give a value close to 0 when the vectors are not alike. Therefore, when thresholding the similarity matrix obtained with these similarity functions, we only get recurrences associated to the useful signal. We have already used the dot product and PCC as a similarity function in previous studies relative to the use of RPA for detection purposes [23]. These similarity functions have already proven to be very interesting solutions when studying various kinds of simulated and real signals with RPA [23–25].

The dot-product between two phase space vectors is given by:

$$d_{i,j} = \sum_{k=1}^{m} x_k(i) \times y_k(j) \tag{14.5}$$

where $x_k(i)$ and $y_k(j)$ are the kth components of vectors $\overrightarrow{x_m(i)}$ and $\overrightarrow{y_m(j)}$, respectively.

The PCC between two phase space vectors is given by:

$$d_{i,j} = \frac{\sum_{k=1}^{m} \left(x_k(i) - \overline{x_k(i)} \right) \left(y_k(j) - \overline{y_k(j)} \right)}{\sqrt{\sum_{k=1}^{m} \left(x_k(i) - \overline{x_k(i)} \right)^2} \sqrt{\sum_{k=1}^{m} \left(y_k(j) - \overline{y_k(j)} \right)^2}} \tag{14.6}$$

where $x_k(i)$ and $y_k(j)$ are the kth components of vectors $\overrightarrow{x_m(i)}$ and $\overrightarrow{y_m(j)}$, respectively, and $\overline{x_k(i)}$ and $\overline{y_k(j)}$ are the empirical means of $\overrightarrow{x_m(i)}$ and $\overrightarrow{y_m(j)}$, respectively.

Finally, in the last step, the CRP is obtained by comparing each coefficient of the similarity matrix to a threshold ε. Therefore, the CRP is a binary matrix, where the coefficient of index (i, j) is 1 if $\overrightarrow{x_m(i)}$ and $\overrightarrow{y_m(j)}$ are considered as similar, and it is 0 otherwise. In mathematical formalism, the CRP is defined as follows:

$$CRP(i,j) = \Theta\left(Sim\left(\overrightarrow{x_m(i)}, \overrightarrow{y_m(j)}\right) - \varepsilon\right) \tag{14.7}$$

where Θ is the Heaviside function.

14.2.2 Recurrence Quantification Analysis

When calculating the CRP of signals received by two spatially separated sensors, a recurrence pattern shows up on the binary image obtained. As shown in Fig. 14.1, the recurrence pattern, which is composed of 15 parallel diagonal lines in our example, is only related to the useful signal (cosine function) and not to the noise. The noise is not supposed to produce any recurrence pattern. The position of the recurrence pattern of the useful signal depends on the TDOA. For example, in Fig. 14.1, receiver 1 is taken as the reference and three different TDOAs are simulated by translation of the signal of receiver 2 from –70 samples to +70 samples. These translations of the cosine on receiver 2 lead to horizontal translation of the associated recurrence pattern on the CRP (Fig. 14.1). If receiver 2 was taken as the reference and the cosine was translated on receiver 1, then the recurrence pattern associated to the cosine would be translated vertically on the CRP.

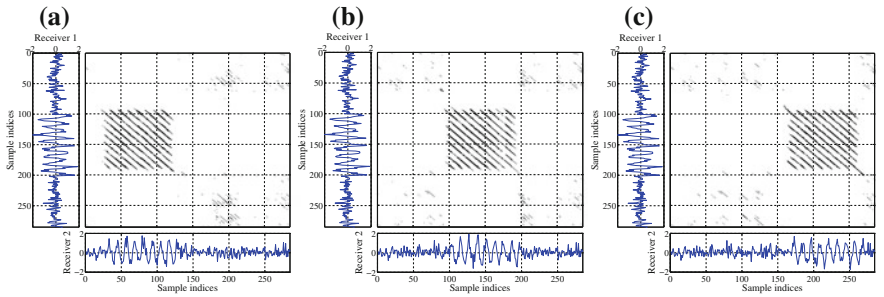

Fig. 14.1 Examples of CRP matrices in the case of a 100-sample cosine mixed with 300-sample additive WGN received by two sensors with three different TDOAs: **a** TDOA = –70 samples; **b** TDOA = 0 samples; **c** TDOA = +70 samples. The SNR is 6 dB. The CRP is built with m = 16, $\tau = 1$, $Sim(.\ ,\ .)$ = dot product, $\varepsilon = 0.075$

The automatic quantification of this pattern, like the diagonal or vertical structures, and of its position, can be performed using the measures known as RQA [26–28]. The RQA measures can be computed on either the entire CRP or separately on each diagonal parallel to the main diagonal [17, 18].

For the TDOA estimation we want to perform in this chapter, we use these latter modified RQA measures, which are calculated for each diagonal of the CRP taken separately, and are therefore based on recurrences that form diagonal lines parallel to the main diagonal [17, 18].

Relying on notations introduced in [17], the RQA measures are expressed as a function of the index of the diagonal $t \in [-T, ..., T]$, where $t = 0$ corresponds to the main diagonal, $t > 0$ are the diagonals above the main diagonal, and $t < 0$ are the diagonals below the main diagonal. T is the maximum allowable TDOA (expressed as samples) with respect to the geometry of the array. Diagonals with indices $t > 0$ represent the positive TDOAs, while diagonals with indices $t < 0$ represent the negative TDOAs.

We recall the expressions of the RQA measures from the literature [17, 18], and we introduce two new RQA measures that we consider for the TDOA estimation. The first four RQA measures from the literature only rely on the length of the diagonal lines of the CRPs, and so they are based only on the binary information contained in the CRPs. We will show that these RQA measures have some limitations and are not appropriate for our purpose. To overcome these limitations. we propose two new RQA measures, which combine the information from the CRP with that of the distance matrix, and which leads to more accurate results.

The first RQA measure is the average diagonal line length given by:

$$L(t) = \frac{\sum_{l=l_{min}}^{N-|t|} lP_t(l)}{\sum_{l=l_{min}}^{N-|t|} P_t(l)} \tag{14.8}$$

where $P_t(l)$ is the distribution of diagonal line lengths for diagonals with index t, l_{min} is the threshold that sets the length of the shortest diagonal line considered, and N is the total number of lines (resp. columns) of the CRP matrix. This RQA measure characterizes the duration of similarities between both signals [17]. Thus, for the diagonal with index t associated to the real TDOA value, we expect that the diagonal lines are long, so that we reach a global maximum for $L(t)$ that indicates that the highest coincidence between both signals is reached for this specific delay t.

The second RQA measure is the determinism given by:

$$DET(t) = \frac{\sum_{l=l_{min}}^{N-|t|} lP_t(l)}{\sum_{l=1}^{N-|t|} lP_t(l)} \tag{14.9}$$

As for the average diagonal line length $L(t)$, the determinism $DET(t)$ characterizes the proportion of recurrence points that form long diagonal structures of all of the recurrence points. A high determinism for the diagonal t indicates that both signals have similar dynamic behavior over a long time span with a certain delay t [17]. Therefore, we expect that the determinism will reach a global maximum for the diagonal with index t associated to the TDOA that we try to estimate.

The third RQA measure is the longest diagonal line within diagonal t, given by:

$$L_{max}(t) = \underset{i \in \{1,\dots N_l\}}{\arg\max}(l_i) \tag{14.10}$$

where $i \in \{1, \dots, N_l\}$ is the index of the diagonal line, and N_l is the total number of diagonal lines within diagonal t. As above, we expect to get a global maximum for $L_{max}(t)$ at the index t associated to the TDOA.

The fourth RQA measure is the recurrence rate, which is defined as:

$$RR(t) = \begin{cases} \dfrac{1}{N+t} \displaystyle\sum_{i=1}^{N+t} CRP(i-t,\ i) & (t < 0) \\ \dfrac{1}{N-t} \displaystyle\sum_{i=1}^{N-t} CRP(i,\ i+t) & (t \geq 0) \end{cases} \tag{14.11}$$

where N is the total number of lines (resp. columns) of the CRP matrix. This RQA measure gives the probability of occurrence of similar waveforms in both signals with a certain time delay t [17]. The greater the number of recurrences in the diagonal t is, the greater the value of $RR(t)$ is. Again, we expect that $RR(t)$ reaches a global maximum for the diagonal t associated to the TDOA value.

$RR(t)$ only relies on the binary matrix obtained with the CRP, and does not take into account all of the information contained in the similarity matrix $d(i, j)$ given by (14.4) that might be valuable. Therefore, we introduce two additional RQA measures that also rely on the similarity matrix $d(i, j)$.

The first proposed RQA measure is the sum of all similarity values $d(i, j)$ from the diagonal t that have led to recurrence points (i.e., $CRP(i, j) = 1$) in the CRP matrix. This is defined as:

$$SS(t) = \begin{cases} \displaystyle\sum_{i=1}^{N+t} CRP(i-t, i) \odot d(i-t, i) & (t < 0) \\ \displaystyle\sum_{i=1}^{N-t} CRP(i, i+t) \odot d(i, i+t) & (t \geq 0) \end{cases} \tag{14.12}$$

where $\odot$ is the Hadamard product.

The last proposed RQA measure is a normalized version of the previous one, and it is given by:

$$SR(t) = \begin{cases} \dfrac{1}{N+t} \displaystyle\sum_{i=1}^{N+t} CRP(i-t,i) \odot d(i-t,i) \;\; (t<0) \\ \dfrac{1}{N-t} \displaystyle\sum_{i=1}^{N-t} CRP(i,i+t) \odot d(i,i+t) \;\; (t \geq 0) \end{cases} \tag{14.13}$$

As for $RR(t)$, $SS(t)$ and $SR(t)$ are high if the number of recurrences in diagonal t and the associated coefficients in the similarity matrix are simultaneously both high. A global maximum is expected for $RR(t)$, $SS(t)$, and $SR(t)$ when t is equal to the TDOA to estimate.

14.2.3 Time-Difference-of-Arrival Estimated with RQA Measures

For the six RQA measures presented in the previous section, we said that their values would be high if both signals have similar waveforms with a certain delay t. Therefore, the estimated TDOA is obtained by looking for the diagonal t that gives the global maximum of each RQA measure. This is given by:

$$\widehat{TDOA} = t \text{ such that } \underset{t}{\operatorname{argmax}} \left(RQA(t) \right) \tag{14.14}$$

where $t \in [-T, \ldots, T]$ and $RQA(t)$ is either $L(t)$, $DET(t)$, $L_{max}(t)$, $RR(t)$, $SS(t)$ or $SR(t)$.

14.3 Results on Simulated Data

In this section, we use simulated signals to study the reliability of CRPs and RQA measures as TDOA estimators. We will show that the RQA measure $SS(t)$ is the most appropriate one. We will also discuss the influence of the embedding dimension m, of the similarity function $Sim(.,.)$, of the SNR and the TDOA value to estimate, on the overall performances of the proposed method. Finally, we will compare the proposed method with the cross-correlation, which is classically used for TDOA estimation in several fields.

14.3.1 Performance Analysis Methodology

To analyze the performances of the proposed method, we use simulated signals as they would be received independently by two sensors. The characteristics of these simulated signals are representative of the real acoustic signals used in Sect. 14.4. The received signals are cosine functions of frequency f_0, where their respective amplitudes $A_1(t)$ and $A_2(t)$ are modulated randomly according to a first-order autoregressive model. The WGN denoted as $n_1(t)$ and $n_2(t)$ is also added to the two receivers, to achieve a certain SNR. A time-delay is added to the second signal to model the TDOA. Therefore, the two simulated signals can be written as follows:

$$s_1(t) = A_1(t) \times cos(2\pi f_0 t) + n_1(t) \tag{14.15}$$

$$s_2(t) = A_2(t) \times cos\big(2\pi f_0 (t + TDOA)\big) + n_2(t) \tag{14.16}$$

We can recall that a first-order autoregressive model is given by:

$$A(t) = \alpha A(t-1) + Y(t) \tag{14.17}$$

where $|\alpha| \in [0,\ 1[$ so that the process remains stationary in a wide sense, and $Y(t)$ is WGN with zero mean and variance of 1.

For the simulations, we take $f_0 = 1$ kHz, which is the frequency of the real acoustic signal used in Sect. 14.4, and a sampling frequency $f_e = 12$ kHz. The SNR varies as follows: $\{-7;\ -5;\ -3;\ 0;\ 3;\ 5;\ 7;\ 10\}$ dB. For the autoregressive model, we arbitrarily choose $\alpha = 0.98$, so that the integration time of the process is about nine cosine periods. Moreover, the peak amplitude of the autoregressive model is normalized to the following values: $A_{peak} = \{0.4;\ 0.7;\ 1;\ 1.3;\ 1.5;\ 1.9\}$, to reach various modulation indices. Figure 14.2 shows an example of the simulated signal as received by the two sensors, with various SNRs.

We study the performances of the proposed estimator for five different theoretical TDOA values; namely [0, 4, 15, 40, 100] samples. This is to determine the reliability of the method when the TDOA corresponds to the diagonals that are either close to or far from the main diagonal of the CRP matrix, and also when the TDOA is lower or higher than the period of the cosine signal.

For each set of sextuple $\big(m,\ \varepsilon,\ Sim(.\ ,\ .),\ SNR,\ A_{peak},\ TDOA\big)$ we repeat the simulations 500 times, to derive a statistical analysis of the results. The performances are assessed by calculation of the ratio of correct estimates over the total number of estimates. An estimated TDOA is considered as correct if it is exactly equal to the simulated TDOA.

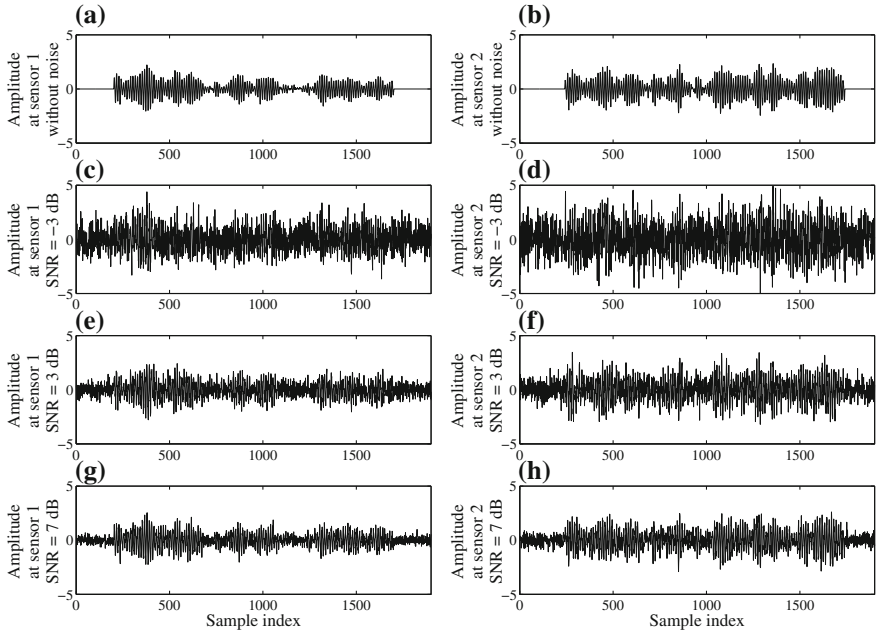

Fig. 14.2 **a**, **b** Simulated signals received by two sensors without additional noise. Each of these is a cosine function where the amplitude is modulated by a first-order auto-regressive model, the peak amplitude of which is $A_{peak} = 1.5$. **c**, **d** Simulated signal added to WGN with SNR = –3 dB. **e**, **f** Simulated signal added to WGN with SNR = 3 dB. **g**, **h** Simulated signal added to WGN with SNR = 7 dB

14.3.2 Performances of the RQA Measures

In this section, we study the relevance of the proposed RQA measures as tools for TDOA estimation. We want to identify the RQA measures that appear to be appropriate for our TDOA estimation problem, and to discard the other ones. We will show that the two proposed RQA measures outperform the classical ones.

Figure 14.3 shows the results of the simulations for three sets of parameters; namely $\tau = 1$, $m = 16$, $Sim(.\ ,\ .) = PCC$, a recurrence threshold ϵ giving a fixed recurrence point density of 10 % over the entire CRP, $A_{peak} = 1$, and $TDOA = 0$, $TDOA = 15$ and $TDOA = 40$. However, the conclusions remain the same whatever the variations of the parameters m, ϵ, SNR, $TDOA$ and A_{peak}.

We notice that the RQA measure SS (Fig. 14.3, solid line) always gives the best performances, as its percentage of correct estimates is always the highest, whatever the SNR, TDOA, similarity function Sim(. , .), and amplitude of the modulations A_{peak}.

The RQA measure SR (Fig. 14.3, line with squares) follows the same trend as SS, but with a lower performance. All of the other proposed RQA measures, i.e., L, L_{max}, DET, and RR, give very poor results and are not suitable for the TDOA estimation.

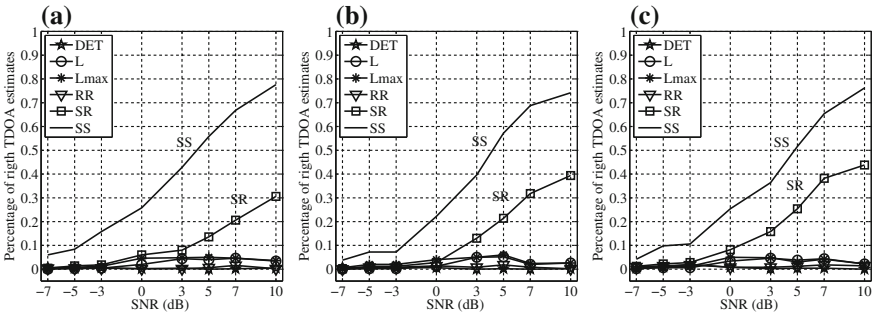

Fig. 14.3 Percentages of good estimates over the total number of estimates as a function of the SNR of the RQA measure and for three simulated TDOAs: **a** TDOA= 0; **b** TDOA = 15; **c** TDOA = 40. The parameters of the simulations are: $\tau = 1, m = 16, Sim(.\ ,\ .) = PCC, A_{peak} = 1$. The RQA measure *SS* almost always gives the highest percentages, and it is therefore the most appropriate RQA measure for our TDOA estimator

A detailed observation of the simulation results shows that these four latter RQA measures give estimates that are far from the true simulated TDOAs.

Finally here, it is important to note that the percentage of correct estimates obtained with *SS* is independent of the TDOA, which is not the case with *SR*. For example, the plots given in Fig. 14.3 show that for SNR = 7 dB the RQA measure *SS* gives a percentage of correct estimates of about 69 % (±1 %) for all three TDOA values, whereas *SR* gives a percentage of correct estimates of about 30 % (±10 %). These observations and conclusions remain valid for other simulated values of amplitude modulations A_{peak} (not presented here). Thus, *SS* gives the most consistent results for all of the sets of parameters, and it is therefore the most reliable RQA measure to estimate TDOAs.

According to these results, in the remainder of this article we only use the RQA measure *SS*, as defined by (14.12) to estimate TDOAs, and all of the results will now be given with this particular RQA measure.

14.3.3 Performances as a Function of the Parameters Used to Build the CRP

In this section, we study the influence of the similarity function $Sim(.\ ,\ .)$, of the embedding dimension m, and of the recurrence threshold ϵ used to compute the CRP on the performances of the TDOA estimator.

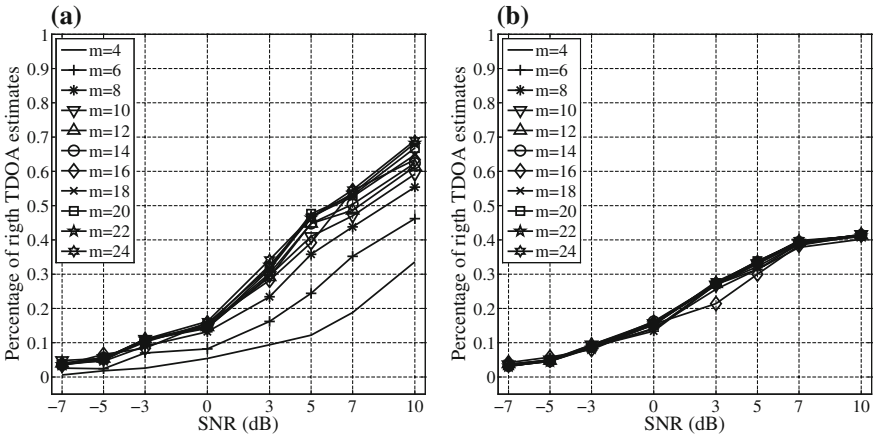

Fig. 14.4 Percentage of good estimates over the total number of estimates as a function of the embedding dimension m, of the SNR, and of the similarity function $Sim(.\ ,\ .)$: **a** $Sim(.\ ,\ .) = PCC$; **b** $Sim(.\ ,\ .) = Dot\ product$. The parameters of the simulations are: $\tau = 1, A_{peak} = 1.5, TDOA = 100$ and RQA measure SS

14.3.3.1 Influence of the Embedding M

Simulations are performed for numerous embedding dimension values within the range of 4–24, while keeping $\tau = 1$ constant, and ε so that the recurrence point density over the entire CRP is always 10 %. Figure 14.4 summarizes the results of these simulations for a given set of parameter: ($A_{peak} = 1.5$, $TDOA = 100$, $RQA = SS$). The conclusions remain similar for all of the other sets of parameters, whatever the values given to A_{peak} and $TDOA$.

We can see in Fig. 14.4a that if the similarity function is the PCC, the percentage of correct estimates increases when m increases. When m increases from 4 to 10 in particular, the percentage of correct estimates is multiplied by three on average (which depends on the SNR), while when m increases from 12 to 24, the percentages of correct estimates increases by just a few percent. For example, from $m = 12$ to $m = 24$, the percentage of correct estimates increases only by about 15 %, whatever the SNR.

If the similarity function is the dot product, then the percentage of correct estimates is very similar for all of the embedding dimensions (Fig. 14.4b).

14.3.3.2 Influence of the Similarity Function

For this subsection, we keep the parameters τ, m, and ε used to build the CRP constant, and we study the influence of the similarity function on the performances of the proposed estimator. Simulations are carried out for various combinations of TDOA and of amplitude modulations A_{peak}. The most significant results are shown in Fig. 14.5.

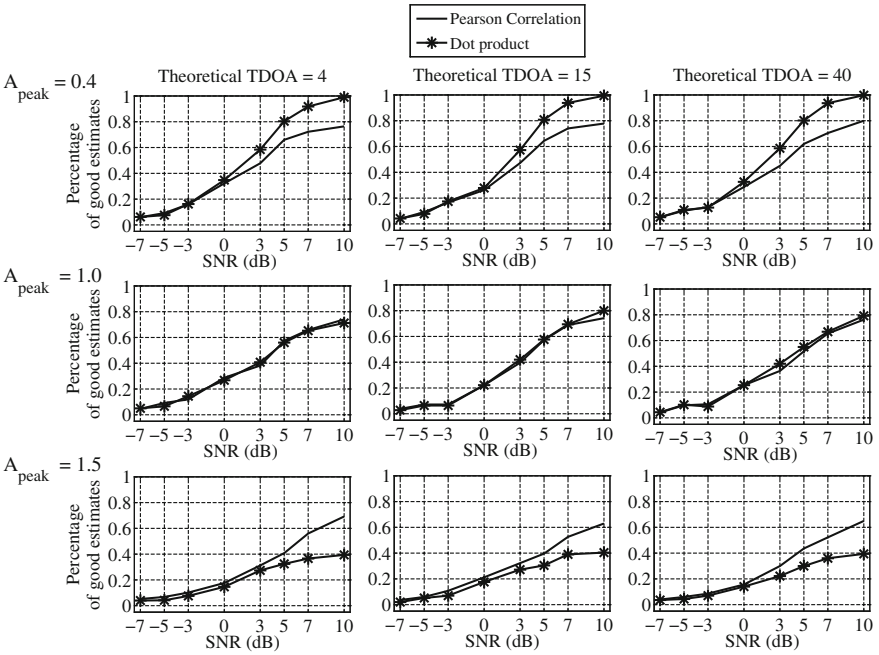

Fig. 14.5 Percentage of correct estimates over the total number of estimates as a function of the SNR, of the TDOA (each column of the panel), and of the amplitude of the modulation A_{peak} (each line of the panel). The results with $Sim(.\ ,\ .) = PCC$ are shown with the *solid line*. The results with $Sim(.\ ,\ .) = Dot\ Product$ are shown with the stars (*). The parameters of the simulations are: $\tau = 1$, $m = 16$, and RQA measure *SS*

We note that when $A_{peak} = 0.4$ (Fig. 14.5, first row of the panel), the estimator using the dot product always gives a better percentage of correct estimates than the estimator with the PCC, whatever the SNR and the TDOA. When $A_{peak} = 1$ (Fig. 14.5, second row of the panel), both similarity functions give very similar results, whatever the SNR and the TDOA. When $A_{peak} = 1.5$ (Fig. 14.5, third row of the panel), the estimator with the PCC has better performances than the one with the dot product.

By examination of the results of the simulations for all six simulated values of amplitude modulation, we can conclude that when the signals are weakly modulated, typically $A_{peak} < 1$, then the estimator using the dot product always outperforms the estimator with the PCC. Then, as the signal is more and more modulated, i.e., the A_{peak} increases, the performances of the estimator using the dot product decreases. Similar performances are obtained with both estimators when $A_{peak} = 1$. Finally, for highly modulated signals with a lot of fading (i.e., $A_{peak} > 1$), the estimator using PCC always outperforms the estimator with the dot product.

Whatever the similarity function used, when A_{peak} increases, the percentage of correct estimates decreases. However, the rate at which performances decrease is

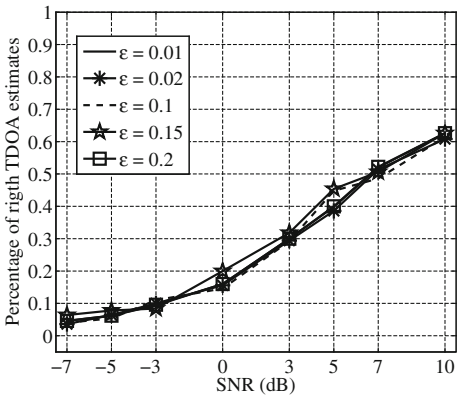

Fig. 14.6 Percentages of correct estimates over the total number of estimates as a function of the SNR and the recurrence densities: 1, 2, 10, 15 and 20 %. The parameters of the simulations are: $\tau = 1$, $m = 12$, $A_{peak} = 1.5$, $TDOA = 40$, $Sim(.,.) = PCC$, and RQA measure SS

not the same for all of the estimators. If the dot product is used, then for all SNRs the percentages of correct estimates are divided by a factor of about 2.5 when A_{peak} goes from 0.4 to 1.5. At the same time, the percentages of correct estimates are divided by a factor of at most 1.5 if the estimator uses the PCC.

Relying on these observations and conclusions, we will choose the PCC as a similarity function to estimate the TDOA on the real acoustic data used in Sect. 14.4.

14.3.3.3 Influence of the Recurrence Threshold

The choice of the recurrence threshold affects the number of recurrence points that appear on the CRP. Simulations are made by keeping τ and m constant, while the values of ϵ are adjusted so that they give a fixed recurrence point density over the entire CRP. Performances are studied for recurrence thresholds with the following recurrence densities: 1, 2, 5, 8, 10, 12, 15, 20 and 25 %. These simulations show that for a given set of values for the quintuple (τ, m, A_{peak}, $TDOA$, SNR), the percentage of correct estimates is nearly the same for all of the recurrence densities used. Figure 14.6 shows an example of these results when $\tau = 1$, $m = 12$, $A_{peak} = 1.5$, $TDOA = 40$, $SNR \in [-7\ ; 10]$ dB. Therefore, we can conclude that our method is invariant with regard to the recurrence threshold.

14.3.4 Comparison with the Classical Cross-Correlation

In this last section, we compare the performances of the proposed estimator with an estimator based on the cross-correlation, which is classically used for this purpose. Unlike the CRP, which divides the signal into multiple smaller pieces that are represented by the phase space vectors to estimate the TDOA, the cross-correlation uses the whole signal at once and the estimated TDOA is then the lag associated to the

global maximum of the cross-correlation function. All of the previous simulations were performed with the cross-correlation as the TDOA estimator.

The results obtained here are shown in Fig. 14.7, and they are compared with those of the proposed estimator. If the received signals are weakly modulated, then the cross-correlation outperforms the CRP-based estimator with the PCC, whereas this latter outperforms the cross-correlation when signals are highly modulated. For a given TDOA, we note that the results of the cross-correlation are very dependent of the amplitude of the modulations, while for a given amplitude of the modulation, the results stay consistent for all simulated TDOAs. Comparing Figs. 14.5 and 14.7, we note that for a given triplet (SNR, A_{peak}, $TDOA$), the cross-correlation gives exactly the same percentage of correct estimates as the CRP with the dot product as a similarity function.

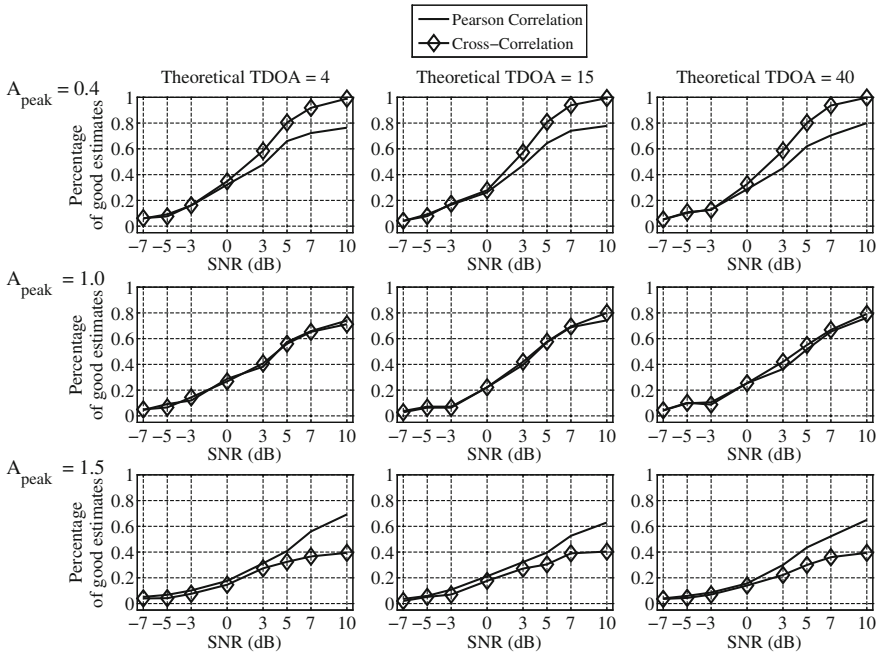

Fig. 14.7 Percentages of correct estimates over the total number of estimates as a function of the SNR, of the TDOA (each column of the panel), and of the amplitude of the modulation A_{peak} (each line of the panel). The results of the proposed method with similarity function $Sim(.\ ,\ .) = PCC$ are given by the *solid line*. The results obtained with classical cross-correlation are shown with the diamonds ($\diamond$). The parameters of the simulations are: $\tau = 1$, $m = 16$, and RQA measure SS

14.4 Validation on Real Data

In the previous subsection, we studied the performances of a CRP-based TDOA estimator and looked at the influence of the parameters used to built a CRP representation using simulated signals. This helped us to define *SS* as the most appropriate RQA measure to estimate the TDOA. We also showed that the PCC is the similarity function that gives the most consistent results for various conditions of the SNR and the modulation of the received signal. In this section, we validate the proposed TDOA estimator on real acoustic signals, and show again that it outperforms the classical cross-correlation.

14.4.1 Materials and Methods

The database used in this section contains underwater sounds that were recorded in the Bay of Calvi (Corsica Island, France) with the support of STARESO team and facilities in the framework of STARECAPMED program. The data were acquired by an autonomous acoustic recorder (RTSys EA-SDA14) that recorded the signals coming from four hydrophones (HTI-92, High Tech Inc., Gulport, MS, USA) simultaneously, at a sampling frequency of 156.25 kHz. The data were digitized at 24 bits, and saved as WAV files on a hard drive. The four hydrophones form an array with a pyramidal shape. The distance between each hydrophone pair was about 1.5 m.

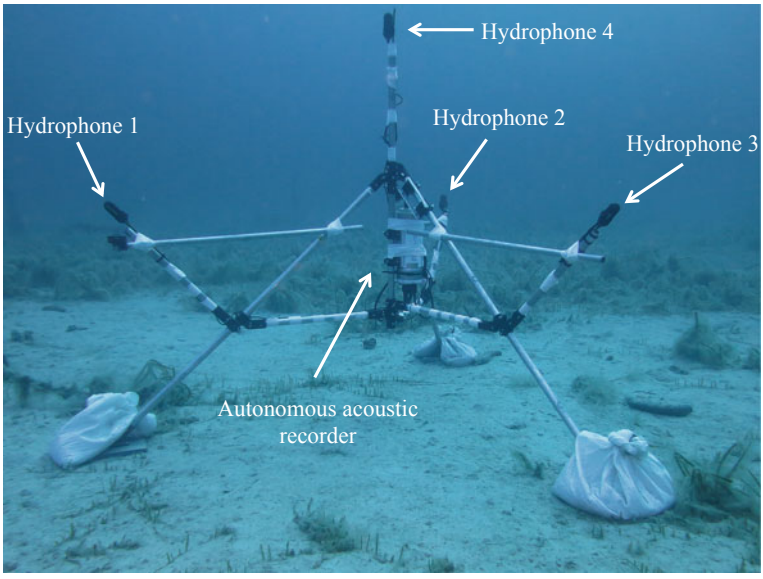

Fig. 14.8 Set-up of the antenna used to record the data. Four hydrophones are connected to four synchronized channels of an autonomous acoustic recorder

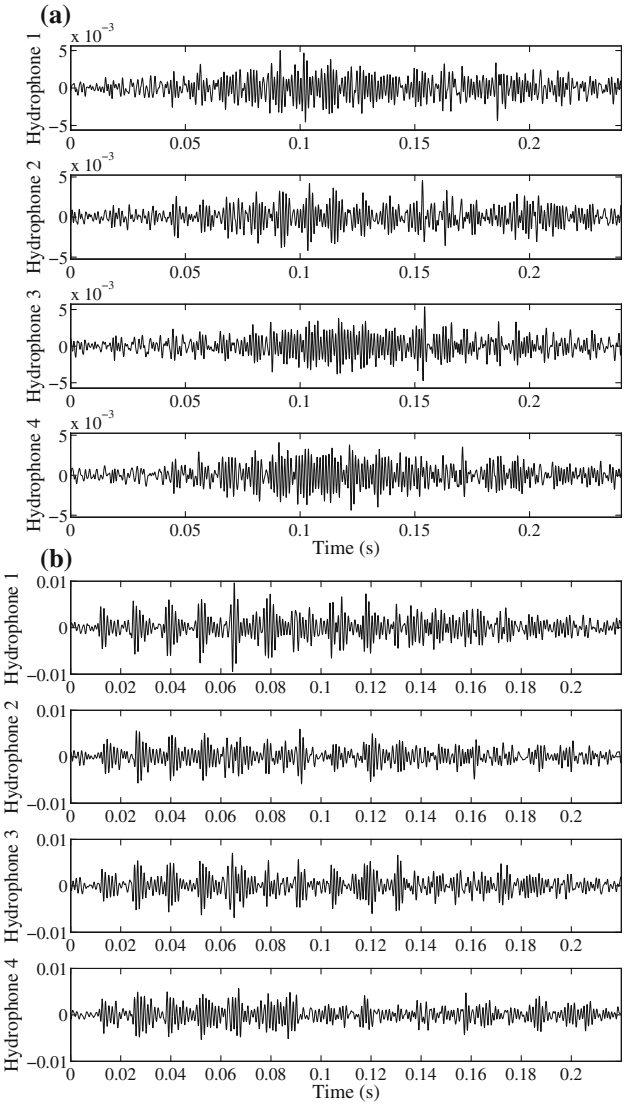

Fig. 14.9 Examples of two distinct recordings of the sounds produced by a fish (undetermined species) and recorded simultaneously by four hydrophones. **a** Record 1. **b** Record 2

The antenna was placed on the seafloor at a depth of 38 m and the hydrophones were between 1 and 1.5 m above the seafloor (Fig. 14.8).

The sounds of interest come from a Mediterranean fish of undetermined species, and they are short frequency modulated signals that last about 0.15 s and have a mean frequency of between 800 and 1000 Hz. Figure 14.9 shows two examples of these acoustic signals, from which we can clearly note that the amplitude modulations

differ significantly from one hydrophone to another due to the effects of the propagation in the seawater. We manually annotated 56 of these sounds and then estimated the TDOAs between all of the possible pairs of hydrophones. As four hydrophones form six separate pairs, we obtained a total of 336 estimated TDOAs to validate the proposed TDOA estimator.

Prior to the TDOA estimation by our CRP-based method, the annotated sounds were down-sampled at 10 kHz to reduce the size of the CRP matrix and to speed up the algorithm, and then they were bandpass filtered between 100 and 2000 Hz. The CRP was then computed with the following parameters: $\tau = 1$, $m = 18$, $Sim(.\ ,\ .) =$ PCC, and a fixed threshold $\varepsilon = 0.75$. The TDOAs were then estimated with the RQA measure *SS* given by (14.12). The TDOAs were also estimated with classical cross-correlation, and the results were compared to those of the CRP-based method, to determine whether this latter was more accurate and efficient.

14.4.2 Methodology to Assess the Results on Real Data

When using real data, we face the problem of knowing the ground truth to validate our results. To assess the validity of the estimated TDOAs given by our method and by the cross-correlation, we used both automatic and visual checking.

The automatic checking was based on the idea that each annotated signal gives six TDOA values, which are not independent of each other, but must respect transitive relations. For example, the TDOA between hydrophones 1 and 4 must be equal to the sum of the TDOA between hydrophones 1 and 2 and hydrophones 2 and 4. Checking all of the possible transitive relations helped us to identify possible mis-estimated TDOAs. Details of the method we followed for this automatic checking can be found in the Appendix.

The visual inspection was performed by superimposing the waveform received by each pair of hydrophones and compensating for the TDOA between both signals with its estimated value. Because of the periodic nature of the received signals, we could quickly assess whether both signals were re-aligned correctly or whether the estimated TDOA was biased by a number of samples that are proportional to the oscillation period of the signal.

As an example, Fig. 14.10a shows the signals received simultaneously by two hydrophones without compensating for the TDOA. It appears that these signals are not synchronized, as their respective waveforms are not superimposed. Figure 14.10b shows the same signals after the TDOA between the two hydrophones was compensated by the estimated value obtained with the proposed method (i.e., CRP with CPP and SS). We note that the TDOA is estimated correctly, as both of the waveforms are perfectly superimposed and oscillate similarly. On the contrary, the cross-correlation misestimated the TDOA by a number of samples equivalent to one oscillation period (Fig. 14.10c). Despite both signals oscillating the same way and at the same speed, we see that they are not perfectly superimposed. Therefore, this visual inspection

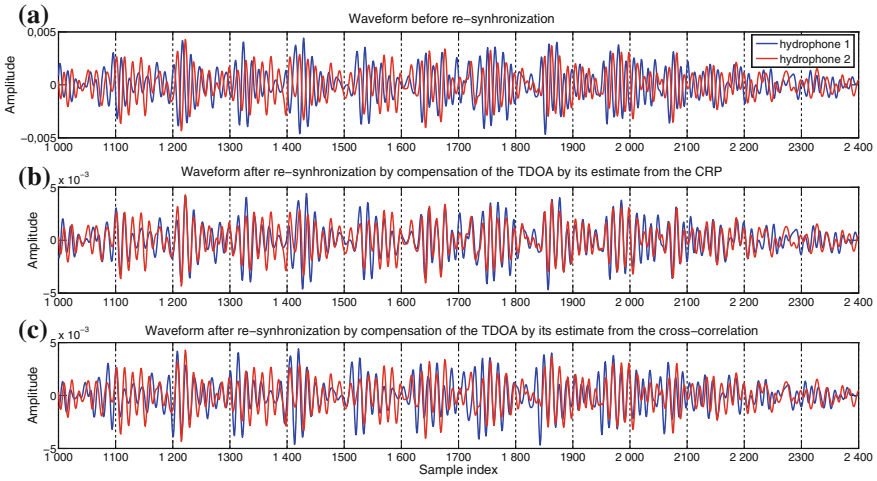

Fig. 14.10 Examples of waveforms received simultaneously by two distinct hydrophones. The signal from hydrophone 1 is shown in *blue*, and the signal from hydrophone 2 is shown in *red*. **a** Before the signals were re-synchronized. **b** After the signals were re-synchronized by compensating for the TDOA by the estimated value obtained with the CRP. **c** After the signals were re-synchronized by compensating for the TDOA by the estimated value obtained with the cross-correlation. We note that for this particular example, the CRP estimated the TDOA well, whereas the cross-correlation failed to estimate the TDOA correctly

allowed validation of the estimated value obtained with the CRP, while discarding that obtained with the cross-correlation.

For most of the annotated sounds, checking simultaneously all of the transitive relations and the alignment of the waveform after compensation of the TDOAs helped to determine the validity of the estimated TDOAs.

14.4.3 Results

Among the 56 annotated sounds, the CRP-based method succeeded in the estimation of all six TDOAs for 18 of them, which represents a success rate of 32.1 %. The classical cross-correlation succeeded in the estimation of all six TDOAs of only 10 annotated sounds, which represents a 17.6 % success rate. This first observation suggests that the proposed TDOA estimator gives better performances than that using the cross-correlation. This apparent low success rate obtained for the estimation of all of the TDOAs of singular annotated sounds can be explained according to the two following reasons: First, the periodic nature of the signal makes it very likely to have mis-estimated the TDOA values. The error is then proportional to the signal period. Secondly, the underwater environment is a highly fluctuating and noisy environment, which heavily affects the form of the acoustic signals received by the

hydrophones of the antenna, even when the distance between these hydrophones is very small. Therefore, despite the apparent good SNR of the annotated signals, the waveforms recorded are very different from one hydrophone to another, as shown on Fig. 14.9, which makes the TDOA estimation difficult.

Consequently, 38 of the annotated sounds had at least one TDOA misestimated by the CRP-based estimator, and 46 sounds had at least one TDOA misestimated by the cross-correlation. For these sounds, the results for the transitive relations were studied closely, and visual checking was used to find the misestimated TDOA values.

The overall results of this analysis include all 336 estimated TDOAs from the 56 annotated sounds, which show that 77.6 % of the TDOAs were correctly estimated by the CRP, while 70.2 % of the TDOAs were correctly estimated by the cross-correlation.

Therefore, the proposed CRP method provides a gain of 7.4 % for the number of correctly estimated TDOAs, with respect to the cross-correlation. Thus, these results validate the use of the CRP as a tool to estimate the TDOA of signals received by spatially separated sensors.

14.5 Conclusion

In this article, we proposed a method based on CRPs and on dedicated RQA measures to estimate the TDOA of signals that arrive at spatially separated sensors.

Instead of computing the RQA measures on the whole CRP, we computed them on each diagonal parallel to the main diagonal to deduce the TDOA. We used four existing RQA measures from the literature, and we proposed two new RQA measures. Among the investigated RQA measures, only the proposed RQA measure *SS*, which gives the sum of all similarity values that lead to recurrence points in a given diagonal of the CRP matrix, appeared to be appropriate and reliable to correctly estimate the TDOAs.

The proposed TDOA estimator was tested and validated on simulated and real data. The simulated data helped to confirm the reliability and the effectiveness of the proposed method with controlled data for which the true TDOA to be estimated was known. Also, these simulations let us study and understand the influence of the parameters, such as the embedding dimension m, the similarity function $Sim(.\,,\,.)$, and the RQA measures, on the performances of the proposed method. In particular, the PCC appeared to be the most appropriate similarity function for our purpose. Moreover, the simulated data were also used to compare the CRP-based TDOA estimator with a cross-correlation based estimator. This comparison showed that for signals with weak amplitude modulations, the cross-correlation was better than the proposed method (from 0 to 25 % higher, depending on the parameters of the simulations and the SNR), whereas for signals with strong amplitude modulations, the CRP using the PCC gave a higher number of correctly estimated TDOAs than the cross-correlation (0–30 % higher depending on the parameters of the simulations and the SNR).

All mis-estimated TDOAs obtained with the proposed method were due to a shift of samples that was proportional to the oscillation period of the signal. The periodicity of the simulated and real signals gave several parallel lines on the CRP, and the distance between these diagonals corresponded exactly to the oscillation period of the signal. Due to the high sensitivity of the proposed method, most of the time we can obtain parts or all of these diagonals, even when the signal is highly modulated or has a poor SNR. However, we then face the problem of picking up the diagonal that corresponds to the true TDOA, with an ambiguity factor related to the oscillation period of the signal. This constitutes the major challenge and the potential downfall faced when using the CRP as a TDOA estimator.

Also, these results appear to be important in the choice and application of a TDOA estimation method on real data. For signals with weak amplitude modulations, the classical cross-correlation or the CRP with the dot product as a similarity function might be sufficient to achieve good performances. On the contrary, more distorted signals would require the use of CRP with PCC to achieve good performances. In the case of underwater acoustics, these concepts of weak and strong amplitude modulations can be related to three different acoustic phenomena: (1) controlled variations in the amplitude of the sound during its emission by the animal; (2) amplitude modulations introduced during the propagation due to the reflections and scattering at the boundaries, the frequency-dependent attenuation in the sea, and the distance between the source and receiver; and (3) the directivity of the sound source that affects the waveform 'viewed' by the different hydrophones.

The proposed method was validated on real acoustic data recorded at sea by a four-hydrophone array. As with the simulated data, we showed that the CRP-based TDOA estimator gave better performances than the cross-correlation, with a global score of 77 % of correct estimates for the CRP, as a gain of 7.4 % over the number of correctly estimated TDOAs with respect to the cross-correlation

Acknowledgments The authors would like to thank the DGA for supporting the postdoctoral scholarship of O. Le Bot, the Water Agency of Rhone-Mediterranean-Corsica for supporting the project SEAcoustic during which the acoustic data were recorded, the research team STARESO based in Calvi, and Julie Lossent for technical support during the recording of the data in the Bay of Calvi (France).

Appendix

This appendix deals with the automatic controls performed to determine the validity of TDOA estimated on real acoustic data.

This automatic checking is based on the concept that each annotated signal gives six TDOA values, which are not independent of each other, but which must respect transitive relations. A four-hydrophone array gives four independent transitive relations between all of the TDOA values (Fig. 14.11). Let $TDOA_{ij}$ be the TDOA between hydrophones i and j, with $(i, j) \in \{1, 2, 3, 4\}$ the index of the hydrophones. The four

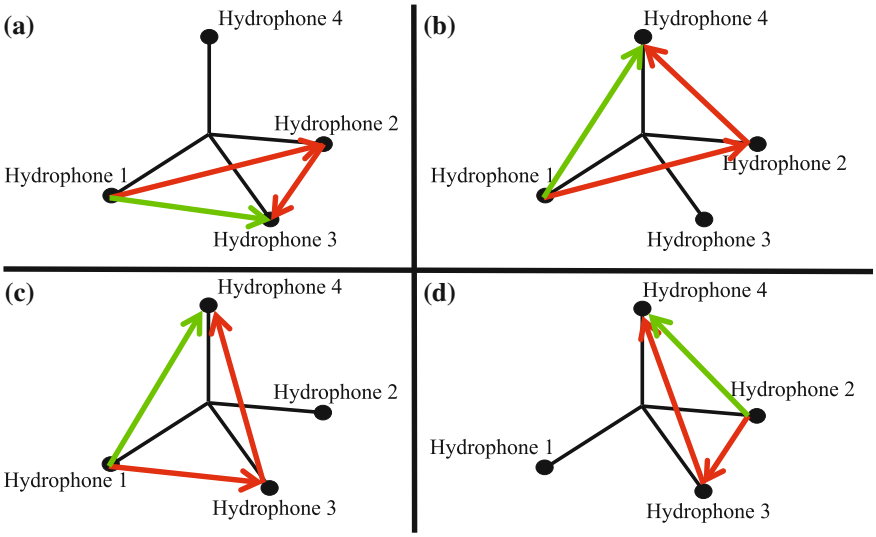

Fig. 14.11 Four examples of independent transitive relations of the TDOAs given by (14.18): **a** $TDOA_{12} + TDOA_{23} = TDOA_{13}$; **b** $TDOA_{12} + TDOA_{24} = TDOA_{14}$; **c** $TDOA_{13} + TDOA_{34} = TDOA_{14}$; **d** $TDOA_{23} + TDOA_{34} = TDOA_{24}$

independent transitive relations between the estimated TDOA (expressed as samples) are given by:

$$\begin{cases} TDOA_{12} + TDOA_{23} = TDOA_{13} \pm 2 \\ TDOA_{12} + TDOA_{24} = TDOA_{14} \pm 2 \\ TDOA_{13} + TDOA_{34} = TDOA_{14} \pm 2 \\ TDOA_{23} + TDOA_{34} = TDOA_{24} \pm 2 \end{cases} \tag{14.18}$$

As each TDOA was estimated with 1-sample precision (equiv. 10^{-4} s), the sum of two TDOAs is estimated with 2-sample precision (equiv. 2×10^{-4} s).

If the four relations given in (14.18) are met, then we can assume that all of the estimated TDOAs are true. Visual inspection is then performed to confirm the validity of the estimates, and particularly to check whether there are any cumulative errors within the relations; i.e., if two or more TDOAs are misestimated but their sum respects the transitive relations.

If one of the transitive relations from (14.18) is not met, we check the transitive relations of the third orders, to determine which TDOA is true and which TDOA is misestimated. The TDOA values give three independent transitive third-order relations, which are given by:

$$\begin{cases} TDOA_{13} + TDOA_{34} - TDOA_{24} = TDOA_{12} \pm 3 \\ TDOA_{12} + TDOA_{23} + TDOA_{34} = TDOA_{14} \pm 3 \\ TDOA_{13} - TDOA_{23} + TDOA_{24} = TDOA_{14} \pm 3 \end{cases} \tag{14.19}$$

For the relations in (14.19), the sum of three TDOAs is estimated with 3-samples precision. If one of the relations in (14.19) is verified, then the four associated TDOA values are assumed to be true and a visual inspection is performed to confirm the results. Again the visual inspection aims to prevent cumulative errors.

Finally, for each transitive relation from (14.19) taken individually, two of the estimated TDOAs are not involved. Either these two estimated TDOAs are false, or only one of them is false. To determine which of these two TDOAs was effectively misestimated, we looked back at the transitive relations given in (14.18) by assuming that the four estimated TDOAs involved in the third-order transitive relations have already been validated.

For example, if the relation $TDOA_{12} + TDOA_{23} + TDOA_{34} = TDOA_{14} \pm 3$ is met, we assume that $TDOA_{12}$, $TDOA_{23}$, $TDOA_{34}$ and $TDOA_{14}$ are true. In this transitive relation, we note that $TDOA_{13}$ and $TDOA_{24}$ have not been used. Then, to check whether $TDOA_{13}$ was well estimated, we look to see whether the following relations are also true: $TDOA_{13} = TDOA_{12} + TDOA_{23} \pm 2$ and $TDOA_{13} = TDOA_{14} - TDOA_{34} \pm 2$. If these are both true, it means that $TDOA_{13}$ was correctly estimated. We can do the same for $TDOA_{24}$ by checking whether $TDOA_{24} = TDOA_{14} - TDOA_{12} \pm 2$ and $TDOA_{24} = TDOA_{23} + TDOA_{34} \pm 2$ are satisfied. We repeat this reasoning for the other two transitive third-order relations.

References

1. B.D. Van Veen, K.M. Buckley, Beamforming: a versatile approach to spatial filtering. IEEE ASSP Mag. **5**(2), 4–24 (1988)
2. R.O. Schmidt, A signal subspace approach to multiple emitter location and spectral estimation, Ph.D. Dissertation, Stanford University, Stanford, CA (1981)
3. R. Roy, T. Kailath, ESPRIT-estimation of signal parameters via rotational invariance techniques. IEEE Trans. Acoust. Speech Signal Process. **37**(7), 984–995 (1989)
4. K.T. Wong, M.D. Zoltowski, Uni-vector-sensor ESPRIT for multisource azimuth, elevation, and polarization estimation. IEEE Trans. Antennas Propag. **45**(10), 1467–1474 (1997)
5. P.R. White, T.G. Leighton, D.C. Finfer, C. Powles, O.N. Baumann, Localisation of sperm whales using bottom-mounted sensors. Appl. Acoust. **67**(11–12), 1074–90 (2006)
6. P. Giraudet, H. Glotin, Real-time 3D tracking of whales by echo-robust precise TDOA estimates with a widely-spaced hydrophone array. Appl. Acoust. **67**(11–12), 1106–1117 (2006)
7. Y. Simard, N. Roy, Detection and localization of blue and fin whales from large-aperture autonomous hydrophone arrays: a case study from the St. Lawrence estuary. Can. Acoust. **36**(1), 104–110 (2008)
8. L. Houégnigan, S. Zaugg, M. van der Schaar, M. André, Space-time and hybrid algorithms for the passive acoustic localisation of sperm whales and vessels. Appl. Acoust. **71**(11), 1000–1010 (2010)
9. C. Knapp, G.G. Carter, The generalized correlation method for estimation of time delay. IEEE Trans. Acoust. Speech Signal Process. **24**(4), 320–327 (1976)

10. J. Chen, J. Benesty, Y. Huang, Time delay estimation in room acoustic environments: an overview. EURASIP J. Appl. Signal Process. 1–19 (2006)
11. W.C. Knight, R.G. Pridham, S.M. Kay, Digital signal processing for sonar. Proc. IEEE **69**(11), 1451–1506 (1981)
12. G. Le Touzé, B. Nicolas, J.I. Mars, P. Roux, B. Oudompheng, Double-Capon and double-MUSICAL for arrival separation and observable estimation in an acoustic waveguide. EURASIP J. Adv. Signal Process. **2012**(1), 1–13 (2012)
13. M. Alam, J.H. McClellan, W.R. Scott Jr., Spectrum analysis of seismic surface waves and its applications in seismic landmine detection. J. Acoust. Soc. Am. **121**(3), 1499–1509 (2007)
14. H. Krim, M. Viberg, Two decades of array signal processing research: the parametric approach. IEEE Signal Process. Mag. **13**(4), 67–94 (1996)
15. J.P. Zbilut, A. Giuliani, C.L. Webber, Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification. Phys. Lett. A **246**(1–2), 122–128 (1998)
16. N. Marwan, M. Thiel, N.K. Nowaczyk, Cross recurrence plot based synchronization of time series. Nonlinear Process. Geophys. **9**(3–4), 325–331 (2002)
17. N. Marwan, J. Kurths, Nonlinear analysis of bivariate data with cross recurrence plots. Phys. Lett. A **302**(5–6), 299–307 (2002)
18. N. Marwan, M.C. Romano, M. Thiel, J. Kurths, Recurrence plots for the analysis of complex systems. Phys. Rep. **438**(5), 237–329 (2007)
19. N. Marwan, M.H. Trauth, M. Vuille, J. Kurths, Comparing modern and Pleistocene ENSO-like influences in NW Argentina using nonlinear time series analysis methods. Clim. Dyn. **21**(3–4), 317–326 (2003)
20. J.P. Eckmann, S.O. Kamphorst, D. Ruelle, A new graphical tool for measuring the time constancy of dynamical systems is presented and illustrated with typical examples. Europhys. Lett. **4**(91), 973–977 (1987)
21. N.H. Packard, J.P. Crutchfield, J.D. Farmer, R.S. Shaw, Geometry from a time series. Phys. Rev. Lett. **45**(9), 712–716 (1980)
22. F. Taken, *Detecting Strange Attractors in Turbulence*. Dynamical Systems and Turbulence, Lecture Notes in Mathematics, vol. 898, pp. 366–381 (1981)
23. O. Le Bot, C. Gervaise, J.I. Mars, Similarity matrix analysis and divergence measures for statistical detection of unknown deterministic signals hidden in additive noise. Phys. Lett. A **379**(40–41), 2597–2609 (2015)
24. F.M. Birleanu, C. Ioana, C. Gervaise, A. Serbanescu, J. Chanussot, Jocelyn: Caractérisation des signaux transitoires par l'analyse des récurrences de phase. XXIIIème colloque GRETSI (2011)
25. O. Le Bot: Détection, localisation, caractérisation de transitoires acoustiques sous-marins. Thése de l'Université de Grenoble (2014)
26. J.P. Zbilut, C.L. Webber, Embeddings and delays as derived from quantification of recurrence plots. Phys. Lett. A **171**(3), 199–203 (1992)
27. C.L. Webber, J.P. Zbilut, Dynamical assessment of physiological systems and states using recurrence plot strategies. J. Appl. Physiol. **76**(2), 965–973 (1994)
28. N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, J. Kurths, Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. Phys. Rev. E **66**(2), 026702 (2002)

Chapter 15
Reservoir-Induced Changes in Dynamics and Synchrony of River Water Temperatures Revealed by RQA and CRQA

Mariola Kędra

Abstract River water temperature is a key environmental determinant for aquatic biota. A range of and dynamic changes in temperature values through different temporal scales from diurnal to annual ones determine conditions of life and survival for fish and other aquatic organisms. As a result of heat exchange with the environment (mainly at the air-water boundary), river water temperature largely follows changes in local weather conditions and is related to the dynamics of local air temperature. A retention reservoir operating in a course of river can alter the dynamics of water temperature and the specific air-water relationship for a given site below that reservoir. But to what extent? Water and air temperatures measured some distance below a group of two reservoirs functioning on a mountain river were studied in two separate time periods—occurring before and after the reservoirs' construction. Recurrence and cross-recurrence quantification analyses revealed that the reservoirs in question have significantly altered the temperature dynamics of water flowing out of them compared to the pre-dam period, and caused the weakening of synchronization between the air and water temperatures studied.

15.1 Introduction

River water temperature is a key environmental factor for aquatic ecosystems. Water temperature directly affects the survival, growth, and reproduction of aquatic biota as well their distribution within a river environment [1, 2]. Water temperature at a given site in a river system is an indicator of a river's thermal budget. It is

M. Kędra (✉)
Department of Environmental Engineering, Institute of Engineering and Water Management, Cracow University of Technology, Warszawska 24, 31-155 Cracow, Poland
e-mail: Mariola.Kedra@iigw.pk.edu.pl

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8_15

determined by the energy transport processes in a river and heat exchange between the river and its surroundings, that is, heat fluxes across the air-water interface and at the streambed [1]. Small streams have less capacity for heat storage than large rivers [3]. Water temperatures in streams are very dynamic and highly dependent on atmospheric conditions [1], and physical characteristics of streams [4], including bank-side vegetation and surrounding topography [3]. Temporal patterns of air and water temperature are often highly correlated [4], and the stream-air temperature relationship resembles an S-shaped function [5]. It was shown [6], that river water temperatures follow air temperatures closely with some lag; the lag time ranged from hours to days, increasing with stream depth. Other factors affecting water temperature include solar radiation, wind speed, relative humidity, water depth, ground water inflow, thermal conductivity of sediments, and artificial heat inputs [6]. As a result, water temperatures are less sensitive to air temperature variations at sites strongly affected by groundwater, river regulation, reservoir and waste water releases, lake outflows, and shading by riparian vegetation [5, 7–10].

Dam reservoirs built on rivers provide numerous socio-economic benefits, but at the expense of the ecological integrity of rivers [11]. The impoundments markedly affect river hydrology, ultimately producing a hydrologic pattern of flow differing significantly from a pre-impoundment natural flow regime [12]. Thermal stratification develops in river water impounded by dam reservoirs, with temperature-explained density differences; and in temperate areas deeper reservoirs (mean depth 20 m) become dimictic, but shallower reservoirs or those with enhanced flow turbulence become polymictic [13]. Reservoir releases from the surface layer (epilimnion) usually closely follow air temperature, as opposed to releases from the bottom layer (hypolimnion) [14]. Therefore, depending on the time of the year, a water body with seasonal thermal stratification causes considerable changes in the thermal regime of rivers, manifested in the warming or cooling of water with respect to conditions undisturbed by reservoir operations [15, 16]. As a result of thermal stratification, a dam reservoir can potentially cause changes in the temperature dynamics of outflow water as well as—for downstream sites—it can alter the existing synchronous air-water relationship. But to what extent? And will these potential changes be statistically significant?

This chapter aims at determining changes in the water temperature *dynamics* of a mountain river caused by reservoir operations as well as assessing the impact of reservoirs on changes in the *synchronous* air-water temperature *relationship*. Water temperature in the Carpathian river Dunajec at the Krościenko site and air temperature in the same locality, approximately 22 km downstream of the Czorsztyn-Sromowce Wyżne complex are subjected to analysis in this chapter. The study period covers two disjoint 15 year intervals before and after the construction of the studied reservoirs, which should facilitate appropriate comparisons and a consistent interpretation of research results. The changes in question will be studied with the use of recurrence plots [17] and cross-recurrence plots [18–20].

15.2 Study Area and Data Description

The Dunajec River, with its source tributary—the Czarny Dunajec—originates in the Western Tatra Mountains. It flows south west to north east across the Outer Western Carpathians in southern Poland (Fig. 15.1) and eventually drains into the Vistula River in its upper course. The Dunajec River is a permanent mountain river with a total length of 247.1 km and a catchment area of 6,804 km^2. The Krościenko site, 98 km from the springs of the Dunajec, is situated at an elevation of 414 m above sea level. The Dunajec sub-catchment at Krościenko covers an area of 1,580 km^2, and its average elevation is 836 m a.s.l. [21].

A complex of two reservoirs (Czorsztyn-Sromowce Wyżne) was built 22–24 km upstream from the Krościenko locality in the period 1994–1997. The Czorsztyn Reservoir, with a holding capacity of 231.9 million m^3 became operational in 1997. Its function is flood prevention and electric energy generation as well as increasing low flows in the Dunajec [22], from 1.5 to 9.0 m^3·s^{-1} (November–March), and to 12.0 m^3·s^{-1} (April–October). In the vicinity of its dam, the reservoir is 50 m deep, but the average depth is about 17.6 m at maximum water level. Nearly 2 km downstream from the dam, the Sromowce Wyżne Reservoir, with a holding capacity of 7.5 million m^3, was constructed in 1994. It is

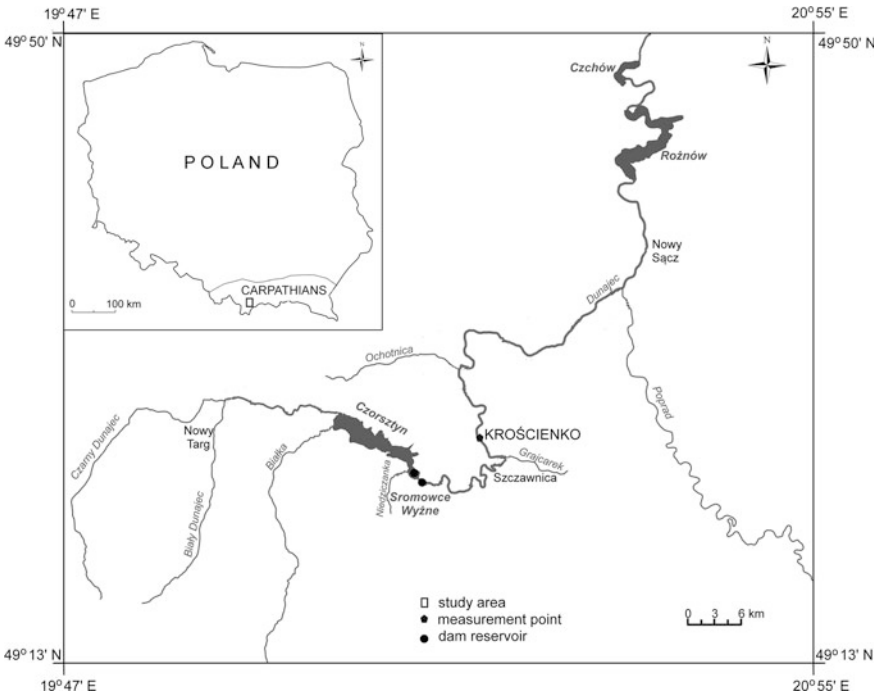

Fig. 15.1 Study area: the Krościenko site, 22 km downstream of the Czorsztyn–Sromowce Wyżne complex, built on the Dunajec River in the Polish Carpathians

Table 15.1 Statistics for the daily temperatures (°C) studied at the Krościenko site for the period 1978–2012 (35 years)

Variable	Period	Minimum	Maximum	Mean	Median	Standard deviation
T_w	1978–2012	0.0	22.1	7.4	7.2	5.8
	1978–1992	0.0	22.1	7.3	6.8	6.1
	1998–2012	0.0	18.9	7.6	7.6	5.4
T_a	1978–2012	−29.2	24.8	4.7	5.3	8.2
	1978–1992	−29.2	23.3	4.5	5.1	8.1
	1998–2012	−28.5	24.8	5.1	5.6	8.2
Annual cycle of T_w	1978–2012	0.4	15.9	7.4	6.8	5.5
	1978–1992	0.2	16.6	7.3	6.9	5.8
	1998–2012	0.6	15.4	7.6	7.7	5.3
Annual cycle of T_a	1978–2012	−6.6	15.3	4.7	4.2	6.8
	1978–1992	−9.1	14.8	4.5	4.2	6.8
	1998–2012	−7.3	16.8	5.1	4.8	6.9

a balancing reservoir for the Czorsztyn Reservoir, with a surface area of 0.9 km^2 and a maximum depth of 9.9 m [22].

The present analysis uses temperature data excerpted from a complete daily observation series collected over a period of 35 years by the Polish Institute of Meteorology and Water Management. Water (T_w) and air (T_a) temperatures (taken at 6^{00} UTC) at the Krościenko locality pertain to the hydrologic years 1978–2012 (November 1st, 1977–October 31st, 2012). The entire 35 year period includes two disjoint 15 year intervals, before and after the construction of the reservoirs, a so-called pre-dam period (1978–1992) and a post-dam period (1998–2012). Water temperature is measured approximately 0.3 m below the surface. The Dunajec River at Krościenko is 1.2 m deep, on average. In winter, when the air temperature is below freezing for a longer period of time and the water temperature drops to 0.1–0.2 °C, the river is partly frozen and frazil ice occurs as well. Descriptive statistics of the data are given in Table 15.1.

15.3 Methods

Changes in the dynamics of the temperatures in question and changes in their synchronous behavior were studied using suitable diagnostic tools: recurrence and cross-recurrence plots. Recurrence plots (RP) [17] and recurrence quantification analysis (RQA) [23–26] allow for the study and quantification of a nonlinear dynamic system on the basis of its recurrence structure. A system's dynamics can be represented by a reconstruction of its phase space trajectory from a given time series $\{x_i\}$, using time delay embedding [27, 28]. RP is a two-dimensional

representation of recurrences in a system's dynamics computed on the basis of its reconstructed trajectory, and is defined as [17]

$$R_{i,j} = H\left(\varepsilon - \|x_i - x_j\|\right),\ x_i \in R^m,\ i, j = 1 \ldots N, \tag{15.1}$$

where ε is the predefined threshold or recurrence criterion, $\|\cdot\|$ is the norm, N is the number of reconstructed points $\boldsymbol{x}_i$ of a trajectory, and $H(x)$ is the Heaviside function. In the study, the following parameters were used: (1) embedding dimension $m = 4$, (2) embedding delay $\tau_e = 91$, (3) recurrence criterion $\varepsilon = 15$ % of nearest neighbors; and without standardization of data. The embedding dimension $m = 4$ was determined based on the false nearest neighbor method [29], with a threshold $R_{tol} = 20$ and the Euclidean norm. The embedding delay was determined on the basis of a mutual information function, with the first minimum criterion [30], and this value coincided with the first zero of the autocorrelation function for each time series considered. In order to avoid serial correlation, the Theiler window [31] was set to 91 as well. The specified set of parameters, if not mentioned otherwise, was used during all subsequent calculations.

Of several RQA measures developed by [23–26], four measures of complexity were taken into account in this study. These are the following: (1) the measure *determinism* (DET), defined as the fraction of the recurrence points forming diagonal lines of at least length l_{min}; (2) the measure *laminarity* (LAM), which is the ratio of the recurrence points forming vertical lines of at least length v_{min}; (3) and the measures $\langle \mathrm{L} \rangle$ and $\langle \mathrm{V} \rangle$, which are the average lengths of diagonal and vertical lines, respectively [23–26]. As a line structure, $l_{min} = v_{min} = 7$ was considered. Moreover, these four employed measures were calculated in a moving window of size w, shifted with a step of size s over the studied time series; yielding so-called time-dependent RQA measures [32]. A moving window was applied to the entire T_w and T_a series independently, with size $w = 2100$ and step $s = 365$. The statistical significance of changes detected in a system's dynamics was calculated using a novel bootstrap-based approach providing confidence levels [32]. For that purpose, test statistics are constructed with B resamplings, and then the confidence intervals are derived on the basis of α-quantiles [32, 33]. This bootstrap approach was employed [32] using $B = 1{,}000$ resamplings, and the 99 % confidence level (taking into account both the upper and lower confidence level).

Cross-recurrence plots (CRP) allow for the study of the synchronous behavior of two time series simultaneously embedded in the same phase space [19]. The cross-recurrence plot is defined as [19]

$$CR_{i,j} = H\left(\varepsilon - \|x_i - y_j\|\right), \tag{15.2}$$

where $\boldsymbol{x}_i$ and $\boldsymbol{y}_j$ are points of the first and second trajectories, respectively. The main diagonal line, called the line of synchronization (LOS), is associated with the frequencies and phases of the systems studied; the slope of the main diagonal line represents the frequency ratio (f_{rat}), and the distance between the axes' points of

origin and the intersection of LOS with the ordinate indicates the phase difference ($\Delta\varphi$) [19] of the systems considered.

Additionally, on the basis of RP, a generalized autocorrelation function $RR^{(\varepsilon)}(\tau)$ [34] is computed for each time series in question; the cross-correlation coefficient (CPR) [34] for these functions allows for assessing the phase synchronization between the two systems. The RQA and CRQA measures, and the resulting plots were obtained with the use of the CRP Toolbox for MATLAB [35].

15.4 Results and Discussion

Figure 15.2a shows an S-shaped relationship between daily water and air temperatures at the Krościenko site for the pre-dam period. This relationship results from complex interactions between two dynamic subsystems that these temperatures are associated with, that is, the river system with its specific characteristics (Dunajec River sub-catchment) and the local weather system at Krościenko site. The shape of this curve is also determined by the fact that these temperatures relate to different media (water and air); hence, water temperature, unlike air temperature, cannot be lower than 0 °C in the winter or arbitrarily high in the summer. However, for the post-dam period (Fig. 15.2b), the slope of the curve has changed (decreased).

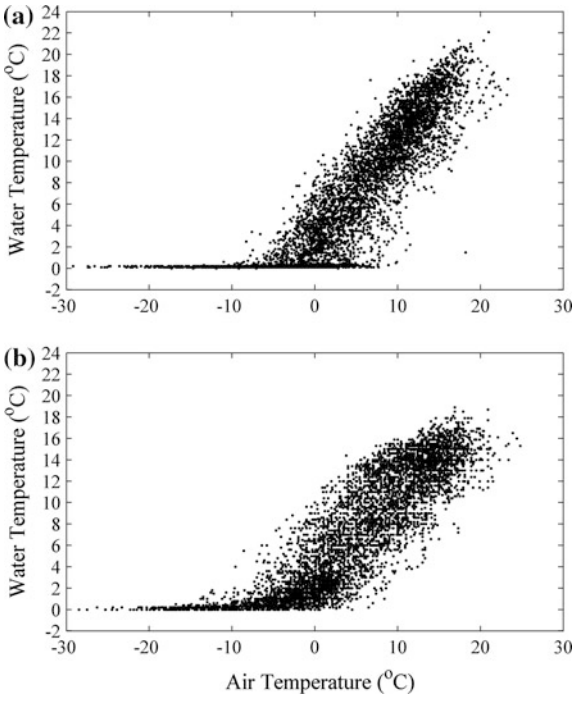

Fig. 15.2 Daily water–air temperature relationship at the Krościenko site for **a** the pre-dam period (1978–1992), and **b** the post-dam period (1998–2012)

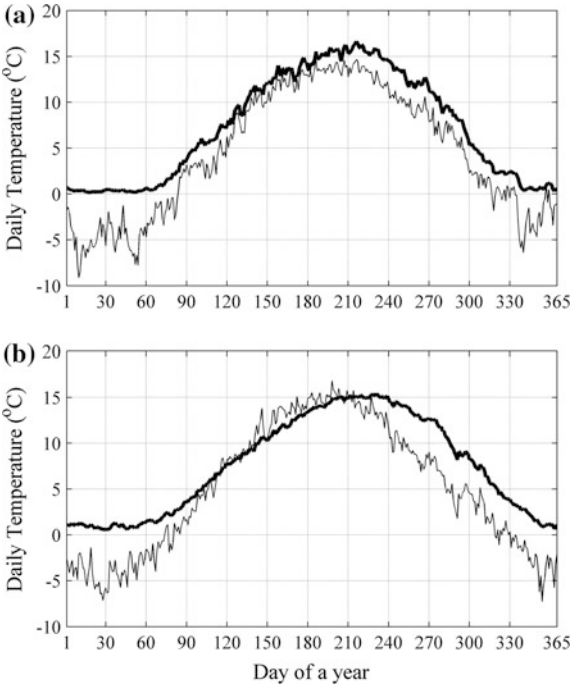

Fig. 15.3 An annual cycle of daily water (*bold line*) and air (*thin line*) temperatures at the Krościenko site for **a** the pre-dam period (1978–1992), and **b** the post-dam period (1998–2012)

The water and air temperature (averaged) pattern for a year in Krościenko for two different 15 year periods is shown in Fig. 15.3, and detailed information about the annual cycles shown is given in Table 15.1. The annual cycles depicted were calculated as mean values for each calendar day for the specified time period. In natural conditions undisturbed by anthropogenic factors, mean water temperature in Carpathian tributaries of the Vistula is higher than mean air temperature, due to groundwater supply [36]. Comparing the upper (for the years 1978–1992) and lower (for the years 1998–2012) panels of Fig. 15.3, it can be seen that, apart from a perceptible change in the range of temperature values for both T_w and T_a (see also Table 15.1), the T_w curve for the post-dam period is shifted to the right with respect to the T_a curve. As a result, the T_w curve is no longer above the T_a curve (as in the pre-dam period, Fig. 15.3a), but intersects the T_a curve (Fig. 15.3b). This most likely indicates a noticeable, underlying change in synchronous air-water temperature behavior.

The time-dependent RQA measures for the water and air temperature series are shown in Fig. 15.4 and Fig. 15.5, respectively. As can be seen in Fig. 15.4, the RQA measures for the T_w series show approximately the same change direction over the entire period considered (1978–2012). RQA measures significantly increase in value exactly in the same time window (15th point in each plot), coinciding in time with the construction of the Czorsztyn and Sromowce Wyżne

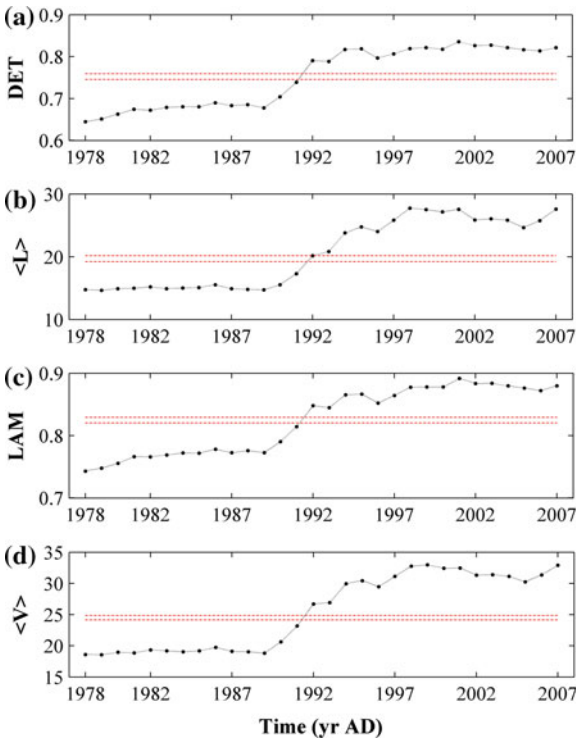

Fig. 15.4 Corresponding RQA measures for water temperature over the period 1978–2012 **a** DET, **b** ⟨L⟩, **c** LAM, and **d** ⟨V⟩. Parameters: $m = 4$, $\tau_e = 91$, $\varepsilon = 15$ % of nearest neighbors, Theiler window = 91, $l_{min} = v_{min} = 7$, moving window $w = 2100$, and $s = 365$; 99 % confidence levels are shown as *red dash-dotted lines*

reservoirs. An increase in DET for the water temperature series most likely indicates an increase in the regularity and autocorrelation [32] of Dunajec River temperature dynamics, which is also accompanied by an increase in the predictability of water temperature (the ⟨L⟩ measure) [37].

An increase in LAM and a corresponding increase in the ⟨V⟩ measure likely indicate an increase in the duration of laminar states [37] of the water temperature dynamics on the river section below the studied reservoirs.

As illustrated in Fig. 15.5, the RQA measures for air temperature in Krościenko show various significant change directions over the entire period considered. DET and LAM significantly decrease in value in the same time windows (points 8–14, corresponding with the years 1985–1995), coinciding with a decrease in ⟨L⟩ and ⟨V⟩ values; this reveals a period of more irregular and stochastic variability [32]. On the other hand, DET and LAM significantly increased in value in time windows 17–19 and 28–30, related to the years 1994–2000 and 2006–2012, respectively. Comparing these time-dependent RQA measures for water temperature (Fig. 15.4) with those for air temperature (Fig. 15.5), one can discern disparate starting points of their significant increasing tendencies. Moreover, the increase in the considered RQA measures for T_w precedes that for T_a, and corresponds with the entire period of the reservoirs' construction and operations.

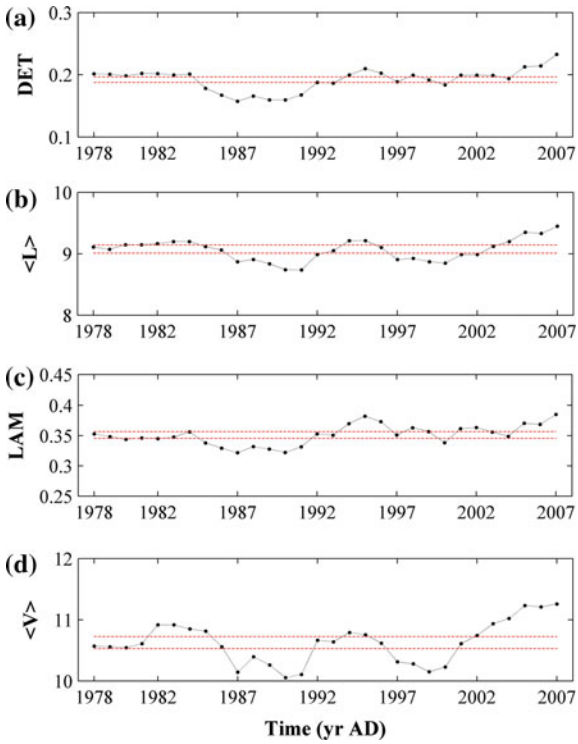

Fig. 15.5 Corresponding RQA measures for air temperature over the period 1978–2012 **a** DET, **b** ⟨L⟩, **c** LAM, and **d** ⟨V⟩. Parameters: $m = 4$, $\tau_e = 91$, $\varepsilon = 15$ % of nearest neighbors, Theiler window = 91, $l_{min} = v_{min} = 7$, moving window $w = 2100$, and $s = 365$; 99 % confidence levels are shown as *red dash-dotted lines*

Cross-recurrence plots for the water and air temperature series pertaining to two different 15 year periods (1978–1992 and 1998–2012) are shown in Fig. 15.6. The reconstructed phase space for each CRP is 4-dimensional, owing to the 2 variables (T_w, T_a) and the embedding dimension $m = 2$ (while $\tau_e = 91$). Each CRP contains the LOS and long diagonal lines; these indicate a similar trajectory evolution [20] of the two temperature series in question. The roughly uniform distance between successive diagonal lines corresponds with an annual cycle.

A linear fit to the LOS (estimated non-parametrically via algorithm [19]) gives the same frequency ratio ($f_{rat} \approx 1.0$) for T_w and T_a both in the pre-dam period (Fig. 15.6a) and in the post-dam period (Fig. 15.6b); however, the phase difference in the latter period ($\Delta\varphi \approx 17.0$) is approximately 6.3 times greater than that in the pre-dam period ($\Delta\varphi \approx 2.7$). The increase in phase difference between T_w and T_a signifies a deterioration of their synchronous behavior.

The obtained results indicate that in the pre-dam period the time lag between the T_w and T_a studied was less than 3 days, and a six-fold increase results in over 17 days of lag time for water temperatures in the Dunajec River relative to air temperatures at the Krościenko site.

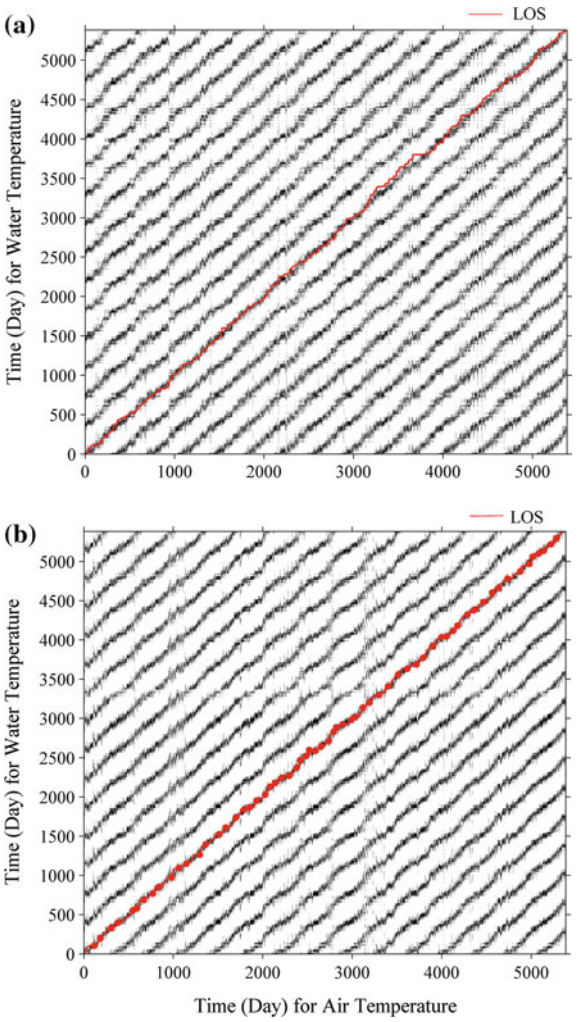

Fig. 15.6 Cross-recurrence plot for daily water and air temperatures at the Krościenko site **a** in the pre-dam period (1978–1992); and **b** in the post-dam period (1998–2012). Parameters: $m = 2$, $\tau_e = 91$, and $\varepsilon = 15$ % of nearest neighbors. A linear fit to LOS: $y = 1.005 \cdot x + 2.715$ and $y = 1.001 \cdot x + 17.002$, for pre-dam and post-dam periods, respectively

Figure 15.7 shows the generalized autocorrelation functions ($RR^{(\varepsilon)}(\tau)$) computed for each time series in question, with the embedding dimension $m = 2$ (while $\tau_e = 91$). The local maxima of these functions occur at $\tau = n \cdot 365$, where n is an integer, and their heights are different for both temperature series, corroborating phase synchronization [37]. The cross-correlation coefficient for these functions is CPR = 0.87 for the pre-dam period (Fig. 15.7a), and CPR = 0.82 for the post-dam period (Fig. 15.7b), which clearly indicates phase synchronization between the two temperatures studied. However, in the latter period, the value of CPR is slightly lower, which points to slightly weaker synchronization.

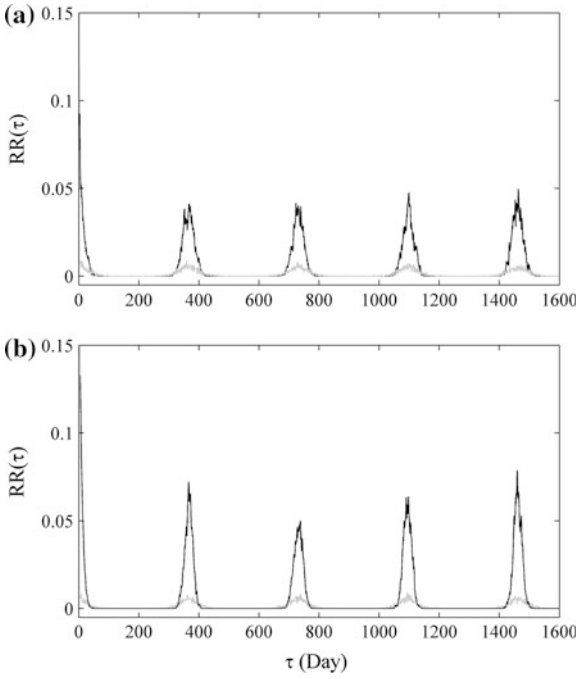

Fig. 15.7 The generalized autocorrelation functions for water (in *black*) and air (in *gray*) temperatures at the Krościenko site, and cross-correlation coefficient CPR for these functions **a** CPR = 0.87 in the pre-dam period (1978–1992); and **b** CPR = 0.82 in the post-dam period (1998–2012). Parameters: $m = 2$, $\tau_e = 91$, and $\varepsilon = 15$ % of nearest neighbors

15.5 Conclusions

On the basis of RP and CRP, the conducted analyses have shown that the water temperature dynamics in the Dunajec River 22 km downstream from the Czorsztyn-Sromowce Wyżne reservoir complex has significantly changed after the construction of the two reservoirs. An increase in the four considered complexity measures was observed for water temperature dynamics, which was not accompanied by similar changes in air temperature dynamics in Krościenko. The increase in the RQA measures for water temperature points to an increase in regularity, autocorrelation, predictability, and laminarity of water temperature dynamics for the river section below the two studied reservoirs.

Phase synchronization (with the same frequency ratio) between water and air temperature was corroborated by LOS appearance in CRP; however, after the construction of the impoundments, the phase difference increased six-fold, on average, in comparison with the pre-dam period, which indicates a weakening of their synchronization. The cross-correlation coefficient between the generalized autocorrelation functions for the studied water and air temperature values was close to 1, which again confirms the phase synchronization of the temperatures considered; though, in the latter period, the value of CPR is slightly lower, which signifies a deterioration of their synchronous behavior.

Acknowledgements The author would like to thank Łukasz Wiejaczka for fruitful discussions and his help in preparing one figure.

References

1. B.A. Sinokrot, H.G. Stefan, Water Resour. Res. **29**(7), 2299–2312 (1993)
2. D. Caissie, M.G. Satish, N. El-Jabi, J. Hydrol. **336**(3–4), 303–315 (2007)
3. G.W. Brown, Water Resour. Res. **5**(1), 68–75 (1969)
4. D. Caissie, N. El-Jabi, M.G. Satish, J. Hydrol. **251**(1), 14–28 (2001)
5. O. Mohseni, H.G. Stefan, J. Hydrol. **218**(3), 128–141 (1999)
6. H.G. Stefan, E.B. Preud'homme, J. Am. Water Resour. **29**, 27–45 (1993)
7. A.P. Mackey, A.D. Berrie, Hydrobiologia **210**, 183–189 (1991)
8. B.W. Webb, D.E. Walling, Freshw. Biol. **29**, 167–182 (1993)
9. B.W. Webb, D.E. Walling, Regulated Rivers Res. Manage. **13**, 463–477 (1997)
10. S.L. Johnson, Can. J. Fish. Aquat. Sci. **61**, 913–923 (2004)
11. F.J. Magilligan, K.H. Nislow, Geomorphology **71**(1), 61–78 (2005)
12. N.L. Poff, J.D. Allan, M.B. Bain, J.R. Karr, K.L. Prestegaard, B.D. Richter, R.E. Sparks, J.C. Stromberg, Bioscience **47**(11), 769–784 (1997)
13. J. Thornton, A. Steel, W. Rast, Reservoirs, in *Water Quality Assessment—A Guide to Use of Biota, Sediments and Water in Environmental Monitoring*, ed. by D. Chapman (London, 1996), p. 651. www.who.int/water_sanitation_health/resourcesquality/wqachapter8.pdf
14. B.A. Sinokrot, H.G. Stefan, J.H. McCormick, J.G. Eaton, Clim. Change **30**(2), 181–200 (1995)
15. B. Liu, D. Yang, B. Ye, S. Berezovskaya, Glob. Planet. Change **48**, 96–111 (2005)
16. R. Soja, Ł. Wiejaczka, Water Environ. J. **28**, 473–482 (2014)
17. J.P. Eckmann, S.O. Kamphorst, D. Ruelle, Europhys. Lett. **4**(9), 973–977 (1987)
18. J.P. Zbilut, A. Giuliani, C.L. Webber Jr, Phys. Lett. A **246**(1–2), 122–128 (1998)
19. N. Marwan, M. Thiel, N.R. Nowaczyk, Nonlinear Proc. Geophys. **9**, 325–331 (2002)
20. N. Marwan, J. Kurths, Phys. Lett. A **302**, 299–307 (2002)
21. W. Chełmicki, in *The Division of the Upper Vistula and its Basin,* ed. by I. Dynowska, M. Maciejewski. The Upper Vistula Basin, part I. (In Polish) (PWN, Warszawa-Kraków, Poland, 1991), pp. 18–29
22. J. Hennig, I. Hennig, A. Roszkowski, in *Retention Reservoirs*, ed. by I. Dynowska, M. Maciejewski. The Upper Vistula Basin, part II. (In Polish) (PWN, Warszawa-Kraków, Poland, 1991), pp. 121–143
23. J.P. Zbilut, C.L. Webber Jr, Phys. Lett. A **171**(3–4), 199–203 (1992)
24. C.L. Webber Jr, J.P. Zbilut, J. Appl. Physiol. **76**, 965–973 (1994)
25. L.L. Trulla, A. Giuliani, J.P. Zbilut, C.L. Webber Jr, Phys. Lett. A **223**(4), 255–260 (1996)
26. N. Marwan, N. Wessel, J. Kurths, Phys. Rev. E **66**(2), 026702 (2002)
27. N.H. Packard, J.P. Crutchfield, J.D. Farmer, R.S. Shaw, Phys. Rev. Lett. **45**, 712–716 (1980)
28. F. Takens, in *Detecting Strange Attractors in Turbulence,* ed. by D.A. Rand, L.S. Young. Dynamical Systems and Turbulence, Lecture Notes in Mathematics, vol. 898 (Springer, Berlin, 1981), pp. 366–381
29. M.B. Kennel, R. Brown, H.D.I. Abarbanel, Phys. Rev. A **45**, 3403–3411 (1992)
30. A.M. Fraser, H.L. Swinney, Phys. Rev. A **33**(2), 1134–1140 (1986)
31. J. Theiler, Phys. Rev. A **34**, 2427–2432 (1986)
32. N. Marwan, S. Schinkel, J. Kurths, Europhys. Lett. **101**, 20007 (2013)
33. S. Schinkel, N. Marwan, O. Dimigen, J. Kurths, Phys. Lett. A **373**, 2245–2250 (2009)
34. M.C. Romano, M. Thiel, J. Kurths, I.Z. Kiss, J. Hudson, Europhys. Lett. **71**(3), 466–472 (2005)
35. N. Marwan, CRP Toolbox 5.17. Platform independent (for Matlab) (2013). http://tocsy.pik-potsdam.de/CRPtoolbox
36. A. Łajczak, in *River Thermal Regime and Ice Cover* ed. by I. Dynowska, M. Maciejewski. The Upper Vistula Basin, part I. (In Polish) (PWN, Warszawa-Kraków, Poland, 1991), pp. 243–249
37. N. Marwan, M.C. Romano, M. Thiel, J. Kurths, Phys. Rep. **438**(5–6), 237–329 (2007)

Chapter 16
Recurrence Analysis of Eddy Covariance Fluxes

Milan Flach, Holger Lange, Thomas Foken and Michael Hauhs

Abstract Measuring energy and matter fluxes between the atmosphere and vegetation using the Eddy Covariance (EC) technique is the state-of-the-art method to quantify carbon exchange between terrestrial ecosystems and their surrounding. The EC equipment is usually mounted onto a *flux tower* reaching higher than the local canopy. Today, more than 600 flux towers are in operation worldwide. The methodological requirements lead to high sampling frequency (20 Hz) and thus to the production of very long time series. These are related to temperature, wind components, water vapour, heat and gas exchange, and others. In this chapter, the potential of Recurrence Analysis (RA) to investigate the dynamics of this atmosphere-vegetation boundary system is elucidated. In particular, the effect of temporal resolution, the identification of periods particular suitable for reliable EC flux calculations, and the detection of transitions between dynamical regimes will be highlighted.

16.1 Introduction

Uptake and release of carbon from vegetation represent two large fluxes of opposite direction. The sign of their difference determines whether an ecosystem, or the biosphere as a whole, acts as carbon sink or carbon source to the atmosphere, and thus the type of feedback to Earth's climate. Every effort to reduce the imprecision of incoming and outgoing carbon fluxes is more than welcome. The problem

M. Flach
Max-Planck-Institute for Biogeochemistry,
P. O. Box 10 01 64, 07701 Jena, Germany
e-mail: milan.flach@bgc-jena.mpg.de

H. Lange (✉)
Norwegian Institute of Bioeconomy Research,
P. O. Box 115, 1431 Ås, Norway
e-mail: holger.lange@nibio.no

T. Foken · M. Hauhs
University of Bayreuth, Bayreuth Center of Ecology
and Environmental Research (BayCEER), 95440 Bayreuth, Germany

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8_16

is intriguing since processes are involved which act on rather disparate time scales, from turbulent exchange of CO_2 molecules at vegetation surfaces, quantified through Eddy Covariance (EC) measurements at a rate of 20 measurements per second, to changes in carbon stocks in forest soils which are significant at decadal scales only.

High frequency measurements from flux towers come with special challenges rendering both reliable calculations of carbon fluxes as well as statistical (time-series) analysis difficult. Specifically, we seek answers to questions relevant to each detailed investigation of EC-related fluxes. Is it possible to:

1. discern periods where the (atmospheric) conditions are particularly suitable to obtain reliable EC fluxes?
2. detect dynamical transitions (different behaviour) beyond those obvious from visual inspection?
3. recommend an optimal time resolution for measuring EC data and analyzing EC fluxes?

Besides seeking answers to these specific questions, Recurrence Analysis (RA henceforth) has the potential to provide insights into the dynamical system at the atmosphere-vegetation boundary not easily obtained otherwise, such as the effective dimension of the (attractor of the) system, delays, memory effects, sensitivity to initial conditions, and more.

16.2 Data and Preprocessing

This section describes the data sets obtained from flux towers. There are more than 600 of these towers in a worldwide net, organized in a network of networks called FLUXNET (Fig. 16.1).

In the near future, the Integrated Carbon Observation System (ICOS) will be operational. ICOS is a mulitnational research infrastructure founded to standardize and harmonize flux tower operation and data gathering with an open access policy (http://www.icos-ri.eu). Under the umbrella of ICOS, a substantial number of new flux tower sites will be established.

To illustrate the potential of RA for this type of observations, we focus on two towers, one with raw data on wind speed and gas concentrations at 20 Hz resolution, serving as basis to compute EC fluxes, and one with multiyear time series of EC fluxes at a temporal resolution of 30 min. An overview of the two sites, their coordinates and climate is given in Table 16.1. Both sites are classified as evergreen needle forests.

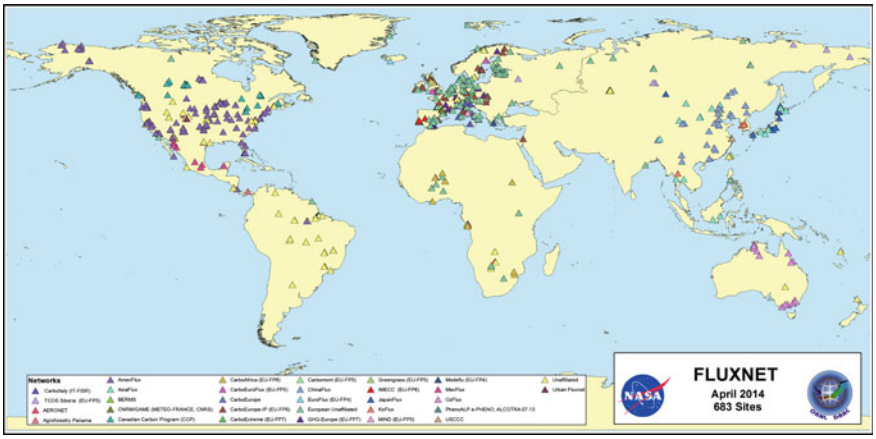

Fig. 16.1 Worldwide net of flux towers, http://fluxnet.ornl.gov/maps-graphics

Table 16.1 Overview of the used sites within Fluxnet

Site-ID	Name	Latitude	Longitude	Climate	Data
DE-Bay	Waldstein–Weidenbrunnen, Germany	50.1419 °N	11.8669 °E	Temperate	20 Hz high resolution
NL-Loo	Loobos, Netherlands	52.1666 °N	5.7436 °E	Temperate	Half hourly fluxes

16.2.1 High Resolution Data from DE-Bay

For this analysis, EC high resolution data obtained in the context of the EGER project [1] at the Waldstein-Weidenbrunnen site (Bavaria, Germany) [2] between June, 24th and July, 1st 2011 are used. Carbon dioxide (CO_2) and water vapour (H_2O) concentrations and vertical wind speed are sampled with 20 Hz at a height of 36 m above a 27 m high spruce forest floor at the slim tower. The period is considered as one with a fairly good data quality (coined "the golden days") according to several quality flags and expert judgements. Unfortunately, we thus analyze only one week of high resolution data, bringing seasonality and long term changes out of scope. However, long term high quality time series at 20 Hz high resolution are almost non-existent, and weather conditions do change of course also within this limited period.

16.2.2 Fluxnet Data from NL-Loo

In Fluxnet, the measurements at the individual towers seek to monitor fluxes across the boundary of the ecosystem. The 20 Hz high resolution data provide the basis to compute fluxes of the quantities of interest. A flux is retrieved by multiplying devia-

tions of vertical wind speed from its mean with the deviations of the respective quantity (e.g. CO_2, H_2O) from its respective mean, then averaging over a time interval of usually 30 min, which is known as covariance (thus the name of the method). Multiplied by the density of dry air, EC fluxes are obtained [3]. Three EC fluxes are used from Fluxnet: Net ecosystem exchange (NEE_f), latent heat flux (LE_f) and sensible heat flux (H_f). NEE_f is the carbon flux, derived from deviations in CO_2 concentrations. This flux quantifies the effects of photosynthesis and respiration of the ecosystem combined. LE_f is derived from H_2O concentrations in the atmosphere, quantifying the process of condensation and evapotranspiration. H_f is derived from the deviation of temperature in the atmosphere, quantifying the transport of energy. In addition, also meteorological variables such as air temperature, precipitation, global radiation and vapour pressure deficit are measured at the towers (typically at much lower temporal resolution than wind speed and concentrations, such as one measurement per minute) and made available via Fluxnet.

The EC method assumes the different measurements to be perfectly synchronized in time (i.e. the same time stamp refers to the very same moment in time), which is difficult to obtain due to limitations of the measurement devices. Thus, small deviations between vertical wind and the quantity of interest are corrected by shifting one of the series until maximal cross correlation at lag zero is obtained.

Another obstacle is purely geometrical: the EC method uses the mass balance equation as physical basis, which assumes the coordinate system to be parallel to the mean horizontal wind and homogeneous topography at its boundaries. Furthermore, small misalignments of the sonic anemometer (measuring wind speed) during assembly are unavoidable. These biases are corrected for with a rotation of the coordinate system (with methods called *planar fit* or *double rotation*, e.g. [3]).

A well-mixed turbulent atmosphere is an additional requirement for obtaining reliable EC fluxes. For instance, low turbulence during nighttime leads to underestimation of the ecosystem sequestration, as storage and advection processes might be of importance as well. Periods of low turbulence are detected by using an heuristically estimated threshold of the friction velocity (u_*), and are removed from the time series [4, 5].

The Fluxnet data set is run through a preprocessing chain, including quality control and assessment. Quality is enhanced by correcting the data for systematic biases, removing spikes and other low quality data and finally gap-filling the time series. This is mostly done by *marginal distribution sampling*. Marginal distribution sampling works like a moving look up table, searching for similar conditions of correlated meteorological variables like global radiation, vapour pressure deficit and temperature in the vicinity of the gap [5].

The interested reader is referred to an overview on the EC method in Aubinet et al. [3] and further details on the Fluxnet standard data processing especially in Papale et al. [4], Reichstein et al. [5], Moffat et al. [6].

16.3 Recurrence Analysis of Eddy Covariance Fluxes

In the following we consider Eddy Covariance fluxes to be part of a dynamical system at the atmosphere-biosphere boundary. Dynamical systems are usually described by one or more differential equations, giving the behaviour of the system for a certain time in the form of state transitions. All possible states of the system are represented in its phase space. Analyzing system behaviour in phase space reveals the crucial characteristics of the system. However, usually only little is known about the other (unmeasured) variables which complete the characterization of the dynamical system and thus its phase space. One way to characterize properties of the phase space of unknown variables, given just one time series of a single variable, is *embedding*.

16.3.1 Embedding

The reconstruction of the m-dimensional phase space is motivated by the embedding theorem [7] and done by considering an observation at time t, evolving from $m-1$ observations, each one step ΔT in the past.

$$\vec{X} = (x_t, x_{t-\Delta T}, x_{t-2\cdot\Delta T}, ..., x_{t-(m-1)\cdot\Delta T}) \tag{16.1}$$

For reconstructing the phase space, it is crucial to find appropriate embedding parameters m and ΔT. m is estimated by searching for a minimal number of *false nearest neighbors* (FNN) [8].

A suitable time delay ΔT can be either selected by accounting for nonlinear correlations with the mutual information or with the autocorrelation of a time series.

The autocorrelation function (ACF) is one simple conventional measure of linear dependence in the dynamics of a time series. In order to obtain independent subsequent $\mathbf{X}$, one has to determine a suitable time delay ΔT. One way would be to choose the first zero-crossing of the ACF [9]. Another possibility is the *length of the autocorrelation* (LAC). LAC is the time a signal needs until the absolute value of its sample autocorrelation function falls below a significance threshold determined by the length of the time series [10]. However, this has the caveat that strongly autocorrelated data, such as many environmental time series, might not exhibit a finite LAC. The same holds even stronger when using the autocorrelation time [11] instead. Here we use the assumption of Gaussian noise and a significance level α with the convention $\alpha = 5\,\%$.

LAC takes only linear correlations into account. Therefore, the second possibility to find ΔT is looking for nonlinear dependencies. The mutual information (*MI*) is a method introduced by Fraser and Swinney [12] which measures the nonlinear dependence of two variables. The first local minimum of *MI* seems to be a good choice for ΔT in time delay embedding [12].

16.3.2 Recurrence Analysis

Ecosystems are considered to show a rather complex behaviour at their boundaries [13]. As EC fluxes quantify the processes between ecosystems and the atmosphere, it is expected that the data sets might display nonlinear behaviour. This requires methods which can deal with chaotic data, like recurrence plots. The current analysis was performed with the *python* package *pyunicorn*.[1]

To quantify the visual structures of the recurrence plot, several measures have been proposed [14–16], known as recurrence quantification analysis (*RQA*). A lot of different *RQA* measures exist. If not referenced otherwise, all definitions of RQA measures are according to Marwan et al. [17] in the following. Most of the measures take the structures of diagonal and vertical lines into account.

For discrete time series, the recurrences in phase space form a recurrence matrix, visualized by the recurrence plot. This matrix is also closely related to the adjacency matrix of a corresponding complex network, by simply excluding the line of identity to avoid self loops [18, 19]. Recurrence networks (*RN*) can be quantitatively characterized taking their vertices, edges and the path between two vertices into account. The definitions of the relevant network measures can be found in [19]. The quantification of recurrence plots (RQA) and recurrence networks (RNA) will be jointly referred to as recurrence analysis (RA) in the following. A comprehensive summary of the RA measures used here is given in Table 16.2.

16.4 High Resolution Data at DE-Bay

As discussed above, high resolution data provide the basis for the computation of EC fluxes. In this section we address the high resolution data of the DE-Bay site.

16.4.1 Fixing Embedding and RA Parameters

To perform RA on the high resolution data, the RA parameters have to be fixed first. For reconstructing the dynamical system's phase space, we choose an appropriate time delay and embedding dimension and use a fixed recurrence rate (RR).

To find an appropriate time delay with the auto correlation function or mutual information, which is not restricted to the normal diurnal cycle of the data, trend and daily periodicity are provisionally removed. We remove the diurnal cycle in EC high resolution data with singular spectrum analysis (a detailed description of the method can be found in [20]). Then, the autocorrelation function and mutual information (MI) is calculated as outlined above for different temporal resolutions of the EC high resolution data.

[1] Available at http://tocsy.pik-potsdam.de/pyunicorn.php.

Table 16.2 Recurrence analysis measures and their interpretation

Measure	Abbreviation	Definition	Interpretation
Laminarity	*LAM*	Proportion of vertical lines	Laminarity of the system
Determinism	*DET*	Proportion of diagonal lines	Determinism of the system
Trapping time	*TT*	Average length of vertical lines	Time the system does not change
Predictability	*PRE*	Average length of diagonal lines	Time the system is predictable, periodic
Entropy	*ENT*	Shannon entropy of the probability to find a line of a specific length	Complexity of recurrent trajectories
Trend	*TRE*	Change in density of recurrent points along diagonals, from the line of identity towards the upper left and lower right corners	Non-stationarity of the system
Assortativity	$\mathcal{A}$	Preferable connection of nodes	Continuity of vectors in phase space
Average path length	$\mathcal{L}$	Average length of all shortest paths connections	Indicator for dynamical transitions
Transitivity	$\mathcal{T}$	Probability that directly connected nodes are also connected via one neighbor	Chaotic versus. regular behaviour (chaotic < 0.75, periodic $= 0.75$, etc.)

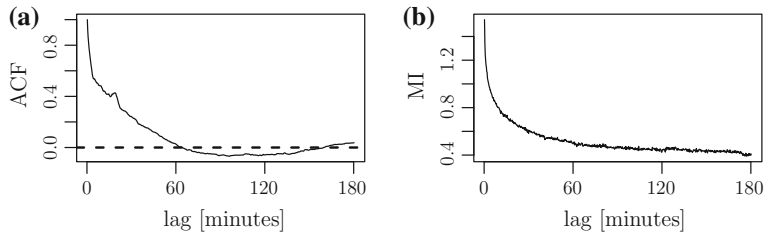

Fig. 16.2 **a** *ACF*, **b** *MI* of vertical wind speed for EC high resolution data

The first zero crossing of the *ACF* varies slightly around 1 h for vertical wind speed (Fig. 16.2a), CO_2 concentration and H_2O concentration. Also *MI* flattens approximately at this time, but does not exhibit an unambiguous local minimum (Fig. 16.2b). Thus, 60 min, the first zero crossing of *ACF*, is used as ΔT for the time delay embedding.

After fixing the time delay parameter, the analysis continues with the original data, including trends and diurnal cycles, with the idea to determine an appropriate embedding dimension and a suitable recurrence rate as well as obtaining recurrence measures.

False Nearest Neighbors calculations did not lead to conclusive results for the choice of the embedding dimension that would fit all time series. This is common for time series containing noise [21]. The long time series make higher dimensions feasible, in contrast to many other applications with shorter data sets. Thus, we choose a general embedding dimension of 5. The whole times series is long enough (8 days) for $m = 5$ and $\Delta T = 1$ h.

A fixed threshold (FTHR) ϵ is known to lead to a substantial variance of the RR [22]. Thus, we choose a fixed recurrence rate (FRR) of 10 %, which is a typical choice for RA in environmental applications [22]. Distances in phase space are computed with the Euclidean norm. RA measures are calculated in moving windows of 6 h length to take diurnal fluctuations of the data into account. Windows are overlapping with a window distance of 4 h to avoid high computation times.

16.4.2 Changing the Temporal Resolution

Changes in temporal resolution have not been analyzed yet in a systematic way with RA, although this happens often when processing high-resolution data. This is relevant for EC data, as in typical EC data processing software, a cutoff frequency of 2 Hz is used [23], i.e. the time series is decimated 10-fold. [24] even stated that sampling with intervals of several seconds suffices for calculating EC fluxes.

One crucial aspect when changing the temporal resolution is the handling of the embedding parameters of the time series. Both embedding parameters are fixed to a certain dimension (m) and a certain time delay (ΔT) in units of time. The corresponding lag of the time delay embedding is changed respectively. Thus, analysis of the temporal resolution is restricted to the temporal resolutions of the (integer) set of divisors of ΔT, which is 1 h in this case.[2] This ensures that changes of RA measures with changing temporal resolution are not an artifact of time delay embedding parameters. Temporal resolution can be changed by applying two basic methods, aggregation and decimation. Aggregation of data is done by forming the average or the sum, depending on the type of variable. Then aggregated values are assigned to the point in time referring to the end of the aggregated period [1]. Decimation is a technique of downsampling. Decimation by the factor n is done by keeping every nth sample of the time series, yielding a new time series of a lower temporal resolution. Decimation in high resolution data is of special interest for the EC method, as some gas concentrations cannot be measured at the high temporal resolution of the sonic anemometer (10–20 Hz). A technique known as disjunct eddy accumulation method

[2]We used temporal resolutions of 1, 2, 4, 6, 8, 10, 12, 15, 18, 20, 25, 30, 36, 40, 45, 50, 60, 72, 80, 90, 100, 120, 144, 150, 180, 200, 225, 240, 300, 360, 400, 450, 600, 720, and 900 s.

makes use of concentrations from lower temporal resolutions to compute fluxes of it, and is intended to minimize the error induced by the lower temporal resolution [25, 26].

As aggregation and decimation are two completely different techniques, one would expect large differences between RA measures computed after decimating or aggregating the data. To quantify the differences between aggregation and decimation, the data are binned in classes of temporal resolution, correlated and the R^2 values between them are compared.

When aggregation or decimating the data up to 900 s, we can only observe as few as 24 observations in one 6 h window. To judge whether results with such short time series are stable and different from noise, a small simulation is done. According to Kaimal et al. [27], the high frequency vertical wind time series might be resembling pink noise processes with a spectral exponent of $\beta = -\frac{5}{3}$. Thus, pink noise of different time series length is produced with the R-package of Ligges et al. [28], embedded and *TRE* calculated. 1000 realizations of a pink noise process are used to achieve confidence intervals (*CI*) of 5 %, 95 % and the median.

16.4.3 Results and Discussion

An important auxiliary information to understand the dynamics of EC high resolution data is an overview of meteorological conditions during the period of interest. At the DE-Bay site during the selected period in 2011, the conditions are changing between Day of the Year (*DoY*) 177 (June, 24th, 2011) and 179: Wind direction (*dir*) from north-east winds to more changing winds coming from west to south between *DoY* 177–179 noon (Fig. 16.3a). Air temperatures increase as well as the amplitude of the diurnal cycle (Fig. 16.3b). An indicator of atmospheric stability shows negative values (unstable stratification) from early morning to late afternoon for these nice days (*DoY* 177–179 noon), nighttime values (grey shaded) indicate neutral to slightly stable stratification (Fig. 16.3b). In the following, we show that these changes in meteorological elements (weather) can also be obtained with recurrence analysis.

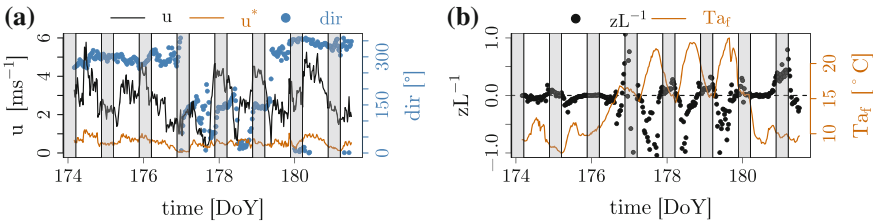

Fig. 16.3 Meteorological overview (June, 24th to July, 1st, 2011) of **a** wind velocity (u) and friction velocity (u_*, *left axis*), wind direction (dir), **b** stability parameter (zL^{-1}) and air temperature (Ta_f). Night time (between sunset and sunrise) is *grey-shaded*

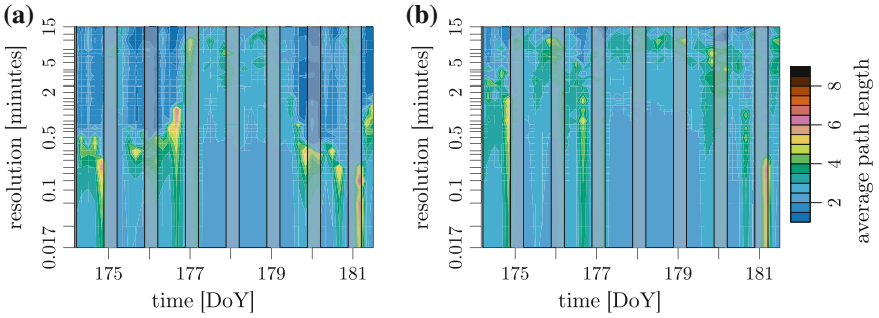

Fig. 16.4 Average path length $\mathcal{L}$ of vertical wind velocity at the Waldstein site in 2 h overlapping moving windows of 6 h length, changing temporal resolution with **a** aggregation and **b** decimation

Table 16.3 Quantiles of average path length with simulated pink noise data on different length of the time series (N) corresponding to 15, 10, 5, 2, 1 and 0.5 min temporal resolution of a 6 h window

	N = 24	N = 36	N = 72	N = 180	N = 360	N = 720
5 %	1.36	2.24	2.54	2.56	2.51	2.46
50 %	2.09	3.19	3.03	2.76	2.64	2.60
95 %	3.48	4.70	3.86	3.32	2.97	2.89

RA measures are presented as contour of time in overlapping moving windows (x-axis) by changing the temporal resolution (y-axis), e.g. Fig. 16.5.

Detecting Dynamical Transitions

Average path length ($\mathcal{L}$) of vertical wind velocity varies between one (very simple) and five nodes (very complex networks), needed on average to connect all pairs of nodes in the network (Fig. 16.4). Higher $\mathcal{L}$ values occur in the evening of *DoY* 174, 176, 179 and in the morning of *DoY* 181. Especially with respect to the nice weather period, it is interesting to note that for the aggregation method (Fig. 16.4a), temporal resolutions shorter than one minute change from high to low $\mathcal{L}$, whereas temporal resolutions aggregated longer than one minute tend to do the opposite. With respect to the pink noise simulation, $\mathcal{L}$ is a relatively stable measure even for small recurrence networks shorter than 100 observations (Table 16.3). Most observed changes of $\mathcal{L}$ in vertical wind speed exceed the confidence intervals of the pink noise simulation and are indicating a following change in the dynamics of the time series. As $\mathcal{L}$ is known to be a very sensitive indicator for dynamical transitions [18, 29] and with the changes of the meteorological conditions in mind (Fig. 16.3), it seems that increased values of $\mathcal{L}$ at *DoY* 176 indicate already in the evening the change into the nice weather period, coming ahead with changed wind directions of the three following days. Interestingly, these changes can be derived only from the analysis of vertical wind speeds' phase space. The visual interpretation of a set of standard meteorological variables supports this indication of change in the system's behaviour.

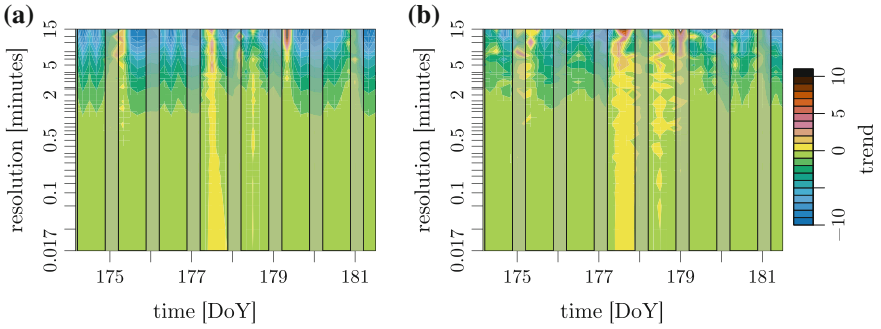

Fig. 16.5 *TRE* of vertical wind velocity at the Waldstein site in 2 h overlapping moving windows of 6 h length, changing temporal resolution with **a** aggregation and **b** decimation

Discerning Stationary Periods

The *trend* (*TRE*) is an interesting RA measure for EC high resolution data as it is interpreted as instationarity of the time series. Stationarity is one of the prerequisites of the EC method, determining the commonly used fluxes from the high resolution data of vertical wind speed and other variables [3].

TRE shows vertical structures with increased *TRE* for daytime values of *DoY* 177–179. These are days with very low wind velocity and thus unstable stratification. This is more pronounced for aggregated and decimated temporal resolutions longer than 2 min (upper part of Fig. 16.5).

Temporal resolutions shorter than 2 min have only slightly negative trend values, thus can be considered to be quasi stationary (Fig. 16.5); thus, this prerequisite of flux derivation with the EC method is met for small temporal scales. Longer temporal resolutions result in *TRE* values which are more negative for most days or larger than zero for the nice weather period. Typically negative values of *TRE* indicate a drift in the time series and positive values are very unlikely to occur, unless sampled in discordance with periodicity.

TRE is already known to react sensitive on changes of the time series length, as it is done when changing temporal resolution [22]. This can be confirmed by the pink noise simulation (Table 16.4). *TRE* tend to be more negative for short time series. Values of *TRE* for vertical wind speed (Fig. 16.5) are in the range of the pink noise

Table 16.4 Quantiles of TRE with simulated pink noise data on different length of the time series (N) corresponding to 15, 10, 5, 2, 1 and 0.5 min temporal resolution of a 6 h window

	N = 24	N = 36	N = 72	N = 180	N = 360	N = 720
5 %	−8.76	−7.30	−3.63	−1.43	−0.63	−0.29
50 %	−5.51	−4.08	−1.59	−0.59	−0.22	−0.08
95 %	3.07	2.16	1.08	0.40	0.26	0.09

simulation (Table 16.4) up to a temporal resolution of approximately 5 min. Then *TRE* gets mostly more negative, which is in concordance to the simulations, with some exceptions for aggregation (noon of *DoY* 175, 181). Some positive values are out of the upper bounds of *TRE*, especially for days with unstable stratification, indicating that there might be longer periodicities of vertical wind speed, which cannot be quantified with a time window of 6 h. To verify this, the size of the time window has to be varied with a fixed temporal resolution, revealing the time of stationary behaviour of vertical wind speed.

Summarizing the results for stationarity, it can be confirmed that *TRE* can detect stationary parts of the EC high resolution data. In the case at issue, the time series can be considered as stationary when aggregating or decimating up to 5 min. But EC fluxes are typically calculated in 30 min intervals [3]. This interval is outside of the methods limits, due to only 12 values in one 6 h window. Nevertheless, going to the upper limit of the analysis, nonstationary parts are detected for single days and especially for the nice weather days. Vertical wind speed in the 6 h window gets nonstationary on certain days on temporal scales longer than 5 min, whereas it is stationary in the same 6 h window on shorter time scales. This fact is contrary to the expectations for the aggregation method, as vertical wind speed is expected to be stationary and zero, when averaging it over longer periods. Non-zero average vertical wind speed has been reported before especially over tall vegetation, as is the case here [3]. This is accounted for by rotating the coordinate system before calculating the fluxes, e.g. with the *planar fit method* (see [3]). *TRE* might be of interest as it has the potential to complement the quality checks used so far for stationarity of EC high resolution data in the EC community [30].

Recommending Temporal Resolution

According to the results above, RA can be used to recommend optimal temporal resolutions. In the sense of RA, a temporal resolution is sufficient if the used RA measures do not change (much) when aggregating or decimating the time series to achieve another temporal resolution (stability on the y axis of Fig. 16.5). Furthermore, aggregation (Fig. 16.5a) and decimation (Fig. 16.5b) should exhibit different structures. Obviously redundant information is available, if they do not. To quantify this, we correlate all RA measures used (Table 16.2) at different temporal resolutions, achieved by either decimating or aggregating.

A very high correlation close to $R^2 = 1.0$ can be found between aggregation and decimation of vertical wind velocity for temporal resolutions of 6 s or shorter (RA measures minimal $R^2 = 0.98$, Fig. 16.6b). CO_2 has high correlations for temporal resolutions of 20 s or shorter (Fig. 16.6a). Average path length of CO_2 exhibits minimal $R^2 = 0.93$. Excluding average path length, the correlation is very high at these very high temporal resolution ($R^2 = 0.97$). H_2O shows very similar correlation structures as CO_2 (not shown).

Summarizing, RA measures exhibit very similar structures for decimated versus aggregated values for short temporal resolutions. This indicates that these resolutions do not convey additional information. Thus, sampling rates seem to be higher than necessary.

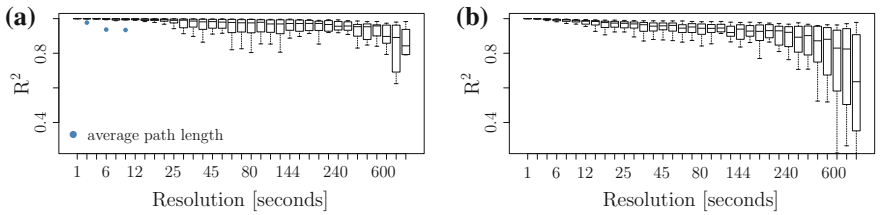

Fig. 16.6 R^2 of all RA measures' correlations between aggregation and decimation for high resolution data of **a** CO_2 and **b** vertical wind velocity

For oversampled data, it would be expected that the aggregation method and the decimation method are on the one to one line without loosing any structures as derived by the RA measures. Normally aggregation and decimation are two very different techniques of changing temporal resolution, thus should exhibit different structures. Temporal resolutions shorter than 6 s (vertical wind) or shorter than 20 s (CO_2) are lying on the one to one line. Thus, sampling rate of vertical wind speed could be reduced to 6 s, maintaining at least 98 % of the structures as derived from 6 s averaged vertical wind speed. CO_2 concentrations maintain at least 93 % of the structures at 20 s temporal resolution and 91 % at 30 s.

These results are in concordance to literature, stating that the correlation between heat fluxes derived from 10 Hz data and disjunct eddy sampling (30 s) is considered to be very good with $R^2 = 0.94$ [25]. In addition, the loss of structures due to decimation (or disjunct sampling) as derived by RA is very low, which leads to the conclusion that disjunct eddy accumulation is a very good alternative to derive fluxes of specific chemical compounds, which cannot be measured with higher sampling rates. As this loss of structures is lower for concentrations (CO_2, H_2O) than for vertical wind speed, it is assumed that only large changes in vertical wind speed, possibly related to larger eddies, are actually transporting concentration differences. It has already been stated by [24] that a sampling frequency of several seconds would suffice for EC data, but it seems currently unlikely that a change in the sampling frequency of EC data will occur anytime soon.

16.5 Fluxes at NL-Loo

In this section, we analyze EC fluxes of the NL-Loo site which are provided at 30 min resolution but extend over several years. Similar to the high resolution data, we first fix the parameters used for embedding and RA.

16.5.1 Fixing Embedding and RA Parameters

We fix the time delay after removing the diurnal and the annual cycle in the data with singular spectrum analysis, to get results which are not constrained by these cycles.

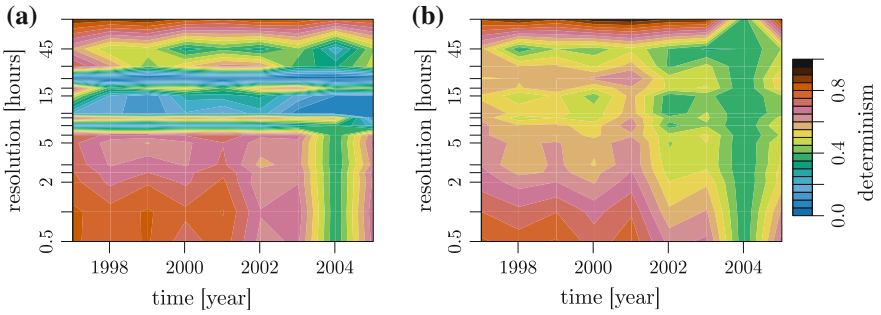

Fig. 16.7 Determinism of NEE_f to infer an optimal temporal resolution (NL-Loo). **a** Aggregation, **b** Decimation

After fixing the delay, the embedding dimension is found and RA performed on the original data, containing the cycles.

As *MI* and *ACF* still show structures depending on the diurnal cycle, and thus indicate its insufficient removal, we use the length of the autocorrelation function (*LAC*) to determine the time delay. *LAC* is the time lag at which the autocorrelation function falls below a 5 % significance threshold of 0.015 taking the length of the time series into account. *LAC* varied slightly around 180 lags. The corresponding time period is used as time delay: $\Delta T = 90$ h $= 3.75$ days. In concordance with the high resolution data, an embedding dimension of 5 and a fixed recurrence rate of 10 % is used.

16.5.2 Results and Discussion

Temporal Resolution

We examine the impact of varying the temporal resolution on the RA measures. A careful analysis, similar to that for the high resolution data (Sect. 16.4.2[3]), revealed that an aggregation up to 6 h resolution does not affect the results of the recurrence analysis: the correlation between RA measures of aggregated and decimated EC fluxes is higher than 95 %. For longer temporal resolutions, however, the correlation shows a steep drop, which is obvious also through visual inspection of the RA measures. As a typical example, we show the *Determinism* of NEE_f for the aggregation method (Fig. 16.7a), which exhibits horizontal structures at certain temporal resolutions. In the case of decimation (Fig. 16.7b), these are less pronounced. These structures occur when the resulting sampling frequency is in disconcordance to the diurnal cycle, which is known as *aliasing*. That there is a stronger aliasing effect for aggregation than for decimation may be due to the fact that decimation quickly

[3]We used divisors of 90 h with half hour accuracy: 0.5, 1, 2, 2.5, 3, 5, 6, 7.5, 9, 10, 15, 18, 22.5, 30, 45, and 90 h.

destroys the correlation structure of the time series (into the direction of random data) whereas aggregation smooths the autocorrelation function. This has not been tested systematically, however.

At high resolution (less than around 2–3 h), for both methods a shift in the dynamical regime is visible around the year 2001. Determinism is systematically lower after this year. The second half of 2003 and the first half of 2004 is singled out as a special period with a rather narrow range of determinism at intermediate values. This could be related to the mid-European heatwave of 2003.

Beyond visual inspection

A more quantitative perspective towards the behaviour of RA measures is provided by producing time series for them by running through moving windows. Here, we aggregated the flux time series to 6 h resolution, fixed the embedding dimension to 5 and the time delay to $\Delta T = 3.75$ days as before. Moving windows of exactly one year length were chosen with 75 % overlap (shifting 91 days each time), and a fixed recurrence rate of 10 % was chosen.

To assess the significance of individual structures in the time series of the RA measures, one should estimate the confidence bounds against a "trivial" case. Here, we calculated confidence bounds based on a bootstrap approach, where the line length probability distributions are kept intact, but data are resampled with replacement otherwise. This approach is available for RQA measures only [31]. The confidence bounds are visualized as grey-shaded areas in Fig. 16.8.

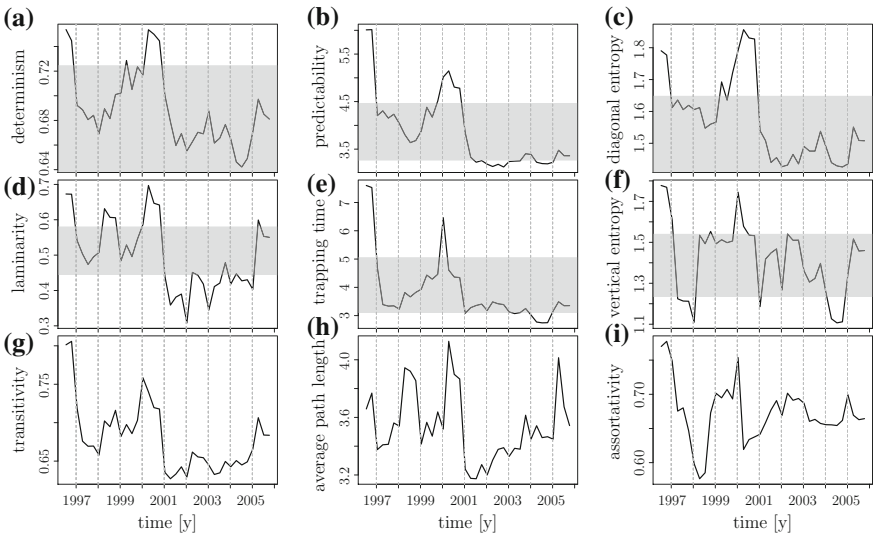

Fig. 16.8 RQA measures *DET*, *PRE*, *ENT*, *LAM*, *TT* with 5 % and 95 % confidence interval after [31] and RNA measures $\mathcal{T}$, $\mathcal{L}$, $\mathcal{A}$ for NEE_f at NL-Loo

Very high values at the very beginning of the investigated period are related to initial implementation problems. The nine RA measures presented share a common peak at the beginning of the year 2000, which is significant in all (RQA) cases. After that, the NEE_f dynamics changes behaviour qualitatively, e.g. predictability or trapping time show rather low values, indicating more random behaviour. These features are not related to visually obvious changes in the time series from the site. Whether they are indicators of proper dynamical transitions in the system or a mere sign of problems with the sensors at the tower, remains an open question at the moment. But Fig. 16.8 demonstrates that RA is capable of detecting subtle changes in time series which easily escapes other methods.

16.6 Conclusions

It could be demonstrated that recurrence analysis is able to describe and quantify various dynamical properties of time series produced from Eddy Covariance measurements at flux towers.

Specifically it was hypothesized, that RA is capable to (1) discern periods where the atmospheric conditions are particularly suitable to obtain reliable EC fluxes, (2) to detect dynamical transitions beyond those obvious from visual inspection and (3) to recommend an optimal time resolution for measuring EC data.

With respect to (1) especially one prerequisite for reliable EC fluxes has to be mentioned, which is stationarity. It has been tested by comparing the sum of six 5-min fluxes with the 30-min flux [30]. A new approach would be to consider the RA measure *trend* (density of recurrence points along diagonals). *Trend* indicates instationarity of the data and could serve as an additional test for stationarity which is more sensitive than the well established method.

Dynamical transitions (2) can be visually detected by inspecting numerous meteorological variables together. In EC high resolution data changes of the behaviour of vertical wind speed are indicated by an increase in average path length of a recurrence network. These increases can be attributed to changes in weather, which are only observable with a number of meteorological variables (air temperature, wind direction, stability of the atmosphere). This RA perspective of dynamical transitions has the advantage of being more objective, directly pointing to the interesting period of time.

For long-term flux data, both unusual periods as well as more long-lasting changes in dynamics can be observed. This may be used to detect changes in system properties as well as deficient measurement devices. As with many other analysis techniques, an attribution cannot be delivered by RA per se.

With respect to changes of temporal resolution (3), RA is capable to recommend an optimal temporal resolution for further analysis. In EC high resolution data, RA on different temporal resolutions (aggregated and decimated data) exhibits only very little differences of structures in phase space between aggregation and decimation.

This is astonishing, as aggregation and decimation are two very different methods of changing the temporal resolution by using the average or sum (aggregation) or leaving values out (decimation). This indicates that EC high resolution data are currently sampled on rates higher than necessary. This oversampling is more pronounced for CO_2 concentrations (<20 s) than for vertical wind speed (<6 s). In the disjunct eddy accumulation method, slow sensors (with a reaction time of 30 s) are already used to determine EC fluxes of chemical compounds like α-pinene (e.g. [32]). Also for standard EC measurements of CO_2 and H_2O, the oversampling can be confirmed by RA. This has already been stated by [24], but seems to be not the case at issue at the moment, as memory and data transfer are rather cheap. Nevertheless, one should keep this in mind, e.g. for EC experiments in remote areas or for analysis of the EC high resolution data, which suffers from long computation times.

Investigating the relations between the patterns found with recurrence analysis and fundamental processes is an interesting option for future research. Another direction to follow is the usage of RA for quality checking and assessment of EC data, a relevant activity e.g. in measurement networks such as ICOS

Acknowledgments We would like to thank C.L. Webber, Jr. for the invitation to contribute a chapter to this book on Recurrence Analysis after one of us (H.L.) gave a presentation at the Recurrence Symposium in Grenoble, France. We would like to acknowledge the FLUXNET data providers and the organizers of the FLUXNET database as outcome of the La Thuile FLUXNET workshop 2007, which would not have been possible without the financial support provided by CarboEurope-IP, FAOGTOS-TCO, iLEAPS, Max Planck Institute for Biogeochemistry, the American National Science Foundation, University of Tuscia, and the U.S. Department of Energy. The Berkeley Water Center, Lawrence Berkeley National Laboratory, Microsoft Research eScience, Oak Ridge National Laboratory developed the database and supported technically. Two anonymous reviewers provided valuable hints for improvement.

References

1. A. Serafimovich, F. Eder, J. Hübner, E. Falge, L. Voß, M. Sörgel, A. Held, Q. Liu, R. Eigenmann, K. Huber, H. F. Duarte, P. Werle, E. Gast, S. Cieslik, L. Heping, T. Foken, in *ExchanGE Processes in Mountainous Regions (EGER): Documentation of the Intensive Observation Period* (IOP3) June, 13th to July, 26th 2011 (Arbeitsergebn, Univ Bayreuth, Abt Mikrometeorol, 2011), pp. 47–135. ISSN 1614-8916
2. P. Gerstberger, T. Foken, K. Kalbitz, in *The Lehstenbach and Steinkreuz Catchments in NE Bavaria, Germany*, ed. by E. Matzner. Ecological Studies, vol. 172 (Springer, 2004), pp.15–41
3. M. Aubinet, T. Vesala, D. Papale, *Eddy Covariance: A Practical Guide to Measurement and Data Analysis* (Springer, Dordrecht Heidelberg London New York, 2012)
4. D. Papale, M. Reichstein, M. Aubinet, E. Canfora, C. Bernhofer, W. Kutsch, B. Longdoz, S. Rambal, R. Valentini, T. Vesala, D. Yakir, Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation. Biogeosciences **3**, 571–583 (2006)
5. M. Reichstein, E. Falge, D. Baldocchi, D Papale, M. Aubinet, P. Berbigier, C. Bernhofer, N. Buchmann, T. Gilmanov, A. Granier, T. Grünwald, K. Havrankova, H. Ilvesniemi, D. Janous, A. Knohl, T. Laurila, A. Lohila, D. Loustau, G. Matteucci, T. Meyers, F. Miglietta, J.-M. Ourcival, J. Pumpanen, S. Rambal, E. Rotenberg, M. Sanz, J. Tenhunen, G. Seufert, F. Vaccari, T. Vesala,

D. Yakir, R. Valentini, On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Glob. Change Biol. **11**(9), 1424–1439 (2005)
6. A.M. Moffat, D. Papale, M. Reichstein, D.Y. Hollinger, A.D. Richardson, A.G. Barr, C. Beckstein, B.H. Braswell, G. Churkina, A.R. Desai, E. Falge, J.H. Gove, M. Heimann, D. Hui, A.J. Jarvis, J. Kattge, A. Noormets, V.J. Stauch, Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. Agricul. Forest Meteorol. **147**(3–4), 209–232 (2007)
7. F. Takens, Detecting strange attractors in turbulence, in *Dynamical Systems and Turbulence*, ed. by D.A. Rand, L.-S. Young, Lecture Notes in Mathematics, vol. 898 (Springer, 1981), pp. 366–381
8. M.B. Kennel, R. Brown, H.D.I. Abarbanel, Determining embedding dimension for phase-space reconstruction using a geometrical construction. Phys. Rev. A **45**(6), 3403–3411 (1992)
9. C.L. Webber Jr., N. Marwan (eds.), *Recurrence Quantification Analysis* (Springer, Understanding Complex Systems. Cham Heidelberg New York Dordrecht London, 2015)
10. K.W. Hipel, A.I. McLeod, *Time Series Modelling of Water Resources and Environmental Systems* (Elsevier, Amsterdam, 1994)
11. L. Cao, Practical method for determining the minimum embedding dimension of a scalar time series. Physica D **110**(1), 43–50 (1997)
12. A.M. Fraser, H.L. Swinney, Independent coordinates for strange attractors from mutual information. Phys. Rev. A **33**, 1–7 (1986)
13. H. Lange, Are Ecosystems dynamical systems? Int. J. Comput. Anticipatory Syst. **3**, 169–186 (1998)
14. J.P. Zbilut, C.L. Webber Jr., Embedding and delays as derived from quantification of recurrence plots. Phys. Let. A **171**, 199–203 (1992)
15. C.L. Webber Jr., J.P. Zbilut, J. Dynamical assessment of physiological systems and states using recurrence plot strategies. J. Appl. Physiol. **76**, 965–973 (1994)
16. N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, J. Kurths, Recurrence-plot-based measures of complexity and their application to heartrate—variability data. Phys. Rev. E **66**(2), 026702 (2002)
17. N. Marwan, M. Carmen Romano, M. Thiel, J. Kurths, Recurrence plots for the analysis of complex systems. Phys. Rep. **438**(5–6), 237–329 (2007)
18. R.V. Donner, Y. Zou, J.F. Donges, N. Marwan, J. Kurths, Recurrence networks—a novel paradigm for nonlinear time series analysis. New J. Phys. **12**(3), 033025 (2010)
19. R.V. Donner, J.F. Donges, Y. Zhou, J.H. Feldhoff, Complex network analysis of recurrences, in *Recurrence Quantification Analysis*, ed. by C.L. Webber Jr., N. Marwan (Springer, Understanding Complex Systems. Cham Heidelberg New York Dordrecht London, 2015), pp. 101–163
20. N. Golyandina, A. Zhigljavsky, *Singular Spectrum Analysis for Time Series* (Springer, Springer Briefs in Statistics. Heidelberg New York Dordrecht London, 2013)
21. H. Lange, Recurrence quantification analysis in watershed ecosystem research. Int. J. Bifurcat. Chaos **21**(04), 1113–1125 (2011)
22. H. Lange, S. Boese, Recurrence quantification and recurrence network analysis of global photosynthetic activity, in *Recurrence Quantification Analysis*, ed. by C.L. Webber Jr., N. Marwan (Springer, Understanding Complex Systems. Cham Heidelberg New York Dordrecht London, 2015), pp. 349–374
23. G. Fratini, M. Mauder, Towards a consistent eddy-covariance processing: an intercomparison of EddyPro and TK3. Atmos. Meas. Tech. **7**(7), 2273–2281 (2014)
24. D.H. Lenschow, J. Mann, I. Kristensen, How long is long enough when measuring fluxes and other turbulence statistics. J Atm Oceanic Techn **11**, 661–673 (1993)
25. H.J.I. Rinne, A.C. Delany, J.P. Greenberg, A.B. Guenther, A true eddy accumulation system for trace gas fluxes using disjunct eddy sampling method. J. Geophys. Res. **105**(24), 791–798 (2000)
26. J. Rinne, C. Ammann, Disjunct eddy covariance method, in *Eddy Covariance: A Practical Guide to Measurement and Data Analysis*, ed. by M.Aubinet, T. Vesala, D. Papale (Springer, Dordrecht Heidelberg London New York, 2012), pp. 291–307

27. J.C. Kaimal, J.C. Wyngaard, Y. Izumi, O.R. Cot, Spectral characteristics of surface-layer turbulence. Q. J.Royal Meteorol Soc. **98**, 563–589 (1972)
28. U. Ligges, S. Krey, O. Mersmann, S. Schnackenberg, tuneR: analysis of music (2014), http://r-forge.r-project.org/projects/tuner/
29. N. Marwan, J.F. Donges, Y. Zou, R.V. Donner, J. Kurths, Complex network approach for recurrence analysis of time series. Phys. Let. A. **373**(46), 4246–4254 (2009)
30. T. Foken, B. Wichura, Tools for quality assessment of surface-based flux measurements.Agric. For. Meteorol. **78**, 83–105 (1996)
31. N. Marwan, S. Schinkel, J. Kurths, Recurrence plots 25 years later—gaining confidence in dynamical transitions. EPL (Europhysics Letters) **101**(2), 20007 (2013)
32. S. Haapanala, H. Hakola, H. Hellén, M. Vestenius, J. Levula, J., Rinne, Is forest management a significant source of monoterpenes into the boreal atmosphere?' Biogeosciences **9**(4), 1291–1300 (2012)

Chapter 17
Recurrence Plots for the Analysis of Combustion Dynamics

Lipika Kabiraj, Aditya Saurabh, Holger Nawroth, C.O. Paschereit, R.I. Sujith and Nader Karimi

Practical applications involving combustion suffer from serious issues such as combustion instabilities and sudden loss of flame: flame flashback and blowout. These phenomena are related to dynamical changes in the combustion system. Here, we summarize our recent studies on the application of recurrence-based methods to identify such dynamical transitions as well as for the characterization of combustion dynamics in laboratory combustors.

17.1 Introduction

Natural phenomena are governed by laws that are ubiquitously nonlinear, and their dynamics are often characterized by—and often hidden due to—the presence of inherent noise. We now know that the dynamics of unrelated naturally-occurring and artificial dynamical systems can be strikingly similar to each other in nonlinear fea-

L. Kabiraj (✉) · A. Saurabh · H. Nawroth · C.O. Paschereit
Chair of Fluid Dynamics, Hermann Föttinger Institute,
Technische Universität Berlin, Berlin, Germany
e-mail: lipika.kabiraj@tu-berlin.de

A. Saurabh
e-mail: aditya.saurabh@tu-berlin.de

H. Nawroth
e-mail: h.nawroth@tu-berlin.de

C.O. Paschereit
e-mail: oliver.paschereit@tu-berlin.de

R.I. Sujith
Department of Aerospace Engineering,
Indian Institute of Technology Madras, Chennai, India
e-mail: sujith@ae.iitm.ac.in

N. Karimi
School of Engineering, University of Glasgow, Glasgow, UK
e-mail: nader.karimi@glasgow.ac.uk

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8_17

tures such as bifurcations, scenarios (routes to chaos), and the effects of noise. This success in characterization and identification of dynamical similarities between nonlinear systems has been made possible by techniques from dynamical systems analysis, most particularly, recurrence analysis, which is the subject of this book. What is more important is that recurrence analysis and related techniques have allowed us to extract useful information from data that seem highly irregular or noisy upon direct observation.

A subset of natural phenomena that we are dealing with involves complex coupling between hydrodynamics, acoustics and combustion processes. More specifically, the results in this chapter are concerned with combustion dynamics—and related issues—observed in combustors such as those employed in practical systems like the gas-turbine engine.

Three particular cases of combustion dynamics form the focus of this chapter: (a) acoustic noise emitted by turbulent flames, (b) the nature of combustion (thermo)-acoustic coupling, and (c) undesirable extreme events: flame blowout and flashback. These phenomena obtain their characteristic features from a combination of deterministic and non-deterministic processes. In the past, studies have been based primarily on statistical analysis of the time series of observations and frequency-domain analysis. The last of the three cases—extreme events—has been particularly difficult to handle using these methods. In this chapter, we are summarizing recent advances in characterizing complex cases of combustion dynamics through the incorporation of recurrence-based methods [1–4]. We will be seeing how this new strategy helps obtain new information and brings a new perspective of looking at the mentioned issues.

The reason it is important to go further into the details of the physical processes and (nonlinear) mechanisms related to our system is not only academic; our research also has a huge practical motivation. A better understanding of combustion dynamics and an improved ability to tackle related issues—thermoacoustic instability, flame blowout and flashback—will enable the enhancement of the operating envelope of combustors towards a regime that is favorable for higher efficiency and lower pollutant emissions, and help the objective of controlling noise emissions [5–7]. Our research is hence a step in the direction of sustainable development.

The implementation of nonlinear time series on combustion dynamics has provided several important results in the past. Strahle and co-authors [8–10], for instance, have explored fractal analysis of time series obtained from experiments on combustion noise. They found that such nonlinear analysis is able to provide useful information from highly noisy time series. Furthermore fractal characterization of time series yielded quantitative measures that could be used to characterize system dynamics and could even be used for modeling, analysis, and validation. Their work on turbulent combustion noise was inspired by previous studies on turbulent flows and flames [11, 12]. Nonlinear methods, notably phase space based characterization [13, 14] are increasingly being incorporated in studies on laminar [15] and turbulent combustion [16–21], internal combustion engines [22] as well as related fields [23, 24]. These techniques allow investigation through an alternate perspective: the phase space. For this reason, they have been quite successful in gaining addi-

tional insight into system dynamics from scalar time series acquired experimentally or from numerical analysis. Here, we look at results from our studies on combustion dynamics, which deal particularly with recurrence analysis.

17.2 Case Studies

In the following section, we have presented the case studies to illustrate where and how the application of recurrence plots and recurrence quantification analysis, together with the conventional analysis such as proper orthogonal decomposition and power spectral density estimates, has led to additional information on the system. First we have dealt with the case of the noise emission from a turbulent flame: combustion noise. In the second section we have presented a study of nonlinear thermoacoustic oscillations in a prototypical laboratory combustor, which consists of a confined, laminar premixed flame. We have discussed dynamical transitions of the self-excited instability from limit cycle to an intermittent state (and subsequently blowout). In the last section, we have discussed a case where thermoacoustic instability induces flame flashback. Thermoacoustic instability, flame blowout, and flame flashback are serious engineering issues for combustors. In worst cases, these have even led to complete system failure.

17.2.1 Unveiling the Spatio-Temporal Nature of Combustion Dynamics for Turbulent Flame

Flames can be laminar or turbulent depending on the base flow. A steady candle flame is laminar, whereas, combustion within an aircraft engine involves turbulent flames. While turbulent flames are essential whenever high power output from combustion is required (turbulent fluctuations enhance mixing and support flame stabilization), the turbulent nature of the flame is the source of acoustic emissions, which must be suppressed (consider that combustion is a dominant source of noise in aircraft engines).

Local fluctuations in the heat release rate in a turbulent flame lead to unsteady volumetric expansion and contraction at the flame front: acoustic noise. Noise emission from turbulent flames depends on several factors such as the flame type, shape and the fuel-oxidizer ratio, but most importantly, it depends on the base flow turbulence, whose spatial and temporal signature consists of random as well as coherent fluctuations [25]. The broadband and intermittent nature of combustion noise [7] has been well-established for a long time now. However, turbulent combustion and the noise emitted is a complex subject involving interactions between multiple processes, and therefore, modeling of turbulent combustion continues to be under active investigation. Investigations have primarily focused on understanding the physics underlying

combustion noise through scaling laws and empirical relations governing the sound pressure level and the spectral characteristics of far-field combustion noise.

By using recurrence techniques, we are essentially trying to incorporate the advantages of recurrence plots in distinguishing the distribution of deterministic and stochastic components in the signature of combustion noise—how the simultaneous presence of periodic components and noise is manifested in the temporal signature of combustion noise. Typically, strong coupling overshadows noise from turbulent combustion and the base flow. If the coupling is feeble, details of flow-induced acoustics in the presence of the flame can be studied. Recurrence plots were also found to reveal the intermittent presence of different deterministic processes characterized by different time scales.

The investigated case involves the analysis of acoustic emission from the experiments on an open bluff-body stabilized, turbulent premixed flame [26]. Experiments were carried on in an anechoic environment to minimize wall-reflections, and acoustic emission measurements were carried out using a 1/4 in. pressure microphone mounted at a distance of nine burner diameters from the burner axis. To observe changes in the acoustic emission due to the presence of the turbulent flame, the parameter varied was equivalence ratio (air-fuel ratio), ϕ. As the equivalence ratio was varied, changes in the flame shape were observed (the flame shapes are shown in Fig. 17.1). In the range of conditions investigated, the flame assumes two shapes: a planar V and an M shape. The V shape exists because of the presence of the wake of the bluff body, a horizontal tungsten rod. The flame assumes this shape at lower equivalence ratios ($\phi < 1.10$). As equivalence ratio is increased ($\phi > 1.2$), the flame attaches to the bluff body, as well as to the burner rim and thus, forms an M shape. At intermediate equivalence ratios, the flame shape switches between V and M shape. In addition to pure turbulent flame dynamics, the noise emissions acquired in these experiments include intermittent coupling of the flame with upstream acoustics. This, however, does not affect the analysis strategy or the demonstration of its importance. In addition to recurrence analysis of pressure time series, the coherent dynamics of the flame that leads to the deterministic structures in the recurrence plot are studied through Proper Orthogonal Decomposition, POD, of the high-speed flame images sequences. Additional results on the spatio-temporal characteristics of noise emission from the investigated rig have been discussed in other reports [27].

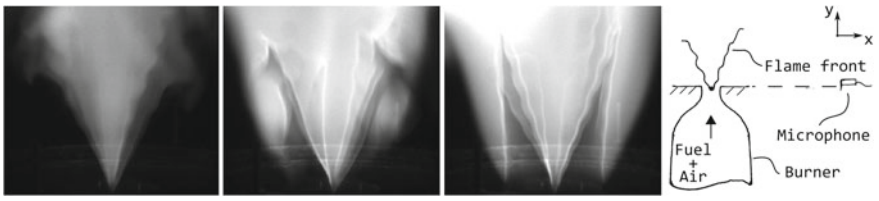

Fig. 17.1 Instantaneous high speed flame images showing various shapes observed during experiments. *Left* to *right* $\phi = 0.9$ (*V flame*), 1.15 (transition to *M shape*), 1.2 (*M flame*). A schematic outline of the experimental setup is shown in the *right* end frame

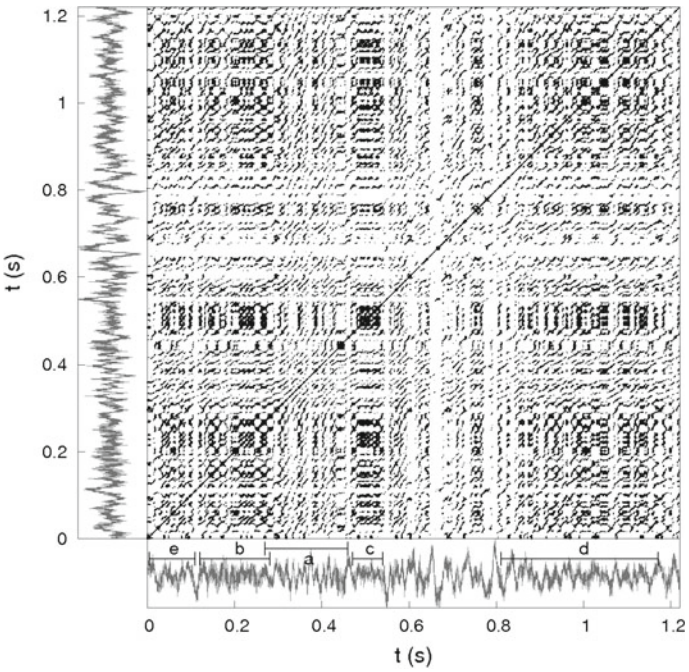

Fig. 17.2 RP for the isothermal flow [26] ($m = 5$, $\delta = 0.015$). Regions *a*–*e*, marked on the time series, indicate regions in the RP with characteristic patterns

To begin with, we will show the RP for the isothermal case (without the flame). The isothermal flow is expected to consist of noise due to turbulent nature of the base flow. However, the acoustic signal also has periodic components due to flow-induced acoustics in the burner at its resonance frequencies. The RP in Fig. 17.2 contains these features. The parameters, embedding dimension (m) and recurrence threshold (δ, in the same units as the corresponding pressure oscillations), used for constructing the RP are indicated in the figure captions. The chosen threshold results in 1 % recurrence rate [3].

The RP for noise emission during isothermal flow contains short line segments of varying lengths as well as isolated points. The resulting characteristic patterns have been marked in the corresponding time series as a, b, c, d and e. Along the diagonal, region 'a' contains diagonally-aligned line segments, which indicate the presence of periodic oscillations. Additionally, it can be seen that while entering and exiting this periodic stretch, these segments converge and diverge with respect to the diagonal. This could result from frequency modulation before and after a relatively periodic window. The average separation between diagonal lines is about 12–13 ms (i.e. 70–80 Hz). A high density of recurrence points is present in regions 'b' and 'c'. The lack of deterministic structures (line segments) indicates that these, and similar, regions are dominated by noise and high frequency dynamics. Regions 'd' and 'e' are also noise dominated. In addition, however, the presence of strong, relatively

slow modulations can be inferred from the large spacing between diagonally oriented patterns.

Thus, the recurrence plot obtained for a duration of ∼1.25 s detailed the different types of transient dynamics present in the pressure fluctuations. Periodic oscillations due to flow-induced acoustics in the burner at its resonance frequencies appear intermittently. This is most likely related to the unsteady nature of the inlet air flow and turbulence. In the following, we will discuss the signature of combustion noise at different equivalence ratios and compare the observations of individual cases with each other, and to the isothermal flow.

Figure 17.3 presents the RPs for four different cases in the order of increasing equivalence ratio. Corresponding spectra (see Ref. [26]) indicates that at $\phi = 0.9$ and

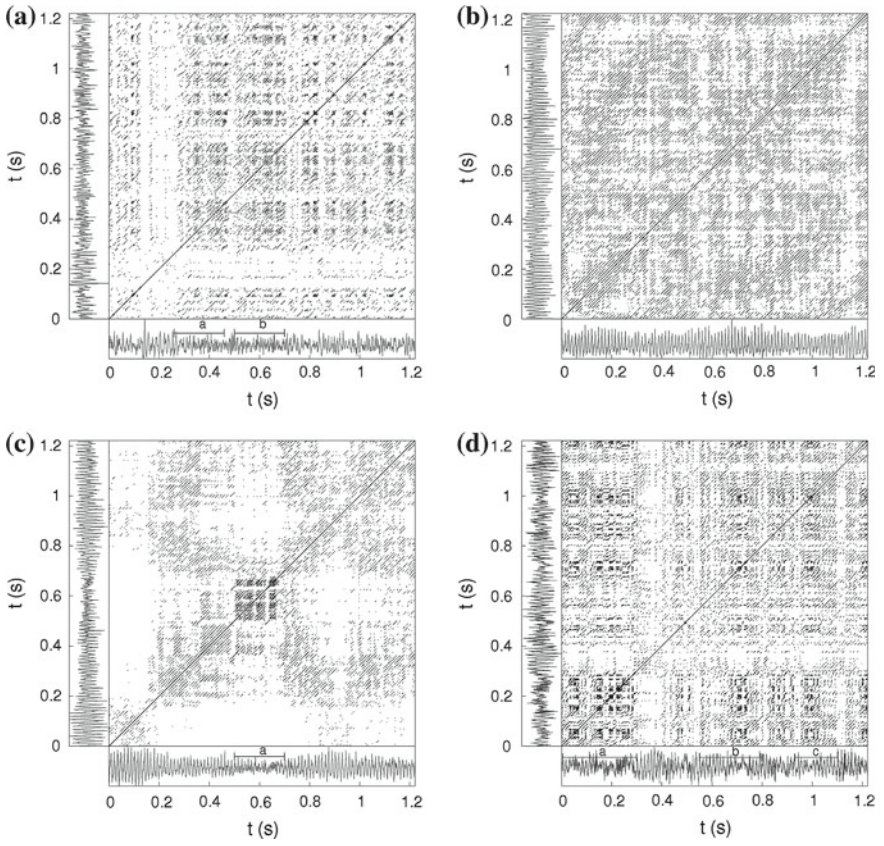

Fig. 17.3 RPs for reacting cases at different equivalence ratios [26]. It can be noticed that each plot clearly shows the contribution of periodic oscillations, noise and transitions between noise and periodic behavior as well as between different periodic states. Regions with characteristic patterns have been marked in the corresponding time traces. **a** Recurrence plot for $\phi = 0.9$ ($m = 5$, $\delta = 0.1$). **b** Recurrence plot for $\phi = 1.1$ ($m = 5$, $\delta = 0.15$). **c** Recurrence plot for $\phi = 1.15$ ($m = 5$, $\delta = 0.15$). **d** Recurrence plot for $\phi = 1.2$ ($m = 5$, $\delta = 0.05$)

higher equivalence ratio cases, $\phi = 1.2$ and higher, weak flame-acoustic coupling is present. For intermediate values, the coupling is stronger. We also have identified that the flame undergoes intermittent shape transitions between V and M shape (see Fig. 17.1 for flame shapes). With the help of RPs, more details of the coexistence of periodic and noisy windows in combustion noise data are obtained. In addition, transient behavior, such as intermittent flame shape transitions that are expected to be accompanied by modifications in noise emission could also be clearly identified as characteristic patterns in RPs.

At $\phi = 0.9$, the presence of a minor periodic component is manifested in the form of groups of short, diagonally-aligned line segments along the main diagonal (Fig. 17.3a). Noise interrupts the weak flame-acoustic coupling and leads to broken diagonal line segments. Evidence of small frequency variations, similar to the isothermal case, still exist (seen, for instance, in the regions 'a' and 'b' in Fig. 17.3a). Compared to the isothermal (Fig. 17.2), this case contains a smaller count of regions with high density recurrence points. This is a result of the impact of flame-acoustic coupling at 67 Hz. At $\phi = 1.1$, the periodic behavior is even more pronounced. This feature is identified in the RP in the form of a higher density of relatively long, equally-spaced, diagonal segments. The coupling is still unsteady and intermittent, as indicated by the dominance of broken diagonal lines and isolated points. For $\phi = 1.15$ (Fig. 17.3c), a V-M transition event is captured in the region demarcated as 'a'. While the plot contains the general features of the noise emission with a V flame (Fig. 17.3a, b), in region 'a'the flame assumes an M shape. This is due to the attachment of the flame tips to the burner rim, assisted by vortical structures formed due to flame-acoustic coupling, mentioned earlier. While the flame assumes an M shape, noise emission features low amplitude noisy periodic (high-frequency) behavior. This short change can clearly be identified in the RP.

Periodic events becomes less pronounced and high density recurrence regions begin to emerge again as the equivalence ratio is increased further. This is observed in the RP for $\phi = 1.2$ (Fig. 17.3d). Frequency variation with time, which was present in the isothermal case, can also be identified in this case within regions marked as 'a' 'b' and 'c' in Fig. 17.3d. The decrease in the periodic component observed in the RPs is consistent with the frequency spectra results (not shown). At $\phi = 1.2$, the flame assumes an M shape and the flame-acoustic coupling is weak. The patterns and features of the RP for this case bear similarity to those found in the RP for the isothermal case. This similarity between the RPs of the isothermal case and the high equivalence ratio case, $\phi = 1.2$, indicates that the noise emission with the M-flame configuration is largely due to flow induced noise. The most apparent difference between the two cases that can be identified is that the dark patches are relatively larger in size for isothermal flow. This tells us that the duration of episodes of high frequency noise are longer in the isothermal than in the reactive case. These dark patches obtained in the RP (Fig. 17.3d) actually originate from high frequency fluctuations that occur along the flame surface in the bluff-body-stabilized flame. The high frequency generates the closely spaced diagonal lines, which appear as dark patches when zoomed-out. These oscillations correspond to a frequency ~385 Hz,

which could be a higher acoustic mode of the burner or due to vortex shedding at the bluff body, whose characteristic frequency is also close to this frequency.

To identify fluctuations in the flame that could be responsible for the acoustic emission discussed above, Proper Orthogonal Decomposition (POD) of sequences of instantaneous, line-of-sight flame images was performed. POD is a routine technique for extracting coherent oscillations from scalar and vector field data obtained for turbulent flows and flames. Specifics of the implementation can be found in several references, for instance in the detailed article by Berkooz et al. [28]. A short description is as follows: Given a sequence of snapshots (instantaneous acquisition) of a scalar or vector field, the POD routine is used to extract mutually-orthogonal modes with the criteria that a linear combination of a given subset of the modes captures the maximum energy of the fluctuating field, which is possible using the number of modes in the subsets: If each snapshot is given by $u(x, t)$, the time-dependent field can be decomposed using POD as $u(x,t) = \sum_i a_i(t)\boldsymbol{\Phi}_i(x)$, where $a_i(t)$ and $\boldsymbol{\Phi}_i(x)$ are orthogonal temporal mode coefficients and the spatial mode shape respectively, such that, when the modes are arranged in a decreasing order according to the fluctuating energy content, mode 1 will represent the most dominant fluctuations, followed by mode 2, and so on. With such decomposition, a spatially propagating fluctuation will be captured by an orthogonal mode pair. For the results that follow, the sequence of instantaneous flame images was fed to the POD routine and the resulting spatial mode shapes, $\boldsymbol{\Phi}_i(x)$, have been discussed. The most energetic/dominant/relevant modes have been discussed and have been referred to by their corresponding numbers, i.

For the discussion of the dominant/coherent flame fluctuations, we chose $\phi = 1.2$, as this case was found to contain the characteristic fluctuations and the advantage of using RPs as a complementary analysis method emerges clearly. Dominant coherent oscillations identified for $\phi = 1.2$ are shown in Fig. 17.4. The first two mode pairs are presented. The spectra of the corresponding mode coefficients, $a_i(t)$ (see Ref. [26]), reveals that they correspond to low frequency (8 and 16 Hz respectively) vortical oscillations convecting along the flame.

The point to be noted is that although these modes are the most dominant coherent oscillations identified through POD, they do not contribute much to noise emissions. In the corresponding recurrence plot, Fig. 17.3d, the visible deterministic patterns correspond to a higher frequencies (identified as ∼70 Hz), which comes from the weak thermoacoustic coupling of the flame with the burner. These are higher order modes—specifically modes 5 and 6. While these low order modes might capture most of the fluctuations of the flame, there are several other components that are either intermittent, or low amplitude—hence, low in fluctuation energy—but still contribute to the noise emission and its characteristics. For instance, the dark patches in the RP Fig. 17.3d correspond to an even higher frequency (∼385 Hz). Coherent flame oscillations corresponding to this frequency are seen in relatively higher (less dominant) POD modes (modes 45–50), and are shown in Fig. 17.5. In fact, indication of such intermittent oscillations could also be seen in the instantaneous image (Fig. 17.1, third frame). Owing to the asymmetric structure, these modes could be related to vortex shedding at the bluff body.

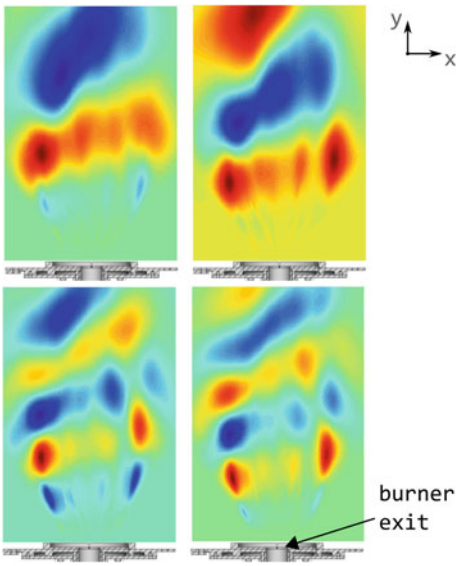

Fig. 17.4 The spatial structure (streamwise plane; see Fig. 17.1 for the burner schematic) of the most dominant flame surface fluctuation mode pairs ($\phi = 1.2$), extracted via POD (*red*—positive, *blue*—negative). *Top* modes 1, 2; *bottom* modes 3 and 4

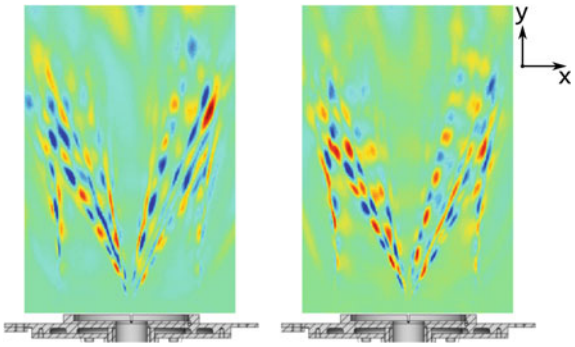

Fig. 17.5 The streamwise spatial structure of the higher order POD mode pair (mode numbers 48 and 49) for the same conditions as in Fig. 17.4. The corresponding (high frequency) oscillatory behavior was identified as dark patches in the RP (Fig. 17.3d). Oscillations corresponding to these modes appear intermittently for short durations. Accordingly, neither spectral analysis nor POD analysis could capture the significance of this mode

From the corresponding RP, Fig. 17.3d, it can clearly be seen that these modes are short and intermittent. The modes are clearly relevant, but due to their nature, POD and time series methods fail to identify their contribution. Hence, using POD, the mode structures can be identified. Simultaneous analysis of the RP constructed from

the time series of acoustic emission reveals which modes are relevant for combustion noise—including those that appear only as short bursts.

Thus, we see that recurrence analysis can be advantageous for the investigation of acoustic emissions from the turbulent flame, which is a result of unsteady heat release rate, hydrodynamically generated noise (shear layers, vortex generated noise) and acoustics generated in the burner (flow-induced and/or due to flame-acoustic coupling). Specific to the presented study, recurrence analysis has provided: (a) a basis for characterizing combustion noise and comparing cases among each other and to a reference case (the isothermal flow in this study) such that mode information about the underlying physics is obtained, and (b) identifying features that are hidden to conventional techniques.

17.2.2 RP Structures for Thermoacoustic Coupling and a Precursor to Flame Blowout

In this section, we discuss a very specific application of recurrence plots: identifying an apparently irregular dynamical state of combustion instability as type-II intermittency [29]. Intermittency [30] has traditionally been identified via the statistical features of the irregular "bursts" observed during intermittent dynamics. Due to limitation of the data acquisition with respect to the time-scales of the bursts (discussed in the following) that were observed in our system, this was not practical. Instead, following the report by Klimaszewska and Żebrowski [31], where the intermittency types were associated with specific patterns in the recurrence plots, we have been able to assign a type to the observed dynamical state of combustion instability. As we will see later, this particular state was found to be associated with repeated flame lift-off. Accordingly, the intermittent state also provides insights into how, in a practical setting, flame blowout would be induced by thermoacoustic coupling. The occurrence of the state thus provides an indication of an incipient flame blowout event—a precursor to flame blowout.

Thermoacoustic coupling (and the resulting instability) is associated with discrete acoustic modes of the combustor—unlike combustion noise, which leads to a broadband spectrum. Thermoacoustic coupling is governed by a feedback coupling between unsteady heat release rate from the flame and the acoustic field of the combustor, conventionally represented by a second-order differential equation involving a time-delay [32]: Fluctuations in the flame lead to acoustic oscillations, which disturb the flame after being reflected from the boundaries of the combustor. The time-delay involved is in the response of the flame to the perturbations. In addition to the time-delay, the feedback coupling is also nonlinear; The non-linearity is associated with the flame. In a practical system, the coupling also involves complex interactions between combustion, hydrodynamics, and acoustics.

Previously, the nature of thermoacoustic system was understood to be associated with Hopf (subcritical or supercritical) bifurcation and the resulting limit cycles. However, recently it has been shown in several reports that the dynamics of thermoa-

coustic coupling can be extremely complex including states such as quasi-periodicity and chaos. That a thermoacoustic system exhibit routes to chaos has been shown recently [15, 33]. We will show in the next section that the complexity in flame dynamics is what supports such nonlinear behavior of thermoacoustic coupling.

Even in the case study that we will discuss here, secondary bifurcation of the limit cycle leading to quasi-periodicity was observed. The quasi-periodic state then undergoes a subsequent bifurcation to another state characterized by intermittent, irregular bursts of oscillations.

17.2.2.1 Combustion Instability, Intermittency and Flame Blowout

The experiments were conducted on a simplified combustor configuration consisting of a laminar conical premixed single flame inside a closed-open cylindrical glass duct that acted as the combustion chamber [29]. Here, the control parameter for the experimental bifurcation analysis that was varied was the location of the flame with respect to the duct (see Fig. 17.6 for a schematic of the configuration). In this section, the focus is to illustrate the importance of recurrence analysis to our investigation; Only the intermittent oscillations that were observed and were found to be associated with the occurrence of flame blowout, are discussed in this part of the chapter.

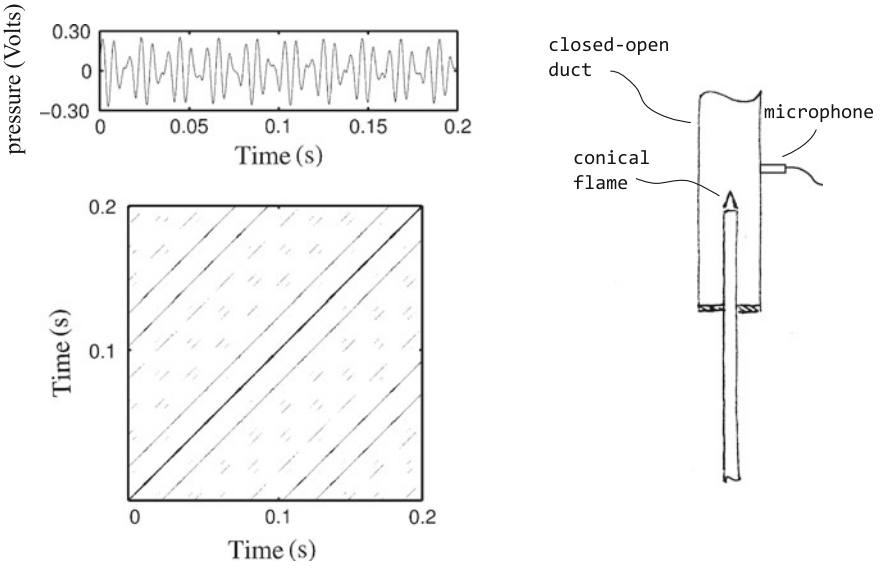

Fig. 17.6 For quasi-periodic oscillations it is seen that diagonal line segments are separated by unequal *vertical* spacings; a manifestation of irrationally related frequencies comprising the quasi-periodic state. A four dimensional space was used to construct the recurrence plot, with a specified recurrence threshold (δ) of 0.3 and 0.7 V for quasi-periodic oscillations. Figure reproduced from Ref. [29], with permission from the Journal of Fluid Mechanics. A schematic outline of the experimental setup is shown in the *right frame*

Further details of the experiment and extended discussion on the analysis and results can be found in [29, 34].

Limit cycle oscillations, which had appeared in the system as a result of subcritical Hopf bifurcation when the flame location was varied, underwent a secondary Hopf bifurcation, leading to a quasi-periodic state. The recurrence plot for the corresponding quasi-periodic oscillations is shown in Fig. 17.6. The presence of a second frequency, which is incommensurate to the frequency of the former limit cycle, in the quasi-periodic oscillations appears in the form of diagonally line segments separated by unequal vertical (and horizontal for the symmetric RP) spacings [35]. The thermoacoustic coupling—the source of these oscillations—involves feedback between acoustic fluctuations in the duct air column, and modulation of the flame surface area. While, for the limit cycle, the flame featured periodic flame surface modulations, the quasi-periodic state introduced new features: A sequence of instantaneous, line-of-sight, high-speed flame images for an approximate cycle acquired during this state are shown in Fig. 17.7. Flame surface area modulations during quasi-periodic oscillations involve flame elongation (b), neck-formation (c), pinch-off (d, e), and subsequent cusp formation (f), in addition to the overall modulation of the flame that was present during limit cycle oscillations. Arrows and circles mark the evolution of flame elongation and pinch-off events during the quasi-periodic oscillations.

The control parameter, the relative flame location, was varied further, and an apparently random bursting behavior was observed where oscillatory behavior repeatedly switched between large amplitude oscillations and relatively *quiet* states. The oscillating flame was again simultaneously recorded and was observed to constantly undergo a change in state from being attached to the burner tube, to leaving its stabilization—lifting-off—and oscillating a short distance from the burner (shown in Fig. 17.8). The sequence of instantaneous flame images is shown in Fig. 17.8. A physical reasoning behind this could be that, as the control parameter was being changed, the amplitude of quasi-periodic thermoacoustic oscillations was gradually increasing. At one parameter value, the amplitude increased sufficiently for the flame to loose stabilization. For this parameter, the flame constantly switched between a relatively steady attached state and an oscillatory lifted state. The switch occurred

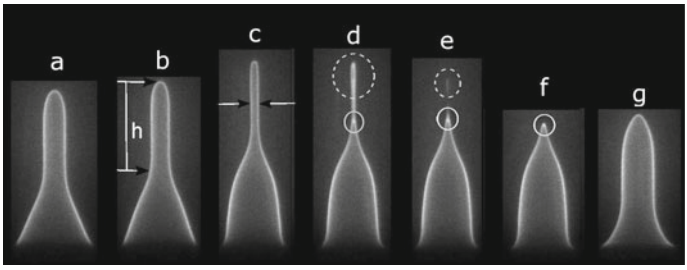

Fig. 17.7 Sequence of instantaneous flame images (framing rate, 5 kHz) at quasi-periodic oscillations. $a = 0$ ms, $b = 2.6$ ms, $c = 4$ ms, $d = 4.2$ ms, $e = 4.4$ ms, $f = 4.6$ ms and $g = 8$ ms. Figure reproduced from Ref. [29], with permission from the Journal of Fluid Mechanics

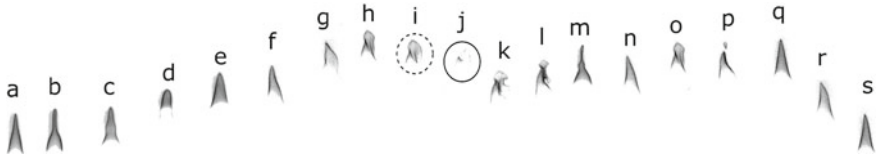

Fig. 17.8 Flame dynamics during intermittent oscillations are shown through instantaneous images. Images are arranged in the order of their occurrence (*left* to *right*). *Circles* on the instantaneous images highlight the characteristic stretching, folding (*i*), and local extinction (*j*) in the flame

intermittently, and at irregular intervals. The flame continued to be involved in thermoacoustic coupling even while it was lifted. The lifted flame oscillates in the circular jet transition region, which is usually around 5 burner tube diameters downstream from the burner exit plane. This indicates that jet flow dynamics could have a role to play during the lifted state. The irregularity in the oscillations is also clearly discernible.

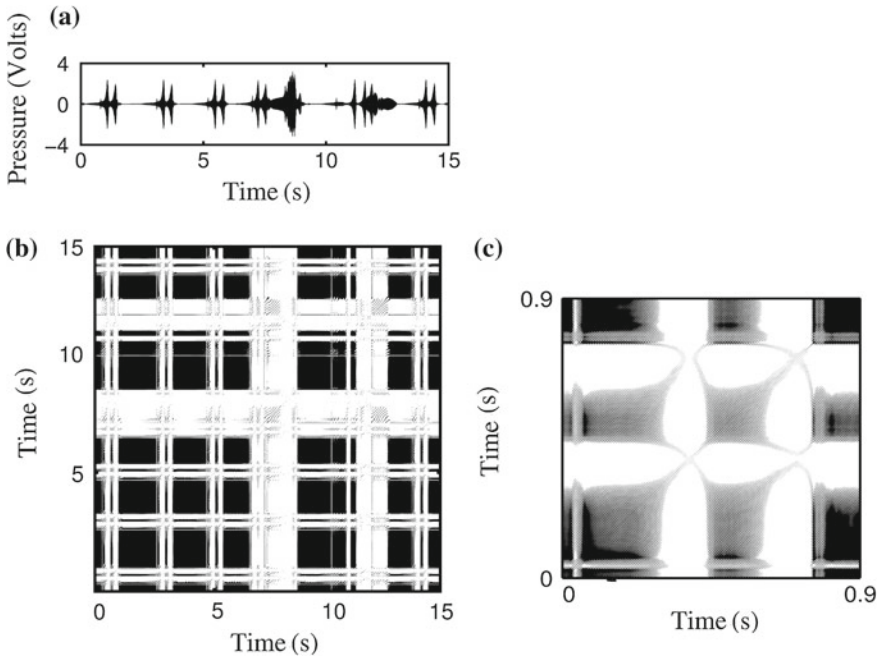

Fig. 17.9 Recurrence plot corresponding to intermittent oscillations (**b**). Embedding dimension = 4, $\varepsilon = 0.4$ V. In (**c**), a closer look into the patterns comprising the recurrence plot is presented. The end of laminar phases correspond to elongated structures, whose kite-like appearance indicates type-II intermittency [31]. Acoustic pressure amplitude in the time series (**a**) corresponding to recurrence plots is in Volts. Figure reproduced from Ref. [29], with permission from the Journal of Fluid Mechanics

This state appeared as a bifurcation of the quasi-periodic oscillations. It is of interest to identify whether the observed new state corresponds to a particular dynamical state; Intermittency [30], emerges as a possible candidate. More specifically, type-II intermittency, owing to the nature of the preceding bifurcations, could explain the observed change in behavior. Its presence however, is identified, in terms of the statistical characteristics of the observed bursts and the laminar state. The data, however, was not sufficient for statistical tests. Instead, we tested for the presence of type-II intermittency through the identification of features in the corresponding RP: We found that in our case the structure associated with the laminar state of the burst has resemblance with the kite-like structure previously identified for type-II intermittency [31]. This is illustrated in the RP in Fig. 17.9c. A closer look into the structure of the RP is given in Fig. 17.10. It is seen that before the burst, transition occurs from limit cycle behavior to quasi-periodic patterns—a trend similar to the bifurcation scenario. Transition between the burner-attached and the lifted states of

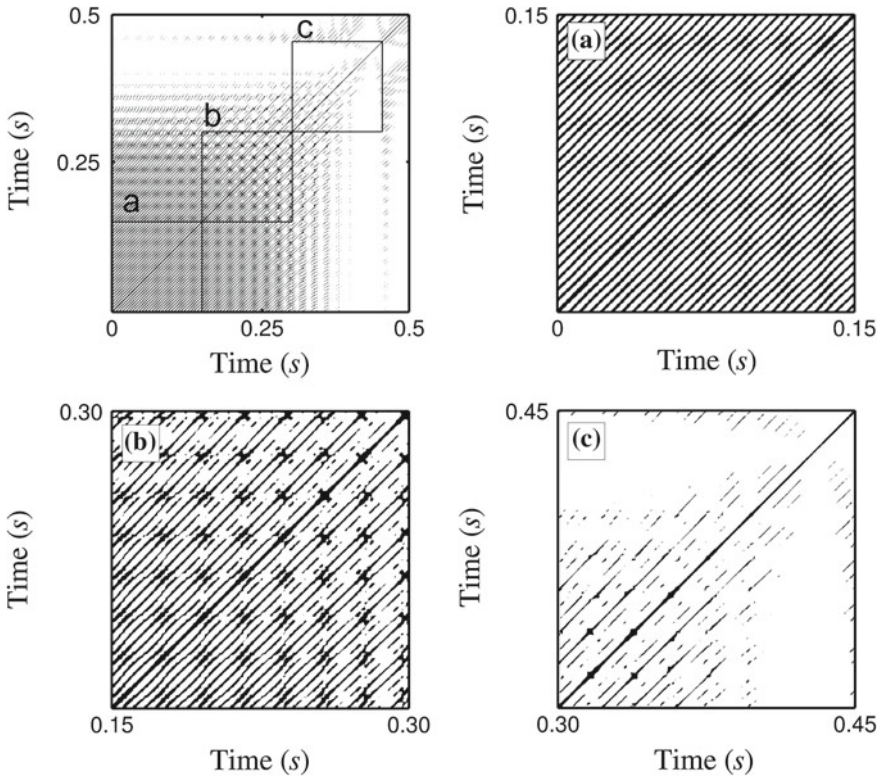

Fig. 17.10 Recurrence behavior of the system prior to a burst (*top-left* frame). The evolution of the system entering a burst state is analyzed by following the main diagonal. Windows *a–c* as marked indicate the transition of the system from limit cycle to quasi-periodic oscillations before the occurrence of a burst. Figure reproduced from Ref. [29], with permission from the Journal of Fluid Mechanics

the flame appears simultaneously. Further change in the control parameter results in an increased rate of burst occurrence and accordingly, reduced duration of laminar states—typical features of intermittency [30].

This identification of a known type of intermittency in the dynamics of thermoacoustic instability is an important result for at least three reasons: (a) the finding leads to the argument that the fundamental (nonlinear) interaction mechanisms behind the phenomenon are not unique to the system, but share significant features with other naturally-occurring and model nonlinear systems, (b) the dynamical state is related to, and could be considered a precursor to the extreme event of flame blowout in practical systems (i.e. the dynamical state could be used as a warning), and (c) the result demonstrates a previously unknown fact about thermoacoustic coupling.

In a recent analysis of a practical combustor configuration [36], the presence of type-II intermittency was identified for a various sets of operating conditions. The identification was again based on RP characterization. Intermittency in this case was proposed to be associated with the presence of multiple modes of combustion and multiple flames in the combustor.

The reason we were able to identify bifurcations and connect the intermittent oscillations to blowout is because the configuration we have studied does not include the noise and complexity of a practical system such as gas turbine combustor. Thus, it has been easier to obtain the dynamical state of intermittency and relate the same to flame-blowout. If the same investigation was to be performed on a more realistic configuration, in the early stages in the bifurcation scenario, where the burst frequency is low, flame destabilization would have caused complete flame extinction.

17.2.2.2 Flame Flashback

Flashback of a premixed flame is the unwanted upstream propagation of the reactive front. As a result, the flame moves into the premixing section of the burner. Once in the premixing section, the flame can become anchored, for example, in the wake of the fuel injection jets. Sustained combustion in the premixing section is undesirable because it results in increased NOx emissions and can also severely damage the burner [38]. Hence, it is important to ensure flashback does not occur. Ideally, this is easy to achieve as long as the combustor operation is limited to “safe” regimes. However, deviations from the prescribed operating range can practically occur. In addition, the current and foreseeable development trends in combustor technology for improved emissions and efficiency demand the prescribed operating regime to be expanded in the direction where the system is more prone to flashback. Thus, there has been a strong motivation to understand, detect, and control the issue of flashback. We have progressed only so far as to propose potential mechanisms for flashback to occur. However, flame flashback in turbulent flows involves complicated, nonlinear physical and chemical processes including aerodynamics and turbulence, chemical kinetics and unsteady transport of heat and mass. It is one of the most complex combustion dynamics problems and is yet to be sufficiently understood. In addition, flashback is a single transient event with a time span of a fraction of a second.

Flashback, like other phenomenon discussed above has been traditionally dealt with using time series and spectral analysis techniques. As expected, the short temporal span of the transitional period, most conventional methods, including spectral techniques, become inaccurate in analyzing the system. More importantly, such techniques are inherently incapable of detecting subtle features associated with the onset of flashback.

Motivated by the success of recurrence analysis in analyzing complex data [39], we have—with significant success—investigated time series associated with flashback. Through the implementation of recurrence plots and recurrence quantification analysis, we have found that prior to the actual flashback event, precursors can be detected. A part of our work in this direction is presented in this section.

For this study, a more generic combustor configuration [21, 37] involving a premixed turbulent flame is employed. Figure 17.11 shows a typical time series acquired during the flashback tests. The flashback event can be identified by the jump in the amplitude. Flashback, in the investigated configuration, occurs while the combustor is thermoacoustically unstable; It is likely that the instability has a contribution in inducing flame flashback. We would like to find out whether the onset of flashback can be identified prior to the observed jump in the amplitude. Also shown in the figure are zoomed-in views (b and c in Fig. 17.11) of the time series at two instants: one far away from the event and the other close to it.

Results of the recurrence quantification analysis (RQA) of the time series are presented in Fig. 17.12. The results have been obtained by sliding a 0.25 s window along the time series and calculating RQA measures for each window. The standard deviation is plotted for comparison. We find that the RQA measures of acoustic pressure oscillations start showing signs of change before the flashback event. In

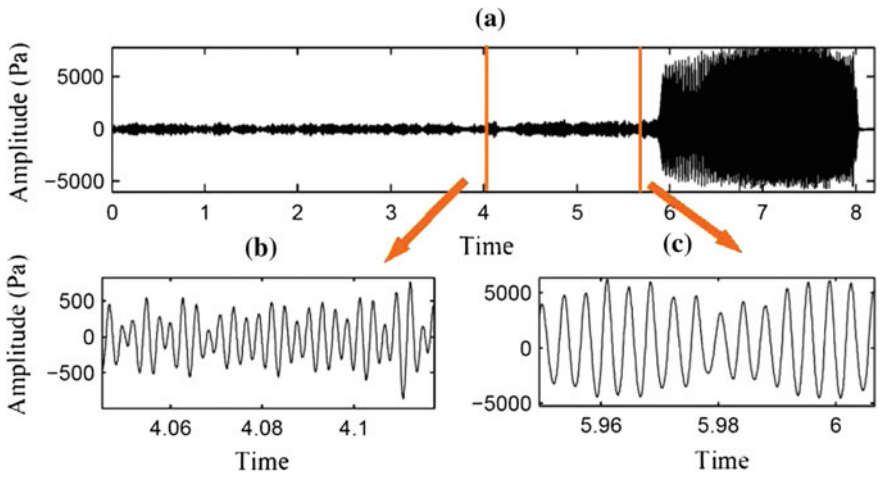

Fig. 17.11 **a** A typical acoustic signature observed during the combustion induced flame flashback event. **b**, **c** Zoomed-in views indicating a change in behavior prior to flashback

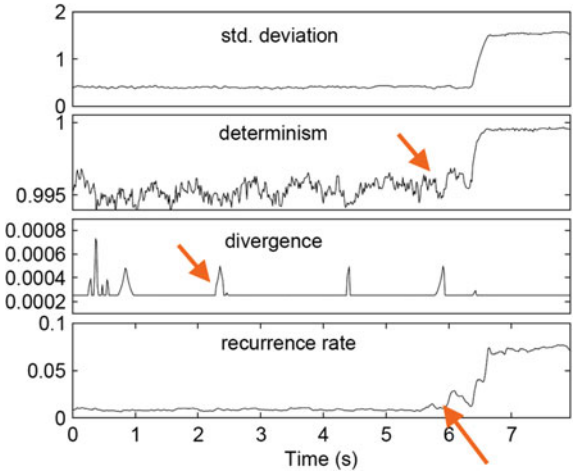

Fig. 17.12 RQA measures obtained with a sliding window of 0.25 s, for the time series in Fig. 17.11. Precursors/variation in the general trend of the measures are indicated by *arrows*

particular, the determinism and the recurrence rate start increasing gradually from the moment indicated by the arrows. The system is already under a rather deterministic state due to thermoacoustic instability, as indicated by the high value of determinism. RQA indicates that as flashback is approached, the determinism and—consequently—the recurrence rate increases. This result also supports the hypothesis that flashback in the studied configuration is induced by thermoacoustic coupling. It is not clear whether the jumps in the divergence measure are associated with dynamical changes in the system related to its proximity to flashback.

Thus, we find that recurrence analysis measures can be used to obtain additional information about system dynamics compared to conventional time series analysis methods. In specific, we find qualitative changes in the determinism of system dynamics, divergence of trajectories and the recurrence rate prior to flashback. In conclusion, our work illustrates that nonlinear methods, such as the RQA, can effectively identify the onset of extreme events such as flashback in combustion systems. This study was specific to the case of thermoacoustically-induced flashback. It remains to be seen if such results can be obtained for other cases, when thermoacoustic coupling is absent.

17.3 Conclusions

In this chapter we have discussed recent results dealing with the incorporation of recurrence plots and quantification analysis for select, outstanding issues in the investigation of combustion dynamics. Specifically, we evaluate the additional insights obtained by using recurrence methodologies in the process of (a) turbulent combustion noise; (b) thermoacoustic coupling and induced blowout; and (c) flame flashback. For combustion noise, recurrence plots provide a novel tool for characterizing

the signature of combustion noise, which consists of deterministic dynamics with multiple scales as well as non-deterministic dynamics, and comparing noise emission from different cases. Through recurrence analysis we could identify a specific, apparently irregular dynamical state of thermoacoustic coupling as type-II intermittency. The most promising application of recurrence analysis appears to be in the detection and characterization of transient events such as flashback and blowout—combustion phenomena where conventional methods are disadvantaged. In the flashback case investigated, recurrence quantification analysis is able to detect the onset of flashback in advance. Such analysis involving nonlinear time series techniques, in particular recurrence analysis, is new in the field and results indicate that its incorporation can enable a better description of combustion dynamics. In general, research on combustion dynamics will benefit significantly by following further developments in such nonlinear analysis techniques.

Acknowledgments L. K would like to acknowledge the research grant (No. KA 3968/1-1) granted by the German Research Foundation (DFG) to continue her research at the Technische Universität Berlin, Germany. The authors (L.K and A.S) would like to thank the organizing group of the Sixth Symposium on Recurrence Plots and particularly Charles L. Webber for initiating this interdisciplinary collection on the application of Recurrence Analysis, and inviting us to contribute.

References

1. J.P. Eckmann, S.O. Kamphorst, D. Ruelle. Europhys. Lett. **4**(9), 973 (1987)
2. J.P. Zbilut, A. Giuliani, C.L. Webber, Phys. Lett. A **237**, 131 (1998)
3. N. Marwan, J. Kurths, Phys. Lett. A **336**, 349 (2005)
4. N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, J. Kurths, Phys. Rev. E **66**, 026702 (2002)
5. H. Chiu, M. Summerfield, Acta Astronaut. **1**(7–8), 967 (1974)
6. S. Candel, Durox, S. Ducruix, Birbaud, N. Noiray, T. Schuller, Int. J. Aeroacoustics **8**(1), 1 (2009)
7. A.P. Dowling, Y. Mahmoudi, Proc. Combust. Inst. **35**(1), 65 (2015)
8. W.C. Strahle, J.I. Jagoda, Symp. (Int.) Combust., [Proc.] **22**(1), 561 (1989)
9. M.K. Ramachandra, W.C. Strahle, AIAA J. **21**(8), 1107 (1983)
10. W.C. Strahle, AIAA J. **29**(3), 409 (1991)
11. K.R. Sreenivasan, R.A. Antonia, Annu. Rev. Fluid Mech. **29**(1), 435 (1997)
12. F.C. Gouldin, K.N.C. Bray, J.Y. Chen, in *Western States Section of the Combustion Institute Paper* (1988), pp. 88–37
13. H. Kantz, T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge Press, 2004)
14. H.D.I. Abarbanel, R. Brown, J.B. Kadtke, Phys. Rev. A **41**(4), 1782 (1990)
15. L. Kabiraj, A. Saurabh, P. Wahi, R.I. Sujith, Chaos **22**(2), 023129 (2012)
16. J.D. Sterling, Combust. Sci. Technol. **89**, 167 (1993)
17. M. Gorman, M. el Hamdi, B. Pearson, K.A. Robbins, Phys. Rev. Lett. **76**, 1 (1996)
18. H. Gotoda, T. Miyano, I.G. Shepherd, Phys. Rev. E **81**(2) (2010)
19. V. Nair, R.I. Sujith, J. Fluid Mech. **747**, 635 (2014)
20. S. Domen, H. Gotoda, T. Kuriyama, Y. Okuno, S. Tachibana, Proc. Combust. Inst. **35**(3), 3245 (2015)
21. L. Kabiraj, A. Saurabh, N. Karimi, A. Sailor, E. Mastorakos, A.P. Dowling, C.O. Paschereit, Chaos **25**(2), 023101 (2015)

22. L.P. Yang, S.L. Ding, G. Litak, E.Z. Song, X.Z. Ma, Chaos **25**(1), 013105 (2015). doi:10.1063/1.4899056
23. J.P. Eckmann, D. Ruelle, Rev. Mod. Phys. **57**(3), 617 (1985)
24. A. Brandstaeter, J. Swift, H.L. Swinney, A. Wolf, J.D. Farmer, E. Jen, P.J. Crutchfield, Phys. Rev. Lett. **51**, 1442 (1983)
25. K.R. Sreenivasan, Annu. Rev. Fluid Mech. **23**(1), 539 (1991)
26. L. Kabiraj, A. Saurabh, H. Nawroth, C.O. Paschereit, AIAA J. **53**(5), 1199 (2015)
27. H. Nawroth, A. Saurabh, C. Paschereit, F. Zhang, P. Habisreuther, H. Bockhorn, in *42nd AIAA Fluid Dynamics Conference and Exhibit* (American Institute of Aeronautics and Astronautics (AIAA), 2012). doi:10.2514/6.2012-3072
28. G. Berkooz, P. Holmes, J.L. Lumley, Annu. Rev. Fluid Mech. **25**, 539 (1993)
29. L. Kabiraj, R.I. Sujith, J. Fluid Mech. **713**, 376 (2012)
30. Y. Pomeau, P. Manneville, Commun. Math. Phys. **74**, 189 (1980)
31. K. Klimaszewska, J.J. Żebrowski, Phys. Rev. E **80**, 026214 (2009)
32. P. Subramanian, R.I. Sujith, P. Wahi, J. Fluid Mech. **715**, 210 (2013)
33. L. Kabiraj, R.I. Sujith, P. Wahi, J. Eng. Gas Turbines Power **134**(3), 031502 (2012)
34. L. Kabiraj, R.I. Sujith, P. Wahi, Fluid Dyn. Res. **44**(3), 031408 (2012). doi:10.1088/0169-5983/44/3/031408
35. Y. Zou, Exploring recurrences in quasi-periodic systems. Ph.D. thesis, University of Potsdam (2007)
36. J.M. Wilhite, B.J. Dolan, L. Kabiraj, R.V. Gomez, E.J. Gutmark, C.O. Paschereit, Analysis of combustion oscillations in a staged MLDI burner using decomposition methods and recurrence analysis, in *54th AIAA Aerospace Sciences Meeting*, AIAA SciTech, (AIAA 2016–1156). http://dx.doi.org/10.2514/6.2016-1156
37. L. Christodoulou, L. Kabiraj, A. Saurabh, N. Karimi, Characterizing the signature of flame flashback precursor through recurrence analysis. Chaos **26**, 013110 (2016). doi:http://dx.doi.org/10.1063/1.4940154
38. A.C. Benim, K.J. Syed, in *Flashback Mechanisms in Lean Premixed Gas Turbine Combustion* (Elsevier BV, 2015), pp. 1–3
39. C.L. Webber, N. Marwan (eds.), *Recurrence Quantification Analysis* (Springer International Publishing, 2015)

Chapter 18
Recurrence Analysis of Turbulent Fluctuations in Magnetically Confined Plasmas

R.L. Viana, Dennis L. Toufen, Z.O. Guimarães-Filho, I.L. Caldas, K.W. Gentle and I.C. Nascimento

Abstract Recurrence plots and their quantification became a modern tool in nonlinear data analysis, currently being used in a myriad of scientific disciplines, a diversity that characterizes the Recurrence Plot Workshop since its beginning. In this work we review some applications of recurrence quantification analysis to data analysis in a fusion plasma, namely turbulent fluctuations in the plasma edge of Tokamak Chauffage Alfvén Brésilien tokamak, which is a magnetic confinement plasma machine. A similar analysis was performed in a low-density plasma device called Texas Helimak (University of Texas at Austin). Our results point out that the deterministic content of the fluctuations tends to increase as we approach the plasma edge, indicating a concentration of the recurrences therein. This favours the use of fluid models to describe the physics of the plasma edge turbulence.

R.L. Viana (✉)
Department of Physics, Federal University of Parana, Curitiba, Parana, Brazil
e-mail: viana@fisica.ufpr.br

D.L. Toufen
Instituto Federal de Educação, Ciência e Tecnologia. Campus de Guarulhos, Guarulhos, São Paulo, Brazil
e-mail: dennistoufen@yahoo.com.br

Z.O. Guimarães-Filho · I.L. Caldas · I.C. Nascimento
Institute of Physics, University of São Paulo, São Paulo, Brazil
e-mail: zwinglio@if.usp.br

I.L. Caldas
e-mail: ibere@if.usp.br

I.C. Nascimento
e-mail: inascime@if.usp.br

K.W. Gentle
Department of Physics and Institute for Fusion Studies, The University of Texas at Austin, Austin, TX 78712, USA
e-mail: k.gentle@utexas.edu

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8_18

18.1 Introduction

Recurrence plots are an extremely useful tool in nonlinear data analysis, particularly when the data series are short, noisy and even non-stationary, a very typical situation in realistic applications [1–3]. One promising field of study where recurrence plots are potentially advantageous is the analysis of fusion plasmas such those generated by tokamaks [4]. These advantages come from the complex nature of experimental signals from magnetically confined high-temperature plasmas, where we have a mixture of deterministic and stochastic features, and which can be sorted out by recurrence-based methods, unlike linear methods like Fourier analysis. Recurrence plots and recurrence quantification analysis have been applied to a wealth of problems involving technological low-temperature plasmas [5–7], space and geophysical plasmas [8] and fusion high-temperature plasmas [9–12].

Tokamaks are toroidal schemes of magnetic confinement of high-temperature plasmas, and it is thought to be the basis of a future thermonuclear fusion reactor, as in large-scale projects like the ITER (International Thermonuclear Experimental Reactor), currently being built in France by an international consortium [13]. ITER, which is to be completed in the next decade, will be the largest tokamak in the world, expected to generate 500 MW of power, ten times more than the power used to generate and confine the plasma in which the fusion reactions take place [14].

Once a plasma column is generated in a tokamak, one wants to confine it for a sufficiently large time, and for this to happen it is necessary that cross-field transport of particles and energy be controlled [15]. A major goal in the study of magnetically confined fusion plasmas has been to understand the causes and associated rates of anomalously large cross-field transport, which is thought to be caused by plasma turbulence [16]. One experimental signature of plasma turbulence in the plasma edge of a tokamak is the fluctuating behavior of the electrostatic floating potential, which often displays a broad fluctuation spectrum [17].

The deterministic content of the electrostatic fluctuations in the plasma edge has been assigned to physical mechanisms governed by nonlinear systems, like the interaction between drift waves which appear due to the steep density gradients in the plasma edge region of a tokamak [17, 18]. These drift wave interactions are known to depend critically on the radial position such that the electrostatic turbulent fluctuations should also exhibit some radial dependence.

However, the radial variation of the dynamical properties of the plasma turbulence may not be immediately apparent from the experimental data obtained in tokamaks. We used recurrence plots as a tool to quantify the recurrence properties of such series to provide a precise diagnostic of the dynamical properties of the system and other related measures of recurrence quantification analysis and their dependence on the radial position near the plasma edge [4, 19]. It turned out that this radial dependence, formerly elusive in conventional analysis using, e.g. spectral methods, is best observed using recurrence-based techniques.

We also have used recurrence-based techniques to investigate the effect of high MHD activity in the electrostatic fluctuations in TCABR. The MHD activity increases at different instants of time during the discharge, and reaches high amplitudes with a narrow wave-number spectrum and a well-defined peak on the Mirnov frequency [20].

This chapter is organized as follows: In Sect. 18.2 we present the experimental setting and the kind of floating potential fluctuation data considered in our recurrence-based analyses. Section 18.3 presents results for recurrence-based diagnostics of potential fluctuations in different positions of the plasma column. Section 18.4 considers the effect of bias voltage on the recurrence-based diagnostics. The last section is devoted to our conclusions.

18.2 Electrostatic Potential Fluctuations

The experiments were performed in a hydrogen circular plasma in the Brazilian tokamak TCABR operating at the Institute of Physics of University of Sao Paulo (Fig. 18.1a), whose main parameters are listed in Table 18.1 [21]. Langmuir probes

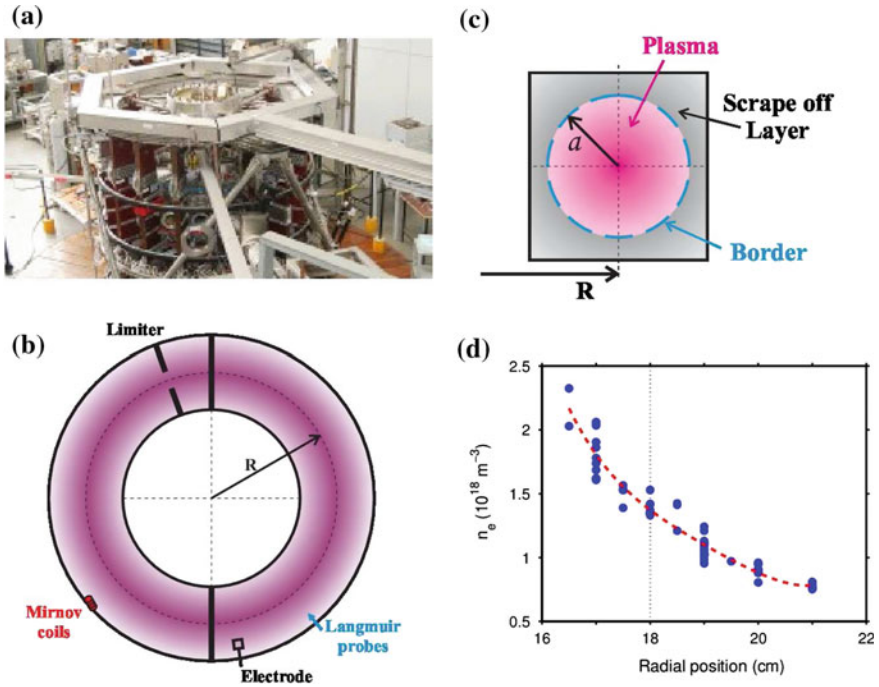

Fig. 18.1 **a** Picture of the TCABR tokamak. **b** Top view of the toroidal chamber and the Langmuir probes used to measure fluctuations of the floating electrostatic potential at the plasma edge. **c** Cross section of the toroidal chamber. **d** Plasma density radial profile

Table 18.1 Basic parameters of the TCABR tokamak

Major radius (R)	61 cm
Minor radius (a)	18 cm
Maximum plasma current ($(I_p)_{max}$)	100 kA
Plasma duration (τ_p)	100 m
Hydrogen filling pressure (p_0)	3×10^{-4} Pa
Toroidal magnetic field ($B_{\phi 0}$)	1.1 T

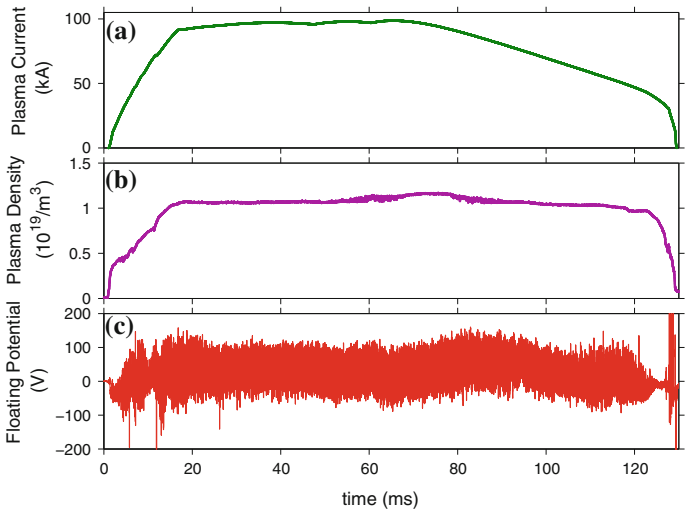

Fig. 18.2 Time evolution of the **a** plasma current, **b** plasma density, and **c** floating potential for a discharge of TCABR tokamak

were used to measure the floating potential whose fluctuations are mainly due to the the fluctuations of the plasma electrostatic potential (Fig. 18.1b) [22, 23]. The probes are mounted on a movable shaft that can be displaced radially with respect to the center of the plasma column, so as to cover both the plasma edge and the so-called scrape-off layer, the latter comprising part of the vacuum layer existent between the plasma column and the vessel wall (Fig. 18.1c, a and R stand for the minor and major radius of the vessel, respectively, see Table 18.1). The probe displacement, however, occurs only for separate discharges [4, 19]. The plasma density in the region surrounding the plasma edge is shown in Fig. 18.1d, showing a significative reduction as we go outwards to the scrape-off layer.

Figure 18.2 shows the time evolution of a typical tokamak plasma discharge in TCABR. The plasma current (Fig. 18.2a) grows rapidly in the first dozen of milliseconds and reaches a plateau where the current stays at a nearly constant level, decaying slowly during the second half until the end of the discharge. The plasma density evolution, indicated by Fig. 18.2b, exhibits a similar evolution. The time

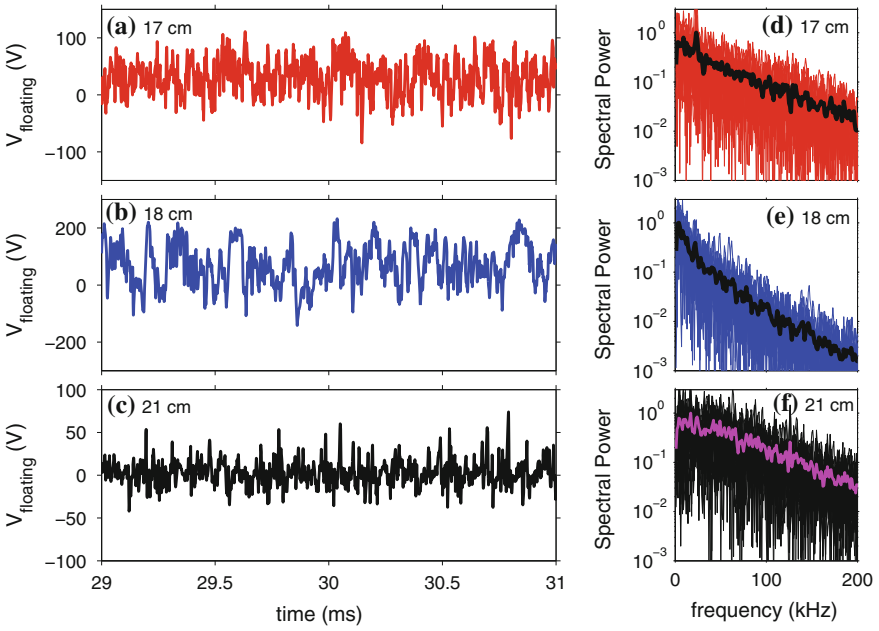

Fig. 18.3 Time evolution of the floating potential measured by Langmuir probes placed at radii **a** 17 cm, **b** 18 cm, and **c** 21 cm. *Light lines* in (**d**), **e** and **f** are the power spectral densities in each 1 m of the discharges corresponding to (**a**), (**b**) and (**c**), respectively. The *bold lines* in **d**–**f** are the average power spectra

signals we are particularly interested to study are for the floating electrostatic potential, a representative example being depicted by Fig. 18.2c, which shows highly irregular fluctuations.

The nature of the electrostatic potential fluctuation is dependent on the radial location where the probe is placed, as shown by Fig. 18.3, where two milliseconds of the time window in the plasma current plateau have been selected. Figure 18.3 also shows the corresponding power spectra in each 1 m (light lines) and the average power spectra in a 10 m window. Within the plasma column (Fig. 18.3a) the potential fluctuations present a −50 to +50 V range. As we move to the edge of the plasma column (Fig. 18.3b) the amplitude of such fluctuations increases by a factor of four, indicating that the fluctuation level augments as we approach the plasma radius. Outside the plasma column (but inside the scrape-off layer of vacuum magnetic field) (Fig. 18.3c) the floating potential amplitude range decreases to an in-between level. Hence the turbulent fluctuations become weaker as we move both outside and inside the plasma edge. This increase of amplitude at the plasma edge has been observed in other Tokamaks [24].

The radial dependence of the electrostatic fluctuations level at the vicinity of the plasma radius is a signature of the role played by radial density gradients in the generation of drift waves, which is the essential cause of turbulence in the plasma

edge [18]. In fact, the presence of steep density gradients in the plasma edge can give rise to fully developed wave turbulence, which is considered a likely candidate for explaining anomalous transport observed in experiments [25].

However, characterizing turbulence in order to quantify its level and radial dependence is a difficult task, what can be illustrated by Fig. 18.3, where we also show the power spectra of the potential fluctuations measured at three radial positions just analyzed [23, 26]. All of them are broadband, which is already expected from the chaotic behavior related to turbulence but, apart from some unessential rippling, those spectra do not show a distinguish feature which could be used to quantify the fluctuation level and specific different dynamical regimes [27, 28].

18.3 Recurrence-Based Analysis of Turbulent Fluctuations

It turns out that recurrence quantification analysis can help us to characterize the fluctuation level observed in tokamak experiments, especially because it does not impose stationarity nor long series length as necessary conditions, and can also work satisfactorily with moderate noise levels in data acquisition system. Basically the major source of noise is thought to be the due to wave turbulence. Moreover, it also yields reliable estimates of the Shannon entropy and can also indicate the amount of determinism in a given time series, what gives us an idea of the noise level added to the chaotic signal [29].

The recurrence plots for the first five hundred points (corresponding to half a millisecond interval) of the time series previously considered are depicted in Fig. 18.4. We choose the embedding dimension as four and the time delay was selected by considering the first local minimum of the autocorrelation function. We can

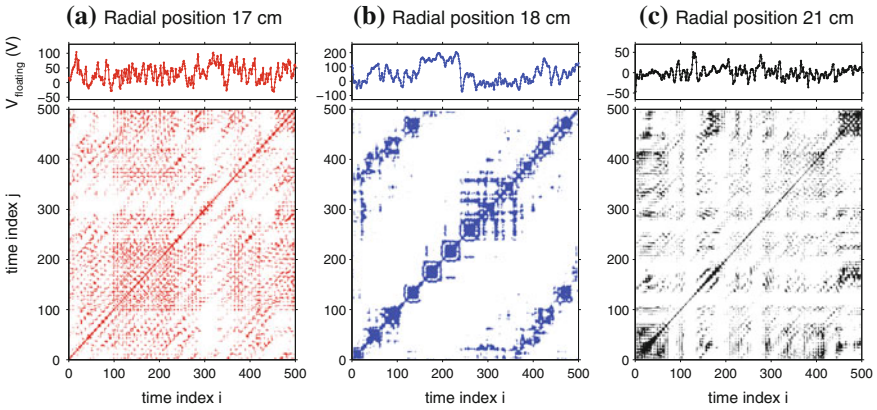

Fig. 18.4 Recurrence plots for the time series of the floating potential for radial positions **a** Radial position 17 cm; **b** Radial position 18 cm; and **c** Radial position 21 cm

recognize the changes in the turbulent behavior at the plasma radius by comparing the diagonal and vertical structures of the time series at the plasma edge (middle of Fig. 18.4) with the scattered nature of the recurrence plots depicted in left and right of Fig. 18.4, which suggest a pronounced stochastic effect, probably related to noise and/or other mechanisms not accounted for in a deterministic theory.

From the recurrence plots of the time series of the floating potential for different positions we have computed a number of recurrence-based diagnostics. One of them is the determinism (DET) of the signal. By sliding the time window used to compute this quantity in each radial position we obtain a statistical distribution of values of DET. In Fig. 18.5 we plot histograms which stand for numerical approximations of the probability distribution function of the values of DET for different radial positions, from 17 to 21 cm. This range was chosen not just because it is accessible for measurements but also because it is one relevant region for cross flow transport. We clearly see that the average value of DET increases as we approach the edge of the plasma column, but decrease afterwards, with a nearly constant dispersion.

In Fig. 18.6a we plot the radial profile of DET from these average values. The line joining those points is a polynomial fit just to guide the eye, i.e. it is not intended to give a radial profile DET (r) but rather a trend: the degree of determinism increases significantly as we approach the plasma radius, and decreases afterwards. By sliding the time window used to compute this quantity in each radial position we obtain a statistical distribution of values of DET in each 1 m (1000 points, without

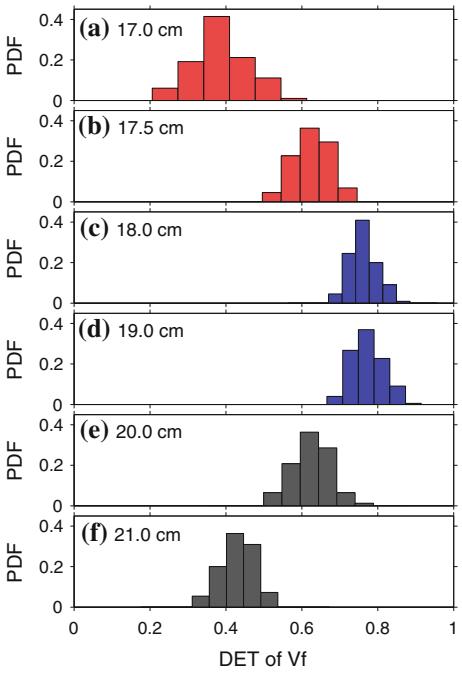

Fig. 18.5 Numerical approximations of the probability distribution function of the determinism for floating potential fluctuations measured at different radial positions

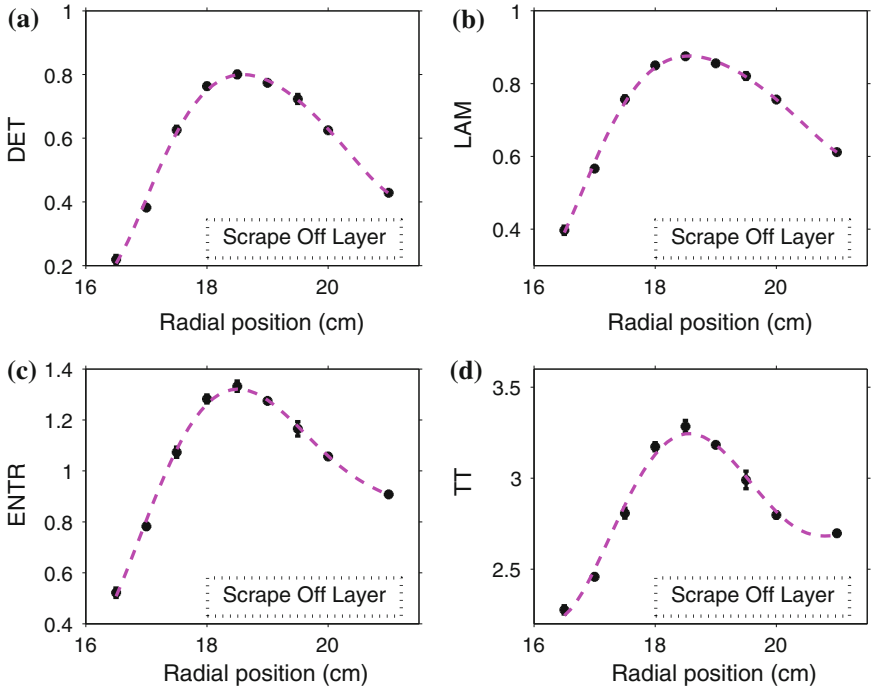

Fig. 18.6 Radial profiles for several recurrence-based diagnostics: **a** determinism; **b** laminarity; **c** entropy; **d** trapping time. The position of the scrape-off layer between the plasma column and the tokamak wall is indicated by a *horizontal bar*

superposition). This suggests a clear modification on the dynamical properties of the plasma floating potential fluctuations in the analyzed region. The same trend is exhibited by other recurrence-based diagnostics like the Laminarity LAM (Fig. 18.6b), entropy of the diagonal line lengths (Fig. 18.6c), and trapping time (Fig. 18.6d). These results reinforce our conclusion that the dynamical properties of the turbulent fluctuations changes substantially in this region, with the simultaneous peak of the DET, LAM and TT indicating that the recurrences are most concentrated around the plasma edge than elsewhere.

18.4 Effects of a Bias Radial Electric Field

The investigation we just reported was the first work about recurrence plots in fusion plasmas and one of the first on plasmas in general [4]. We have, since then, continued our pursuit of applying recurrence-based techniques to characterize a wealth of dynamical phenomena related with turbulent plasma fluctuations. In order to control plasma turbulence, external electric potentials have been applied, in several devices,

to change the electric field profile. This turbulence reduction decreases the particle loss at the plasma edge and, consequently, improves plasma confinement [14, 26]. In a recent paper [19] we applied recurrence quantification analysis to the electrostatic fluctuations at the edge of the plasma in TCABR tokamak with and without the application of a radial electric bias in order to control the turbulence and improve plasma confinement. The latter is a form of creating internal transport barriers in the plasma so as to decrease the radial transport. Hence the quality of plasma confinement is expected to improve through this procedure.

In order to keep the recurrence plots comparable among signals with very different standard deviations, we fixed the recurrence rate. In other words, we consider two points as being recurrent or not depending on the overall recurrence rate of the series be equal to some given value. Hence, if most points are so close that they would be considered recurrent if a fixed threshold be given, then with a fixed recurrence rate they will be recurrent only if their contribution falls into that percentage. We observed substantial changes in the turbulent behavior at the plasma edge by comparing the diagonal and vertical structures of Fig. 18.7a—obtained without biasing—with the recurrence plot with biasing depicted in Fig. 18.7b. Such comparison suggests a reduction in the deterministic effect, probably related to the existence of a lower number of large-scale structures due to biasing.

In Fig. 18.8 we show the radial profile of the determinism (from potential fluctuations time series) with and without electric bias. The recurrence rate was kept at 1, 2.5 and 10 % in Fig. 18.8a–c, respectively. The overall trends are the same for the three cases here considered, apart from presenting higher values of DET as the recurrence rate is larger, as expected from the higher density of recurrent points. The radial

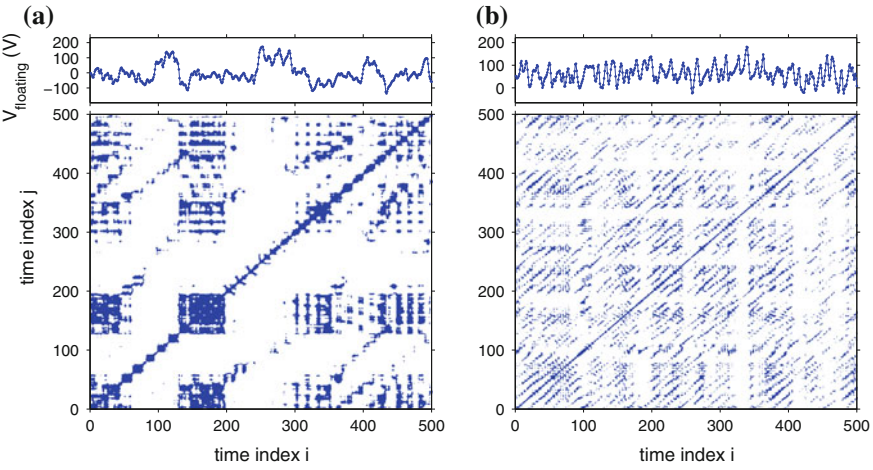

Fig. 18.7 Recurrence plots for the time series of the floating potential for radial positions 19 cm without bias (**a**) and with bias (**b**)

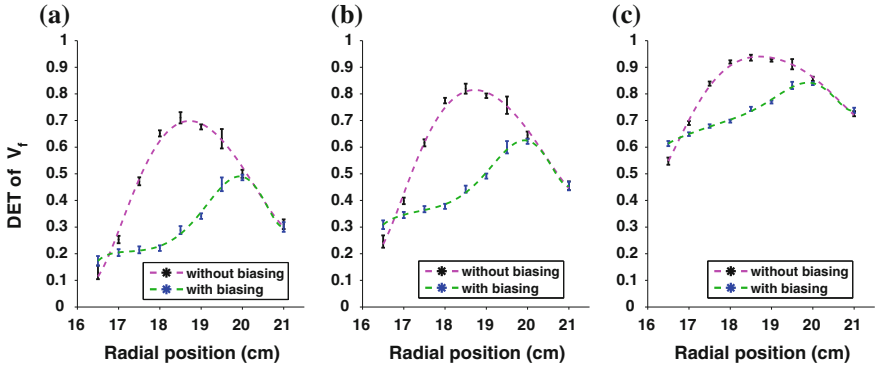

Fig. 18.8 Radial profile of determinism for fluctuations with (*red*) and without (*blue*) radial electric field bias. The recurrence rate was fixed on 1.0, 2.5 and 5.0 %. **a** RR = 1 %, **b** RR = 2.5 %, **c** RR = 10 %

profile of determinism shows an increase as we approach the plasma radius with a maximum value just after that location already, as we shown in the previous Section.

With biasing, however, the radial profile of determinism suffers a radical change, since it decreases the most where it is more intense, namely at the vicinity of plasma edge, an effect possibly caused by the breakup of vertical and horizontal large-scale structures in that region. This structure reduction decreases the particle transport, so improving plasma confinement. Moreover, this alteration in the recurrence properties occurs over a wide radial interval, indicating that a broad region is modified due to the biasing. The results are qualitatively similar as recurrence rate has been varied, what suggests that our results are not artifacts of a too small or too large threshold radius. Hence, as a general trend, the determinism decreases after biasing. We speculate that this is due to the destruction of highly recurrent regions within the plasma that enhance particle and heat transport.

Another experimental investigation of the effects of external biasing on turbulence was performed in the Texas Helimak, a plasma toroidal device located at the University of Texas at Austin [30]. It has a vacuum vessel with rectangular cross section of external radius 1.6 m, internal radius 0.6 m and height 2.0 m (Fig. 18.9a). In this machine the combination between the toroidal and a small vertical field creates a helical magnetic field with curvature and shear, with magnetic field lines with long (*circa* 40 m) connection lengths (in such a way that end effects can be neglected). One of the goals of this machine is to investigate plasma edge turbulence, with a diagnostic system of more than 700 Langmuir probes mounted at four sets of bias plates, where a bias electric field can be applied (Fig. 18.9b) [30]. For negative biasing, turbulence control has been investigated and states of greatly reduced turbulence have been achieved [17]. On the other hand, for positive biasing turbulence shows enhanced broadband spectran and non-gaussian probability distribution functions (PDF) with extreme events associated to bursts [17].

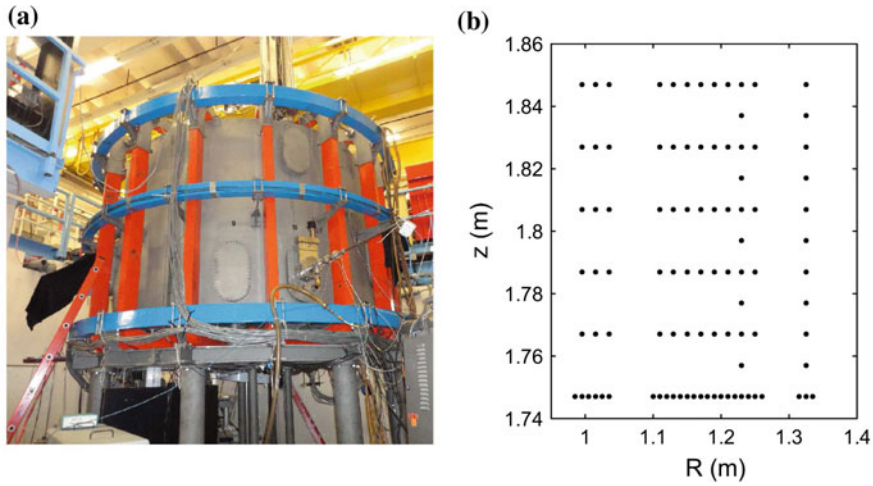

Fig. 18.9 **a** Picture of the Texas Helimak. **b** Langmuir probe distribution on the top plates

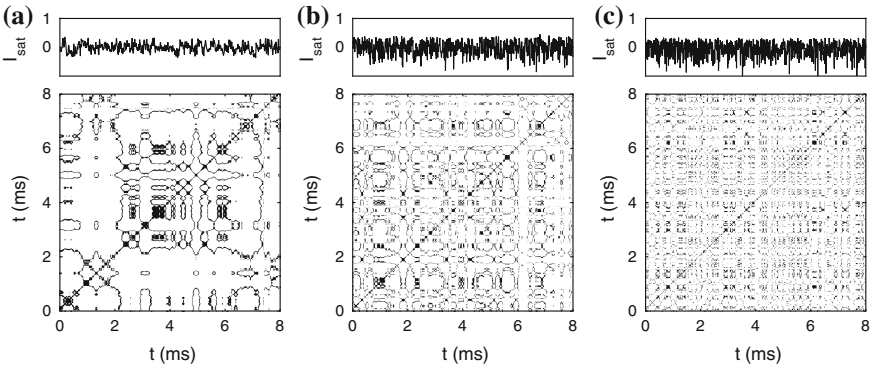

Fig. 18.10 Recurrence plots for plasma turbulence data from Texas Helimak discharges with voltage bias **a** +20 V, **b** zero, and **c** −25 V

In Fig. 18.10, we present the recurrence plots obtained from the measured fluctuation data for short time intervals (of 4 m, corresponding to 2.000 points) with and without bias [31]. We see that for negative bias the size of the recurrence structures in the RPs is bigger than for positive bias. Hence the regularity is lower at positive bias than at negative one, which can be interpreted as an increase of the turbulence. Moreover the determinism of the corresponding time series increases from 0.61 (Fig. 18.10a), for a bias +20–0.85 V (Fig. 18.10c), for a bias −25 V). Hence the value of determinism decreases as we go from negative to positive bias. These results confirm those previously obtained from TCABR discharges in Brazil [19].

18.5 Conclusions

The dynamical characterization of plasma turbulence is one of the outstanding theoretical and experimental problems in modern fusion plasma physics, specially taking into account the current large-scale projects for fusion tokamaks like ITER. It is thought that at least part of the turbulent phenomena are related to deterministic mechanisms, such as the existence of fluid wave turbulence [32]. On the other hand there must be a variety of other factors that influence turbulent data, mostly from a stochastic nature and hence only amenable to statistical analysis.

We have used recurrence plots as a major tool on analysing turbulent fluctuation data so as to give a precise indication about the dynamical properties of them, in order to understand how long we can advance using theoretical models to characterize plasma edge turbulence in tokamaks. In this article we review past work on this direction using data from the Brazilian tokamak TCABR [4, 19] and the Texas Helimak [31].

Our results show that the recurrence-based diagnostics like determinism, laminarity, entropy and trapping time increase as we approach the plasma edge. This suggests that the concentration of the recurrences is higher at the plasma edge than in its neighbourhood, where the stochastic content is comparatively higher. From the theoretical point of view this means that fluid models like, for example, drift-wave or interchange turbulence, are expected to explain better the features at the plasma edge than elsewhere.

Mier et al. [9] also made a recurrence quantification analysis of fluctuating potential data (of the TJ-II stellarator), with focus on the radial dependence of the determinism and other recurrence-based diagnostics, in a way similar to our analysis. Their results indicate that the deterministic content of the fluctuations increase as we approach the plasma edge, confirming our findings. Moreover, they show an interesting increase of the determinism with the magnetic shear, a feature which we did not consider up to now.

We have performed recurrence-based investigations on the application of a bias electric field in the radial direction. We show that biasing is responsible to the concentration of the recurrences in most of the radial positions considered. Indeed, in the Texas Helimak we found that the determinism actually increases as we go from a positive to a negative bias.

In retrospect we see a wealth of potential applications for recurrence-based diagnostics in plasma physics, since the physical setting involved in plasma phenomena are so complex that it is quite often the case in which the experimental data have both deterministic and stochastic contents.

Acknowledgments The brazilian authors would like to thank CNPq, CAPES and FAPESP for partial financial support.

References

1. J.P. Zbilut, C.L. Webber Jr., Phys. Lett. A **171**(3–4), 199 (1992). doi:10.1016/0375.9601(92)90426-M
2. C.L. Webber Jr., J.P. Zbilut, J. Appl. Physiol. **76**(2), 965 (1994)
3. N. Marwan, M.C. Romano, M. Thiel, J. Kurths, Phys. Rep. **438**(5–6), 237 (2007). doi:10.106/j.physrep.2006.11.001
4. Z.O. Guimarães-Filho, I.L. Caldas, R.L. Viana, J. Kurths, I.C. Nascimento, Y.T. Kuznetsov, Phys. Lett. A **372**, 1088 (2008). doi:10.1016/j.physleta.2007.07.088
5. P. Alex, S. Arumugam, K. Jayaprakash, K.S. Suraj, Results Phys. **5**, 235 (2015)
6. V. Mitra, A. Sarma, M.S. Janaki, A.N.S. Iyenger, B. Sarma, N. Marwan, J. Kurths, P.K. Shaw, S. Saha, S. Ghosh, Chaos. Solit. Fract. **69**, 285 (2014)
7. C. Stan, C.P. Cristescu, D.G. Dimitriu, Phys. Plasmas **17**, 042115 (2010)
8. R.O. Dendy, S.C. Chapman, Plas. Phys. Contr. Fusion **48**, B313 (2006)
9. J.A. Mier, R. Sánchez, M.A. Pedrosa, C. Hidalgo, L. Garcia, J. Varela, D.E. Newman, *Proceedings of the 37th EPS Conference on Plasma Physics*, P1.0168p (2010)
10. J.A. Mier, R. Sánchez, L. Garcia, J. Varela, D.E. Newman, Phys. Plasmas **18**, 062306 (2011)
11. M. Rajkovic, T.H. Watanabe, M. Skoric, Phys. Plasmas **16**, 092306 (2009)
12. T. Zivkovic, K. Rypdal, Phys. Rev. E **77**, 037401 (2008)
13. https://www.iter.org
14. C.W. Horton, S. Benkadda, *ITER Physics* (World Scientific, Singapore, 2015)
15. R.D. Hazeltine, J.D. Meiss, *Plasma Confinement* (Addison Wesley, New York, 1992)
16. C.W. Horton, *Turbulent Transport in Magnetized Plasmas* (World Scientific, Singapore, 2012)
17. C.W. Horton, Rev. Mod. Phys. **71**, 735 (1999). doi:10.1103/RevModPhys.71.735
18. I.L. Caldas, H. Tasso, Plasma Phys. **20**, 1299 (1978)
19. Z.O. Guimarães-Filho, I.L. Caldas, R.L. Viana, I.C. Nascimento, Y.K. Kuznetsov, J. Kurths, Phys. Plasmas **17**, 012303 (2010). doi:10.1063/1.32800100
20. Z.O. Guimarães-Filho, G.Z.S. Lima, I.L. Caldas, R.L. Viana, I.C. Nascimento, Y.K. Kuznetsov, J. Phys.: Conf. Ser. **246**, 012014 (2010). doi:10.1088/1742-6596/246/1/012014
21. I.C. Nascimento, I.L. Caldas, R.M.O. Galvão, J. Fusion Energy **12**, 529 (1993)
22. R.M.O. Galvão et al., Plasma Phys. Control. Fusion **43**, 1181 (2001)
23. A.A. Ferreira, M.V.A.P. Heller, I.L. Caldas, E.A. Lerche, L.F. Ruchko, L.A. Baccalá, Plasma Phys. Control. Fusion **46**, 669 (2004)
24. ChP Ritz et al., Phys. Rev. Lett. **62**, 1844 (1989)
25. A.A. Ferreira, M.V.A.P. Heller, I.L. Caldas, Phys. Plasmas **7**, 3567 (2000)
26. I.C. Nascimento, Y.K. Kuznetsov, J.H.F. Severo, A.M.M. Fonseca, A. Elfimov, V. Bellintani, M. Machida, M.V.A.P. Heller, R.M.O. Galvão, E.K. Sanada, J.I. Elizondo, Nucl. Fusion **45**, 796 (2005)
27. C. Hidalgo et al., Nucl. Fusion **45**, S266 (2005)
28. G. Van Oost et al., Plasma Phys. Control. Fusion **45**, 621 (2003)
29. M. Thiel, M.C. Romano, J. Kurths, R. Meucci, E. Allaria, F.T. Arecchi, Physica D **171**, 138 (2002)
30. K.W. Gentle, H. He, Plasma Sci. Technol. **10**, 284 (2008)
31. D.L. Toufen, Z.O. Guimarães-Filho, I.L. Caldas, J.D. Szezech, S.R. Lopes, R.L. Viana, K.W. Gentle, Phys. Plasmas **20**, 022310 (2013). doi:10.1063/1.4793732
32. A.M. Batista, I.L. Caldas, S.R. Lopes, R.L. Viana, C.W. Horton, P.J. Morrison, Phys. Plasmas **13**, 042510 (2006). doi:10.1063/1.2184291

Chapter 19
Recurrence Quantification Analysis as an Approach for Ultrasonic Testing of Porous Carbon Fibre Reinforced Polymers

Carsten Brandt

Abstract A first investigation of Recurrence Quantification Analysis (RQA) for the assessment of porosity in ultrasonic testing of Carbon Fibre Reinforced Polymers (CFRP) is presented. The standard method for detecting porosity with Non-Destructive Testing (NDT) using ultrasonic pulse-echo inspection is the evaluation of the back-wall echo (BWE) from the side opposite to the ultrasonic probe. The work presented aims at determining a BWE-equivalent out of the ultrasonic intermediate echoes from the inner of the part if a BWE cannot be evaluated, as for e.g. CFRP sandwich structures. Ultrasonic measurements on three CFRP samples with artificial porosity were performed. A delay embedding to reconstruct the state space for the intermediate echo time series and a subsequent creation of Recurrence Plots are carried out. The features Recurrence Rate *RR* and determinism *DET* are calculated with Euclidean and angular distance as metric. RQA parameters are largely varied and the results are evaluated on best correlations of *RR* and *DET*, respectively, with the BWE. The feature *DET* presents an appropriate BWE-equivalent. When using Euclidean distance, higher values of determinism are obtained for higher porosity, based on reduced amplitude due to additional reflections not going back to the ultrasonic transducer. This effect can be obtained with simpler evaluation (Quartile Coefficient of Dispersion), too, with a correlation similar to RQA. For angular distance, determinism decreases with higher porosity based on the randomness introduced by the pores. This first step of ongoing research will be followed by investigations on CFRP samples of other material and with "natural" porosity.

C. Brandt (✉)
Centre for Industrial Mathematics, Bremen, Germany
e-mail: carsten.brandt@uni-bremen.de

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8_19

19.1 Motivation

Carbon Fibre Reinforced Polymers (CFRP) play an increasing role in aeronautic industry due to their high ratio of mechanical strength to weight. Their potential for saving of weight and, thus, fuel burn and costs has led to an increasing usage in aircrafts, for example from 25 % in the Airbus A380 to 53 % in the Airbus A350XWB [1]. 100 % Non-Destructive Testing (NDT) of produced CFRP laminates is standard for material used for example in structural components of Airbus aircrafts, generally performed using ultrasonic testing in pulse-echo mode [2]. Defects like delaminations (a material separation) or porosity have to be detected [3]. Porosity is evaluated via the height of the so-called back-wall echo (BWE), the reflection of the ultrasonic wave at the opposite side of the part. However, there are special situations as complex geometries, bonded components or sandwich structures (consisting of a honeycomb core between two CFRP laminates), in which this analysis is not possible, cf. [2, 4] and Sect. 19.2. Parts have to be designed as if porosity were present. It is therefore beneficial to create an evaluation method without needing the back-wall echo, solely using the echoes out of the inside of the part to be inspected (intermediate echoes) to create a "back-wall echo equivalent" (BWE-equivalent). The aim of this chapter is to present first work on exploring the use of Recurrence Quantification Analysis (RQA) to provide a BWE-equivalent.

The use of state space methods and RQA has been reported for the evaluation of ultrasonic waves with the use of guided as well as diffuse ultrasonic waves for Structural Health Monitoring with permanently installed sensors [5–7]. Cacciola et al. [8] report about usage of Recurrence Plots for the NDT method eddy current inspection on steel samples. An interesting work is presented in [9]. Signals from ultrasonic pulse-echo testing of cement pastes are evaluated using state space based prediction features and RQA to assess the porosity content. The authors assume that porosity has a stochastic influence and thus decreases the determinism in the system.

For this chapter, two hypotheses are followed for the data analysis:

(1) Pores will lead to a perturbation of the otherwise almost periodic (decaying) signal from the layered structure of the material.
(2) Reflections at the pores will also lead to changes of signal amplitudes, hence comparing the time-localised energy of the signal between porous and non-porous structures will reveal a change of energy.

See Sect. 19.3.3 for further discussion.

Up to the author's knowledge, RQA has neither been used before for the ultrasonic pulse-echo inspection of CFRP nor for the evaluation of porosity in CFRP. Other approaches to generate information about porosity out of ultrasonic intermediate echoes are presented in [4, 10], based on linear approaches.

This chapter contains an initial assessment of the ability of the two basic RQA features Recurrence Rate and determinism to provide a back-wall echo equivalent for the evaluation of porosity, and presents a first step of ongoing research. Euclidean and as recently proposed angular distance are evaluated as metric for the

Recurrence Plots. Results are processed from measurements of three CFRP samples containing artificial porosity.

The remainder of this chapter is organised as follows:

- Section 19.2 contains some background to CFRP and ultrasonic pulse-echo inspection,
- the theory of state space reconstruction and Recurrence Quantification Analysis and the underlying hypotheses for the presented application are outlined in Sect. 19.3,
- Section 19.4 describes the measurements and
- the outcome of evaluations is shown in Sect. 19.5.

Some theoretical considerations are compiled in the appendix after conclusion and outlook in Sect. 19.6.

19.2 CFRP and Ultrasonic Inspection

CFRP is a composite material consisting of a resin matrix with reinforcing carbon fibres. One method of production is the build-up of laminates by placing of resin pre-impregnated layers of carbon fibres ('prepregs') in a mould on top of each other. The prepregs consist either of unidirectional material (tape) or of fabric material containing fibres in several directions. They are then cured under pressure and temperature in an autoclave to form a laminate. The "[…] cure parameters […] affect laminate porosity", being "[…] trapped pockets [or bubbles] of air, gas, or vaccum […] [11]". For reference to CFRP and its production, see also [12].

Different CFRP-parts can be joined using the co-bonding procedure, where prepregs are placed on a cured ('hard') part to form a second ('wet') part, and the joining of the hard and wet partner is done in one autoclave cycle together with the curing of the second part. For a more detailed explanation see [11], and refer to the same reference for information about CFRP sandwich structures mentioned in Sect. 19.1.

Ultrasonic testing in pulse-echo mode involves one transducer (also called probe) sending an ultrasonic pulse wave into the part under inspection and receiving the response of the part, the reflected echoes. The transducer can be directly coupled to the part's surface with water as a coupling medium (contact technique) or both transducer and part are in a water tank (immersion technique). At boundaries between materials that differ in the product of density and ultrasound velocity, the acoustic impedance [13], part of the wave is transmitted and part of the wave is reflected [14]. The highest reflections occur at the interface between water and surface (surface echo), at the back-side of the part (back-wall echo) and, if present, at material imperfections like delaminations. Furthermore, in the case of CFRP inspection, many internal reflections occur due to the layered laminate build-up and due to the inhomogeneity of the (composite) material.

The received echoes are converted to voltage in the transducer due to the direct piezoelectric effect [15]. These can be shown against time as a so-called amplitude

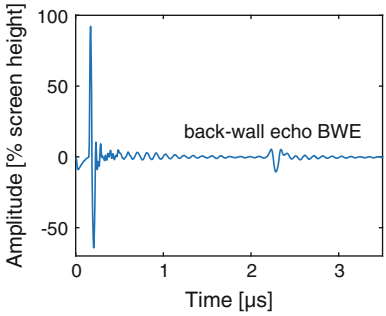

Fig. 19.1 Example of A-scan out of measurements of sample CG070M15 (cf. Sect. 19.4)

scan or A-scan [15] on a screen, similar to an oscilloscope. Figure 19.1 shows one example of an A-scan out of measurements performed for the present work.

A calibration has to be performed to determine the correct gain (amplification) of the ultrasonic device before carrying out an inspection; a common approach is to have a constant BWE of 80 % screen height. For the present investigation, the gain was chosen such as to have the complete surface echo recorded with signal height less than 100 % (on the very left of the A-scan in Fig. 19.1), thus leading to a back-wall echo much lower than usual for serial inspection. The large difference between BWE and surface echo is due to the many internal reflections in the layered CFRP material, which cause a higher attenuation of the signal after a greater time of flight (i.e. in greater depth).

Nowadays, ultrasonic inspections are often carried out using phased array technology [3], as utilised for the present investigations, too. Phased array ultrasonic probes consist of several (e.g. 64 or 128) elements, which can be individually electronically controlled. This enables for example an electronic scanning, where a certain number of elements (the aperture), e.g. element 1–8, is fired simultaneously, then element 2–9, then element 3–10 and so on.

In automatic inspection, a part is scanned and a lot of A-scan data is generated. A common way to display such results is the so-called C-scan [14]. Here, one characteristic variable per A-scan is shown, such as the height or depth (where the time scale is converted into depth by knowing the ultrasonic velocity) of a defined echo, with different shades of grey or in different colors. The defined echo can be for example the back-wall echo.

Evaluation of industrial ultrasonic inspection uses often the appearance, height and depth of direct echoes [16], for example occurring from large defects like delaminations. For porosity a different way has to be gone. The pores do not cause significant direct echoes, which could be evaluated. They reflect only very small echoes back to the ultrasonic probe or they reflect the ultrasound in other directions than that of the probe, due to the geometrical appearance of the pores. That is why the back-wall echo is evaluated in addition to direct echoes. The ultrasonic energy of the reflections from the pores, however not significantly visible in an A-scan, is lost in the portion of the ultrasonic wave travelling to the opposite side of the part,

the back-wall. The back-wall echo decreases. At Airbus, a reduction of the BWE of 6 dB, 12 dB or 18 dB, respectively, is considered critical [2].

As mentioned in Sect. 19.1, this standard way to evaluate CFRP on porosity when inspected with ultrasonic testing via the back-wall echo fails in several special situations:

- For complex geometries like wedge or curved parts, the opposite surface of the part deviates significantly from being parallel to the upper side. Such deviations, starting from a few to several degrees, cause the ultrasound to be not any more directly reflected to the ultrasonic transducer. The back-wall echoes decrease. Especially for spherical structures, for which the angle between the surfaces varies from point to point on the surface, it is very difficult to make out whether changes in the BWE originate from this geometrical effect or from porosity.
- For CFRP co-bonded parts, it may be beneficial to make a statement solely for one bonding partner. This is difficult because the internal interface between the partners after bonding may not deliver a stable echo comparable to the BWE.
- For sandwich parts, it is not yet possible to make a statement about porosity in the CFRP skin, because the interface between the skin and the core does not deliver a stable echo, similar to the situation with bonded parts above.

Still, intermediate echoes can be received in all these situations. The information from these echoes shall be used to generate a back-wall echo equivalent in the current chapter with the use of Recurrence Quantification Analysis.

19.3 Background and Basics of Recurrence Quantification Analysis

The basis for Recurrence Quantification Analysis (RQA) is a presentation of a dynamical system in a state space (also: phase space). Dynamical systems [17, 18] are technical or economical or biological systems which evolve (deterministically) in time. They are modelled via a set of d first order differential equations and represented by d variables. These variables are called state or phase variables; their evolvement over time can be represented in a d-dimensional space, the state or phase space. A very simple example of a dynamical system is a frictionless (undamped), not driven pendulum in earth's gravity field [18]. Its state is completely determined by the pendulum's angle and angle velocity. The state space is two-dimensional in this example. The pace the variables follow over time is called orbit or trajectory [19].

19.3.1 State Space Reconstruction

With the rise of chaos theory (starting in the 1970's and 1980's of the last century and being concerned with systems apparently stochastic but indeed deterministic)

powerful methods for the analysis of time series evolved [20]. These include state space reconstruction [21–23]. In a normal experimental setting, it is not possible or not economical to measure all state variables. Only a limited number or often, as in the present investigations, merely one variable is recorded. Now, the reconstruction method of delay embedding enables experimenters and scientists to reconstruct, from the few or one measured variable(s), trajectories in a reconstructed state space that are one-to-one to the ones in the original state space. One-to-one means that distinct points of the original trajectories map to distinct points of the reconstructed trajectories [24]. Takens [21] includes a mathematical proof for this, the reconstruction theorem (cf. also [18, 25]).

As mathematically sophisticated as Takens' proof is, as relatively simple is the employment of delay embedding. In the following, we treat discrete systems: These are easier to handle and are the general case in practise, since the measured signal is digitised (as is the case with the ultrasonic signal). Here, the delay embedding is presented with summation symbol as in [26], and d unit vectors $\mathbf{e}_l$ or $\mathbf{e}_{k+1}$, respectively, assumed, representing the axes of the reconstructed state space of dimension d:

$$\mathbf{x}_i = \sum_{k=0}^{d-1} x_{i+\tau k} \mathbf{e}_{k+1}. \tag{19.1}$$

One point in state space, i.e. one point in time, is reconstructed by using the value of the time series from this point in time and $d-1$ points in the future, each one shifted by a time delay τ for creating the other state variables.

Figure 19.2 visualises the method of delay embedding, where ξ_l is understood as the coordinates of one point in time in direction of the unit vectors $\mathbf{e}_l$. For further details, refer to [2, 18].

It is interesting to ask for which cases this delay embedding truly delivers a reconstruction which is one-to-one to the trajectories of original state space.

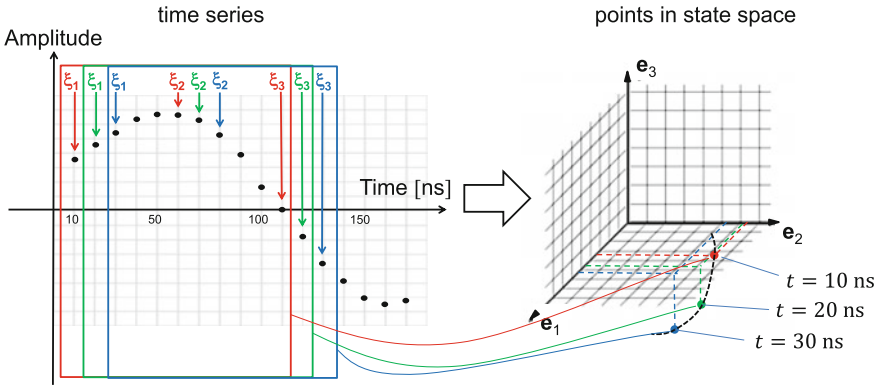

Fig. 19.2 Sketch for visualising the method of delay embedding [2]

Mathematically, the function(s) describing the dynamical system as well as the measurement function (mapping the trajectory in original state space on the real line, i.e., creating one real number from one point in state space) have to be generic. For further details refer to [18, 27, 28]. According to [29], "A mathematically generic observation, by definition, monitors all degrees of freedom of the system." For the present case, it is assumed that the ultrasonic measurement at the surface of the part gives sufficient information about the influence of porosity on the wave propagation inside the part [2]. Refer to Sect. 19.3.3 for the effects expected.

First applied on chaotic systems, state space reconstruction and subsequent investigation of the reconstructed dynamical systems have become a versatile tool for evaluation of signals originating from nonlinear systems, known as Nonlinear Time Series Analysis [30]. A part of this is based on the recurrence of dynamical systems.

19.3.2 Recurrence Plots and Recurrence Quantification Analysis

The formal concept of recurrences of dynamical systems was introduced by Henri Poincaré [26]. In his paper about celestial three body movement [31] he introduced what is now known as the Poincaré Recurrence Theorem [32, 33].

In 1987, Recurrence Plots (RPs) for the evaluation of dynamical systems were proposed [34]. The mathematical representation of a Recurrence Plot is the recurrence matrix [26]. In the recurrence matrix, each state of the system $\mathbf{x}_i$ at each instant in time $t = i\Delta t$ (Δt is the time between two subsequent measurements and corresponds to the reciprocal of the sampling rate) is compared with all states $\mathbf{x}_j$. Each entry i, j is either 1 in case that $\mathbf{x}_i$ and $\mathbf{x}_j$ are recurrent or 0 if they are not [26]. A Recurrence Plot is a two-dimensional plot in which recurrence points are plotted with a dot (often a black dot whereas the non-recurrent points are white [34, 35]).

Recurrence is defined as two states of the system being similar, which is again defined by having a distance lower than a threshold ε. The measure or metric for the distance as well as the threshold ε have to be defined; one of the common metrics used for Recurrence Plots is the Euclidean distance [35]

$$D_{Euc,ij} = ||\mathbf{x}_i - \mathbf{x}_j||; i, j = 1, \ldots, \mathrm{N}, \tag{19.2}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are points in time and N is the number of points in state space. $||\cdot||$ represents here the Euclidean metric (induced by the Euclidean norm, [36])

$$||\mathbf{x}_i - \mathbf{x}_j|| = \sqrt{\sum_{m=1}^{d} (\xi_{m,i} - \xi_{m,j})^2}. \tag{19.3}$$

The index m represents the coordinate of the d-dimensional state space. Alternatively, the angular distance is utilised:

$$D_{ang,ij} = \arccos\left(\frac{\mathbf{x}_i \mathbf{x}_j}{||\mathbf{x}_i|| ||\mathbf{x}_j||}\right), \tag{19.4}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are now understood as vectors pointing from the origin to the points $\mathbf{x}_i$, $\mathbf{x}_j$. Equation (19.4) is nothing else but the angle between these two vectors. Usage of this metric for Recurrence Plots has been proposed in [37] because of its independence of scaling effects—(19.4) delivers the same result no matter what the lengths of the vectors are. This is considered as great asset for NDT to be less dependent on the adjustment of gain in the course of calibration. Because this often involves a human factor, it is advisable to achieve an independence or reduced dependence on this calibration. On the other hand, the amplitude may also contain information about the porosity, cf. Sect. 19.3.3. Hence, both Euclidean and angular distance are investigated in the present work.

A further method not yet investigated for the current application is the combination of order patterns with Recurrence Plots [38].

Using the distances of two points in state space according to (19.2) or (19.4), respectively, a distance matrix $\mathbf{D}$ can be derived, consisting of the distance from each to every point in state space:

$$\mathbf{D} = D_{ij}. \tag{19.5}$$

With recurrence threshold ε, and utilising the Heaviside function to derive a "1" for recurrence points and a "0" for non-recurrent points, the recurrence matrix $\mathbf{R}$ is created from the distance matrix:

$$\mathbf{R} = R_{ij} = \theta\left(\varepsilon - D_{ij}\right), i,j = 1, \ldots, \mathrm{N}, \tag{19.6}$$

ε is taken constant over the whole recurrence matrix. Recurrence matrices can also be calculated with ε varied in a way such that the neighbourhood (the ball of radius ε) around a point always contains the same number of points [34, 35]. Thus, every row and every column in the recurrence matrix consists of the same amount of ones and the Recurrence Rate (see below) has a fixed value.

Since their first occurrence in 1987, a whole field of time series analysis involving Recurrence Plots has been developed, now known as Recurrence Quantification Analysis (RQA) [35]. In the present chapter, two basic out of several common RQA features are calculated [35] to check their suitability as back-wall echo equivalent: Recurrence Rate *RR* and determinism *DET*. The Recurrence Rate is the number of recurrence points divided by the number of all points,

$$RR = \frac{1}{N^2} \sum_{i,j=1}^{N} R_{ij}, \tag{19.7}$$

and can have values between 0 and 1. Equation (19.7) includes the main diagonal line (line of identity) (which trivially contains only ones since a state is always recurrent with itself) in the definition of *RR* as for example in [26].

The determinism *DET* is the number of all recurrence points that occur in diagonal lines of a length equal to or greater than a number l_{min} (minimum line length), divided by the number of all recurrence points. For its mathematical representation, we consider first the number of all lines containing exactly l recurrence points

$$H_D(l) = \sum_{i,j=1}^{N} (1 - R_{i-1,j-1})(1 - R_{i+1,j+1}) \prod_{k=0}^{l-1} R_{i+k,j+k}, \tag{19.8}$$

which leads then to determinism *DET*

$$DET = \frac{\sum_{l=l_{min}}^{N} l H_D(l)}{N^2 RR}. \tag{19.9}$$

l_{min} has to be at least 2; for $l_{min} = 1$ the determinism equals 1.

Diagonal lines in Recurrence Plots represent states where (at least when considering Euclidean distance) a section of the trajectory of the dynamical system runs nearly parallel to another section of the trajectory, or more exactly, in a region with the shape of a tube (of radius ε) around the other section. This is an indication of deterministic (or predictable) behaviour of the underlying dynamical system [26]. A system with a pure periodic motion of one frequency has a Recurrence Plot with only diagonal lines [26], as for example a frictionless, non-driven pendulum in earth's gravity field.

19.3.3 RQA for the Detection of Porosity in CFRP

The author's original motivation to use Recurrence Quantification Analysis for the present application is the complexity of the elastic wave propagation in CFRP with nonlinear multiple scattering [39] in the presence of pores [2]. The hypotheses now followed in this chapter are

(1) Pores will lead to a perturbation of the otherwise almost periodic signal (which has decaying amplitude due to the attenuation of the CFRP, Fig. 19.4). The intermediate echoes obtained from the investigated samples are almost periodic due to the layered structure of the material. Porosity leads to a perturbation of these signals, bringing higher randomness into the system by

randomly distributed scatterers as assumed in [9], thus decreasing determinism.

(2) Reflections at the pores will also lead to changes of signal amplitudes, hence comparing the time-localised energy of the signal between porous and non-porous structures will reveal a change of energy. This may be a reduction due to reflections not going back to the ultrasonic transducer or an increase due to additional reflections measured by the transducer.

The presented work shall be the first step gaining knowledge towards or against these hypotheses for porous CFRP.

19.4 Experiments

The measurements were performed on a set of three CFRP samples out of fabric material 913C-926-35%F with artificially introduced porosity fields (using the blowing agent Freon) of different amounts [40, 41]. These samples have a thickness of approximately 7 mm and a size of 120 mm times 100 mm. Their designations are CG070M15, the part with the highest porosity, on which most of the evaluations are performed, and CG070M13/14 with lower contents of porosity. All samples are plane-parallel and have thus a proper back-wall echo. The intermediate echo time series is used for the attempt to generate a back-wall echo equivalent; the intermediate echoes are no different whether or whether not the back-wall behind represents a proper reflector. This back-wall echo equivalent is then directly compared with the back-wall echo of one and the same A-scan.

The measurement equipment consists of commercially available ultrasonic electronics "Multi2000 32x128" by M2M including a laptop for controlling and recording and an Olympus "5L64" 5 MHz phased array probe (Fig. 19.3).

The probe consists of 64 elements, covering an area of 46 mm times 10 mm. It was driven with 5 MHz frequency and a voltage of 40 V with an aperture of eight

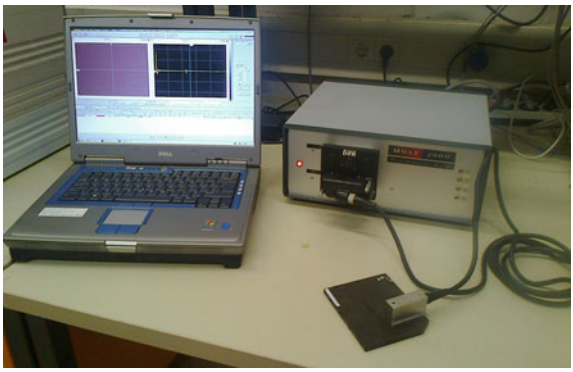

Fig. 19.3 Ultrasonic measurement equipment: electronics (*top right*), one sample and phased array probe (*bottom right*) and controlling laptop (*left*)

elements firing simultaneously. For all measurements, a constant gain of 18.9 dB was used.

The samples were inspected manually in contact technique. Two scans per sample were recorded. One scan covers approximately half of the sample.

No filters were applied and the received signals were sampled with a frequency of 100 MHz, thus every 10 ns an amplitude value was recorded.

The measurement data obtained with the M2M equipment was read into the evaluation software "Ultis" [42] and stored in data format ".nka". This data format was then read into Matlab, with which all following evaluations are carried out, using self-programmed functions.

19.5 Results and Optimal Parameters

The choice of optimal parameters for the Recurrence Quantification Analysis presented in this work is very much driven by the RQA itself (a similar approach to the one proposed in [43]). ε is varied between its extremes for both Euclidean and angular distance (Table 19.1).

To have a starting point for time delay τ, a quick check using the autocorrelation [44] is performed, which leads to values between 5 and 6 (time steps) for the different time series; quite similar to the "good first guess" [30] to take a quarter of the dominating wave length (if present). For the embedding dimension d, the method of false nearest neighbours [30, 35] is merely checked on one representative time series using the CRP-toolbox for MATLAB [26, 45]. This check delivered a value of 4. Based on these values, τ and d are both varied in a range from 1 to 10.

The minimum line length l_{min} for the RQA feature determinism is initially set to 2 for Euclidean distance. For angular distance, $l_{min} = 2, 4, 8$ is applied.

For 29 times 65 A-scans, thus 1885 time series, of one volume scan of sample CG070M15, the RQA features *RR* and *DET* are generated for all parameter values shown in Table 19.1. The linear correlation coefficient r (or Pearson's r, [44]) is determined between *RR* and BWE (r_{RR}) and between *DET* and BWE (r_{DET}).

Table 19.1 Applied range of parameters for Recurrence Quantification Analysis

RQA parameter	Min	Max	Steps/values used
Time delay τ	1	10	1
Embedding dimension d	1	10	1
Recurrence threshold ε for Euclidean distance	0.1	10	0.1, 0.2, 0.3, 0.4, 0.5, 0.8, 1, 1.2, 1.5, 2, 4, 6, 8, 10
Recurrence threshold ε for angular distance	0.01π	0.75 π	0.01π, 0.025π, 0.05π, 0.075π, 0.1π, 0.25π, 0.5π, 0.75π
Minimum line length l_{min}	2	8	Euclidean distance: $2 \leq 15$ Angular distance: 2, 4, 8

19.5.1 Euclidean Distance

The time series in these investigations have a dominating frequency of approximately 5 MHz (Fig. 19.4). Theoretical considerations for Euclidean distance based on this fact lead to an assumption of optimum Recurrence Rate of approximately 0.055 (see Appendix).

Extrema of correlation coefficients for all variations of RQA parameters and in the optimum Recurrence Rate range are determined (Table 19.2).

Three sets of parameters (row 3, 5 and 6) lead to Recurrence Rates extremely low or high, respectively. The minimum r_{RR} occurs for $RR = 0.004$. The according Recurrence Plots merely consist of the line of identity for non-porous areas. In the case of time series measured from an area with porosity, this line simply gets thicker on some sections. It is obvious that no serious RQA is possible with such spurious results.

The maximum correlation coefficients go along with a Recurrence Rate of almost 1 and with a range of Recurrence Rate R_{RR} (difference between minimum

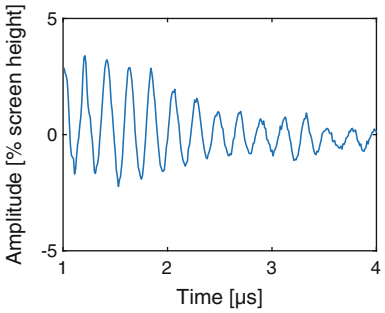

Fig. 19.4 A-scan of Fig. 19.1, intermediate echoes from 1 to 4 µs magnified: the time series of 301 points used for RQA

Table 19.2 Optimum coefficients of linear correlations between BWE-equivalent and BWE in dependence of RQA parameters for $l_{min} = 2$ for Euclidean distance

Row nr		Correlation coefficient r	Mean RR/R_{RR}	Mean DET/R_{DET}	ε	d	τ
1	RR driven	$r_{RR} = -0.60$	0.055/0.218	0.931/0.101	1	6	3
2	RR driven	$r_{DET} = -0.75$	0.055/0.147	0.729/0.263	0.3	2	1
3	min r_{RR}	$r_{RR} = -0.65$	0.004/0.005	0.975/0.267	0.4	10	6
4	max r_{RR}	$r_{RR} = 0.84$	0.996/0.048	0.9999/0.0023	10	4	1
5	min r_{DET}	$r_{DET} = -0.81$	0.142/0.277	0.904/0.112	0.8	3	1
6	max r_{DET}	$r_{DET} = 0.78$	0.995/0.073	0.9998/0.0019	6	2	5

and maximum) of only 0.2 % (the linear correlation coefficient makes no statement about the magnitude of the slope of the regression line [44]). The Recurrence Plots are almost black, recurrence and determinism decrease by a very small amount in the case of porosity. These results are not further considered either.

The minimum correlation coefficient for *DET*, reaching a large absolute value of −0.81, goes along with a reasonable Recurrence Rate of 0.142. The range of *DET* is still rather low (0.112). The Recurrence Plots for the situation without porosity (Fig. 19.5 left) as well as with porosity (Fig. 19.5 right) show lines perpendicular to the main diagonal, which may be caused by non-optimal embedding [46].

This choice of embedding parameters does not prove to be very robust: a reduction of the embedding dimension from 3 to 2 causes the correlation coefficient to change from −0.81 to −0.35, an increase of the time delay from 1 to 2 goes along with a change to $r_{DET} = -0.50$.

The optimum correlation coefficients taking optimum *RR* into account (row 1 and 2 in Table 19.2) do not lead to satisfying results either. The best r_{RR} takes a value of merely −0.60. The best $r_{DET} = -0.75$ is achieved via RPs again with a checkerboard-like structure as in Fig. 19.5 and the results are again sensitive against change of embedding parameters as for the minimum r_{DET}.

Earlier work by the author with exemplary evaluations revealed that increasing l_{min} can lead to significant improvements. In [2], results for $\varepsilon = 2$, $d = 3$, $\tau = 6$ and $l_{min} = 10$ are presented. The Recurrence Plots for this choice of recurrence threshold and embedding parameters show for the situation without as well as with porosity partly diagonal thick lines and, especially on the upper right, almost black regions (Fig. 19.6). The decaying amplitude of the original time series apparently leads to a trajectory which falls for later instants in time completely into the ball with radius ε in state space.

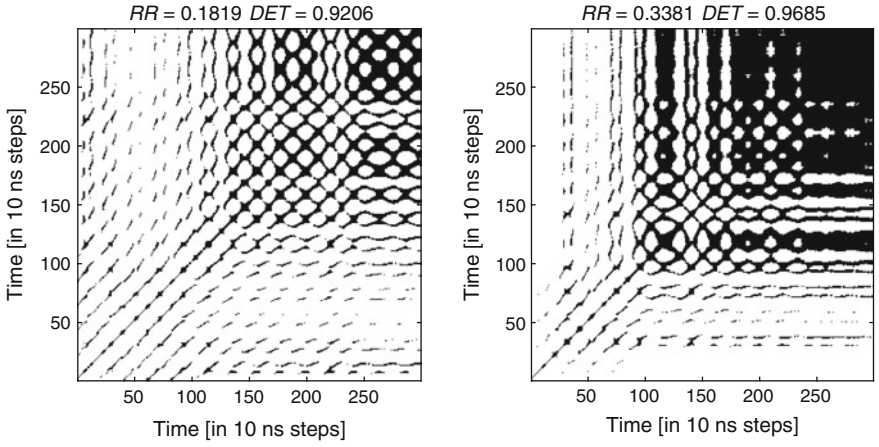

Fig. 19.5 Recurrence Plots for $\varepsilon = 0.8$ (Euclidean distance), $d = 3$, $\tau = 1$ from first scan of CG070M15, non-porous (*left*, A-scan with coordinates 2,20) and porous area (*right*, A-scan with coordinates 2,50)

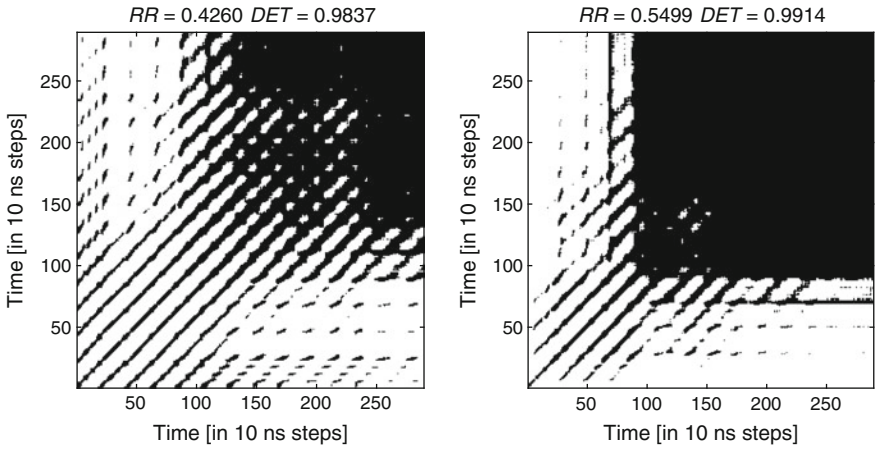

Fig. 19.6 Recurrence Plots for $\varepsilon = 2$ (Euclidean distance), $d = 3$, $\tau = 6$ from first scan of CG070M15, non-porous (*left*, A-scan with coordinates 2,20) and porous area (*right*, A-scan with coordinates 2,50)

Porosity has a decreasing effect on the amplitudes of the intermediate echoes in the current scans. The pores reflect the ultrasonic wave in other directions than to the ultrasonic transducer and thus reduce the measured echoes coming from the CFRP material. This loss in ultrasonic energy is reflected in an increase of recurrence, especially in the upper right of the Recurrence Plots (Fig. 19.6 right). This increase is only severe for the area of highest porosity, from which this RP is taken: the correlation between Recurrence Rate and BWE over all the A-scans is merely $r_{RR} = -0.12$. The feature determinism however leads to $r_{DET} = -0.75$ with $l_{min} = 2$. A significantly better value -0.87 is achieved for a minimum line length from 7 to 15 (with the highest in absolute value, precision of 4 decimals, for $l_{min} = 8$). Without the non-recurrent gaps in the Recurrence Plots due to the reduced amplitude for the situation with porosity, considerable more diagonal lines are detected: the feature *DET* increases with higher porosity (Fig. 19.7). This effect may also be detected with the RQA features laminarity or trapping time, not investigated here, which are related to the occurrence of vertical lines in the Recurrence Plot [35].

The range of *DET* equals 0.242 for $l_{min} = 8$, which is considered as sufficient. The choice of embedding parameters is insensitive to variations in comparison to results presented above: The correlation coefficient decreases to a minimum in absolute value of -0.70 for a variation of either d or τ by ± 1.

The optimum embedding parameters determined by autocorrelation and false nearest neighbours method (Sect. 19.5, $d = 4$ and $\tau = 5$ to $\tau = 6$) are similar to the ones for which the above described good correlation on sample CG070M15 was achieved.

This correlation on CG070M15 with RQA parameters $d = 3$ and $\tau = 6$, recurrence threshold $\varepsilon = 2$ (Euclidean distance) and $l_{min} = 8$ is checked on a second scan

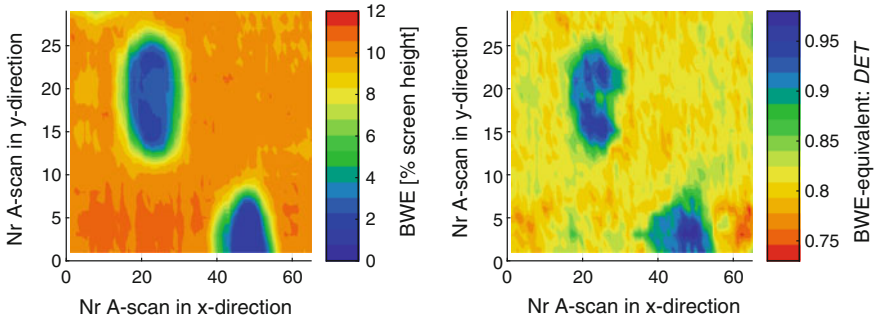

Fig. 19.7 C-scans of CG070M15 first scan: BWE (*left*) and *DET* as BWE-equivalent with RQA parameters $\varepsilon = 2$ (Euclidean distance): $d = 3$, $\tau = 6$ and $l_{min} = 8$ (*right*)

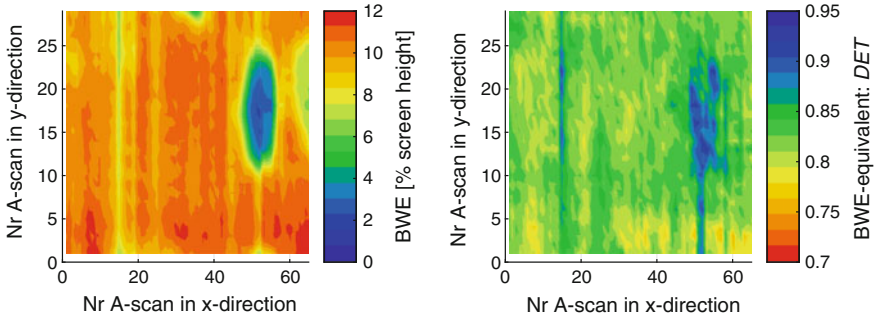

Fig. 19.8 C-scans of CG070M13: BWE (*left*) and *DET* as BWE-equivalent with RQA parameters $\varepsilon = 2$ (Euclidean distance): $d = 3$, $\tau = 6$ and $l_{min} = 8$ (*right*)

of CG070M15 ($r_{DET} = -0.89$), on a scan of CG070M13 ($r_{DET} = -0.69$) as well as of CG070M14 ($r_{DET} = -0.76$). The poorest correlation coefficient of CG070M13 may be related to the fact that the amount of porosity is not as high as in CG070M15 (cf. C-scans Fig. 19.8; the range of the color scale is identical to Fig. 19.7). The poorer scan quality may also negatively affect the correlation. However, even the vertical line obviously showing a slip-stick effect of scanning at approximately $x = 15$ is represented in the *DET*-C-scan.

The amplitude effect observed with Recurrence Quantification Analysis turned out to be detectable with the Quartile Coefficient of Disperson *QCD* [47], too. For its determination, the time series is sorted in ascending order of amplitude values. It is then in the present case calculated as

$$QCD = \frac{x_{sort,226} - x_{sort,76}}{x_{sort,226} - x_{sort,76} - 2x_{min}}. \tag{19.10}$$

An offset was added to the time series to make all values nonnegative, the lowest value equalling zero. This is reflected by subtracting x_{min} twice from the denominator in (19.10).

Using *QCD* as back-wall echo equivalent, a correlation coefficient $r_{QCD} = 0.81$ for the first scan of CG070M15 is achieved, slightly lower in absolute value than with Recurrence Quantification Analysis. Refer to Table 19.4 in Sect. 19.5.2 for further results of r_{QCD} for the other scan of the same part and scans of the other investigated parts.

Whether this amplitude effect exists for other CFRP materials and "natural" porosity, whether it can be detected by RQA and whether it can be determined using the Quartile Coefficient of Dispersion or other simple means will be investigated in further steps of this ongoing research.

To summarise results for Euclidean distance, the RQA feature determinism has been found to provide a good back-wall echo equivalent for the investigated sample set. This is achieved with embedding parameters near to the ones determined with autocorrelation and false nearest neighbours method, leading to higher determinism for higher porosity. The cause is an amplitude effect, corresponding to hypothesis (2) (see Sects. 19.1 and 19.3.3) with a loss of energy when porosity is present. This effect can also be observed with the simpler means of Quartile Coefficient of Dispersion, with a correlation almost as good as with RQA.

Next, angular distance is investigated, a metric independent of amplitude effects by definition.

19.5.2 Angular Distance

The varied RQA parameters for angular distance (Table 19.1) are checked on extrema of the coefficients of correlation of *RR* or *DET*, respectively, with the back-wall echo. With the experience from Euclidean distance, $l_{min} = 4$ and $l_{min} = 8$ are included into the investigated RQA parameters. Three of the four extrema of the coefficients are less than 0.7 in absolute value (Table 19.3) and are thus not further considered.

The maximum coefficient is $r_{DET} = 0.75$ for the correlation between *DET* and BWE. The range of determinism $R_{DET} = 0.338$, larger than any range of optimum

Table 19.3 Optimum coefficients of linear correlations between BWE-equivalent and BWE in dependence of RQA parameters for $l_{min} = 2, 4, 8$ for angular distance

Extrema r	Correlation coefficient r	Mean RR/R_{RR}	Mean DET/R_{DET} for $l_{min} = 8$	ε	d	τ	l_{min}
min r_{RR}	$r_{RR} = -0.64$	0.754/0.111	0.947/0.005	0.75π	8	1	–
max r_{RR}	$r_{RR} = 0.64$	0.101/0.063	0.551/0.102	0.1π	6	1	–
min r_{DET}	$r_{DET} = -0.55$	0.012/0.028	0.549/0.705	0.075π	10	7	8
max r_{DET}	$r_{DET} = 0.75$	0.504/0.201	0.910/0.338	0.5π	4	8	8

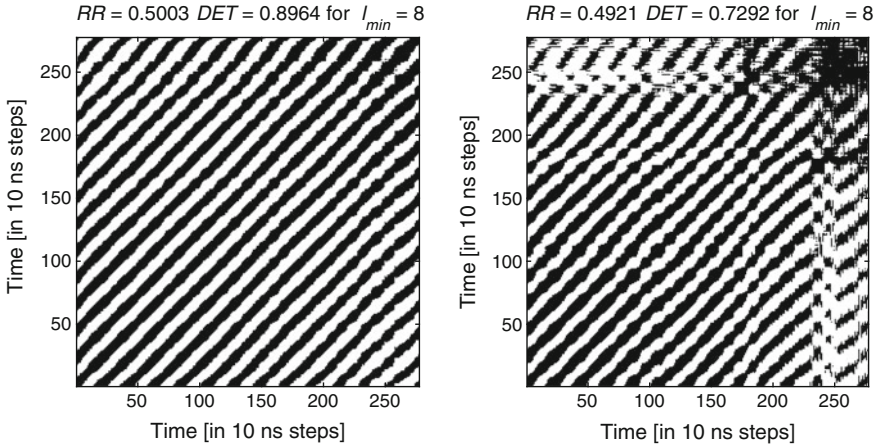

Fig. 19.9 Recurrence Plots for $\varepsilon = 0.5\pi$ (angular distance), $d = 4$, $\tau = 8$ from first scan of CG070M15, non-porous (*left*, A-scan with coordinates 2,20) and porous area (*right*, A-scan with coordinates 2,50)

correlation coefficients for Euclidean distance. The according Recurrence Plot derived from an A-scan from a non-porous area exhibits as expected no special dependence on time and consists almost exclusively of thick diagonals (Fig. 19.9 left). For porosity, the Recurrence Plots loses partly its structure, which shows the reduced periodicity of the signal. The Recurrence Rate is almost the same as for the non-porous situation, but the determinism is significantly decreasing (Fig. 19.9 right).

The according C-scan for determinism *DET* displays the correlation (Fig. 19.10), which is not as good as for the (in absolute value) greater minimum r_{DET} for Euclidean distance.

The optimum embedding dimension here equals the one determined by false nearest neighbours method, whereas the time delay is 2–3 time steps higher than the one determined with autocorrelation (Sect. 19.5).

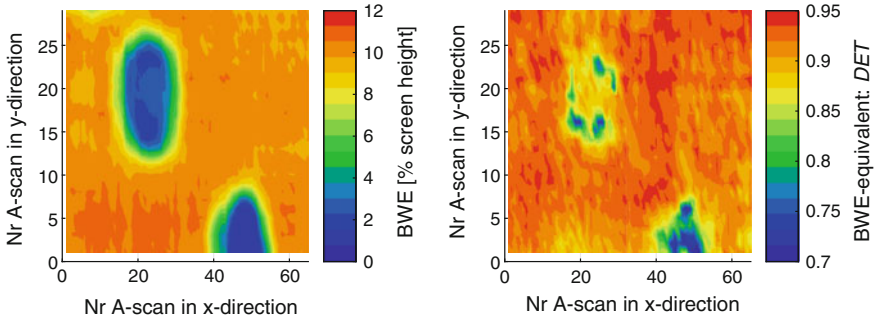

Fig. 19.10 C-scans of CG070M15 first scan; BWE (*left*) and *DET* as BWE-equivalent with RQA parameters $\varepsilon = 0.5\pi$ (angular distance), $d = 4$, $\tau = 8$ and $l_{min} = 8$ (*right*)

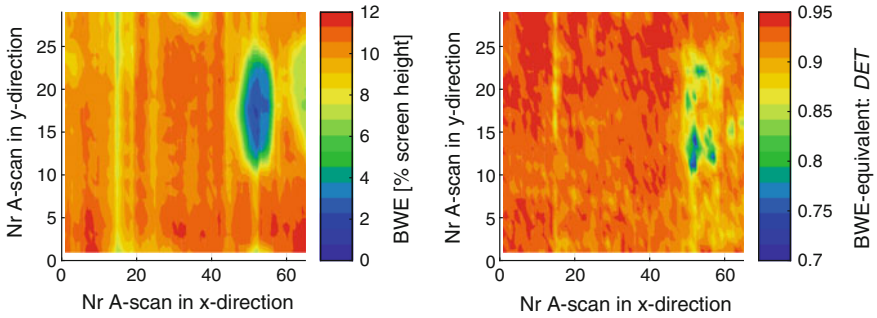

Fig. 19.11 C-scans of CG070M13: BWE (*left*) and *DET* as BWE-equivalent with RQA parameters $\varepsilon = 0.5\pi$ (angular distance), $d = 4$, $\tau = 8$ and $l_{min} = 8$ (*right*)

A variation of embedding parameters by ±1 decreases r_{DET} to a minimum of 0.65, which is considered as rather robust. The minimum line length has an influence of the same magnitude: $r_{DET} = 0.72$ and $r_{DET} = 0.69$ for $l_{min} = 4$ and $l_{min} = 2$, respectively.

For the optimum RQA parameters, which led to the maximum $r_{DET} = 0.75$ for CG070M15, the correlation coefficient is determined at a second scan of CG070M15 and at one scan each of CG070M13 and CG070M14. The correlation coefficients are $r_{DET} = 0.73$, $r_{DET} = 0.57$ (C-scans in Fig. 19.11) and $r_{DET} = 0.69$, respectively.

Summarising for angular distance, the RQA feature determinism has been found as an appropriate back-wall echo equivalent. In contrast to the results with Euclidean distance, the feature decreases with increasing porosity. Amplitude effects are not taken into account with angular distance by definition, and the decreasing determinism shows a lower periodicity in the ultrasonic wave propagation, caused by the pores, corresponding to hypothesis (1) in Sects. 19.1 and 19.3.3.

The results with best parameters for angular distance in this section, originating from the reduction of determinism, as well as the results for best parameters with Euclidean distance and the hence derived Quartile Coefficient of Dispersion (Sect. 19.5.1), both originating from amplitude effects, are summarised in Table 19.4.

Table 19.4 Correlation coefficients r with the optimum RQA parameters for Euclidean distance and angular distance and Quartile Coefficient of Dispersion

	CG070M15 first scan	CG070M15 second scan	CG070M13	CG070M14
$r_{DET,Euc}$ with $\varepsilon = 2$, $d = 3$, $\tau = 6$, $l_{min} = 8$	−0.87	−0.89	−0.69	−0.76
r_{QCD} with time series offset	0.81	0.81	0.67	0.76
$r_{DET,ang}$ with $\varepsilon = 0.5\pi$, $d = 4$, $\tau = 8$, $l_{min} = 8$	0.75	0.73	0.57	0.69

19.6 Conclusion and Outlook

Recurrence Quantification Analysis (RQA) is proposed as a tool to derive a "back-wall echo equivalent" for evaluating porosity in Carbon Fibre Reinforced Polymers with ultrasonic pulse-echo testing. Three samples with local porosity fields of different severity have been measured. The trajectory in state space has been reconstructed from the measured time series and Recurrence Plots have been generated, using Euclidean and angular distance. The RQA features Recurrence Rate and determinism *DET* have been determined out of the Recurrence Plots. For the sample with highest porosity, embedding parameters and the recurrence threshold, and partially the minimum line length have been varied. The results of Recurrence Rate and determinism have been correlated with the back-wall echo to assess their suitability as back-wall echo equivalent. Theoretical considerations about optimum Recurrence Rate for Euclidean distance and the robustness of results have been taken into account.

The feature determinism has been found to present an appropriate back-wall echo equivalent. When using Euclidean distance, the coefficient of correlation with the back-wall echo on the sample with highest porosity is −0.87; *DET* increases with increasing porosity. This is caused by amplitude effects; measured intermediate echoes and thus the time-localised energy decrease due to reflections from the pores not reaching the ultrasonic transducer. This effect can be shown with simpler evaluations (quartile coefficient of dispersion) as well, leading to a correlation coefficient of 0.81.

Using angular distance, the coefficient of correlation of the feature determinism with the back-wall echo takes the value 0.75 for the sample with highest porosity. Angular distance excludes amplitude effects by definition; the measured effect is based on lower periodicity of the ultrasonic wave propagation with higher porosity content, leading to lower values of determinism.

The presented first step of ongoing research will be continued by exploitation of other features out of recurrence plots and include other samples of different CFRP material and with "natural" porosity.

Acknowledgement This work represents (together with [2]) the first step of my doctoral thesis. I am very grateful to my employer Airbus Operations for supporting this outside-my-working-time project, including allowance to use lab equipment for all measurements and evaluations.

I sincerely appreciate that Dr Norbert Marwan, the reviewer team of this book and my doctoral supervisor Professor Dr Peter Maaß led me in the right direction to interpret and to carve out the observed effects, especially when using Euclidean distance.

Appendix

Considerations About Optimum Recurrence Rate

Based on the dominating 5 MHz frequency in the time series (Fig. 19.4) the Recurrence Rate for an embedding of a sine wave with a wavelength of 20 time steps (corresponding to 5 MHz at a sampling rate of 100 MHz) is considered. With any recurrence threshold $\varepsilon > 0$ (because of the ideal motion exactly equalling itself after one wave length) and less than a value depending on the amplitude of the sine wave and the embedding dimension (e.g. 0.3167 or 0.6258 times the amplitude for $d = 3$ or 8, respectively) this leads to a Recurrence Rate of 0.05, i.e. 5 %. The according Recurrence Plot consists of diagonal lines ([26] Fig. 1A) occurring every 20 time steps. Thus it can be assumed that Recurrence Rates around 0.05 are appropriate for Recurrence Quantification Analysis in the present case and may reveal changes in determinism caused by porosity.

The Recurrence Rate differs significantly with the embedding dimension, as the example with the ideal sine wave above shows (this is due to the larger distances between points in state space with a larger embedding dimension for one and the same time series).

Consequently, the recurrence thresholds, for which the mean Recurrence Rate (i.e. the average of the Recurrence Rates of all measured time series per scan) is nearest to a value of 0.05 (Fig. 19.12 right), are determined (Fig. 19.12 left).

For the values of ε shown in Fig. 19.12, the coefficients of linear correlation between Recurrence Rate *RR* and back-wall echo BWE are shown in Fig. 19.13 left hand side. The coefficients for the according correlation between determinism *DET* and BWE are displayed in Fig. 19.13 right hand side.

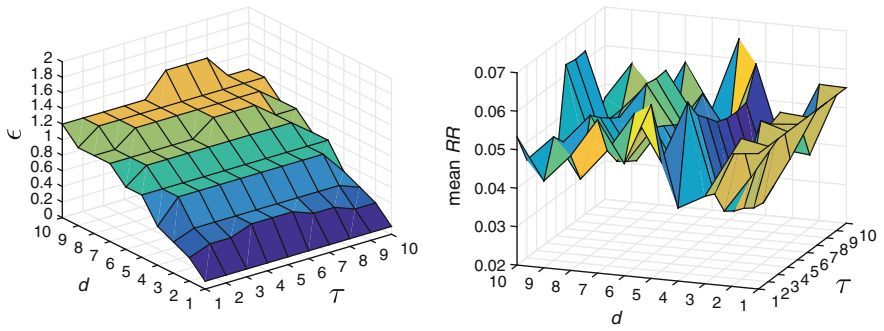

Fig. 19.12 *Left* Recurrence thresholds for which mean values of the Recurrence Rate are nearest to 0.05 (Euclidean distance); *Right* Mean values of the related Recurrence Rate

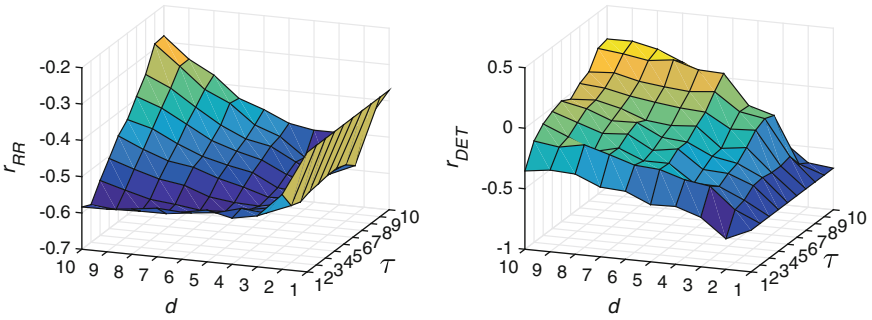

Fig. 19.13 Coefficients of linear correlation between *RR* and BWE (*left*) and between *DET* and BWE (*right*) for ε as in Fig. 19.12, i.e. for *RR* near to 0.05 (Euclidean distance)

References

1. C. Brandt, "Schichtdickenmessung auf CFK-Oberflächen", in *DGZfP-Jahrestagung 2010 (annual conference of German society for NDT), Online-Proceedings on ndt.net*, 10.-12.05.2010, Erfurt, Germany, (2010)
2. C. Brandt, P. Maaß, A state space approach for the non-destructive evaluation of CFRP with ultrasonic testing, in *7th Symposium on NDT in Aerospace, Online-Proceedings on ndt.net,* 16.-18.11.2015, Bremen, Germany, (2015)
3. U. Schnars, R. Henrich, Applications of NDT methods on composite structures in aerospace industry, in *CDCM 2006—Conference on Damage in Composite Materials 2006, Online-Proceedings on ndt.net*, 18./19.09.2006, Stuttgart, Germany (2006)
4. A. Lozak, C. Boller, A. Bulavinov, R. Pinchuk, J. Kurz, D. Sednev, Phase statistics and spectral analysis of ultrasonic signals for CFRP component assessment, in *EWSHM—7th European Workshop on Structural Health Monitoring*, 8.-11.7.2014, Nantes, France (2014), https://hal.inria.fr/file/index/docid/1022992/filename/0442.pdf [2015, 10/12]
5. J.M. Nichols, S.T. Trickey, M. Seaver, Damage detection using multivariate recurrence quantification analysis. Mech. Syst. Signal Process. **20**(2), 421–437 (2006)
6. Y. Lu, J.E. Michaels, State space feature extraction applied to diffuse ultrasonic signals using simulated chaotic excitations, in *Review of Progress in Quantitative Nondestructive Evaluation*, ed. by D.O. Thompson, D.E. Chimenti, American Institute of Physics (2006), pp. 625–632
7. T.R. Fasel, M.D. Todd, G. Park, Piezoelectric active sensing using chaotic excitations and state space reconstruction, in *Proceedings of SPIE,* vol. 5768, no. -, pp. 253–263 (2005)
8. M. Cacciola, D. Costantino, C.M. Francesco, M. Versaci, Dynamical analysis for flaw shape identification in non linear eddy current tests. COMPEL **26**(4), 1081–1094 (2007)
9. A. Carrión, R. Miralles, G. Lara, Measuring predictability in ultrasonic signals: an application to scattering material characterization. Ultrasonics **54**(7), 1904–1911 (2014)
10. N. Dominguez, B. Mascaro, Ultrasonic non-destructive inspection of localised porosity in composite materials, in *ECNDT 2006—9th European Conference on NDT 2006 Online-Proceedings on ndt.net*, 18./19.09.2006, Berlin, Germany (2006)
11. *Composite Materials Handbook Volume 3: Polymer Matrix Composites* 2012, [Homepage of Department of Defense, USA], [Online]. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.232.439&rep=rep1&type=pdf [2015, 10/06]
12. K.K. Chawla, *Composite Materials: Science and Engineering* (Springer, New York, NY [u.a.], 2013)
13. J.L. Davis, *Mathematics of Wave Propagation* (Princeton University Press, Princeton, 2000)

14. J. Krautkrämer, H. Krautkrämer, W. Grabendörfer, *Werkstoffprüfung mit Ultraschall* (Springer, Berlin [u.a.], 1986) (also available as English edition: J. Krautkrämer, H. Krautkrämer, *Ultrasonic Testing of Materials,* 4th edn. (Springer, Berlin, 1990))
15. J. Blitz, G. Simpson, *Ultrasonic Methods of Non-destructive Testing* (Chapman & Hall, London [u.a.], 1996)
16. G.L. Gooberman, *Ultrasonics: Theory and Application* (The English Universities Press Ltd., London, 1968)
17. D. Hinrichsen, A.J. Pritchard, *Mathematical Systems Theory,* 1, corrected softcover printing edn. (Springer, Heidelberg u.a, 2010)
18. H.W. Broer, F. Takens, *Dynamical Systems and Chaos* (Springer, New York, NY [u.a.], 2011)
19. E. Ott, *Chaos in Dynamical Systems* (Cambridge University Press, Cambridge [u.a.], 1993)
20. C. Diks, *Nonlinear Time Series Analysis: Methods and Applications* (World Scientific, Singapore, 1999)
21. F. Takens, Detecting strange attractors in turbulence, in *Lecture Notes in Mathematics: Dynamical Systems and Turbulence, Warwick 1980: Proceedings of a Symposium Held at the University of Warwick 1979/8*0, ed. by D. Rand, L. Young (Springer, Berlin [u.a.], 1981), pp. 366–381
22. N. Packard, J. Crutchfield, J.D. Farmer, R. Shaw, Geometry from a time series. Phys. Rev. Lett. **45**(9), 712–716 (1980)
23. D. Aeyels, Generic observability of differentiable systems. SIAM J. Control Optim. **19**(5), 595–603 (1981)
24. T. Sauer, J.A. Yorke, M. Casdagli, Embedology. J. Stat. Phys. **64**(3/4), 579–616 (1991)
25. J.P. Huke, *Embedding Nonlinear Dynamical Systems: A Guide to Takens' Theorem* (2006) [Homepage of University of Manchester], [Online]. https://www.eprints.ma.man.ac.uk/175/ [2015, 04/08]
26. N. Marwan, M. Carmen Romano, M. Thiel, J. Kurths, Recurrence plots for the analysis of complex systems. Phys. Rep. **438**(5–6), 237–329 (2007)
27. E. Ott, T. Sauer, J.A. Yorke, *Coping with Chaos: Analysis of Chaotic Data and the Exploitation of Chaotic Systems* (Wiley, New York, NY [u.a.], 1994)
28. D.R.J. Chillingworth, *Differential Topology with a View to Applications* (Pitman, London [u.a.], 1976)
29. T.D. Sauer, Attractor reconstruction. Scholarpedia **1**(10), 1727 (2006), http://www.scholarpedia.org/article/Attractor_reconstruction [2014, 09/21]
30. H. Kantz, T. Schreiber, *Nonlinear Time Series Analysis,* 2, repr. edn. (University Press, Cambridge [u.a.], 2005)
31. H. Poincaré, Sur le problème des trois corps et les équations de la dynamique. Acta Mathematica **13**, 1–270 (1890)
32. H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory* (Princeton University Press, Princeton, 1981)
33. A. Katok, B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems* (Cambridge University Press, Cambridge [u.a.], 1996)
34. J.-P. Eckmann, S.O. Kamphorst, D. Ruelle, Recurrence plots of dynamical systems. EPL (Europhys. Lett.) **4**(9), 973 (1987)
35. C.L. Webber, N. Marwan, *Recurrence Quantification Analysis: Theory and Best Practices* (Springer, Cham [u.a.], 2015)
36. E. Kreyszig, *Introductory Functional Analysis with Applications* (1989), reprint from 1978 edn. (Wiley, New York [u.a.])
37. C. Ioana, A. Digulescu, A. Serbanescu, I. Candel, F. Birleanu, Recent advances in non-stationary signal processing based on the concept of recurrence plot analysis, in *Translational Recurrences*, ed. by N. Marwan, M. Riley, A. Giuliani, C.L. Webber Jr. (Springer International Publishing, 2014), pp. 75–93
38. S. Schinkel, N. Marwan, J. Kurths, Order patterns recurrence plots in the analysis of ERP data. Cogn. Neurodyn. **1**(4), 317–325 (2007)

39. J. Groenenboom, R. Snieder, Attenuation, dispersion, and anisotropy by multiple scattering of transmitted waves through distributions of scatterers. J. Acoust. Soc. Am. **98**(6), 3482–3492 (1995)
40. F. Damaschke, *Quantitative Bewertung der inneren Struktur von Faserverbundwerkstoffen mittels Ultraschallsignalanalyse*, Universität Oldenburg (1996)
41. A. Gausmann, *Quantitative Bewertung von Poren in Faserverbundwerkstoffen mittels Ultraschalltechnik*., Fachhochschule Osnabrück, Fachbereich Werkstofftechnik (1995)
42. S. Barut, V. Bissauge, G. Ithurralde, W. Claassens, Computer-aided analysis of ultrasound data to speed-up the release of aerospace CFRP components, in *18th World Conference on Nondestructive Testing, Online-Proceedings on ndt.net,* 16.-20.4.2012, Durban, South Africa (2012)
43. J.P. Zbilut, C.L. Webber, Embeddings and delays as derived from quantification of recurrence plots. Phys. Lett. A **171**(3–4), 199–203 (1992)
44. W.H. Press, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, Cambridge [u.a.], 2007)
45. N. Marwan, *Cross Recurrence Plot Toolbox for Matlab®, Ver. 5.18 (R29.3*) (2015), http://tocsy.pik-potsdam.de/CRPtoolbox [2015, 10/1/2015]
46. N. Marwan, How to avoid potential pitfalls in recurrence plot based data analysis. Int. J. Bifurcat. Chaos **21**(04), 1003–1017 (2011)
47. D.G. Bonett, Confidence interval for a coefficient of quartile variation. Comput. Stat. Data Anal. **50**(11), 2953–2957 (2006)

Index

C.L. Webber, Jr. et al. (eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons*, Springer Proceedings in Physics 180,
DOI 10.1007/978-3-319-29922-8