University of Reading

School of Systems Engineering

Doctor of Philosophy thesis

KurSL: a model of coupled oscillators based on Kuramoto's coupling and Sturm-Liouville theory

Dawid Laszuk

Supervisors:

Prof. Sławomir J. Nasuto

Prof. Ingo Bojak

Dr. Oswaldo Cadenas

March 2018

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Acknowledgements

My deepest gratitude goes to the Society, without which I would not be able to devote my time to intellectual pursuit. It has spared me from having to spend time searching for food sources or fearing for my life on daily basis. It has also provided me some level of independence from the achievements and mistakes of my relatives. I am not sure if I have seen much further, but everything I have seen so far was by standing on the shoulders of giants whose existence and well-being was supported by people of all sizes.

I greatly appreciate the support of my supervisor and personal giant — Prof. Sławomir Nasuto. If it were not for him, my research journey would have been different. Busy as he was, he always found time to discuss any problems that I had. I am thankful for all the interesting and stimulating conversations, even though, perhaps, some of them could have been shorter. I regret not to follow on his more philosophical discussion baits which, unfortunately, happened in my meta-philosophical interest period during which I considered all strictly philosophical discussions to be meaningless.

I would also like to thank Dr. Oswaldo Cadenas for his support and motivation to execute any ideas, even those small ones. He was able to introduce a positive twist to my self-criticality. Where I saw failure and insignificance, he saw learning opportunity and practical experience. It is thanks to him that many of my side projects have been publicly shared and well received.

Last, but definitely not least, I would like to point out how grateful I am to my parents — Jolanta Danielewicz & Wiktor Laszuk. They managed to put up with me for a very long time and, despite shattering their dreams of reaching high income position, they still invite me over for Christmas. Their constant reminders about other peoples' successes kept me working even when I really wanted to sleep. Unfortunately,

they do not understand English, so hopefully seeing their names here will please them and the plentiful of colourful graphs will provide some entertainment for at least a few minutes. *Rodzice: dziękuję za ciągłe wsparcie i za to że byliście!*

Abstract

Methods commonly used to analyse oscillatory systems, such as short-time Fourier or wavelet transforms, require predefined oscillatory structures or fine-tuning of method's parameters. These limitations may be detrimental for an adequate component description and can introduce bias to the interpretation. This thesis addresses the challenge of identifying interacting components in a signal by introducing a model of coupled oscillators. The proposed model consists of two parts: Sturm-Liouville self-adjoint ordinary differential equation (ODE) and Kuramoto's coupling model. The resulting model, *KurSL*, is described by a set of coupled ODEs producing general amplitude- and frequency-modulated mutually interacting oscillations. The complexity of these equations depends on the definition of the coupling function, the number of oscillators and the initial state of each oscillator. Thus, the performance of the *KurSL* decomposition can be characterised in terms of the model parameters optimisation. After introducing the model, the thesis provides analysis and discussion of the *KurSL* with examples of its usage. The method is firstly tested on various synthetic data that were generated from simulated stationary and dynamical processes. Such testing allows capturing various characteristics that are desirable in coupled oscillatory components such as phase and amplitude dynamics. Subsequently, experiments were performed on empirical EEG signals recorded from patients with epilepsy. Validation of these experiments is through comparisons to different orders of the *KurSL* and to other time-frequency methods. Overall results indicate that the *KurSL* method provides a more detailed description of oscillatory processes than the Huang-Hilbert transform and it provides insights comparable to manually tuned short-time Fourier transform and Morlet-based wavelet time-frequency representations. However, the advantage of the *KurSL* is that the similar results can be achieved with a finite number of components. Moreover, in contrast to the mentioned representations which, due to finite resolution, are unable to localise time-frequency events precisely, the *KurSL* provides an instantaneous description. This exactness allows to identify any modulations in both time and frequency domains and thus better describe the behaviour of the analysed system.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

One of the inherent human urges is to understand the underlying mechanics of the surrounding environment. Since ancient times people have been trying to understand the behaviour of all observable events [1, 2]. A significant realisation, which allowed for further developments, was that with every physical object a particular type of information could be associated and expressed as a set of values. In many cases, the more one knows about the previous and current state of the system the more precise estimation can be deduced [3]. In general, any data have little value unless they are recorded and analysed. Once processing is finished, one can attempt to describe the evolution of a system through the *modelling* [4]. Such description allows simulating the system's behaviour and thus predict its state at the desired time. These predictions often simplify the reality and project it onto a few dependent parameters which best predict future response within a given error margin.

A general approach to modelling the reality can be presented in an iterative framework of the experiment-theory cycle. To explain some phenomenon, scientists have first to propose a hypothesis and experiments which could validate it [5, 6]. Then, depending on the results, they can update their understanding and design even more precise experiments confirming a more refined hypothesis. Further advancements in modelling came with the development of Physics and Mathematics which lead to the creation of *mathematical modelling*. Using mathematics to describe phenomena gave modelling a more rigorous foundation. These models can have different levels of complexity that depend on the amount of information available and the expected precision of predic-

tion. For example, to send a satellite into an orbit, it is enough to consider gravity from a Newtonian perspective. However, when trying to determine the position of an object via a GPS system, it is essential to include Einstein's special relativity.

When modelling a system, it is essential to consider its complexity, which can be dependent on the type and number of objects within. Those objects can be, for example, separated into temporal or spatial domains [7]. Such distinctions, however, do not forbid any interactions between the objects or with the surrounding environment. In fact, these interactions are always present, but in the majority of cases, communications can be omitted as they have too small an effect on the system. The main difficulty of constructing a model is to determine which interactions are essential for predictions with a defined accuracy.

A particular interest in the scientific community is given to systems which periodically repeat their behaviour or visit a specific state [5, 8]. This attention is due to an abundance of examples present in nature ranging from the Earth's day-night and seasonal behaviour, through laser generation [9] and fireflies flashing synchronisation to the cardiovascular system [10]. The periodicity of a system does not necessarily mean that it will repeat the same behaviour over time. Perturbation to its state may be due to outside influences and their effect will vary depending on the scale at which they are analysed. For example, Earth's rotation around the Sun is continuously disturbed by other astronomical objects and thus its period continually changes. However, the scale of this phenomenon is outside of human general perception and, unquestionably, year length is considered constant. In this case, for the majority of calculations, one can assume objects on Earth are isolated from the influence of other astronomical objects. In some systems, however, interactions are much more pronounced and can have a dominant impact on the overall behaviour. One such example is the brain [9, 11]. It is common to model brain regions as oscillatory objects that interact with each other, when explaining the behaviour of neural activities [12, 13].

As it happens, many methods commonly used to extract information from oscillatory systems are not suitable for such analysis [14, 15]. Some of these, like Fourier transform, assume mathematically idealistic behaviour of the system, such as linearity and stationarity. These assumptions are in contrast with most systems in nature,

including the brain or the climate, which are non-linear and non-stationary [16] (see Appendix A.1). Moreover, these methods often neglect possible interactions between a system and its surrounding. Such disadvantages have long been understood and the attempts to mitigate them have influenced the development of data-driven methods [17, 18].

Data-driven classification refers to a method's ability to shape its underlying model to fit the provided data. These methods often assume a general behaviour, which is then limited depending on the input's characteristics. One of the reasons for increased interest in data-driven methods is the progress in technology, mainly in storage capacity and computational power. Such developments allowed for shifting away from computational paradigms where data had to be manually selected and limited to only the essential characteristics. The memory limitations are currently much less stringent, which allows one to analyse a significant amount of data, and with increasingly high computational throughput the preliminary analysis takes little time. Overall, technological development enables including more factors in analysis, for example, those responsible for interactions, and producing results in reasonable time. However, this trend also caused a shift in analytical approaches towards applying algorithms in a black-box manner where the emphasis is more on their output and less on their meaning. An example of such an approach is the empirical mode decomposition (EMD, Section 2.3). Since its proposal, the method has been empirically verified by many researchers to provide physically meaningful results [19–21]. Nevertheless, despite many attempts, its exact behaviour is yet to be discovered as the method is lacking a mathematical framework.

The goal of this project was to develop a robust method which would be able to extract meaningful information from oscillatory systems. The proposed method, *KurSL*, aims to be general purpose, including non-linear and non-stationary processes, and thus its properties are data-driven. The work was inspired by EMD's idealised features such as extraction of physically meaningful oscillatory components. However, in contrast to EMD's development, the core emphasis of the KurSL is to start from a mathematically sound framework. Such a framework is based on a combination of two components: Kuramoto's coupling model and Sturm-Liouville theory. This

combination allows for modelling systems in which objects interact with one another and maintain oscillatory behaviour. The KurSL aspires to provide components that are both mathematically well-defined and providing physically meaningful results.

Outline

The following chapter (Chap. 2) describes popular data analysis methods. It begins (Sec. 2.2) by introducing and discussing classical time-frequency transformations such as Fourier and wavelet transformations. After inspecting their advantages and limitations, it follows (Sec. 2.1) with an introduction of a few decomposition methods out of which a particular interest is given to the empirical mode decomposition (EMD) which is explained in greater detail in Section 2.3.

Model-based representation as a complementary approach to data analysis is discussed in Chapter 3. Its primary emphasis is on models that describe oscillatory systems. The first Section 3.1 discusses Sturm-Liouville theory which refers to an oscillatory system defined by a second order differential equation. The following Section 3.2, however, presents a model of a system in which oscillators are allowed to interact with each other.

In Chapter 4 more thorough investigations of some of the EMD's properties are presented. These include analysis of computational stability under different data formats (Sec. 4.1) and defining an objective metric in the results space (Sec. 4.2). Additionally, Section 4.3 focuses on an analysis of the frequency-mixing behaviour. This phenomenon describes the appearance of a component when input components have similar frequencies.

Inspired by the properties of the EMD and its idealised behaviour, the KurSL model is introduced in Chapter 5. The proposed model is based on a combination of two models as mentioned earlier, i.e. Sturm-Liouville and Kuramoto. After presenting the formal definition in Section 5.1, its general properties and a few examples are presented in the following Section 5.2. A generalisation of the model is described in Section 5.3, where the order of the model denotes the number of applied harmonics terms in the coupling.

Chapter 6 presents an application of the KurSL model as a data-driven adaptive decomposition method. It begins (sec. 6.1) with a description of the process of obtaining parameters, which are extracted based on the input's properties. This is additionally summarised in the form of an algorithm and presented in Section 6.2. The following sections discuss the method's convergence (sec. 6.3) and parametric stability (sec. 6.4). Moreover, Section 6.4 discusses how traversing through KurSL's parameter space affects its components in time series space.

Examples of the method's applications are presented in Chapter 7. These include decomposition of known synthetic data (sec. 7.1) performed for different orders of KurSL model. Additionally, Section 7.2 describes a dynamic approach to KurSL data decomposition in which the change of parameters over time is considered. As a summary and final example, the method is presented on empirical EEG signals in Section 7.3.

The final Chapter 8 of the document contains conclusion of the research. Section 8.1 inspects obtained results and acknowledges potential limitations. Finally, it discusses possible improvements to the model and the method, which are presented in Section 8.2.

Chapter 2

Data analysis

The analysis of data starts even before the recording is started; it is initiated with the design of an experiment. In the first step, one has to identify the system and understand how it can be measured considering all its impurities. These include other objects not contained within the system. In nature, it is highly unlikely to record a signal which originated from a single source or sources that are not coupled with a noisy environment. Even if one can isolate them physically, the measurement can still be affected by the noise either from the measuring device's imperfections or through natural variations in the observing system. Knowing the limitations of experiment one should adjust methods of analysis appropriately. For this reason, it is commonly convenient to consider data X to be composed of noise ξ and many components c_n, which can correspond to different features or phenomena. In case of time t series, $X(t)$, with noise $\xi(t)$ contamination data can be expressed as

$$X(t) = \sum_{n=1}^{N} a_n c_n(t) + \xi(t), \tag{2.1}$$

where N is the number of the components and a_n is a scale of the contribution of component c_n to whole data. The process of extraction and identification of such components is called *decomposition*, and it is vital for analysing the system.

Unfortunately, there is not a unique approach on how to decompose data; conversely, any real dataset can be represented in infinitely many different ways. In a simplified example, number 4 can be expressed both as $1 + 3 = 4$ and $10 - 6 = 4$, or

any other combination of two or more numbers resulting in 4. Typically, to extract some meaningful information, one has to specify conditions or characteristics which are sought in the data. In case of many non-interfering sources this might be looking for components' independence, or in case of observing repeatable events, one would be interested in periodicity. In this thesis, the primary interest is in the latter situation, i.e. when phenomena have oscillatory behaviour.

This chapter briefly introduces some methods of data analysis. The first Section (2.1) describes two general decomposition methods and their assumptions. In Section 2.2 the emphasis is made on methods for time-frequency analysis. Final Section 2.3 presents a thorough study of an *empirical mode decomposition*, which is a general decomposition method that can produce time-frequency representation of a signal.

2.1 Data decomposition

In this section, two decomposition methods are introduced and discussed. The first method, matching pursuit (MP), focuses on an iterative matching of all user-defined components to the input data. This approach is in contrast to principal component analysis (PCA) extraction, which is defined to decompose a signal into uncorrelated components. These methods are presented in mentioned order in Sections 2.1.1 and 2.1.2, respectively.

2.1.1 Matching pursuit

One of the methods that decompose data into a set of functions is *matching pursuit* (MP). These components (*atoms*) are often selected to possess desirable properties dependent on the nature of the input [22]. Mallat and Zhang in their article introducing MP [23] used sine waveforms with Gaussian modulated envelopes, which in the signal processing community are often called *Gabor* functions and are defined as

$$g_\gamma(t) = K(\gamma) e^{-\pi\left(\frac{t-u}{s}\right)^2} \cos\left(\omega(t-u)+\phi\right), \tag{2.2}$$

where u is time displacement, s is scale, ω is frequency and ϕ is the phase shift. These parameters are often collectively represented as a set $\gamma = \{u, s, \omega, \phi\}$ and allow to compute normalisation factor $K(\gamma)$. An argument for choosing Gabor definition of an atom is its good behaviour under Wigner's transformation [14, 24]; it has the best localisation in such time-frequency representation. Once atoms are defined, they are stored in a large and redundant set called *dictionary*. Ideally, this set should contain all possible positions in the γ-parameter space; however, due to storage and computational limitations, one has to specify a finite set of these parameters.

The MP decomposition is performed in an iterative manner. At each iteration, atoms are compared to the previous iteration's residual to find the best match. The cost function varies, but it is often assumed to minimise $\mathcal{L}^2$ metric [23, 25]. An atom that minimises cost function is considered as a component and is subtracted from the input signal $R^1X = X - \langle R^0X, g_{\gamma_0}\rangle$. The residue is then carried over to the next iteration, and the process continues on a dictionary with a removed atom. The algorithm for MP process can be described more formally

$$\begin{cases} R^0X = X, \\ R^nX = \langle R^nX, g_{\gamma_n}\rangle g_{\gamma_n} + R^{n+1}X, \\ g_{\gamma_n} = \arg\max_{g_{\gamma_i}\in D} |\langle R^nX, g_{\gamma_i}\rangle|. \end{cases} \tag{2.3}$$

As a result, the set of obtained components approximates the original signal

$$X \approx \sum_{n=1}^{N} \langle R^nX, g_{\gamma_n}\rangle g_{\gamma_n}, \tag{2.4}$$

where the number of components N can be either defined in advance or will depend on the decomposition criteria.

In this method, the size of the *dictionary* and the form of *atoms* are essential. The bigger the set of available functions, the more likely it is that it contains the exact representation of the feature. However, the size also affects computation time since all atoms should be compared to the data at each iteration. Moreover, due to the greediness of the algorithm, it always tries to fit function to match data globally. Such fitting can introduce artefacts when, for example, the data consists of few components

with the similar parameter. In such case, the method instead of perfectly matching few parts it will assign a single mediocre fit as its dot product will be more significant. Subtracting non-existing element will introduce false structure into data that can be later wrongly identified as a component.

2.1.2 Principle component analysis

Another example of decomposition method is the *principle component analysis* [26]. Depending on imposed conditions and the source of the input data this method may be called differently, for example, empirical orthogonal functions (EOF) or Karhunen-Loeve decomposition [27]. The process tries to find a base that will highlight specific characteristics within a provided dataset. In case of PCA or EOF, the imposed condition is to find vectors with the smallest correlation [28]. This can be easily calculated for multivariate signals where the covariance matrix Ξ is estimated considering activities on all possible dimensions. For single variable time series, however, such approach is not possible. To overcome this problem, time series will often be extended with additional $N-1$ time-shifted duplicates, creating an artificial N-variate signal. Such preparation and appliance can also be referred to as *singular spectrum analysis* [29]. In this process, data matrix Ξ is of form

$$\Xi_{ij} = \frac{1}{T-|i-j|} \sum_{t=1}^{t=N-|i-j|} X(t)X(t-|i-j|), \tag{2.5}$$

where T denotes the length of time series $X(t)$. Decomposition of Ξ under PCA conditions will generate a set of orthogonal functions $c_i(\tau)$ such that

$$X(t) = \sum_{n=1}^{N} a_n c_n(t), \tag{2.6}$$

where the significance of each component is denoted by its scaling factor a_n. The orthogonality of components c_n is computed via eigenproblem decomposition, i.e. by finding eigenfunctions fulfilling

$$Ac_n = \lambda_n c_n, \tag{2.7}$$

where A and λ_n are the conditional operator and an eigenvalue, respectively. Produced components of this eigen-decomposition can be ordered depending on the value of eigenvalue. The larger the value, the stronger presence of the component in for of eigenvector.

The advantage of PCA is in its general approach. Extraction of components with the smallest correlation allows to identify and retrieve noise easily as it by definition should have little correlation with the measured data. Moreover, since for time series the covariance is based on a lagged signal, the PCA will also highlight components that have periodic structure and thus oscillatory features. However, difficulty arises when there are many phenomena within the system, and they are coupled. PCA is not able to distinguish these as they are highly correlated. Moreover, this method is heavily dependent on the definition and size of matrix operator A. It means that components will contain only those features that are specified by covariance and there is no guarantee that they are physically meaningful.

2.2 Time-frequency analysis

Time series can be considered as functions in the time domain. Such representation is convenient when one is interested in the changes of a system over time. However, in cases when it is known that some periodic behaviours are present, it can be more suitable to represent a signal in a *frequency domain*. Such representations emphasise how strong is the presence of periodic components in the signal of interest, although the definition of these elements depends on the applied method. In this section two representations are discussed: Fourier transform and spectrogram in Section 2.2.1 and wavelet transformation wit scalogram in Section 2.2.2. Discussion of their limitations is then presented in Section 2.2.3.

2.2.1 Fourier transform

Fourier transform (FT), named after French mathematician Joseph Fourier, is an important transformation used in data analysis. It provides reversible conversion between two reciprocal domains such as time and frequency. Given data as a time series $s(t)$,

it allows to represent them in frequency f domain via kernel integration

$$S(f) = \mathcal{F}(s)(f) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} s(t) e^{-2\pi i t f} dt, \tag{2.8}$$

where $\mathcal{F}$ denotes Fourier transform which kernel integration is over the whole time domain. This representation can be inverted back to the time domain by similar integration

$$s(t) = \mathcal{F}^{-1}(S)(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} S(f) e^{2\pi i t f} df, \tag{2.9}$$

which is over frequency domain. In both cases integration is performed using harmonic kernel $\mathcal{K} = \exp(i\omega t) = \cos(\omega t) + i \sin(\omega t)$, which also suggests that these are the canonical oscillations in FT. Acting on functions defined on whole domains, FT allows for lossless and reversible transformation. Since the integration is with a complex kernel, its result is also complex where the amplitude and phase describe sinusoid of a particular frequency. The function which is describing strengths of each component $|S(f)|^2$ is called Fourier spectrum, which due to FT popularity is often shortened to the *spectrum*.

One of the properties which make FT very useful is its linearity. Directly from definition 2.8 it can be seen that when applying to a superposition of two or more components the transformation $\mathcal{F}$ results in

$$\mathcal{F}\left(\alpha s_1(t) + \beta s_2(t)\right) = \alpha S_1(f) + \beta S_2(f), \tag{2.10}$$

where $S_1(f) = \mathcal{F}_1 s(t)$ and $S_2(f) = \mathcal{F}_2 s(t)$. This property allows performing inference about multiple sinusoidal oscillations present in a signal. Regardless of this fact, the spectrum of the superimposed signal may not be the same as a sum of both spectra separately. A simple case is when both signals contain an oscillation with the same amplitude and frequency but in opposite phases.

Whenever FT acts on time series $s(t)$, it produces a representation of harmonic oscillations with constant over time amplitude and phase for all frequencies that have ever occurred in the signal. This representation can be misleading as unless there is exact sinusoid in the signal any transient component will be falsely described as an

infinitely long oscillation. A practical improvement in localisation can be performed by analysing shorter windows of a signal, e.g. in a range $t \in [-T/2, T/2]$

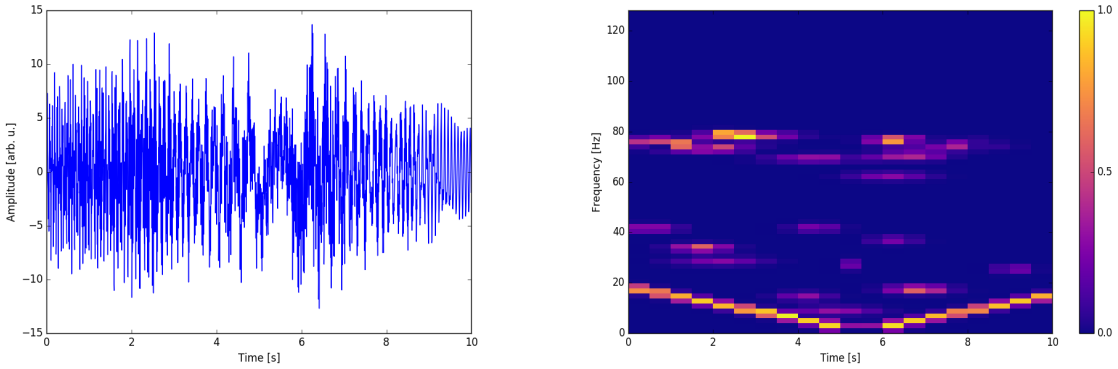
$$S(f) = \frac{1}{\sqrt{2\pi}} \int_{-T/2}^{T/2} s(t) e^{-2\pi i t f} dt, \tag{2.11}$$

which is often referred to as short-time Fourier transform (STFT). Such operation is the same as performing FT on a signal masked with window function $w_{(\tau,T)}(t)$

$$S(\tau, f) = \mathcal{F}(s_{(\tau,T)})(f) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} s(t) w_{(\tau,T)}(t) e^{-2\pi i t f} dt, \tag{2.12}$$

where $w_{(\tau,T)}(t) = 1$ for $t \in [\tau - T/2, \tau + T/2]$ and 0 otherwise. In general, parameters τ and T refer to window's position and its width, respectively. These properties and window's actual shape significantly modify the analysed signal, and thus they should be selected carefully. It can be shown that multiplying two functions in one domain, $g(t) = w(t) \cdot s(t)$, is equivalent to performing convolution on independently transformed signals in reciprocal domain, $G(f) = W(f) * S(f)$. In case of aforementioned rectangular function in the time domain, for the frequency domain, this is represented as a $\mathrm{sinc}(x)$ function, convolution with which can distort the *true* spectrum. Two examples of popular window functions are Hann and Hamming windows. Both of these are defined as a lifted cosine on finite support with the highest value being in the centre. It is a common approach to observe how frequency content changes in time by shifting window's position, sometimes even leading to overlaps. Visual representations of computed spectra for different window positions is called *spectrogram*. Figure 2.1 presents an arbitrary example of the spectrogram, where the x- and y-axis represent time and frequency domains, respectively, and intensity of colour indicates amplitude.

Most of the current data analysis is done using computers, which forces recorded signals to be discretised at certain intervals. Such processing imposes restrictions on what can be inferred. Just the process of sampling limits the maximum frequency possible to extract, i.e. *Nyquist frequency*. If a signal is sampled with constant frequency f_s then the fastest frequency possible to extract is $f_{\max} = f_s/2$ [Hz] or $\omega_{\max} = \pi f_s$ $[\frac{\mathrm{rad}}{\mathrm{s}}]$ [14]. Moreover, the minimum frequency is dictated by the length of the signal T to $f_{\min} = 1/T$ [Hz] or $\omega_{\min} = 2\pi/T$ $[\frac{\mathrm{rad}}{\mathrm{s}}]$. Having discrete input signal



(a) Arbitrary signal.

(b) Spectrogram of arbitrary signal.

Figure 2.1: Example of time-frequency representation. An arbitrary signal 2.1a of 10 s length was generated with sampling rate of 256 Hz. It is composed of a chirp function and 20 harmonic oscillations with Gaussian envelope. Its spectrogram 2.1b was computed using 1 s Hamming window with 0.5 s overlap. In both Figures x-axis refers to the time domain and y-axis is amplitude and frequency for 2.1a and 2.1b, respectively. Both amplitudes where scaled such that the maximum per graph is one and the scale is presented by colour intensity.

requires the use of discrete Fourier transform (DFT) which is in the form of

$$S[k] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} s[n] e^{-i\frac{2\pi n}{N}k}, \tag{2.13}$$

and the inverse DFT is defined as

$$s[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} S[k] e^{i\frac{2\pi k}{N}n}, \tag{2.14}$$

where for both cases N refers to the number of samples $N = T/dt$. With limits on the frequency range and the number of samples, one can see that the spectrum also has a limited number of points. Such constraint means that not all values of frequencies are possible to extract and thus one should be careful when applying DFT to time series. If these contain frequencies spaced closer than $\Delta f = f_{\min}$, then DFT spectrum will not be able to separate them. In case of the simulated signal, one can always increase resolution by increasing length of the signal. However, this is not possible if the signal was already recorded.

2.2.2 Wavelet transformation

In *Wavelet transformation* (WT) the *wavelet* refers to a small wave [30]. This notion is in contrast with sinusoids which are called the "big waves" due to their infinite length. The purpose of WT is to highlight transient oscillations by progressive localisation in time and frequency domains. It does that by using an adaptive time window which is stretched and compressed accordingly. For a function to be called wavelet $\psi(u)$ it has to have the following two properties:

1. The integral of $\psi(u)$ over the whole domain is zero:

$$\int_{-\infty}^{\infty} \psi(u)du = 0. \tag{2.15}$$

2. Squared absolute wavelet $|\psi(u)|^2$ integrates over the entire domain to a unity:

$$\int_{-\infty}^{\infty} |\psi(u)|^2 du = 1. \tag{2.16}$$

The first property emphasises the oscillatory behaviour of wavelets, imposing that there has to be the same amount of function above as it is below some reference point (zero). The square integrability, also called the finite support property, indicates that the function can be normalised. For this to happen, the wavelet has to reach zero value while going to both positive and negative infinities. This highlights the physical sense of these functions; recorded signals had to start and stop at specific points, or quickly converge to zero.

Unlike the Fourier transform which is strictly defined by sinusoids, the WF refers to a general class of functions. These can be described in both the real and complex spaces, although in practice when dealing with measurable signals only the real part of wavelets are chosen. Despite the vastness of wavelet class some functions are more popular and commonly used in the analysis [31]. Figure 2.2 presents few of these examples. The top left graph (Figure 2.2a) shows Haar [32] wavelet, i.e. a square function, which is the most straightforward wavelet used for analysis. Despite sharp edges, such form not only has compact support but also allows to identify and understand component's frequency easily. Another class of wavelets with compact support are

Daubechies wavelets [33], which are characterised by the maximum number of vanishing moments for given support width. Examples of these components are presented in Figures 2.2d and 2.2e, which have 2 (*db2*) and 6 (*db6*) vanishing moments, respectively, and Figure 2.2f shows more symmetrical version of *db6*, i.e. symlet 6. In the central and right positions of the top row are presented functions with infinite support, but quickly converging to zero. Both are defined using an exponential decline, which for Ricker wavelet (Fig. 2.2c), also known as *Mexican hat* wavelet, is

$$R_\omega(t) = (1 - \omega^2 t^2) \exp\left(-\frac{\omega^2 t^2}{4}\right), \tag{2.17}$$

where ω denotes a peak in Fourier spectrum. Morlet wavelet, however, is defined [34] as harmonic oscillation with Gaussian envelope

$$\Psi_\omega(t) = \frac{1}{\sqrt[4]{\pi}} \left(\exp(i\omega t) - \exp\left(-\frac{\omega^2}{2}\right)\right) \cdot \exp\left(-\frac{t^2}{2}\right), \tag{2.18}$$

where ω is the base frequency of modulation. The advantage of this wavelet is in its optimal volume of uncertain bound in time-frequency representation [35, 36]. Figure 2.2b shows real part of Morlet wavelet with the centre frequency $\omega = 5$, which is chosen as the smallest value such that the second term in Eq. (2.18) is negligible.

Given *wavelet* Ψ and time series $s(t)$ the *wavelet transformation* is defined as

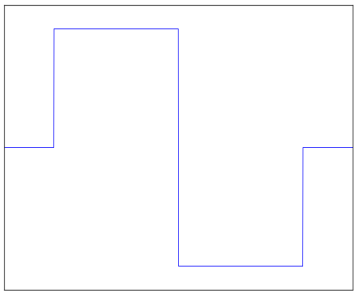
$$\mathcal{W}_\Psi(a, b) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{a}} \Psi^* \left(\frac{t - b}{a}\right) s(t) dt, \tag{2.19}$$

where Ψ^* denotes complex conjugate of Ψ and parameters $a > 0$ and b define *scale* and time *shift*, respectively. The transformation is commonly shorted with notation of *mother wavelet* Ψ and its *child wavelets* $\psi_{a,b}$ by defining

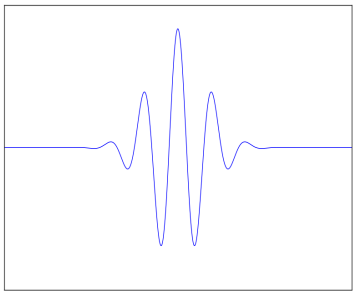
$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \Psi \left(\frac{t - b}{a}\right), \tag{2.20}$$

which allows representing WT in the form of a dot product

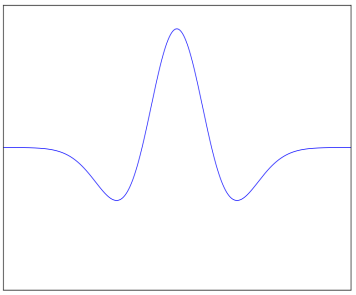
$$\mathcal{W}_\Psi(a, b) = \langle \psi_{a,b}(t), s(t) \rangle = \int_{-\infty}^{\infty} \psi^*_{a,b}(t) s(t) dt, \tag{2.21}$$



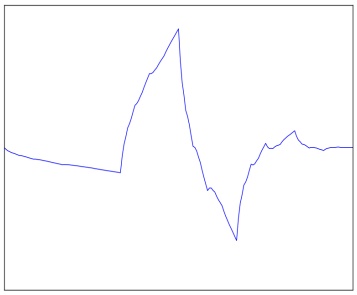
(a) Haar wavelet



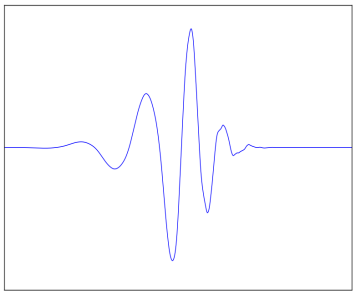
(b) Morlet wavelet



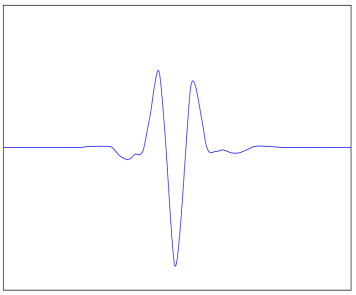
(c) Ricker wavelet



(d) Daubechies (db2) wavelet



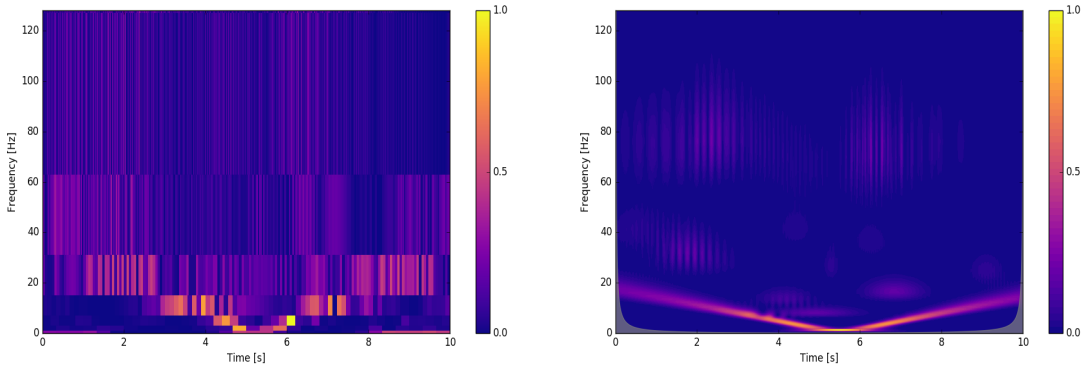
(e) Daubechies (db6) wavelet



(f) Symlet (sym6) wavelet

Figure 2.2: Gallery of selected popular wavelets. Each panel contains function as indicated by its label. For ease in shape comparison all functions have been scaled such that the maximum deflection have the same value 1.

where $\langle f, g \rangle$ is a dot product in Hilbert $\mathcal{L}^2$ space. Scale factor a refers to the width of wavelet which indicates how much of a signal a wavelet can explain. The smaller the scale value, the shorter segment is used to analyse, and the faster frequencies are associated. Selection of the appropriate set of scales is dependent on chosen mother wavelet. In case of orthogonal wavelets it is argued [37] that a dyadic dilation $a = 2^i$ with discrete shifts $b = j2^i \; i, j \in \mathbb{Z}$ are preferred. For other wavelets it is convenient [31] to use scales with fractional powers of two $a_i = a_0 2^{i\delta_i}, i = \{1..I\}$, where I determines the largest scale, $I = \log_2 \left(T/a_0\right)/\delta_i$. As it can be seen parameter δ_i determines the resolution in spectral space and its optimal value is wavelet specific, e.g. $\delta_i = 0.5$ is the largest value preferred for Morlet wavelet [31]. Limitations on shift b value are imposed by the edges of the finite-length input signal. Although it is possible to extend time series with zero on both ends, such augmentation impacts transformation values. The region where these edge effects are pronounced is called the *cone of influence* (COI) and is quantitatively defined as an area where the wavelet power decreases e^{-2}-fold due to imposed discontinuities.



(a) Discrete scalogram

(b) Continuous scalogram

Figure 2.3: Results of WT with discrete *sym6* (Fig. 2.3a) and continuous Morlet ($\omega_0 = 5$) (Fig. 2.3b) wavelets computed on the arbitrary signal (Fig. 2.1a). For ease of comparison with spectrogram both produced scalograms where transformed from scale a and shift b representations into the time-frequency domain. Both representations were scaled independently such that the maximum amplitude value is one, with colour progression according to included scales. All axes are linear. Greyed area visible in bottom corners of 2.3b indicates region outside of COI.

Similarly to spectrogram obtained using STFT one can display results of WT in terms of scale a and shift b. Although WT produces scale-shift representation often called *scalogram* in the thesis these representations are transformed into time and frequency values, allowing for more straightforward comparison with other time-frequency methods. The actual relation depends on used wavelet and its properties. In case of Morlet function, an association between frequency ω and scale factor a can be directly computed as

$$\omega = \omega_0 / a, \tag{2.22}$$

where ω_0 is the central frequency in Morlet's definition. Using the arbitrary signal introduced in FT section (Fig. 2.1a), two transformed scalograms are computed and presented. For discrete transformation (Fig. 2.3a) wavelet *sym6* was used, whereas continuous transformation (Fig. 2.3b) was calculated with Morlet function of centre frequency $\omega = 5$. Despite the different focus of these representations they both seem to highlight similar activity in time-frequency space. These are also consistent with Fourier spectrogram presented in Fig. 2.1.

In practice, a type of wavelet is usually chosen depending on features that one

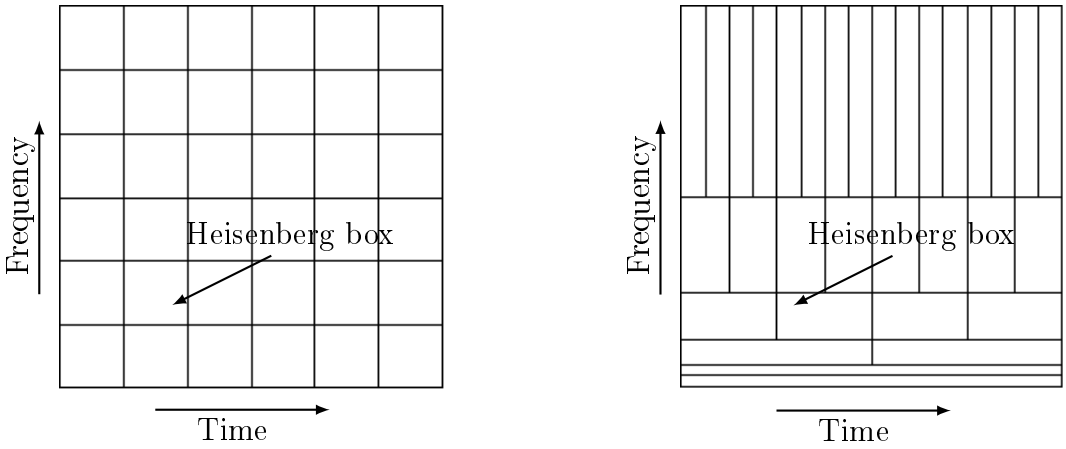
intends to extract. As it can be observed in Fig. 2.2, there is a variety of shapes which will highlight different characteristics. For example, when dealing with signals of brain origin, a popular choice is the Morlet wavelet [38–41].

2.2.3 Limitations

This section introduced two methods for providing time-frequency representation. Fourier spectrogram and wavelet scalogram allow one to observe the content of time series in the form of components with specific periods or frequencies. Although these methods provide insightful representation close to what is expected from the signal, they also suffer from some limitations.

As it has been already highlighted, one of the limitations it is the abstract definition of component's form and the oscillation that it represents. In case of FT, the type of oscillation is already defined as a sinusoid or segmented sinusoid in case of STFT. Such representation is dictated by the formula to preserve specific mathematical properties. However, it is unlikely that the signal itself would contain components of such nature. In case of WT the structure of components, and thus the definition of oscillation, is more flexible. This allows constructing oscillations that better reflect possible physical processes undergoing within the time series. However, this also introduces selection bias forcing the signal to be projected onto observer's assumptions. Under such conditions, it is possible to falsely interpret the presence of an event with assumed structure, despite that its true form might be different and not singular. Even if a signal is composed of noise both TF representations would highlight some activity. Furthermore, both spectrogram and scalogram assume a single form of a periodic component within a signal; they reject the possibility that event might be a mix of few distinct forms of oscillations.

Another constraint that needs to be acknowledged is a finite resolution in both time and frequency domains. This limitation is often called an *uncertainty principle* [14]. It refers to a situation when two properties are so tightly bound, that making more precise determination of one parameter increases the uncertainty of the other one. In other terms, the product of variances for both time Δt and frequency $\Delta\omega$ resolutions cannot be smaller than certain value σ, $\Delta t \cdot \Delta\omega \geq \sigma$. This relation manifests



(a) Fourier spectrogram.

(b) Wavelet spectrogram.

Figure 2.4: Time-frequency plot for Fourier 2.4a and Wavelet 2.4b Transformation. Both in Fourier spectrogram and wavelet scalogram each box has the same area.

itself as a grid of available values — so-called *Heisenberg boxes* — in time-frequency representation. Such boxes can be seen in Figure 2.4 which shows Fourier spectrogram (2.4a) and wavelet scalogram (2.4b). Any expansions or contractions to box's width Δt or height $\Delta\omega$ will cause a counter transformation in the reciprocal domain keeping the volume constant. In order to obtain better resolution in frequency, one needs to gather more extended signal, which prevents from localising event in the time domain. Hence, it is impossible to determine the exact frequency value present at a particular time, and one needs to balance these properties depending on the expected observations.

Presented time-frequency representations also can mislead with an assumption about causality in dynamics. One might expect that observing high activity at a particular time and frequency ranges preceded by similar high activity at previous time block could be its continuation. Although such extension is possible, it is not necessarily true. Both in Fourier spectrum and wavelet scalogram each Heisenberg box is computed independently from others, and thus a sudden change in signal might occur. To cover this sudden change often significant overlaps in segments are taken and appropriate adjustments of scale and shift parameters. This, however, smears TF representation and, depending on the approach, may not be able to detect immediate variation.

Concluding, FT and WT can provide insightful information about frequency content of time series. However, there are limitations to what such time-frequency repre-

sentation can describe. These representations are susceptible to researcher's bias and methods' limitations. Although being insightful, they should be used with precautions.

2.3 Empirical mode decomposition

Previous sections discussed methods for extracting components of general form (Sec. 2.1) and highlighting how frequency content of time series changes in time (Sec. 2.2). Given that many systems in nature have oscillatory characteristics, it is essential to be able to extract these intrinsic periodical components. This process requires both identifying data specific oscillations and highlighting their dynamics. Such endeavour was taken by Huang et al. [17] who proposed the *empirical mode decomposition* (EMD) which is a method for non-stationary and non-linear processes. The EMD aims to decompose data into general oscillatory functions, i.e. amplitude- and frequency-modulated components. The decomposition is dependent only on the shape of the input signal, thus making the method purely data-driven. Resulting components are considered to represent intrinsic oscillations within the system, and for this reason, they are called *intrinsic mode functions* (IMF).

The popularity of EMD has been steadily increasing since the introduction. Most of the attention has been focused on application to geophysical signals, especially to analyse wind and earthquake data. As it has been shown in [42], it is possible with EMD to capture different time-scale patterns embedded in the data. In their findings, each IMF corresponds to events with different frequency, e.g. daily, weekly or monthly. Similarly, patterns related to physical events were discovered in other studies, e.g. daily river flows [43] or natural wind phenomena as a diurnal cycle, frontal passages or baroclinic instability [44–46].

Many of studies are also performed in biomedical science. Some authors [47] have shown synchrony between the cardiac and respiration signals. Some research, however, has been done concentrating on brain waves, that is the electroencephalogram (EEG) recordings. Researchers have studied the phase synchronisations between different sensors [48, 49] or detecting significant features in EEG [50, 51]. EMD was also discovered to be a valuable tool for removing eye movements signal from the EEG [20, 52–54].

Moreover, EMD has been successfully applied in many scientific and engineering fields, e.g. [55–57], producing insightful results. Such wide popularity and general accomplishments necessitate investigation and understanding of the method.

The emphasis of this section is on introducing EMD. Presented here content is structured as follows. In the first Subsection 2.3.1, the EMD original algorithm is presented, and the following Subsection 2.3.3 discusses variations proposed by other researchers. Final Subsection 2.3.2 highlights and discusses some properties of the EMD and its components.

2.3.1 The algorithm

The original EMD [17] has proposed algorithmically, without any underlining mathematical framework. The method acts on time series $S(t)$ producing N intrinsic mode functions (IMFs) which forms depend strictly on signal's shape, i.e. change in amplitude of time series. The algorithm for EMD can be described in few steps as follows:

1. Store the input signal $S(t)$ as the initial iteration $s_0(t) := S(t)$.

2. Identify all local extrema (both minima and maxima) in time series $s_j(t)$. Extrema are defined by locations where the derivative $\frac{ds_j(t)}{dt} = 0$.

3. If the number of extrema is less or equal than 2 then $s_i(t)$ is considered to be a trend — a low frequency modulation — and the algorithm stops ($R(t) = s_j(t)$).

4. Estimate top $e_{\max}$ and bottom $e_{\min}$ envelopes of $s_j(t)$ by interpolating respectively local maxima and local minima with natural cubic splines (Figure 2.5).

5. Calculate instantaneous mean $m(t)$ of both envelopes, $m(t) = \frac{1}{2}\left(e_{\max}(t) + e_{\min}(t)\right)$.

6. Subtract the mean from focused time series $h_j(t) = s_j(t) - m_j(t)$.

7. If $h_j(t)$ fulfils the stopping criteria, then it is considered an intrinsic mode function (IMF) (a component $c(t)$) and the procedure is repeated for a modified signal $S(t) := S(t) - c(t)$ from the first point. Otherwise, the algorithm starts from second step with $s_{j+1}(t) := h_j(t)$.

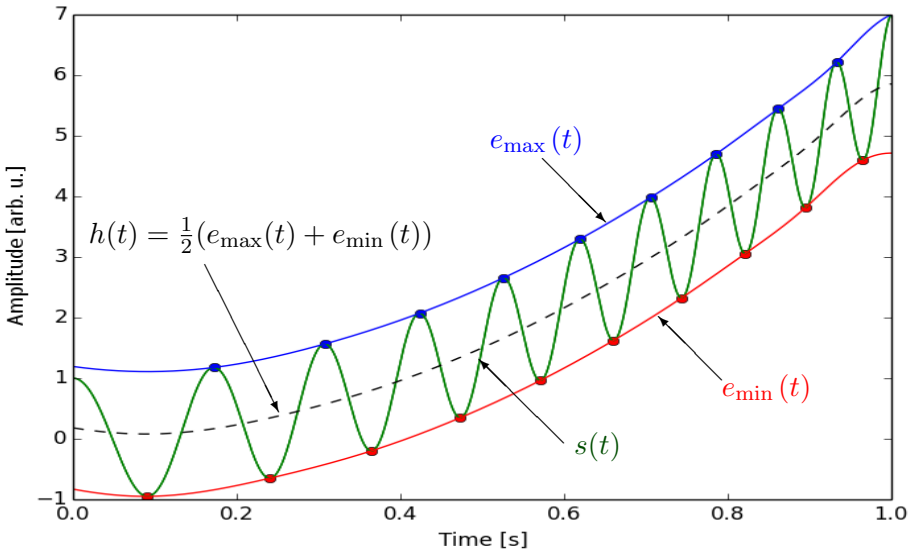


Figure 2.5: Signal's features identification in EMD sifting process. Top ($e_{\max}$) and bottom ($e_{\min}$) envelopes of an input signal $s(t)$ composed of slowly changing trend and a riding wave. Averaging both envelopes creates local mean $h(t)$ (dashed line).

The critical element of the algorithm is to calculate local mean and subtract it from the signal. These steps are often referred to as *sifting process* [17]. As a result of EMD one obtains a decomposition in the form of

$$S(t) = \sum_{n=1}^{N} c_n(t) + R(t), \tag{2.23}$$

where c_n are IMFs and R is a residue, which is a slowly varying trend. The number of components N is finite and dependent on the input data's complexity. It has been shown that for Gaussian noise EMD behaves as a dyadic filter bank [58] producing on average $N \approx \log_2(f_s)$, where f_s is the sampling frequency.

The stopping criterion introduced in the original paper [17] tests IMF's time series convergence in consecutive iterations. It states that sifting process is convergent if the Cauchy's standard deviation, SD, is smaller than a predefined threshold value, σ,

$$SD = \sum_{t=0}^{T} \left[\frac{|h_k(t) - h_{k-1}(t)|^2}{h_{k-1}^2(t)} \right] < \sigma, \tag{2.24}$$

where the threshold value σ is suggested to be between 0.2 and 0.3 [17]. The justification

given by the authors is based on Fourier spectra. When comparing two spectra, one of which has 5 data points (out of 1024) shifted, their SD value will be in a range of 0.2–0.3 (calculated point-by-point).

2.3.2 Properties

As it has been mentioned before, EMD decomposes a signal into a set of oscillatory functions called intrinsic mode functions (IMFs). In the original paper, the authors have defined IMF as a function which fulfils two conditions:

1. In a given domain, the number of extrema and zero-crossings must differ at most by one, and

2. the local mean spanned by the average of top and bottom envelopes should be zero everywhere.

These conditions are meant to guarantee the oscillatory behaviour of IMF. The necessity to cross the zero-value between consecutive extrema refers to the narrowband requirement for a stationary Gaussian process [17]. The second condition refers to the symmetrical shape of a component. Such property is desired when performing Hilbert transform (see Appendix A.2) on the signal as it allows for its extraction. Authors deliberately imposed such condition to extract meaningful instantaneous frequency.

Figure 2.6 presents two oscillatory functions out of which only the top function fulfils IMF's properties. Despite being periodic with a period of 1 s, the bottom function is not an IMF. It possesses two frequencies: slow wave (2 Hz) and a fast wave (22 Hz). Often the fast component in such combination is referred as a *riding wave*. Top function, however, is a single oscillation in a general form, i.e. having modulations in both amplitude and frequency. Although in presented example modulations have apparent structures, in general, they can be more complicated.

An exemplary set of IMFs obtained via EMD is presented in Figure 2.7. The top graph (red colour) displays the input signal, which is a normalised Gaussian noise filtered with a moving average window of 5 samples. Consecutive rows contain 5 first IMFs starting with the first on top. The order of IMFs also corresponds to the order of instantaneous frequencies averaged over time, i.e. average frequencies.

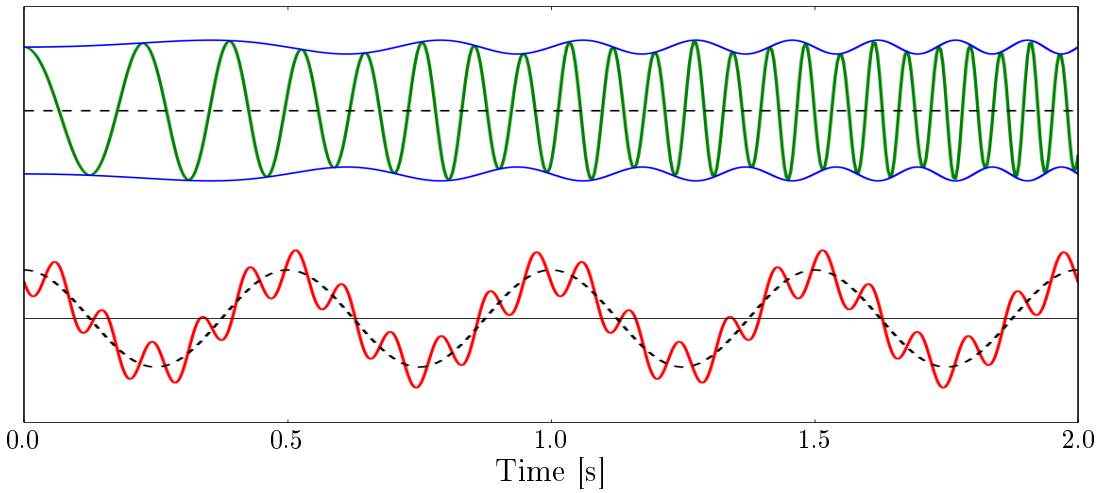


Figure 2.6: Examples of oscillatory functions. Only the top oscillation, presented in green with blue envelopes, fulfils IMF conditions. The bottom function described in red colour is composed of slowly oscillating trend (dashed) with much faster wave.

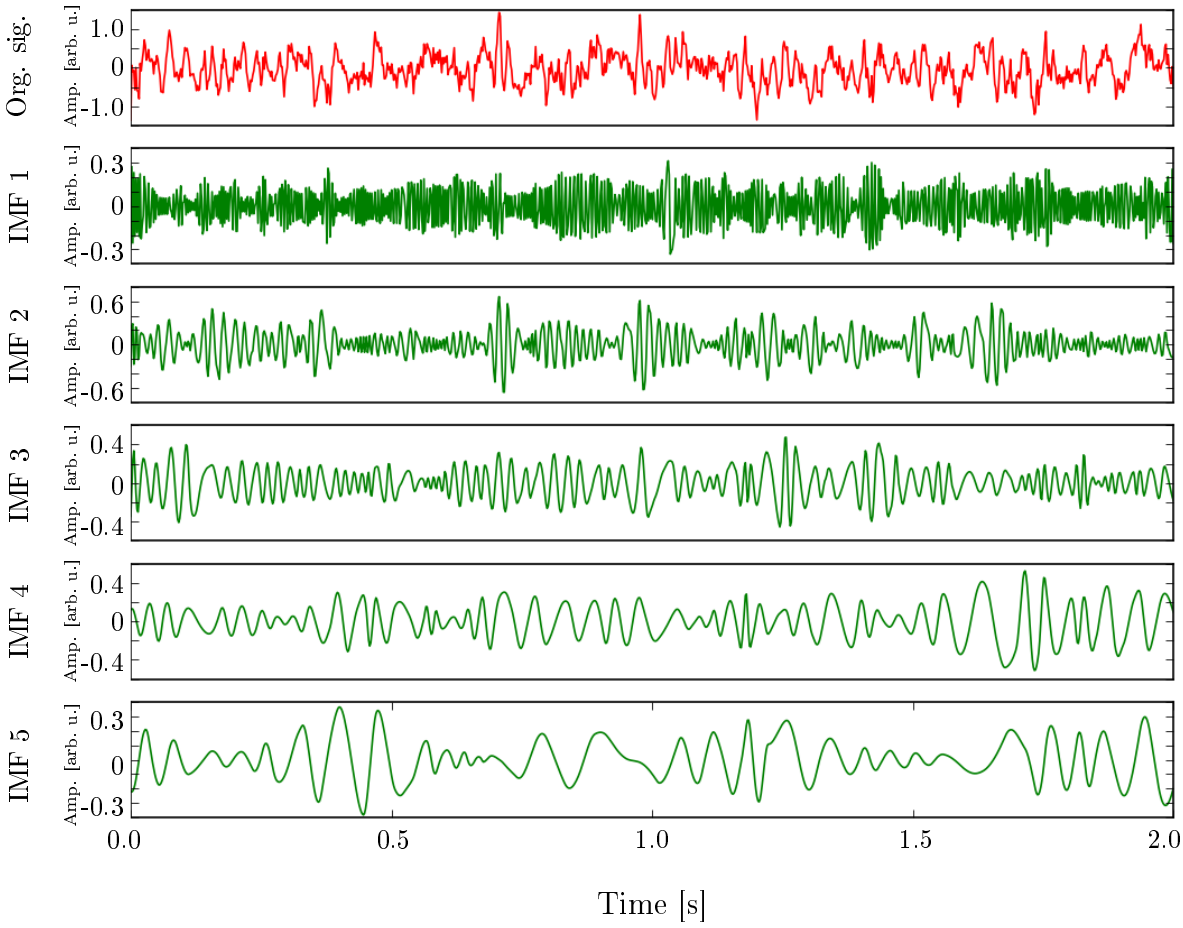


Figure 2.7: An example of a set of IMFs obtained using EMD. An input signal (top graph, red) is a moving-average filtered Gaussian noise. The following graphs (green) represent first 5 obtained IMFs.

Common practice in signal processing community is to apply the Hilbert transform to extract amplitude and phase components from the signal. It has been proven, that such transformation will result in a complete decomposition when the signal fulfils the Bedrosian theorem (see Appendix A.3). Unfortunately, without the mathematical framework, it is impossible to conclude decisively whether IMFs fulfil theorem's conditions. Some researchers [21, 59, 60] have started developing framework centred around Bedrosian theorem. Such approach allows for construction of components with easily extractable amplitude and phase features.

2.3.3 Extensions

The lack of formal mathematical definition was a great stimulus for developing EMD extensions. These either come as a variation on the algorithm itself or from assuming some mathematical framework. Overall, one could consider dividing most of the suggested modifications into few groups.

- **Technique modifications**. These modifications relate to changing at least one of the steps used within the original EMD algorithm. An example of the most common modification is proposing a different spline interpolation technique, e.g. in [61, 62]. Other modifications can relate to interpolation of extrema positions and values. Since the execution of the EMD is only possible on computers, which can only store discrete signals, only available data points should be considered. However, the EMD uses continuous signals in its definition, and the discretisation is only a result of framework limitations. This indicates challenges in both determining extrema's related values, e.g. estimating through cubic interpolation [63], and determining envelopes values at the edges of the signal of interest [64].

- **Procedural modification**. In contrast to the first instance, these modifications relate to particular elements of the EMD procedure. Most commonly they suggest modifying stopping criterion related to the sifting procedure either proposing a different convergence metric [58, 65] or advocating for a replacement step [66]. Some researchers have proposed modifying sifting procedure by either subtracting scaled or masked [67, 68] local mean. Moreover, inclusive to this group are

modifications that allow merging different parts of the analysed signal, e.g. using a sliding window technique [69, 70].

- **Method variations**. Changes in this group have the biggest effect on the outcome of the method. Often they modify significant part of the algorithm and include additional steps. The most popular variation is Ensemble Empirical Mode Decomposition (EEMD) [71], which creates an ensemble of noise-added signals and performs EMD on each one of them. The expected result is the grand average of all ensembles. Other modifications include defining EMD method in multi-dimensional problems [72–76]. Similarly, some variations are only loosely related to EMD and they try to explain its behaviours through other approaches, e.g. filtering [68], diffusion [77], or infinitesimal local mean [78].

As previously mentioned, the list of EMD applications and achievements suggests that the method has been recognised to provide meaningful results. Nevertheless, the number of proposed variations also suggests that there is capacity for improvements. For this to happen the method would require a mathematical framework; otherwise, any advancements cannot be objectively compared and stated what the benefit is. The popularity of the EMD and drawbacks mentioned above indicate there is a lack of data analysis method which has properties similar to EMD and a well-defined mathematical foundation.

Chapter 3

Model-based analysis

Systems can be described and analysed from different perspectives. Along the data analysis approach presented in Chapter 2 one can focus on constructing the underlying model behind the system. This is called *model-based analysis* and is described in this chapter. The main difference between mentioned approaches, is that with modelling one tries to initiate discussion by presenting a mathematical model of the system. One starts with principles in order to analyse system's properties and then validate the hypothesis with observations.

By constructing models, one not only can explain the underlying behaviour, but also can analyse the system theoretically. This allows to forecast and simulate data, instead of performing potentially time consuming experiments. In fact, any physical phenomena can be modelled. In case of oscillations two specific approaches are very important: Sturm-Liouville theory and Kuramoto model for coupling in phase. Both describe dynamics of the system via differential equations, i.e. in terms of system's function and its changes. The general form of a ordinary differential equation (ODE) of function $y(x)$ can be written as

$$\sum_{n=0}^{N} a_n(x) \frac{d^n y(x)}{dx^n} = q(x), \tag{3.1}$$

where $q(x)$ and $a_n(x)$ are coefficient functions of an independent variable, x, and the summation limit, N, determines the order of the system. As it can be seen ODE in equation (3.1) are linear in respect to $y(x)$ and its derivatives.

The Sturm-Liouville theory describes general behaviour of oscillators in terms of 2^{nd} order ODE. It states how a single oscillator will behave given its initial conditions and some function of surrounding. The latter model, i.e. Kuramoto model, refers to interactions between a number of oscillators. It emphasises that when objects are communicating, their properties may change. It focuses on these changes in phases, i.e. delaying and extending periodic behaviour.

3.1 Sturm-Liouville eigenvalue problem

The general form of the second order ODE can be written as

$$a_2(x)\frac{d^2y}{dx^2}(x) + a_1(x)\frac{dy(x)}{dx} + a_0(x)y(x) = q(x), \tag{3.2}$$

where the notation is as for the Equation (3.1). A particular form of this equation was named after *Jacques Charles François Sturm* [79] and *Joseph Liouville* [80]. Their research focused on ODEs of form

$$-\frac{d}{dx}\left(p(x)\frac{dy(x)}{dx}\right) + q(x)y(x) = \lambda w(x)y(x), \tag{3.3}$$

where coefficient functions $p(x)$, $q(x)$ and derivative $p'(x)$ belong to Hilbert space $\mathcal{L}^2$. Equation (3.3) is, in fact, an eigenvalue problem with λ and $w(x)$ being an eigenvalue and weighting function, respectively. These relations are more apparent once the left-hand side is presented as a linear operator L acting on function y, i.e. $Ly(x) = \lambda y(x)$, where

$$L = \frac{1}{w(x)}\left[-\frac{d}{dx}\left(p(x)\frac{d}{dx}\right) + q(x)\right]. \tag{3.4}$$

Often, in practice, a simplified form of the equation is used [80]. Changing variables from x to ξ such that

$$\xi(x) = \int_{x_0}^{x} \frac{d\tilde{x}}{p(\tilde{x})}, \tag{3.5}$$

allows to define the first derivative as

$$\frac{d}{dx} = \frac{d\xi}{dx}\frac{d}{d\xi} = \frac{1}{p(x)}\frac{d}{d\xi}, \tag{3.6}$$

and the second as

$$\frac{d^2}{d^2x} = \frac{1}{p^2(x)}\frac{d^2}{d\xi^2} - \frac{p'(x)}{p(x)}\frac{d}{d\xi}. \tag{3.7}$$

This conversion results in an operator with changed variable ξ as

$$\tilde{L} = \frac{1}{\tilde{w}(\xi)}\left(\frac{d^2}{d\xi^2} + \tilde{q}(\xi)\right). \tag{3.8}$$

Replacing dependent variable again by x and omitting weighting function $w(x)$, the *Sturm-Liouville* (SL) equation can be written as

$$y''(x) + Q(x)y(x) = \lambda y(x). \tag{3.9}$$

This form is commonly used in literature. It is especially popular when analysing inverse Sturm-Liouville problem [80].

The popularity of SL equations comes from their wide appearance when analysing physical problems. For reasons that are explained later, these equations describe wave-like behaving functions often met in nature. For example, wave propagation in materials like strings or drums can be postulated in SL forms [6]. An equation of the form (3.9) is especially prevalent in quantum mechanics as it represents Schrödinger equation [81], which describes the movement of particles.

Properties and solutions for some particular coefficient functions, i.e. $Q(x)$ or $\{p(x), q(x)\}$, have already been intensively studied. Two of the most popular equations are Bessel type with ($p(x) = -x$, $q(x) = (x^2 - \nu^2)$) and Airy ($p(x) = -1$, $q(x) = -x$) [82]. Solutions to such equations are special functions, which often appear in quantum mechanics [35]. Another set of equations which have been thoroughly studies is the set with a constant coefficient, $Q(x) = 0$. In such case, often called Fourier type, equation (3.9) is transformed into

$$y''(x) = \lambda y(x), \tag{3.10}$$

which for $\lambda < 0$ is solved by sinusoids with period of $T = 2\pi/\sqrt{|\lambda|}$, i.e. $y(x) = A\sin(\sqrt{|\lambda|}x + \phi)$.

An SL problem which has separated boundary conditions on a finite interval $[0,1]$[1] of form

$$\begin{cases} \alpha_0 y(0) + \beta_0 y'(0) &= 0 \qquad\qquad (\alpha_0^2 + \beta_0^2 > 0), \\ \alpha_1 y(1) + \beta_1 y'(1) &= 0 \qquad\qquad (\alpha_1^2 + \beta_1^2 > 0), \end{cases} \tag{3.11}$$

and $p(x), w(x) > 0$ is said to be *regular*. The meaning and properties of the regular Sturm-Liouville problem are described in the Theorem 1.

Theorem 1. *The regular Sturm-Liouville problem has a countable number of discrete and simple eigenvalues λ_n, which accumulate only at ∞. The corresponding normalised eigenfunction $\tilde{y}_n$ can be chosen real-valued and form an orthonormal basis for the considered interval I, i.e.*

$$f(x) = \sum_{n=0}^{N} \langle \tilde{y}_n, f \rangle \, \tilde{y}_n(x), \tag{3.12}$$

for all $f \in \mathcal{L}^2$. Moreover, if f belongs to a linear subspace of Hilbert space $\mathcal{L}^2$, then the series is uniformly convergent.

It can be shown [82] that the operator L is a self-adjoint operator. The meaning of this property is that if there exists a non-empty set of solutions all their corresponding eigenvalues are real. Moreover, respective eigenfunctions span basis in the Hilbert's space $\mathcal{L}^2_{\mathcal{I}}$ on a segment $\mathcal{I}$ with the inner product defined as

$$\langle f, g \rangle_w = \int_{\mathcal{I}} f^*(x) g(x) w(x) dx. \tag{3.13}$$

In the equation 3.13 asterisk * denotes complex conjugation and functions $f, g, w \in \mathcal{L}^2_{\mathcal{I}}$. Another essential property of solutions to the SL problem is that their eigenfunctions can be sorted by the increasing number of roots. This is stated in Theorems 2 and 3 [83], where the notation of an eigenfunction $y(x, \lambda)$ with an eigenvalue λ was used. The first Theorem 2 states that for the same SL eigenproblem, the eigenfunction related to a more prominent eigenvalue has a higher frequency of zero-crossings. Oscillation Theorem 3, however, refers to the number of possible solutions below a certain value of eigenvalue. It implies that the n$^{\text{th}}$ eigenfunction has exactly n zero-crossings. Their combined results are stated in another Theorem 4.

[1]The interval is general, because one can always apply a linear mapping $T : [0,1] \to [a,b]$.

Theorem 2 (Sturm comparison theorem). *For $j = 1, 2$, let y_j be eigenfunction of L_j with eigenvalue λ_j. Suppose $0 \leq a < b \leq 1$, $y_1(a) = y_1(b) = 0$ and $\lambda_1 < \lambda_2$. Then y_2 has a zero in (a, b).*

Theorem 3 (Oscillation theorem). *The number of eigenvalues of (3.9) strictly below λ is equal to the number of zero values $y = 0$ of $y(x, \lambda)$ in $(0, 1)$.*

Theorem 4. *Let $\lambda_0 < \lambda_1 < \ldots$ be the eigenvalues of L (3.9) in $\mathcal{L}^2_{[0,1]}$ with boundary conditions $y(0) = y(1) = 0$. Then $y(x, \lambda_n)$ has exactly n zeros in $(0, 1)$.*

The main implication of these theorems is the possibility to create a set of orthogonal functions with oscillatory properties. Each eigenfunction has a number of zero-crossings related to its ordinality. Thus with increasing order, the number of zero-crossings also increases allowing to relate that property to the frequency of solution and uniquely decompose functions in such constructed basis. As the equation (3.12) implies, scaling for each fundamental component, i.e. eigenfunction, can be obtained by projecting the data onto the respective component. An example of such decomposition is Fourier series, which is obtained while solving the SL eigenproblem under a constant coefficient (3.10). In such case eigenvalues have straightforward interpretation as they correspond to the frequency of individual sinusoids.

Sturm-Liouville theory has also been connected to the structure of IMFs (Section 2.3). Vatchev & Sharpley in their articles [84, 85] have related conditions for IMFs to the oscillation theorem. Their modified definition of IMF, which they called a *weak IMF*, is postulated concerning a solution to the SL problem with $\{p(x), q(x)\}$ such that

$$f(x) = \frac{1}{q(x)} h'(x), \qquad h(x) = -q(x) f'(x), \tag{3.14}$$

where $f(t)$ is the weak IMF and $h(x)$ is an associated function from C^2_I. The obtained eigenfunction f is an oscillation, i.e. the numbers of extrema and zero-crossings differ at most by one.

3.2 Interacting oscillators and the Kuramoto model

The oscillation theory presented in the previous section refers to a single oscillator. In nature, however, it is difficult to observe an isolated object, i.e. one which does not interact with its surroundings. An object can be either influenced by a dynamically changing environment or interact with other components creating a network or a system. There are many examples where such interactions between objects have been observed [86–92]. The abundance of these problems resulted in plenty of studies in this field [10, 93, 94]. Interactions in an obvious manner affect each component and make the whole system behave differently than a simple superposition of all components. Although the behaviour of the whole system depends on its configuration, some systems will exhibit the shared phenomena. Commonly described physical effects are the synchronisation and the oscillation death. The former can be commonly observed in nature and thus has been widely studied [95]. The synchronisation is defined as an adjustment of rhythms (oscillation patterns) in oscillating objects through their weak interactions [95]. In contrast, the oscillation death describes the process of mutual extinction of either amplitude or phase.

A common distinction of interactions is to classify them as either structural or functional. The structural interaction indicates the existence of a direct connection between the components. Some components, however, can influence each other despite not having a visible connection. Components that exhibit statistically significant interactions, e.g. are correlated, are considered to be functionally connected. However, often interaction is more complex than presented connections. Some components can have greater impact on others, and the coupling does not need to be symmetrical. In extreme cases, there can be components that influence others, but they are immune to changes within the network. For these reasons, networks often distinguish directionality and strengths between nodes. In case of the brain, the structural connectivity would be described by the neuronal connections between regions whereas the functionality is through observing correlated activity within the brain. When describing a system with analytical methods, it is easier to differentiate three types of connectivity, i.e. by adding the effective connectivity [92] to the structural and functional connectivities. The effective connectivity provides more in-depth qualitative knowledge about

the system such as coupling strengths and their directions. These, in turn, provide more knowledge about the information flow in the system. Understanding the dynamics of the network can provide information on what are specific contributions of each node and which regions have the most significant impact on the network. Furthermore, understanding the dynamics allows building appropriate models, through which it might be possible to understand the general mechanisms of the physical entity, such as the brain. For these reasons, there have been many attempts on creating methods to identify couplings.

Solving a general dynamical system of coupled oscillators is difficult; physical oscillators will interact on all dimensions simultaneously. With an increase of dimensionality/complexity, there are more variables to solve for in a model of these interactions. As with any solving approach, one needs to divide the problem into smaller parts. With coupled oscillators, it is common to consider the focus on either *strong* or *weak* interactions. The difference is in coupling regime and its effect on the network. Although the boundary depends on the specific system, some properties can be generalised.

The strong interaction is considered when the coupling affects the whole network, especially oscillators' amplitude [94]. The research on the coupled systems with strong coupling is ongoing with many interesting open questions such as synchronisation or oscillation death. The analysis of the synchronisation has been significantly moved forward by the introduction of the master-stability function (MSF) [95, 96]. The class of strong interactions is difficult to analyse as it affects many states creating a large problem space. To ease the analysis a common step is to reduce the dimensionality of the problem or consider the system under particular conditions.

In the weak interactions regime, the coupling does not affect oscillators amplitude. Even though it is a smaller class than the strong coupling, it is nevertheless a substantial class with many applications [10, 18, 92, 97]. Such reduction allowed for a significant research influx with valuable results. The vastness of the problem space is typically approached by assuming certain conditions of the system such as frequency distributions or specific types of the coupling function. In special cases, it is possible to find constraints on global solutions or provide approximate local solutions with loosen constraints.

Theoretical attempts to model these modulations have started with Winfree who proposed a generic model for dynamics in the phase space. Movements in this space describe any period activity where the phase denotes current position in reference to the period. The beginning of an event is at phase $\phi = 0$ and similarly, the phase $\phi = \pi$ (or 180°) indicates the median position. The first systems [8, 10, 98] which were designed to consider phase dynamics of an oscillator were proposed in terms of

$$\dot{\phi}_i = \omega_i + \left(\sum_{j=1}^{N} X_{ij}(\phi_j) \right) Z(\phi_i), \tag{3.15}$$

where the $X(\phi_i)$ is the phase-dependent influence on others oscillators, and the $Z(\phi)$ denotes the sensitivity function. Such form indicates that oscillators are expected to interact indirectly, through an interaction with the environment. Another change to the model incorporated the potential direct coupling between oscillators

$$\dot{\phi}_i = \omega_i + \sum_{j=1}^{N} \Gamma_{ij} \left(\phi_j - \phi_i \right), \tag{3.16}$$

where Γ is a general function of a difference between each pair of coupled oscillators. Such definition, however, was too general for practical use. Although it could be applied to the majority of the real systems, most of them would have too complicated interaction, making them not solvable analytically. A simpler, yet still general approach to strictly oscillatory components [9, 89, 99, 100] is to consider a system in the form of

$$\dot{\phi}_i = \omega_i + \frac{K}{N} \sum_{j=1}^{N} h(\phi_j - \phi_i), \tag{3.17}$$

where the coupling function $h(\cdot)$ is 2π periodic and is the same for all components. A specific version of this model was proposed by Kuramoto in 1975 [101]. He suggested to use $h(\cdot)$ function in the form of a scaled sinusoid, i.e.

$$\dot{\phi}_i = \omega_i + \frac{1}{N} \sum_{j=1}^{N} K \sin \left(\phi_j - \phi_i \right), \tag{3.18}$$

where ϕ_i is a phase of i^{th} oscillator and ω_i is its natural frequency often called *intrinsic*

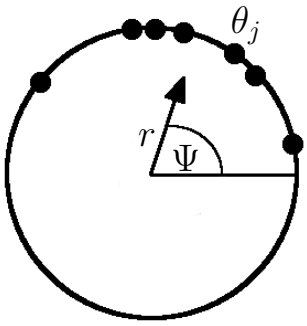


Figure 3.1: Geometric representation of oscillators' phases θ_j as points on a circle. Mean of all these vectors describes the mean-field vector of a length r and phase Ψ. (Source: [10])

frequency. This equation means that phase dynamics are dictated by a phase shift relative to another oscillator's phase. These modulations determine the oscillations instantaneous frequency ω. Commonly in the analysis of the model, it is assumed that ω is drawn from a distribution $g(\omega)$, which is typically a unimodal distribution, symmetric around some value Ω. In this context, Ω can be thought of as an overall average frequency. Due to rotational symmetry made by periodic coupling function in the model subtracting Ω from all ω_i does not change the overall dynamic. Such transformation allows for $g(\omega_i)$ to be substituted with $\hat{g}(\omega_i) = g(\omega_i - \Omega)$. Additionally, coefficient K in formula (3.18) denotes coupling strength, as it scales the impact of the interaction between oscillators.

The advantage of Kuramoto's model is that it can be solved analytically for a large number of oscillators. The solution is performed by averaging oscillators or, conceptually equivalent, by analysing them using a moving reference frame, i.e.

$$r \sin(\psi - \phi) = \frac{1}{N} \sum_{j=1}^{N} \sin(\phi_j - \phi_i), \tag{3.19}$$

where r and ψ are the amplitude and the phase of the frame, respectively. Geometric representation of such situation is represented in Figure 3.1. The phase of each oscillator is represented as a dot on a circle, and a vector r gives their average position at angle Ψ. Moving frame transformation allows writing the formula (3.18) in the form of

$$\dot{\phi}_i = \omega_i + rK \sin(\psi - \phi_i), \tag{3.20}$$

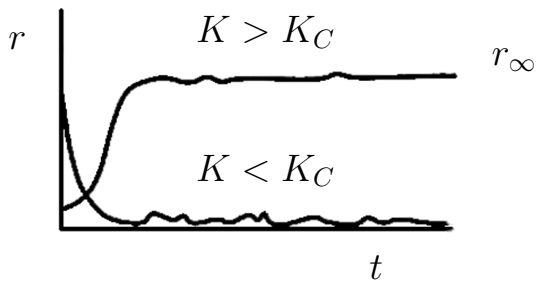


Figure 3.2: A typical dynamic of coherence r value depending on initial coupling K value in reference to critical coupling value K_C. (Source: [10])

which highlights that the phase of each oscillator is pulled towards the mean-field phase. Moreover, the positive feedback loop means that the more oscillators are coupled (i.e. big Kr), the more influential they will be on the other oscillators, thus making the coefficient Kr even bigger. Eventually, depending on the initial value of K, all oscillators will either be coupled entirely or incoherent, with phases uniformly distributed in the domain. The divergence can be seen in Figure 3.2, which represents changes of coherence value r over time-depending on the initial value of K. Typically a set of coupled oscillators can be divided into two groups: completely synchronised and partially synchronised. The first group tends to follow the global trend and thus stays in relatively equal distances. The partially synchronised group, however, is less rigid and allows for irregular migration in and out of the group.

Discussed behaviour refers to an ideal situation without any external influence on the system. In a more realistic scenario, however, the system can be affected by some non-oscillatory input from the environment. The perturbation can introduce time-dependent modulations into natural frequencies and couplings [102]. For simplicity, such a disturbance is often only considered as a non-stationary influence on the natural frequency leading to a modified Kuramoto equation

$$\dot{\phi}_i = \omega_i + \xi_i(t) + \frac{1}{N}\sum_{j=1}^{N} K \sin\left(\phi_j - \phi_i\right), \tag{3.21}$$

in which ξ is the noise component. Under such conditions, the activity of components is described by the Fokker-Planck model and is often used to model physical systems [9]. Some recent development has indicated that such systems might not be stochastic at all but instead belong to a new class of *chronotaxic systems* [103, 104]. These systems

are allowed to interact with the environment through dissipative dynamics. Such interaction, however, is stabilised and resist changes with the support from the internal source of energy. It is important that models which aspire to deal with measurements should belong to either of mentioned classes since the surrounding will often influence physical systems [105–107].

Once the model is defined, one can proceed to extract valuable information from the gathered data. As mentioned, the area of coupled oscillators is fast evolving. Earlier methods focused only on the existence or directionality of coupling [94]. The information extraction was based on the data characteristics such as the time-frequency representation [108] or information-based measures like the mutual information [109] or the Granger causality [110]. Improvements in measuring techniques and advances in computations, however, allowed for more expressive methods and their fit, shifting the trend towards a model-based inference. Few examples of recently introduced methods for reconstructing the coupling functions include the model fitting through a metric optimisation with the least-square fitting, statistical maximum likelihood estimation or the phase resetting.

Methods based on the least-square fitting [111, 112] attempt to fit the data a defined model with free parameters. The simplicity of the approach made it be the first to derive coupling functions, and thus the effective connectivity, from observed oscillations in the phase of recorded data [94]. As the name suggests, their approach starts by constructing a metric and then attempting to find a set of parameters which minimises the metric. This approach is based on minimising the error for the explanation when trying to fit a model. Conceptually somehow different optimisation is with the Bayesian statistics approach of maximising the likelihood estimation (MLE). In the Bayesian philosophy, the focus on the likelihood of whether given data could have been produced with the assumed model [3]. An example of a method that uses MLE to fit the model is MLE-MS [113]. The abbreviation is expanded to MLE multiple shooting (MS) [114] which focuses on fitting through stating a multipoint boundary-value problem for which all these points are nonlinear constraints in the optimisation process. Such definition decreases the chances of an ill initial starting point and allows to better estimate the actual maximum a priori (MAP) values.

Interesting usage of the Bayesian approach is in the Dynamical Bayesian inference (DBI) [115, 116]. The method encompasses the time-variable dynamics of the stochastic differential equation

$$\dot{\phi}_i = \mathbf{f}\left(\phi|\mathbf{c}\right) + \sqrt{\mathbf{D}}\xi_i, \tag{3.22}$$

where the function $\mathbf{f}\left(\phi|\mathbf{c}\right)$ defines independent evolution and interactions with other oscillators both of which are adjustable through parameter $\mathbf{c}$. Additionally, parametric matrix $\mathbf{D}$ characterises the noise diffusion within the network. The core assumption is that the noise ξ is white which allows describing the problem in terms of the Gaussian process. Such a definition allows expressing the optimisation as a set of equations that can be evaluated iteratively. The authors suggest that the reformulated approach is converges within only a few cycles even when initiated with a non-informative flat distribution and parameter $\mathbf{c}_0 = \mathbf{0}$ [116]. The dynamicity of the method comes from the possibility to propagate the estimates along the signal and observe how they change. These steps depend on the system's assumed structure and the certainty of extracted parameters as a priori probability distribution for the following segment depends on the previous' posterior.

A different approach is proposed by Z. Levnajić and A. Pikovsk [117] who aim at reconstructing both the topology and coupling functions of a general oscillatory network. They have introduced the Random Phase Resetting which utilises an ensemble of oscillators that were repeatedly initiated with randomly drawn starting phases and instantaneous frequencies. The core of the method is based on the appropriate formulation of the test function in terms of 2π-period dissipative function. Such definition allows to expand it with the Fourier series and optimise for their coefficients, although the number of harmonics was suggested to be defined based on the assumptions about the system or empirical observations of the data. After enough reruns of the system one can take all results and with appropriate usage of a kernel smoother create an ensemble solution. This approach is more robust than inferring effective connectivity from a single system execution which might overfit to a particular state. The drawback of the method, however, is its invasiveness; it assumes that one is able to repeat experiments with different initial conditions which in case of physical oscillators is commonly impossible.

Presented here methods utilise models for which certain assumption needs to be made. These can be regarding the number of oscillators, the type of noise or the way the oscillators interact with one another. Certain systems, such as cardiovascular [91, 108, 118] or some electrochemical oscillators [86, 88, 113, 119] have been studied thoroughly allowing to incorporate the expert knowledge into the research; however, the majority of systems is yet to be well defined. Moreover, discussed methods focus only on the weak interactions leaving a gap for further research in data-driven methods that describe a general coupled oscillatory system. The advances are also expected to be accompanied with more empirical approach due to the significant improvements in the computational performance allowing for shorter feedback and quicker validation of research ideas.

Chapter 4

EMD analysis

As it has been discussed in previous chapters, EMD is a promising method for obtaining oscillatory features. This chapter details the analysis of the decomposition method, as well as some of its properties and limitations. In Section 4.1 focus is put on EMD's performance limitations under different data formats. Then, in Section 4.2 the problem of validating EMD's results is addressed. This issue is taken care through proposing objective metrics that would satisfy claims made by Huang et al. [17]. Finally, Section 4.3 discusses the issue of components' and their frequencies' mixing. All Sections of this chapters are based on author's published papers [120–122].

4.1 EMD performance

Modern computers used for the signal processing have such a high computational performance that researchers do not think about the efficiency of data handling nor the used format. This simplicity often makes the calculations to be performed in a very high precision formats like double floating point (DFP). Such precision rarely is necessary, but since most analysis is performed in a relatively short time, the configuration is kept for the sake of high precision. For some systems, however, changing the format into a single floating point (SFP) can significantly reduce computation time. An example of such device is graphical processing unit (GPU), which benefits from parallel processing and is reported to work several times faster using SFP precision instead of DFP [123].

EMD can be computationally intensive and not suitable for real-time analysis [57]. Moreover, due to its empirical nature, the method might be very susceptible to the data format. An example of such sensitivity is presented by Rilling et al. in [124], where they studied the smallest sampling frequency for the EMD to work.

In this section, the question of EMD's performance is addressed. This is done by conducting experiments and comparing results under DFP and SFP conditions.

4.1.1 Experiments

Series of experiments were conducted with the purpose to analyse whether there is a difference between two floating point formats and, if so, what is the scale of this discrepancy. EMD used in these experiments was configured with natural cubic spline technique to interpolate envelops on local maxima and minima for the top and bottom envelops, respectively. Each extremum was defined as a peak of a parabola interpolated on three consecutive samples, where the central sample is below (minimum) or above (maximum) its closest surrounding [63]. The stopping criterion for the decomposition was ten consecutive sifting iteration for all of which proto-IMF had the number of extrema and zero-crossings different at most by one.

All examples were generated and analysed using *Python* programming language. The source code of the EMD implementation used in these experiments is freely available from the author's web-page [125]. The numerical manipulations were performed using NumPy scientific package [126].

Worth noting is the fact that the interpolation techniques depend both on points' values and their positions. This means that the difference between two sets will be even greater when comparing values at different positions. When analysing signals, it is advised to scale appropriately independent variable, so that it has exact numerical representation. For binary floating point precision, this means to assign a step value to be a multiple of the power of 2 ($m2^p$).

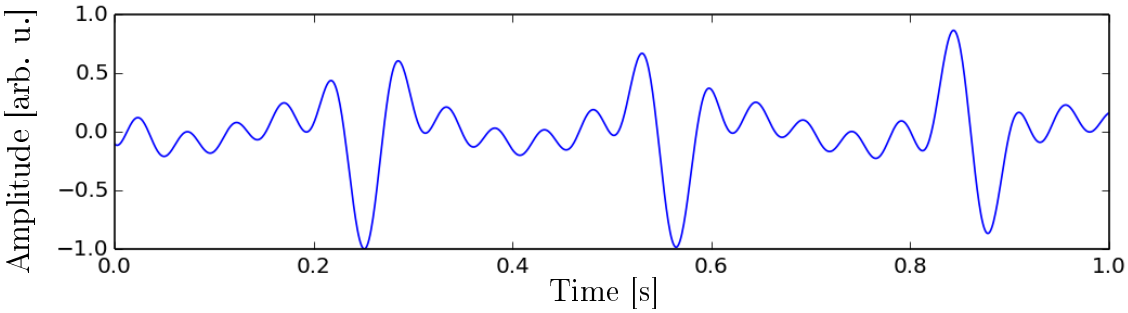


Figure 4.1: Signal used in example 1 and generated according to the formula (4.1).

Example 1

In the first experiment, signal $S(t)$ was generated as a sum of cosines with different frequencies and phases, i.e.

$$s_1(t) = A \sum_{i=1}^{5} \cos\left(2\pi f_i t + \phi_i\right), \tag{4.1}$$

where frequencies and phases are, respectively, $f_i = \{6.1, 9.4, 12.7, 16, 19.3\}$ Hz and $\phi_i = \{0, 1, 2, 3, 4\}$ rad. The particular set of frequencies and phases was chosen so that components are not harmonics of one another and their initial value are different. The amplitude value A was assigned such that the $\max(|S|) = 1$. This normalisation was performed for easier comparisons between presented examples. The signal was generated with time t in the range $[0,\ 1]$ s with a sampling frequency of 1024 Hz and is visualised in Fig. 4.1. Its EMD decomposition is shown in Fig. 4.2, where the solid blue line and dashed green indicate DFP and SFP, respectively. As it can be seen, two sets are visually ideally overlapping each other. To visualise the difference more clearly, the set obtained with SFP was projected onto DFP and subtracted from it, as it has a higher precision. The difference between corresponding IMFs is presented in Figure 4.3. The biggest difference is in order of 10^{-6} which is only one magnitude larger than the machine epsilon for the SFP. Moreover, the discrepancy between compared time series is still about five orders smaller than the magnitude of signal $s_1(t)$. Thus, unless such small values are expected from analysis of the experiment, it can be considered as a negligible noise; they have no meaningful effect on the results.

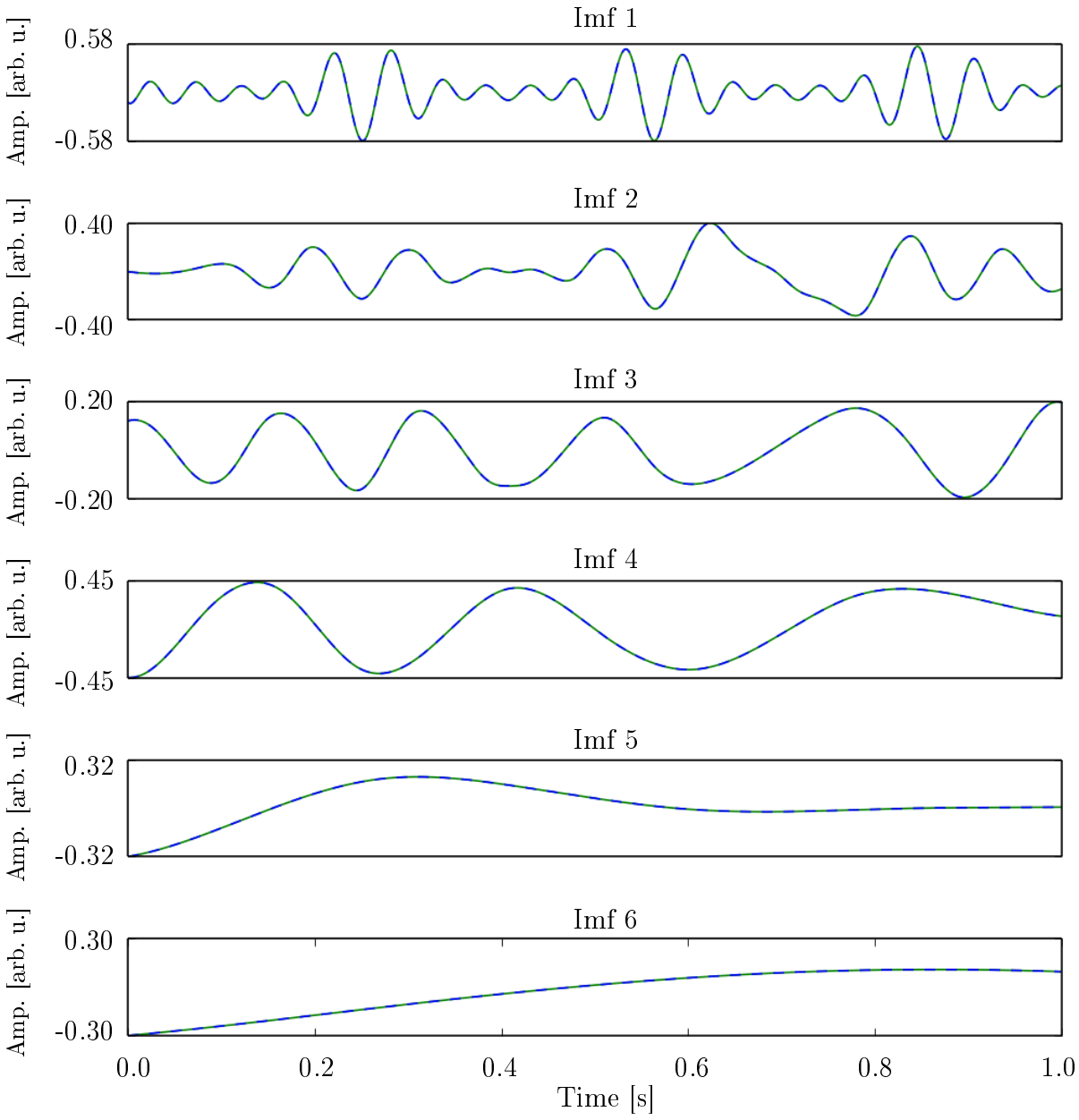


Figure 4.2: EMD decomposition of example 1 signal (Fig. 4.1). Overlapping results for DFP and SFP were plotted with solid blue and dashed green lines, respectively. All functions have the same amplitude scale with arbitrary units.

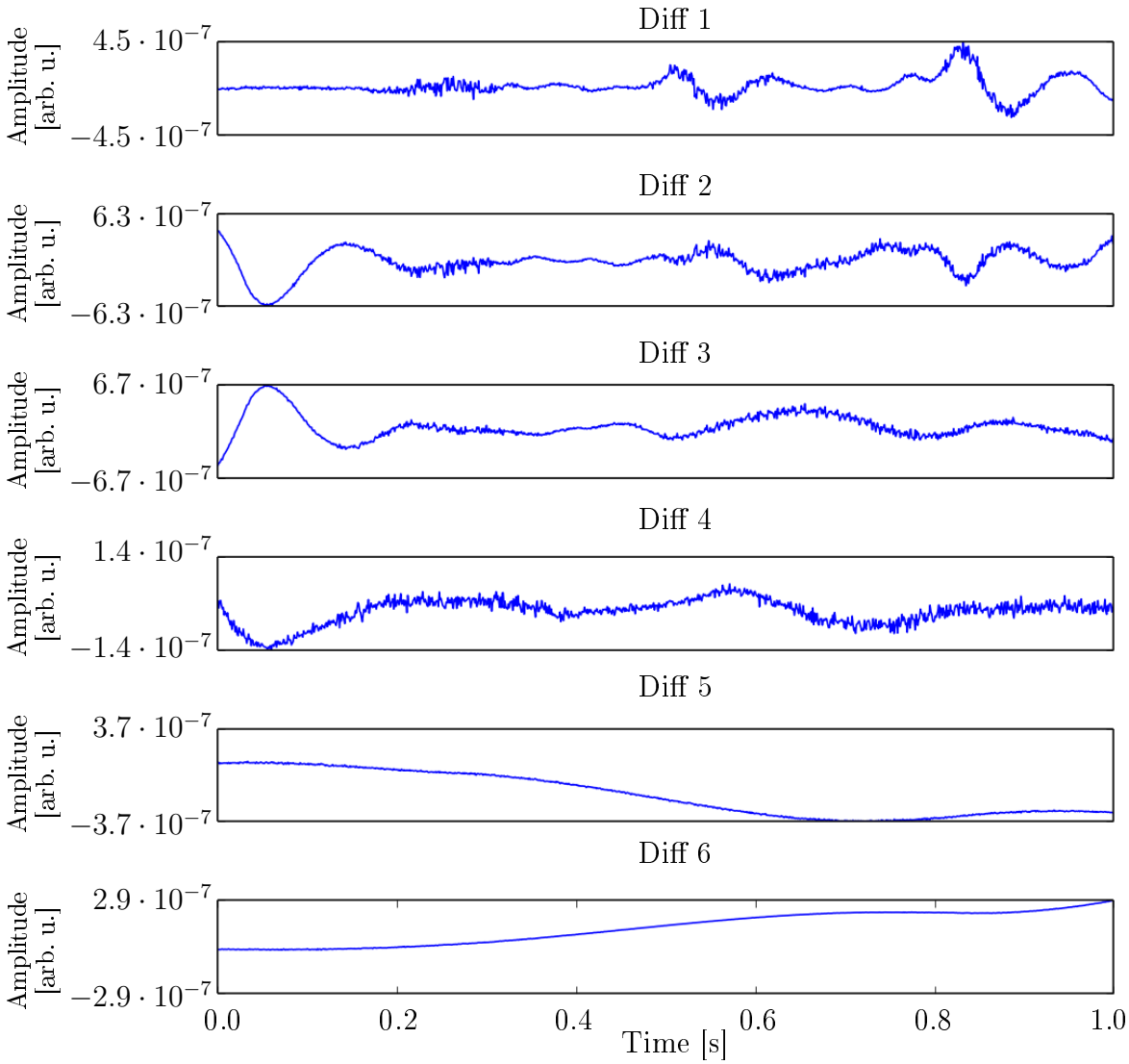


Figure 4.3: Pointwise differences between EMD sets obtained for SFP and DFP from example 1. These of SFP were first projected onto double precision and then subtracted from EMD DFP set. All functions have the same amplitude scale with arbitrary units.

Example 2

For the second example, a signal was generated using 1024 data points sampled from a random Gaussian process, i.e.

$$s_2(t) = \mathcal{N}(\bar{x} = 0, \sigma = 1), \tag{4.2}$$

with a zero mean and a standard deviation of 1. The signal (Fig. 4.4) was additionally scaled so that the biggest amplitude value was one. The EMD decomposition of $s_2(t)$ is shown in Figure 4.5, where solid blue lines and dashed green lines represent DFP and SFP, respectively. Similarly to the previous example, at a presented scale, not much difference between the two sets is visually noticeable. In order to emphasise the discrepancies between these time series an additional figure was generated (Fig. 4.6) in which a pointwise difference for respective IMFs is presented. In this example the biggest variance in visible for IMF 6, where the range of values is of magnitude five. Curiously, the pointwise differences between the initial 5 IMFs have much more variation than those later IMFs. This is due to mismatch in exact representations for extrema in the respective data formats. For high frequency components, there are more extrema and thus the higher chance for a difference in representation. One can also see, that the variance is bigger for time $t \approx 1$ s, where the numerical representation is more sparse.

Comparing results for signals $s_2(t)$ and $s_1(t)$, it seems that there is a bigger difference and more variance between the two data formats in this example. Such discrepancy is expected as the signal $s_2(t)$ has more complex structure and has a larger number of extrema. Nevertheless, when comparing magnitudes the discrepancies between two sets, i.e. magnitude five, to the amplitude of the input signal $s_2(t)$ these differences are magnitude six times smaller and thus can be considered as noise.

Example 3

The final example uses time series which represent a single channel of real EEG data. These recordings were obtained during resting state, i.e. when a person was not involved in any physical, nor mental activity. For analysis, a four seconds segment of the signal, sampled at the rate of 128 Hz, was chosen randomly. Before the EMD decomposition

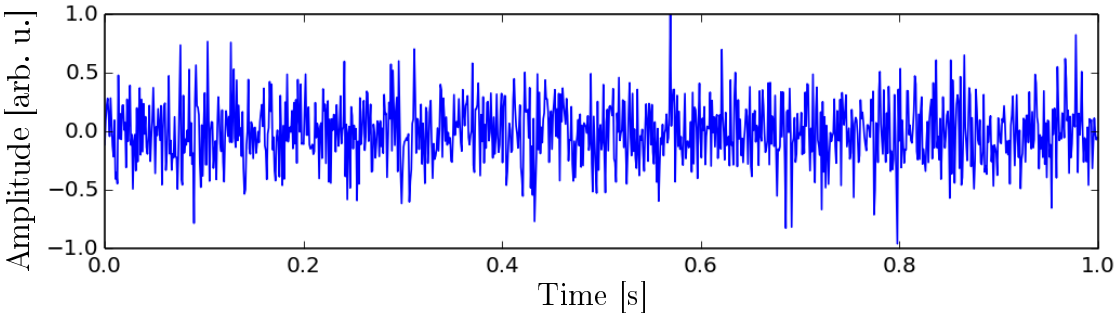


Figure 4.4: The generated signal used in example 2. It consists of 1000 random points drawn from a Gaussian distribution with mean 0 and standard deviation of 1.

was performed the signal was preprocessed, i.e. the mean value was removed and the amplitude was scaled so that the highest amplitude was 1. Also, to decrease the error along the time axis, values were scaled into range $t \in [-1, 1]$ s with sampling frequency 256 Hz. The signal used for decomposition is presented in Fig. 4.7.

Set of IMF components obtained from EMD is shown in Fig. 4.8 using solid blue lines and dashed green lines for DFP and SFP, respectively. The difference between corresponding IMFs is displayed in Fig. 4.9. From this figure, one can see that the difference between obtained sets has the biggest absolute value and variance when $|t| \approx 1$ s, i.e. where the numerical representation is more sparse. Such result is consistent with the previous examples (Figs. 4.3 and 4.6). Nevertheless, similarly to the previous two examples, the difference between both decomposition is relatively small. The magnitude of difference has not changed significantly from the one obtained for $s_2(t)$ (Example 2). Again, the range of differences has the order of magnitude -6 and thus can be ignored when compared to the input signal.

4.1.2 Conclusion

As reported in Section 4.1.1, there is a difference between decomposition obtained for different precision formats, namely single and double floating point precisions. Those differences are tiny and negligible when observing at the input signal scale. However, these differences can be seen clearly when comparing the differences between obtained sets for different floating point formats, i.e. Figures 4.3, 4.6 and 4.9. As it has been pointed out, both absolute values and variance of error are small near $t = 0$ s and increase when approaching $|t| = 1$ s. This is because extrema positions are deter-

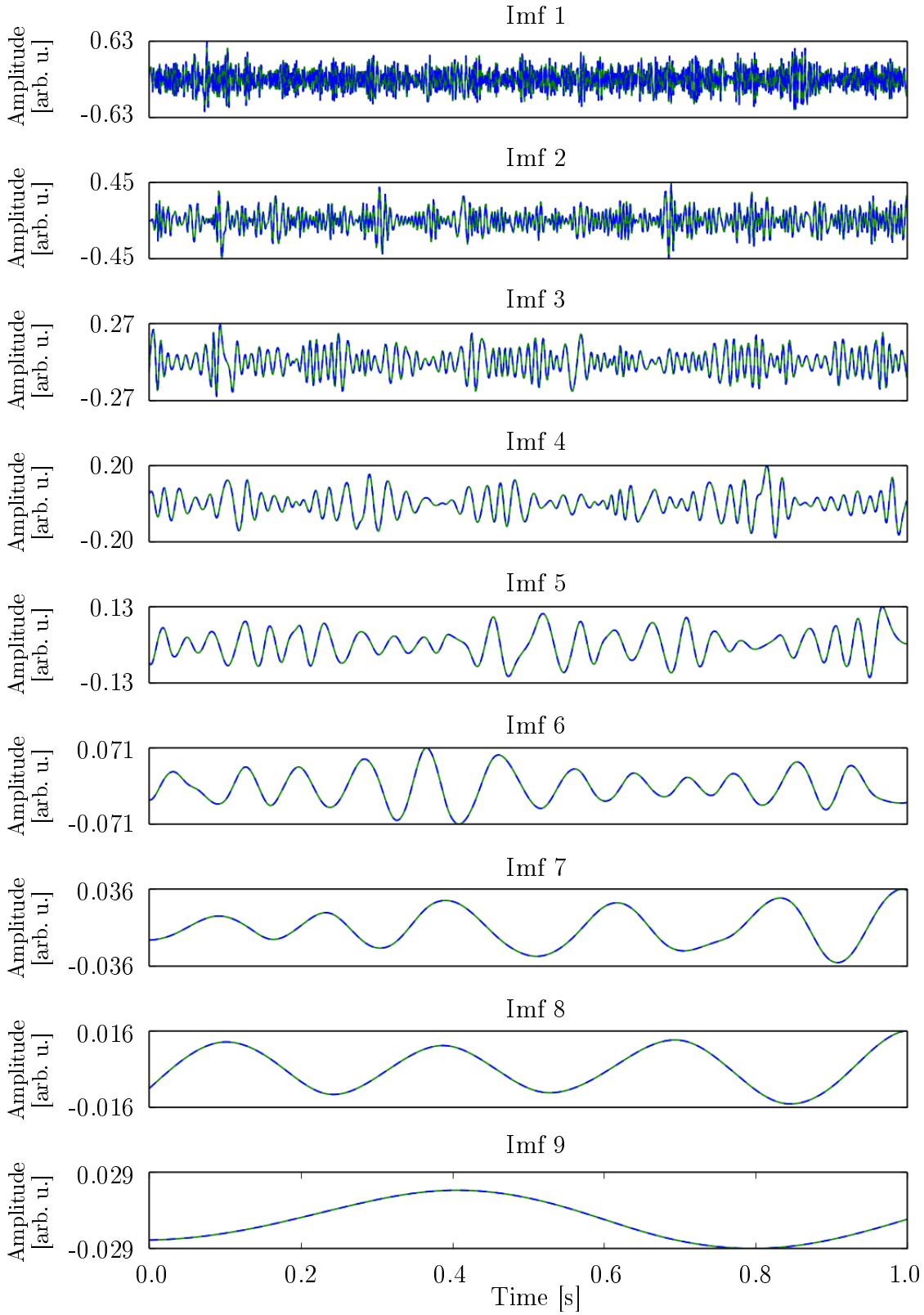


Figure 4.5: EMD decomposition of the signal from example 2 (Fig. 4.4). Decompositions for DFP and SFP are drawn overlapping with solid blue and dashed green lines, respectively.

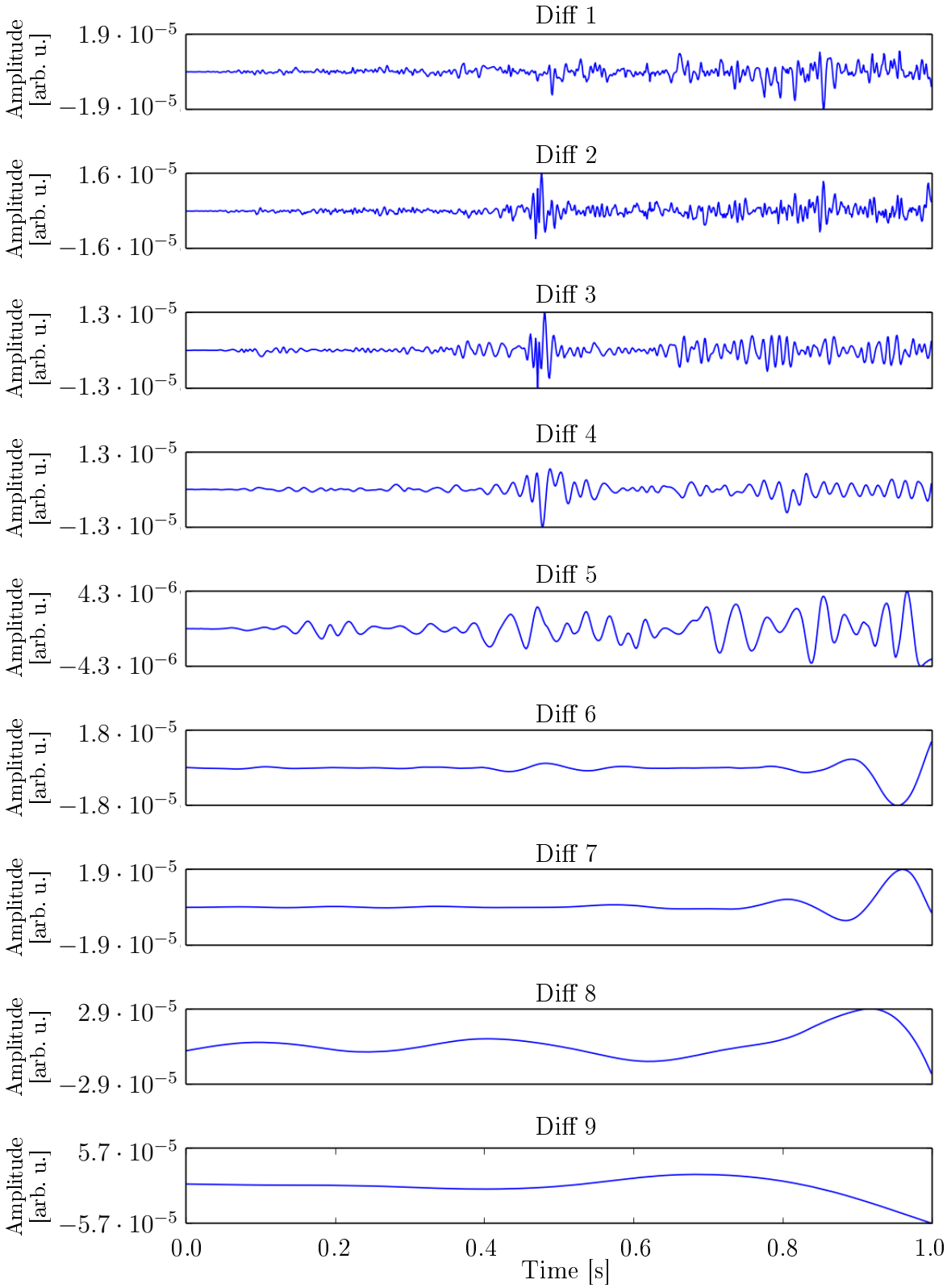


Figure 4.6: Pointwise differences between EMD sets obtained for SFP and DFP from example 2. These of SFP were first projected onto double precision and then subtracted from EMD DFP set. All functions have the same amplitude scale with arbitrary units.

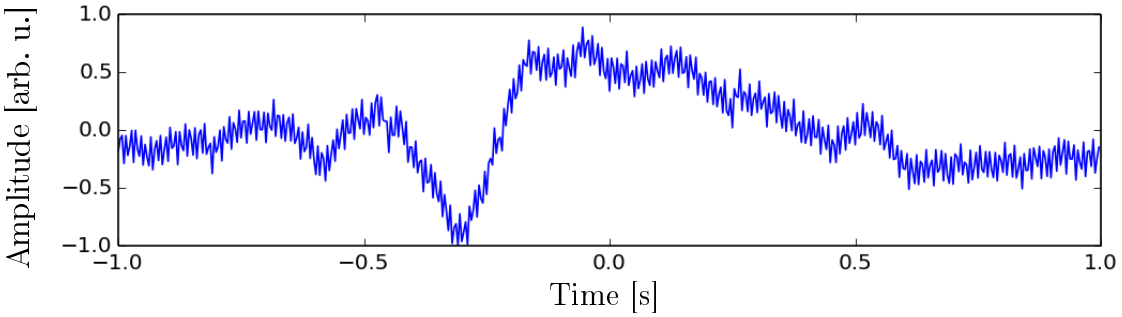


Figure 4.7: EEG data used in the third example. Processing involves removing the mean and scaling amplitude so that the maximum deflection is 1. Timescale changed to span from -1 to 1 with sampling frequency 256 Hz.

mined with parabolic interpolation, thus not necessarily falling onto the exact numerical representation grid. Such pronounced effect is a result of binary floating point representation, which has much bigger resolution close to zero and it decreases with distance [127].

In summary, in all three experiments obtained differences are minimal compared to the average amplitude of each component. Corresponding IMFs produced in two different data formats are visually indistinguishable. Such similarity means that using systems or devices, such as NVIDIA GPU [123], which perform faster on a single floating point compared to double floating point precision, one should be able to decrease computational time without a loss of meaningful content.

4.2 EMD metric

In the original paper on EMD [17] authors noted that small perturbations to the input signal results in different outputs. This is a highly undesired effect, but unfortunately, due to the heuristic nature of the EMD, it is impossible to determine which set of IMFs is better. Nevertheless, researchers usually can make, and do, assessments of method's performance based on their knowledge and experience. Few authors [17, 128, 129] have made attempts to formulate rules based on which sets of IMFs should be chosen. All of them, however, are based on the assumption that IMFs belong to Hilbert's $\mathcal{L}^2$ function space, which is not necessarily true. As stated in [17] obtained orthogonality (in Hilbert sense) is purely by coincidence and should not be expected from the method.

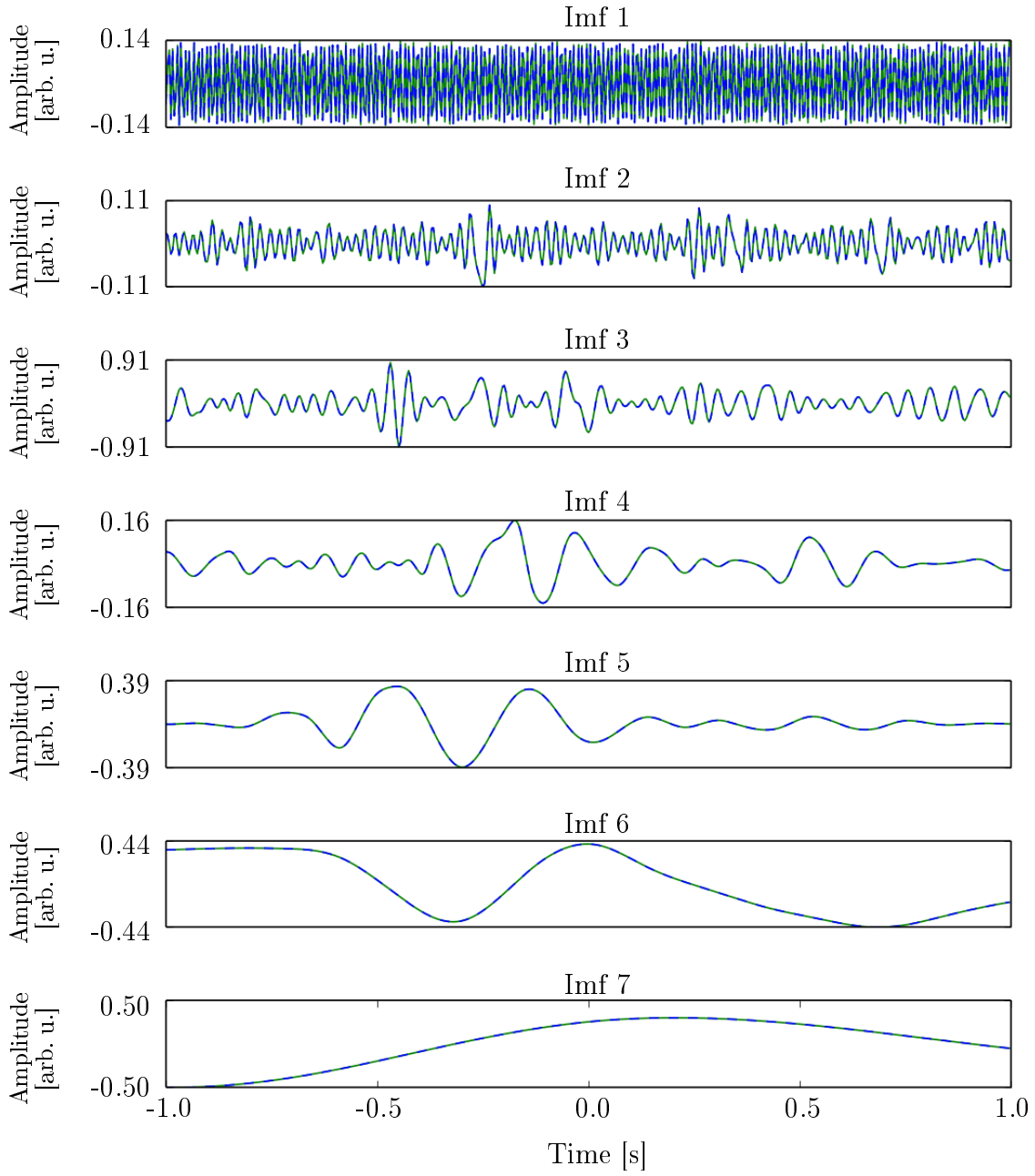


Figure 4.8: EMD decomposition of the EEG signal from example 3 (Fig. 4.7). Decompositions for DFP and SFP are drawn overlapping with solid blue and dashed green lines, respectively.

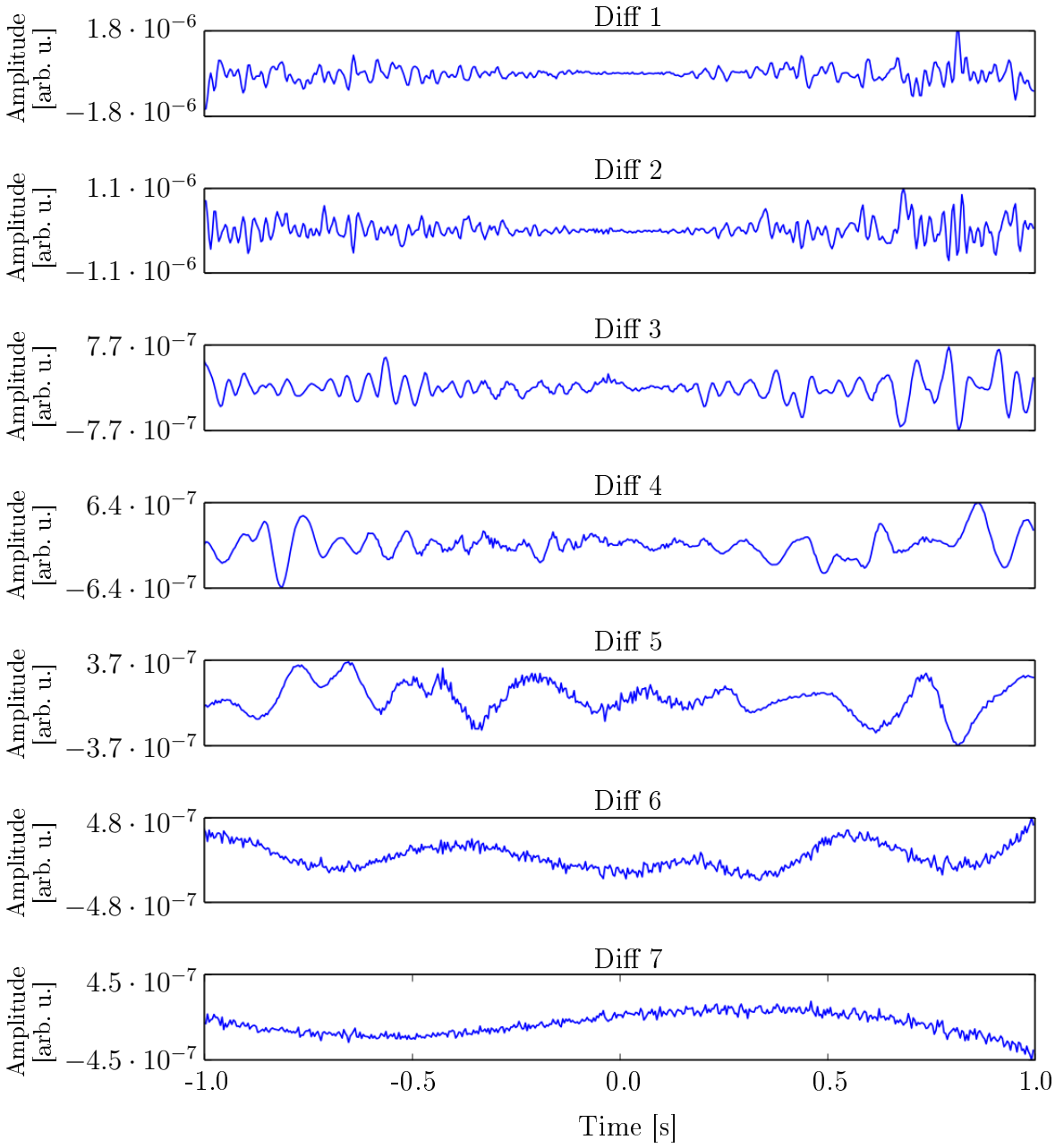


Figure 4.9: Pointwise differences between EMD sets obtained for SFP and DFP from example 3. These of SFP were first projected onto double precision and then subtracted from EMD DFP set. All functions have the same amplitude scale with arbitrary units.

The most popular validation method is based on a visual inspection of the results [61, 63, 128]. Such manual process introduces the subjective opinion into the process. Conclusions from their research cannot be generalised since they have looked at different, often single, features of output. In this chapter objective methods for validating the decomposed sets are introduced. The proposals are based on the features that IMFs are expected to possess. Each variant focuses on different characteristics of the data.

Section 4.2.1 describes proposed validating methods. Then, Section 4.2.2 introduces two numerical experiments for which the conclusions are presented in Section 4.2.3.

4.2.1 Proposed validation methods

The main reason for method proposed in this chapter is to avoid the use of subjective judgement in evaluation of EMD performance. This is achieved by relying on the intrinsic features reportedly possessed by the IMFs, or on those that would help in the future analysis [17]. The main characteristics considered are: 1) decrease of average frequency with the increase of IMFs index, 2) distinct instantaneous frequency for each IMF and 3) disjoint Fourier spectra support for IMF's amplitude and phase.

In this Section, IMFs are represented in polar form, i.e. time series of the j^{th} indexed IMF are assumed to have amplitude a and phase ϕ modulations, i.e. $\text{IMF}_j(t) = a_j(t)\cos\left(\phi_j(t)\right)$. Although, all proposed metrics are designed for continuous functions, in most cases, change to the discrete domain is a straightforward operation. Such processing requires exchanging integration operator over time period T into sum over all data points P.

Validation method I

This metric is based on the empirical evidence for the decrease of average instantaneous frequency, simply referred to as the average frequency, with the increase of IMF's index number. Although the order with which IMFs are constructed corresponds in general to the order of average frequencies, there are instances when the instantaneous frequencies cross over other components and temporally break the order. Since it has been

claimed that each IMF has a significant and non-mixing instantaneous frequency [17], such behaviour is undesired, and hence it is penalised by this metric. Penalties are introduced when instantaneous frequency of an IMF with a lower number (high average frequency) is smaller than the instantaneous frequency of any IMF with a higher number. The penalty value is proportional to the length of the crossing over effect, i.e.

$$m_j^{\mathrm{I}} = \sum_{k=j+1}^{N} \int_{\dot{\phi}_k > \dot{\phi}_j} \frac{dt}{T}, \tag{4.3}$$

where k, j are IMFs' indices. Formula (4.3) compares functions of instantaneous frequencies of two IMFs and returns the total duration over which the IMF with higher index has a lower frequency. The crossing over effect has been presented in Figure 4.10. It shows instantaneous frequency of each IMF as a function of time. Coloured regions indicate where the crossing over occurred. Summing over all pairs of IMFs allows us to assess results for a particular EMD. Metric value for the whole set is given as

$$M_{\mathrm{I}} = \sum_{j=1}^{N} m_j^{\mathrm{I}}, \qquad M_{\mathrm{I}} \in \left[0, \frac{N(N-1)}{2}\right]. \tag{4.4}$$

According to this measure, the best IMF set is the one for which $M_{\mathrm{I}} = 0$, i.e. there are no crossing-over parts in the instantaneous frequency domain. The worst case, $M_{\mathrm{I}} = N(N-1)/2$, is when the order of all IMFs is reversed, i.e. when the first IMF is under all others and the last IMF is above all others. However, this theoretical upper limit is very unlikely and the corresponding IMF set could be still considered upon index reversal.

Validation method II

Another validating measure is based on the Bedrosian Theorem [130] (see Appendix A.3). It refers to the necessary conditions for the signal's amplitude, $a(t)$, and phase, $\phi(t)$, to be exactly recoverable using Hilbert transform. For signal $s(t) = a(t)\cos(\phi(t))$ these conditions require the support of amplitude and phase Fourier spectra to not overlap. In other words, for the amplitude function, $f(t) = a(t)$, and the phase function,

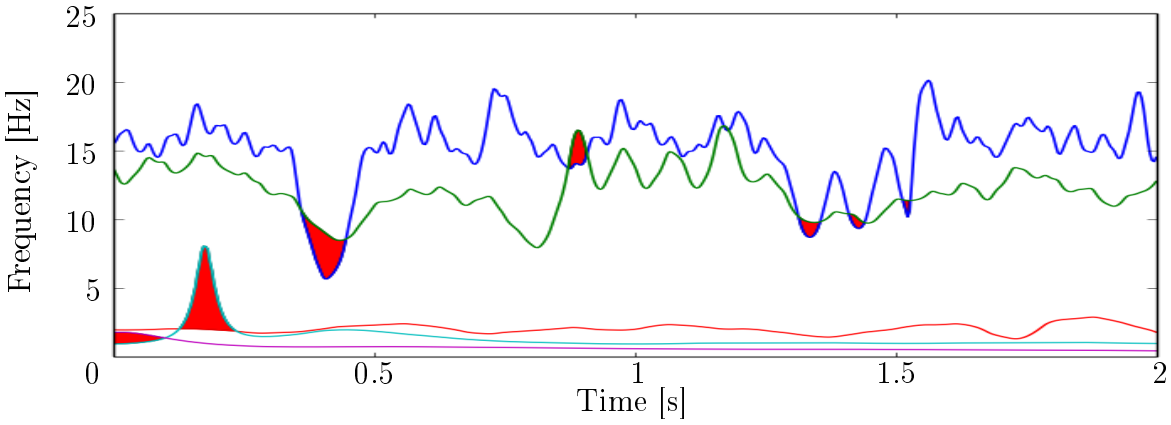


Figure 4.10: A Plot of instantaneous frequency as a function of time for each IMF of an arbitrary signal. Each instantaneous frequency is displayed with different colour and red-coloured regions indicate where the frequency crossing over occurs. Metric M_{I} penalises based on the length of highlighted regions.

$g(t) = \cos(\phi(t))$, the following is required

$$\langle \mathcal{F}(f), \mathcal{F}(g) \rangle = 0, \tag{4.5}$$

where $\mathcal{F}$ represents the Fourier transform and $\langle h(t), l(t) \rangle = \int h^*(t) l(t) dt$ is the dot product. Here it is assumed, that all functions belong to $\mathcal{L}^2$ normed space.

Let $F_j^a = |\mathcal{F}(a_j(t))|$ and $F_j^\phi = |\mathcal{F}(\cos(\phi_j(t)))|$ be absolute values of Fourier transforms of a_j and $\cos(\phi_j)$, respectively, for j^{th} IMF. Their normalised measure of overlapping spectra is given as

$$m_j^{\mathrm{II}} = \frac{\left\langle F_j^a, F_j^\phi \right\rangle}{\sqrt{\|F_j^a\| \|F_j^\phi\|}}, \tag{4.6}$$

where $\|h\| = \langle h, h \rangle$ is a norm of a function h. Assumptions of Bedrosian theorem are completely fulfilled when spectra are not overlapping, thus the minimum value of m_j^{II} is zero. This allows for different definitions of metric for the whole IMF set, depending on application of EMD. The first definition is based on a biggest value of overlap m_j in considered decomposition, i.e.

$$M_{\mathrm{II}} = \max_j \{m_j^{\mathrm{II}}\}, \qquad M_{\mathrm{II}} \in [0, 1], \tag{4.7}$$

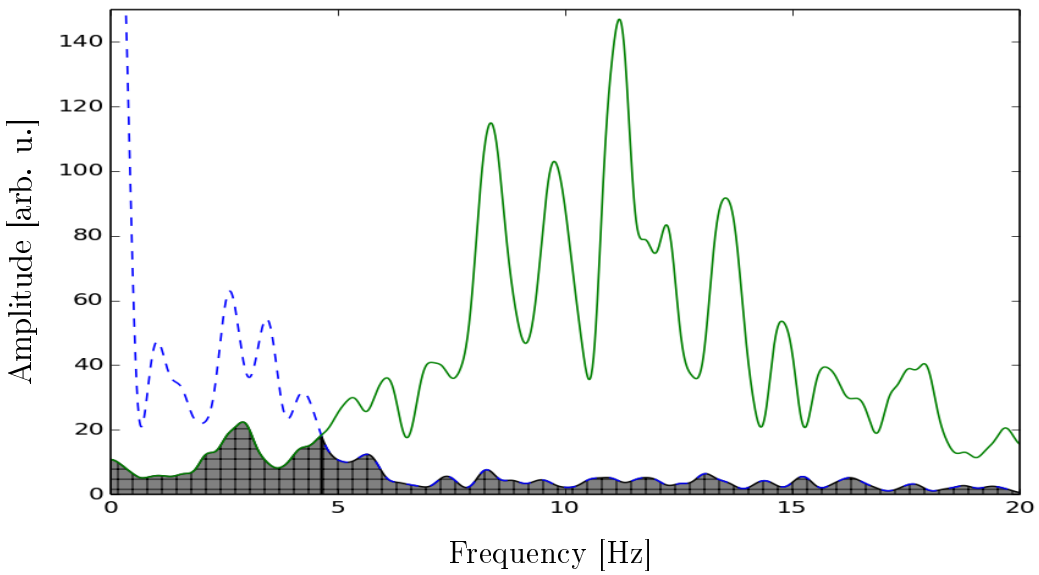


Figure 4.11: Example of comparing Fourier spectrum of amplitude component F^a (blue dashed line) with a spectrum of phase component F^ϕ (solid green line) for an arbitrary signal. Gray-striped area indicates where two components overlap.

and the second refers to the cumulative overlap within the decomposed set, i.e.

$$M_{\mathrm{III}} = \sum_{j=1}^{N} m_j^{\mathrm{II}}, \qquad M_{\mathrm{III}} \in [0, N], \tag{4.8}$$

where in both cases N is the number of extracted IMFs. Zero for both metrics implies no overlap between amplitude's and phase's spectra in any of IMFs.

Visual interpretation of the validation measure (4.6) is presented in Figure 4.11. It shows example Fourier spectra of slowly changing amplitude (dashed line) and higher frequency phase (solid line). Gray-striped region indicates an overlapping area of both spectra. Proposed value is a measure of the ratio of the overlapping area to the total area under both functions.

Since metric M_{III} is a sum over all IMFs, it also contains the one which maximises value m_j^{II} (Eq. (4.6)). This means that M_{III} for each decomposition has to be equal or higher than M_{II}.

Application of the validation measures

Each of the presented metrics highlights different properties of the decomposition. Computing all three values is equivalent to finding a point $M = (M_{\mathrm{I}}, M_{\mathrm{II}}, M_{\mathrm{III}})$ in a 3-dimensional space, where each dimension relates to the specific metric. The best decomposition corresponds to the minimum over all the metrics, i.e. $M = (0, 0, 0)$, and the worst decomposition to $M = (\frac{N(N-1)}{2}, 1, N)$. For any other point, one has to decide on the importance, or weight, for each of the proposed metrics, on the basis of the problem being considered. Although the distance in the M-space is not strictly defined, it can be any $\mathcal{L}^p$ norm. It is suggested using the weighted Manhattan metric, i.e.

$$\|M\| = w_1 M_{\mathrm{I}} + w_2 M_{\mathrm{II}} + w_3 M_{\mathrm{III}}, \tag{4.9}$$

where w_i are respective weights. Their values should reflect the relative importance of features one is concentrated on.

4.2.2 Experiment

Measures proposed in Section 4.2.1 quantify characteristics of well behaved IMFs. The smaller those metrics are, the better the IMF decomposition set represents the desired properties of the EMD-based decomposition. One could also extend the EMD method into an optimisation problem for any parameter, where M_i measures defined cost. In the following examples, optimisation was performed to choose the best value of a parameter H_F for an input signal. The parameter H_F indicates when to stop sifting procedure; it refers to the number of consecutive iterations of sifting for which the residue fulfils the definition of an IMF. EMD was performed for each value of the parameter and the decomposition which minimised its value was chosen as the best.

Two experiments were conducted for an illustration. First one is performed on synthetic signal constructed of four sinusoidal components and the second on filtered Gaussian noise. All signals utilised in the following experiments were generated with a single floating point precision. As it was shown in Section 4.1, in most cases this does not influence the quality of the decomposition but can increase computation performance. Boundary effect introduced by using Hilbert transform was removed by

symmetrically truncating the signal to 80% of the original, i.e. removing initial and final 10% of samples. Additionally, the signal was smoothed by adding mean of each sample's neighbours.

Experiment 1

The first experiment was conducted on a synthetic signal composed of harmonic components. The test signal was generated according to the following formula

$$S_1(t) = \sum_{j=1}^{5} A_j \sin(2\pi f_j t + \phi_j) + \mathcal{N}(0, 0.1), \tag{4.10}$$

where values for the amplitude (A_j), the frequency (f_j) and the phase shift (ϕ_j) are included in Table 4.1. These values were sampled from uniform random distributions with $[0, 2\pi]$ rad range for phase ϕ, range of integers $[0, 5]$ for amplitude A and $[1, 50]$ Hz range for frequency f. For simplicity of analysis and to minimise the effect of mode-mixing [66, 124], an additional constraint was imposed on frequencies such that distance between any two values would not be less than 4. Moreover, in Eq. (4.10) the symbol $\mathcal{N}(\mu, \sigma)$ denotes noise in the form of the normal distribution with a mean μ and a variance σ^2. The graphical representation of the signal can be seen in Fig. 4.12.

The experiment was conducted as follows:

1. Generate test signal $S_1(t)$.
2. Set value range of the parameter — H_F spanning from 1 to 20.
3. For each value H_F, decompose the signal with EMD and calculate all metrics ($M_{\mathrm{I}}, M_{\mathrm{II}}$ and M_{III}).
4. The best decomposition set is the one with the smallest sum of all metrics $M = M_{\mathrm{I}} + M_{\mathrm{II}} + M_{\mathrm{III}}$ (all weights equal).

The signal was decomposed 20 times with the EMD under different stopping criteria conditions, i.e. H_F ranged from 1 to 20. All validating metrics, computed for each decomposition, are presented in Table 4.2. Decomposition sets are assessed based on a total of all metric values for a parameter. The best set is the one with the smallest sum; likewise, the worst set is one with the largest value.

Table 4.1: Parameters for amplitude (A_j), frequency (f_j) and phase shift (ϕ_j) used to generate the signal in experiment 1 according to formula 4.10.

j	1	2	3	4	5
A [arb. u.]	1	1	3	2	3
f [Hz]	35	25	19	15	4
ϕ [rad]	2.0	4.0	0.0	3.4	5.7

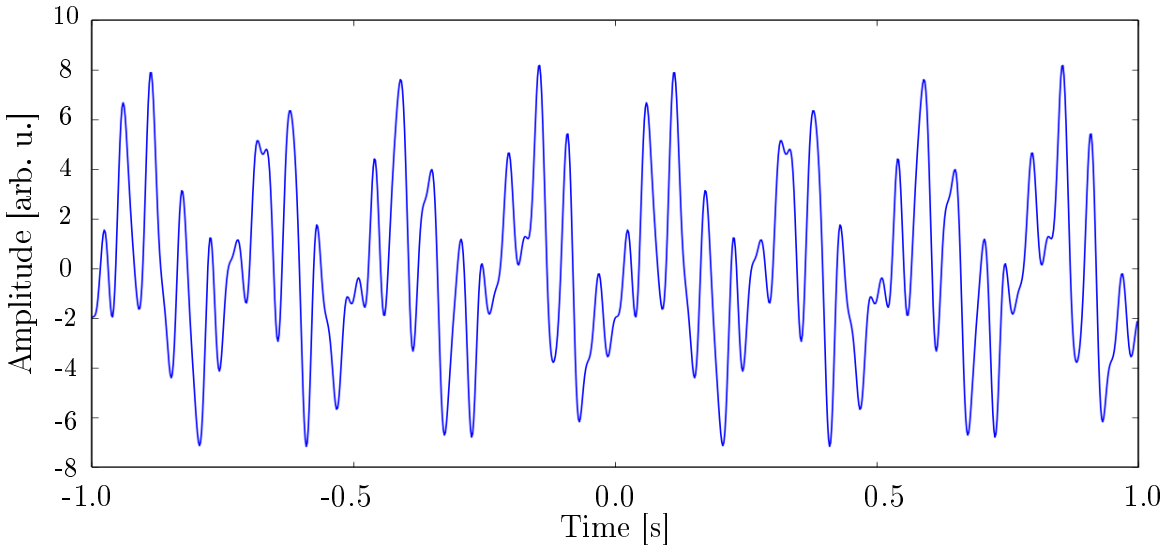


Figure 4.12: Test signal $S_1(t)$ used in the EMD metric experiment with synthetic data which was generated according to Eq. (4.10).

For the provided signal, the best decomposition was obtained with the parameter $H_F = 17$. All extracted IMFs under such condition are shown in Figure 4.13. In contrast, the worst decomposition set is obtained for $H_F = 1$ (Fig. 4.14). Although these results appear similar, especially when comparing the first IMFs of both decompositions, there are few differences between obtained sets. The main difference is the number of produced components. The set with a higher value of metric M has one additional IMF. It might not be evident at which stage it was added, but there seem to be more components with a low number of extrema. Moreover, the difference is also apparent when analysing position and amplitude of extrema. Comparing second and third IMFs one can see that in the best decomposition position of local extrema are relatively evenly spaced. The same observation holds for the amplitude of extrema — there is a clear, stable modulation in the amplitude. The worst decomposition, however, has less visible structure, even though it also contains an apparent repeating pattern.

By design, the discrepancy in obtained sets is due to the different value of parameter H_F. This parameter denotes the number of consecutive sifting iteration for which IMF conditions have to be fulfilled. As explained in the EMD Section (sec. 2.3) each subtraction of the mean removes slowly varying trend, leaving only single fast oscillation. In an idealised scenario, this would mean that the more sifting iterations, the better representation of a single oscillations. However, the sifting operation is not ideal. Each mean is estimated on envelops which are arbitrary interpolated within signal's region and additionally have to be extrapolated onto the boundaries. Any of the imperfections in the estimating process will contaminate the signal and will be emphasised with each sifting operation. Thus, the larger the number of iteration the more visible effect of the inaccurate estimations of the actual mean signal. This means that for small and large values of H_F, the EMD will perform worse than for midrange values. The exact progress of decomposition's wellness depends on all its parameters and applied algorithm.

Table 4.2: Metric values obtained by performing the EMD on $S_1(t)$ with varying value of H_F parameter.

H_F	M_{I}	M_{II}	M_{III}	$M_{\mathrm{I}} + M_{\mathrm{II}} + M_{\mathrm{III}}$
1	1.17	0.63	0.87	2.66
2	1.13	0.57	0.77	2.47
3	1.33	0.48	0.53	2.27
4	1.17	0.61	0.88	2.65
5	0.91	0.22	0.51	1.64
6	0.92	0.22	0.51	1.65
7	0.93	0.22	0.50	1.64
8	1.15	0.35	0.69	2.19
9	1.12	0.32	0.66	2.10
10	1.12	0.35	0.72	2.19
11	1.18	0.30	0.54	2.02
12	1.20	0.29	0.54	2.03
13	0.98	0.14	0.31	1.44
14	0.98	0.16	0.32	1.45
15	0.96	0.16	0.31	1.43
16	0.97	0.16	0.32	1.45
17	0.96	0.16	0.30	1.42
18	1.44	0.48	0.64	2.56
19	1.45	0.36	0.50	2.32
20	1.56	0.27	0.45	2.24

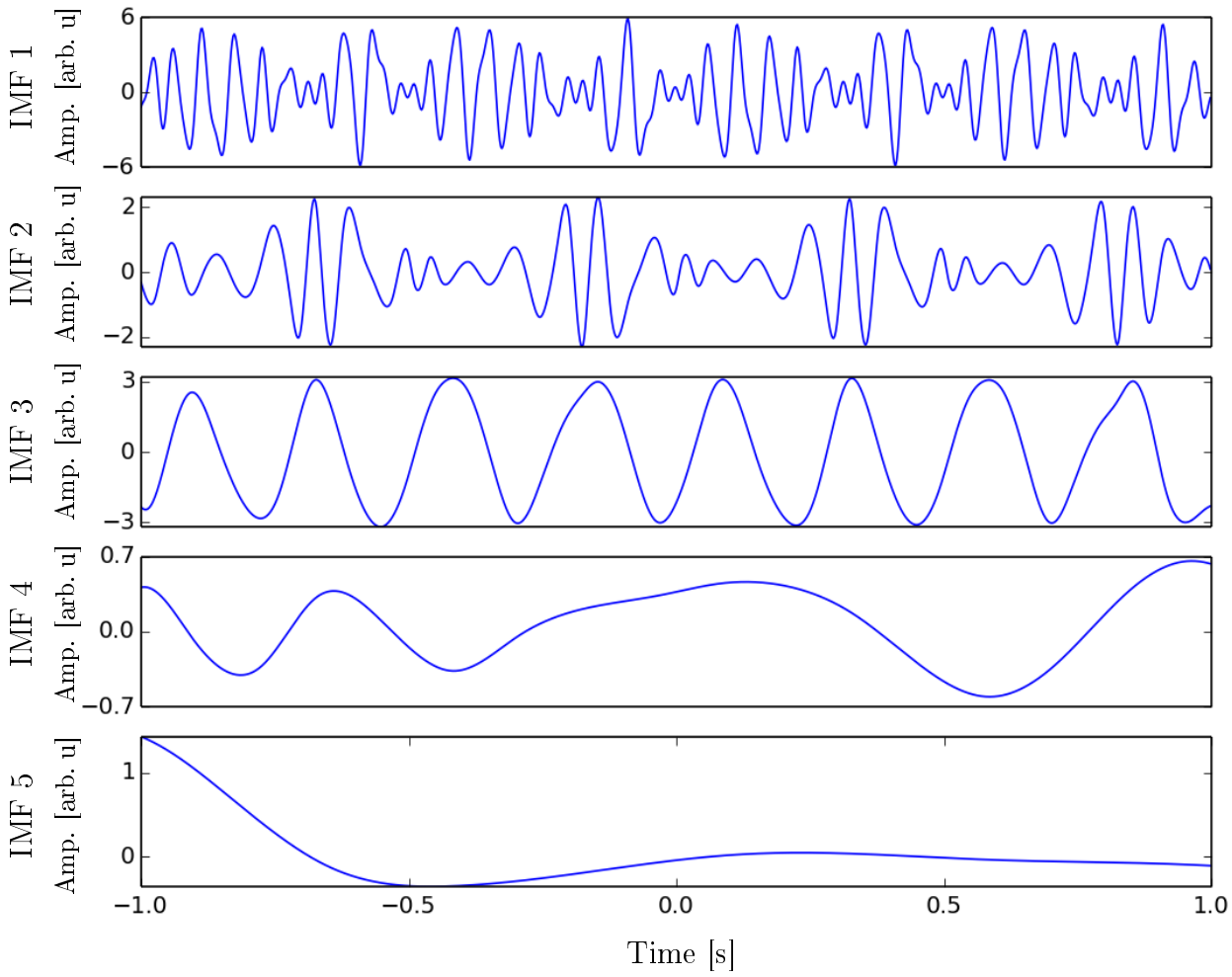


Figure 4.13: The best EMD decomposition set, i.e. producing the smallest metric M value, for the synthetic signal given the range of H_F parameters. Decomposition obtained from signal $S_1(t)$ with H_F=17.

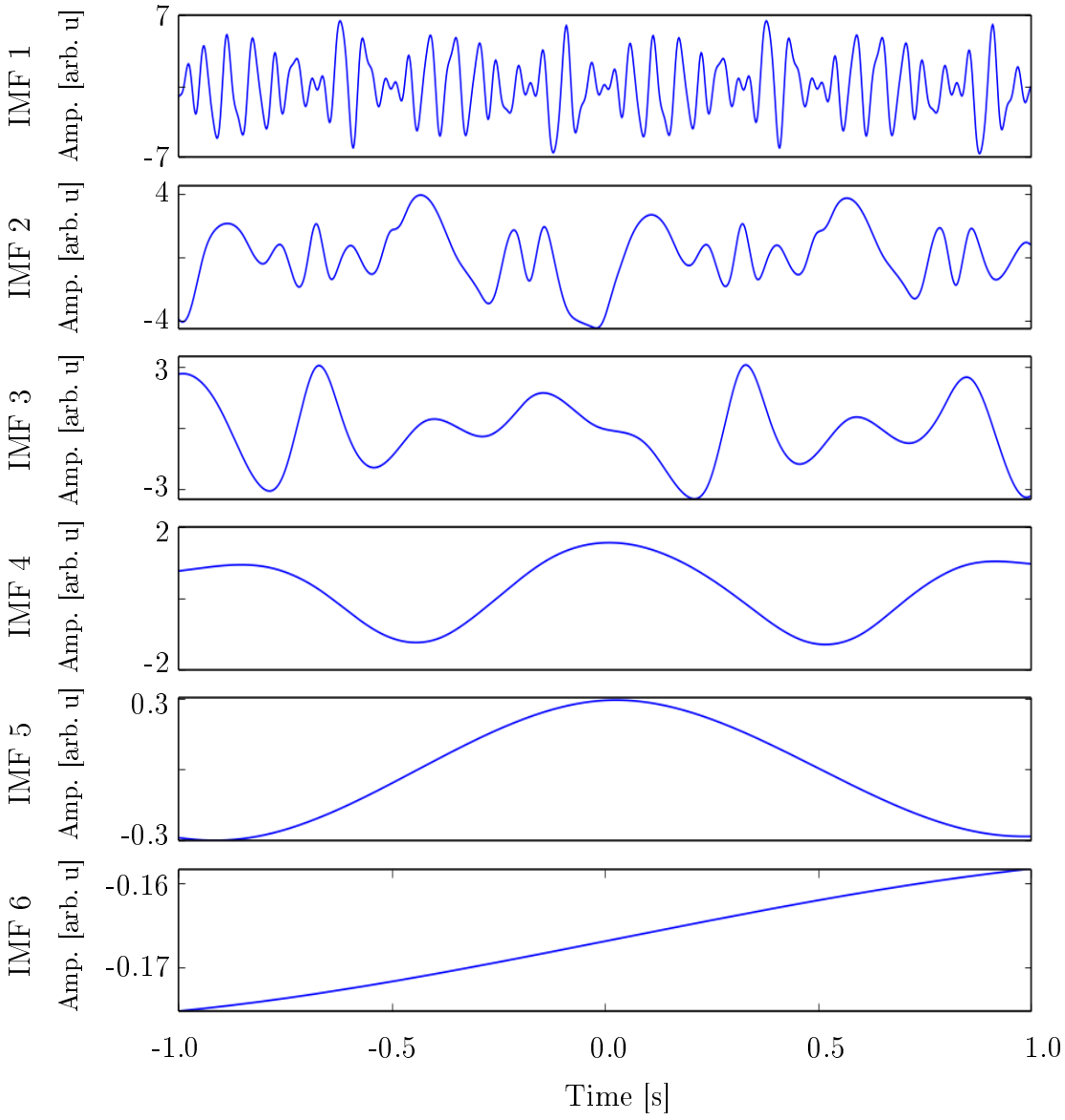


Figure 4.14: The worst EMD decomposition set, i.e. producing the biggest metric M value, for the synthetic signal given the range of H_F parameters. Decomposition obtained from signal $S_1(t)$ with H_F=1.

Experiment 2

For this experiment signal was generated using Gaussian noise with mean value 0 and standard deviation 1

$$S_2(t) = \mathcal{N}(0, 1). \tag{4.11}$$

As it has been suggested in [129], in order to obtain meaningful decomposition it is necessary to have ratio of signal's sampling frequency, f_s, to the highest Fourier frequency, f_f, of at least 10 ($f_s/f_f \geq 10$). This means that when sampling with frequency 500 Hz, the signal has to be low-pass filtered with the cut off frequency of 50 Hz. For the experiment, we used the zero-phase Butterworth low-pass filter of order 4. The resulting signal is visualised in Figure 4.15.

The experiment was conducted similarly to the Experiment I, with the difference in weights used to calculate the metric. The exact steps of execution were:

1. Generate the test signal, $S_2(t)$.
2. Set the value range of the parameter — H_F spanning from 1 to 20.
3. For each value H_F, decompose the signal with the EMD and calculate all metrics ($M_{\mathrm{I}}, M_{\mathrm{II}}$ and M_{III}).
4. The best decomposition set is the one with the smallest sum of all metrics $M = 2 \cdot M_{\mathrm{I}} + M_{\mathrm{II}} + 0.5 \cdot M_{\mathrm{III}}$.

Such a choice of weights puts more emphasis on selecting IMFs with more mutually separate instantaneous frequencies. It also increases the significance of the component with the most overlapping amplitude and phase Fourier spectra; the sum of all measuring values (Eq. (4.6)) has lower priority. Overall, the metric, M, is meant to select a decomposition with the most distinct frequencies. This effect should be visible when analysing location of extrema, as they should be spaced more evenly.

All calculated values of metrics are presented in Table 4.3. The last column contains a weighted sum of all other metrics for each value of the parameter H_F. The smallest and the largest values are obtained for H_F equal to 2 and 12, respectively. The best decomposition can be seen in Figure 4.16, whereas the worst in Figure 4.17. Comparing the two decompositions one can see the difference in the number of IMFs

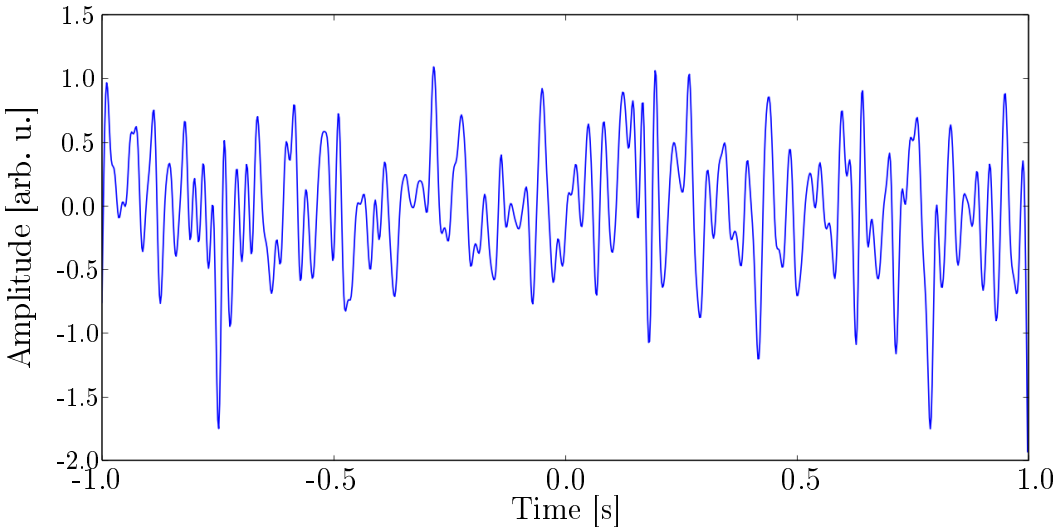


Figure 4.15: Test signal $S_2(t)$ used in the EMD metric experiment with filtered Gaussian noise generated.

— two more in the worst EMD set. In this case, two first IMFs seem to be similar, or at least without any apparent differences. Analysing position and value of the third component's extrema, one can see that there are more extrema for $H_F = 12$. Additionally, IMF 6 from the worst decomposition does not seem to have any close counterpart in the best decomposition. Its small amplitude suggests that the component might be hidden within amplitude modulation of any other IMF.

Compared to the previous example, in this case, the decomposition optimum was obtained with a relatively small value of sifting parameter, i.e. for $H_F = 2$. As it can be seen in Table 4.3 the main effect on the metric value has M_I as the other two columns have very little spread of values. Moreover, one can see that with the increase of H_F metric values M_I increase gradually, although not monolithically. The difference between this example and the previous one is in the structural complexity. Signal, $S_2(t)$, used in this experiment is more complicated and thus requires more sifting iterations to extract each component. Since each IMF depends on decomposition process of all previous components, the estimation contamination can spread between IMFs. It can be seen that IMF sets for $H_F \geq 12$ have similar values in all metrics. For the worst decomposition, i.e. $H_F = 12$, all metrics give the largest value except for M_I which is 0.01 behind the largest value.

Table 4.3: Metric values obtained by performing the EMD on $S_2(t)$ with varying value of H_F parameter.

H_F	M_{I}	M_{II}	M_{III}	$2M_{\mathrm{I}} + M_{\mathrm{II}} + 0.5M_{\mathrm{III}}$
1	1.64	0.60	0.84	4.30
2	1.62	0.59	0.70	4.18
3	1.78	0.58	0.85	4.56
4	2.17	0.56	0.67	5.23
5	2.27	0.53	0.82	5.48
6	2.17	0.45	0.69	5.14
7	2.25	0.47	0.79	5.37
8	2.07	0.57	0.89	5.15
9	2.21	0.53	1.04	5.47
10	2.20	0.55	0.86	5.38
11	2.25	0.55	0.87	5.48
12	2.58	0.60	1.04	6.28
13	2.56	0.60	1.00	6.22
14	2.17	0.51	0.74	5.22
15	2.54	0.61	0.81	6.09
16	2.51	0.60	0.80	6.02
17	2.52	0.60	0.80	6.04
18	2.51	0.60	0.78	6.01
19	2.52	0.60	0.78	6.03
20	2.59	0.58	0.99	6.26

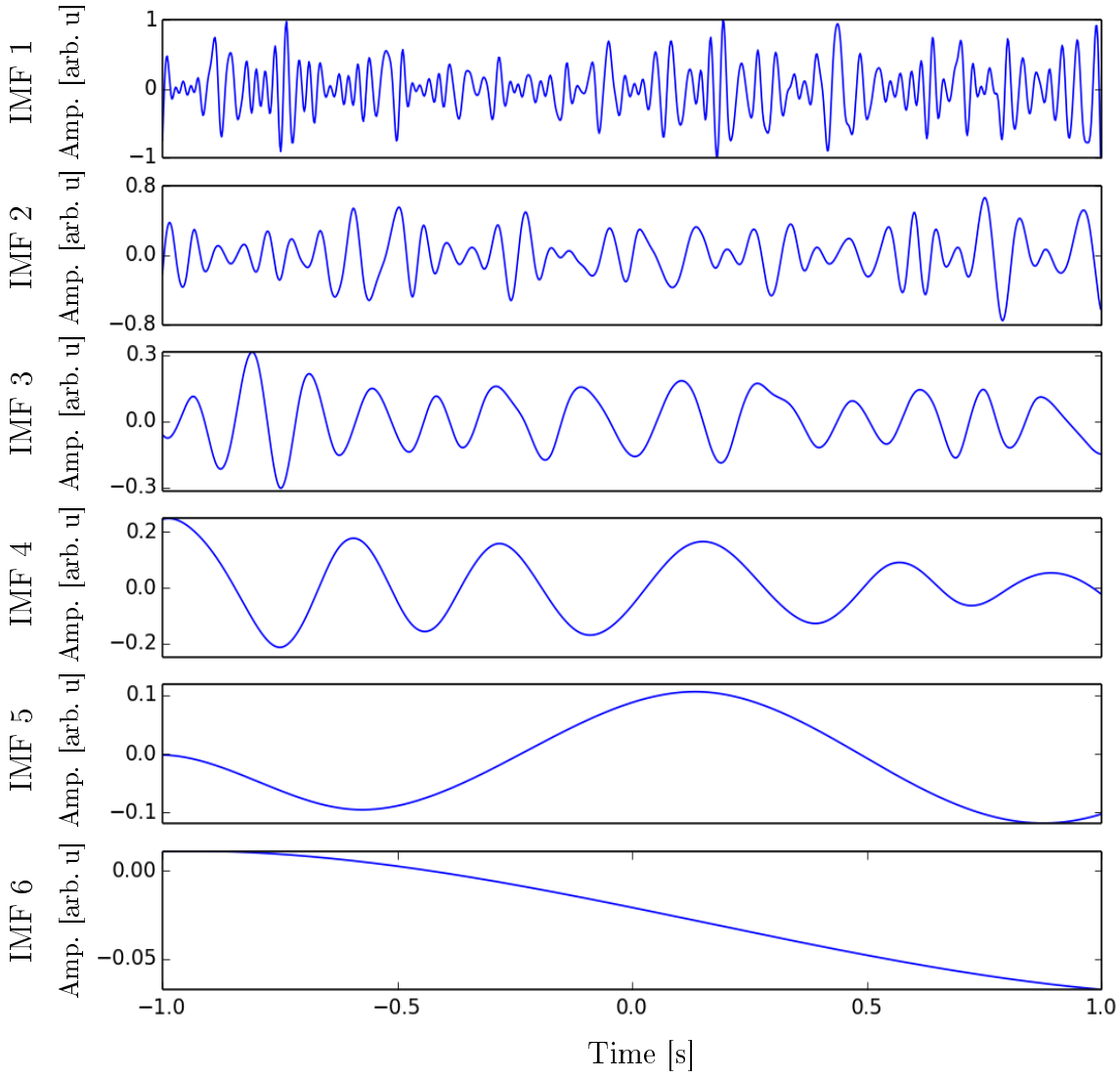


Figure 4.16: The best EMD decomposition set, i.e. producing the smallest metric M value, for the Gaussian noise signal given the range of H_F parameters. Decomposition obtained from signal $S_2(t)$ with H_F=2.

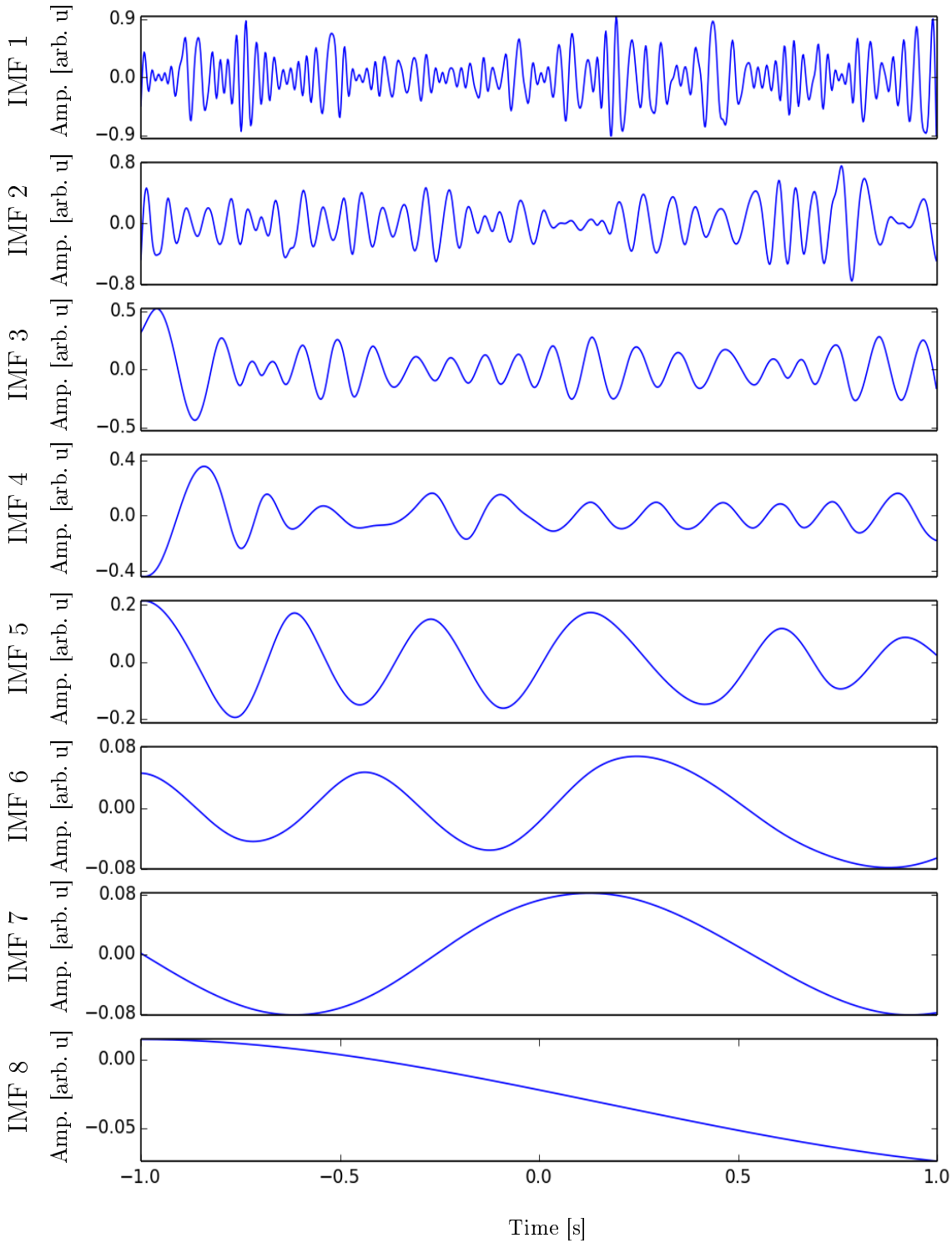


Figure 4.17: The worst EMD decomposition set, i.e. producing the biggest metric M value, for the Gaussian noise signal given the range of H_F parameters. Decomposition obtained from signal $S_2(t)$ with H_F=12.

4.2.3 Conclusion

Metrics proposed in this chapter refer to idealised characteristics of EMD — either suggested or intended while creating the method. In contrast to the most validation methods discussed in the literature, those proposed in this chapter are objective, independent from user's subjective judgement of the decomposition. Since EMD creates a set of oscillatory components, the properties on which these metrics were based are related to their frequencies. Validation of the whole IMF decomposition set is performed either by significant and mutually separate instantaneous frequencies, or on the attempt to fulfil or be close to the Bedrosian conditions.

Analysis of examples presented in Section 4.2.2 seems to support the usefulness of the proposed metrics. Although the behaviour of the decomposition can be changed by adjusting weights w_i of the desired feature (Eq. (4.9)), it should still provide good results. The best decomposition always produces fewer IMFs, which suggests having more compact information representation of the original signal. Moreover, visual inspection confirms that the best decompositions have better structured (evenly spaced locations of extrema) than the worst ones.

Nevertheless, despite the foregoing discussion, until there is a mathematical foundation of the EMD, it is impossible to create a single metric, which would capture all required features. Proposed measures should be considered as assistance for an inexperienced user, providing him/her with additional arguments for used parameters choices.

4.3 Frequency mixing

In the original article, Huang et al. [17] have argued that functions which fulfil IMF properties (discussed in Section 2.3) have significant modes and well-behaved instantaneous frequencies obtained via the Hilbert transform. The authors call these components physically meaningful, as typically their instantaneous frequencies are non-negative and they have modulations in amplitude and frequency.

Due to the absence of a mathematical framework for EMD, it can only be analysed empirically. As it has been observed by many [66, 67, 131], the decomposition process

suffers from the mode-mixing phenomena, which describes a situation, when a mode contains more than one scale. Another deficiency of EMD is its decomposition stability problem. Whenever acting on a single IMF $I(t)$ EMD returns the same function, i.e.

$$I(t) \xrightarrow{EMD} I(t). \tag{4.12}$$

However, although EMD decomposition of a signal $S(t)$ gives a set of k IMFs, i.e.

$$S(t) \xrightarrow{EMD} \mathcal{C}^k = \{I_1(t), I_2(t), \dots I_k(t)\}, \tag{4.13}$$

a signal composed from a set $\mathcal{C}^l$, $\hat{S}(t) = \sum_i^l I_i(t)$, which is a subset of $\mathcal{C}^k$, will produce another set of IMFs,

$$\hat{S}(t) \xrightarrow{EMD} \hat{\mathcal{C}}^n = \{\hat{I}_1(t), \hat{I}_2(t), \dots \hat{I}_n(t)\}, \tag{4.14}$$

but there is not necessarily any correspondence between any $\hat{I}_i(t)$ and $I_j(t)$. A thorough study of frequency mixing for two components was presented in [132]. The authors compared IMFs obtained from signals composed of two cosines, i.e. $S(t) = \cos(t) + a\cos(ft + \phi)$, generated with different values of amplitude, $a \in \mathbb{R}$, and frequency, $f \in (0, 1)$, values. They found that for $a < 0.5$ the quality of the decomposition, i.e. similarities of IMFs and the original cosine components, depends only on the frequency, f. Moreover, they found that the smaller the frequency ($f \approx 0$), the better the recovery of initial modes. They also noted that the transition from almost perfect decompositions to near impossible increases monotonically with f. This, however, poses a question: what is responsible for mixing when frequencies of input components have similar values? Since only the frequency varies, this means that the mixing phenomenon depends only on that parameter. As the system is closed, i.e. there are only two components, any added modulation to the first IMF is the same as removing that modulation from the second IMF. Given the results obtained in [132], it has been hypothesised [122] that this phenomenon is due to the mutual relationship between instantaneous frequencies of the sources. To validate this hypothesis, harmonic components are used and tested whether the coupling between them can account for observed IMFs' frequency

dependencies. Kuramoto model (Section 3.2) was chosen to model this behaviour as it assumes phase coupling between the oscillators as a function of their instantaneous phase differences.

This Section attempts to answer whether frequency mixing observed in EMD can be due to EMD decomposition efficiently mimicking the Kuramoto coupling between oscillators. Section 4.3.2 presents all conducted experiments, and their results are discussed in 4.3.3.

4.3.1 Parameter estimation

Kuramoto model (Section 3.2) describes interactions between oscillators. Each oscillator has an intrinsic frequency and its observed frequency modulated by the difference between each pair of phases [10]. Mathematical form of the model for i^{th} oscillator is given as

$$\dot{\theta}_i = \omega_i + \sum_{j=1}^{N} k_{i,j} \sin\left(\theta_j - \theta_i\right), \tag{4.15}$$

where the dot above variable, i.e. $\dot{x}$, denotes the time derivative, ω_i is the intrinsic frequency and $k_{i,j}$ are coupling strength parameters. To fully solve these coupled differential equations N initial phase values, N intrinsic frequencies and $N(N-1)$ values for coupling strengths parameters are required. This means that in total $N(N+1)$ parameters fully describe a system and these parameters have to be estimated from data.

The best fit of the model was performed using particle swarm optimisation (PSO) [133]. It is an optimisation method where many agents iteratively search through parameter space. They interact with each other by exchanging their fitness and their position. Movement of the i^{th} particle is dictated by the formula

$$\begin{aligned} \vec{V}_i(t) = \phi_V \vec{V}_i(t-1) \quad & +\phi_L \left(\vec{B}_i(t) - \vec{X}_i(t)\right) \\ & +\phi_G \left(\vec{G}(t) - \vec{X}_i(t)\right), \end{aligned} \tag{4.16}$$

where X_i, V_i and B_i are the particle's current position in the parameter space, its velocity and its best position until time t, respectively. G is the best global po-

sition discovered by any particle until time t. Communication is performed after each iteration, when all particles update their positions based on the velocity, i.e. $X(t+1) = X(t) + V(t)$. At $t = 0$ all particles have a randomly chosen position and velocity.

In the experiments described below, swarms consisted of 400 particles. The swarm size was chosen to be large in comparison to standard recommendations to ensure more thorough coverage of the search domain. Each particle traverses a 6 dimensional parameter space, where each location is a vector consisting of all initial values, i.e. $X_i(0) = [\theta_{01}, \theta_{02}, \omega_1, \omega_2, k_{1,2}, k_{2,1}]$, which fully determines the Kuramoto system with two oscillators. The intrinsic frequencies for the oscillators were drawn from Gaussian distributions, where the expected values and standard deviations were equal to those of the IMFs' instantaneous frequencies. Phases and couplings values k also were drawn from Gaussian distributions; however, their absolute values were used. The expected values and standard deviations were π and $\pi/4$ for phases, and 0 and 5 for k values as it has been observed that these parameters cover most of the relevant parameter space.

The optimising fitness function is given as

$$M = \sqrt{\frac{1}{N}\frac{1}{T}\sum_{n=1}^{N}\left(\sum_{t=1}^{T}\left(\dot{\Theta}_n(t) - \dot{\theta}_n(t)\right)^2\right)}, \tag{4.17}$$

where Θ_n and θ_n are phases of n^{th} IMF and reconstructed oscillator respectively. Summation goes through all t timestamps and there are N oscillators of length T time points. The optimisation procedure terminates when, after 100 initial iterations, the cost value is the same for 20 consecutive iterations.

4.3.2 Experiments

To show phase coupling between components of the input signal, a set of experiments was conducted. They all were based on synthetic data constructed according to the formula

$$S_f(t) = \cos(13 \cdot 2\pi t) + 2\cos(f \cdot 2\pi t + \phi), \tag{4.18}$$

where frequency ($f \in [1, 7]$ Hz with step 0.5) and initial phase ($\phi \in [0, 2\pi]$ rad with step $\frac{2\pi}{15}$) of the second component were varied. To ensure robustness of the results against the influence of implementation details, for each f and ϕ combination, EMD was performed many times with different spline techniques (natural cubic or Akima [134] spline) and parameters related to stopping criteria [65]. Out of these options a set was chosen that minimises defined in previous Section (Sec. 4.2) metric $M = \sqrt{M_1^2 + M_2^2}$, where M_1 quantifies the pairwise crossover of instantaneous frequencies between IMFs, and M_2, which penalises based on the overlap between IMF's amplitude and phase spectra. After the decomposition was performed, the Hilbert transform of each IMF was used to obtain its instantaneous phase and amplitude. Due to the error created by boundary effects, each component's first and last 0.5 s had to be removed leaving 2 s of the signal.

A typical EMD decomposition obtained in the experiment is presented in Figure 4.18, in which case the varied component had frequency $f = 4$ Hz. The top graph contains input signal, whereas the second and the third rows are respectively first and second IMFs. Each component (solid line) was scaled (scale in the top left corner) so that its maximum value was one. In the same figure, dashed lines were used to display cosine function of instantaneous phase ($\cos \Phi(t)$) obtained via Hilbert's transformation of the corresponding IMF. Almost complete overlap of the presented functions suggests that there is very little amplitude modulation. For this reason, in further analysis, only phase modulations are considered.

It has been observed that there were modulations in the obtained instantaneous frequencies of the IMFs. To better understand these changes for each instantaneous frequency time series, a Fourier spectrum was obtained. If there were no modulations of instantaneous frequency, one would expect zero-valued spectrum everywhere except for frequency 0. In contrast, a predominant periodic modulation would manifest itself as a single spike in the Fourier spectrum. Figure 4.19 displays spectra for all f for the first IMF. For comparison purpose, each spectrum was scaled such that the largest value was set to one. In the figure, one can observe that peaks are aligned. This additionally is emphasised by overlaying results with a line ($F = 13 - f$). Similar results are visible for the second IMF, presented in Figure 4.20. However, in this case, peaks are aligned along a different line, i.e. $F = 2 \cdot (13 - f)$.

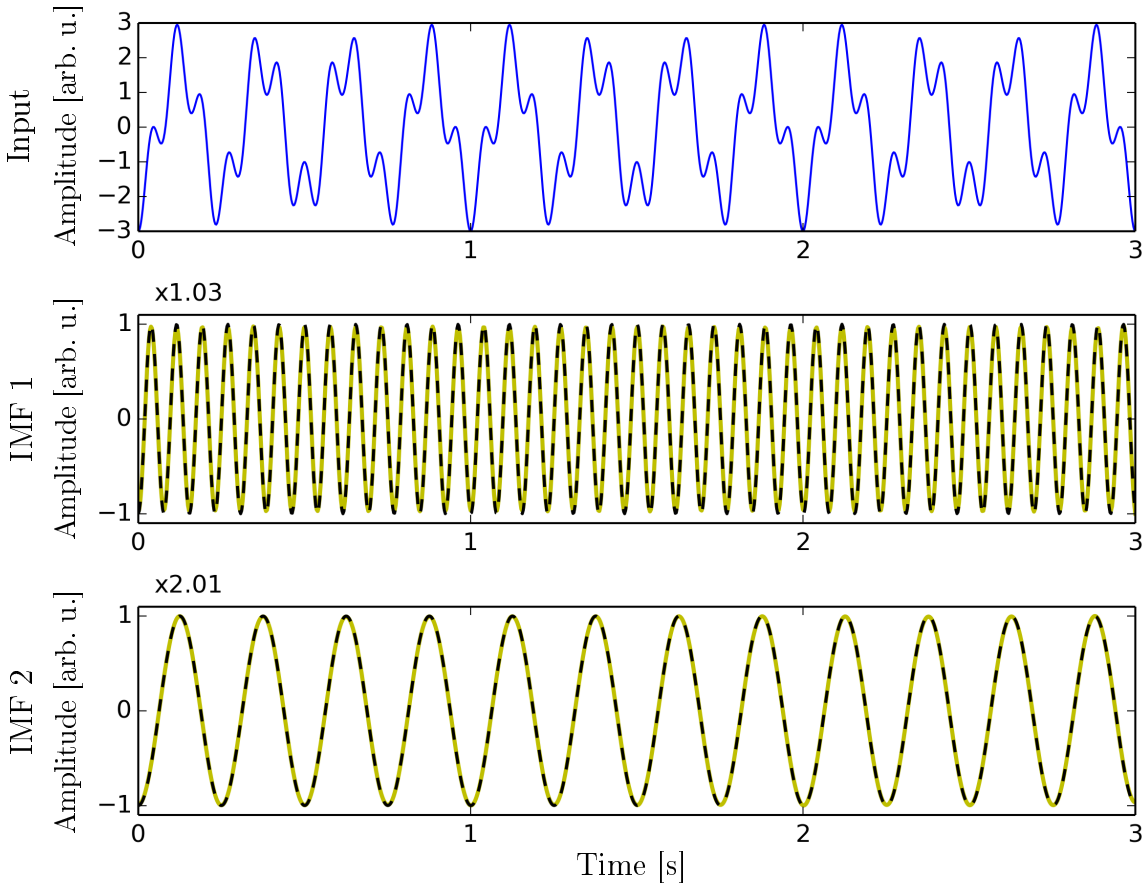


Figure 4.18: Top row presents synthetic signal generated with $f = 4$ Hz and the following rows are its EMD decomposition. IMFs (solid line) are scaled (scale in the top left corner) such that their maximum value is one. For comparison, dashed lines indicate cosine functions with constant amplitude and phase equal to IMF's instantaneous phases.

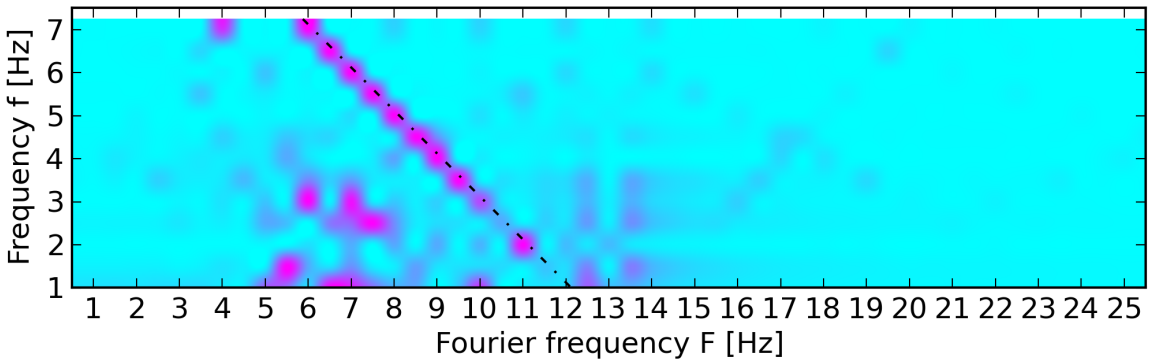


Figure 4.19: A depiction of Fourier spectra obtained for the first IMF for different values of frequency f (Eq. (4.18)). Each row relates to a different frequency f and presents Fourier spectrum with colour-coded amplitudes scaled such that the maximum is one. The dashed line which is going through the figure highlights trend which is given by the function $F = 13 - f$.

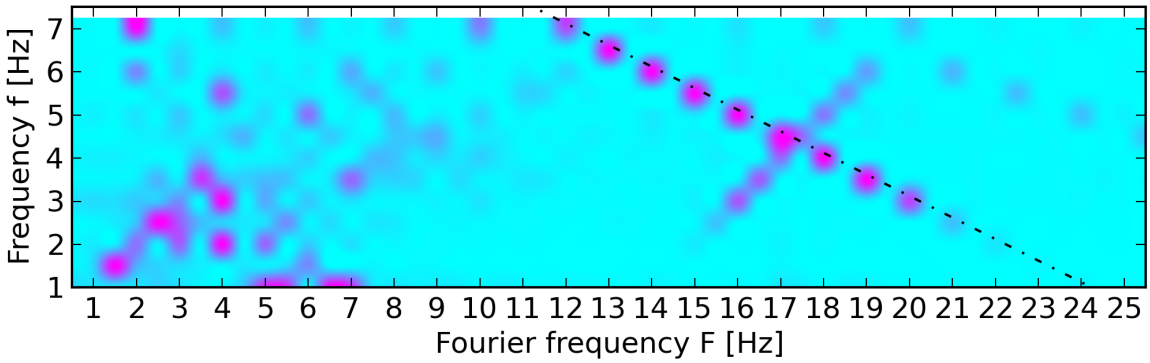


Figure 4.20: A depiction of Fourier spectra obtained for the second IMF for different values of frequency f (Eq. (4.18)). Each row relates to a different frequency f and presents Fourier spectrum with colour-coded amplitudes scaled such that the maximum is 1. The dashed line which is going through the figure highlights trend which is given by the function $F = 2 \cdot (13 - f)$.

The highlighted dependency between peaks of Fourier spectra and initial frequency f can also be observed when analysing cross-correlations between all pairs of IMFs' instantaneous frequencies. Such analysis emphasises modulations common for both components. Since there are only two IMFs, those modulations must be a product of their interaction, most likely a mixture of frequencies. An example of such mixing is shown in Fig. 4.21, where the top plots represent the instantaneous frequencies centred at zero by mean subtraction. In the same figure, the middle graph presents cross-correlation between the instantaneous frequencies of both IMFs, whereas the bottom graph shows the Fourier spectrum of the correlation signal. The vertical line indicates the value equal to the difference of the IMFs' mean frequencies, which for this example is $\Delta f = 13\text{Hz} - 4\text{Hz} = 9\text{Hz}$. Cumulative result for all pairs of IMFs is shown in Figure 4.22, where for each f the Fourier spectrum of the IMFs' correlations is plotted. Each spectrum was normalised so that the biggest value is one. This step allows for visual comparison of the results, as the maximum amplitude of cross-correlation depends on the f value and varies by a factor of 10^5 when comparing results for $f = 1$ Hz and $f = 7$ Hz. Again it can be observed that there exist two channels of peaks along lines $F_1 = 13 - f$ and harmonic $F_2 = 2 \cdot (13 - f)$. This suggests a strong coupling between the instantaneous frequencies of the IMFs at some f.

A quantitative attempt to explain visible effects was performed by fitting Kuramoto coupling model to the obtained instantaneous frequencies of IMFs. An example of the reconstruction is presented in Figure 4.23. On this graph, the left column contains the instantaneous frequency of the IMF (solid line) and the reconstructed one via Kuramoto model (dashed line). The right column shows the difference between the two instantaneous frequencies for each IMF, which are the first and second for the top and bottom rows, respectively. The values of obtained parameters and measure of fitness (Eq. (4.17)) are presented in Table 4.4. Parameters f_1 and f_2 relate to the intrinsic frequencies of Kuramoto model for the first and the second IMFs respectively. As it can be seen, they are relatively close matches to the input signal's modes. For small values of f, i.e. when there is a big difference between input components' frequencies, there is little coupling, i.e. k_1 and k_2 are small. Although coupling values seems to increase with f, they do not necessarily lead to a better reconstruction. Graphical depiction of

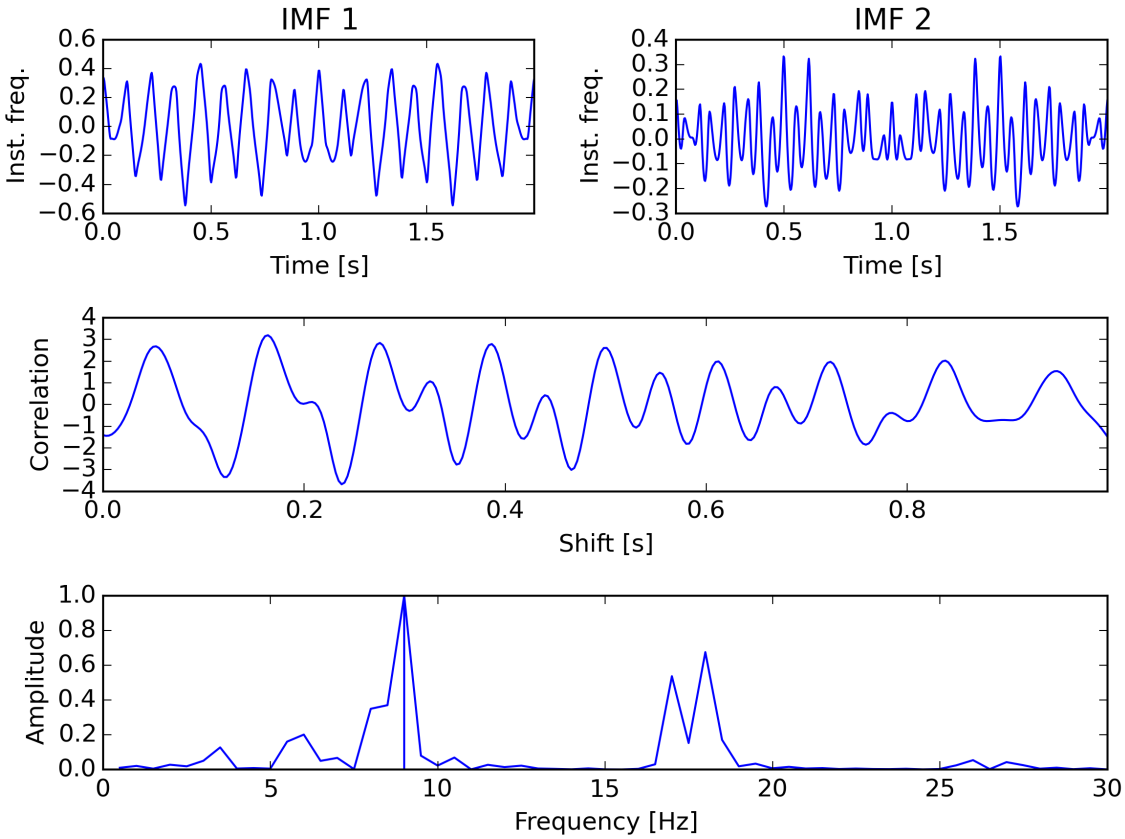


Figure 4.21: Correlation between IMFs' instantaneous frequencies centred at zero (case when $f = 4$ Hz). Top plots represent the instantaneous frequencies, central plot displays their cross-correlation, and the bottom graph shows Fourier spectrum of their cross-correlation. The vertical line marks value equal to the difference of IMFs' mean frequencies, i.e. $13 - 4 = 9$ Hz.

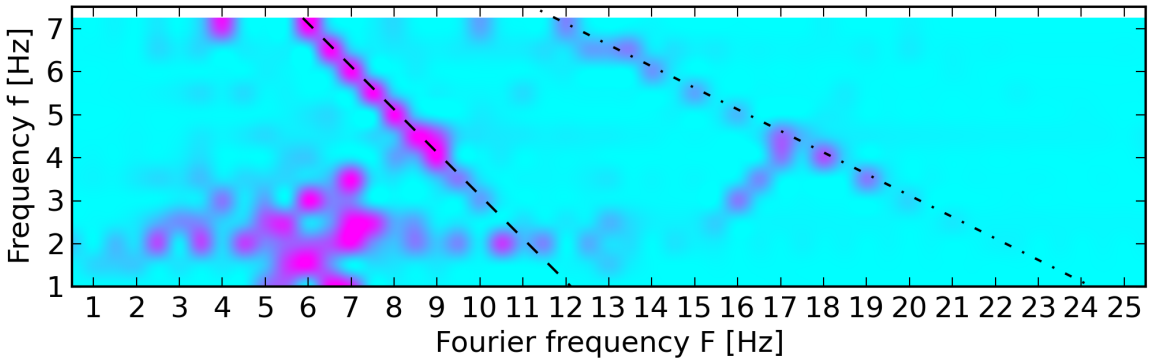


Figure 4.22: A depiction of Fourier spectra obtained for correlation functions between each pair of IMFs for different values of frequency f. The intensity of colour depicts value of amplitude which all were normalised, such that for given row frequency f the maximum amplitude is equal to one. A single horizontal slice for $f = 4$ is presented in Figure 4.21. This Figure is overlaid with two lines — $F_1 = 13 - f$ (dashed line) and $F_2 = 2 \cdot (13 - f)$ (dash-dotted line), which highlight the trend of observable peaks.

the dependencies between k values and the frequency f is presented in Fig. 4.24, where the left and right graphs represent $|k_1|$ and $|k_2|$ respectively. Additional variables M_r and M_w in Table 4.4 refer to the mean square error (Eq. (4.17)) of the Kuramoto model fit to IMFs' instantaneous frequencies when using coupling (M_r) and without coupling (M_w). The last column represents how much percentage-wise the reconstruction explains the variation. For $f \approx 6$ Hz the Kuramoto model has successfully explained more than 50% of the variation. However, for small values of f_2, i.e. where $f_2 \leq 3$ Hz, the mean square error has not decreased significantly. The reason is that IMFs almost perfectly match the input signal components and there is no need to include coupling k factors. A special case is $f = 2$ Hz for which coupling k_1 is relatively large and the error is decreased by 25%. It can be observed that for all examples, the coupling k_1 is bigger than k_2. This means that instantaneous frequency of the first IMF has more modulation proportional to the difference of the source's frequencies. This is in accordance with Figures 4.19 and 4.20, where it can be seen that dominant frequencies in instantaneous frequencies are $\Delta_{f_1} = 13 - f$ and $\Delta_{f_2} = 2 \cdot (13 - f)$ for the first and the second IMFs, respectively.

Table 4.4: Parameters obtained for Kuramoto's model to fit the data for different input modes. Frequencies f_1 and f_2 refer to the mean instantaneous frequencies, k_1 and k_2 are the coupling values, and M_r and M_w refer to fitness (see Eq. 4.17) with and without couplings. Indices 1 and 2 refer to the first and the second IMFs, respectively.

f [Hz]	f_1 [Hz]	f_2 [Hz]	$\lvert k_1 \rvert$	$\lvert k_2 \rvert$	M_r	M_w	$(M_w - M_r)/M_w$ [%]
1	13.000	1.000	0.001	0.000	0.020	0.020	0.01
1.5	13.013	1.501	0.031	0.000	0.069	0.069	0.07
2	13.013	2.002	0.696	0.004	0.250	0.337	25.88
2.5	13.013	2.502	0.003	0.000	0.081	0.081	0.10
3	13.013	3.003	0.046	0.009	0.082	0.084	2.81
3.5	13.013	3.504	0.180	0.020	0.076	0.097	21.48
4	13.001	4.000	0.273	0.031	0.087	0.133	34.96
4.5	13.013	4.504	0.509	0.027	0.150	0.227	33.96
5	13.005	5.011	2.225	0.023	0.300	0.950	68.41
5.5	13.027	5.511	2.162	0.178	0.364	0.878	58.50
6	13.030	5.989	5.687	0.036	0.700	2.203	68.24
6.5	13.006	6.492	2.411	0.014	0.442	0.962	54.07
7	13.044	7.021	4.930	0.164	2.285	3.280	30.35

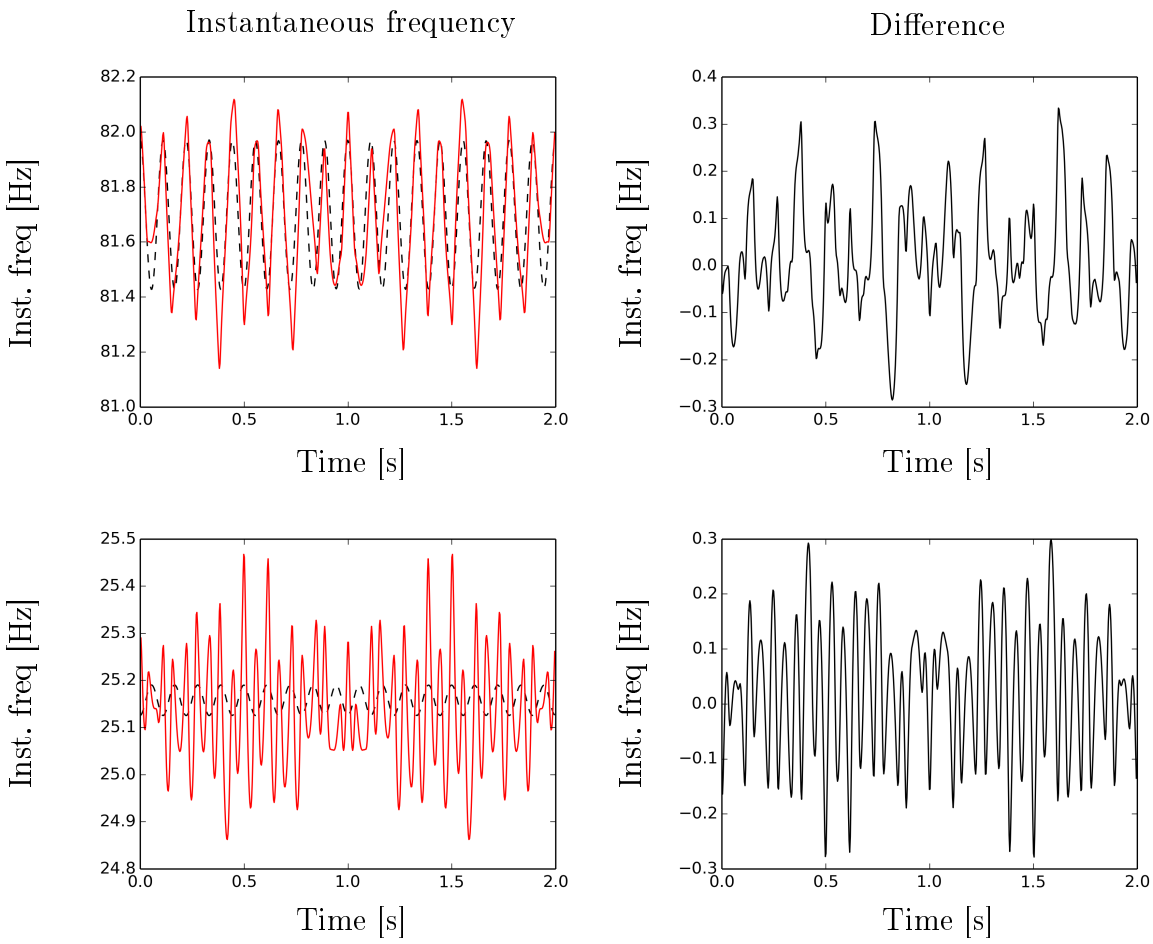


Figure 4.23: Comparison of instantaneous frequencies for $f = 4$. The left column contains IMF's instantaneous frequency (solid red line) and the reconstructed one (dashed line), whereas the right column shows their differences. Top and bottom rows correspond to first and second IMFs, respectively.

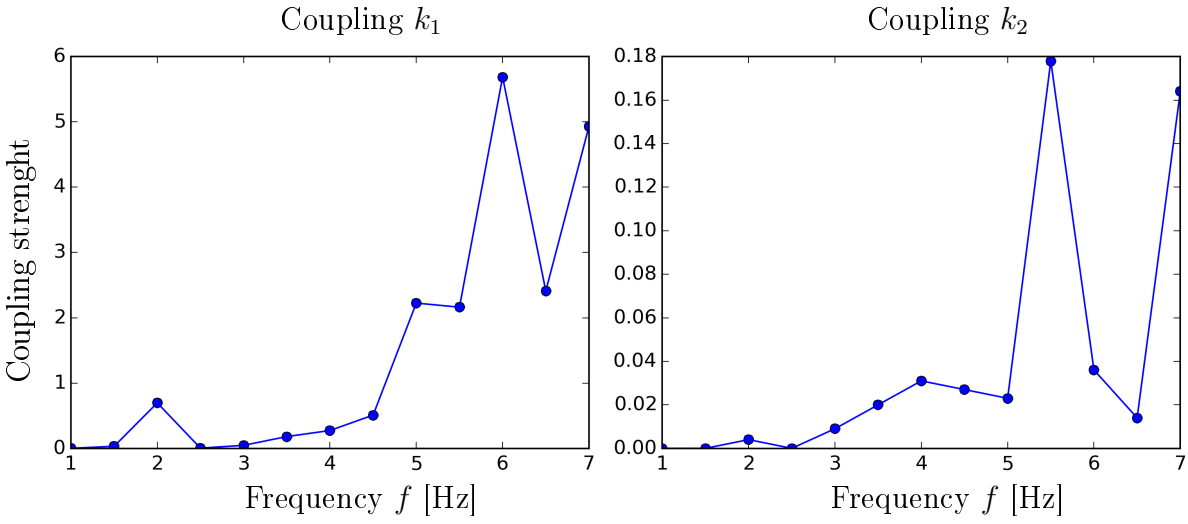


Figure 4.24: Relation between respective coupling values k and the frequency f.

4.3.3 Conclusion

As observed in all the experiments in Section 4.3.2, the instantaneous frequencies of each IMF have some modulations. These modulations, both their amplitudes and frequencies, seem to depend on the difference between the input's modes. This means that there is some interaction between the components, which depends on their frequencies and phases. An attempt to explain this frequency mixing was performed by assuming Kuramoto type phase coupling between the modes that is proportional to the sine of their differences. The results presented in Table 4.4 suggests that in many cases the fit was good. In some cases, however, including the coupling only reduced the error fit by 20%. This implies that there is more complex behaviour between the modes than a pure sine coupling. A possible solution would be to allow for additional components in Kuramoto's coupling function, e.g. including harmonic modulations in Eq. (4.15).

This study focused on simple signal composed of two sinusoidal oscillations. Although obtained IMFs were close to the input components, such behaviour is not expected in general [132, 135]. For more complex signals it is unlikely that the interaction would be only visible in phase domain. This suggests that more general model is required. Such model would not only describe couplings between phases but would further incorporate interactions between components' amplitude dynamics. The next chapter introduces a model with such properties.

4.4 Limitations and inspiration

The focus of this Chapter is on the analysis of EMD's properties. As it has been discussed many times and as the number of applications suggests this decomposition method is promising for extracting a finite set of components in general oscillatory form. Moreover, these components are supposed to be physically meaningful due to their non-negative instantaneous frequency and a single mode. Such properties are beneficial; however, as it has been shown, they might not hold for EMD in its current form which suffers from some limitations.

One of the disadvantages is that EMD is susceptible to small changes in signal. Modifying input's length can result in a different decomposition highlighting a different

feature set. Reasons for this are uncertain and depend on the actual definition of the EMD algorithm, but one of them is related to the boundary effect. The most popular EMD algorithm [17, 136] uses cubic splines that depend on all extrema points, and a mirroring technique for extrapolating at the boundaries. These conditions mean that any change close to the edges will shift extrema and thus affect the whole signal through iterative local mean extraction emphasising small changes regardless of the data precision format. Such susceptibility to modifications is even more pronounced when the signal is augmented with small in amplitude noise. Again, since envelopes are spanned by cubic spline adjustments to extrema position or value will intensify after a large number of iterations and propagate through all components. This event could also be considered in terms of frequency-mixing as discussed in the previous section. Since the noise does not have any structure, it is also considered to have all frequencies which will mix with modes of all components. Regardless of the reason, such effect is far from desirable. Ideally one would expect the method to be robust to small adjustments, preventing from changes in interpretation of the internal dynamics of a system.

EMD's main disadvantage is its lack of mathematical framework. Despite successful application in many fields [20, 47, 52–57], it is difficult to explain what the results represent precisely. Although there have been some attempts to formalise the algorithm and its outcome [21, 59, 60], it is still unknown what is the impact of the decomposition. It is difficult to objectively assess whether modifications and variations significantly improve the method as there is no foundation allowing for such validation. Nevertheless, many variations to EMD have been empirically proven to improve the method, suggesting that it is possible.

A high number of EMD applications suggests that under certain circumstances EMD produces meaningful results. Similarly, its popularity advocates for the existence of EMD-like method. All these limitations and demand from the research community lead us towards investigating the EMD in depth and understand areas of its possible improvement. The result of this investigation is a decision that a new method with robust mathematical framework has to be developed. This method would be inspired by the EMD and its philosophy to extract oscillations directly from data; however, it

will start with a general definition of an oscillation. The following part of this thesis describes construction and analysis of the EMD-inspired method for characterising a set of mutually interacting oscillators.

Chapter 5

The KurSL model

As it has been discussed in Chapter 2, there are limits to currently existing methods that aim at extracting oscillatory features. Those methods are typically either lacking strict mathematical framework — making their results difficult to interpret — or contain predefined mathematical features, unlikely to emerge in real systems. The core purpose of this research is to fill the gap in-between both ends of the mentioned spectrum.

One of the promising methods to extract oscillatory features is EMD. Its properties and purely data-driven approach make it an exciting attempt at data decomposition. In the previous chapter, EMD method and its properties were analysed. As it has been shown, its formal analysis is difficult to execute due to the lack of mathematical framework. Until now comparing results of EMD was performed via visual inspection. Although more objective measures for comparison were proposed [121] (sec. 4.2), it is still based on heuristics, which may not necessarily be true. Even effects like mode-mixing or frequency mixing cannot be explained without knowing what mechanism is behind it.

This chapter introduces a model, which is a framework for a *KurSL* method — the main result of this thesis. The model is based on two approaches: Kuramoto's coupling model (see Section 3.2) and Sturm-Liouville self-adjoint ordinary differential equation (ODE) (Section 3.1).

5.1 Joint model

Models, in general, have a particular purpose: they try to provide meaning to extracted data. A signal by itself can represent any phenomenon; it is the context that matters. Two models presented and discussed in Chapter 3 refer more specifically to oscillations. Sturm-Liouville equation (sec. 3.1) focuses on a single oscillator and tries to analyse its overall behaviour. It focuses on oscillation in a broad sense, explaining how amplitude changes over time. The coupling models (sec. 3.2), however, give quantitative analysis only of phase dynamics of mutually coupled oscillators. Despite being very valuable and used to explain many phenomena they are approximations, as it is impossible to observe an isolated oscillator or measure only phases. A combination of both approaches seems to be a natural extension. Sturm-Liouville eigenvalue problem introduces all possible forms of functions that have oscillatory property. Unfortunately, the family of solutions is too big to be analysed collectively. Addition of Kuramoto synchronisation model not only gives meaning to solutions but also restricts the number of possible solutions. This section describes a joint model of Sturm-Liouville theory and Kuramoto coupling model, which respectively are responsible for an amplitude- and frequency-modulated components.

Motivated by EMD results, we have focused on constructing method which would explain oscillatory signals. For simplicity of discussion and analysis, those oscillations can be described in forms of

$$y(t) = r(t)\cos(\phi(t)), \tag{5.1}$$

which is a product of two functions: an amplitude $r(t)$ and phase-related $\cos\phi(t)$ component. As mentioned earlier, SL describes oscillations in general form. Substituting (5.1) into SL equation (3.9), one obtains

$$\cos(\phi)\left(\ddot{r} + (Q - \lambda - \dot{\phi}^2)r\right) - \sin(\phi)\left(2\dot{r}\dot{\phi} + r\ddot{\phi}\right) = 0, \tag{5.2}$$

which due to mutual instantaneous orthogonality of sine and cosine functions leads to

two coupled equations

$$\begin{cases} 2\dot{r}\dot{\phi} + r\ddot{\phi} &= 0, \\ \ddot{r} + \left(Q - \dot{\phi}^2\right) r &= 0. \end{cases} \tag{5.3}$$

These equations describe a relationship between an amplitude and a phase for a given function Q. For simplicity of notation, in the equation (5.3) λ value was omitted since it can be treated as an offset for function the Q.

Another constraint is obtained by imposing phase dynamics through the Kuramoto system. In its simplest form, where the coupling weights are equal K, the relation between phases is introduced as

$$\dot{\phi}_i = \omega_i + \frac{K}{N} \sum_{j=1}^{N} \sin\left(\phi_j - \phi_i\right), \tag{5.4}$$

where indices refer to different oscillators from a set of N ($i, j \in \mathbb{N}_N$). Combination of both models leads to a coupled system of $3N$ equations in total: N for amplitudes r, phases ϕ and potentials Q, respectively. Their explicit forms are

$$(1) \qquad \dot{\phi}_i = \omega_i + \frac{K}{N} \sum_{j=1}^{N} \sin\left(\phi_j - \phi_i\right), \tag{5.5}$$

$$(2) \qquad 2\dot{r}_i\dot{\phi}_i + r_i\ddot{\phi}_i = 0, \tag{5.6}$$

$$(3) \qquad \ddot{r}_i + \left(Q_i - \dot{\phi}_i^2\right) r_i = 0, \tag{5.7}$$

where index i indicates that each set is for a single i^{th} oscillator.

5.1.1 Two oscillators

In case of two oscillators ($i \in \{1, 2\}$), the problem is well defined and can be solved analytically. Let intrinsic frequencies be ω_1 and ω_2 for respective oscillators. Expanding formula (5.4) for all oscillators leads to

$$\dot{\phi}_1 = \omega_1 + K/2\left(\sin(\phi_1 - \phi_1) + \sin(\phi_2 - \phi_1)\right) = \omega_1 + K/2\sin(\phi_2 - \phi_1) \quad , \tag{5.8}$$

$$\dot{\phi}_2 = \omega_2 + K/2\left(\sin(\phi_1 - \phi_2) + \sin(\phi_2 - \phi_2)\right) = \omega_2 - K/2\sin(\phi_2 - \phi_1). \tag{5.9}$$

These equations can be better presented by introducing variables $\mu = \phi_1 + \phi_2$ and $\nu = \phi_1 - \phi_2$, which redefine phases and frequencies as

$$\begin{aligned} \phi_1 &= \tfrac{1}{2}(\mu + \nu), \implies \dot{\phi}_1 = \tfrac{1}{2}(\dot{\mu} + \dot{\nu}), \\ \phi_2 &= \tfrac{1}{2}(\mu - \nu), \implies \dot{\phi}_2 = \tfrac{1}{2}(\dot{\mu} - \dot{\nu}), \end{aligned} \tag{5.10}$$

Such operations allow for general simplification of coupled equations. Adding equations (5.8) and (5.9) side-by-side one obtains

$$\dot{\mu} = \omega_1 + \omega_2, \tag{5.11}$$

and the subtraction of (5.9) from (5.8) leads to

$$\dot{\nu} = \omega_1 - \omega_2 + K \sin \nu. \tag{5.12}$$

Such defined problems have easy solutions. Equation (5.11) is a time-independent function to which the solution is simply

$$\mu(t) = (\omega_1 + \omega_2)t + \mu_0, \tag{5.13}$$

where $\mu_0 = \mu(0)$ is the initial value. Additionally, it can be shown, that solution to equation (5.12) is

$$\nu(t) = 2 \arctan\left(\frac{\beta \tan\left(\frac{1}{2}\beta t\right)}{\omega_1 - \omega_2} + \nu_0\right), \tag{5.14}$$

where $\beta^2 = (\omega_1 - \omega_2)^2 - K^2$. Returning to the initial forms for phases, ϕ_i, they can be represented as

$$\phi_1(t) = \arctan\left(\frac{\beta \tan\left(\frac{1}{2}\beta t\right)}{\omega_1 - \omega_2}\right) + \frac{\omega_1 + \omega_2}{2}t + \theta_1, \tag{5.15}$$

and

$$\phi_2(t) = -\arctan\left(\frac{\beta \tan\left(\frac{1}{2}\beta t\right)}{\omega_1 - \omega_2}\right) + \frac{\omega_1 + \omega_2}{2}t + \theta_2, \tag{5.16}$$

where θ_1 and θ_2 are initial phase values.

A solution for an amplitude can be sought by using a modified form of equa-

tion (5.6). Rewriting it into a form of $2\dot{r}_i(t)/r_i(t) = -\ddot{\phi}_i(t)/\dot{\phi}_i(t)$ it can be easily integrated on both sides leading to

$$2\log(r_i(t)/\rho_i) = -\log(\dot{\phi}(t)/\dot{\phi}_{i0}), \tag{5.17}$$

or in different form

$$r_i^2(t) = \frac{\rho_i^2\dot{\phi}_{i0}}{\dot{\phi}_i(t)}, \tag{5.18}$$

in both cases $\rho_i = r_i(0)$ and $\dot{\phi}_{i0}$ are i^{th} oscillator's initial amplitude and initial instantaneous frequency, respectively. Using Kuramoto's equation (5.4) and substituting phases with calculated form gives

$$r_1(t) = \rho_1\sqrt{\frac{\omega_1 + \frac{1}{2}K\sin(\theta_2-\theta_1)}{\omega_1 + \frac{1}{2}K\sin\left(2\arctan\left(\frac{\beta\tan\left(\frac{1}{2}\beta t\right)}{\omega_1-\omega_2}\right)\right)}}, \tag{5.19}$$

and

$$r_2(t) = \rho_2\sqrt{\frac{\omega_2 - \frac{1}{2}K\sin(\theta_2-\theta_1)}{\omega_2 - \frac{1}{2}K\sin\left(2\arctan\left(\frac{\beta\tan\left(\frac{1}{2}\beta t\right)}{\omega_1-\omega_2}\right)\right)}}. \tag{5.20}$$

With obtained equations for phases, i.e. (5.15) & (5.16), and amplitudes, i.e. (5.19) & (5.20), solving for the function Q is a straightforward operation by substituting amplitude and phase expressions into equation (5.7).

5.1.2 N oscillators

Due to Kuramoto's coupled nature, when considering more than two oscillators, the KurSL model needs to be solved numerically. However, a simplification can be made to present whole set dependent only on a single function. It can be shown that (5.6) and (5.7) can be transformed into simpler forms either dependent on the amplitude $r(t)$

$$\dot{\phi}_i(t) = \frac{r_{i0}^2\dot{\phi}_{i0}}{r_i^2(t)}, \tag{5.21}$$

$$Q_i(t) = -\frac{\ddot{r}(t)}{r(t)} + \frac{r_0^4\dot{\phi}_0^2}{r^4(t)}, \tag{5.22}$$

or the instantaneous frequency $w = \dot{\phi}$,

$$r_i^2(t) = \frac{\rho_i^2 w_{i0}}{w_i(t)}, \tag{5.23}$$

$$Q_i(t) = -\frac{\ddot{w}_i}{w_i} + \frac{3}{4}\frac{\dot{w}_i^2}{w_i^2} + w_i^2, \tag{5.24}$$

in both cases ρ_i and $\dot{\phi}_{i0} = w_{i0}$ are the initial values of the amplitude and the instantaneous frequency, respectively. Since synchronisation in phase cannot be simplified, the canonical representation of the system is dependent on the phase function, i.e.

$$w_i = \omega_i + \sum_{j=1}^{N} k_{ij} \sin(\phi_j - \phi_i), \tag{5.25}$$

$$r_i^2(t) = \frac{\rho_i^2 w_{i0}}{w_i(t)}, \tag{5.26}$$

$$Q_i(t) = -\frac{\ddot{w}_i}{w_i} + \frac{3}{4}\frac{\dot{w}_i^2}{w_i^2} + w_i^2. \tag{5.27}$$

Such a coupled set of ODEs fully describes the KurSL system. Recall that each oscillator is composed of an amplitude and a phase-related function (5.1). With this one can present component in a form dependent only on a phase, i.e.

$$y_i(t) = \rho_i \sqrt{\frac{\dot{\phi}_{i0}}{\dot{\phi}_i(t)}} \cos(\phi(t)), \tag{5.28}$$

or in an expanded version

$$y_i(t) = \frac{\rho_i \sqrt{\dot{\phi}_{i0}} \cos(\phi(t))}{\sqrt{\omega_i + \sum_{j=1}^{N} k_{ij} \sin(\phi_j - \phi_i)}}, \tag{5.29}$$

which depends on all the initial values and the coupling strengths k between all the oscillators.

5.2 Model's properties

5.2.1 Method classification

One of the applications for the KurSL model is as a signal decomposition method. In essence, when applied to signal, the method will try to find parameters for which the reconstruction will match the input closely. Specificity of the method will be discussed in the next chapter; here are presented the properties of oscillators.

As noted in the previous section, the result of the method is a simple superposition of all oscillators. This, however, does not necessarily mean that the method is linear (see Appendix A.1). Only in the particular case when all coupling strengths k are zero, i.e. there is no coupling between oscillators, the method imitates Fourier series and is treated as linear. Otherwise, adding new oscillator with any $k_{ij} \neq 0$ will affect the whole reconstruction with effect dependent on other components. Despite having a mathematical framework, coupled ODE form makes it difficult to determine the impact of each parameter analytically. Such complexity forces the method to be treated explicitly defined by parameters.

5.2.2 The KurSL example

To discuss some properties of the method and its components, in this subsection few examples are produced numerically. The primary focus is to present behaviour of KurSL's components in time and frequency domains.

A few time series were generated in the performed experiment. The difference between consecutive executions was the number of oscillators used for generating oscillations. The first example has two oscillators, the second three and the third has four oscillators. Table 5.1 presents all parameters used in the experiment. However, not all were used for each experiment. Bounded regions indicate which parameters were used for which simulation. Moreover, for each experiment, a visual representation of obtained results was constructed. They are presented in Figures 5.1, 5.2 and 5.3 for experiments with 2, 3 and 4 oscillators, respectively. Each figure is composed of two columns: the left side contains time series whereas the right side has a Fourier spectrum normalising the highest value to be 1. The order of components is from the

Table 5.1: Parameters used in the example in all experiments. Corner boxes indicate which parameters were chosen for each experiment, with the first having only two oscillators and the third experiment using four oscillators. Values k_{ij} indicate coupling strengths between respective oscillators as indicated by row and column ordinals.

n	$\omega/2\pi$ [Hz]	θ [rad]	ρ [arb. u.]	k_{ij} [$\frac{\text{rad}}{\text{s}}$]			
				1	2	3	4
1	2	1.2	1.0		2.0	1.5	0.5
2	5	2.5	1.5	-2.0		-7.0	4.1
3	10	0.0	1.0	2.2	3.2		3.8
4	13	2.0	1.0	0.1	10.0	0.0	

smallest intrinsic frequency on top to the biggest on the bottom. One should notice that although first two components for all experiments are always used with same values, their time series differ significantly. This can be easily noticed either by studying their amplitude's modulations in time series or based on the width of frequency peak. If there were no coupling, one would expect time series to have a constant amplitude and a single delta Dirac like peak in Fourier spectrum. Interestingly, the KurSL model can produce more than one distinct peak in component's spectrum. This means that in addition to the main frequency, there are distinct patterns of repetition. Those additional peaks seem to be located around the main peak, with the distance approximate to all possible frequencies' f differences. For example, the third oscillator with frequency $f = 10$ Hz has additional peaks in locations $\nu = \{2, 5, 7, 13, 15\}$ Hz, which can be explained as $10 \pm \{|10 - 2|, |10 - 5|, |10 - 13|\}$ Hz.

In addition, the cumulative representations, both in time and frequency domains, are presented in Figures 5.4a and 5.4b, respectively. As expected, all central frequencies are visible in the spectrum. However, those peaks are more spread than it would be expected if there were no couplings between oscillators. Such behaviour shows that it is possible to obtain complex signals with only a few coupled oscillators. The complexity of these components is also visible in the spectrogram 5.5 which was computed using Tukey window with tapering parameter $\alpha = 0.25$ used on 1 second windows with 90% overlap. It can be observed that all components are modulated in frequency through

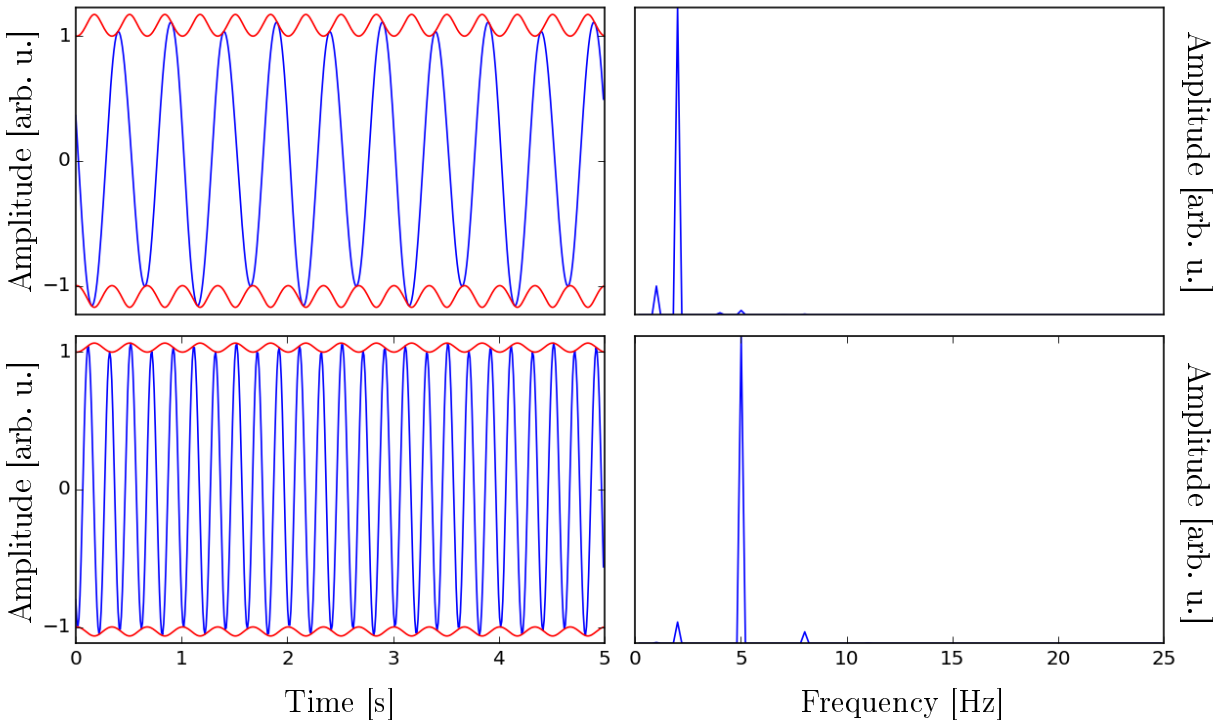


Figure 5.1: Simulation of KurSL model assuming 2 oscillators and parameters according to the table 5.1. The left column displays all components with their amplitudes in red, whereas the right column has respective component's Fourier transformation normalised to the highest value being 1.

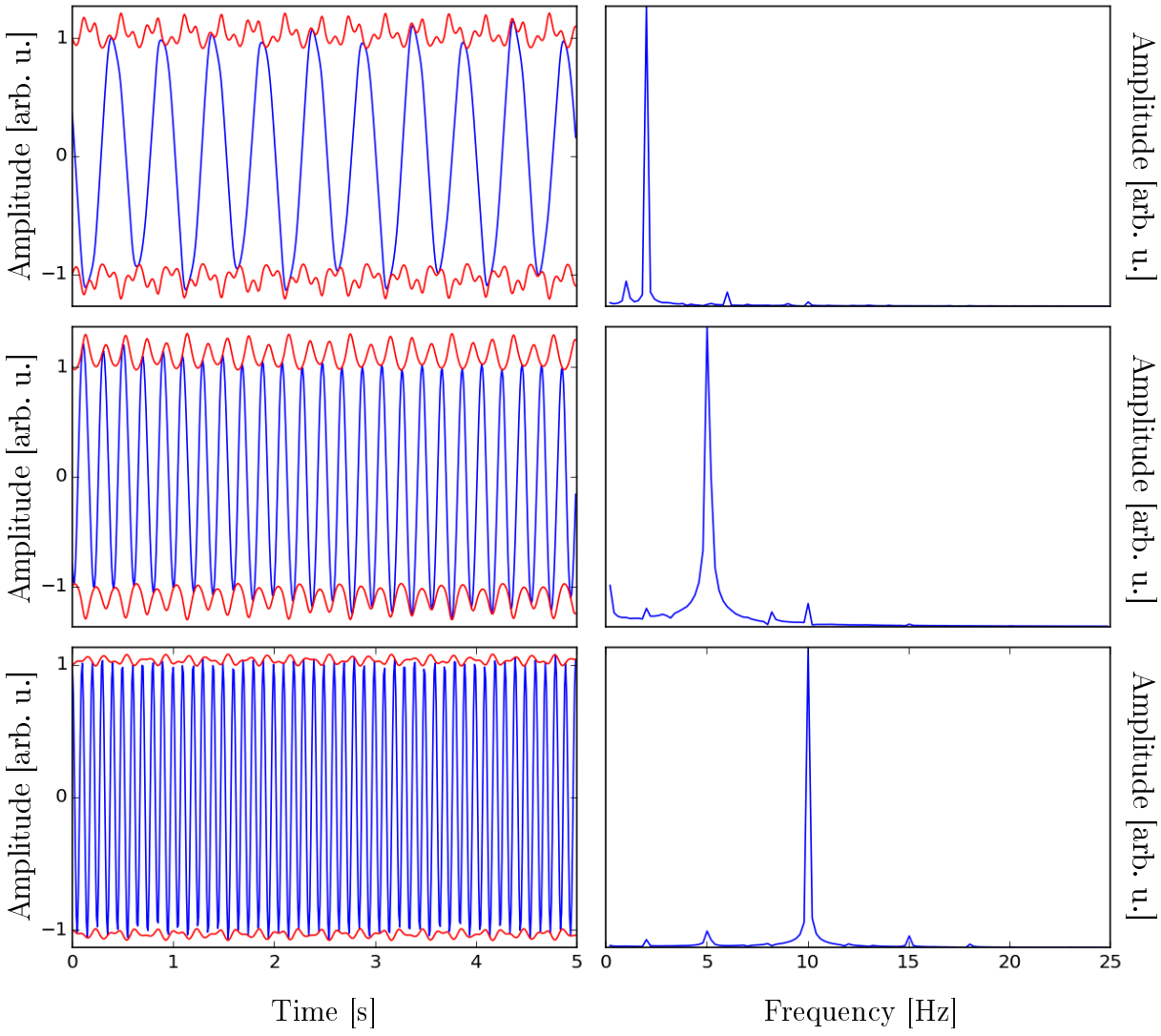


Figure 5.2: Simulation of KurSL model assuming 3 oscillators and parameters according to the table 5.1. The left column displays all components with their amplitudes in red, whereas the right column has respective component's Fourier transformation normalised to the highest value being 1.

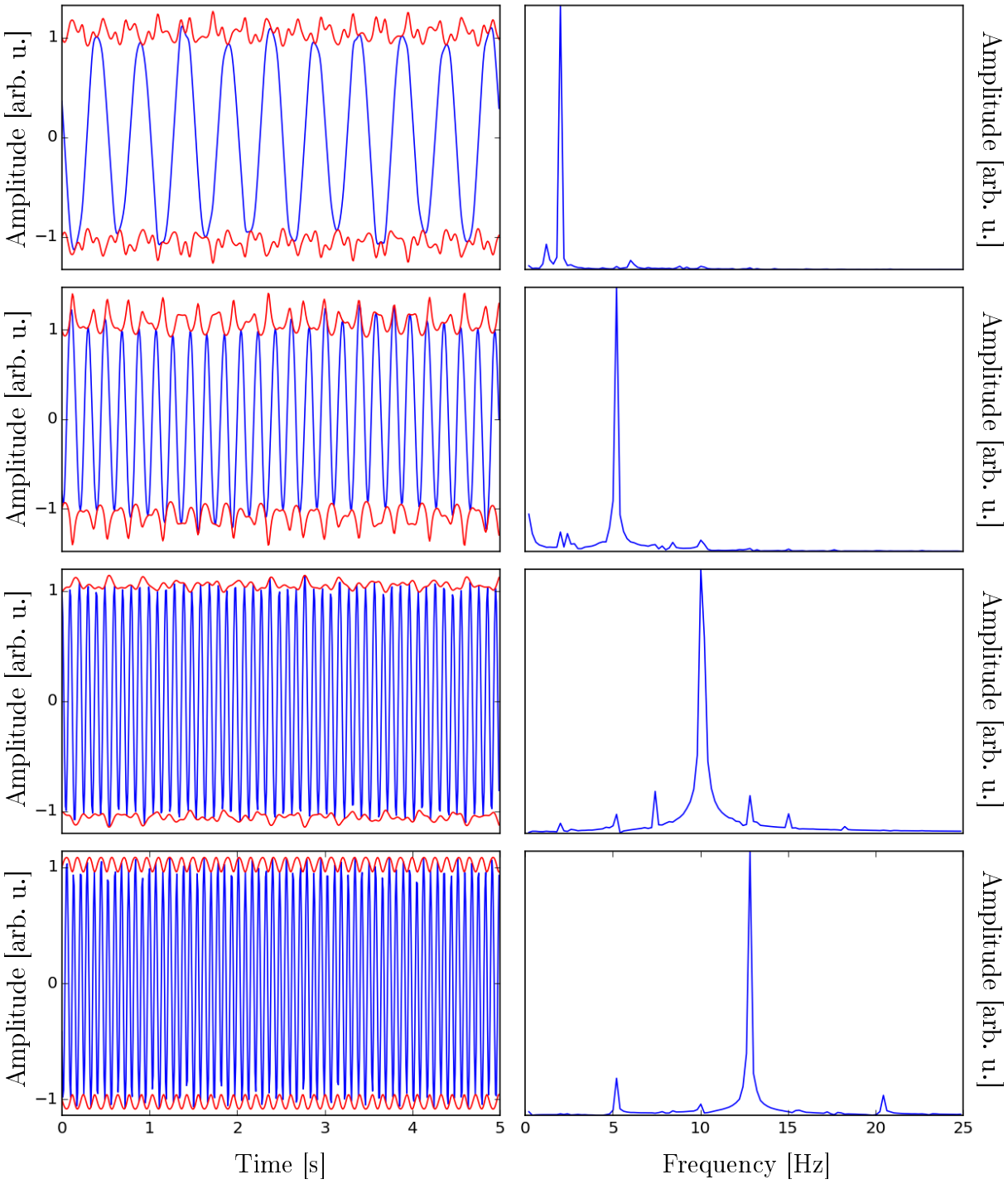
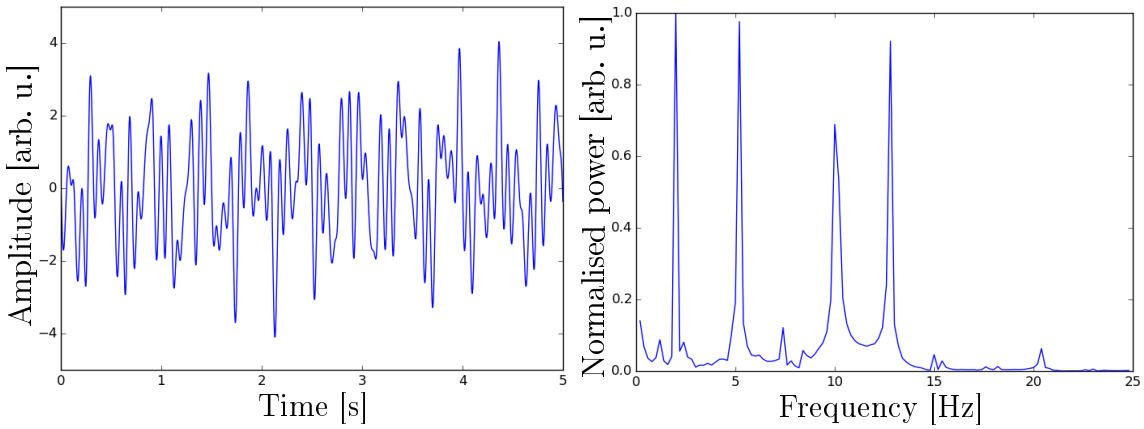


Figure 5.3: Simulation of KurSL model assuming 4 oscillators and parameters according to the table 5.1. The left column displays all components with their amplitudes in red, whereas the right column has respective component's Fourier transformation normalised to the highest value being 1.



(a) Time series

(b) Fourier spectrum

Figure 5.4: Time (a) and frequency (b) domain representations of a collective sum for all generated components using parameters from Tab. 5.1. In case of the Fourier spectrum all values were normalised such that the highest peak has amplitude one.

a coupling with other oscillators. The interaction is especially visible between the two oscillations with the highest frequencies which seem to connect periodically. Such result is due to indirect coupling; although both oscillators are loosely connected (Table 5.1) they have high coupling strength with oscillator $n = 2$. In addition, this component has more significant variation towards higher frequencies. These methods could help to estimate the KurSL parameters and once that is done the KurSL can obtain more detailed representation.

5.2.3 Meaningful instantaneous frequency

Although there is no universally agreed definition of what can be considered as a physically meaningful component, many researchers have been defining it in terms of instantaneous frequency [21, 135, 137, 138]. More strictly, they are connecting physical meaningfulness to non-negativity of the instantaneous frequency. Such definition is often troublesome since the evolution of the instantaneous frequency depends on the method which extracted it. Some researchers [139, 140] have strongly advocated for using an analytic signal via Hilbert transform to estimate the instantaneous frequency properly. Unfortunately, dealing with discrete signals of a finite length, as is the case with empirical data, introduces some artefacts. In comparison, KurSL by definition has always positively defined instantaneous frequency making components physically

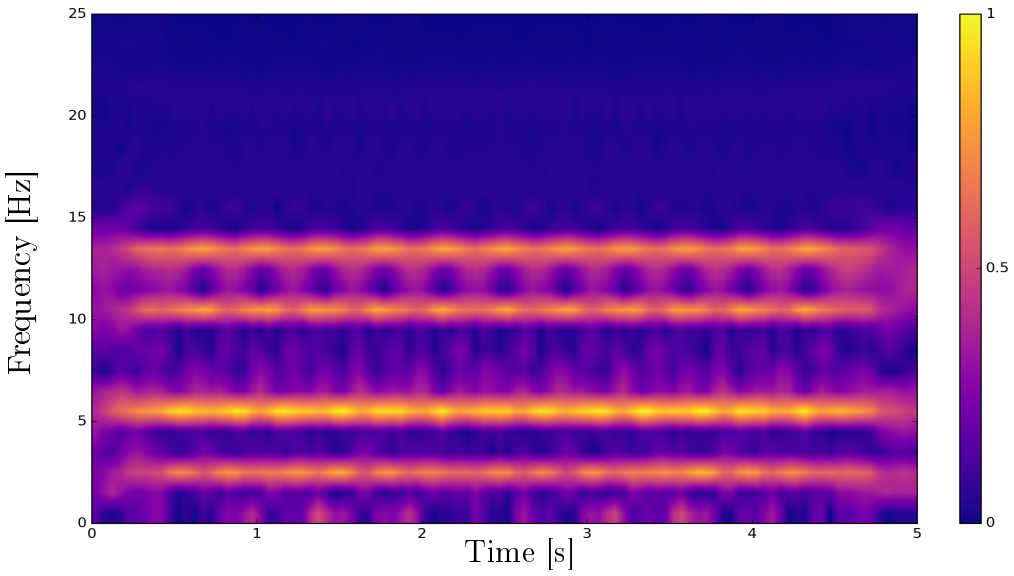


Figure 5.5: Spectrogram of a collective sum for all generated components using parameters from Tab. 5.1. It was computed using Tukey window with tempering parameter $\alpha = 0.25$ of length 1 s and 0.95% overlap.

meaningful. This can be explicitly observed in component's defining equation, i.e.

$$c_i(t) = \rho_i \sqrt{\frac{\dot{\phi}_{i0}}{\dot{\phi}_i(t)}} \cos(\phi(t)), \tag{5.30}$$

where the instantaneous frequency scaled by the initial value appears in the denominator of a square root. Although this allows the frequency to start negative, the direction of "rotation" is just a convention and one can simply revert the notation. Regardless, this condition prevents frequency from reaching value zero, and thus forbids *oscillation death* phenomenon [141]. It additionally forces the coupling function to be always below the respective value of the intrinsic frequency ω. This imposes inequality constraint on all coupling values

$$\sum_{j=1}^{N} k_{ij} \sin(\phi_j - \phi_i) \leq \omega_i, \ \ \forall t, i \tag{5.31}$$

which is true if and only if

$$\sum_{j=1}^{N} k_{ij} \leq \omega_i, \ \ \forall i \tag{5.32}$$

for all oscillators.

A special situation is when the oscillator has zero frequency all the time, i.e. in the equation (5.5) the instantaneous frequency $\dot{\phi}_i = 0$. This results in a single constraint

for the amplitude as

$$\ddot{r}_i + Q_i r_i = 0, \tag{5.33}$$

which is an equation for oscillation under the force Q_i and thus can be further decomposed using substitution (5.1). The only difference is in the semantics as they would refer to amplitude's "amplitude-" and phase-modulations.

5.2.4 On Bedrosian theorem

One of the useful features of the EMD is its ability to extract an instantaneous phase of the signal. Such extraction typically is performed through the Hilbert transformation with careful consideration of the Bedrosian theorem (see Appendix A.3). For a component to fulfil related conditions, its amplitude, $r(t)$, and phase related function, $\cos\phi(t)$, must have disjoint Fourier spectra. Such requirement translates to the dot product of Fourier spectra, i.e.

$$d = \mathcal{F}(r) \cdot \mathcal{F}(\cos\phi), \tag{5.34}$$

to yield $d = 0$ for the Bedrosian theorem to hold. Fulfilment of such condition, however, does not hold for the KurSL method as it can be seen in Figure 5.6 which presents Fourier spectra of oscillators amplitude component, $r(t)$, and phase related function, $\cos\phi(t)$. These components were obtained using parameters from Table 5.1, except for amplitudes ρ which increased threefold. For ease of comparison, functions were detrended and scaled such that the largest value of either component's spectrum is 1. A quantitative measure of the overlap, d (Eq. 5.34), is presented in the top right corner of each graph.

Figure 5.6 shows that even in this simple example Bedrosian condition is not fulfilled. Nevertheless, such result does not decrease meaningfulness of the KurSL method. Bedrosian theorem is essential in the context of EMD, where it would allow for a correct extraction of an instantaneous frequency. In the KurSL case, however, the instantaneous frequency can be precisely calculated provided initial conditions for the model and thus these conditions are not necessary.

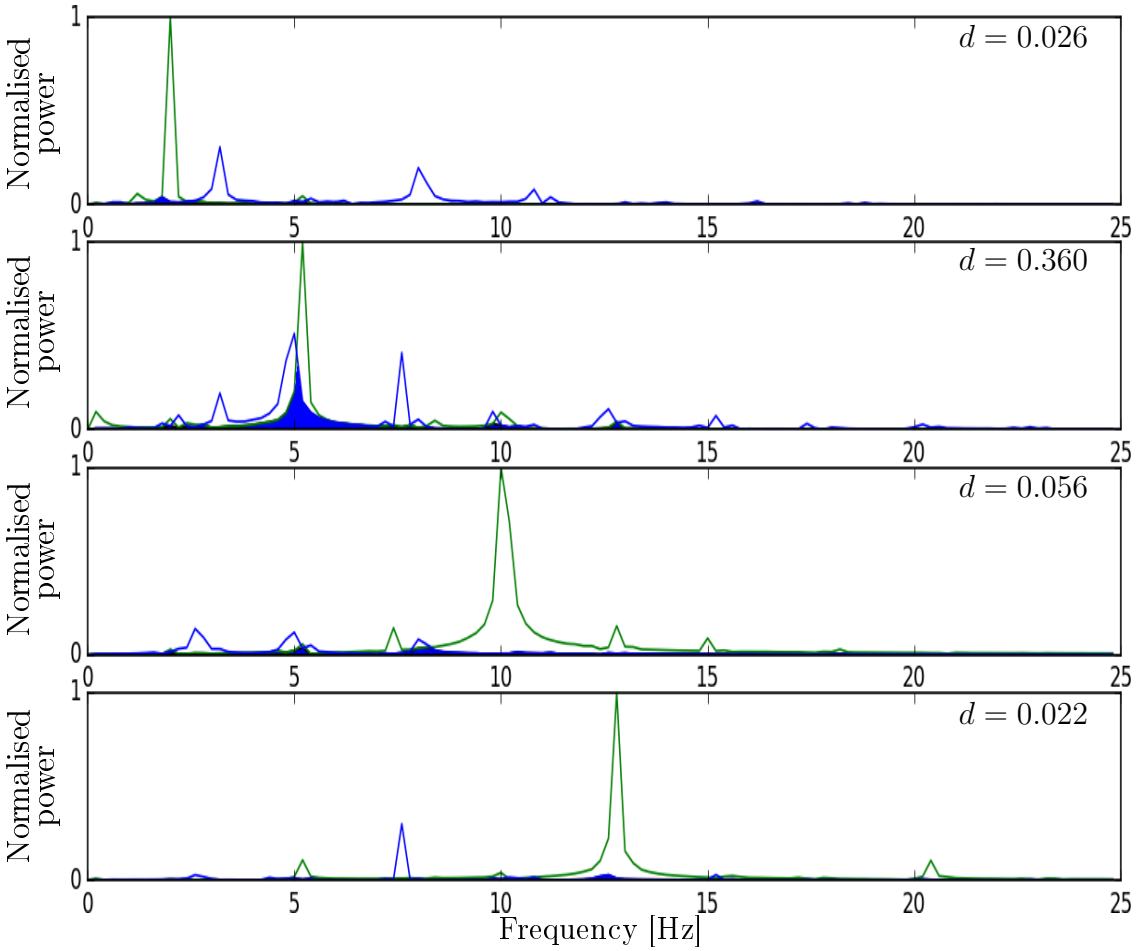


Figure 5.6: Normalised Fourier spectra of amplitude $r(t)$ and phase related component $\cos\phi(t)$ presented in blue and green, respectively. Components were created using parameters from Tab. 5.1 except for amplitudes ρ which were three times larger. Spectra are presented in decreasing order of intrinsic frequency with the top having. Coloured areas indicate where components are overlapping with the metric d (Eq. 5.34) presented in the top right corner.

5.3 The Mth order model

In the simplest form the Kuramoto model assumes coupling between each oscillator as a single sine function, i.e.

$$K(\phi_i, \phi_j, t) = k \sin(\phi_j(t) - \phi_i(t)), \tag{5.35}$$

where $\phi.$ are phases for corresponding oscillators and k is the scaling factor. However, this function does not need to have a form of a single sine. As discussed in Section 3.2 the coupling function can be of any periodical form. This means that any coupling function can be presented in the form of Fourier series. In case of series consisting of N components, one can formulate it as

$$K_M(\phi_i, \phi_j, t) = \sum_{m=1}^{M} k_m \sin\left(m(\phi_j(t) - \phi_i(t))\right), \tag{5.36}$$

where k_m indicates the strength of m^{th} harmonic component. Case for $m = 0$ is omitted as it refers to no coupling between oscillators — $K_0 = 0$.

The Kuramoto model with M coupling components can be treated as the Kuramoto model of the M^{th} order. Likewise, KurSL that utilises Kuramoto's M^{th} order model can be considered as the KurSL model of the M^{th} order. This modifies equations (5.5), (5.6) and (5.7) into

$$(1) \qquad \dot{\phi}_i = \omega_i + \sum_{m=1}^{M} \sum_{j=1}^{N} k_{ij}^m \sin\left(m(\phi_j - \phi_i)\right), \tag{5.37}$$

$$(2) \qquad 2\dot{r}_i\dot{\phi}_i + r_i\ddot{\phi}_i = 0, \tag{5.38}$$

$$(3) \qquad \ddot{r}_i + \left(Q_i - \dot{\phi}_i^2\right) r_i = 0, \tag{5.39}$$

which differs in the phase function from the model introduced in Section 5.1. Such definition means that the model of an order M, $KurSL_M$, incorporates all possible models up to its order. The transition from order M to $M' < M$ is performed by setting all intermediate coupling factors $k^m = 0$, where $m \in (M', M]$. It follows that a solution space of M^{th} order model, $\mathcal{S}_M$, is a subspace of solution space of any higher

Table 5.2: Parameters used in all experiments in Section 5.3 using third order model. Corner boxes indicate which parameters were chosen for each experiment, with the first having only two oscillators and the third experiment using 4 oscillators. The index of coupling strength indicates by oscillator number with omitting itself. Columns K,L and M refer to first, second and third harmonic, respectively.

i	$\omega/2\pi$ [Hz]	θ [rad]	ρ [arb. u.]	K_1	L_1 $[\frac{\text{rad}}{\text{s}}]$	M_1	K_2	L_2 $[\frac{\text{rad}}{\text{s}}]$	M_2	K_3	L_3 $[\frac{\text{rad}}{\text{s}}]$	M_3
1	2	1.2	1.0	2.0	0.5	-1.2	1.5	0.2	-0.4	0.5	1.8	0.2
2	5	2.5	1.5	-2.0	3.5	-3.3	-7.0	2.1	4.2	4.1	1.2	1.2
3	10	0.0	1.0	2.2	0.5	1.7	3.2	0.3	0.1	3.8	-2.2	6.0
4	13	2.0	1.0	0.1	1.9	1.9	10.0	-2.1	10.7	0.0	9.1	-1.5

order model, $\mathcal{S}_M \subset \mathcal{S}_{M+m} \forall m \in \mathbb{N}$. In this thesis, unless specified otherwise, KurSL is assumed to be in its simplest form, i.e. of order 1.

For demonstratory purpose an experiment similar to the one presented in 5.2.2, where the first order KurSL was presented. This experiment uses the KurSL of order 3 for which the parameters are contained in Table 5.2. Those parameters were chosen to match the ones in the previous example where possible. Results for 2, 3 and 4 oscillators are presented in Figures 5.7, 5.8 and 5.9, respectively. Moreover, Figure 5.10a presents consolidated results for the time domain and in the Figure 5.10b for the frequency domain. Unsurprisingly, KurSL with a higher order has more complex structure. This is especially visible when comparing spectra with the same number of oscillators. Those produced with the order 3 have broader main peaks and additional variability in the whole spectrum.

5.4 Conclusion

This chapter introduced the KurSL model. The purpose of the model is to describe systems with oscillators that communicate with each other. In the KurSL model, the oscillators are assumed to be harmonic, i.e. when isolated from the system they would oscillate with a constant frequency. However, due to mutual interaction with other oscillators, their behaviour in amplitude and phase dynamics are modified accordingly.

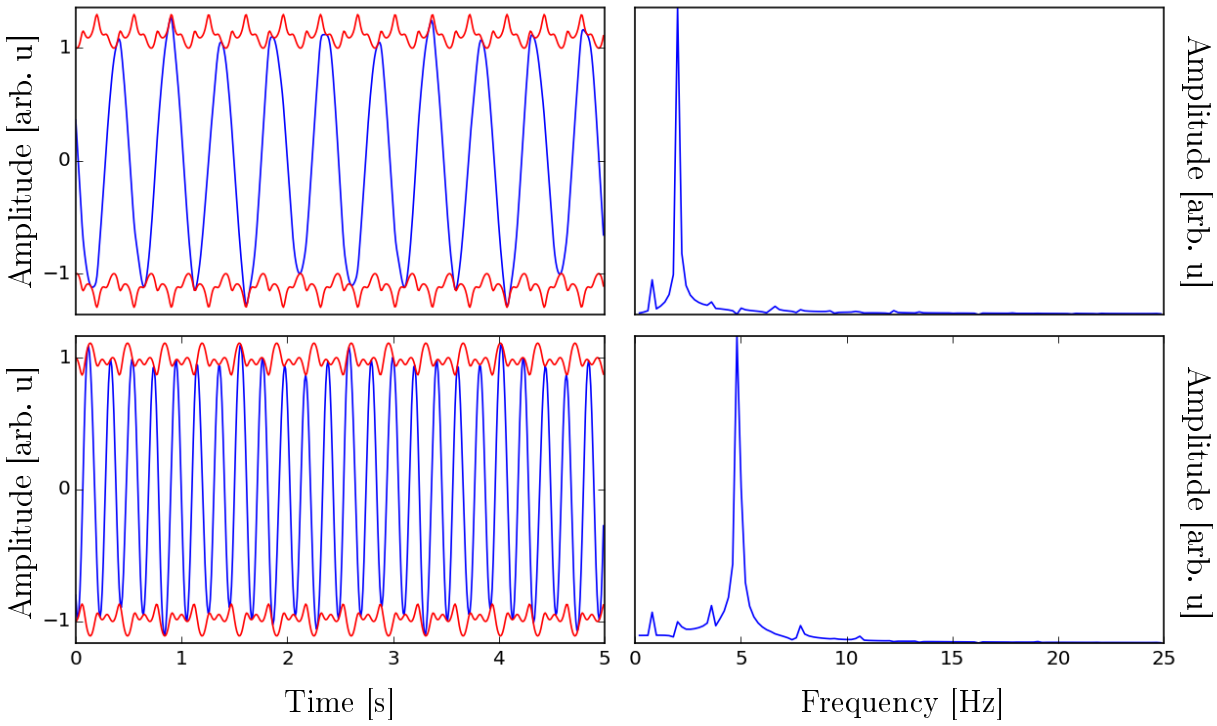


Figure 5.7: Simulation on KurSL model of order 3 assuming 2 oscillators and other parameters according to the table 5.2. The left column displays all components with their amplitudes in red, whereas the right-hand side has respective component's normalised Fourier transformation.

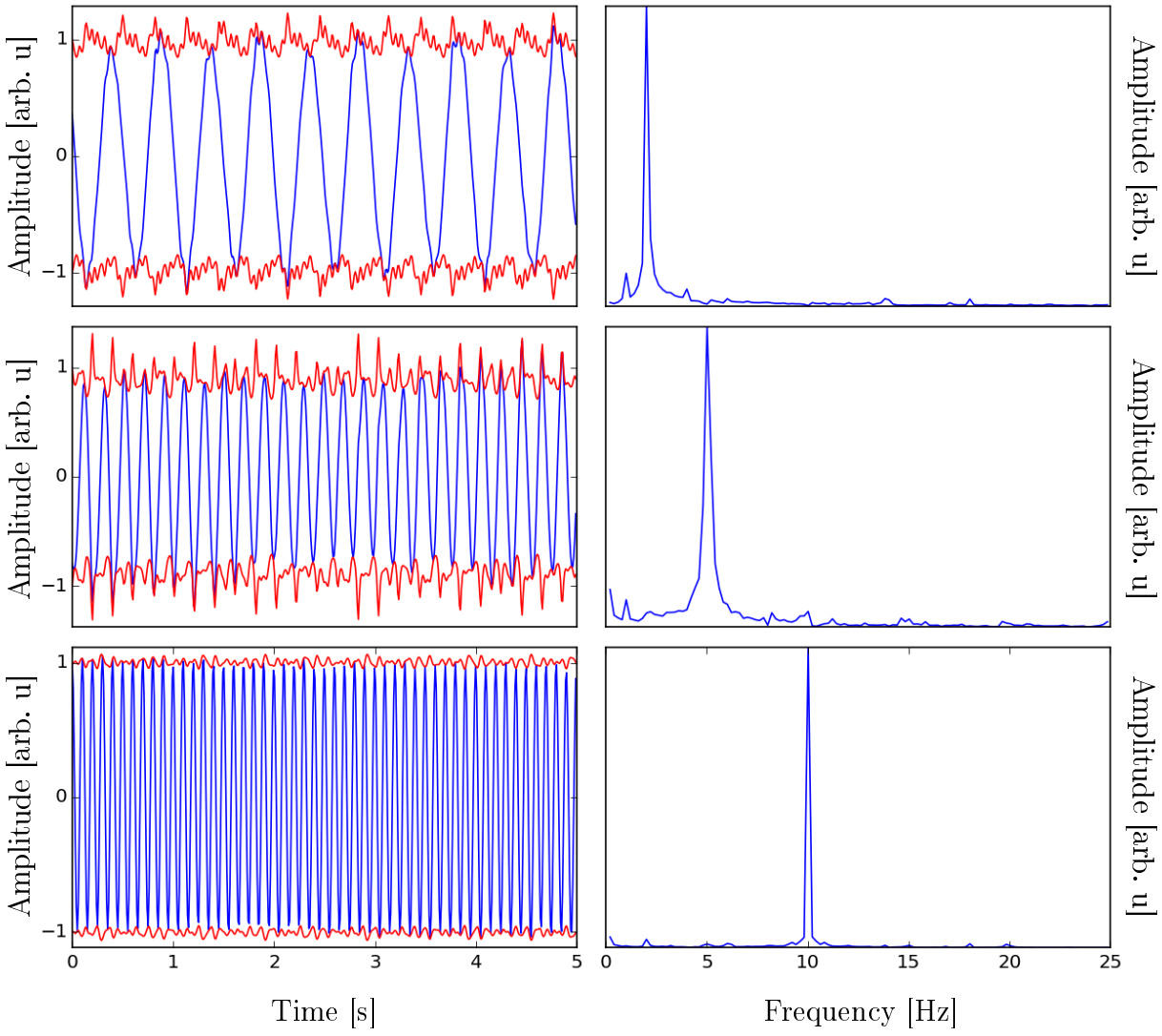


Figure 5.8: Simulation on KurSL model of order 3 assuming 3 oscillators and other parameters according to the table 5.2. The left column displays all components with their amplitudes in red, whereas the right-hand side has respective component's normalised Fourier transformation.

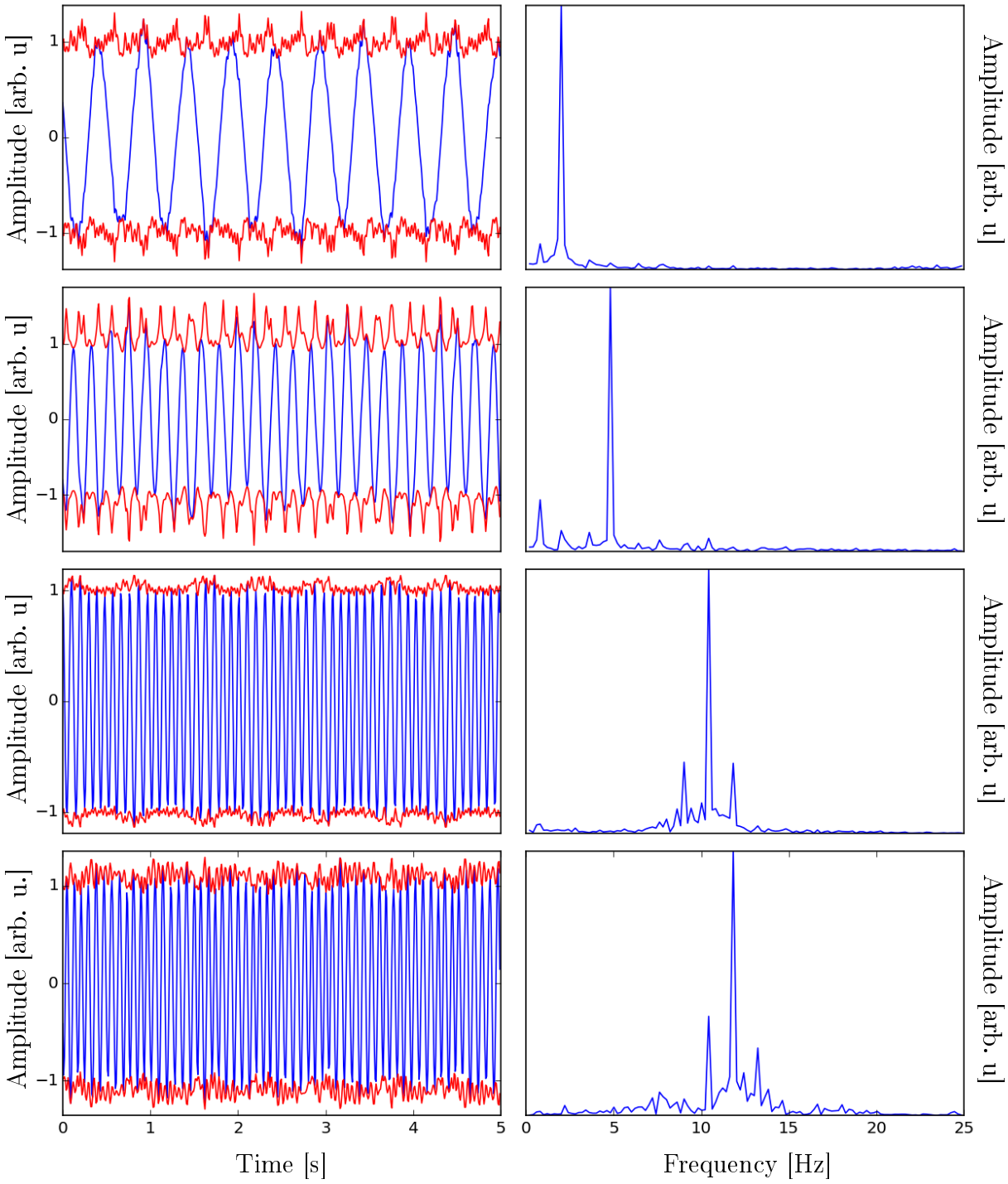
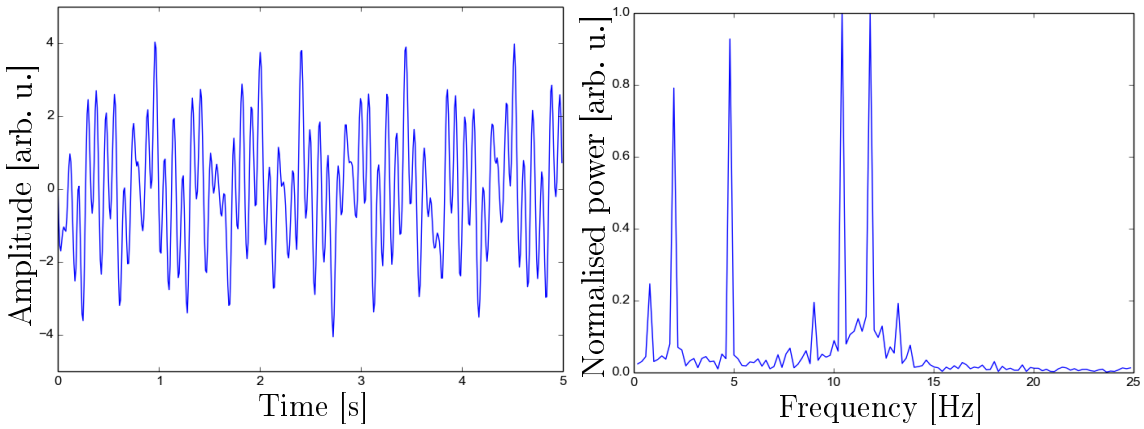


Figure 5.9: Simulation on KurSL model of order 3 assuming 4 oscillators and other parameters according to the table 5.2. The left column displays all components with their amplitudes in red, whereas the right-hand side has respective component's normalised Fourier transformation.



(a) Time series

(b) Fourier spectrum

Figure 5.10: Time (a) and frequency (b) domain representations of a collective sum for all generated components using parameters from Tab. 5.2. In case of the Fourier spectrum all values were normalised such that the highest peak has amplitude one.

These modulations are explained by incorporating a Kuramoto coupling model and Sturm-Liouville oscillation theory which mathematically is described in terms of coupled ordinary differential equations. According to this model, the dynamic of the whole system is described by both the initial state, as well as coupling functions which definition can be dependent on pair and directionality of oscillators. Such definition allows for variability in these properties leading to a general amplitude- and phase-modulated components. Additionally, the model ensures that components will always have physically meaningful characteristics such as positively defined instantaneous frequency at any time.

The interaction between any two objects is defined by order of the model and the strength values of coupling factors. However, since these oscillators are not isolated from others, their dynamics are also affected by the environment. This means that there is a non-unique outcome of the coupling process. Some of the effects can be observed in the Fourier spectrum as a widening of the peak functions or shifting their centres respective to the interactions. Moreover, as references Figs. 5.1 to 5.3 suggest, the coupling between oscillators can introduce modulations that are visible in the spectrum as small satellite peaks. It has been observed that their height is proportional to the absolute coupling strength between oscillators, and their frequency locations f can be estimated as a combination of coupled oscillators' locations, i.e. $|f_i \pm f_j|$.

An estimate of the complexity of the model is through the number of parameters which is mainly impacted by the number of oscillators, $\mathcal{N}$. Each new oscillator in the system adds a degree of freedom for all existing oscillator, thus scaling the complexity quadratically with the number of oscillators, $O(\mathcal{N}^2)$. Another property that can affect the model's complexity is the form of assumed interaction. Depending on the type of oscillators and their surrounding, the coupling functions can be of different forms. The variety of these has been denoted in the model by its order, M, which refers to the highest non-zero term of the Fourier series of the coupling function. Such expansion has a linear $O(M)$ impact on the complexity. Therefore, the KurSL model is fully described by stating its order M and the number of oscillators $\mathcal{N}$ in the system. Under these conditions, there are $n = \mathcal{N}(3+M(\mathcal{N}-1))$ parameters spanning the model. This also means, that in general, the complexity of the model follows $O(M\mathcal{N}^2)$ asymptotic.

Chapter 6

The KurSL method

The model presented in Chapter 5 describes a general system of coupled oscillators. Such definition means that this model can characterise any signal with oscillatory components. The KurSL method, which utilises the KurSL model, focuses on extracting oscillatory information from a signal by finding the best fitting model. As a result, a set of coupled oscillators with fully defined amplitudes and phases are obtained.

This Chapter introduces the KurSL method. First, Section 6.1 proposes a method for model's parameters estimation. The primary technique used for parameter fitting is the Markov Chain Monte Carlo (MCMC, Sec. 6.1.1). This technique uses Bayesian statistics by providing estimates for parameters' probability distributions based on a large number of simulations. An algorithm for obtaining initial conditions for the process is described in Section 6.1.2, and the complete algorithm for the KurSL method is presented in Section 6.2. The Chapter finishes with an analysis of the method. Section 6.3 discusses the convergence property of the method, whereas Section 6.4 focuses on the relation between the time domain and the parameter space.

6.1 Determining parameters of the model

The KurSL model in its standard form, i.e. the first order KurSL, requires adjusting $\mathcal{N}(\mathcal{N}+2)$ parameters, where $\mathcal{N}$ is the number of oscillators and is also an adjustable parameter. In general, fitting M^{th} order model requires finding a parameter set $\mathbf{p}_{\gamma_N} = \{\vec{\omega}, \vec{\theta}, \vec{\rho}, \mathbb{K}\}$, where the frequency, $\vec{\omega}$, the initial phase, $\vec{\theta}$, and the initial amplitude, $\vec{\rho}$,

vectors are of length $\mathcal{N}$, and $\mathbb{K}$ is the coupling matrix of size $\mathcal{N} \times M(\mathcal{N}-1)$. This means that $n = \mathcal{N}(3+M(\mathcal{N}-1))$ values fully define the M^{th} order model. This section presents techniques with which these parameters are determined.

6.1.1 Markov Chain Monte Carlo

Markov Chain Monte Carlo (MCMC) is an optimisation method for finding a set of parameters and their probability distribution functions (pdfs) [142]. Its name is derived from a combination of two methods: a *Markov chain* and a *Monte Carlo* simulation. Markov chain characterises a random process, where the transition to the next state depends only on the current state, regardless of its history. Commonly it refers to a discrete process with a discrete time step, and in such cases, it is called a discrete-time Markov chain (DTMC). The Monte Carlo process refers to a broad class of numerical simulation algorithms. The development of the method was inspired by an observation that for some problems it is easier and quicker to perform simulations than doing a thorough mathematical analysis of the system [143]. Most Monte Carlo methods compute results based on a randomly generated samples and then infer properties of obtained distributions. Despite the method being computationally demanding, it is used with problems where the analytical solutions are difficult to determine, such as in numerical integration, optimisation and sampling probability distributions [3].

Markov Chain Monte Carlo as a combined method fully utilises both components. It is often used to sample unknown probability distribution. Typically this is performed in an iterative manner where each next sample depends on some previous sample's states. Prior knowledge of the system is reflected in stating prior probability distributions for each parameter. These parameters then evolve in sampling process to best describe desired probability distributions.

There are many different MCMC methods of which the most popular are Metropolis-Hastings [144, 145] and Gibbs sampling [146]. The method used in this algorithm is called *ensemble sampler with affine invariance* and was proposed by Goodman & Weare [142]. The method was selected due to its excellent general performance and being unaffected by affine transformations of space. These conditions are preferable for inappropriately scaled distributions, which are possible in KurSL case due to its

possibly highly nonlinear behaviour. The implementation of this MCMC method is called *emcee Hammer* and is freely available from [147]. The method utilises the so-called *walkers*, which are particular points a parameter space. These walkers change their location depending on a position of other walkers, their previous states and their current state. Since the result is dependent on the communication and cumulative exploration of the parameter space, there should be a relation between the number of walkers and their initial positions. It is advised [142] to have at least twice as many walkers as the number of parameters and generate at least hundred movement iterations. A fraction of the initial iterations is often referred as a *burn-in* period. Since the walkers are just starting to learn the space, such results can be meaningless and it is suggested to discard this period. Walkers' position should eventually converge to a real probability distribution for which the *maximum a posteriori* (MAP) estimates the best fit of parameters.

In case of the KurSL, MCMC assumes that the input signal $S(t)$ can be explained as a sum of all KurSL components $c(t)$ and some additional noise

$$S(t) = \sum_{i=1}^{\mathcal{N}} c_i(t) + \xi(t), \tag{6.1}$$

where $\xi(t)$ is a Gaussian noise and all $\mathcal{N}$ components $c_i(t)$ sum up to $c(t)$. In case of discrete signals with constant sampling rate where $t_n = t_0 + n\Delta t$ and $c_i(t_n) = c_{ni}$ the equation (6.1) modifies to

$$S_n = \sum_{i=1}^{\mathcal{N}} c_{ni} + z_n, \tag{6.2}$$

where z_n denotes a discrete sample from a Gaussian distribution. Considering that each z_n is an independent and identically distributed sample, it is possible to write a joint probability distribution function as

$$P\left(\{\mathbf{z}_n\}\right) = \prod_{n=0}^{N-1} \frac{d\mathbf{z}_n}{\sqrt{(2\pi\Delta_t)^L |\hat{\mathbf{D}}|}} \exp\left(-\frac{1}{2\Delta_t}\mathbf{z}_n^T \hat{\mathbf{D}}^{-1}\mathbf{z}_n\right), \tag{6.3}$$

where in this case $\hat{\mathbf{D}}$ is an identity matrix and $\mathbf{z}_n$ is a noise vector with individual

values being

$$z_n = S_n - \sum_{i=1}^{\mathcal{N}} c_{ni}. \tag{6.4}$$

Considering such probability is equivalent to minimising L_1 error

$$E = \sum_{n=0}^{N} \left| S_n - \sum_{i=1}^{N_{osc}} r_{ni} \cos(\phi_{ni}) \right|. \tag{6.5}$$

Calculating negative log-likelihood function $\mathcal{L}_z = -\log\left(P\left(\{\mathbf{z}_n\}\right)\right)$ of Eq. 6.3 one obtains

$$\mathcal{L}_z = \frac{1}{2\Delta_t} \sum_{n=0}^{N-1} |\mathbf{z}_n|^2 + \frac{NL}{2} \log\left(2\pi\Delta_t\right), \tag{6.6}$$

where only the first term depends on the parameters, thus being the only significant part.

An additional step has to be performed for coupling values. Due to the requirement imposed by inequality (5.32) sum of all couplings $K_i = \sum_j k_{ij}$ has to be smaller than the respective intrinsic frequency ω_i. Thus, at each iteration parameters are checked for violation of this condition. If the sum of couplings exceeds ω_i, then each coupling value k_{ij} is scaled by a factor $s = 0.95\omega_i/K_i$, which results in new coupling values $\hat{k}_{ij} = sk_{ij}$ and their sum $\hat{K}_i = 0.95\omega_i$.

6.1.2 Determination of priors for MCMC

As mentioned in Section 6.1.1, MCMC requires *a priori* probability distribution functions (pdfs) for the walkers to initiate their positions. If none is available or the model can not infer their shape, one can assume a uniform distribution. Such solution is, however, suboptimal and one should take advantage of any information available. Studying examples provided in Section 5.2, e.g. Figures 5.1 to 5.3, one can infer some relations between parameters and the observed behaviour. Very suggestive is the structure of component's Fourier spectrum. There is a relation between spectrum's mode positions and the value of intrinsic frequency ω. Moreover, the spread of spectrum's peak is correlated with the sum of all coupling values for a given oscillator. As for the value of amplitude parameter, one can utilise Parseval's identity [148], which states that the total energy calculated in time and frequency domains are equal. For simplicity and

as a first approximation, one can assume that the component is a simple harmonic oscillator. Such hypothesis implies that the amplitude of a peak in the spectrum is a squared value of the actual amplitude r.

Based on these observations, an algorithm to determine the number of components and their initial parameters is proposed. The algorithm 1 is based on input signal's Fourier spectrum. It iteratively tries to fit a predefined peak function (triangular, Gaussian or Lorentz) to the highest peak and then subtract it. It does that until specified ratio E_ϵ of residual energy to the initial signal's energy is obtained, or until the number of obtained oscillators is larger than a predefined maximum. Sometimes subtracting fitted peak can result in a power spectrum with negative values. This artefact is dealt with by converting at each iteration all negative values to zero.

The default configuration of the algorithm uses for peak definition a Gaussian function with a standard deviation of 1 Hz. Unless stated otherwise the algorithm uses energy ratio $E_\epsilon = 0.1$ and no limitation on the number of oscillators. It was empirically determined that such configuration performs oscillation detection well on most tested signals. However, as with many threshold parameters, these values should not be expected to give the best possible results all the time and they should be adjusted appropriately depending on input signal's features.

Visualisation of the main part of the algorithm is presented in Figure 6.1. For a given signal (top graph), the algorithm tries to fit a peak function to the maximum of its spectrum (second row). Each column on the figure relates to different peak function type, i.e. triangular, Gaussian and Lorentz type of peaks, respectively from left. Red colour lines represent fitted peak functions. Subtracting the fitted peak (third row) could leave some regions of the spectrum with negative values. Such situations are dealt with by replacing negative spectrum values with zero (bottom row).

Selected optimisation method, MCMC, requires pdf for each parameter and thus once all initial values are determined, a joint probability function needs to be created. Here it is assumed that except for coupling strengths k, all pdfs are Gaussian with their mean values equal to initial values. Only phase variance is fixed to be $\mathrm{Var}(\theta) = 1$ for which 99.8% of the distribution lays within the range $[-\pi, \pi]$ which covers the whole phase domain. The variance of amplitude and intrinsic frequency are dependent

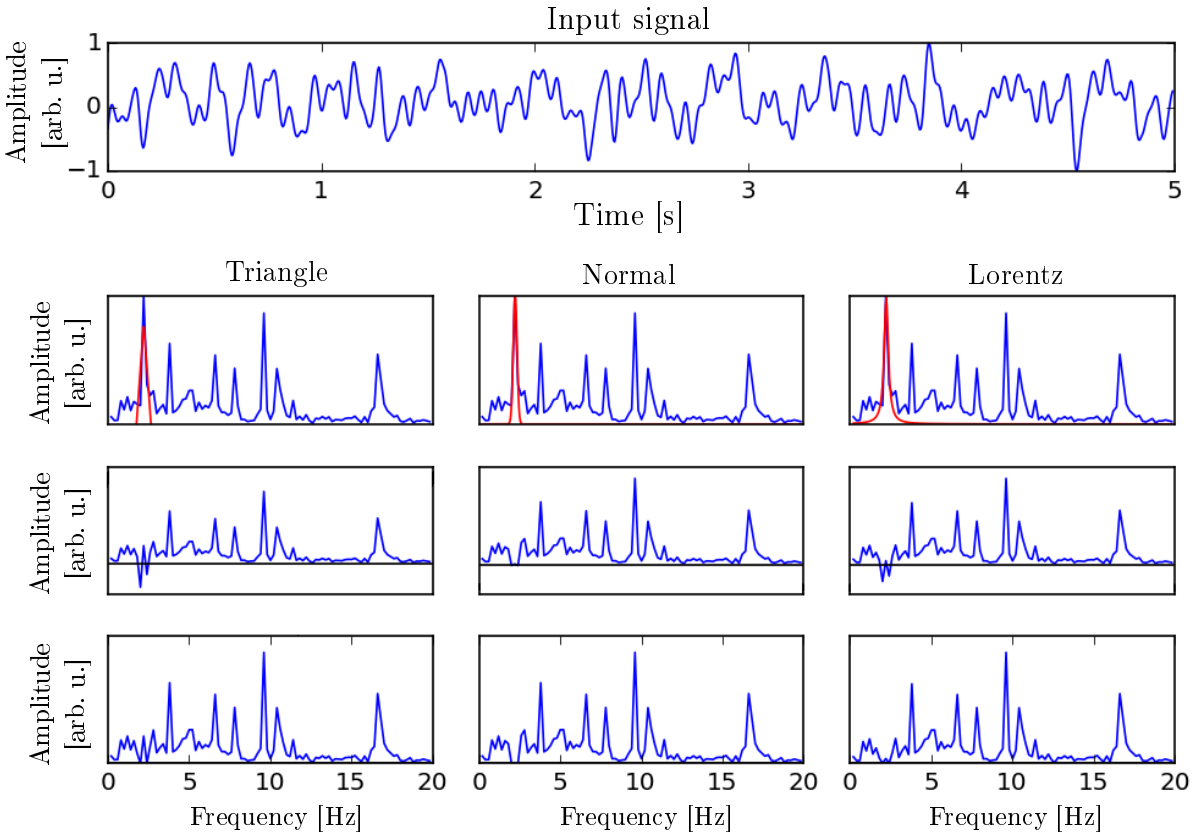


Figure 6.1: Illustratory example of peak removal algorithm. Columns refer to different type of a fitted peak, starting from left being triangular, Gaussian and Lorentz types of peaks. Red function on the graph represents a fitted peak. Graph below denotes spectrum with subtracted peak, whereas the bottom figures have normalised negative parts.

Data: Signal $S(t)$

Result: Estimated number of oscillators and a set of initial parameters

Compute Fourier spectrum of input signal S;

Calculate energy E_S of the spectrum;

Assign residual energy $E_{res} = E_S$;

Specify energy target ratio E_ϵ and maximum number of oscillators $N_{\max}$;

while *Ratio* $E_{res}/E_S > E_\epsilon$ & *number of oscillators* $\mathcal{N} < \mathcal{N}_{\max}$ **do**

    Find position and amplitude of residue's spectrum $\mathcal{F}_{res}(S)$ maximum;

    Fit peak of preselected type spectrum at given position;

    Subtract the peak function from residual spectrum;

    Assign zeros to any negative part of spectrum;

    Calculate energy E_{res} of residual spectrum;

    Calculate the ratio $E = E_{res}/E_S$;

end

Algorithm 1: Algorithm for initial estimation of KurSL parameters, i.e. finding the number of oscillators $\mathcal{N}$ and initial values.

on initial values and they are $\mathrm{Var}(\rho) = 0.1\rho$ and $\mathrm{Var}(\omega) = \log(\omega + 1)$, respectively. Such assignment is due to the fact, that the algorithm is typically more precise in determining the value of frequency ω rather than amplitude ρ.

The difficulty with estimation of the coupling k value is related to its non-unique impact on the signal. Depending on its value and the term it scales coupling strength can affect peaks in Fourier spectrum by widening them, creating satellite peaks or shifting their positions. Although there are many documented attempts of estimating coupling factors, the majority of the research assumes that the input is presented as multivariate observations of coupled oscillators [97, 112, 149]. In the KurSL case, however, the observation is assumed to be a univariate time series composed of scaled observations and to the best of our knowledge, no method would work with such conditions. Due to limited understanding of the general impact of these factors their pdfs are defined as unity distribution with a relatively wide range of $k \in [-5, 5]\ \frac{\mathrm{rad}}{\mathrm{s}}$. Once the coupling values are sampled, they are additionally scaled such that the sum of individual oscillator does not exceed intrinsic frequency as equation (5.32) suggests (positive instantaneous frequency constraint).

The result of the algorithm is an initial estimate of the KurSL system which matches the provided signal the best. Since the number of oscillators depends on the energy ratio threshold E_ϵ and the type of the peak, one should choose them carefully. Although the system consisting of more oscillators, i.e. more degrees of freedom, would provide with better estimate this does not mean, that the result would be meaningful. This could lead to overfitting the system. Moreover, from the pragmatic point of view, the more oscillators are used, the longer it takes for the optimisation to finish.

6.2 The algorithm

The algorithm for the KurSL method consists of a few steps. Initially, parameter estimates are obtained via the algorithm provided in Section 6.1.2. Based on these results initial joint probability distribution is created so that the MCMC can fit the model. At each iteration and for each walker, the reconstruction is compared to the initial signal. If computed error is below the predefined threshold, the algorithm terminates returning a set of parameters which is associated with the smallest error. Otherwise, each walker's position is updated accordingly to the MCMC rules. The best set is assigned as a *maximum a priori* (MAP) vector for *a posteriori* joint pdf. The whole procedure is presented in the algorithm below (Algorithm 2).

6.3 Convergence

The KurSL method aims to decompose a signal into oscillatory components. Its convergence is defined by the process of obtaining a limit for a sequence made out of reconstruction components. Such problem can also be considered as minimising the difference between the input signal $\mathcal{S}$ and the reconstruction signal $\mathcal{K}$ obtained via KurSL. The sequence in question is enumerated by the number of oscillators used for reconstruction,

$$R_N = \mathcal{S} - \mathcal{K}_N = \mathcal{S} - \sum_{n=1}^{N} c_n, \tag{6.7}$$

```
Data: input signal
Result: KurSL oscillatory components
Determine initial parameters using peak fitting algorithm (Alg. 1);
Create a priori pdfs for each parameter;
foreach Iteration of MCMC do
    foreach Walker do
        Compute likelihood;
        Update location;
        if likelihood below threshold then
            Return set as the best set;
            Terminate for loops ;
        end
    end
end
if no likelihood below threshold then
    Assing global MAP as the best set;
end
Reconstruct and return oscillators generated for the best set;
```

Algorithm 2: The KurSL algorithm.

where $\mathcal{K}_N$ denotes the decomposition set of N components c. The difference is also related to the total error of reconstruction, which in case of the Hilbert space $\mathcal{L}^2$ is

$$E_N = \sqrt{\int_0^T \left| \mathcal{S}(t) - \sum_{n=1}^{N} c_n(t) \right|^2 dt}. \tag{6.8}$$

For the $\mathcal{K}_N$ sequence to be convergent, its error E_N has to decrease to 0 with the increase of N.

The convergence for the KurSL is guaranteed by the Fourier series, which are a specific reconstruction set given by the KurSL method. These series are obtained when there are no couplings between the oscillators, making spectra to be delta functions at particular frequencies. The proof of uniform convergence is based on the Weierstrass M-test (Definition 1) [150], which states that if a convergent sequence can be limited for every positive index by the sequence in question, then such sequence is also convergent.

As it has been proven, Fourier series $\mathcal{F}$ of a function will converge uniformly to a given limit [151], which in this case is the initial function $\mathcal{S}$. Let $\mathcal{M}_n$ be the error of

reconstruction at index N, i.e.

$$\mathcal{M}_N = \sqrt{\int_0^T |\mathcal{S}(t) - \mathcal{F}_N(t)|^2 dt}. \tag{6.9}$$

The KurSL method tries to optimise the metric (6.8) for each index N. Although it is not guaranteed to find the global minimum, the obtained optimum can always be compared with a case without coupling between oscillators, and thus it will always have an error at most equal to the Fourier series. This observation means that the Fourier series limits the KurSL sequence

$$\|\mathcal{S}(t) - \mathcal{K}_N(t)\| \leq \|\mathcal{S}(t) - \mathcal{F}_N(t)\| = \mathcal{M}_N, \tag{6.10}$$

for all lengths T, and since the former is convergent, the latter is also convergent.

Definition 1 (Weierstrass M-test)**.** *Suppose that* $\{u_n\}$ *is a sequence of real- or complex-valued functions defined on a set* D*, and that there is a sequence of positive numbers* $\{\mathcal{M}_n\}$ *satisfying*

$$\forall n \geq 1, \forall x \in D : \|u_n(x)\| \leq \mathcal{M}_n, \quad \sum_{n=1}^{\infty} \mathcal{M}_n < \infty. \tag{6.11}$$

then the series

$$\sum_{n=1}^{\infty} u_n(x) \tag{6.12}$$

converges uniformly on D*.*

6.4 Parametric stability

This section addresses the question of stability of the mapping function from the parametric space into the time series space. Although in general for the number of oscillators larger than two the problem is difficult to solve analytically, it can still be analysed through numerical simulations. Often knowing how traversing through one of the spaces is projected onto the other can provide some insights on how these spaces are shaped. Although the KurSL model is defined by both the number of oscillators,

$\mathcal{N}$, and the order, M, the following discussion focuses mainly on the above parameter. Such a choice is considered to have a much bigger qualitative impact on the general performance. For this reason, in this section spaces are annotated as $\mathcal{P}_\mathcal{N}$ and $\mathcal{S}_\mathcal{N}$ for parameter and time series spaces, respectively.

The parameter space $\mathcal{P}_\mathcal{N}$ refers to space with all possible input parameters for the KurSL method. The position $\mathbf{p}_\mathcal{N}$ in this space depends on the order M of the model and is defined by all the values of the set, i.e.

$$\mathbf{p} = [\omega_1, \ldots, \omega_\mathcal{N}, \rho_1, \ldots, \rho_\mathcal{N}, \theta_1, \ldots, \theta_\mathcal{N}, k^1_{1,1}, \ldots, k^1_{2,\mathcal{N}}, \ldots, k^{N_h}_{\mathcal{N},\mathcal{N}}], \tag{6.13}$$

where the total number of parameters is $N = \mathcal{N}(3 + N_h(\mathcal{N} - 1))$. Under the condition of having a discrete set of parameters, it is proposed to define distance with metric M_2, i.e. Euclidean distance. This means that for two vectors $\mathbf{p}_1$ and $\mathbf{p}_2$ the distance between them is given as

$$\mathcal{M}_\mathcal{P}(\mathbf{p}_x, \mathbf{p}_y) = \sqrt{\sum_{i=1}^{n} (p_{x_i} - p_{y_i})^2}, \tag{6.14}$$

with x and y being all corresponding parameters for vectors $\mathbf{p}_x$ and $\mathbf{p}_y$, respectively. For example, if two vectors differ only at θ value by $\Delta\theta$ the distance will be $\mathcal{M}_\mathcal{P}(\mathbf{p}_x, \mathbf{p}_y) = |\Delta\theta|$.

In case of time series space in general, one would require space for continuous functions. However, since all signals are recorded and discretised, this means one can utilise similar metric as in parameters space. Each oscillation can be stored in an array of a length T_N and since there are $\mathcal{N}$ oscillations in reconstruction, the whole time series vector $\mathcal{S}$ can be described as

$$\mathbf{s} = [s_1(t_0), \ldots, s_1(t_{T_N}), s_2(t_0), \ldots s_2(t_{T_N}), \ldots s_\mathcal{N}(t_{T_N})]. \tag{6.15}$$

This means that all oscillatory components were concatenated to create a single vector

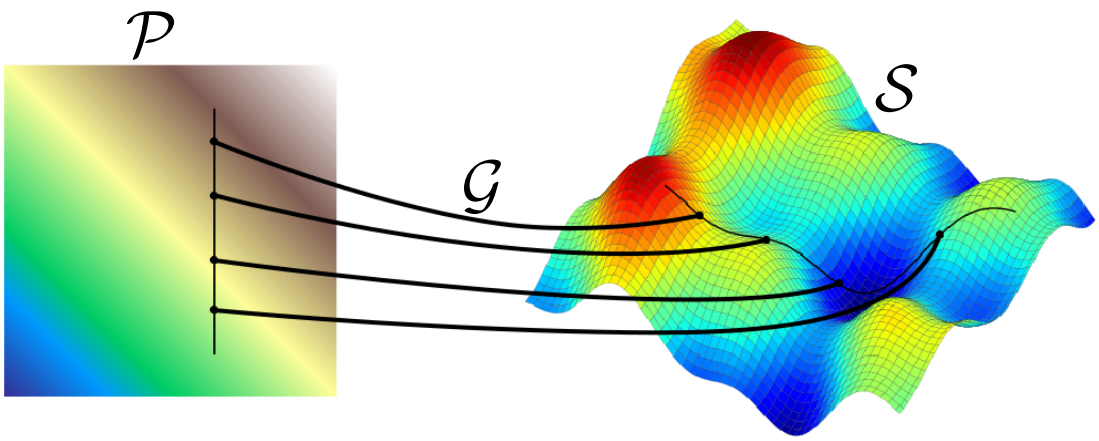


Figure 6.2: Graphical representation of a mapping $\mathcal{G}$ between the parameter space $\mathcal{P}$ and the time series space $\mathcal{S}$. The function $\mathcal{G}$ maps points (dots) in parameter space $\mathcal{P}$ onto respective positions in time series space $\mathcal{S}$.

of length $\mathcal{N} \times T_N$. It also means that the distance between two vectors $\mathbf{s}_1$ and $\mathbf{s}_2$ is

$$\mathcal{M}_{\mathcal{S}}(\mathbf{s}_1, \mathbf{s}_2) = \sqrt{\sum_{i=1}^{n} (s_{x_i} - s_{y_i})^2}, \tag{6.16}$$

where, again, x and y are values corresponding to vectors $\mathbf{s}_1$ and $\mathbf{s}_2$.

An illustration of a mapping function $\mathcal{G}$ from the parameter space $\mathcal{P}$ onto the time series space $\mathcal{S}$ is presented in Figure 6.2. It highlights that a straight line in one space may not necessarily map onto a straight line in the other. Not only the curvature can differ between spaces, but also the distance between consecutive points can vary. Such behaviour indicates nonlinearity of the mapping function. A series of experiments were performed to present how each parameter affects the model and whether the KurSL is a nonlinear model. In each experiment, all parameters except for one were kept constant and for each position of the free parameter, a time series were generated using the KurSL model. The core parameters used in all experiments are presented in Table 6.1. Although the number of oscillators was chosen to be $\mathcal{N} = 4$, obtained results can be generalised.

In the first experiment, the parameter of interest was amplitude. All parameters except for the amplitude of the 2^{nd} oscillator, i.e. ρ_2, where set according to the Table 6.1. The range of modification values is given $\rho_2 \in [1, 10]$ with step $\Delta\rho = 0.2$.

Table 6.1: Initial parameters used for all experiments in this section. In each experiment, a single parameter was chosen and modified accordingly.

n	$\omega/2\pi$ [Hz]	θ [rad]	ρ [arb. u.]	k_{ij} $[\frac{\text{rad}}{\text{s}}]$			
				1	2	3	4
1	30	π	2		0.1	2.2	4.2
2	25	0	3	1.1		2.0	1.1
3	17	0	5	0.2	2.2		-0.9
4	10	2	1	0.1	0.1	0	

Obtained results in the form of time series are cumulatively presented in Figure 6.3. These graphs display changes in time series depending on the value of the parameter. Each row represents a different oscillator with the horizontal and vertical axes corresponding to time and component's amplitude values, respectively. The whole figure is colour-coded with legends on the side showing their numerical values.

As it can be seen from these graphs, only time series of the second oscillator were affected. Such result is due to the amplitude ρ_i manifesting itself only as a simple scaling value in the KurSL model (Eq. 5.28). Such behaviour is additionally expected to produce monotonous mapping function from $\mathcal{P}$ into $\mathcal{S}$ space. Indeed, this can be observed in Figure 6.4, where the distances in time series space $\mathcal{S}$ are presented. The top graph shows the absolute distance, i.e. $\|\mathbf{s_p}\|$, whereas the bottom shows relative distance to the previous parameter $\mathbf{p}$, i.e. $\mathcal{M}_{\mathcal{S}}(\mathbf{s}_{\mathbf{p}_i}, \mathbf{s}_{\mathbf{p}_{i+1}})$. The behaviour of both functions can be explained by noticing that in this case, the metric function behaves like

$$f(\rho) = \sqrt{a\rho^2 + b}, \tag{6.17}$$

with respect to the initial amplitude ρ. Investigating its changes with a constant step $\Delta\rho$ one can see from Equation (6.16) that the result has the form of

$$\mathcal{M}_{\mathcal{S}}(\mathbf{s}_{\mathbf{p}_i}, \mathbf{s}_{\mathbf{p}_{i+1}}) = \sqrt{a(\Delta\rho)^2 + b}, \tag{6.18}$$

where all variables are constant and independent from ρ. This means that in this case expression $\mathcal{M}_{\mathcal{S}}(\mathbf{s}_{\mathbf{p}_i}, \mathbf{s}_{\mathbf{p}_{i+1}})$ is constant.

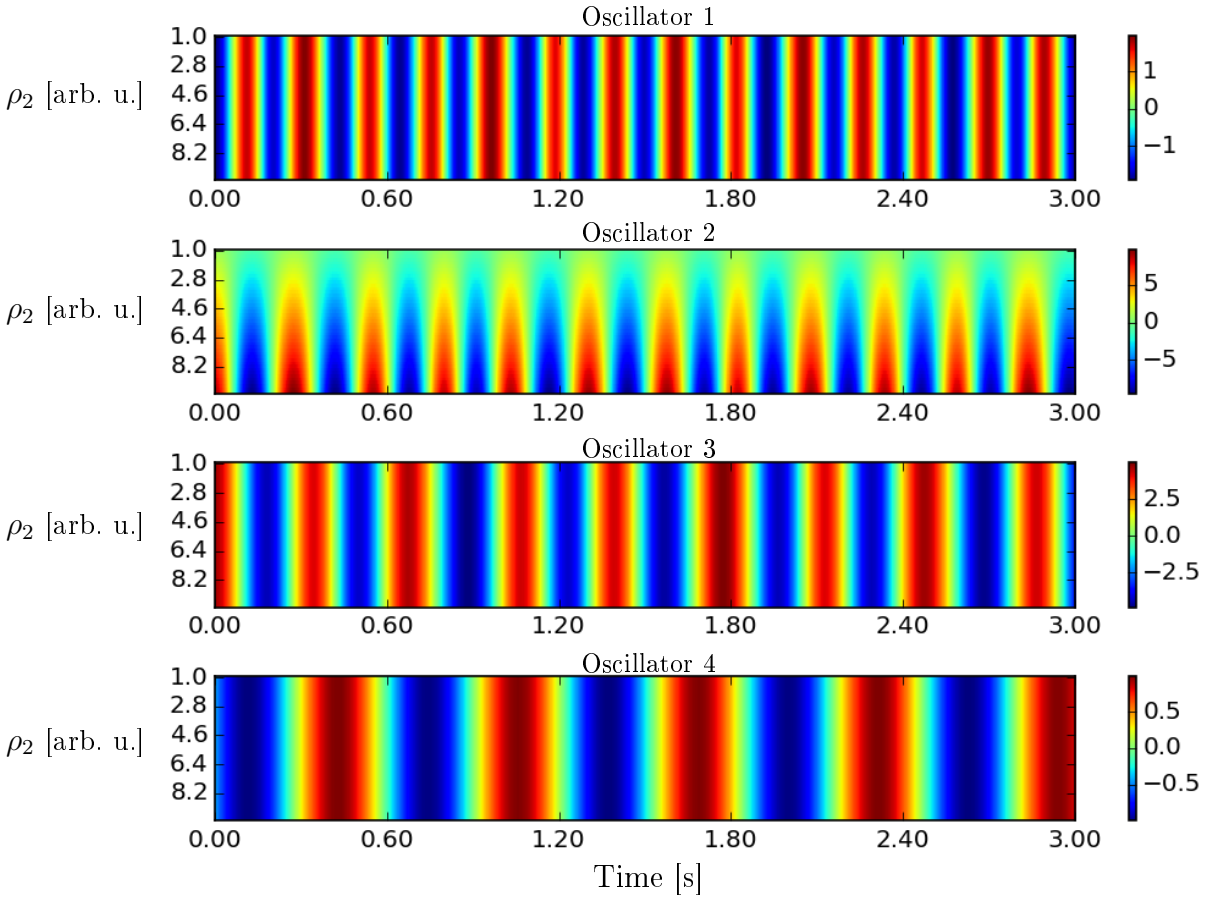


Figure 6.3: Colour-coded representation of time series obtained when modifying amplitude parameter, ρ. Each graph correspond to a different oscillator (initial values Tab. 6.1). The horizontal and vertical axes correspond to time and ρ values, respectively.

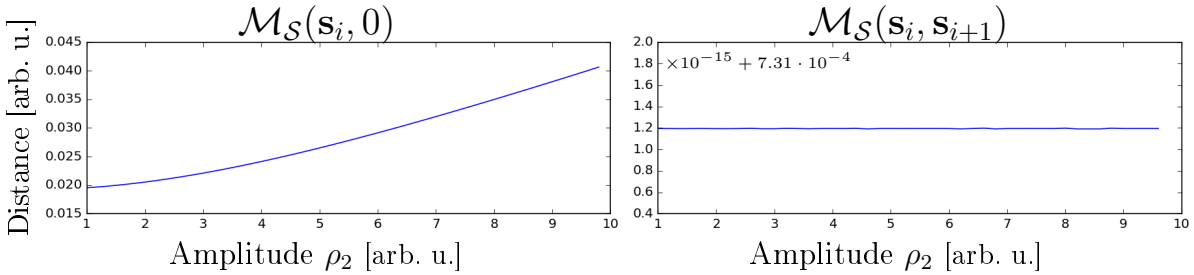


Figure 6.4: Distance values of time series as a function of amplitude parameter, r. The top plot displays the absolute distance of the vector $\|\mathbf{s}_\mathbf{p}\|$, whereas the bottom one is a distance to the previous vector $\mathbf{p}$ in parameter space, i.e. $\mathcal{M}_{\mathcal{S}}(\mathbf{s}_{\mathbf{p}_i}, \mathbf{s}_{\mathbf{p}_{i+1}})$.

The second experiment tested the behaviour when changing initial phase. The parameter has been modified in a range of $\theta_2 \in [0, 2\pi]$ rad with a step $\Delta\theta = 0.1$ rad. All obtained signals are presented in Figure 6.5, which displays the same information as Figure 6.3 with the difference that vertical axis now corresponds to the phase values. In this example, the effects of modulation are much more pronounced. The noticeable effect is the shift in phase of the whole second oscillator, which directly corresponds to the parameter. However, the shift is not monotonous; when phase $\theta_2 \approx 1.5$ rad there is a change in the progression. Due to the strong coupling between the second and third oscillators, this transition is also visible in the third oscillator and indirectly in the first. Effects in the fourth row are unnoticeable.

The aforementioned transition is more understandable when analysing distance graphs in Figure 6.6. On both plots, there is a peak near $\theta_2 \approx 1.5$ rad. This means that with a constant change of phase value in parameter space steps in time series space $\mathcal{S}$ suddenly increase and they differ more significantly. These modulations do not seem to affect the frequency, but they are visible as phase modulations, which also can be observed as a variation in the width of red and blue columns (Fig. 6.5).

In another experiment, the coupling strength between the first and the second oscillator has been modulated. Again, all values were kept constant (Tab. 6.1) except for k_{21}, which modulation range was set to $[-6, 6]$ $\frac{\text{rad}}{\text{s}}$, with step $\Delta k = 0.1$ $\frac{\text{rad}}{\text{s}}$. Collective results and distance values are presented in Figures 6.7 and 6.8. In this case, modulations in both frequency and amplitude are visible. On the first sight it seems that effects of k modulations are symmetrical — the further from $k = 0$ the more modulations on all components. This is especially pronounced in Figure 6.7 and the relative distance plot 6.8. It seems that there are two maxima for approximately $k = -5$ $\frac{\text{rad}}{\text{s}}$ and $k = 5$ $\frac{\text{rad}}{\text{s}}$. When coupling value reaches these points, the frequency of the second oscillator is closely matching the frequency of the first one. With the increase of coupling strength $|k| > 5$ $\frac{\text{rad}}{\text{s}}$ the distance is decaying and stabilising. Based on the shape of the peak in Fig. 6.8 one can see that there is a different behaviour depending on the sign of the coupling strength k; for large positive k, the signal converges faster to the final frequency.

It has to be pointed out, that although the coupling is between the first and

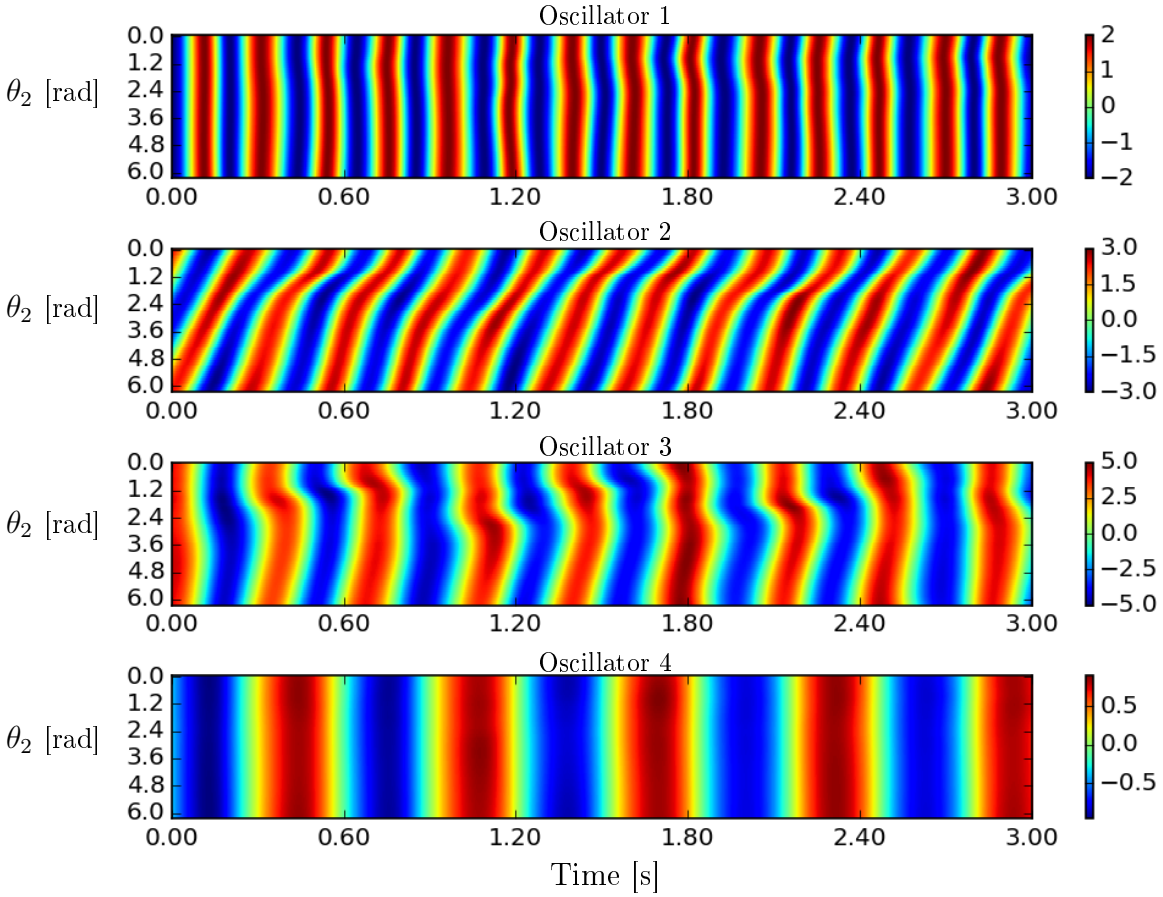


Figure 6.5: Colour-coded representation of time series obtained when modifying phase parameter, θ_2. Each graph corresponds to a different oscillator (initial values Tab. 6.1). The horizontal and vertical axes correspond to time, and θ_{02} values, respectively.

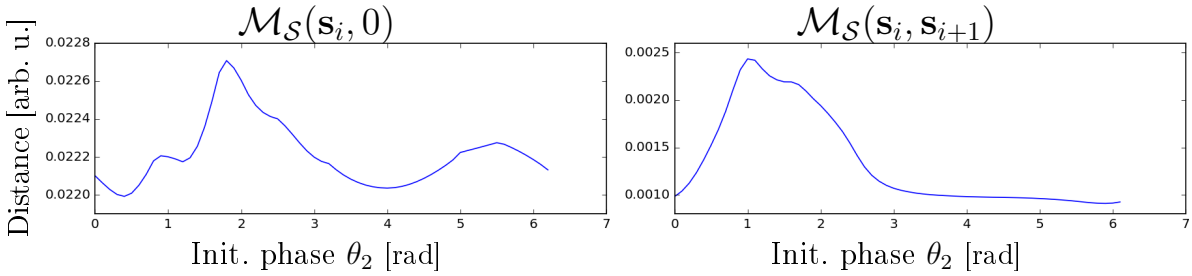


Figure 6.6: Distance values of time series as a function of the phase parameter, θ_2. Top plot displays the absolute distance of the vector $\|\mathbf{s_p}\|$, whereas the bottom one is the distance to the previous vector $\mathbf{p}$ in parameter space, i.e. $\mathcal{M}_{\mathcal{S}}(\mathbf{s}_{\mathbf{p}_i}, \mathbf{s}_{\mathbf{p}_{i+1}})$.

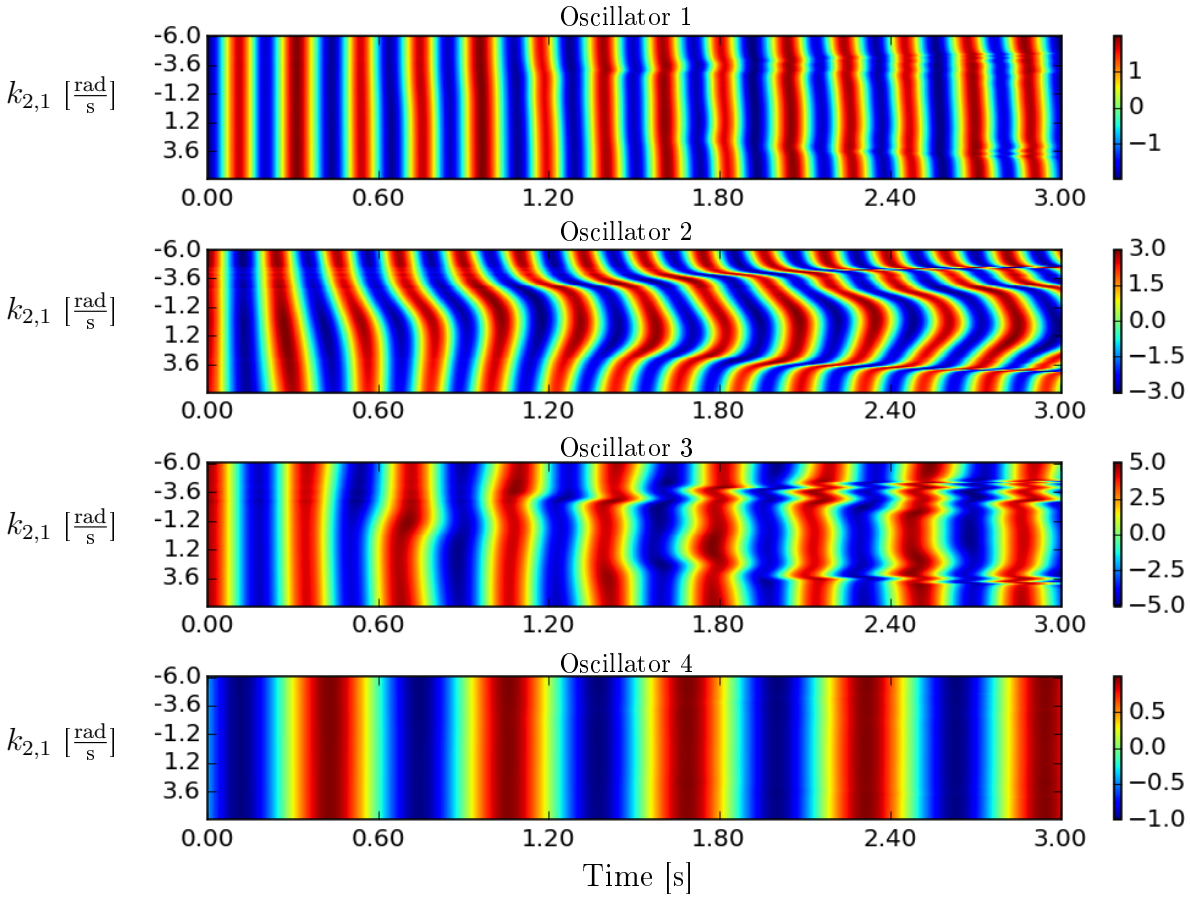


Figure 6.7: Set of components obtained by varying coupling strength between the first and second components, $k_{2,1}$. The range of changes is from -6 to 6 with step 0.1. Each graph corresponds to a different oscillator (initial values Tab. 6.1). The horizontal axis is the time, and the vertical axis is the value of the coupling, $k_{2,1}$.

second components, the first component has not been visibly affected. It is the third oscillator which seems to be affected the most. Such behaviour is explained by the effect of coupling not being normalised to the intrinsic frequency. From equation (5.5) it seems clear that the bigger intrinsic frequency, the more signifiant coupling strength has to be to affect phase modulations. Third's component ω is about two times smaller than the first ones leading to larger sensitivity in modulations, even if indirect.

The parameter that seems to have the most significant impact on the time series is the intrinsic frequency, ω. In the fourth experiment, the frequency of the second component ω_2 has been varied in the range $[5, 45]$ $\frac{\text{rad}}{\text{s}}$, with step $\Delta\omega = 0.1$ $\frac{\text{rad}}{\text{s}}$. As in previous examples, Figure 6.9 and 6.10 correspond to colour-coded graphs of time series and computed distance metrics, respectively. In this case, all components have

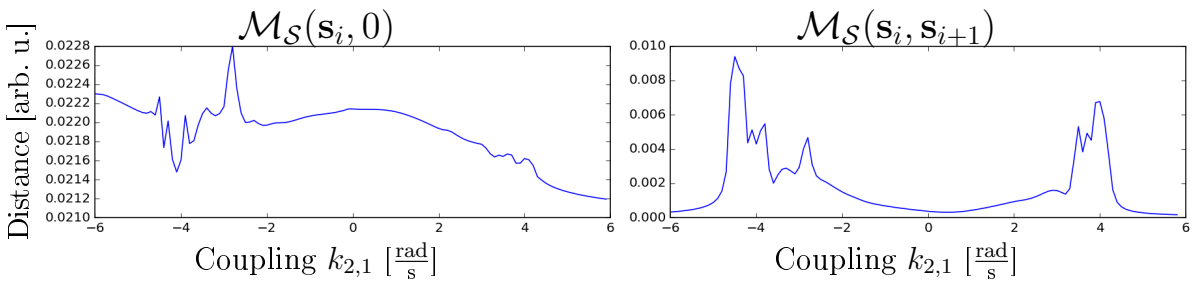


Figure 6.8: Distance values of time series as a function of coupling strength parameter, $k_{2,1}$. The top plot displays the absolute distance of the vector, whereas the bottom one is the distance to the previous vector $\mathbf{p}$ in parameter space.

been visibly affected. Major distortions are in areas of $\omega \approx 30$ $\frac{\text{rad}}{\text{s}}$ for the first, $\omega \approx 8$ for the fourth oscillator and in the range $[13, 22]$ $\frac{\text{rad}}{\text{s}}$ for the second oscillator. This seems to correspond directly to the frequency of other components, suggesting that there is resonance effect. These sudden changes can also be observed when analysing distance plots in Fig. 6.10. In regions close to the frequency of any other oscillator there are dynamical changes. However, in regions relatively far, i.e. when $\omega \approx 15$ $\frac{\text{rad}}{\text{s}}$ and $\omega > 33$ $\frac{\text{rad}}{\text{s}}$, changes in metric values are much more gradual.

These experiments and analyses provide with some insight on how the model behaves. They show that mapping from a parameter space $\mathcal{P}$ into time series space $\mathcal{S}$ is not a simple linear function, but rather that its shape depends on both absolute and relative values of all parameters. Results obtained when manipulating intrinsic frequency and coupling strength suggest that there can be regions with a sudden change of gradient. Based on the model's structure and regions in which these transitions occurred, it is expected that such sharp modulations are due to resonance effect between oscillators. Although the exact mechanism for this effect is uncertain, it highlights what behaviour should be considered with further studying of the method.

6.5 Conclusion

In this Chapter, it has been presented how to obtain a method for extracting oscillatory features based on the KurSL model. Such transition is done by incorporating a scheme for finding a parameter set by which the model describes the input data the best. The difficulty with analytically solving the KurSL model forces the usage of numerical

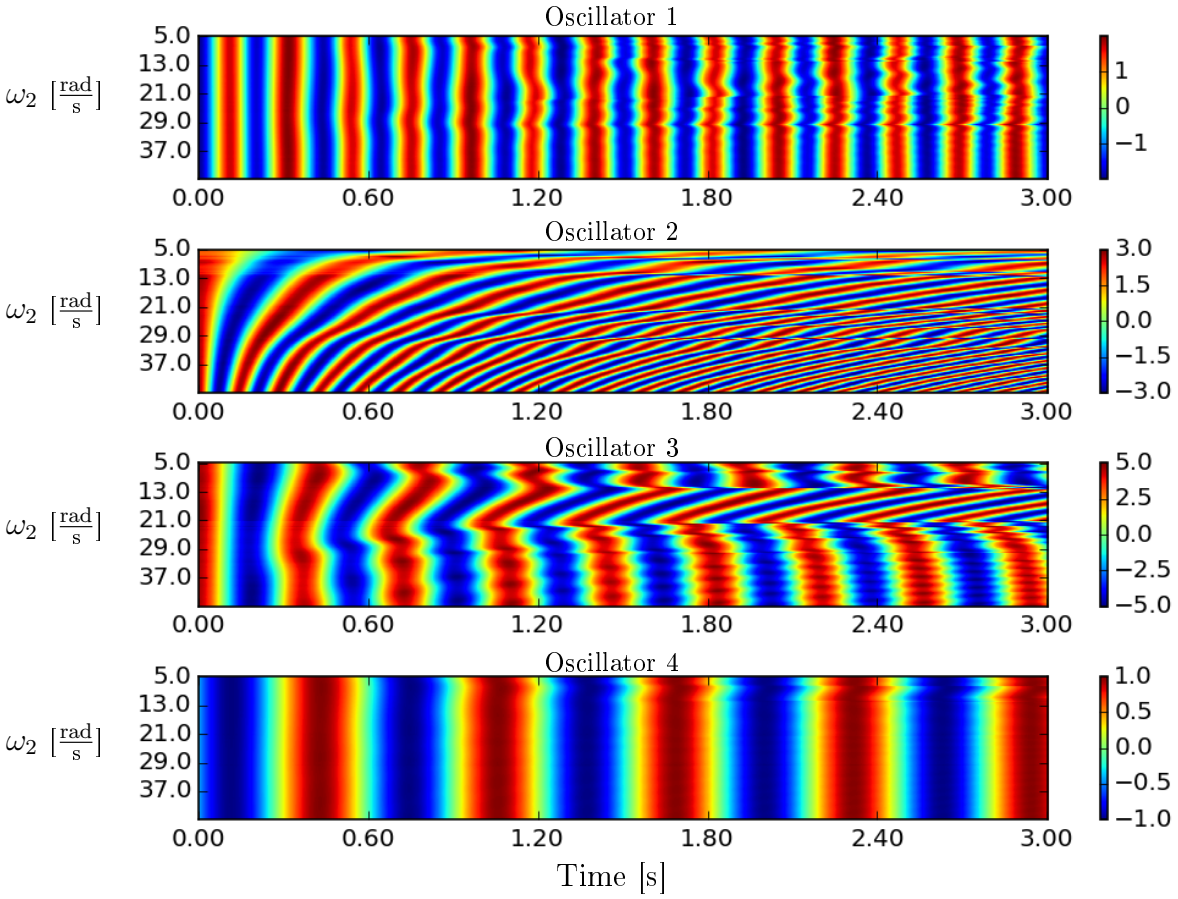


Figure 6.9: Colour-coded representation of time series obtained when modifying intrinsic frequency parameter, ω_2. Each graph corresponds to a different oscillator (initial values Tab. 6.1). Horizontal and vertical axes corresponds to time and ω_2 values, respectively.

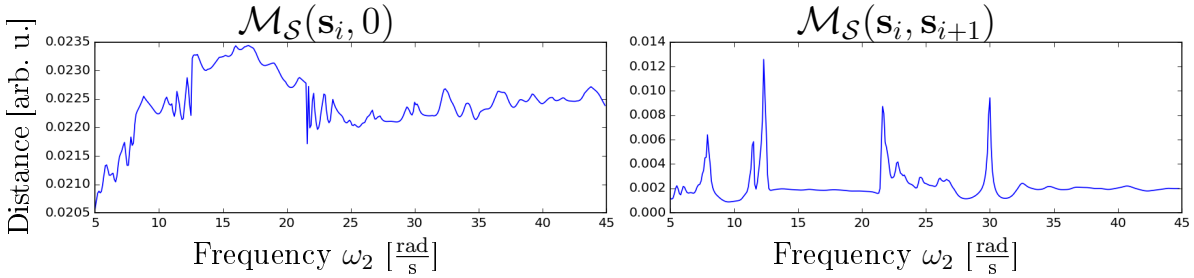


Figure 6.10: Distance values of time series as a function of intrinsic frequency parameter, ω_2. The top plot displays the absolute distance of the vector, whereas the bottom one is distance to the previous vector $\mathbf{p}$ in parameter space.

fitting methods. Choice of methods, however, are limited due to the non-convex behaviour of the cost function. The presented solution uses Monte Carlo Markov Chain (MCMC) to find the global optimum in a heuristic manner. This optimisation is done by updating the global probability density function (pdf) of parameters based on initial estimations. These estimates are determined based on the signal's shape and model's properties. Some of the initial values can be computed based on the spectrum of the input signal. For this reason, the emphasis is put on the peaks in the Fourier spectrum which correspond to decoupled oscillators. Through correspondence, these peaks, i.e. their location and amplitude, can be connected to values of intrinsic frequencies and component's amplitudes. Additionally, it has been observed that the width of the peak function could be linked to the coupling values. These observations have been combined into an algorithm for estimating the number of oscillators and their initial values. The algorithm greedily tries to remove the most energy from the spectrum by iteratively removing peak functions where each peak corresponds to a single oscillation. The algorithm does that until it meets a certain energy threshold. Once these parameters are obtained the MCMC method computes the best global parameters for the model.

The KurSL method in such form can be computationally demanding. The time of convergence and whether it obtains the best solution, depends greatly on the optimisation process. However, as it has been proven in Section 6.3, for a given method's order M and a set number of oscillators $\mathcal{N}$, the possible reconstruction of the method is bounded by the Fourier series reconstruction with the same number of components. This means that in the worst case the method converges to the input signal at least as quick as the Fourier series sequenced by the number of oscillators $\mathcal{N}$. In practice, however, one would expect the convergence to be quicker, due to a broader family of the possible components.

The richness of the solutions has also been demonstrated through studying the mapping between parameter space $\mathcal{P}$ and time series space $\mathcal{S}$. As it has been shown in Section 6.4, traversing through one of the spaces along a line does not necessarily create a linear path in the other. The analysis shows that the mapping of distance from parametric space $\mathcal{P}$ onto a time series space $\mathcal{S}$ has a non-linear and local behaviour;

it depends on the number of oscillators and their positions. Moreover, based on the results of the mapping on intrinsic frequency, there seems to be an effect which can be described as *resonance*. This means that the mapping function behaves differently when any pair of intrinsic frequencies have similar values.

It has to be emphasised that the KurSL is not a general purpose signal processing technique that extracts any type of oscillations from any given dataset. The method focuses is on systems composed of mutually interactive oscillators which would otherwise have harmonic characteristics. Through such definition, the method tries to include only physically meaningful systems. Moreover, by definition, one is not expected to obtain a single (isolated) oscillation other than harmonic. Without any external forces acting upon the object its dynamic ought not to be changed; likewise, the frequency of the oscillation should also be preserved. Nevertheless, this does not mean that only harmonic oscillators can be obtained. The KurSL treats the system as a whole. Any changes in one oscillator are induced by the other oscillators, regardless of their frequency range and directional influence. As an example, an oscillator spinning with friction can be modelled as two coupled oscillators where one heavy object has a strong asymmetric influence on the other. As another example, the KurSL is not expected to state that a chirp signal is a single oscillator. Instead, it would be described as a system with strong interaction between components. The KurSL method is assumed to be applied to data generated through communication processes and may not be suitable for an abstract signal.

Chapter 7

The KurSL application examples

As discussed in the previous chapter, the KurSL model can be used as a signal decomposition method. This chapter aims to demonstrate such usage by decomposing and analysing both synthetic and empirical (EEG) data. The first section (7.1) describes the decomposition of stationary signals using the KurSL method. In the second section (7.2) dynamical analysis of extracted parameters is presented. Finally, in Section 7.3 two empirical EEG signals are analysed and decomposed with the KurSL method both in stationary and dynamical manners.

All computations in this chapter were performed using the Python programming language. The implementation of the KurSL algorithm 2 (sec. 6.2) with which all simulations were performed was written using the *NumPy* [152] and *the MCMC Hammer* [147] libraries. Computation of spectrograms and scaleograms were performed using the SciPy [126] and PyWavelets [153] packages, respectively. Although each experiment required parameter tuning due to having different focal time series, all scaleograms were computed using the Morlet function as a mother wavelet. Such a choice is popular for processing brain signals due to its optimal concentration simultaneously in both time and frequency [35, 36]. The source code for the KurSL and the Huang-Hilbert transformation is available from the author's webpage [125].

7.1 Analysis of a simulated signal

This section presents and discusses an application of the KurSL decomposition method on synthetic data. The first demonstration (Sec. 7.1.1) introduces decomposition based on a signal generated using the 1^{st} order KurSL method with four oscillators. Its results are presented in a form of probability distributions, which indicate the most common and most likely parameter values. The following Section 7.1.2 compares the performance of the KurSL method with two different orders. For this experiment, a signal was generated using the 3^{rd} order KurSL and then it was decomposed with the 1^{st} and 3^{rd} order KurSL methods.

7.1.1 Simple example

This case study is based on a synthetic signal generated with the KurSL model. Parameters for the model were sampled from uniform distributions with ranges $f \in [5, 15]$ Hz, $\theta \in [0, 2\pi)$ rad, $\rho \in [1, 3]$ and $k \in [-5, 5]$ $\frac{\text{rad}}{\text{s}}$. Frequencies had an additional constraint such that they had to be at least 0.5 $Hz \approx 3.14$ $\frac{\text{rad}}{\text{s}}$ apart. All selected values are presented in Table 7.1. The time series are generated for $t \in [0, 5]$ s with the sampling frequency $f_s = 1000$ Hz which allows for the spectrum frequency resolution $df = 0.2$ Hz. Their visual representation and computed Fourier spectrum are presented in Figures 7.1a and 7.1b, respectively. Despite having a visibly simple structure in the time domain, the spectrum appears to have many peaks smeared in the frequency domain. If all coupling factors $k_{i,j}$ were set to 0, the signal would be composed only of four harmonic oscillators for which the Fourier spectrum would contain only four peaks. Contrasting results imply that the coupling between oscillators has to strengthen the smearing effect and to introduce satellite peaks. Details about individual components are presented in Figure 7.2 which shows their time series, Fourier spectrum and spectrogram from the left, respectively.

As discussed in the previous chapter, the coupling between oscillators introduces modulations in both time and frequency domains. These effects make it difficult to determine the optimal window's properties for the spectrogram. In the case of the KurSL system, modulations are proportional to pairwise differences between the frequency val-

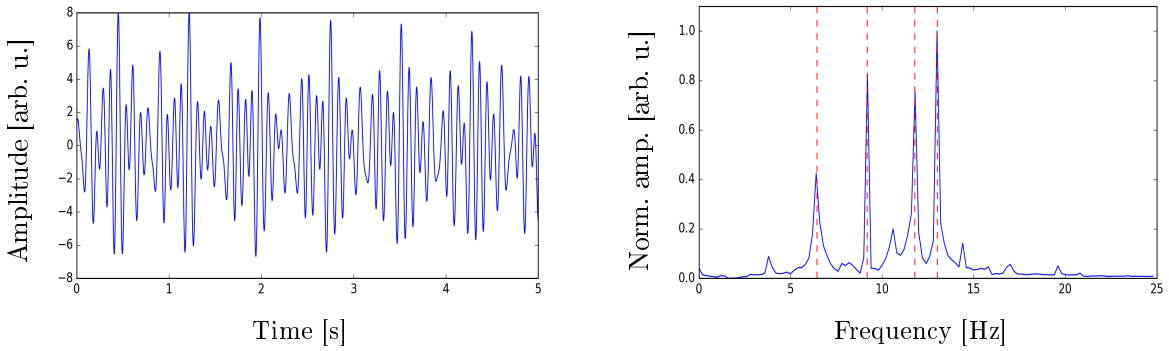
Table 7.1: Model's parameters used to generate simple KurSL example (Sec. 7.1.1).

n	$\omega/2\pi$ [Hz]	θ [rad]	ρ [arb. u.]	$k_{i,j}$ $[\frac{\text{rad}}{\text{s}}]$			
				1	2	3	4
1	13.03	0.23	2.67		2.18	1.97	-4.03
2	11.85	3.32	2.47	-4.8		-4.65	3.82
3	8.98	4.64	2.21	1.93	-4.76		4.09
4	6.46	0.52	1.58	2.95	-0.37	-4.85	

ues; thus, the expected resolution should be at least equal to the pairwise distances. The smallest frequency differences are between pairs of oscillators (1,2) and (3,4) which are $\Delta f_1 = |13.03 - 11.85| = 1.18$ Hz and $\Delta f_2 = |8.98 - 6.46| = 2.52$ Hz, respectively. To capture changes in the frequency domain at a given resolution the time window length needs to be reciprocal which gives the lengths equal to $l_1 = 1/\Delta f_1 = 0.85$ s and $l_2 = 1/\Delta f_2 = 0.35$ s. However, the window length of 0.85 s is too large to detect any modulations in the time domain. For this reason, the window was halved and the frequency resolution was artificially increased by zero-padding segments with an array equal to the length of the window. Since the padding only adds intermediated points, interpolated by a sinc function, an increased overlap would provide a smoother display of the translation. A result of these constraints is a Tukey window with the tapering parameter of $\alpha = 0.25$, an overlap of $p = 80\%$ and window lengths of $l_1 = 0.425$ s for the two oscillators with the highest frequencies and $l_2 = 0.35$ s for two with the lowest frequencies.

From these graphs, one can see that all components exhibit modulations in both time and frequency domains. Both Fourier spectra and spectra indicate that all components have a primary mode which is modulated through interactions with other components. In all spectrograms, these interactions produce an additional transient oscillation close to the main mode, but with a significantly smaller amplitude. The Fourier spectrum, however, shows these as satellite peaks around the position of the intrinsic frequency.

The KurSL analysis of the signal was performed in accordance with the method described in Chapter 6. The algorithm for determining initial parameters was set



(a) Time series

(b) Fourier spectrum

Figure 7.1: The signal generated using the KurSL model with parameters from Table 7.1. Figure (a) displays the sum of time series from all oscillators, whereas (b) is its Fourier spectrum. Vertical dashed lines on the spectrum denote peak positions determined by the KurSL method.

to extract a maximum of 4 oscillators, which is equivalent to using the energy ratio $E_\epsilon = 0.32$. Such configuration allows to demonstrate the results of the KurSL and compare these with the parameters that generated the signal. The method has detected oscillators which frequency and amplitude pairs (f_i, ρ_i) equal to {(13.01 Hz, 1.37), (11.78 Hz, 3.19), (9.18 Hz, 4.66), (6.43 Hz, 1.71)}. For visual comparison, these positions have been imposed as vertical dashed lines on the input's Fourier spectrum (Fig. 7.1b). The KurSL fitting mechanism was performed with MCMC which was deployed using 200 walkers each of which explored the 24-dimensional parameter space. The number of iterations was set to 300 out of which the initial 60 samples (20% of total) were discarded to mitigate the influence of incorrect initial conditions. This burn-in process made the overall sample size for all parameters' distributions to be 480000 points. Based on these points, a probability distribution function (pdf) was estimated using the Kernel Density Estimator (KDE) with a Gaussian kernel of width computed using Silverman's method. These KDEs are presented using grey colour in Figures 7.3, 7.5, 7.4 and 7.6 which pertain to intrinsic frequencies, initial phases, amplitudes and couplings strengths, respectively. In all figures, vertical lines indicate a representative value for a given distribution. Yellow and magenta colours represent median and mean values of all samples, respectively. Blue colour refers to the *maximum a posteriori* (MAP) value for a given marginal distribution which does not

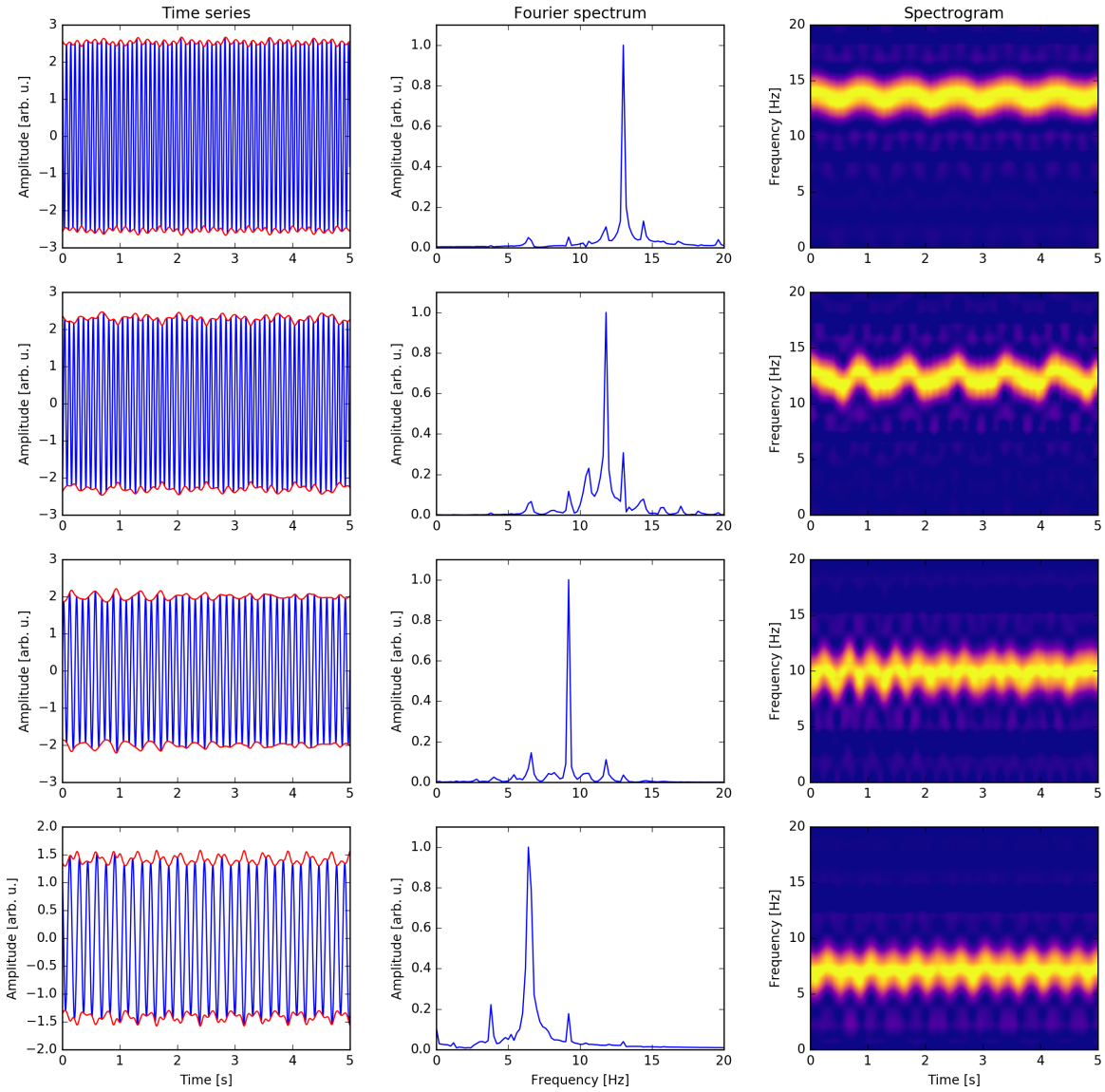


Figure 7.2: All components generated using the KurSL model with parameters taken from the Table 7.1. The left column contains time series with their instantaneous amplitude highlighted using red colour. Fourier spectra normalised such that the largest value is one, are presented in the central column. The right column contains spectra which were computed using the Tukey window with tempering parameter $\alpha = 0.25$, overlap $p = 80\%$ and lengths of 0.425 s, 0.425 s, 0.35 s and 0.35 s from the top, respectively. All spectra are computed with zero-padding equal to the length of window.

necessarily optimise the joint distribution; the best result is indicated by the global MAP, which is denoted with a vertical red line. All these lines can be compared to the parameters (black) with which the signal was generated (Table 7.1). The closeness of the black and red lines for all parameters suggests that the search has been performed successfully. Such result might be surprising since according to obtained distributions certain parameters are unlikely, e.g. initial phase θ_2 lays on the 91^{st} percentile. The multi-modality and skewing in many distributions towards the red line indicate that these values could have been obtained late in the walker's search. These could suggest that the MCMC search has not converged fully to the actual distributions and further computation could improve the results.

Let $\mathbf{p}_{median}$, $\mathbf{p}_{mean}$, $\mathbf{p}_{margin}$ and $\mathbf{p}_{map}$ refer to median, mean, marginal MAP and global MAP parameter sets, respectively, and their time series be denoted by $s(t)$ with the same index notation. As can be seen from these graphs, on each distribution these parameters $\mathbf{p}$ are relatively close to each other. Nevertheless, as discussed in Section 6.4, a small distance in the parameter space does not necessarily imply closeness in the time series space. For each aforementioned parameter set $\mathbf{p}$ a reconstructed signal was created and compared with the generated input. These comparisons are presented in Figure 7.7, where rows correspond to a different set of values; from the top they are: median $\mathbf{p}_{median}$, mean $\mathbf{p}_{median}$, marginal MAP $\mathbf{p}_{margin}$ and global MAP $\mathbf{p}_{map}$ reconstructions. For each row, the left column displays an overlaid reconstruction for a given parameter set $\mathbf{p}$ with an input signal, whereas their pointwise differences shown in the right column. Additionally, each reconstruction was labelled with values of pointwise mean square error (MSE) and residual energy (RE) defined as

$$\text{RE} = \frac{E(y_{in} - y_{rec})}{E(y_{in})}, \tag{7.1}$$

where y_{in} and y_{out} refer to input and reconstructed signals, respectively, and $E(\cdot)$ indicates energy of its argument. These comparisons show that the best parameter set provides a reconstruction which closely matches the input signal (RE=0.035, MSE=0.286 [arb. u.]). Nevertheless, as indicated previously, these results can be still improved as the significant difference for the other parameters suggests.

Further time-frequency analysis and the comparison with other methods is pre-

sented in Figure 7.8 which contains four time-frequency representations. The top row of the Figure contains two instantaneous frequency representations; the KurSL representation is on the left and the Huang-Hilbert transformation (HHT), i.e. using EMD decomposition, is on the right. Two heatmaps are presented in the bottom row where the left and right are short-time Fourier transformation (STFT) and the wavelet transformation (WT), respectively. These time-frequency representations were obtained using the following settings.

The EMD algorithm in the HHT has been configured to use the cubic spline interpolation and to accept an IMF only after five consecutive sifting when the number of zeros and extrema differed at most by one. Parameters used to generate presented STFT and WT were selected based on visual comparison. It was expected that the best representation for both methods would provide a similar result and that they would be able to highlight at least three regions of increased activity since the two components with the highest frequencies have overlapping frequency bands. In the case of spectrograms, a set of candidates was obtained by generating all possible combinations of a variable w_γ which consists of the length l, the type τ and the overlap percentage p, i.e. $\gamma = \{l, \tau, p\}$. These parameters are defined as $l \in [0.3, 2.0]$ (seconds) with 0.1 s step, $\tau \in$ {Hann, Tukey ($\alpha = 0.25$), Tukey ($\alpha = 0.5$)} and $p \in \{50\%, 75\%, 80\%\}$. Similarly, for the WT, a set of scaleograms was computed where the central frequency ω_0 was selected from a range [5, 15] $\frac{\text{rad}}{\text{s}}$ with a step 0.25 $\frac{\text{rad}}{\text{s}}$. Such range was chosen as it covers the majority of frequency bands shown in the individual components' spectrogram (Fig. 7.2). It was concluded that for the STFT the best set uses 0.9 s long Tukey type window with the taper fraction $\alpha = 0.25$ and overlap step $p = 75\%$. In the case of the Wavelet transform the central frequency was determined to be $\omega_0 = 13$ $\frac{\text{rad}}{\text{s}}$ as it describes a periodic increase in high frequencies activity with simultaneous, although subtle, indication that this is due mixing of two components. Worth highlighting is also representation obtained when using central frequency $\omega_0 = 6.5$ $\frac{\text{rad}}{\text{s}}$ presented with other computed spectrograms and scaleograms in Appendix B.1. Representation with $\omega_0 = 6.5$ $\frac{\text{rad}}{\text{s}}$, even more, emphasises the modulations in frequency, indicating the interaction between all components, and yet preserving bands with close to constant activity. In all scaleograms, the grey region indicates an area outside of the cone of

influence. For comparison purpose, the STFT and WT results were normalised such that the maximum value in a region of interest is one, as indicated by the attached scale bar.

When comparing time-frequency representations for all methods, the one obtained with HHT stands out as it is the only one which indicated more than four components in the signal. Although the EMD decomposed the signal into 8 IMFs, only two are used to highlight events where the other representations indicate four components, i.e. in the frequency range of $f \in [5, 15]$ Hz. An interesting behaviour displayed in the HHT is the number of deep peaks observed for the first component. Excluding the one that starts the component, all these dips correspond to positions where, in the KurSL representations, the two fastest components crossover, e.g. at time $t \approx 0.9$ s. Similarly, the STFT and WT have shown a merge and an expanse of two components in the frequency content, respectively. These methods were also able to identify the other two components, i.e. at frequencies $f \approx 9$ Hz and $f \approx 6$ Hz. Both representations highlight modulations in these components; in case of STFT variations are in the frequency domain, whereas the WT show changes in component's amplitude. All these phenomena agree with the KurSL instantaneous frequency representation which visibly displays variations in all components.

Concluding this experiment, for a simple nonlinear signal the KurSL method resulted in a parameter set, which is close to the one used to generate the signal. Since the signal was generated using the KurSL model, one would expect that the fitting mechanism would find the optimal values. Although provided results are suboptimal, the positioning of the best-detected parameters onto the tails of obtained distributions suggests that the optimisation method has not yet converged and further iterations would improve the results. Overall, all compared time-frequency representations display comparable results and highlighting similar behaviour.

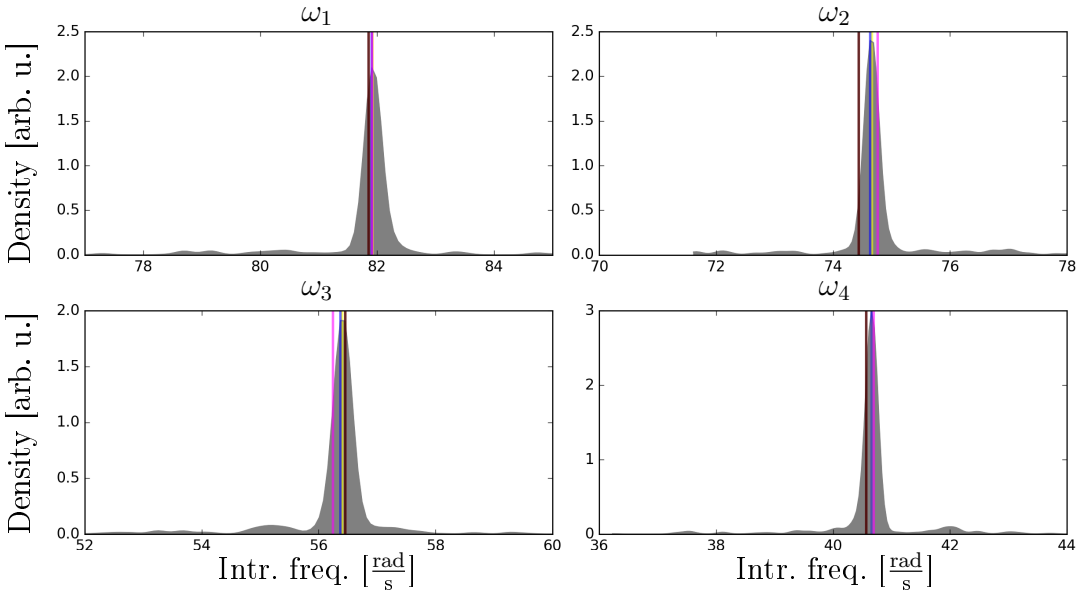


Figure 7.3: The probability distribution of intrinsic frequencies obtained using Gaussian KDE. Vertical lines indicate representative values of distributions. Blue, yellow and magenta lines code the maximum value of KDE, median and mean values, respectively. Red vertical line denotes the maximum a posteriori value of a joint distribution.

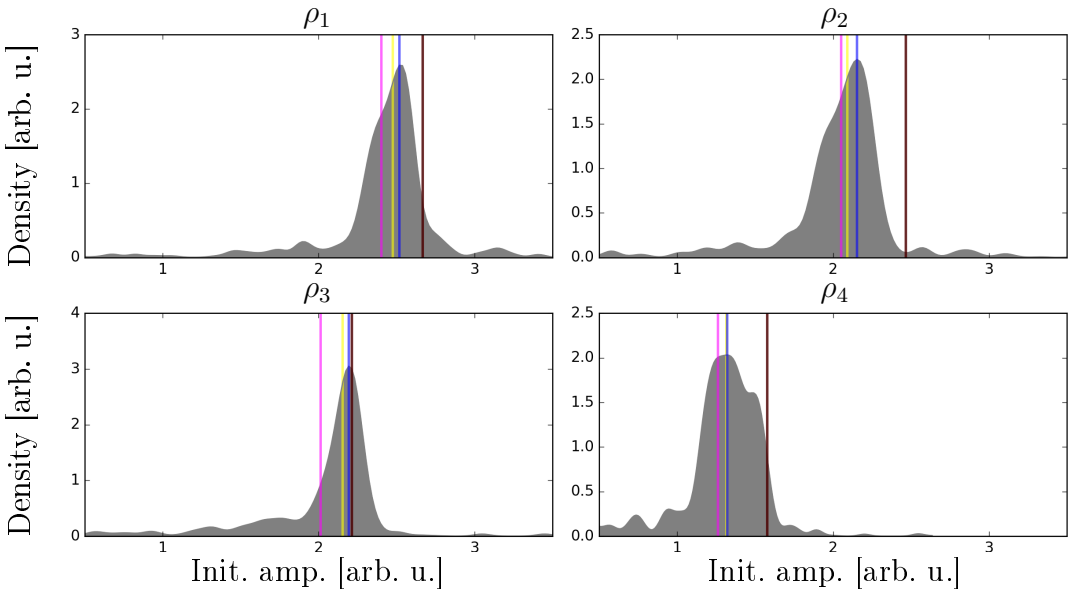


Figure 7.4: The probability distribution of amplitudes obtained using Gaussian KDE. Vertical lines indicate representative values of distributions. Blue, yellow and magenta lines code the maximum value of KDE, median and mean values, respectively. Red vertical line denotes the maximum a posteriori value of a joint distribution.

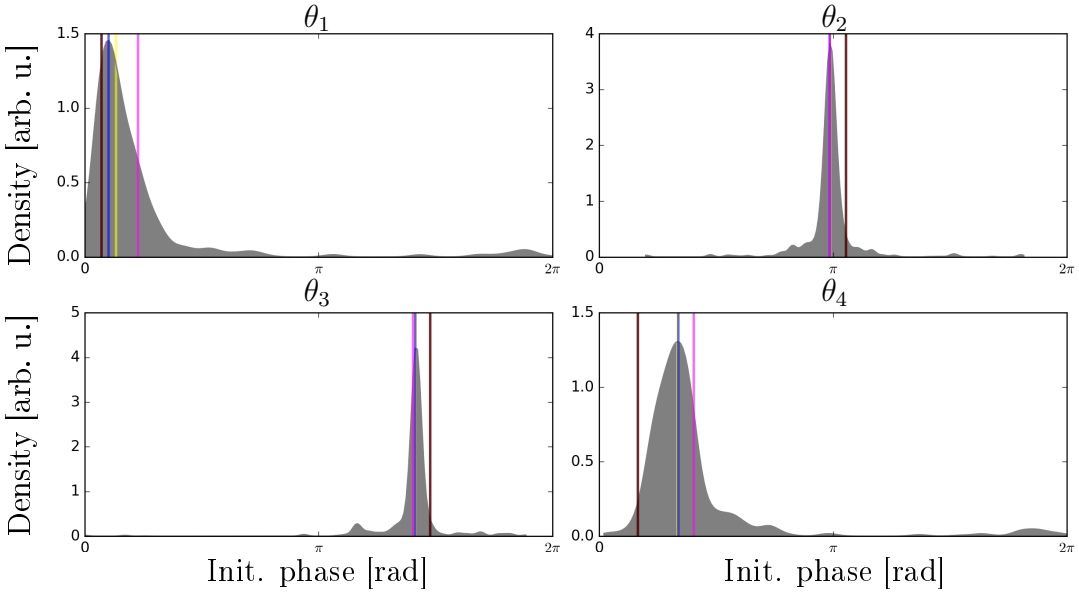


Figure 7.5: The probability distribution of phases obtained using Gaussian KDE. Vertical lines indicate representative values of distributions. Blue, yellow and magenta lines code the maximum value of KDE, median and mean values, respectively. Red vertical line denotes the maximum a posteriori value of a joint distribution.

7.1.2 Order comparison

The purpose of the second experiment is to demonstrate how the order of the KurSL method affects its performance and representation. For this experiment, an exemplary synthetic signal was generated. It consists of five oscillators that are coupled together with the 3^{rd} order KurSL process. Values used to create the input signal are presented in Table 7.3 and were sampled from uniform distributions with ranges $\omega \in [1, 150]$ $\frac{\text{rad}}{\text{s}}$, $\theta \in [0, 2\pi)$ rad, $\rho \in [1, 2.5]$ and $k \in [-15, 15]$ $\frac{\text{rad}}{\text{s}}$. The signal was generated for the time t range $t \in [0, 3]$ s with the sampling rate of 200 Hz which relates to resolution in time $dt = 0.005$ s and frequency $df = 0.33$ Hz. Visual representation of the time series and its Fourier Spectrum are presented in Figures 7.9a and 7.9b, respectively.

The choice of the model and its parameters means that the best reconstruction should be obtained using the KurSL of at least 3^{rd} order. For the purpose of studying effects of the KurSL order, the experiment has been conducted twice: initially using the first order KurSL and then the third. Both times the algorithm for extracting initial parameters was executed with the default configuration, i.e. $E_\epsilon = 0.1$ and no limit on the number of oscillators. For such configuration, the method was initiated with five

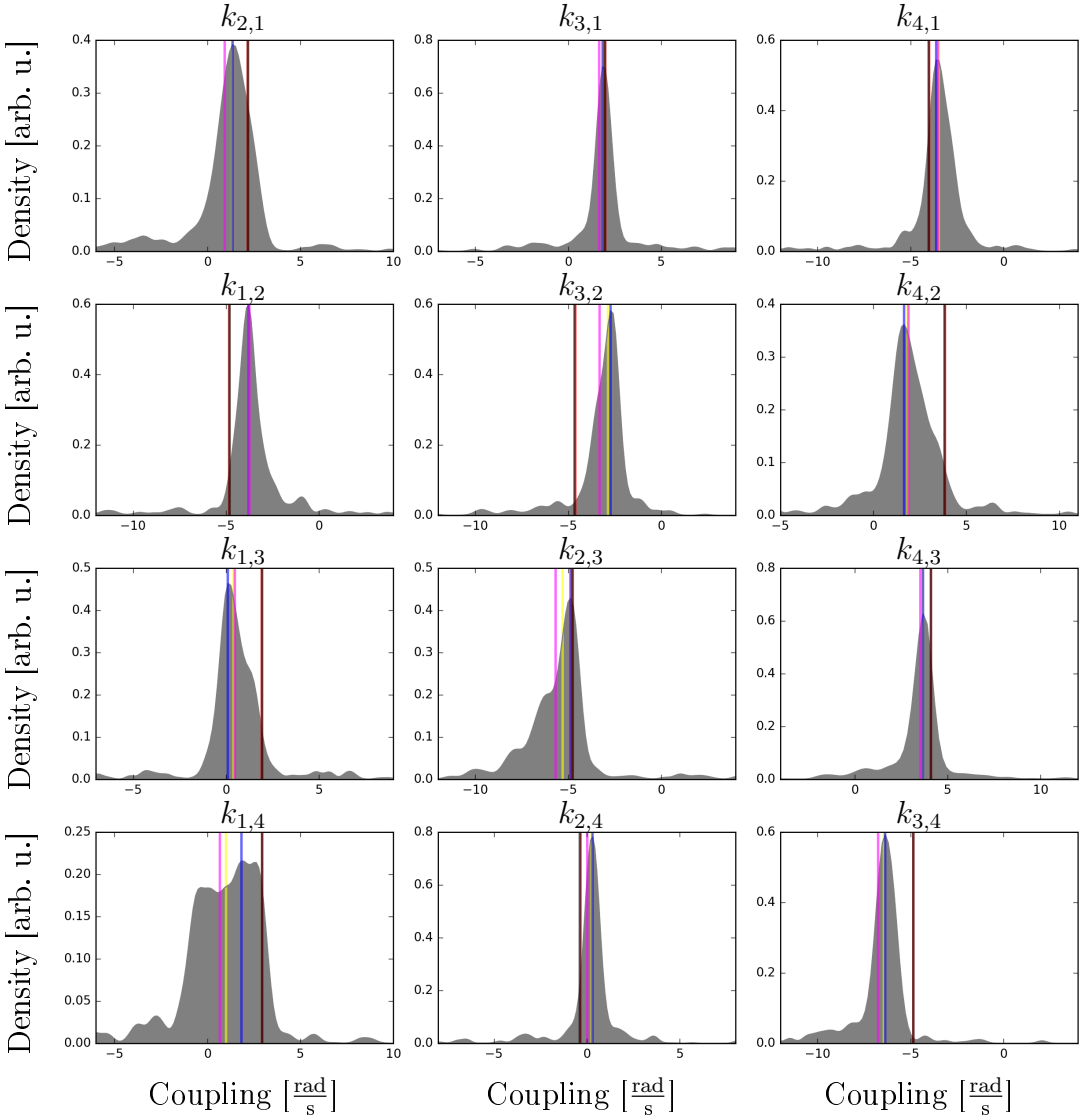


Figure 7.6: The probability distribution of coupling strengths obtained using Gaussian KDE. Vertical lines indicate representative values of distributions. Blue, yellow and magenta lines code the maximum value of KDE, median and mean values, respectively. Red vertical line denotes the maximum a posteriori value of a joint distribution.

n	1	2	3	4	5
ω [$\frac{\text{rad}}{\text{s}}$]	12.32	31.42	39.77	60.63	77.66
f [Hz]	1.96	5.00	6.33	9.65	12.36
ρ [arb. u.]	2.00	1.72	2.59	1.64	2.36

Table 7.2: Initial frequencies f and amplitudes r obtained for the simulated signal with five oscillators.

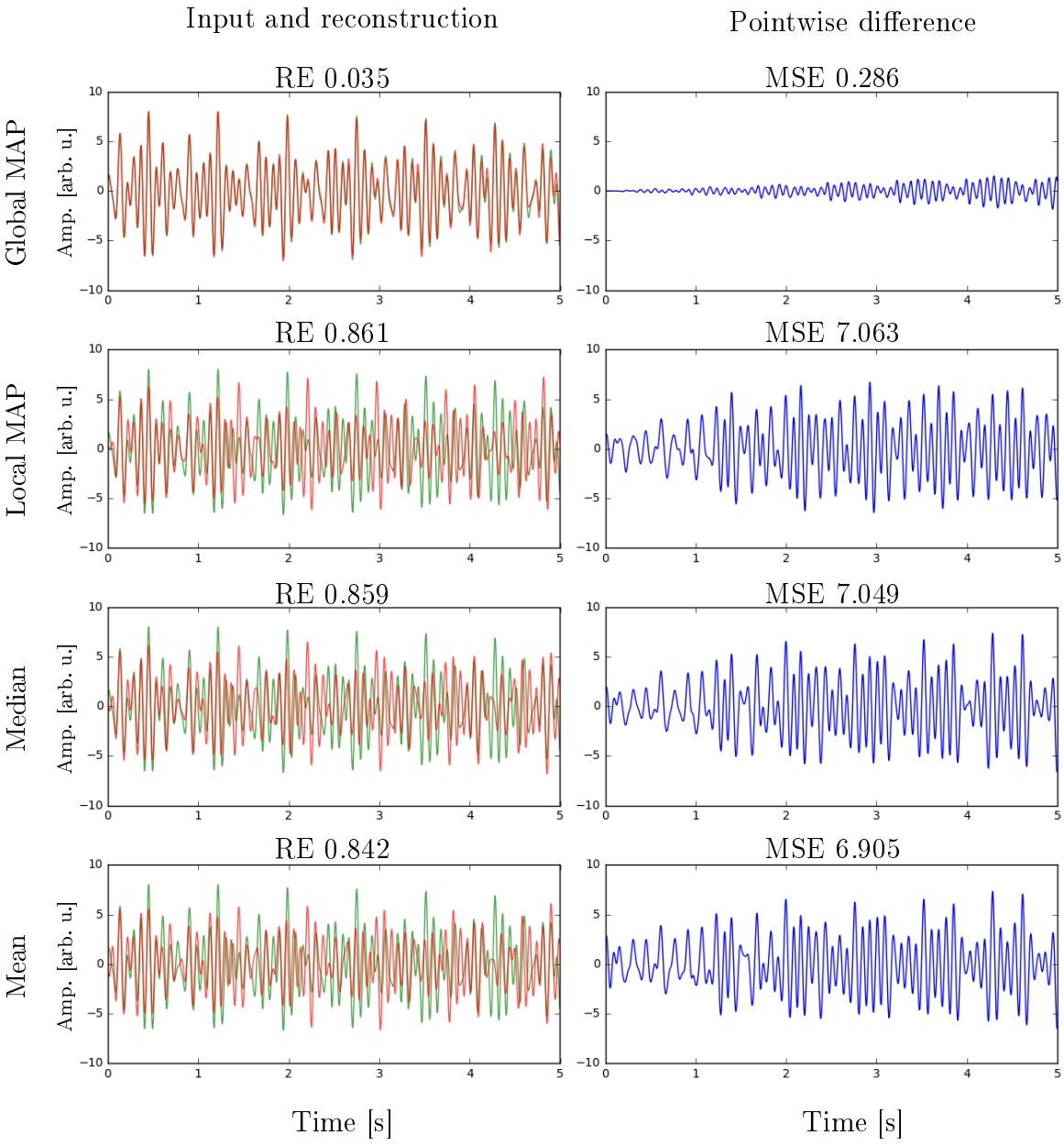


Figure 7.7: Comparisons of reconstructions obtained for sets corresponding to all obtained parameter estimates. Each row represents reconstruction for a set, which from the top are global MAP, marginal MAP, median and mean. Left column displays overlaid reconstructions (red) with an input signal (green). The right column shows a pointwise difference between those signals. Titles denote reconstruction measures, where residual energy (RE) is defined as a mean square error (MSE) divided by the sum of squares of the input signal.

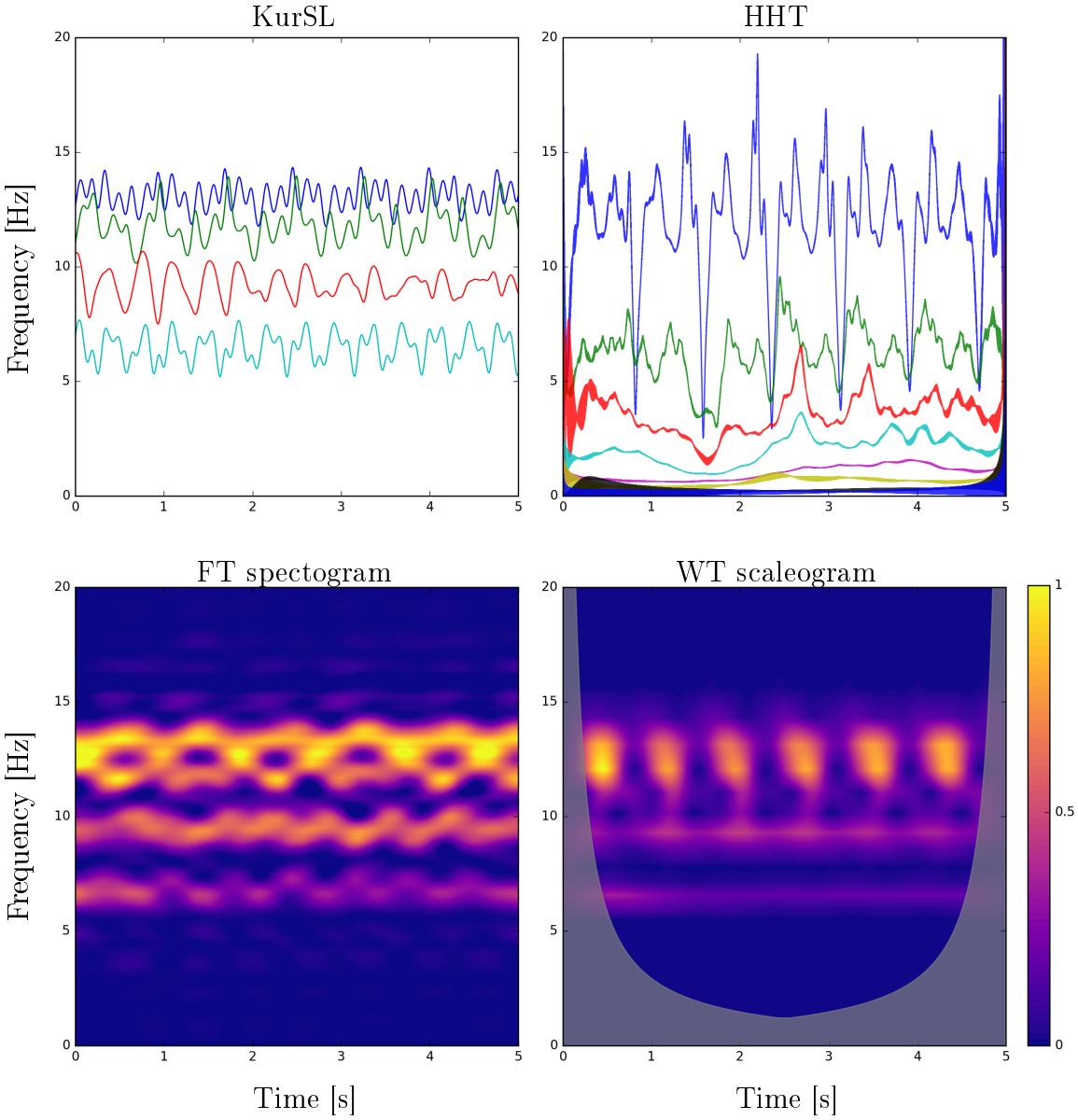
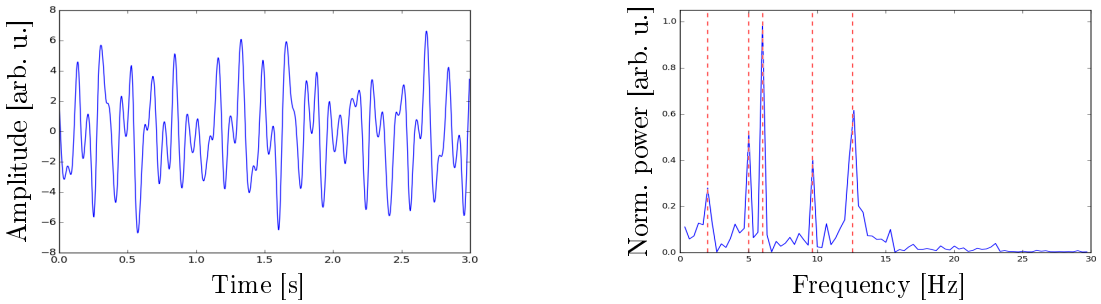


Figure 7.8: A comparison of different time-frequency representations computed on the simple KurSL signal (sec. 7.1.1). The top row contains instantaneous frequency dynamics obtained via the KurSL and the Huang-Hilbert transformations, respectively from the left. The bottom row, however, contains time-frequency Fourier (left) and wavelet (right) spectrogram heatmaps, which were normalised such that the maximum value is one. Additionally, on all graphs, the black line denotes the instantaneous frequency of the input components.



(a) Time series

(b) Fourier spectrum

Figure 7.9: Signal generated for order comparison experiment. Exact values for the KurSL model are presented in Table 7.3 (sec. 7.1.2). Figure (a) displays time series of the sum of all oscillators, whereas (b) is the Fourier spectrum. Vertical dashed lines on the spectrum denote peak positions determined by the KurSL method.

Table 7.3: Parameters for the KurSL model used to generate the signal in the experiment with order comparison (sec. 7.1.2). Units of initial frequency ω and coupling strength k are rad/s, and initial phase and amplitude are radians and arbitrary units, respectively.

	ω	θ	ρ	k^1_{ij} [$\frac{\text{rad}}{\text{s}}$]					k^2_{ij} [$\frac{\text{rad}}{\text{s}}$]					k^3_{ij} [$\frac{\text{rad}}{\text{s}}$]				
n	[$\frac{\text{rad}}{\text{s}}$]	[rad]	[arb. u.]	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	81.68	2.0	2.0		5.10	1.90	1.90	10.10		-2.10	10.70	-1.90	9.10		-1.50	0.50	0.20	-1.10
2	62.83	0.0	1.8	7.90		13.27	4.99	8.55	4.08		0.06	8.94	-0.78	3.88		2.78	3.69	-2.65
3	37.70	0.5	3.2	-7.04	-1.32		-1.02	2.35	1.26	1.44		2.71	8.42	-5.54	-5.11		0.06	-0.66
4	31.42	2.5	1.5	-0.64	3.02	-1.05		-2.23	10.22	1.34	1.62		1.02	3.66	3.09	2.32		-0.57
5	12.57	1.2	1.3	7.57	-0.24	-0.08	-0.10		0.69	-0.03	0.51	-0.80		-1.72	0.14	0.28	-0.16	

Table 7.4: Parameters for the KurSL model used to generate the signal in the experiment with dynamical analysis (sec. 7.2). Units of initial frequency ω and coupling strength k are rad/s, and initial phase and amplitude are radians and arbitrary units, respectively.

	ω	θ	ρ	k^1_{ij} [$\frac{\text{rad}}{\text{s}}$]						k^2_{ij} [$\frac{\text{rad}}{\text{s}}$]						k^3_{ij} [$\frac{\text{rad}}{\text{s}}$]					
n	[$\frac{\text{rad}}{\text{s}}$]	[rad]	[arb. u]	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
1	147.6	3.01	1.7		12.08	13.70	9.92	-7.98	12.65		-11.01	13.93	10.56	0.08	0.72		10.67	11.95	13.89	3.63	-7.21
2	92.82	5.09	0.8	8.25		8.22	-8.25	6.80	10.97	-3.12		3.12	12.78	3.99	-0.38	1.96		0.56	5.94	-3.42	-5.70
3	85.31	1.37	0.7	2.38	-6.77		7.29	-4.83	-1.36	-1.50	-0.50		6.14	-3.96	6.19	9.39	14.15		-0.88	9.04	3.37
4	69.08	3.18	0.9	8.65	3.02	-5.26		-2.29	-4.53	9.85	-0.49	-0.51		8.53	-1.77	-2.45	-2.16	4.44		8.89	-3.24
5	39.96	5.93	0.6	1.39	-2.04	-2.52	1.56		3.23	-0.05	6.31	2.90	-4.29		-2.18	-1.13	-2.96	1.44	2.34		2.35
6	25.73	0.28	1.1	-2.14	-0.35	2.31	-0.62	-1.64		-0.88	1.14	1.67	-0.57	1.48		3.92	0.53	0.21	-2.74	1.93	

oscillators parameters of which are presented in Table 7.2. The MCMC computation was set to perform 200 iterations out of which the initial 10% were discarded. Moreover, despite having much larger parameter space than in the previous experiment, i.e. optimising over 75-dimensional parameter space, 200 walkers were deployed to search for the best solution. Cumulative results from both the first and third KurSL order experiments are shown in Appendix in the form of probability distributions. Four sets of parameters, i.e. intrinsic frequencies, initial phases, signal's amplitudes and all coupling strengths, are presented, respectively, in Figures B.2, B.3, B.4 and B.5 for the 1^{st} order KurSL and Figures B.6, B.7, B.8, B.9, B.10 and B.11 for the 3^{rd} order. Vertical lines in these plots indicate parameter values which either maximise global likelihood (red) or were used to generate input signal (black). When comparing probability densities, one can notice that in case of the 3^{rd} order KurSL there is more significant variability resulting in a bigger spread in the density. One of the reasons for such behaviour is higher dimensionality of the parameter space. Size difference causes changes within a single parameter to have a smaller effect on the global reconstruction. As it has been mentioned in Section 5.3, increased number of parameters and iterations can help find the global optimum, but it comes with the computational time cost.

The experiment demonstrates the difference between both obtained parameter sets and the one used to generate the input signal. One can see that for the 1^{st} order KurSL the distance to the true set is much higher than for the 3^{rd} order KurSL, which seems to fit the input signal parameters very closely. Reconstructed time series and their mean square errors (MSE) when compared to the input are presented in Figures 7.10 and 7.11 for the first and third orders, respectively. Rows in these Figures contain time series for corresponding oscillators with the top row being a sum of all oscillators. In all graphs, the green colour relates to the generated signal, whereas the red indicates reconstructed data. As it can be seen, results obtained with the third order KurSL (MSE=0.244 [arb. u.]) are similar to the input signal. Moreover, despite not obtaining a perfect fit with the first order KurSL, reconstructed time series are visually comparable to the generated signal (MSE=0.799 [arb. u.]). Noticeable difference in the fifth oscillator suggests that these results are not necessarily the best fit in the entire parameter space. Similarly to the previous experiment, increased number of walkers

or the iterations should result in a closer fit.

In this experiment, the higher order method has provided better results. Such behaviour is expected to be more general since the increase of the order expands the parameter space granting additional degrees of freedom. The increase of degrees of freedom makes it more likely for the model to fit the data. However, as has been discussed in Section 5.3, a better fit does not necessarily mean that results are more significant, especially when data contain noise. In a trivial case, when the underlying data generation process is simple, there is no need for adding extra complexity to the model. Likewise, even when dealing with more structurally complex signals, one might not be interested in extracting its exact representation. It can be argued that in this experiment the first order KurSL has provided with close representation. Although, the definition of closeness depends on the used metric, which itself can be dependent on the experiment. As provided Figures highlight, a high degree method produces more prominent variability in parameter distributions. Further improvement under such conditions requires an increased number of walkers and performing more iterations, both of which has an apparent computational disadvantage and can significantly increase the experiment duration. The choice of the method's order can have a significant impact on the representation. Although it is expected that structurally complex signals will have a higher order of the underlying generating process, this does not necessarily require applying complex methods. As discussed in this experiment, when choosing the order, one needs to balance between the reconstruction's expected fidelity, the precision of estimations and the computational complexity.

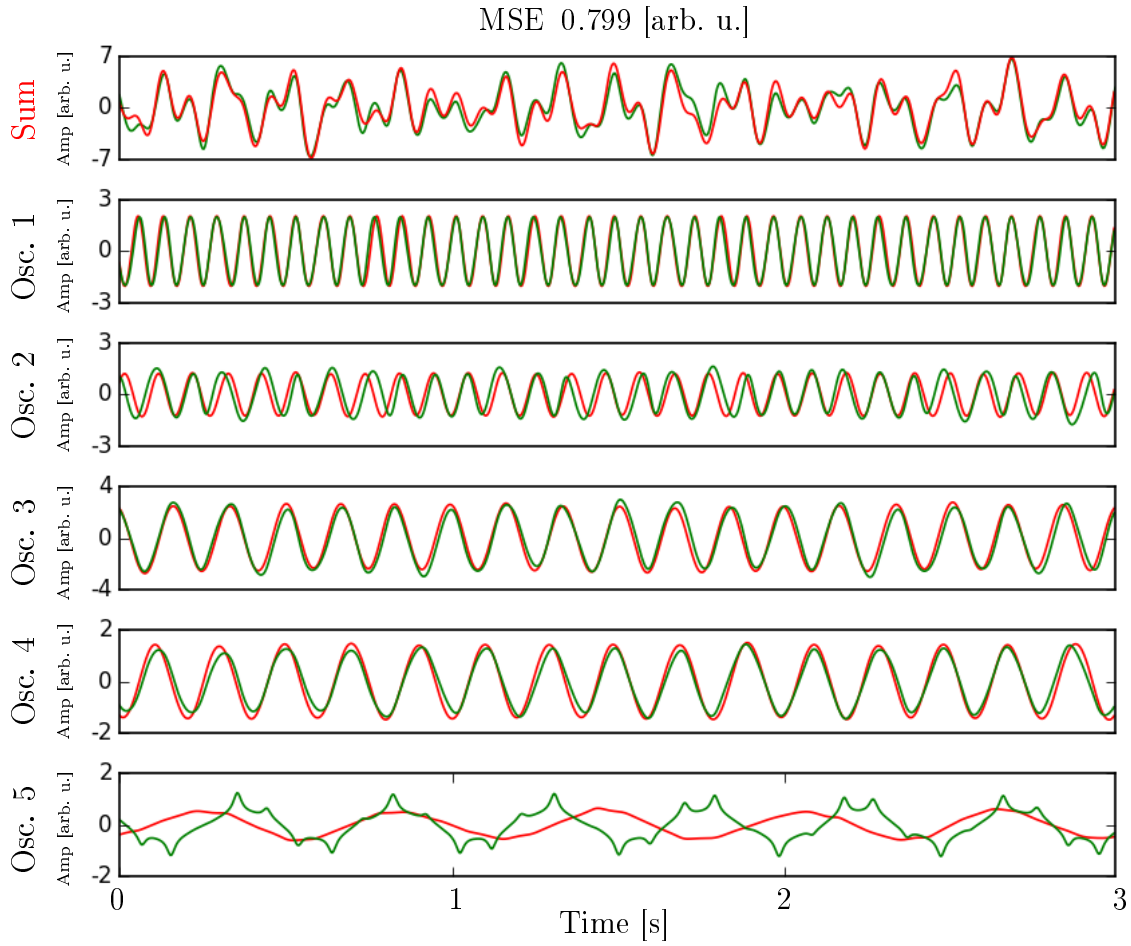


Figure 7.10: Reconstructions obtained using the 1st order KurSL model in the comparison experiment. The first row contains the original time series and its reconstruction, whereas following rows present comparison between respectively generated oscillators. Green and red colours indicate original and reconstructed time series, respectively.

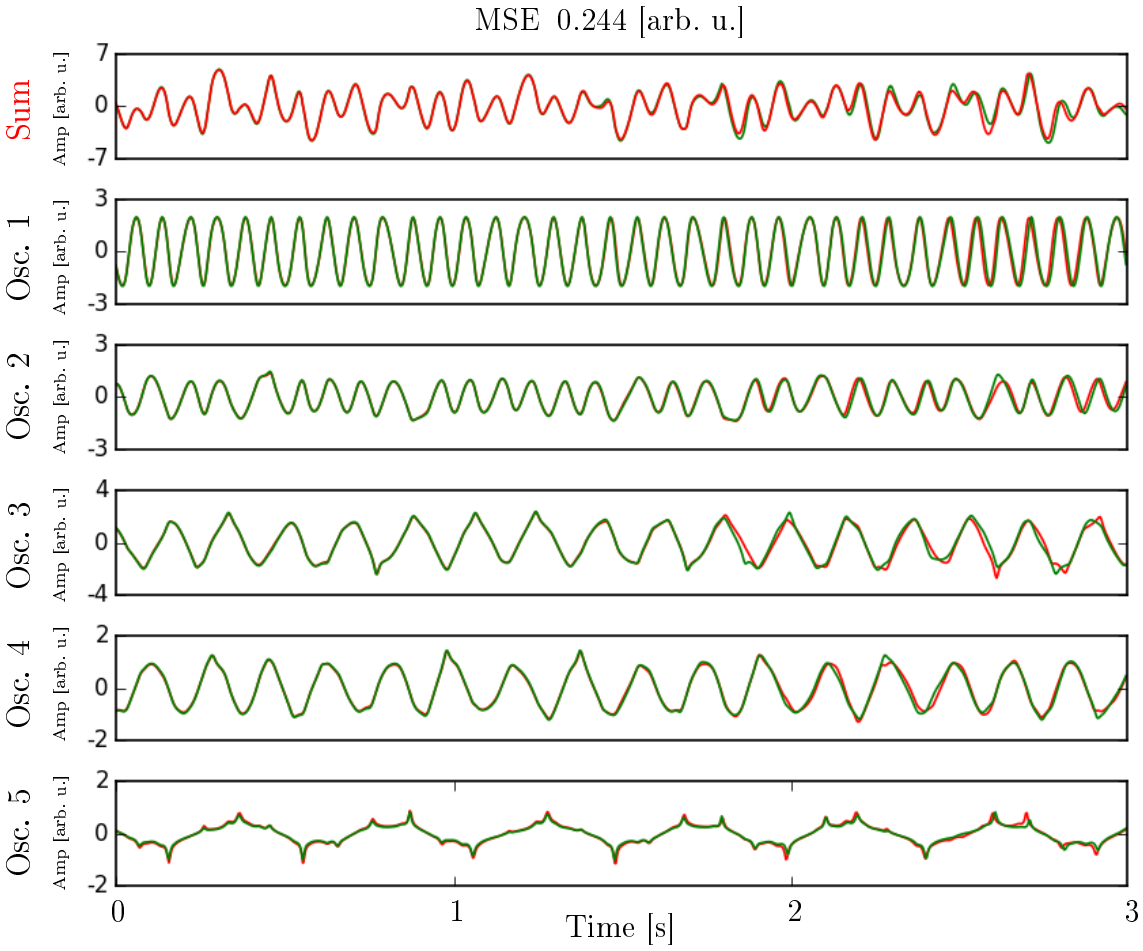


Figure 7.11: Reconstructions obtained using the 3rd order KurSL model in the comparison experiment. The first row contains the original time series and its reconstruction, whereas following rows present comparison between respectively generated oscillators. Green and red colours indicate original and reconstructed time series, respectively.

7.2 Windowed analysis of a simulated signal

As it has been mentioned in Chapter 3, all systems are affected by the environment in which they are placed. Sometimes such interaction can have a significant effect on the system by changing its properties over time in unpredictable manner. Even if the system is known to have certain characteristics it might be too difficult to describe it based on the whole signal. In these cases, it can be beneficial to consider short time frames for which analysis is easier to perform and observe how model parameters vary in time. One of such techniques is based on performing a sliding window analysis, where the whole signal is divided into segments and for each one of them a model is fit. Assuming that the system does not change its properties abruptly, small time shifts should not have a significant impact on the ideal representation[1] and thus parameters are also expected not to change significantly in time. Furthermore, parameters associated with the best fit for a given segment can be used as initial parameters for the following segment.

Demonstration of such analysis was performed on a synthetic signal generated with the 3^{rd} order KurSL model. Parameters used in the simulation were sampled from uniform distributions with ranges $\omega \in [10, 200]$ $\frac{\text{rad}}{\text{s}}$, $\theta \in [0, 2\pi)$ rad, $\rho \in [0.5, 2]$ and $k \in [-15, 15]$ $\frac{\text{rad}}{\text{s}}$ and all obtained values are presented in Table 7.4. Signal was generated for time array spanned from 0 to 7 seconds with sampling frequency of 200 Hz. Visualisation of the input time series and its Fourier spectrum are presented in Figures 7.12a and 7.12b, respectively. Dashed vertical lines on spectrum Figure relate to 8 most impactful peaks, which were selected by the KurSL algorithm configured to explain energy ratio of $E_\epsilon = 0.1$. For this experiment it was decided to modify algorithm's parameter to limit it only to the first 6 oscillators (red vertical lines) or equivalently energy ratio $E_\epsilon = 0.26$. Justification for such subjective modification is based on the purpose of this experiment, which is to present behaviour of dynamic KurSL. Nevertheless, it is acknowledged that the algorithm for determining parameters could be improved, e.g. by taking into account a coupling between oscillators which seems to be responsible for additional wide satellite peaks. Values of all detected peaks

[1] It is assumed here that the step or changes to the system are not big. If they were significant they should be identified and included in the model.

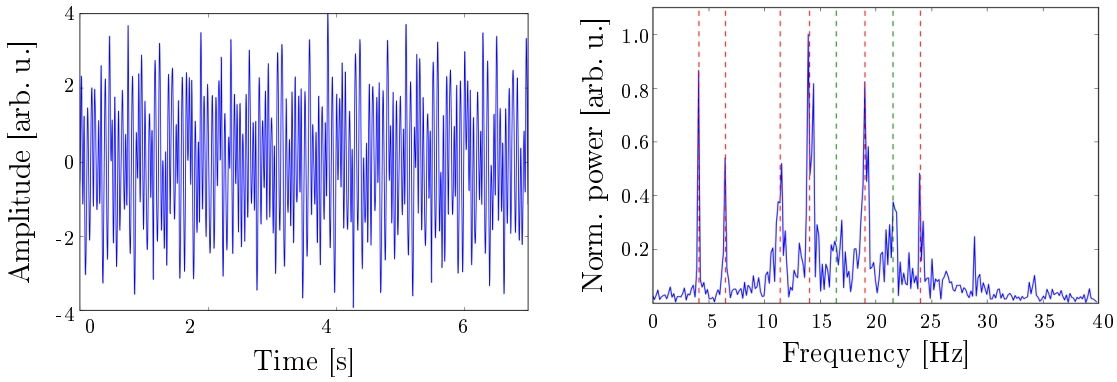
n	1	2	3	4	5	6	7	8
ω [$\frac{\text{rad}}{\text{s}}$]	88.6	25.9	119.7	40.8	71.8	150.9	135.6	103.4
$\omega/2\pi$ [Hz]	14.1	4.1	19.1	6.5	11.4	24.0	21.6	16.5
ρ [arb. u.]	2.7	0.6	2.3	1.8	1.8	1.1	1.5	1.3

Table 7.5: Initial intrinsic frequencies ω, frequencies f and amplitudes ρ obtained for dynamical KurSL with simulated signal experiment. Parameters are sorted based on their detection order. Only the first 6 oscillators were used as initial values for the KurSL method.

are presented in Table 7.5 where they are sorted by the order of detection.

For dynamical analysis signal was divided into 3 s long segments with 1 s step starting from 0 s. All MCMC optimisation were set to have 300 walkers searching the parameter space in 300 iterations. For each segment prior parameters' distributions were assumed to be the best set from the previous segment, except for the first window for which they were determined accordingly to the algorithm assuming $E_\epsilon = 0.26$. Moreover, a reconstruction threshold was implemented indicating when to stop the process for current segment. Threshold T_h value was set to 0.1, meaning that whenever mean square error MSE between the reconstructed S_R and input signal S_I was $\text{MSE}(S_R, S_I) < 0.1$, the simulation for current segment would stop and move to the following time frame.

Reconstructed time series and their Fourier spectrum are presented in Figure 7.13, where each row corresponds to a different analysis segment. These results (red) are compared to the input signal (green) for corresponding segment. For all reconstructions, independently for time series and spectra a measure of residual energy (RE) has been calculated as a mean square error (MSE) of reconstruction divided by the total energy of the input. Interestingly, it can be observed that the RE values differ significantly when comparing results for Fourier power spectra and time series. In case of the former the highest RE_F value is 0.432, whereas for time series in two cases the RE_T value is above 1. Since RE refers to the amount of energy that has not been explained by the reconstruction, RE > 1 means that additional variability has been introduced into signal. Such outcome is expected to be a result of difference in phases



(a) Time series

(b) Fourier spectrum

Figure 7.12: Signal generated using the 3rd order KurSL model with parameters from Table 7.4. Figure 7.12a displays time series of a sum of all oscillators, whereas 7.12b is its normalised Fourier spectrum. Vertical lines on spectrum Figure indicate the first 8 detected; red colour denotes 6 first peaks and the following 2 are in green.

when comparing components in the input system with corresponding components in the reconstruction. The reasoning is in accordance with observed similarities in presented Fourier power spectra which contain only information about the amount of given frequency without any notion of its phase. Nevertheless, it is noticeable that with the increase of segment's order the fitting performance also increases. This is a result of parameters estimation propagation from previous to the following segments. Quantitatively one can compare time series and spectrum RE values of the first segment for which $\mathrm{RE}_T = 2.094$ and $\mathrm{RE}_F = 0.432$ with the last segment for which $\mathrm{RE}_T = 0.619$ and $\mathrm{RE}_F = 0.124$. The decrease is more than threefold implying significant improvement.

Increase of performance as a function of segment's position order can be also observed in dynamics of all parameters. From all reconstructions the best parameter sets were compared and their results are presented in Figures B.12 for intrinsic frequencies, B.13 for phases, B.14 for amplitudes and B.15, B.16, B.17 for, respectively, the first, second and third coupling k strengths. In all these figures and for all parameters black vertical line indicates the actual value used to generate synthetic signal. One can see that at segment starting with time $t = 0$ s almost all estimated values have the biggest deflection from the truth when compared to other segments for which the values are relatively similar. This is especially visible in the amplitude and intrinsic frequency parameters. As it has been already mentioned, the element that distinguishes the first

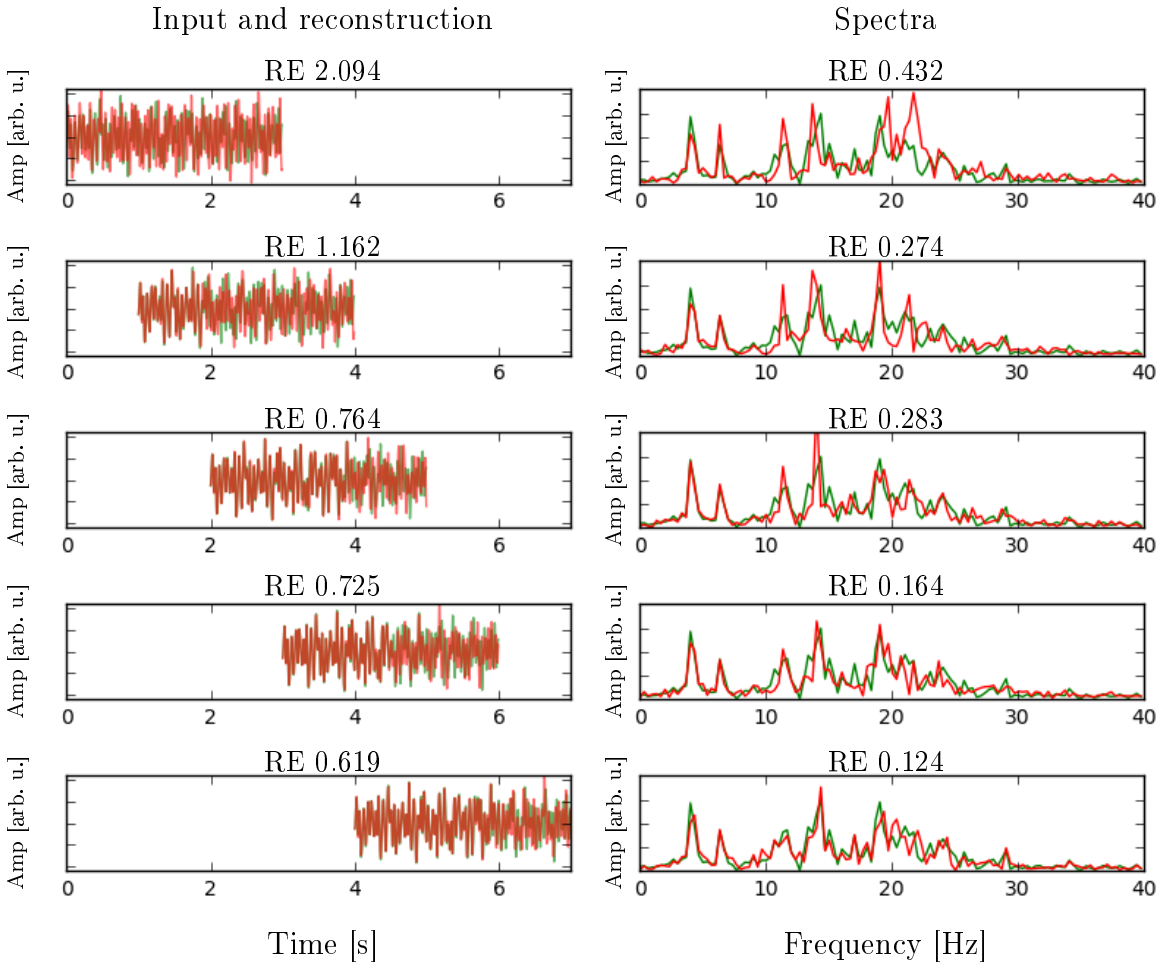


Figure 7.13: Comparison between reconstructed (red) and the synthetic input (green) signals presented for all analysed segments (sec. 7.2). Left and right columns contain overlaid time series and their Fourier spectra for each segment respectively.

segment is the process in which the initial parameters are estimated. The following segments take their estimates as the most optimal values from the previous. For the initial segment, however, estimates are based on the peak fitting algorithm 1 which focuses on finding oscillator's intrinsic frequencies' positions and amplitudes. Estimating coupling factors is difficult as they can have many effects on the spectrum such as widening peak, introducing satellite peaks or shifting coupled oscillators towards each other. Unfortunately, as mentioned in Section 6.1.2, most coupling estimating methods are unfit for the KurSL purpose as they assume the input to be multivariate. For this reason, unless parameters are input with specified ranges, the optimisation method considers relatively wide scope of search and big variance when proposing initial values. Under such procedure it is unlikely to quickly find a good global estimate and thus more iterations and walkers are necessary. One of possible approaches to increase performance of optimisation method is to increase sample size of input data, which would impose additional constrains on cost function. However, extending scope of input invalidates the assumption on which the dynamic approach is based, i.e. that whole signal has too complex structure and dividing it into smaller segments allows for more accurate representation. This then allows to analyse how parameters change in time, enhancing general understanding of the system. Overall one has to balance the costs of applying any method depending on availability of all resources.

Presented experiment shows that for more structurally complex signals it is advantages to perform dynamical analysis. This allows for improved reconstruction performance due to more accurate initial estimations of parameter. Additionally, this suggests that in cases when analysis in stationary manner does not produce satisfying results and visual inspection suggests that these results could be improved, one could attempt analyse the data again using the best estimates. As it has been discussed, such procedure should mitigate imperfections of the estimating algorithm.

7.3 Analysis of empirical EEG signals

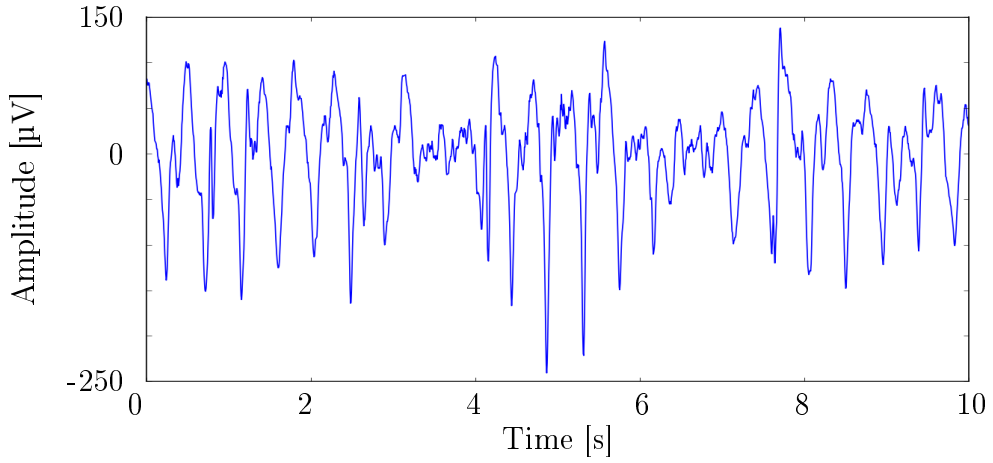
In this section, two experiments on empirical data are presented and discussed. The first one focuses on a short stationary signal and its analysis performed using the KurSL of 3^{rd} order. Section 7.3.1 explains the experiment and presents obtained results. In the second experiment (sec. 7.3.2) input signal is analysed with a moving window KurSL 3^{rd} order method. All data used in the experiments are EEG signals provided by [154] and are freely available to access[2]. Each signal from this dataset was digitised at a sampling frequency of 173.61 Hz and has the length of 4096 samples. The authors have also filtered all included signals using a band-pass filter with cut off frequencies 0.53 and 40 Hz (12 dB/oct). The characteristic of the data is that they were recorded from patients suffering from epilepsy and some of the recordings were recorded during an epileptic seizure. Such signals are interesting from the analysis perspective, as they often have visibly periodic structure, significantly different to normal state EEG. The precise mechanism behind this phenomenon is yet to be discovered; however, it is commonly considered to be a result of abnormal synchronisation and increase of coupling between specific brain regions [155–157].

The purpose of this section is to provide an exemplary scheme on how empirical signals can be analysed using the KurSL method. Despite having some intuition regarding the nature of the data generating system, the exact underlying process is unknown. This makes such analysis difficult to validate objectively and thus no claims regarding the true decomposition should be made. Nevertheless, some general characteristics, such as increased activity at a particular frequency, are expected to be observed.

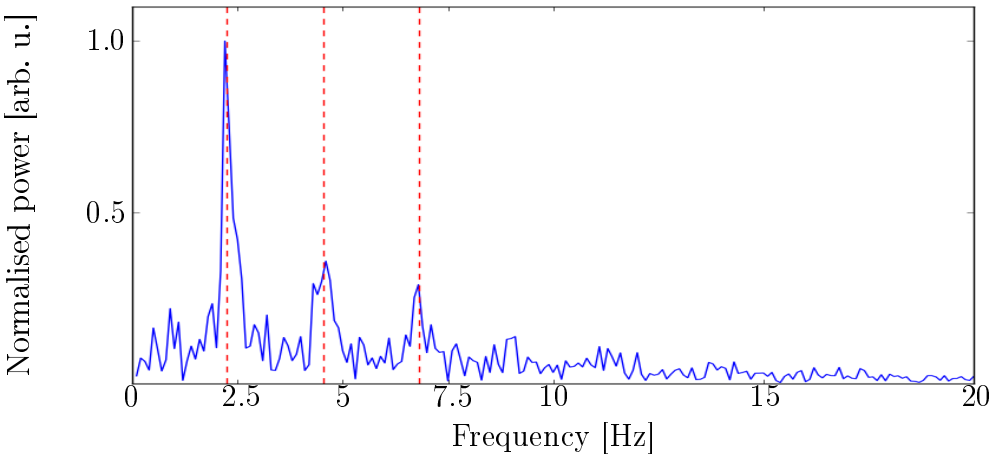
7.3.1 Static EEG analysis

In the first experiment on an empirical signal, EEG recordings are analysed in a stationary manner. The signal comes from a dataset provided by authors of [154], where it has a code F11. It was recorded from within the epileptogenic zone in a relaxed and awake state with eyes opened. The analysis concentrates on the first 10 seconds of the EEG signal for which the Figure 7.14 presents its time series and frequency power

[2]Web address to the data is provided in the main article and on the author's webpage.



(a) Time series



(b) Fourier spectrum

Figure 7.14: Empirical EEG time series (Fig. 7.14a) and their Fourier spectrum (Fig. 7.14b) used in the experiment with the 3$^{\text{rd}}$ order KurSL method (sec. 7.3.1).

spectrum. In the figure, one can see that the spectrum consists of a slowly decaying background noise and three distinct peaks at frequencies $F = \{2.24, 4.55, 6.82\}$ Hz. The second and third values appear to be very close to integer multiples, 2 and 3 respectively, of the first frequency indicating that they are its harmonics. This suggests that within the signal there is a pattern which repeats itself with the primary frequency, although its shape is not sinusoidal. Under such circumstances, this behaviour could be explained either by proposing a periodic pattern and attempting to find its structure or as many strongly coupled harmonic oscillators. The KurSL method focuses on the latter approach.

The KurSL method was initiated with parameters estimated according to the

algorithm provided in Section 6.2 with the energy ratio set to $E_\epsilon = 0.2$. An increase of the energy ratio E_ϵ compared to a typical value of 0.1 is due to a high amount of background noise spread throughout the spectrum. Under such conditions, 3 oscillators were obtained with their initial values presented in Table 7.7. Identified oscillators could be harmonics as the second and third oscillator have frequency values close to 2 and 3 times of the first (fundamental) frequency, respectively. As noted previously this suggests that there is a strong coupling within the system.

Unfortunately, the strength and form of coupling are difficult to determine without knowing much about the system which generated the signal. A higher value of the KurSL order means that the model is more flexible and can explain a broader range of signals; however, it is also more susceptible to conform to the noise. For this reason, the processing algorithm in this experiment was updated with an extra step to find the most suitable order value for this signal. The fitting procedure started with initial parameters determined using the peak matching algorithm and assumption that there is no coupling between oscillators. Once the optimisation algorithm finished the KurSL model its complexity has been modified by increasing the order number and fitting updated model again. The new model was initiated with the best parameters from the previous run and added coupling parameters were set to 0. The experiment stopped upon finding optimum for order $M = 4$.

For each parameter optimisation MCMC run for 600 iterations and 6 walkers per parameter, where the number of parameters is given by

$$\#_{\text{param}}(\mathcal{N}, M) = \mathcal{N}(3 + M(\mathcal{N} - 1)), \tag{7.2}$$

where $\mathcal{N}$ and M are the number of oscillators and the order, respectively, which for this experiment evaluates to 54 for 0 order KurSL and 252 walkers for KurSL of order 4. Since every fitting starts with the best estimate from the previous step and only accepts better reconstruction, it is natural for the error to go down with a higher order number. Difficulties here are twofold: not only with the increase of order number the flexibility increases, which can lead to focusing on reconstructing noise, but also the computational cost raises. For this experiment, which is to determine a parameter set for oscillatory empirical signal, more emphasis should be given to derivative of the

Table 7.6: Cost values obtained for different orders M of KurSL model.

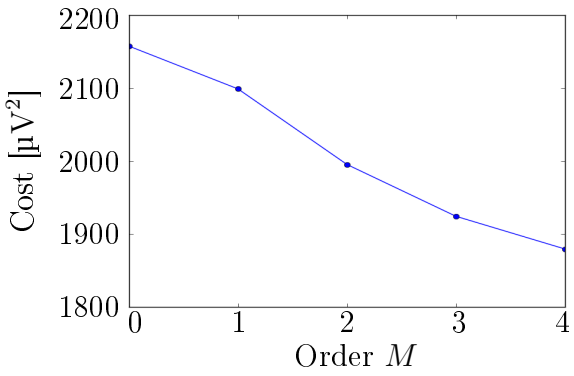
M	0	1	2	3	4
# Parameters	9	15	21	27	33
Cost [μV^2]	2157.4	2098.8	1995.2	1924.2	1879.4
Difference [μV^2]	—	-58.6	-103.6	-71.0	-44.8

Table 7.7: Initial parameters used in the stationary EEG experiment.

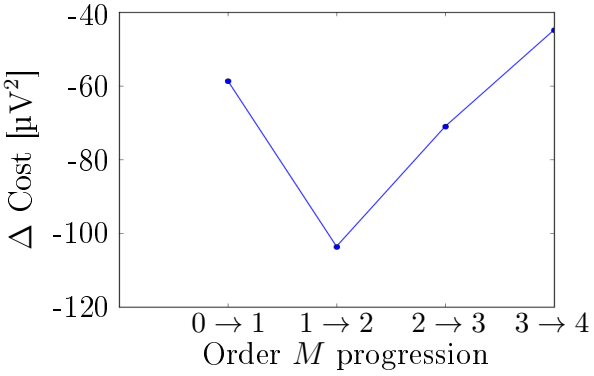
n	ω [$\frac{\text{rad}}{\text{s}}$]	θ [rad]	ρ [μV]
1	42.86	3.51	20.52
2	28.60	1.93	24.04
3	14.32	6.09	26.35

cost curve and not their absolute values. Reconstruction errors and their differences divided by the change in the number of parameters are presented in Table 7.6 and Figure 7.15. These resources confirm that the reconstruction error indeed is monotonically decreasing with the increase of the model's order. The difference, however, has visibly the biggest increase in error reduction when moving from the 1^{st} to 2^{nd} order. Such reduction suggests that the rate of improvement decreases for orders bigger than 2. For this reason KurSL of order $M = 2$ was considered to be sufficient model, capturing the most of signal's characteristics with the least necessary number of parameters.

The best set of parameters, i.e. the one that maximises joint probability distribution and minimises cost function, is presented in Table 7.8. Components produced with these parameters are shown in Figure 7.17 where the first row compares the input signal with model's reconstruction, and the following rows correspond to respective KurSL oscillator (red) and its instantaneous amplitude (blue). It can be seen that all amplitudes have significant variations. Of all oscillators, the first one has the least modulations. For two other oscillators, the amplitude changes about up to 10 µV, which for the oscillator 2 is about 50% of its maximum deflection. Both these oscillators have sharp peaks which appear in pair with increased activity in the input signal. Nevertheless, despite similarities in behaviour, the reconstruction does not adequately capture the structure of the input time series.



(a) Absolute cost values per order M.



(b) Cost difference with order M increase.

Figure 7.15: Cost values (Fig. 7.15a) and their piecewise differences divided by the change in the number of parameters (Fig. 7.15b). These results were obtained for a range of order M values in a stationary EEG experiment.

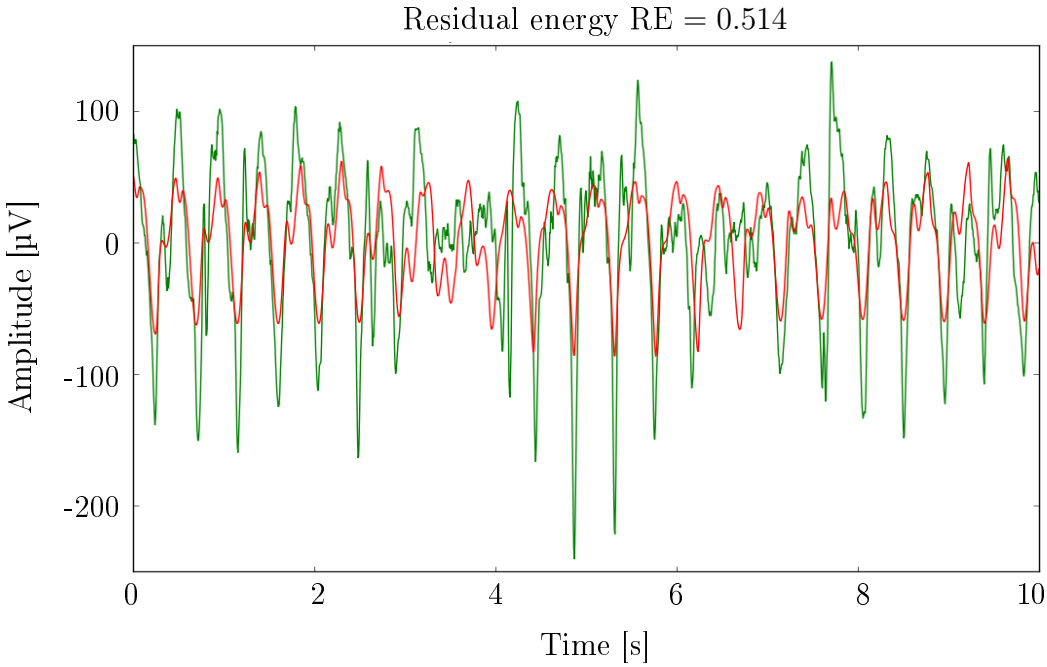
Table 7.8: Parameters that minimised the cost function when used 3rd order KurSL model on EEG stationary signal.

				k_{ij}^1 [$\frac{\text{rad}}{\text{s}}$]			k_{ij}^2 [$\frac{\text{rad}}{\text{s}}$]			k_{ij}^3 [$\frac{\text{rad}}{\text{s}}$]		
n	ω [$\frac{\text{rad}}{\text{s}}$]	θ [rad]	ρ [µV]	1	2	3	1	2	3	1	2	3
1	44.79	0.76	15.28		-1.73	-3.97		-6.09	0.53		-6.20	-0.48
2	27.51	3.47	27.69	-1.52		-3.63	6.38		-0.76	-4.38		3.14
3	13.44	6.28	55.97	-1.81	1.98		-1.62	-0.11		-0.02	0.50	

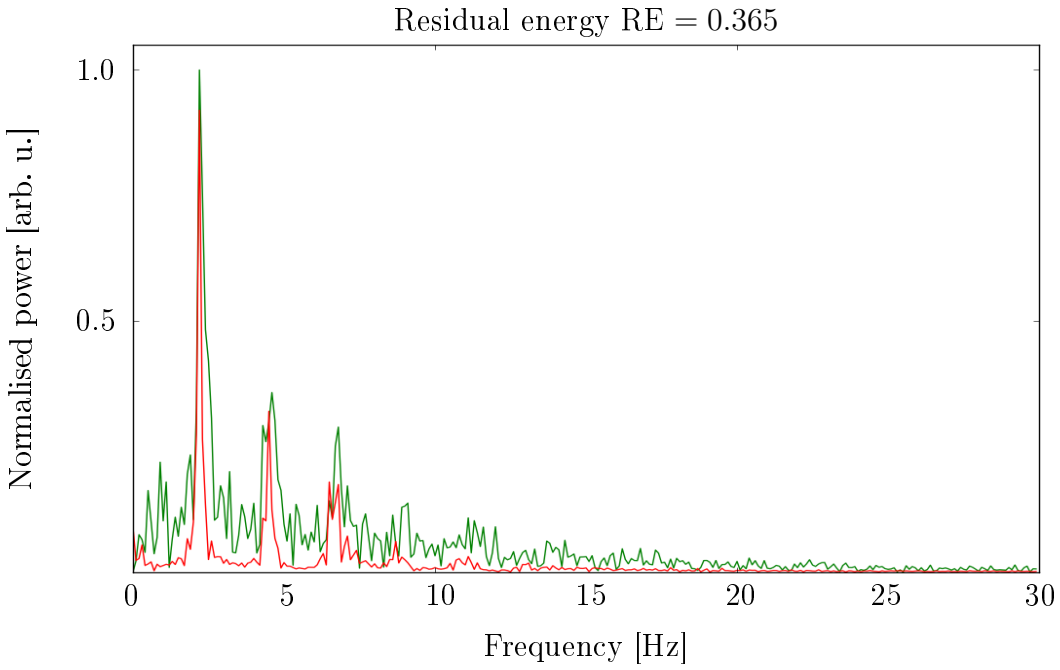
Unfortunately, as it has been stated before, the exact mechanism behind the system generating EEG epilepsy data is unknown. Lack of this knowledge limits the methods performance validation and whether the decomposition and obtained parameters have empirical meaning. Nevertheless, one can infer something about the method's properties by comparing its results to the input signal itself, and to results obtained with other data analysis methods.

The general approach of validating obtained results is to compare their fit to the input signal. Such comparisons are presented in Figure 7.16, where overlaid reconstruction (red) and the original (green) signals are presented in both time (Fig. 7.16a) and Fourier frequency (Fig. 7.16b) domains, respectively. The visual comparison indicates that the KurSL method has proposed a signal which follows the general oscillatory trend of the input. Reconstruction has identified all dips, although in many instances their depths are not exact. The difference is due to a skewed structure of the input signal, making it difficult to fit both crests and sharp troughs. An attempt for quantitative comparison is through the residual energy (RE) metric. Despite imperfect reconstruction, a model with only three oscillators was able to explain about half (RE = 0.514) of the energy in time series. Relatively good fit can also be observed when comparing Fourier spectra. The location of all peaks has been successfully identified with amplitudes of the fundamental component and its first harmonic closely matching their counterparts. In case of the second harmonic, it seems that peak's amplitude could be possibly bigger, which might suggest that obtained results are not necessarily the best global fit. However, with only these three components the visible residual energy is RE = 0.365. Quantitative difference between both comparisons is expected as the power spectrum only reflects the existence of particular frequency and loses all information about its phase and modulations. Nevertheless, additional components could be included to reconstruct the signal better. Further error reduction might include modelling the noise separately and removing it from the data. Using the KurSL model, this could be done by assuming many strongly coupled oscillators with intrinsic frequencies evenly spaced within the observed frequency space.

As expected from harmonic series there is an additional peak located close to frequency 9 Hz. Its small amplitude, comparable with the level of surrounding noise,



(a) Comparison in time domain.



(b) Comparison in Fourier domain.

Figure 7.16: Comparison between input EEG signal and its KurSL reconstruction (sec. 7.3.1). In both time (Fig. 7.16a) and Fourier (Fig. 7.16b) domains, EEG and reconstructed signals are indicated by green an red, respectively. For both comparisons computed residual energy (RE), i.e. energy of piecewise difference divided by energy of the input (7.1), is included in the title for a particular figure.

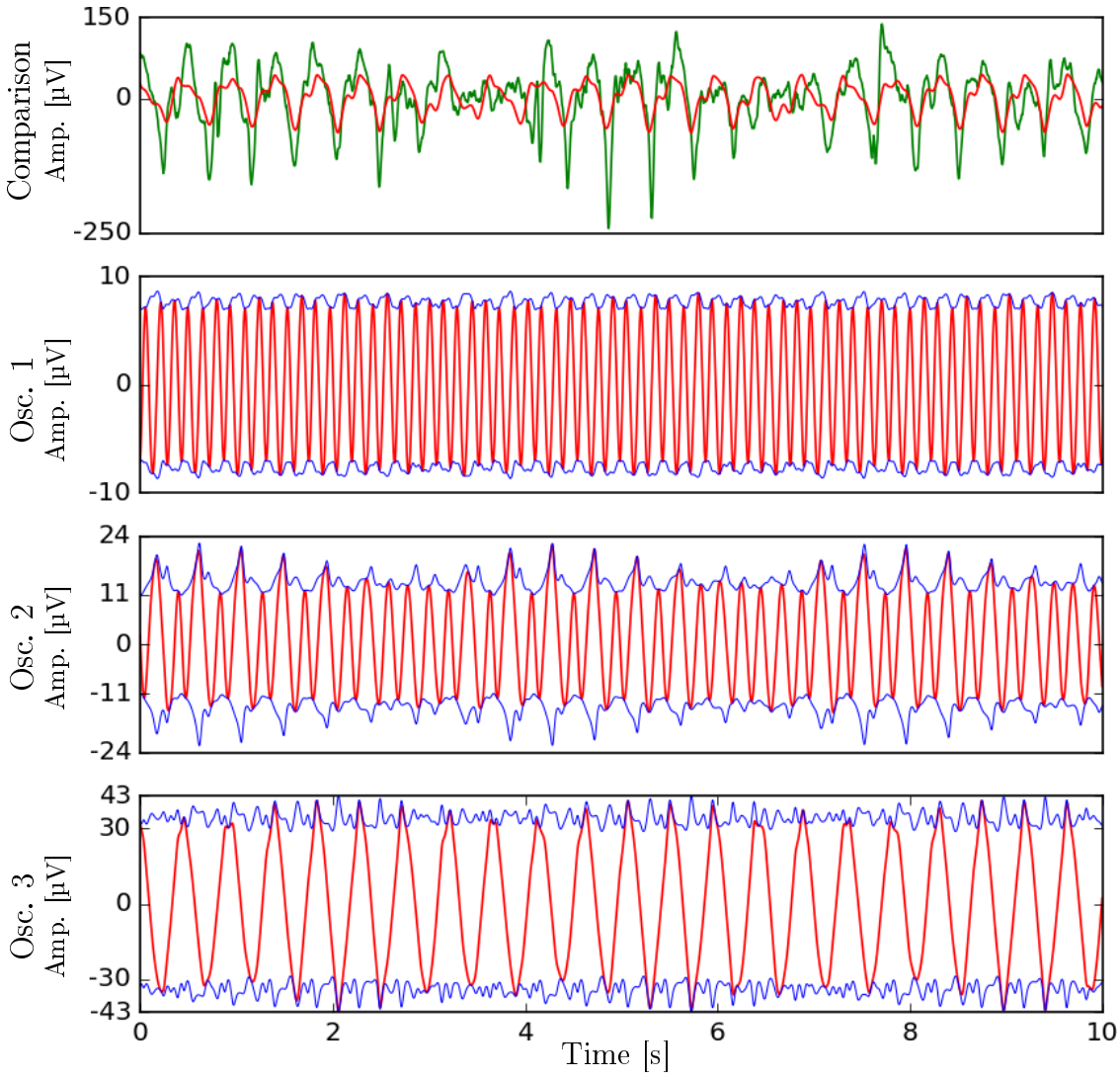


Figure 7.17: The figure represents the KurSL reconstruction of the signal in stationary EEG experiment (sec. 7.3.1). The first row contains the EEG time series in green and its KurSL reconstruction in red. Following rows present in frequency decreasing manner fitted oscillators (red) with their instantaneous amplitudes (blue).

makes it difficult to be noticed. Nevertheless, the proposed KurSL reconstruction has indicated a small peak at a location close to 9 Hz. This peak is a result of a strong coupling which can produce behaviour noticeable in the Fourier spectrum as satellite peaks to the coupled oscillators. It has been observed that relative position of these satellite peaks to the component peak is approximately equal to the difference in intrinsic frequencies of strongly coupled oscillators. In a case when two consecutive harmonics are interacting, e.g. $f_n = nf_0$ and $f_{n+1} = (n+1)f_0$, their difference is equal to the fundamental frequency f_0, which is making the satellite peak f_s to appear as another harmonic, $f_s = f_{n+1} + f_0 = (n+2)f_0 = f_{n+2}$. Such behaviour could also explain why there are two peaks reconstructed visible close to 2^{nd} harmonic ($f_2 \approx 6.5$): one is due to the explicit inclusion of the third oscillator, and the other is a result of coupling between the first and second oscillators.

Another validation of the method that can be done is to compare its results with those obtained through other methods. Here, the KurSL time-frequency (TF) representation is compared with the short-time Fourier transform (STFT), wavelet transformation (WT) and Huang-Hilbert transformation (HHT). The window length for the spectrogram was chosen to maximise the frequency resolution but with a constraint to acknowledge the slow variation in the amplitude. The shortest visible length of the amplitude modulation, and thus the longest considered window, starts and ends around 4 s and 6 s, respectively, which accounts for about 2 s (sample size $N_{FT} = 2\lfloor f_{fs} \rfloor = 346$). Keeping such length constant, it was visually verified that the spectrogram computed using the Hann window type and 75% overlap provides the clearest representation. In the case of the WT, the scaleogram was obtained using the Morlet wavelet with the central frequency set to $\omega_0 = 7\ \frac{\text{rad}}{\text{s}}$. Such central frequency, when used with a dyadic scale a, allows to match events at Fourier frequencies $f = \frac{\omega_0 + \sqrt{2+\omega^2}}{4\pi a} \approx \frac{2.25}{a}$ Hz [31], which closely match those present in the epilepsy F11 signal at 2.24 Hz and 4.55 Hz for scales $a = 1$ and $a = 0.5$, respectively. For comparison purpose some additional examples of spectrograms and scaleograms computed with different parameters are presented in Appendix B.18. In case of the HHT representation, EMD was configured to accept IMFs after they have been sifted $H_F = 5$ consecutive times with the difference in the number of zero-crossings and extrema to be at most one. Such setting produced

the smallest variance in the frequency domain when contrasted with other values of H_F.

Figure 7.18 presents the comparison of all computed TF representations. The top row from the left shows instantaneous frequencies for the KurSL and HHT methods, whereas the bottom row in the same order contains spectrogram and scaleogram. Additionally, the HHT representation is presented without the first component, as its high variance hinders readability by overlapping other components. In fact, the variance of the second (blue) and third (green) components is significant enough to obstruct any inference about the content with high frequencies. Its fourth component (red), however, manifests existence of an event with a smaller mean frequency (about 2 Hz) and less variability. This component agrees with the other representations. Interestingly both STFT and WT highlight a sudden drop in the amplitude of that component just before the fourth and after the sixth second. In case of the KurSL, these events are indicated by the change in variation, which is not visible in the time domain (Fig. 7.17). Although this might appear as an indication of different behaviours, the decrease in spectral amplitude and change in the variation of instantaneous frequency highlight the same phenomena. STFT and WT have a finite resolution in TF; thus, they highlight the density within defined ranges, which also increases when the variation increases. Similar behaviour is observed by these three methods for components located closely to frequencies 4.5 Hz and 7 Hz. Again, both STFT and WT indicate a sudden drop in amplitude. They also agree on the decrease of the overall intensity in the first segment, i.e. before 4^{th} s, is smaller than the rest. This event, however, was not captured by the KurSL either in the TF representation or in the time domain (Fig. 7.17). Given that STFT and WT indicate extra activity for higher frequencies, it is speculated that an injection of additional oscillators into the KurSL system could better reflect mentioned behaviours and improve overall reconstruction.

In conclusion, the result provided by the KurSL method agrees overall with those obtained through STFT and WT. They all indicated activities at similar positions in time-frequency representation. All these methods have provided more insightful representation than HHT, which in this particular case provided incomprehensible results. The advantage of KurSL shown in this experiment is its ability to highlight variations

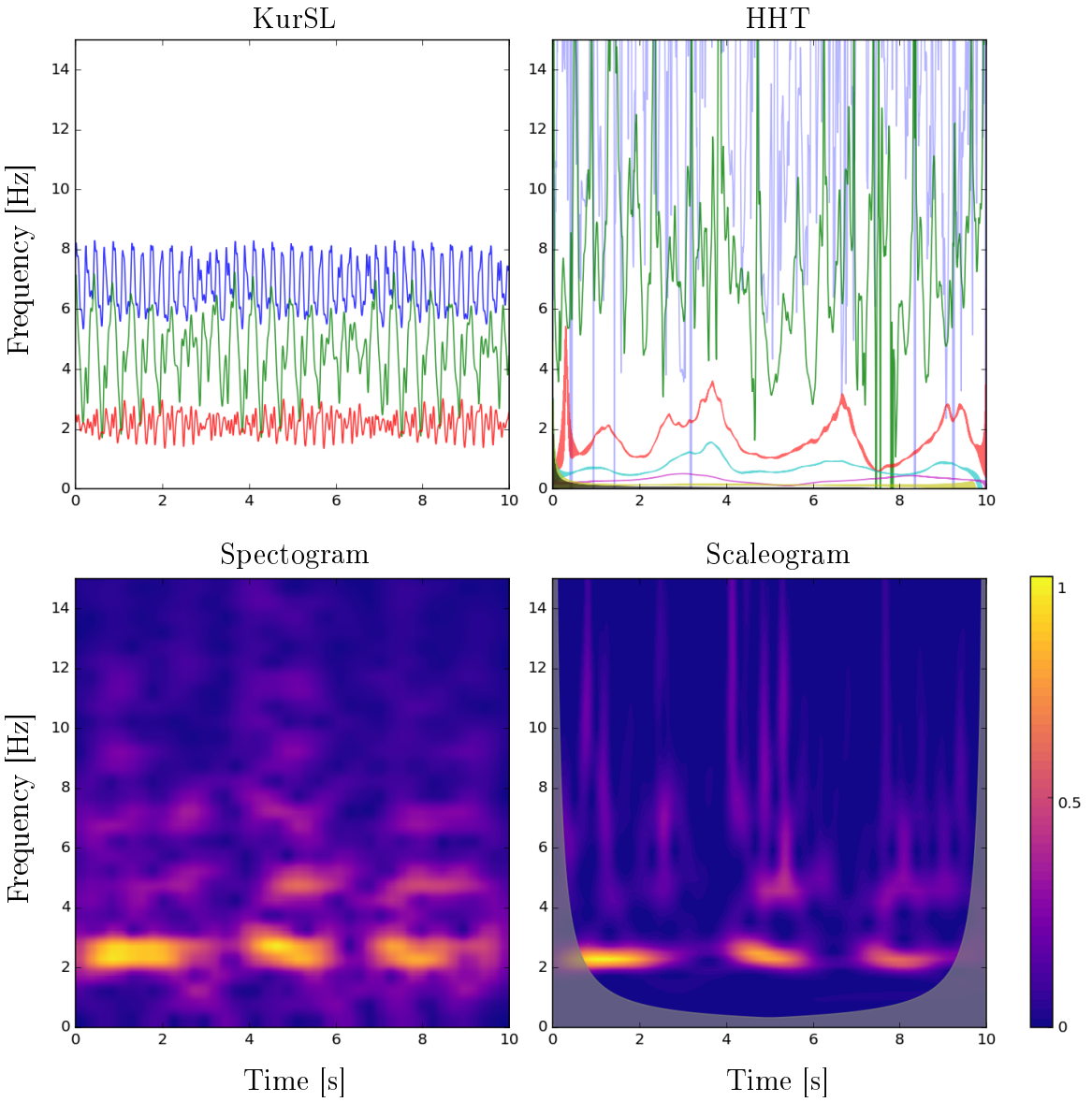


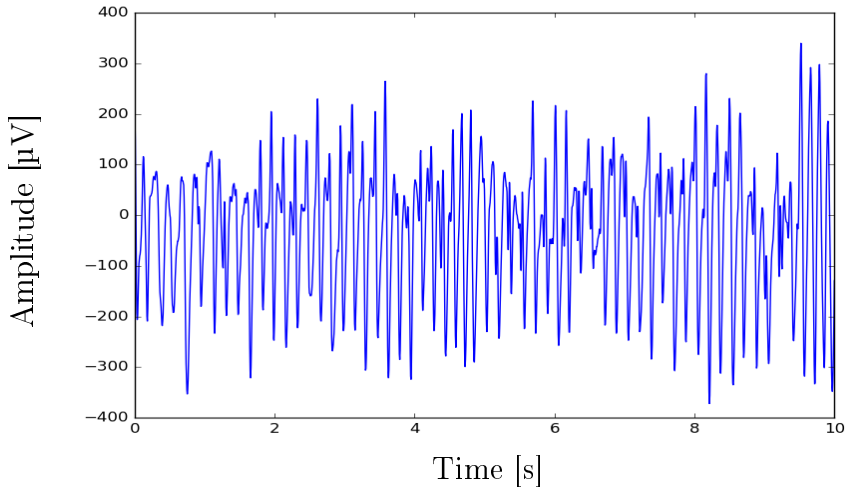
Figure 7.18: A comparison between different time-frequency representations for provided EEG signal (Sec. 7.3.1). The top row from left presents instantaneous frequency dynamics obtained with the 3$^{\text{rd}}$ order KurSL and the Huang-Hilbert transformation, where the EMD was configured with $H_F = 5$. The bottom left panel holds spectrogram obtained using STFT with about 2 s window and 75% overlap. The scaleogram presented in the bottom right corner was obtained with the Morlet wavelet of $\omega_0 = 7\ \frac{\text{rad}}{\text{s}}$. Both spectrogram and scaleogram were normalised such that the maximum value is 1 and the progression scale is presented on the right.

in the system through a finite and relatively small number of oscillators. Although such representation does not explain entirely the behaviour of the system, it is expected that by increasing the KurSL model's complexity one would be able to provide more detailed representation.

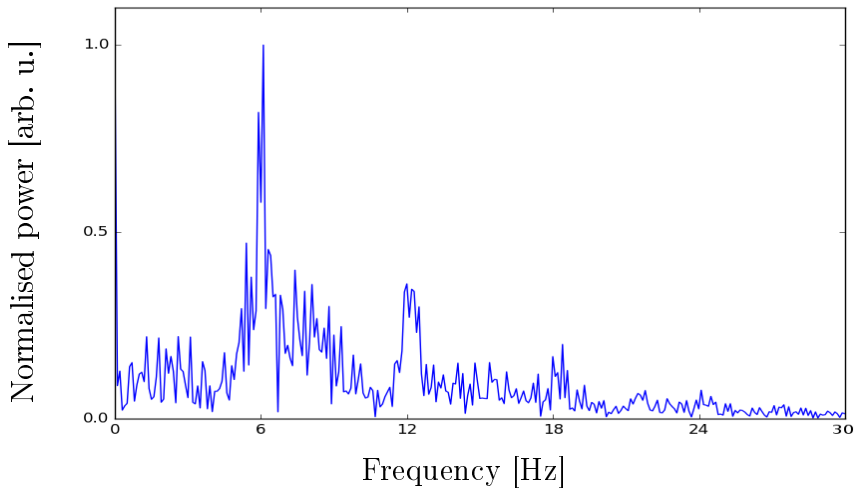
7.3.2 Dynamic EEG analysis

This section describes another experiment with an empirical EEG signal. In this case, the analysis is performed using a moving window to capture changes in the system. The EEG recording comes from the same database as for the previous example, i.e. database [154], where its code is *S63*. The measurement comes from a scalp during collected during an epileptic seizure. Such behaviour presents itself as a strong modulation of a particular period. Graphical representation of these time series and their Fourier spectrum are shown in Figures 7.19a and 7.19b, respectively. From these graphs, one can see that the patter have well-defined periodic structure manifested as an oscillation of 6 Hz with two less visible harmonics at 12 Hz and 18 Hz. Similarly to the previous experiment, presented here analysis assumes that such composition can be explained as a few tightly coupled oscillators in a noisy environment. Here it is explicitly considered that the KurSL is not able to fully explain all of the observed behaviours within the system. Such limitation might be because of the presence of non-oscillatory objects that influence the rest of environment. Nevertheless, identification could be made by assuming that on short segments the majority of observed behaviour is indeed due to coupled oscillators. Moving such window one would observe small perturbation to the fitted system and thus identifying how the oscillatory model evolves.

The experiment was performed using segments of 4 s length, which allows identifying Fourier frequencies with a resolution of $f_\Delta = 0.25$ Hz. Each of segments started 0.5 s after the previous one making a a total of 12 segments used for analysis. Such settings mean that the first and the last segments describe signal in ranges 0–4 s and 6–10 s, respectively. The initial number of the KurSL oscillators and their parameters were initiated using the peak detection algorithm (Alg. 1). Because of the assumption that the following segments are a result of the slightly perturbed KurSL system and unknown changes in the environment, one should use the first segment to initiate



(a) Time series



(b) Fourier spectrum

Figure 7.19: EEG signal used in the empirical dynamic analysis. Figure 7.19a contains time domain representation, whereas the Fourier spectrum is presented in Figure 7.19b.

the description process. For this reason, initial oscillators were identified using the signal in range 0–4 s for which the time series and Fourier spectrum are presented in Figures 7.20a and 7.20b, respectively. Due to visible noise in the spectrum and explicit assumption regarding system's content, the energy ratio E_ϵ was set to a value $E_\epsilon = 0.25$. Such setting allowed to identify six oscillators which positions are presented in Table 7.9 and graphically overlaid with the Fourier spectrum. The algorithm has correctly matched oscillations with the main peak at 6 Hz and its harmonics. Interestingly, the importance and complex shape of the dominant peak were highlighted by representing it as two oscillators. Such dual representation means that after removing the spike in the first iteration, the residual energy was large enough to consider it again in following iterations but with a smaller amplitude. In case of identified peaks at positions close to 2 Hz and 9 Hz, these are difficult to notice in the original spectrum due to general high background activity. This lack of contrast makes it difficult to provide meaning to these selections. Nevertheless, apparent activity in these regions justifies the attempt to include oscillators of such properties; they these provide additional means to reconstruct the signal.

Same as in the previous experiment with empirical EEG signal (Sec. 7.3.2), the order of the KurSL method was determined by evaluating the reconstruction cost using different orders. Again, the KurSL fitting was performed for 600 iterations and the number of walkers given as $\#_{\text{walkers}} = 54 + 4 \cdot \#_{\text{param}}(\mathcal{N}, M)$, where the number of parameters $\#_{\text{param}}$ is given by (7.2). With six oscillators and the order value M in range 0–4, inclusive, the formula provides values of $\#_{\text{walkers}} = [126, 246, 366, 486, 606]$. Initial parameters for the computation with $M = 0$ were estimated using the peak detection algorithm, and the following order evaluations used the best set obtained for the previous order. As in the previous experiment, added coupling values were initiated with zeros. For each order M and the best parameter set, a *residual energy* metric value was computed. All obtained numerical values are presented in Table 7.10, and their graphical interpretation is displayed in Figure 7.21. These results indicate that the best improvement with an increase of the KurSL M is from $M = 1$ to $M = 2$. Although $M = 2$ does not provide the best result, such order value was decided to be used in further analysis as it improves the system the most with the least number of

Table 7.9: Initial parameters used in the dynamic EEG experiment.

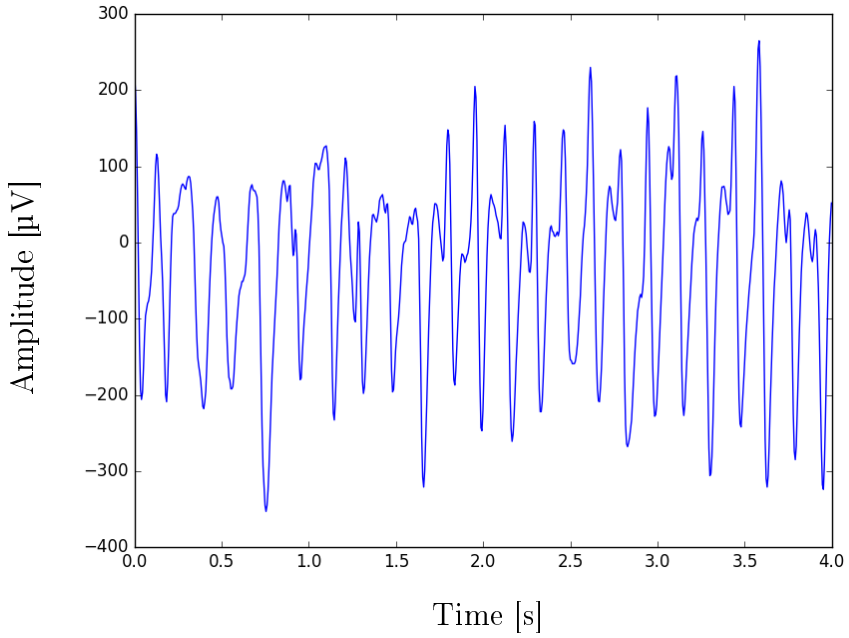
n	1	2	3	4	5	6
ω $[\frac{rad}{s}]$	113.48	76.52	58.26	38.36	34.27	11.88
$\omega/2\pi$ [Hz]	18.06	12.18	9.27	6.11	5.45	1.89
θ [rad]	5.55	0.17	6.18	3.38	3.38	5.48
ρ [µV]	10.81	35.51	15.80	17.66	48.05	15.75

Table 7.10: Cost values obtained for different KurSL model orders M in the dynamic KurSL experiment on EEG signal.

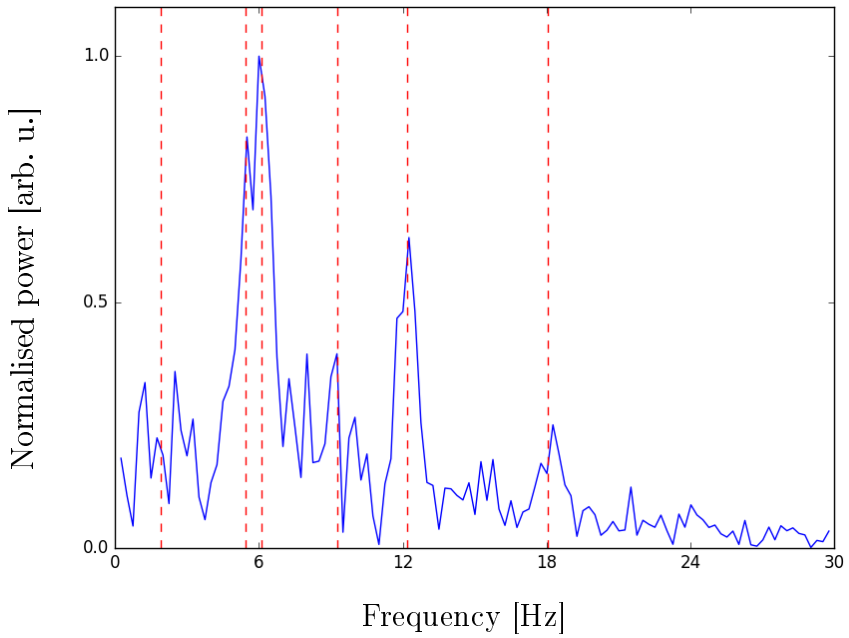
M	0	1	2	3	4
# Parameters	18	48	78	108	138
Cost [µV^2]	8597.8	8274.3	7312.1	7103.2	6502.9
Difference [µV^2]	—	-323.5	-962.3	-208.9	-600.2

parameters.

Obtained results for the selected order $M = 2$ and their comparison with the input in the time and Fourier frequency domains are presented in Figure 7.22. Comparing visually Fourier spectra of both signals one can see a general agreement between the reconstruction and the input. The similarity is additionally confirmed by the residual energy metric RE = 0.307, which is close to assumed surrounding noise level set as an energy ratio 0.25. Such positive result is because the comparison in Fourier spectrum does not take into account instantaneous phase; thus it provides a simplified comparison. The simplification is confirmed when comparing signals directly in the time domain. As the residual energy RE = 0.487 indicates, the KurSL was able to explain only half of the energy. Interestingly, for time $t > 2$ s a sudden significant improvement in fitting occurs. This match is in contrast to the preceding period, for which a difference in amplitudes in noticeable despite preservation of a general oscillatory structure. Figure 7.23 presents time series (red) and instantaneous amplitude (blue) of identified KurSL oscillators in all rows, except for the first row where the comparison between reconstruction and the input is shown. From these graphs, it is difficult to observe significant changes in any individual oscillator. This subtlety suggests that the

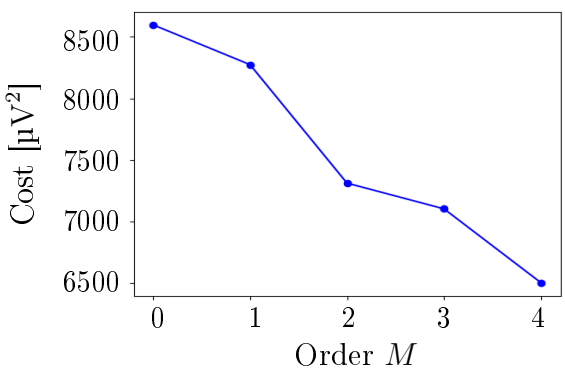


(a) Time series

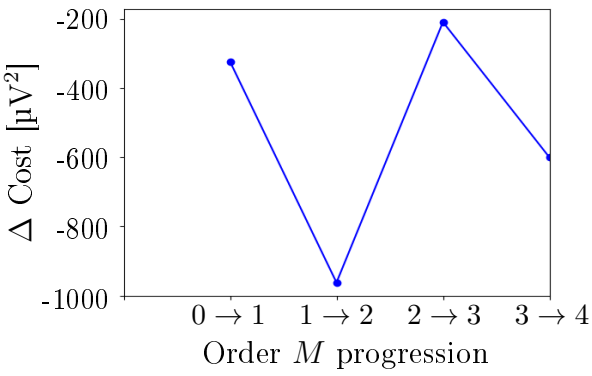


(b) Fourier spectrum

Figure 7.20: EEG signal used in the empirical dynamic analysis. Figure 7.20a contains time domain representation, whereas the Fourier spectrum is presented in Figure 7.20b.



(a) Absolute cost per order M.



(b) Piecewise cost difference.

Figure 7.21: Cost values (Fig. 7.21a) and their piecewise differences divided by the change in the number of parameters (Fig. 7.21b). These results were obtained for a range of order M values used in the dynamic EEG experiment.

improvement in reconstruction is a result of synchronisation between all oscillators. It also suggests that the following segments should observe improvements in fitting.

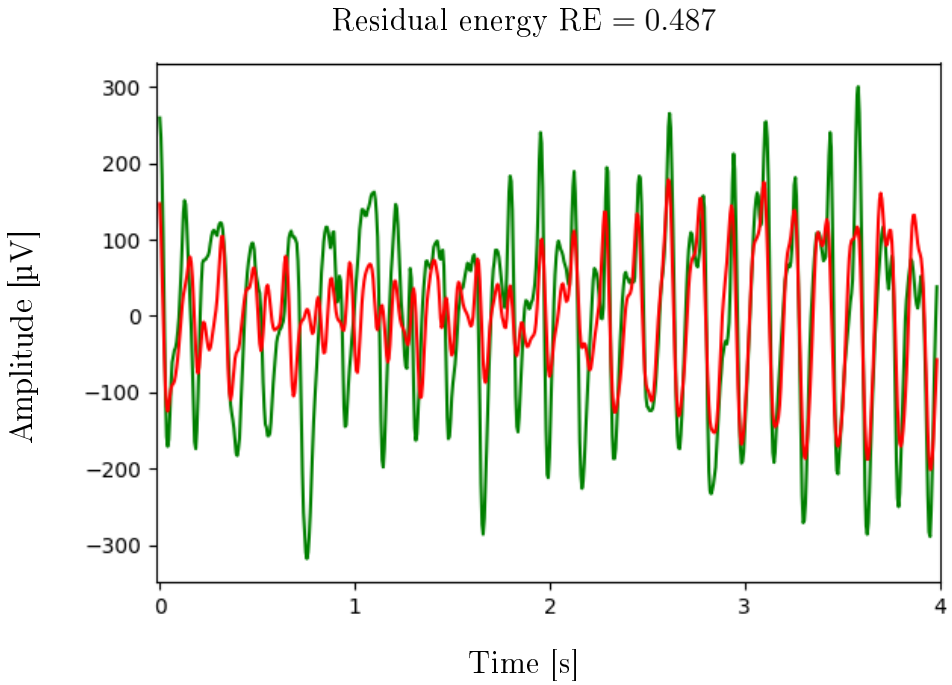
The KurSL analysis for all segments followed the same procedure as for the first one. For each window, the initial parameters were assigned as the best set from the previous segment. With such setting all obtained residual energy numerical values, for both time series RE_T and Fourier frequency RE_F, are presented in Table 7.11. A visual comparison between these values is presented in Figure 7.24 where RE_T and RE_F are indicated with solid blue and dashed red lines, respectively. Despite observing small variations within the domain, these values seem to be bound to a small numerical range. No substantial changes in the residual energy RE suggests that identified oscillators are present in the whole recording, although not necessarily with the same state at each window. These states and their dynamics are presented in Appendix grouped by intrinsic frequencies (Fig. B.19), phases (Fig. B.20), initial amplitudes (Fig. B.21) and coupling strengths of the first (Fig. B.22) and the second (Fig. B.23) orders. In all figures, the x-axis is the initial value of a time window domain, meaning that all figures have 12 points evenly spaced from 0 to 5.5 s inclusive. These results highlight that indeed the observed system is not composed of only coupled oscillators since obtained values changes for each segment. The magnitude of variance depends on the oscillator and inspected segment. Unfortunately, all obtained results are difficult to interpret on their own, without knowing what is exactly happening within the system. The KurSL

Table 7.11: Obtained residual energy (RE) values for all segments in the KurSL window analysis on the EEG signal. Measures RE_T and RE_F refer to RE obtained from comparison in the time and Fourier frequency domains, respectively, at window which initial position is indicated by t_{seg}.

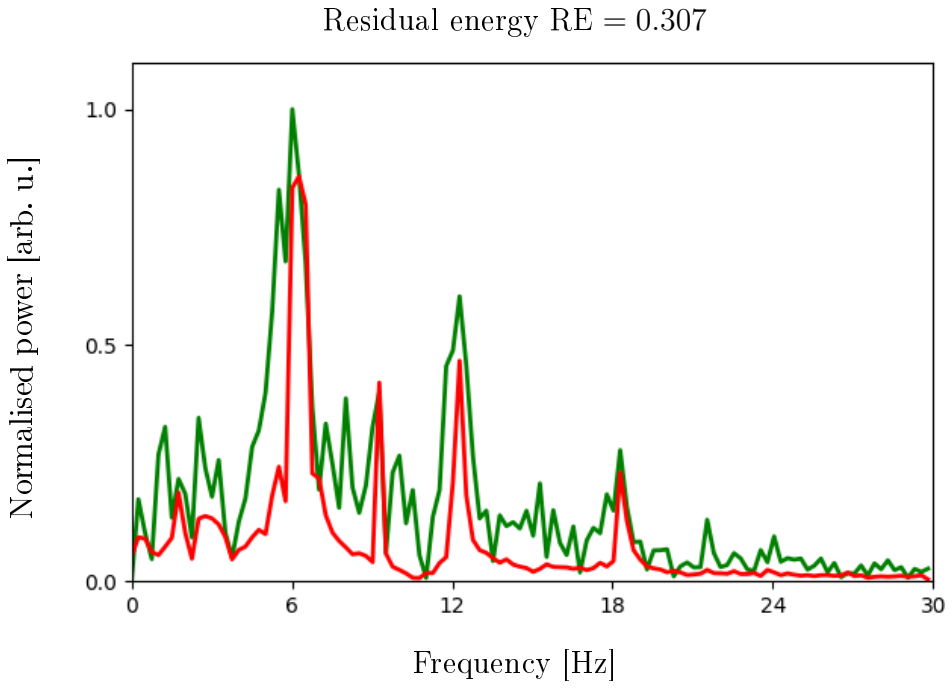
t_{seg} [s]	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5
RE_T	0.487	0.468	0.532	0.467	0.468	0.499	0.615	0.520	0.549	0.544	0.449	0.471
RE_F	0.307	0.300	0.332	0.260	0.253	0.250	0.366	0.338	0.276	0.336	0.239	0.215

method provides a quantitative description of the system, but the domain specialist should perform analysis of its results.

Results obtained through the KurSL were additionally compared with the Huang-Hilbert transform (HHT), short-time Fourier transform (STFT) and wavelet transform (WT). The comparison of all time-frequency (TF) representations are presented in Figure 7.25, where from the top left in the clockwise order panels refer to the KurSL and HHT spectral maps, then WT scaleogram and STFT spectrogram. Parameters for the STFT and WT representations were selected based on a visual comparison, similarly to the approach in the simulated signal (Section 7.1.1). Given three visible regions of increased activity in the Fourier spectrum 7.19b, in this case, it was also expected that the representation for both methods would provide a similar result at three frequency bands. Candidates for the STFT were obtained by generating all possible combinations of w_γ, where $\gamma = \{l, \tau, p\}$ and $l \in [0.5\mathrm{s}, 4.0\mathrm{s}]$ with the step 0.25 s, $\tau \in$ {Hann, Tukey ($\alpha = 0.25$), Tukey ($\alpha = 0.5$)} and $p \in \{50\%, 75\%, 90\%\}$. In the case of scaleograms, they were computed with the central frequency ω_0 in range 3–7 $\frac{\mathrm{rad}}{\mathrm{s}}$ (step 0.2), for which the Fourier frequencies are 4 Hz and 9 Hz, respectively, when the scale $a = 2^{-3}$ [31]. For the STFT the best set was selected to use 2 s long Hann window with a overlap step $p = 75\%$. Such window length allows for the frequency resolution of 0.5 Hz which is enough to highlight increased activity and gaps between them. In the case of the Wavelet transform the central frequency was determined to be $\omega_0 = 5$ $\frac{\mathrm{rad}}{\mathrm{s}}$ for which the Fourier frequency is $f(a = 2^{-3}) = 6.5$ Hz. Such frequency not only closely matches the most dominant peak in the Fourier spectrum (Fig. 7.19b) but also provides similar representation to the spectrogram. Few examples of other computed candidates are shown in Appendix B.24. The HHT was evaluated on the



(a) Comparison in time domain.



(b) Comparison in Fourier domain.

Figure 7.22: Comparison between the first segment of the EEG signal and its KurSL reconstruction (sec. 7.3.2). In both time (Fig. 7.22a) and Fourier (Fig. 7.22b) domains, EEG and reconstructed signals are indicated by green an red, respectively. For both comparisons computed residual energy (RE), i.e. energy of piecewise difference divided by energy of the input (7.1), is included in the title for a particular figure.

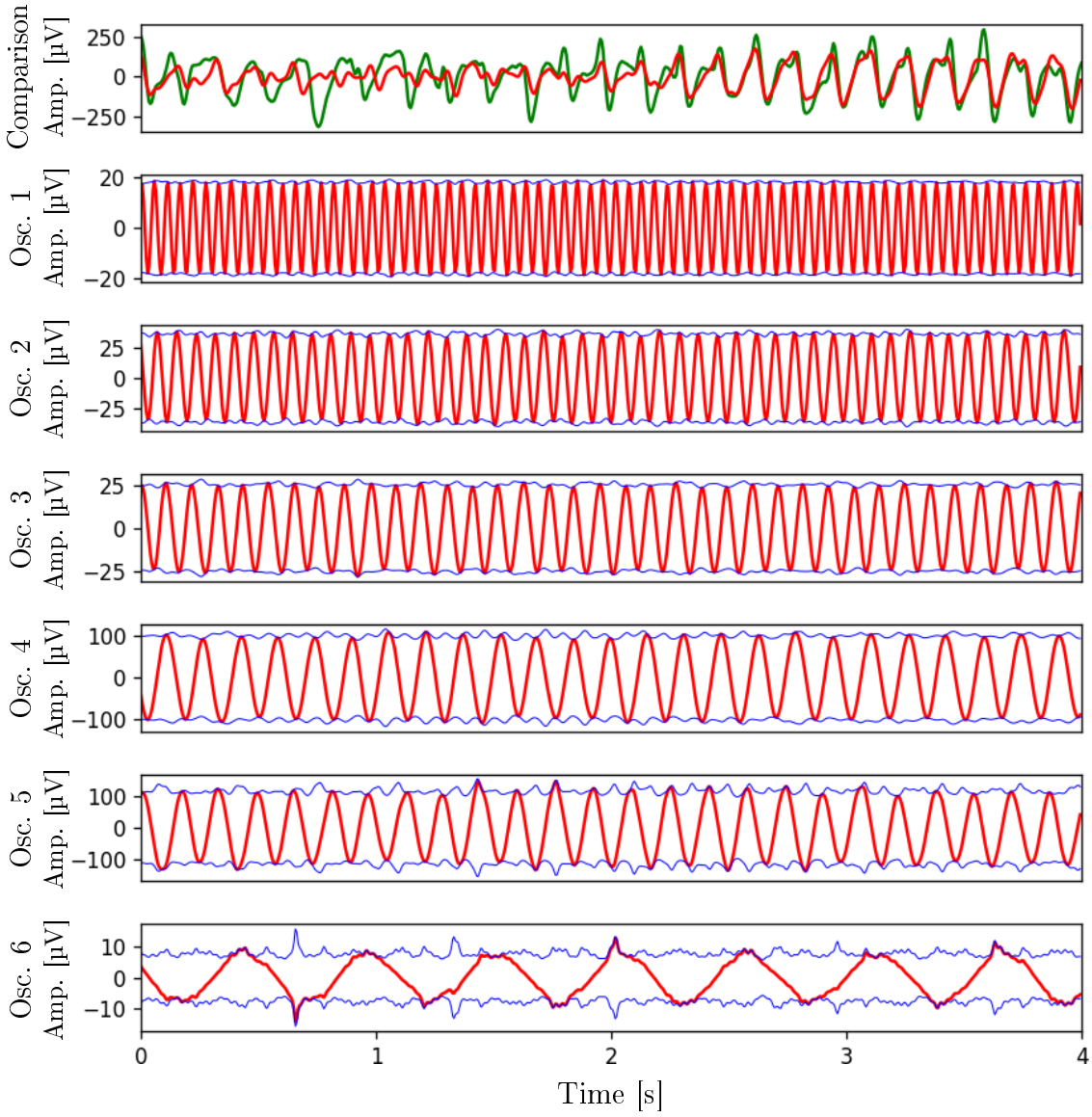


Figure 7.23: The figure represents the KurSL reconstruction of the first segment of the EEG signal in dynamic experiment (sec. 7.3.2). The first row contains the EEG time series in green and its KurSL reconstruction in red. Following rows present in frequency decreasing manner fitted oscillators (red) with their instantaneous amplitudes (blue).

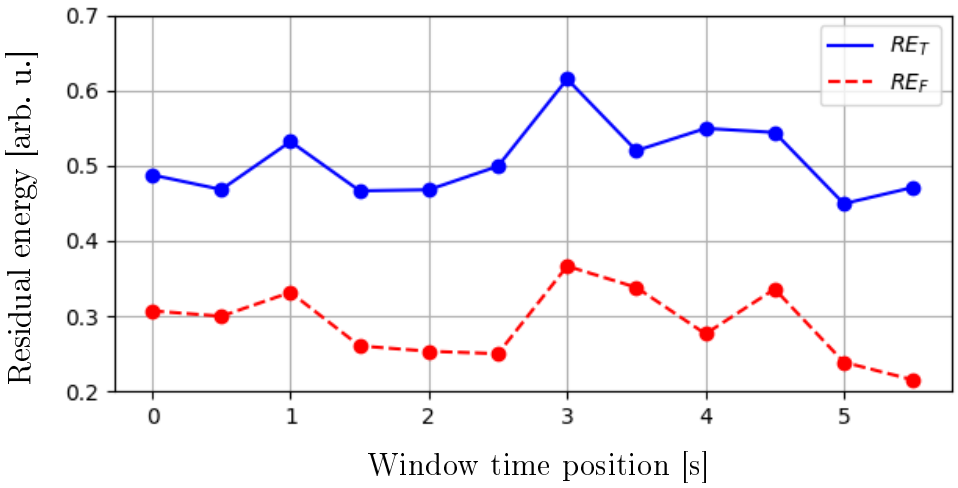


Figure 7.24: All residual energy RE values obtained for the dynamic KurSL approach used with the EEG signals. Solid blue and dashed red lines indicate residual energy for time series RE_T and Fourier frequency RE_F, respectively.

whole 9.5 s signal using cubic splines and the default stopping criteria [17]. For ease of comparison with other methods, obtained instantaneous frequencies were binned into a two-dimensional histogram with time and frequency resolutions of 0.1 s and 0.25 Hz, respectively. With such projection, it is possible to present all components in a readable manner despite high variance in some components. The same steps were applied to the KurSL results. Such projection was required due to overlapping analysis windows, which are causing discontinuities in the KurSL instantaneous frequencies. With such representation, it is possible to observe collective instantaneous frequency in the whole signal.

The only behaviour that is strongly highlighted by all representations is the activity centred around the dominant frequency $f = 6$ Hz where a sudden change in behaviour appears close to the time $t = 5$ s through a variation in the density. Other components are either highlighted differently or missed. For example, only the KurSL and HHT have consistently indicated an activity with slow frequency $f = 2.5$ Hz. Spectrogram and scaleogram's activities in this region are also visible but less pronounced. The difference in intensity of the slow component is because the spectrogram and scaleogram are scaled by the amplitude of the component, whereas constructed representations of the KurSL and HHT only indicate a component's existence.This means that from such presentation of KurSL and HHT methods it is difficult to inter-

pret which, if any, is the base frequency. Further Figure investigation shows that except for the HHTs, all representations have shown a component centred close to frequency $f = 12$ Hz and its sudden change in behaviour at a position close to $t = 5$ s, although it might be difficult to observe from the graph. Additionally, the KurSL and STFT show activity of the second subharmonic with frequency $f \approx 18$ Hz. The reason for this component to have limited visibility in the scaleogram is due to selected dyadic scale resolution. It is expected that with a different scale resolution this component would also be highlighted by WT. In contrast, it is theoretically uncertain whether any modifications to HHT configuration could improve its representation. An empirical attempt to improve the spectral map by modifying EMD configuration has shown little difference in the representation. Because of this, it is acknowledged that for given signal the HHT has provided the worst TF representation with the rest methods providing similar results.

In summary, despite not obtaining complete reconstructions in either time or frequency domains, the KurSL method provided with similar insights about the data as the other time-frequency representation methods. The advantage of using dynamical approach is in observing how the parameters change, which can highlight some properties of the signal. In this example, the analysis confirmed the appropriate choice of dominant frequencies and their behaviour changes. It is difficult to determine whether there are any patterns in such behaviour by merely observing dynamics of all parameters. These changes could be a result of both influences from the non-deterministic behaviour of the environment and incorrect assumption of the model's order. Selection of higher order would improve the reconstruction in both cases; however, it would also increase the chance of overfitting to the noise. A mitigation to the threat, as with other signal processing methods, could be a consultation with an expert in the domain, who would be able to interpret the result by associating them with certain physical phenomena.

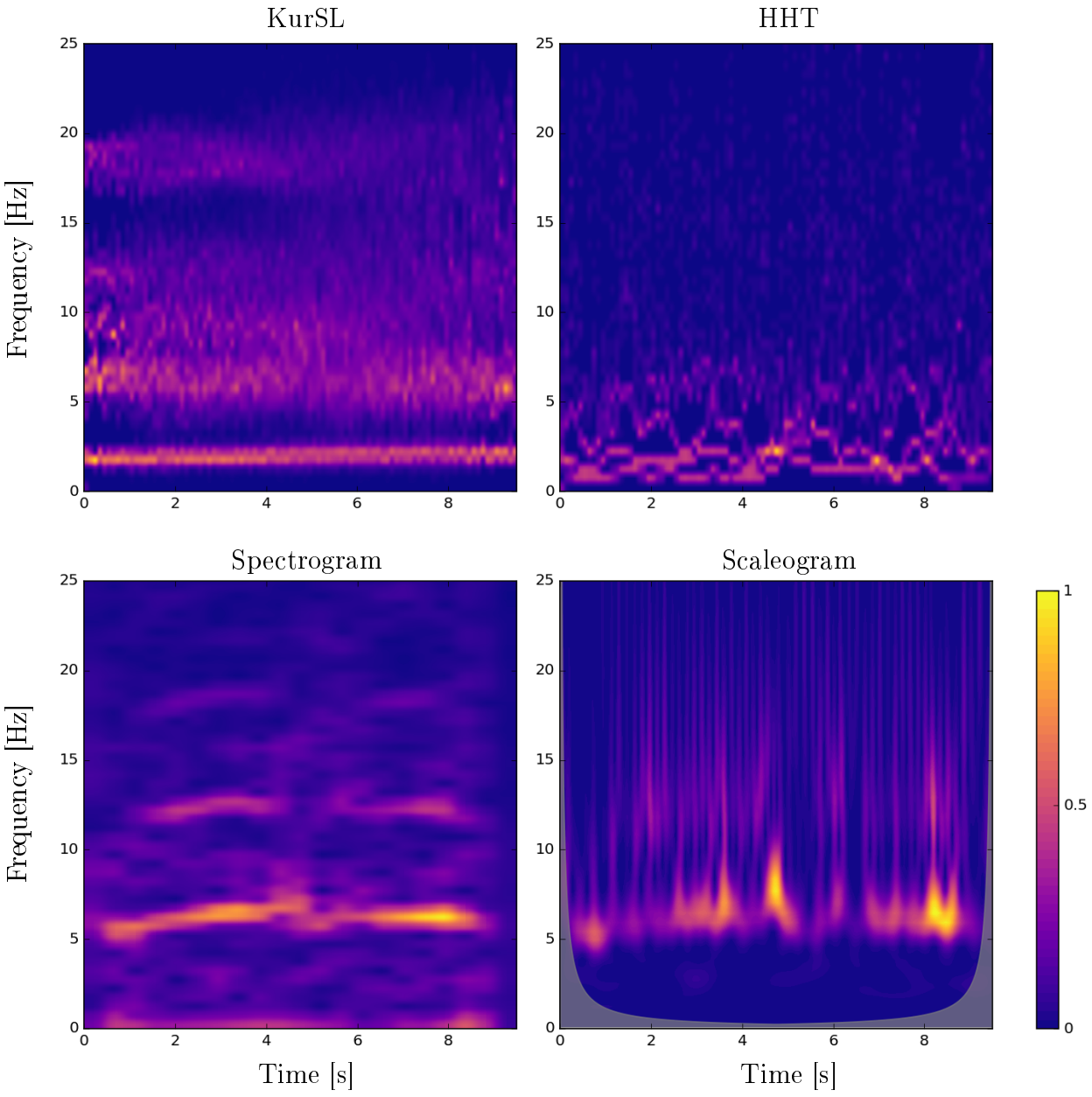


Figure 7.25: A comparison between different time-frequency representations for the EEG signal of 9.5 s length. In a clockwise order starting from the top right, panels present Huang-Hilbert transformation spectral map, scaleogram using Morlet wavelet with central frequency $\omega_0 = 5\ \frac{\text{rad}}{\text{s}}$, STFT with 2 s window and 0.5 s step and a spectral map obtained from all KurSL windows. All representations were scaled separately to the highest value of 1, and they use the same colour dynamic as indicated by the legend. The grey area in the scaleogram indicates regions outside the cone of influence.

7.4 Conclusion

In this Chapter, an application of the KurSL method has been discussed through experiments on synthetic and empirical signals. The purpose of these examples was to demonstrate the usage of the KurSL, its adaptation progress and the meaning of its results. The KurSL model was also used to generate the synthetic signals which allowed to demonstrate and discuss obtained results easily. Such a choice also makes sure that the condition to analyse a system consists of mutually interacting oscillators is met. Although other models could generate such signal, knowing underlying parameters allows to control and validate the method's performance.

The purpose of the first example (Section 7.1.1) was to introduce how the KurSL method works and how one can infer its performance based on the results. The focus of the second analysis (Section 7.1.2) was to show how the order of the KurSL method can affect its performance. In the last experiment on the synthetic data, the analysis was performed in a dynamical approach, i.e. by dividing the signal into segments which benefited from the estimates on the previous window. Based on these results, additional examples were presented on the brain signals. The selection of such origin is due to the common assumption that the underlying system is based on mixing oscillatory components [9, 11–13]. These signals were analysed in two experiments performed in stationary and dynamical approaches.

The first step in the KurSL method is to estimate the initial parameters for the optimisation process. These are obtained by using algorithm 1 (sec. 6.1.2) which iteratively removes the most prominent peak from the input signal's Fourier spectrum. Despite being successful in estimating initial intrinsic frequencies, phases and amplitudes, it is unable to determine the values of coupling factors. As it has been discussed in Section 7.2, the difficulty is due to non-unique effects that the coupling has on the spectrum. Depending on the coupling function, its strength and interacting oscillators it has been observed that the effects can include widening of the peak in the spectrum, the appearance of satellite peaks or shifting positions of these oscillators towards each other. For this reason, oscillators are initially assumed to be isolated, i.e. coupling terms $k = 0$, with assigned relatively wide probability distribution function (pdf) on having interactions. Such an approach provides excellent performance on signals com-

posed of few oscillators with relatively simple interaction. In case of more complex structures, however, as presented in Section 7.2, lack of this information may hinder the performance. The imperfections of the algorithm can be mitigated by reapplying the method to the signal using for initial parameters the best parameter set from the previous execution. Reassigning pdf with new expected locations one shifts and broadens the search space around the previous best result. One might also increase the number of walkers, iterations and the variance of initial pdf. These modifications would increase both the search space of the best parameters and the computational complexity, making the optimisation process run longer.

The KurSL method provides as a result pdf for all parameters. Few experiments attempted to provide meaning to this property by presenting its marginal distributions. In most cases, obtained distributions have a form of a single peak although their shapes can vary significantly. As it has been discussed in the first experiment (sec. 7.1) the global *maximum a posteriori* (MAP) values are not necessarily in the mode position of any marginal distribution. Moreover, it is expected that sometimes these distributions will have multimodal shapes. An example of an observed binomial distribution is presented in Figure 7.6. Multimodal distributions are expected especially when two or more oscillators with similar properties are nearby. Since the method does not make any distinction between oscillators, it is possible that they would occasionally swap positions.

In presented experiments, the results of the KurSL method has also been compared with other time-frequency representation methods, i.e. Huang-Hilbert Transform (HHT), short-time Fourier Transform (STFT) and Wavelet transform (WT). It is acknowledged that direct comparison of a scaleogram and spectrogram with KurSL's results is not possible since the former provides intensity levels within the predefined time and frequency ranges, whereas the latter provides a model to generate components without any resolution limits. Nevertheless, such comparison is beneficial since if there is any activity within the analysed time series, it should appear in any time-frequency representation regardless of its construction. Based on the presented results in this chapter, one can infer that indeed the KurSL indicates components that are also visible in spectrograms and scaleograms. However, as indicated, the advantage of

the KurSL is that it provides a model that can generate instantaneous features and directly explain their behaviour.

Out of all provided time-frequency representations results provided by the HHT stand out. Only when analysing a relatively simple signal, i.e. four coupled oscillators, a few of obtained IMFs have instantaneous frequencies that reflect the content. However, where other methods have suggested only a few components, the HHT showed significantly more components. The additional oscillators contain low frequency, not present in other representations. When acting on more complex signals such as noisy EEG time series presented in the previous section, the HHT produces components with a little match to those presented by other methods. Despite repeated decompositions with a modified setting to select the best decomposition in respect to frequency content some of these components have overlapping or negative frequencies. Obtained sets contain a coupled of IMFs with a high frequency and variance, and many more with low frequency. Such results are consistent with Flandrin and Rilling's observations [58] that the EMD treats highly complex signals as noise and acts on these like a dyadic filter. Since it is a greedy algorithm, all incorrect extractions propagate artefacts to the following IMFs. Due to being an empirical method, the HHT not only has limitations on the number of components related to the number of samples, but the Nyquist frequency limits its frequency range. In contrast, the KurSL does not have such limitations. Since the result is a model fit to the system's properties, if required, one can use it to interpolate between observations. In a case when the initial values are unknown, the KurSL will start by extracting the most impactful oscillators and will attempt to add components until a sufficient energy threshold is reached. This approach reduces chances to overfit to noise and extract the only required number of oscillators. Moreover, the HHT only extract components and their instantaneous frequency which then requires additional steps to identify what these mean.

In the case of empirical signals, it is impossible to assess which method provided the most accurate insights. Such a comparison requires expert knowledge and the understanding of system's functionality to validate the real components. Nevertheless, since three methods have obtained similar representations, it is expected that they are performing comparably well. An advantage that comes with the KurSL is its underly-

ing model which can fully explain oscillatory systems. Knowing the parameters with which the signal was generated allows simulating its dynamics without any resolution constriction. It also allows observing how the behaviour would change if the system were placed in a different environment. Such studies are crucial when learning about the system, and the KurSL tries to help in understanding the oscillatory systems.

Chapter 8

Conclusion

8.1 Summary

The study introduces a new data-driven method for describing systems that contain mutually interacting oscillators. The method uses the KurSL model, which is composed of two well-studied components, i.e. the Kuramoto coupling model and the Sturm-Liouville oscillation theory. As discussed in Chapter 5, both these components describe oscillators, although their approach differs. The Kuramoto coupling model describes how a set of oscillators can interact with each other through complex interactions in the phase domain. The Sturm-Liouville theory, however, characterises possible observations of generally defined oscillations independent from the surrounding. As a combination of these two components, the KurSL model can describe the dynamics of a system with objects expressing periodical behaviour and capable of interacting with one another. The model is essential due to the abundance of such examples, i.e. oscillators with amplitude- and phase-modulations. In fact, it is impossible to observe in nature any object isolated from the rest of the environment or where no external forces are applied, and thus leaving dynamics unaffected. The KurSL model explicitly acknowledges and emphasises such interactions. When the method is applied to a signal, it assumes that it was generated through a physical process. Knowing how the signal was created, i.e. knowing the processes within the system, allows for recreation of the data and in-depth analysis of the system's behaviour.

The advantage of the KurSL over commonly used methods is that the method

is not restricted to linear or stationary signals containing oscillations. Its components are estimated based on the provided data and inherently allow for the possibility of interactions within the system. In the case when there is no coupling between any oscillator, the KurSL will extract components in the form of harmonic oscillators, mimicking the behaviour of the Fourier series and converging as such.

The method's adaptation to data is performed in a two-step procedure. It firstly estimates parameters from data's properties and then, based on these estimates, optimises the cost function with Markov Chain Monte Carlo (MCMC). Such a choice of the heuristic optimisation method is due to the non-convex shape of the cost function. As it has been shown in Section 5.2, some properties of the KurSL model can be estimated through the analysis of the Fourier spectrum. It has been observed that dominant peaks in the spectrum are related to the main oscillators in the system. This allows estimating the number of oscillators and their intrinsic frequencies as well as initial phases and amplitudes. These parameters are automatically detected for a signal by applying an algorithm presented in Section 6.2 which are then passed into the MCMC. Finally, the search for a global optimum in a cost function defined as the difference between the reconstruction of the model and input data.

The whole procedure of applying the method to different types of data has been presented in Chapter 7. Experiments were conducted on both synthetic and empirical data. Their purpose was to highlight the KurSL performance and the implication of different configurations. After demonstrating the method (sec. 7.1.1) and the impact of order selection (sec. 7.1.2), additional analysis has been performed on the empirical EEG signal (sec. 7.3.1). It has been presented that general properties of a signal with a complex structure can be well approximated by assuming a small number of coupled oscillators. Moreover, the usage of the KurSL in a dynamic manner has been demonstrated in Sections 7.2 and 7.3.2. Dividing a signal into segments and utilising the best estimates from the previous window, one can improve results for the segment in focus. Such approach also allows for observing how the system can evolve in time and use these observations to create more precise model. If, for example, one knows the structure of a model at a particular window, shifting its position allows to extrapolate the rest of the signal. Analysis performed in a dynamical manner have been applied

both to synthetic and empirical EEG data with promising results.

In experiments conducted in Chapter 7, the KurSL method was often compared to other time-frequency representation methods, namely Huang-Hilbert transform (HHT), short-time Fourier transform (STFT) and wavelet transform (WT). For a simple input signal which is composed of few oscillators, all methods have provided comparable results, i.e. highlighting activity in similar regions. However, differences emerged with the increase of complexity. In case of empirical signals, the decomposition provided by HHT was heavily contaminated with considerable variance noise, making it difficult to compare visually. The obtained representation disagree with the rest of representations despite removing the most obscure components. For the most of signals, results obtained through the KurSL, STFT and WT methods closely agreed with each other. However, the difficulty of the interpretation has been observed when analysing empirical signals. Due to visual contrasting and finite resolution of the STFT and WT methods, the KurSL is considered to provide a more detailed representation. By knowing the exact construction of the system, the KurSL can explicitly evaluate instantaneous frequency. This allows for better understanding of the state and its reference to other components.

Although the focus in the thesis is mainly on time series decomposition and time-frequency representation, it needs to be emphasised that this can only be achieved by having an effective connectivity model at the method's core. As discussed in section 3.2 such model allows describing properties of connections between components within a system. Approaches discussed in that section similarly attempt to find the parameters for a model to explain the data. In that regard, the biggest similarities can be found between the KurSL and the dynamical Bayesian inference (DBI). Both methods use the Bayesian inference probabilistic approach which allows deducing the most likely parameters under which the data could have been observed. Such probabilistic approach allows these methods to fit a model with a general coupling form and observe its parameters dynamics by updating their probability distributions through window shift. There are also differences between these methods, and the main one is in the assumed interactions. The DBI provides a method to extract a general model of weak interactions, whereas the KurSL explicitly models phase and amplitude dynamics. In

situations when there are no amplitude modulations, they should provide similar results. Otherwise, the KurSL does not require any additional steps for extracting the phase and can be directly applied to amplitude dynamics. However, to the benefit of the DBI is its quick convergence, requiring only a few iterations of evaluating formulas. This is in contrast to computationally demanding KurSL which utilises MCMC and thus needs to perform a significant number of numerical integrations to search the parameter space. These extra operations significantly decrease the chance of overfitting a model and provide a better estimate of posterior distributions.

The KurSL method can provide insights into the system composed of coupled oscillations. It does that both quantitatively and qualitatively, making it especially suitable for research on dynamical systems. The definition on the KurSL explicitly models phase and amplitude modulations, making the method directly applicable to time series without any additional conversion to the phase space. This thesis contributes to science by providing an adaptive and robust tool for studying complex systems of interacting oscillators.

8.2 Open questions

The introduction of a new tool for analysing systems and generated data opens many paths for exciting research. One of the paths can be devoted to better understand the interactions between oscillators within a system. Although it is a broad research area, it is a promising area with the most significant impact on the KurSL. Being able to determine what are the implications of different coupling orders and their values could not only allow for a better estimation of the initial coupling terms $k^m_{i,j}$, but also provide with more precise estimation on the number of oscillators $\mathcal{N}$. Providing these estimates would significantly improve method's robustness and computational performance. Moreover, obtaining narrower estimates for parameters could additionally limit the possibility of overlapping search spaces and swapping of positions by the walkers. An appropriate estimation also reduces the complexity of the method to the necessary and the most suitable. Such restriction should mitigate the problem of overfitting the model.

Additional interesting opportunity lays in creating multivariate method. In this thesis, the KurSL method has only discussed where the input is in the form of a single variable time series observation. Such scope is equivalent to recording a system with a single recorder. Increasing the number of recorders and placing them in unique locations would mean that due to different distances observations would be perceived with different delays. However, they all monitor the same system; therefore, they all should indicate precisely the same parameters. Incorporating multivariate elements into the KurSL method should benefit from utilising an additional source of information under a little cost of including the delay parameters. Additionally, signal gathered by a different recorder would presumably propagate through a different path and thus be influenced differently by the surrounding. Each recording would be contaminated with different noise making it easier to be removed from the data.

An attractive challenge, which could significantly influence the rest of the research, lays in improving the computational complexity and performance. The model is defined by a set of coupled differential equations which complexity scales as $O(N^2M)$, where N and M are the number of oscillators and harmonics, respectively. Current implementation of the KurSL method takes about 20 minutes to perform 10 iterations fitting eight oscillators with three harmonics or six oscillators with five harmonics. One of the main goals is to decrease the computation time, which would allow for research with a fast feedback. Improvement could be obtained by changing the software implementation platform. For studies reported in this thesis, the whole programme was written in Python programming language, which is a high-level programming language. It is expected that implementing the method in a lower-level language, perhaps Go or C++, and utilising parallel computational architecture, such as graphical processing units, would decrease computational time by at least an order of magnitude. Nevertheless, the most improvement should be obtained by removing the need for numerical simulations and providing analytical solutions. The current software implementation is open source and freely available to download from author's homepage [125]. The page will contain all improvements and additional implementations of the method and model.

Bibliography

[1] J. Diamond, *Guns, Germs, and Steel: The Fates of Human Societies*. W. W. Norton, 1997.

[2] N. MacGregor, *A History of the World in 100 Objects*. Penguin Books Ltd., 2012.

[3] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, 3rd ed. CRC Press, 2014, p. 696.

[4] D. Lawson and G. Marion, *Course notes: An introduction to mathematical modelling*, 2008.

[5] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman lectures on physics*, 6th ed. Basic books, 1963, vol. 1.

[6] D. Halliday, R. Resnick, and J. Walker, *Fundamentals of Physics - Extended*, 6th ed. Wiley, 2000.

[7] K. J. H. Law, A. M. Stuart, and K. C. Zygalakis, *Data assimilation: A mathematical introduction*, 2015.

[8] A. T. Winfree, "Biological rhythms and the behavior of populations of coupled oscillators," *Journal of Theoretical Biology*, vol. 16, no. 1, pp. 15–42, 1967.

[9] J. A. Acebrón, L. L. Bonilla, C. J. Pérez Vicente, F. Ritort, and R. Spigler, "The kuramoto model: A simple paradigm for synchronization phenomena," *Reviews of Modern Physics*, vol. 77, no. 1, pp. 137–185, 2005.

[10] S. H. Strogatz, "From kuramoto to crawford: Exploring the onset of synchronization in populations of coupled oscillators," *Physica D: Nonlinear Phenomena*, vol. 143, no. 1, pp. 1–20, 2000.

[11] T. K. Das, P. M. Abeyasinghe, J. S. Crone, A. Sosnowski, S. Laureys, A. M. Owen, and A. Soddu, "Highlighting the structure-function relationship of the brain with the Ising model and graph theory," *BioMed Research International*, vol. 2014, pp. 1–14, 2014.

[12] R. R. Llinas, "The intrinsic electrophysiological properties of mammalian neurons: Insights into central nervous system function," *Science*, vol. 242, no. 4886, pp. 1654–1664, 1988.

[13] M. Abbas, M. Nain, and H. Awang, "Monotonicity preserving interpolation using rational spline," in *Proceeding of the International MultiConference of Engineers and Computer Scientists*, vol. I, 2011.

[14] L. Cohen, *Time-frequency Analysis*. Prentice Hall PTR, 1995, p. 299.

[15] M. X. Cohen and R. Gulbinaite, "Five methodological challenges in cognitive electrophysiology," *NeuroImage*, 2013.

[16] A. Galka, *Topics in nonlinear time series analysis: With implications for EEG analysis*. World Scientific, 2000, p. 360.

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

[18] A. Duggento, T. Stankovski, P. V. E. McClintock, and A. Stefanovska, "Dynamical Bayesian inference of time-evolving interactions: From a pair of coupled oscillators to networks of oscillators," *Physical Review E*, vol. 86, p. 061 126, 6 2012.

[19] C. Junsheng, Y. Dejie, and Y. Yu, "A fault diagnosis approach for roller bearings based on EMD method and AR model," *Mechanical Systems and Signal Processing*, vol. 20, no. 2, pp. 350–362, 2006.

[20] D. Looney, L. Li, T. M. Rutkowski, D. P. Mandic, and A. Cichocki, "Ocular artifacts removal from EEG using EMD," in *Advances in Cognitive Neurodynamics ICCN 2007*, Springer Netherlands, 2008, pp. 831–835.

[21] Y. Xu and H. Zhang, "Recent mathematical developments on empirical mode decomposition," *Advances in Adaptive Data Analysis*, vol. 1, no. 4, pp. 681–702, 2009.

[22] M. Jörn, C. Sielużycki, M. A. Matysiak, J. Zygierewicz, H. Scheich, P. J. Durka, and R. König, "Single-trial reconstruction of auditory evoked magnetic fields by means of template matching pursuit," *Journal of Neuroscience Methods*, vol. 199, no. 1, pp. 119–128, 2011.

[23] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.

[24] E. Wigner, "On the quantum correction for thermodynamic equilibrium," *Physical Review*, vol. 40, pp. 749–759, 5 1932.

[25] P. Durka, "Matching pursuit," *Scholarpedia*, vol. 2, no. 11, p. 2288, 2007.

[26] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 2, no. 1-3, pp. 37–52, 1987.

[27] M. Hožič and A. Stefanovska, "Karhunen-Loéve decomposition of peripheral blood flow signal," *Physica A: Statistical Mechanics and its Applications*, vol. 280, no. 3, pp. 587–601, 2000.

[28] C. Obled and J. D. Creutin, "Some developments in the use of empirical orthogonal functions for mapping meteorological fields," *Journal of Climate and Applied Meteorology*, 1986.

[29] N. Golyandina and A. Zhigljavsky, *Singular spectrum analysis for time series*. Springer Science & Business Media, 2013.

[30] D. Percival and A. Walden, *Wavelet methods for time series analysis*, ser. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2000.

[31] C. Torrence and G. P. Compo, "A practical guide to wavelet analysis," *Bulletin of the American Meteorological Society*, vol. 79, no. 1, pp. 61–78, 1998.

[32] A. Haar, "Zur Theorie der orthogonalen Funktionensysteme," *Mathematische Annalen*, vol. 69, no. 3, pp. 331–371, 1910.

[33] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Communications on Pure and Applied Mathematics*, vol. 41, no. 7, pp. 909–996, 1988.

[34] D. T. Lee and A. Yamamoto, "Wavelet analysis: Theory and applications," *Hewlett Packard journal*, vol. 45, p. 44, 1994.

[35] L. D. Landau and E. M. Lifshitz, *Quantum mechanics: Non-relativistic Theory*. Pergamon Press, 1965, vol. 3, p. 616.

[36] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Transactions on Information Theory*, vol. 36, no. 5, pp. 961–1005, 1990.

[37] M. Farge, "Wavelet transforms and their applications to turbulence," *Annual Review of Fluid Mechanics*, vol. 24, no. 1, pp. 395–458, 1992.

[38] S. Blanco, R. Q. Quiroga, O. A. Rosso, and S. Kochen, "Time-frequency analysis of electroencephalogram series," *Physical Review E*, vol. 51, pp. 2624–2631, 3 1995.

[39] A. Bruns, "Fourier-, Hilbert- and wavelet-based signal analysis: are they really different approaches?" *Journal of Neuroscience Methods*, vol. 143, no. 2, p. 237, 2005.

[40] R. Kuś, P. T. Różański, and P. J. Durka, "Multivariate matching pursuit in optimal Gabor dictionaries: Theory and software with interface for EEG/MEG via Svarog," *Biomedical Engineering OnLine*, vol. 12, no. 1, p. 94, 2013.

[41] U. Malinowska, C. Chatelle, M.-A. Bruno, Q. Noirhomme, S. Laureys, and P. J. Durka, "Electroencephalographic profiles for differentiation of disorders of consciousness," *Biomedical Engineering Online*, vol. 12, no. 1, p. 109, 2013.

[42] Q. Chen, N. E. Huang, S. Riemenschneider, and Y. Xu, "A B-spline approach for empirical mode decompositions," *Advances in Computational Mathematics*, vol. 24, no. 1-4, pp. 171–195, 2006.

[43] Y. Huang, F. G. Schmitt, Z. Lu, and Y. Liu, "Analysis of daily river flow fluctuations using empirical mode decomposition and arbitrary order Hilbert spectral analysis," *Journal of Hydrology*, vol. 373, no. 1-2, pp. 103–111, 2009.

[44] K. T. Coughlin and K. K. Tung, "11-Year solar cycle in the stratosphere extracted by the empirical mode decomposition method," *Advances in Space Research*, vol. 34, no. 2, pp. 323–329, 2004.

[45] D. G. Duffy, "The application of Hilbert-Huang transforms to meteorological datasets," *Journal of Atmosheric and Oceanic Technology*, vol. 21, pp. 599–611, 2004.

[46] Y. Huang, F. G. Schmitt, Z. Lu, and Y. Liu, "Application of arbitrary-order Hilbert spectral analysis to passive scalar turbulence," *Journal of Physics: Conference Series*, vol. 318, no. 4, p. 042003, 2011.

[47] M.-C. Wu and C.-K. Hu, "Empirical mode decomposition and synchrogram approach to cardiorespiratory synchronization," *Physical Review E*, vol. 73, no. 5, pp. 1–11, 2006.

[48] C. M. Sweeney-Reed and S. J. Nasuto, "A novel approach to the detection of synchronisation in EEG based on empirical mode decomposition," *Journal of Computational Neuroscience*, vol. 23, no. 1, pp. 79–111, 2007.

[49] C. M. Sweeney-Reed and S. J. Nasuto, "Detection of neural correlates of self-paced motor activity using empirical mode decomposition phase locking analysis," *Journal of Neuroscience Methods*, vol. 184, no. 1, pp. 54–70, 2009.

[50] R. B. Pachori, "Discrimination between ictal and seizure-free EEG signals using empirical mode decomposition," *Research Letters in Signal Processing*, vol. 2008, p. 5, 2008.

[51] W. Liu, Y. Yan, and R. Wang, "Application of Hilbert-Huang transform and SVM to coal gangue interface detection," *Journal of Computers*, vol. 6, no. 6, pp. 1262–1269, 2011.

[52] T. M. Rutkowski, A. Cichocki, D. P. Mandic, and A. L. Ralescu, "Multichannel spectral pattern separation: An EEG processing application," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, 2009, pp. 373–376.

[53] T. M. Rutkowski, D. P. Mandic, A. Cichocki, and A. W. Przybyszewski, "EMD approach to multichannel EEG data the amplitude and phase components clustering analysis," *Journal of Circuits, Systems and Computers*, vol. 19, no. 1, pp. 215–229, 2010.

[54] M. K. I. Molla, T. Tanaka, T. M. Rutkowski, and A. Cichocki, "Separation of EOG artifacts from EEG signals using bivariate EMD," in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, 2010, pp. 562–565.

[55] A. Erturk, M. K. Gullu, and S. Erturk, "Hyperspectral image classification using empirical mode decomposition with spectral gradient enhancement," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 5, pp. 2787–2798, 2013.

[56] S. Lin, P. Tung, and N. E. Huang, "Data analysis using a combination of independent component analysis and empirical mode decomposition," *Physical Review E*, vol. 79, no. 6, p. 066 705, 2009.

[57] N. E. Huang and Z. Wu, "A review on Hilbert-Huang transform: Method and its applications to geophysical studies," *Reviews of Geophysics*, vol. 46, no. 2, pp. 1–23, 2008.

[58] P. Flandrin and P. Gonçalvés, "Empirical mode decompositions as data-driven wavelet-like expansions," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 2, no. 4, pp. 477–496, 2004.

[59] T. Qian, Q. Chen, and L. Li, "Analytic unit quadrature signals with nonlinear phase," *Physica D: Nonlinear Phenomena*, vol. 203, no. 1-2, pp. 80–87, 2005.

[60] T. Qian, "Mono-components for decomposition of signals," *Mathematical Methods in the Applied Sciences*, vol. 29, no. 10, pp. 1187–1198, 2006.

[61] S. Qin and Y. Zhong, "A new envelope algorithm of Hilbert-Huang transform," *Mechanical Systems and Signal Processing*, vol. 20, no. 8, pp. 1941–1952, 2006.

[62] Z. Xu, B. Huang, and K. Li, "An alternative envelope approach for empirical mode decomposition," *Digital Signal Processing*, vol. 20, no. 1, pp. 77–84, 2010.

[63] R. T. Rato, M. D. Ortigueira, and A. G. Batista, "On the HHT, its problems, and some solutions," *Mechanical Systems and Signal Processing*, vol. 22, no. 6, pp. 1374–1394, 2008.

[64] F. R. Holzinger and M. Benedikt, "Online instantaneous frequency estimation utilizing empirical mode decomposition and Hermite splines," in *2014 22nd European Signal Processing Conference*, 2014, pp. 446–450.

[65] G. Rilling, P. Flandrin, and P. Goncalves, "On empirical mode decomposition and its algorithms," in *IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing*, vol. 3, 2003, p. 5.

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

[67] R. Deering and J. J. F. Kaiser, "The Use of a Masking Signal to Improve Empirical Mode Decomposition," in *Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, vol. 4, IEEE, 2005, pp. 485–488.

[68] L. Lin, Y. Wang, and H. Zhou, "Iterative filtering as an alternative algorithm for empirical mode decomposition," *Advances in Adaptive Data Analysis*, vol. 1, no. 4, pp. 543–560, 2009.

[69] R. Faltermeier, A. Zeiler, I. R. Keck, A. M. Tome, A. Brawski, and E. W. Lang, "Sliding empirical mode decomposition," in *The 2010 International Joint Conference on Neural Networks IJCNN*, 2010, pp. 1–8.

[70] D. Chen, D. Li, M. Xiong, H. Bao, and X. Li, "GPGPU-aided ensemble empirical-mode decomposition for EEG analysis during anesthesia," *IEEE Transactions on Information Technology in Miomedicine*, vol. 14, no. 6, pp. 1417–1427, 2010.

[71] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 1, pp. 1–41, 2009.

[72] J. Nunes, Y Bouaoune, E. Deléchelle, O. Niang, and P. Bunel, "Image analysis by bidimensional empirical mode decomposition," *Image and Vision Computing*, vol. 21, no. 12, pp. 1019–1026, 2003.

[73] C. Damerval, S. Meignen, and V. Perrier, "A fast algorithm for bidimensional EMD," *IEEE Signal Processing Letters*, vol. 12, no. 10, pp. 701–704, 2005.

[74] N. Rehman and D. P. Mandic, "Multivariate empirical mode decomposition," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 466, no. 2117, pp. 1291–1302, 2009.

[75] Z. Wu, N. E. Huang, and X. Chen, "The multi-dimensional ensemble empirical mode decomposition method," *Advances in Adaptive Data Analysis*, vol. 1, no. 03, pp. 339–372, 2009.

[76] D. P. Mandic, N. ur Rehman, Z. Wu, and N. E. Huang, "Empirical mode decomposition-based time-frequency analysis of multivariate signals: The power of adaptive data nalysis," *IEEE Signal Processing Magazine*, vol. 30, no. 6, pp. 74–86, 2013.

[77] E. Deléchelle and J. Lemoine, "Empirical mode decomposition: An analytical approach for sifting process," *IEEE Signal Processing Letters*, vol. 12, no. 11, pp. 764–767, 2005.

[78] E. H. S. Diop, R. Alexandre, and A. O. Boudraa, "Analysis of intrinsic mode functions: A PDE approach," *IEEE Signal Processing Letters*, vol. 17, no. 4, pp. 398–401, 2010.

[79] J. C. F. Sturm, "Mémoire sur les équations différentielles linéaires du second ordre," *Journal de Mathématiques pures et Appliquées*, vol. 1, pp. 106–186, 1836.

[80] G. Freiling and V. A. Yurko, *Inverse Sturm-Liouville problems and their applications*. NOVA Science Publishers New York, 2001.

[81] E. Schrödinger, "An undulatory theory of the mechanics of atoms and molecules," *Physical Review*, vol. 28, no. 6, pp. 1049–1070, 1926.

[82] G. Teschl, *Ordinary Differential Equations and Dynamical Systems*. American Mathematical Society, 2012, vol. 140, p. 356.

[83] B. Simon, "Sturm Oscillation and Comparison Theorems," *ArXiv e-prints*, pp. 1–17, 2003.

[84] R. C. Sharpley and V. Vatchev, "Analysis of the Intrinsic Mode Functions," *Constructive Approximation*, vol. 24, no. 1, pp. 17–47, 2006.

[85] V. Vatchev and R. Sharpley, "Decomposition of functions into pairs of intrinsic mode functions," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 464, no. 2097, pp. 2265–2280, 2008.

[86] Y. Kuramoto, *Chemical oscillations, waves, and turbulence*, ser. Springer Series in Synergetics. Berlin, Heidelberg: Springer Berlin Heidelberg, 1984, vol. 19.

[87] D. C. Michaels, E. P. Matyas, and J Jalife, "Mechanisms of sinoatrial pacemaker synchronization: A new hypothesis," *Circulation Research*, vol. 61, pp. 704–714, 1987.

[88] S. H. Strogatz, *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering*. Addison-Wesley Publishing Company, 1994, p. 499.

[89] H. Daido, "Why circadian rhythms are circadian: competitive population dynamics of biological oscillators," *Physical Review Letters*, vol. 87, no. 2, p. 048 101, 2001.

[90] E. D. Herzog, "Neurons and networks in daily rhythms," *Nature Reviews Neuroscience*, vol. 8, no. 10, pp. 790–802, 2007.

[91] A. Stefanovska, "Coupled Oscillators: Complex but Not Complicated Cardiovascular and Brain Interactions," *IEEE Engineering in Medicine and Biology Magazine*, vol. 26, no. 6, pp. 25–29, 2007.

[92] H.-J. Park and K. Friston, "Structural and Functional Brain Networks: From Connections to Cognition," *Science*, vol. 342, no. 6158, 2013.

[93] A. Pikovsky and M. Rosenblum, "Dynamics of globally coupled oscillators: Progress and perspectives," *Chaos*, vol. 25, no. 9, 097616, p. 097 616, 2015.

[94] T. Stankovski, T. Pereira, P. V. E. McClintock, and A. Stefanovska, "Coupling functions: Universal insights into dynamical interaction mechanisms," *Reviews of Modern Physics*, vol. 89, p. 045 001, 4 2017.

[95] A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization: A Universal Concept in Nonlinear Sciences*, ser. Cambridge Nonlinear Science Series. Cambridge University Press, 2001.

[96] L. M. Pecora and T. L. Carroll, "Master stability functions for synchronized coupled systems," *Physical Review Letters*, vol. 80, pp. 2109–2112, 10 1998.

[97] B. Kralemann, A. Pikovsky, and M. Rosenblum, "Reconstructing phase dynamics of oscillator networks," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 21, no. 2, p. 025 104, 2011.

[98] A. T. Winfree, "Integrated view of resetting a circadian clock," *Journal of Theoretical Biology*, vol. 28, no. 3, pp. 327–374, 1970.

[99] H. Daido, "Generic scaling at the onset of macroscopic mutual entrainment in limit-cycle oscillators with uniform all-to-all coupling," *Physical Review Letters*, vol. 73, no. 5, pp. 760–763, 1994.

[100] ——, "Onset of cooperative entrainment in limit-cycle oscillators with uniform all-to-all interactions: Bifurcation of the order function," *Physica D*, vol. 91, pp. 24–66, 1996.

[101] Y. Kuramoto, "Self-entrainment of a population of coupled non-linear oscillators," English, in *International Symposium on Mathematical Problems in Theoretical Physics*, ser. Lecture Notes in Physics, H. Araki, Ed., vol. 39, Springer Berlin Heidelberg, 1975, pp. 420–422.

[102] S. Petkoski and A. Stefanovska, "Kuramoto model with time-varying parameters," *Physical Review E*, vol. 86, p. 046 212, 4 2012.

[103] Y. F. Suprunenko, P. T. Clemson, and A. Stefanovska, "Chronotaxic Systems: A New Class of Self-Sustained Nonautonomous Oscillators," *Physical Review Letters*, vol. 111, p. 024 101, 2 2013.

[104] G. Lancaster, Y. Suprunenko, K. Jenkins, and A. Stefanovska, "Modelling chronotaxicity of cellular energy metabolism to facilitate the identification of altered metabolic states," *Scientific Reports*, vol. 6, p. 29 584, 2016.

[105] K. B. Lauritsen, R. Cuerno, and H. A. Makse, "Noisy Kuramoto-Sivashinsky equation for an erosion model," *Physical Review E*, vol. 54, no. 4, p. 4, 1996.

[106] R. Sassi and N. J. Balmforth, "A shocking display of synchrony," *Physica D: Nonlinear Phenomena*, vol. 143, no. 1-4, pp. 21–55, 2000.

[107] V. Vlasov, E. E. N. Macau, and A. Pikovsky, "Synchronization of oscillators in a Kuramoto-type model with generic coupling," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 24, no. 2, p. 023 120, 2014.

[108] A. Stefanovska and M. Bračič, "Physics of the human cardiovascular system," *Contemporary Physics*, vol. 40, no. 1, pp. 31–55, 1999.

[109] M. Paluš and A. Stefanovska, "Direction of coupling from phases of interacting oscillators: An information-theoretic approach," *Physical Review E*, vol. 67, p. 055 201, 5 2003.

[110] A. Barrett and L. Barnett, "Granger causality is designed to measure effect, not mechanism," *Frontiers in Neuroinformatics*, vol. 7, p. 6, 2013.

[111] M. G. Rosenblum and A. S. Pikovsky, "Detecting direction of coupling in interacting oscillators," *Physical Review E*, vol. 64, p. 045 202, 4 2001.

[112] B. Kralemann, L. Cimponeriu, M. Rosenblum, A. Pikovsky, and R. Mrowka, "Phase dynamics of coupled oscillators reconstructed from data," *Physical Review E*, vol. 77, no. 6, p. 066 205, 2008.

[113] I. T. Tokuda, S. Jain, I. Z. Kiss, and J. L. Hudson, "Inferring Phase Equations from Multivariate Time Series," *Physical Review Letters*, vol. 99, p. 064 101, 6 2007.

[114] E. Baake, M. Baake, H. G. Bock, and K. M. Briggs, "Fitting ordinary differential equations to chaotic data," *Physical Review A*, vol. 45, pp. 5524–5529, 8 1992.

[115] T. Stankovski, A. Duggento, P. V. E. McClintock, and A. Stefanovska, "Inference of time-evolving coupled dynamical systems in the presence of noise," *Physical Review Letters*, vol. 109, p. 024 101, 2 2012.

[116] ——, "A tutorial on time-evolving dynamical Bayesian inference," *The European Physical Journal Special Topics*, vol. 223, no. 13, pp. 2685–2703, 2014.

[117] Z. Levnajić and A. Pikovsky, "Network Reconstruction from Random Phase Resetting," *Physical Review Letters*, vol. 107, p. 034 101, 3 2011.

[118] C. Schäfer, M. G. Rosenblum, H.-H. Abel, and J. Kurths, "Synchronization in the human cardiorespiratory system," *Physical Review E*, vol. 60, no. 1, p. 857, 1999.

[119] I. Z. Kiss, C. G. Rusin, H. Kori, and J. L. Hudson, "Engineering complex dynamical structures: Sequential patterns and desynchronization," *Science*, vol. 316, no. 5833, pp. 1886–1889, 2007.

[120] D. Laszuk, O. Cadenas, and S. J. Nasuto, "EMD performance comparison: Single vs double floating points," *International Journal of Signal Processing Systems*, vol. 4, no. 4, pp. 349–353, 2016.

[121] D. Laszuk, O. Cadenas, and S. Nasuto, "Objective empirical mode decomposition metric," *International Journal of Advances in Telecommunications, Electrotechnics, Signals and Systems*, vol. 5, no. 1, pp. 8–14, 2016.

[122] D. Laszuk, O. Cadenas, and S. J. Nasuto, "On the phase coupling of two components mixing in empirical eode decomposition," *Advances in Adaptive Data Analysis*, vol. 8, no. 2, pp. 1 650 004–1–1650004–11, 2016.

[123] NVIDIA, *Nvidia cuda toolkit release notes*, 2015 (accessed April 1, 2015).

[124] G. Rilling and P. Flandrin, "On the influence of sampling on the empirical mode decomposition," in *2006 IEEE International Conference on Acoustics Speed and Signal Processing Proceedings*, vol. 3, IEEE, 2006, pp. 444–447.

[125] D. Laszuk, *Python implemention of empirical mode decomposition algorithm*, `http://www.laszukdawid.com/codes`, 2014.

[126] E. Jones, T. Oliphant, P. Peterson, *et al.*, *SciPy: Open source scientific tools for Python*, `http://www.scipy.org/`, 2001.

[127] D. Goldberg, "What every computer scientist should know about floating point arithmetic," *ACM Computing Surveys*, vol. 23, no. 1, pp. 5–48, 1991.

[128] S. Riemenschneider, B. Liu, Y. Xu, and N. E. Huang, "B-spline based empirical mode decomposition," in *Hilbert-Huang transform and its applications*, N. E. Huang and S. S. P. Shen, Eds., WORLD SCIENTIFIC, 2005, ch. 1, p. 324.

[129] N. Stevenson, M. Mesbah, and B. Boashash, "A sampling limit for the empirical mode decomposition," in *Proceedings of the Eighth International Symposium on Signal Processing and Its Applications, 2005.*, vol. 2, IEEE, 2005, pp. 647–650.

[130] E. Bedrosian, "A product theorem for Hilbert transforms," *Proceedings of the IEEE*, vol. 51, no. 5, pp. 868–869, 1963.

[131] B. Tang, S. Dong, and T. Song, "Method for eliminating mode mixing of empirical mode decomposition based on the revised blind source separation," *Signal Processing*, vol. 92, no. 1, pp. 248–258, 2012.

[132] G. Rilling and P. Flandrin, "One or two frequencies? The empirical mode decomposition answers," *IEEE Transactions on Signal Processing*, vol. 56, no. 1, pp. 85–95, 2008.

[133] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Neural Networks, 1995. Proceedings., IEEE International Conference on*, vol. 4, 1995, pp. 1942–1948.

[134] H. Akima, "A new method of interpolation and smooth curve fitting based on Loca Procedures," *Journal of the ACM*, vol. 17, no. 4, pp. 589–602, 1970.

[135] N. E. Huang, Z. Wu, S. R. Long, K. C. Arnold, X. Chen, and K. Blank, "On Instantaneous Frequency," *Advances in Adaptive Data Analysis*, vol. 1, no. 2, pp. 177–229, 2009.

[136] K. Drakakis, N. Tsakalozos, and S. Rickard, "Some remarks on the well-posedness of the EMD algorithm," in *IET Irish Signals and Systems Conference (ISSC 2009)*, IET, 2009, p. 49.

[137] T. Qian, Y.-B. Wang, and P. Dang, "Adaptive decomposition into mono-components," *Advances in Adaptive Data Analysis*, vol. 1, no. 04, pp. 703–709, 2009.

[138] W. Zhu, H. Zhao, D. Xiang, and X. Chen, "A flattest constrained envelope approach for empirical mode decomposition," *PLoS ONE*, vol. 8, no. 4, e61739, 2013.

[139] P. J. Loughlin and B. Tacer, "On the amplitude- and frequency-modulation decomposition of signals," *The Journal of the Acoustical Society of America*, vol. 100, no. 3, p. 1594, 1996.

[140] B. Picinbono, "On instantaneous amplitude and phase of signals," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 552–560, 1997.

[141] G. B. Ermentrout and N. Kopell, "Oscillator death in systems of coupled neural oscillators," *SIAM Journal on Applied Mathematics*, vol. 50, no. 1, pp. 125–146, 1990.

[142] J. Goodman and J. Weare, "Ensemble samplers with affine invariance," *Communications in Applied Mathematics and Computational Science*, vol. 5, no. 1, pp. 65–80, 2010.

[143] M. Richey, "The evolution of Markov chain Monte Carlo methods," *The American Mathematical Monthly*, vol. 117, no. 5, pp. 383–413, 2010.

[144] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *The Journal of Chemical Physics*, vol. 21, no. 6, pp. 1087–1092, 1953.

[145] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970.

[146] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, no. 6, pp. 721–741, 1984.

[147] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, "EMCEE: The MCMC Hammer," *Publications of the Astronomical Society of the Pacific*, vol. 125, pp. 306–312, 2013.

[148] E. W. Weisstein, *Parseval's theorem*, From MathWorld — A Wolfram Web Resource. `http://mathworld.wolfram.com/ParsevalsTheorem.html`, 2016.

[149] C. F. Cadieu and K. Koepsell, "Phase coupling estimation from multivariate phase statistics," *Neural Computation*, vol. 22, no. 12, pp. 3107–3126, 2010.

[150] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists*, 6th ed. Elsevier Inc., 2005, p. 1182.

[151] A. Zygmund, *Trigonometric series*. Cambridge University Press, 1988, p. 747.

[152] S. van der Walt, S. C. Colbert, and G. Varoquaux, "The numpy array: A structure for efficient numerical computation," *Computing in Science Engineering*, vol. 13, no. 2, pp. 22–30, 2011.

[153] G. Lee, R. Gommers, F. Wasilewski, K. Wohlfahrt, A. O'Leary, H. Nahrstaedt, and Contributors, *PyWavelets - Wavelet Transforms in Python*, `https://github.com/PyWavelets/pywt/`, 2006.

[154] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Physical Review E*, vol. 64, p. 061 907, 6 2001.

[155] C. J. Stam and B. W. Van Dijk, "Synchronization likelihood: An unbiased measure of generalized synchronization in multivariate data sets," *Physica D: Nonlinear Phenomena*, vol. 163, no. 3-4, pp. 236–251, 2002.

[156] E. Pereda, R. Q. Quiroga, and J. Bhattacharya, "Nonlinear multivariate analysis of neurophysiological signals," *Progress in Neurobiology*, vol. 77, no. 1-2, pp. 1–37, 2005.

[157] P. Jiruska, M. de Curtis, J. G. R. Jefferys, C. A. Schevon, S. J. Schiff, and K. Schindler, "Synchronization and desynchronization in epilepsy: Controversies and hypotheses," *The Journal of Physiology*, vol. 591, no. Pt 4, pp. 787–797, 2013.

[158] K. J. Engel and R. Nagel, *One-parameter semigroups for linear evolution equations*, ser. Graduate Texts in Mathematics. New York: Springer-Verlag, 2000, vol. 194, pp. 278–280.

[159] Y. Xu and D. Yan, "The Bedrosian identity for the Hilbert transform of product functions," *Proceeding of the American Mathematical Society*, vol. 134, no. 9, pp. 2719–2728, 2006.

Appendices

Appendix A

Additional information

A.1 Systems

A system is a set of distinct states whose temporal change is be determined [158]. The complexity of the states depends on the structure of the system and the environment in which it is embedded. Some systems can be considered as a set of subsystems, i.e. units, which on their own would be considered systems. These subsystems, however, could be influenced by environment and thus would produce different states if isolated. Depending on system's properties they can be classified for example by their *linearity* of *stationarity*.

Linearity

A system is linear if the output of the sum of independent input components is the same as it would be when passed inputs independently. That is, given two independent inputs $x(t)$ and $y(t)$. If system H behaves in such way, that $H(x(t)) = X(t)$ and $H(y(t)) = Y(t)$, then the operator H is called linear if

$$H\left(\alpha x(t) + \beta y(t)\right) = \alpha X(t) + \beta Y(t), \tag{A.1}$$

where α and β are scalar weights.

An example of a linear system can be a simple model of harmonic oscillator

defined by a formula

$$H\left(x(t)\right) = m\frac{d^2x}{dt^2} + kx, \tag{A.2}$$

where m and k are some constants. If the input is sum of two independent states $x(t) = \alpha y(t) + \beta z(t)$ then

$$\begin{aligned} H\left(x(t)\right) &= m\tfrac{d^2x}{dt^2} + kx \\ &= m\tfrac{d^2(\alpha y(t)+\beta z(t))}{dt^2} + k(\alpha y(t) + \beta z(t)) \\ &= \alpha m\tfrac{d^2y}{dt^2} + \alpha ky + \beta m\tfrac{d^2z}{dt^2} + \beta kz \\ &= \alpha H\left(y(t)\right) + \beta H\left(z(t)\right). \end{aligned} \tag{A.3}$$

Stationarity

Stationarity property is determined by a dynamic of states, i.e. a *process*. A stochastic process in which joint probability distribution stays the same regardless of the observation time is called *stationary*. Likewise, if the system is not stationary, it is classified as *non-stationary*. However, such definition is rigorous and in empirical signals rarely meet this condition. Commonly in signal processing community, an approximation of this property is used. The definition for stationarity in weak-sense, referring to only two first moments of a signal, is presented in Definition 2.

Definition 2. *If* $\{X_t, t \in \mathbb{Z}\}$ *(a stochastic process) satisfies properties*

1. $E\{X_{t+\tau}\} = \mu_X, \qquad \forall\tau \in \mathbb{Z};$
2. $Cov\{X_t, X_{t+\tau}\} = s_{X,\tau}, \qquad \forall t, \tau \in \mathbb{Z};$

where E *and* Cov *refer to expected value and covariance, then* $\{X_t\}$ *is said to be stationary (second order).*

A.2 Hilbert transform

The term *analytic* signal refers to the fact that its values are complex numbers. These provide both information about the amplitude and the phase at each time point. One

of the most popular methods for transforming a measured signal, which is real, into an analytic on is a *Hilbert transform*. The linear operator acting on a time series $s(t)$ is defined through the signal's convolution with Cauchy's kernel ($K_C(t) = 1/t$)

$$\mathcal{H}(s)(t) = \frac{1}{\pi} P.V. \int_{-\infty}^{\infty} \frac{s(\tau)}{t-\tau} d\tau, \tag{A.4}$$

where P.V. is Cauchy's principal value. Operator $\mathcal{H}$ can alternatively be defined with limits [151] as

$$\mathcal{H}(s)(t) = -\frac{1}{\pi} \lim_{\varepsilon \to 0} \int_{\varepsilon}^{\infty} \frac{s(t+\tau) - s(t-\tau)}{\tau} d\tau. \tag{A.5}$$

If $s \in L^p(\mathbb{R})$ for $1 < p < \infty$, then the limit of (A.5) is well defined for almost every t.

In the physical sense, this operator removes all negative frequencies and doubles the positive [14]. Hilbert transform can be simplified with the use of Fourier transform as

$$\mathcal{F}(\mathcal{H}(u))(\omega) = (-i\mathrm{sgn}(\omega))\mathcal{F}(u)(\omega), \tag{A.6}$$

where sgn is signum function, and $\mathcal{F}$ is Fourier Transform. The analytic signal, constructed on real-value f time series, is thus defined as

$$Af = f(t) + i\mathcal{H}(f)(t). \tag{A.7}$$

A.3 Bedrosian identity

Computing the Hilbert transform of the product of two functions can be simplified with the help of an identity known as *Bedrosian identity*. The original formulation of the theorem by Bedrosian (also proven in [130]) is as follows:

Theorem 5. *Let $f(x)$ and $g(x)$ denote general complex functions in $L^2(\mathbb{R})$ of the real variable x. If*

1. *the Fourier transform $F(u)$ of $f(x)$ vanishes for $|u| > a$ and the Fourier transform $G(u)$ of $g(x)$ vanishes for $|u| < a$, where a is an arbitrary positive constant, or*
2. *$f(x)$ and $g(x)$ are analytic, i.e. their real and imaginary parts are Hilbert pairs,*

then the Hilbert transform of the product of $f(x)$ and $g(x)$ is given by

$$\mathcal{H}\left[f(x)g(x)\right] = f(x)\mathcal{H}\left[g(x)\right]. \tag{A.8}$$

Over time more generic versions have been proposed. One of them is introduced in terms of the Fourier transform [159] as shown in Theorem 6.

Theorem 6. *Let $f \in W^{1,2}(\mathbb{R})$ and $g \in L^2(\mathbb{R})$. Then the Hilbert transform of function fg satisfies the Bedrosian identity (A.8) if and only if*

$$\int_{-1}^{0}\int_{\mathbb{R}} \frac{\omega}{t^2} e^{2i\pi x\omega(t+1)/t} \hat{f}\left(\frac{\omega}{t}\right) \hat{g}(\omega) d\omega dt = 0, \tag{A.9}$$

where $\hat{f}$ refers to Fourier transform of function f.

All this interest is motivated by the attempts to characterise analytic function constructed from real-valued one. Any real part of an analytic function can be considered as a product of amplitude and phase functions, that is $x(t) = a(t)\cos(\phi(t))$. The Hilbert Transform of the function is $y(t) = \mathcal{H}(x(t)) = -ia(t)\sin(\phi(t))$. This allows to fully restore analytic function $c(t) = x(t) + iy(t) = a(t)\exp(i\phi(t))$ as well as its components $a(t) = |c(t)|$ and $\phi(t) = \arctan\left(\frac{y(t)}{x(t)}\right)$.

Appendix B

Figures

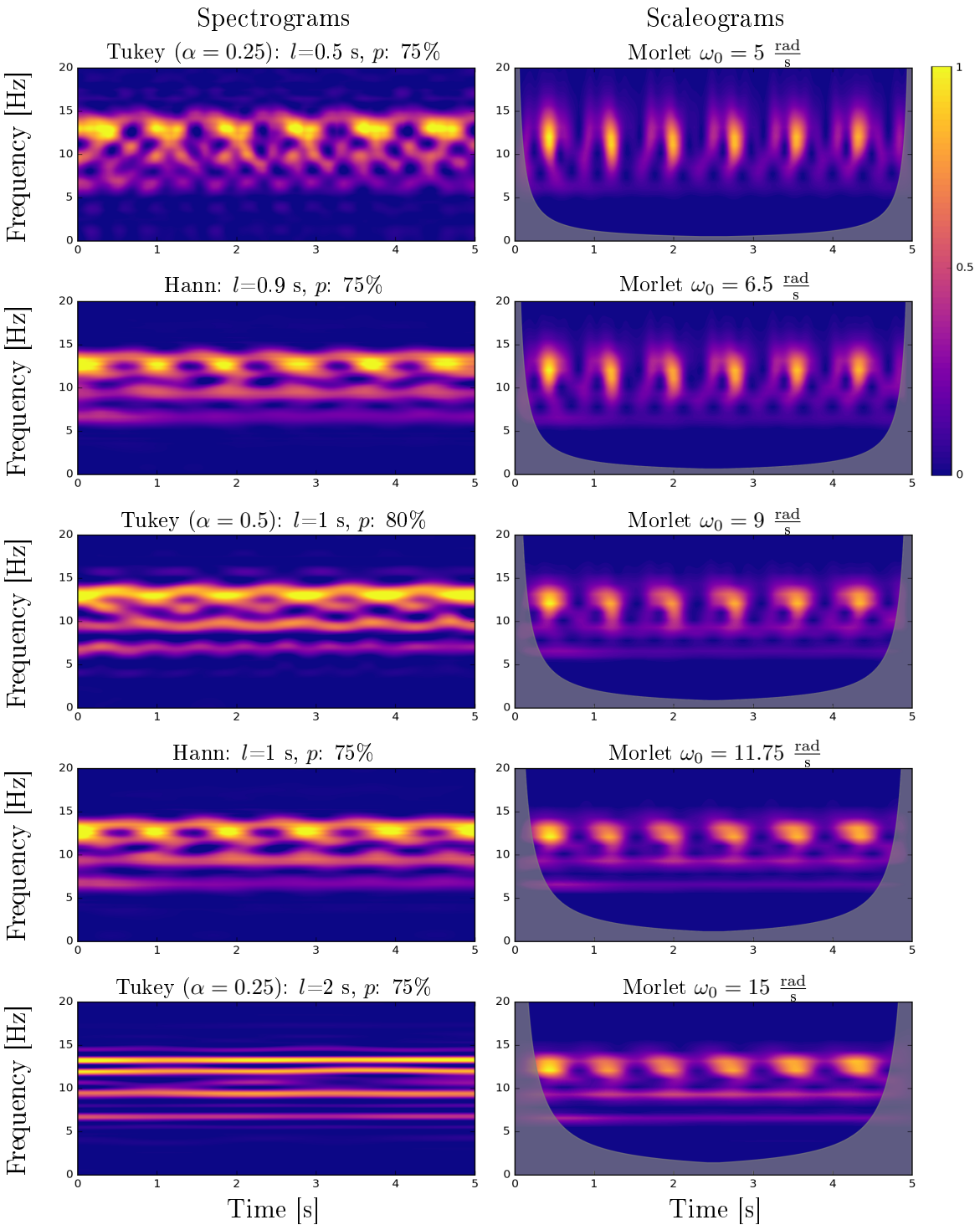


Figure B.1: Different time-frequency representations of the generated signal using the KurSL model with parameters from Table 7.1 which was used in the simple stationary experiment with 4 oscillators (Section 7.1.1). Column on the left represents spectrograms for which respective labels denote window's type, its length l and the percentage overlap p. The right column shows scaleograms in order from the top computed using the Morlet wavelets with central frequencies $\omega_0 = \{5, 6.5, 9, 11.75, 15\}$ $\frac{\text{rad}}{\text{s}}$, respectively. All values were scaled such that the smallest and largest for any plot are zero and one with the colour dynamic presented in the bottom right corner.

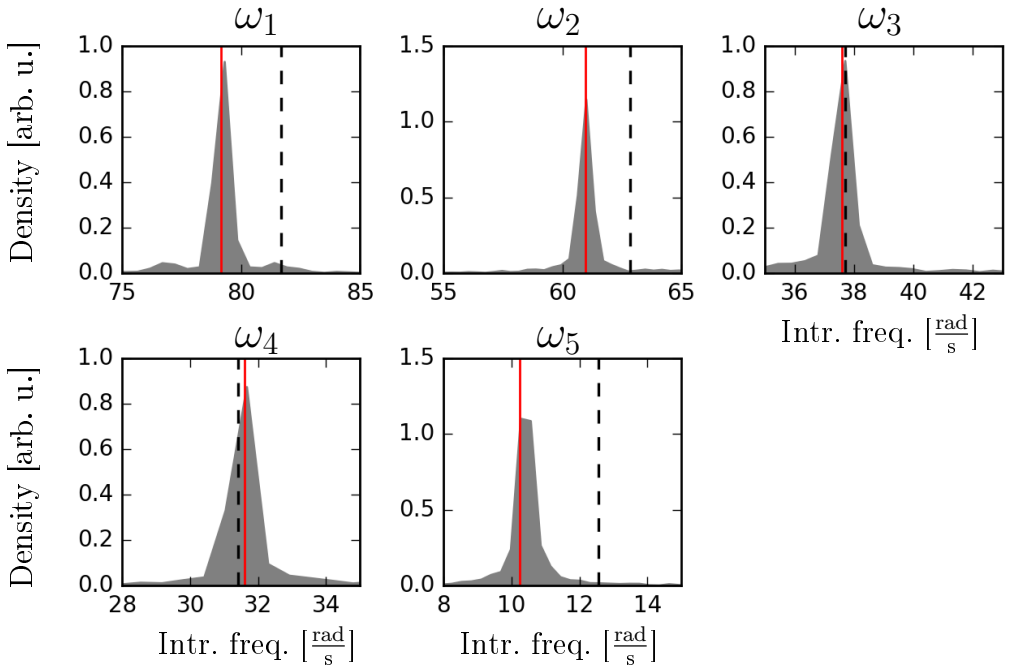


Figure B.2: The probability density for intrinsic frequency values obtained for the 1$^{\text{st}}$ order KurSL in the comparison experiment. All values are in rad/s units. Red vertical lines mark the global optima found for respective parameters. Values used to generate the input signal are indicated by a black dashed line.

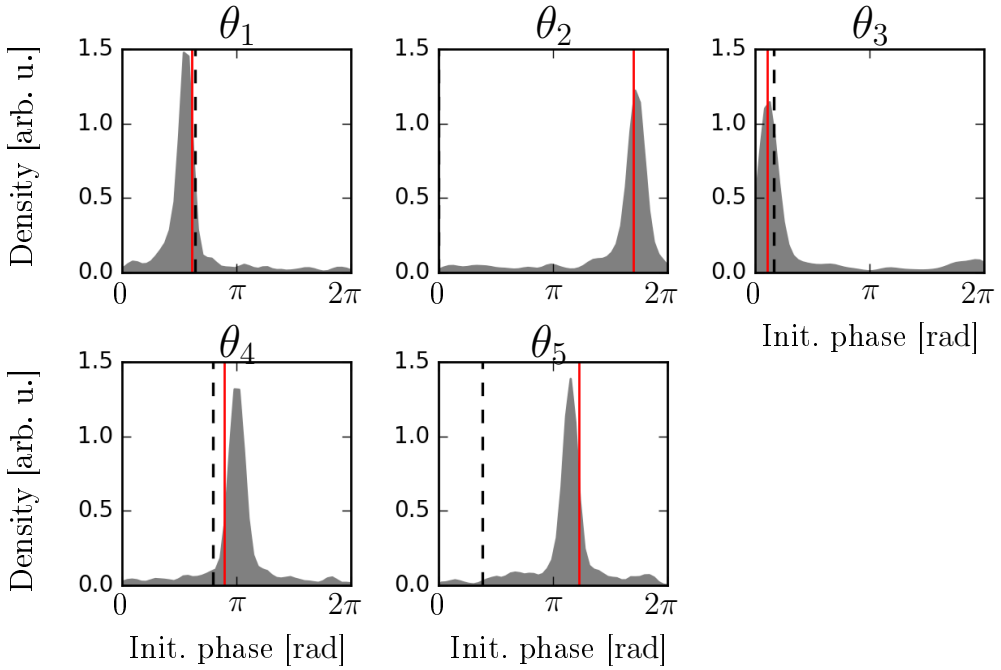


Figure B.3: The probability density for phase values obtained for the 1$^{\text{st}}$ order KurSL in the comparison experiment. All values are in radians. Red vertical lines marks the global optima found for respective parameters. Values used to generate the input signal are indicated by a black dashed line.

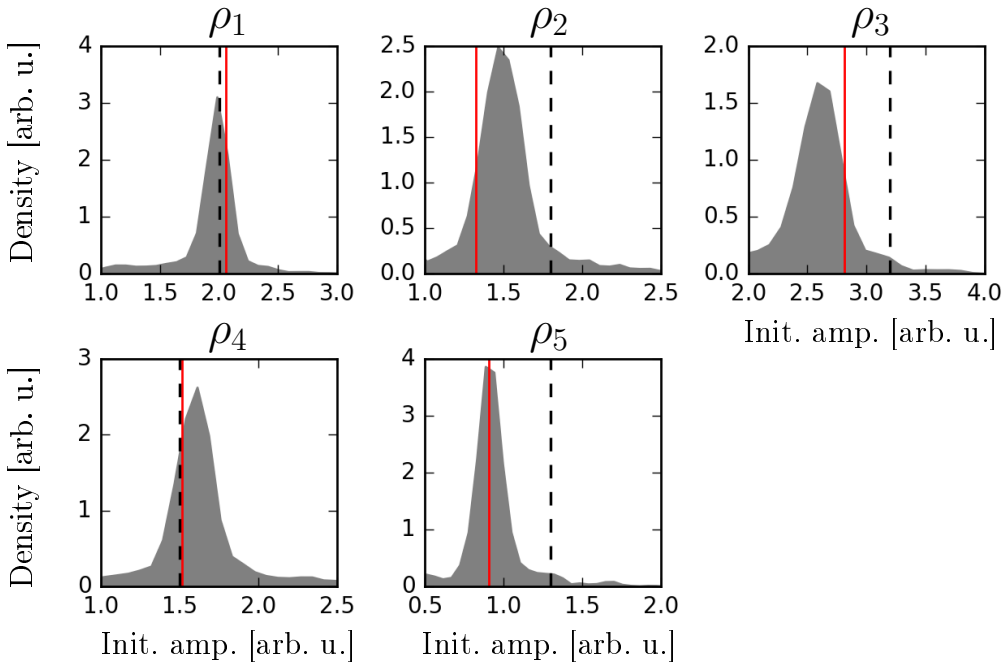


Figure B.4: The probability density for amplitude values obtained for the 1st order KurSL in the comparison experiment. All amplitude values are in arbitrary units. Red vertical lines marks the global optima found for respective parameters. Values used to generate the input signal are indicated by a black dashed line.

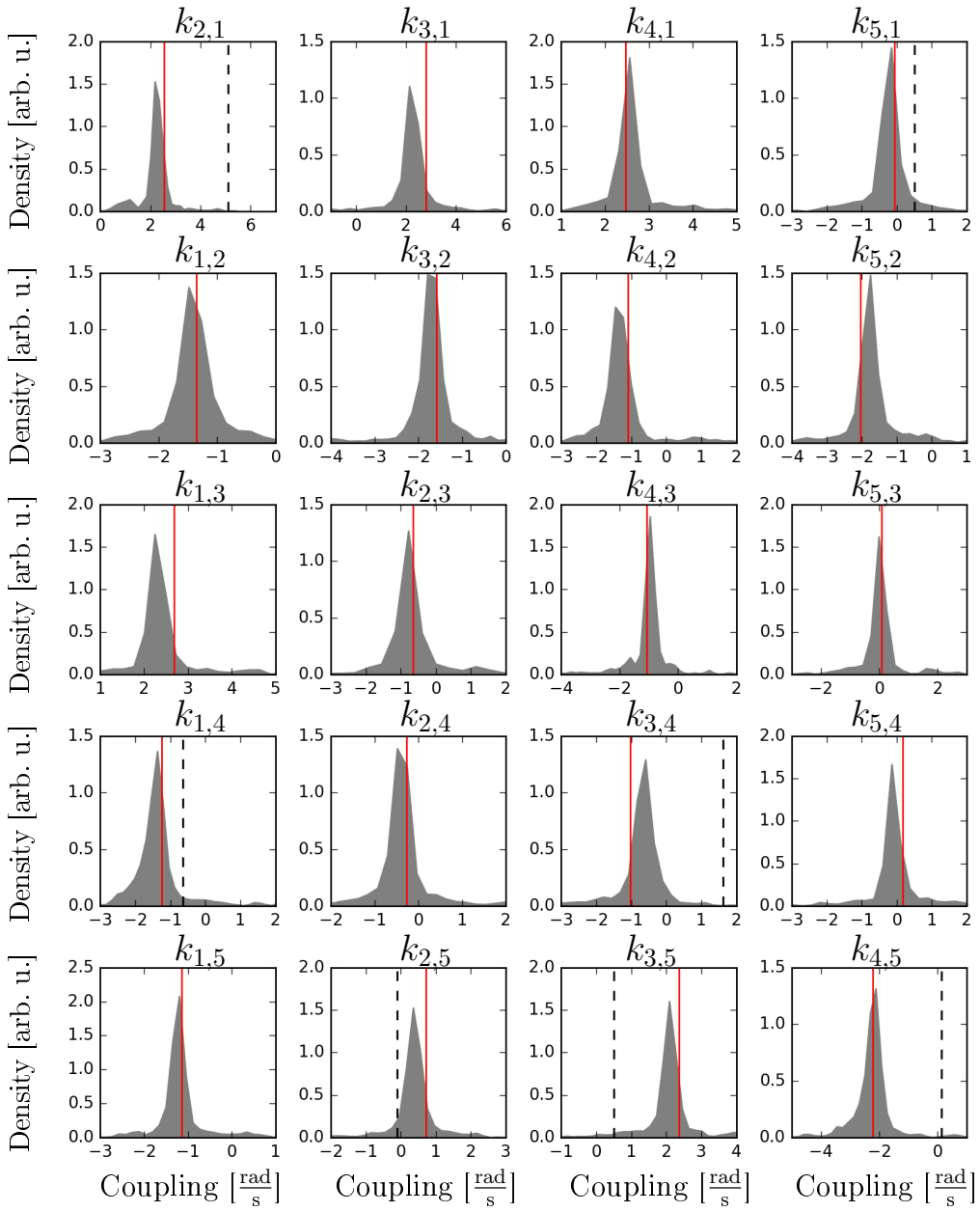


Figure B.5: The probability density for coupling strength values obtained for the 1$^{\text{st}}$ order KurSL in the comparison experiment. All coupling values are in rad/s units. Red vertical lines marks the global optima found for respective parameters. Values used to generate the input signal are indicated by a black dashed line.

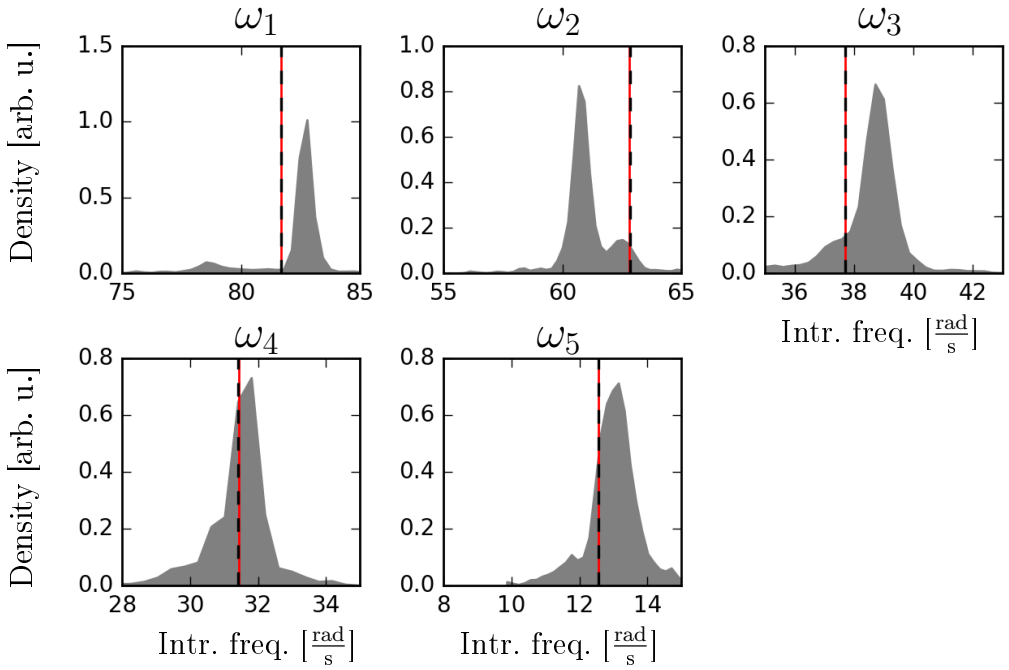


Figure B.6: The probability density for intrinsic frequencies values obtained for the 3$^{\text{rd}}$ order KurSL in the comparison experiment. All values are in rad/s units. Red vertical lines marks the global optima found for respective parameters. Values used to generate the input signal are indicated by a black dashed line.

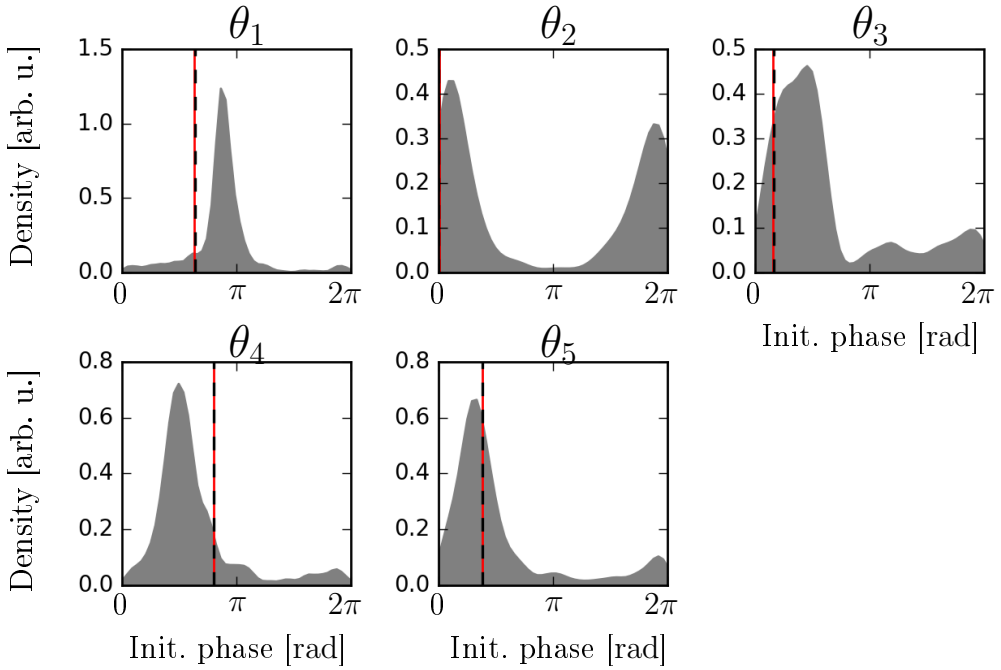


Figure B.7: The probability density for phase values obtained for the 3$^{\text{rd}}$ order KurSL in the comparison experiment. All values are in radians. Red vertical lines marks the global optima found for respective parameters. Values used to generate the input signal are indicated by a black dashed line.

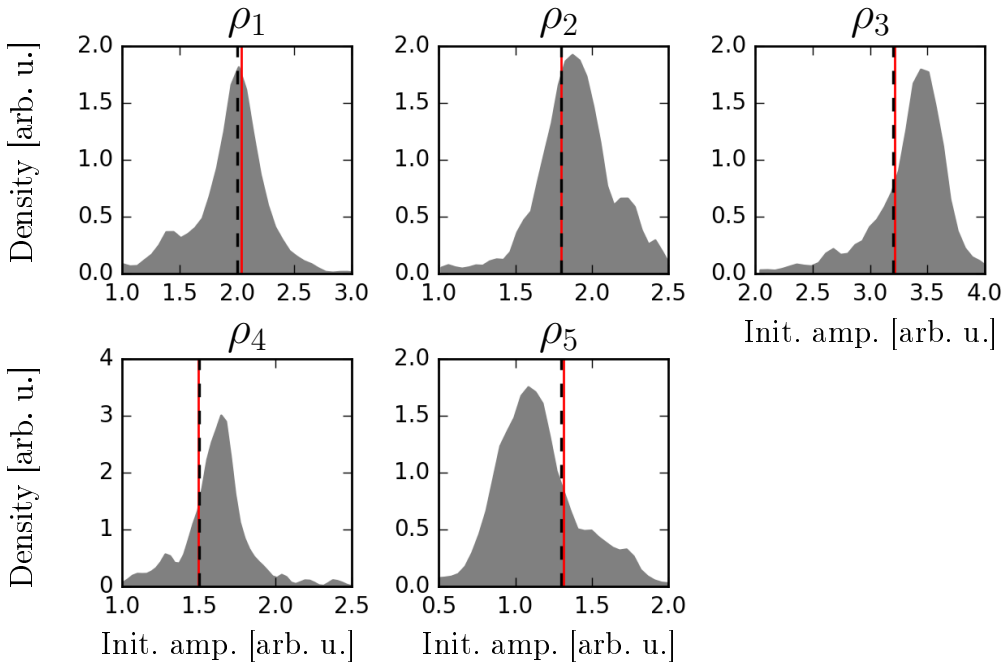


Figure B.8: The probability density for amplitude values obtained for the 3$^{\text{rd}}$ order KurSL in the comparison experiment. All amplitude values are in arbitrary units. Red vertical lines marks the global optima found for respective parameters. Values used to generate the input signal are indicated by a black dashed line.

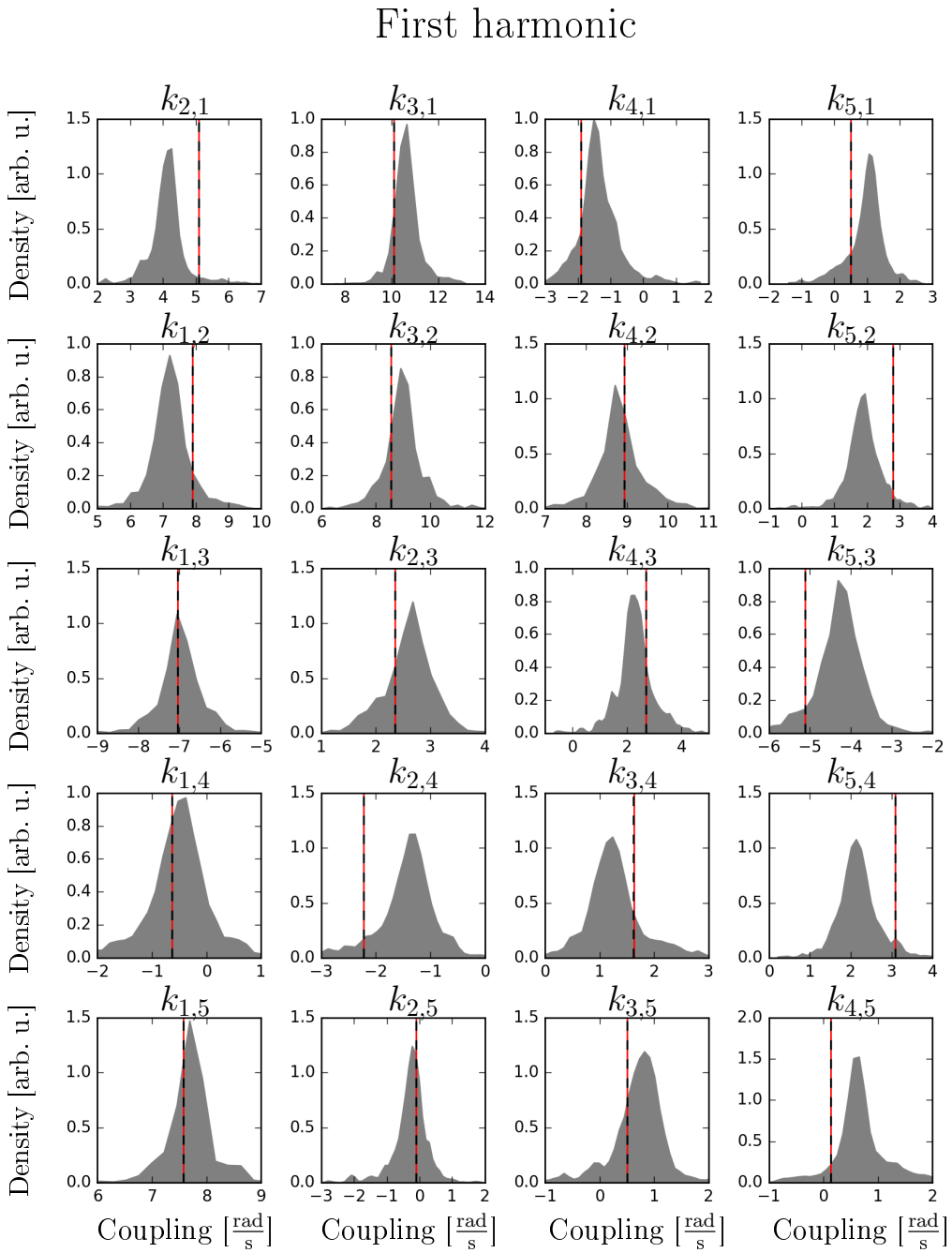


Figure B.9: The probability density for coupling scaling values related to the first harmonic obtained for the 3$^{\text{rd}}$ order KurSL in the comparison experiment. All coupling values are in rad/s units. Red vertical lines marks the global optima found for respective parameters. Values used to generate the input signal are indicated by a black dashed line.

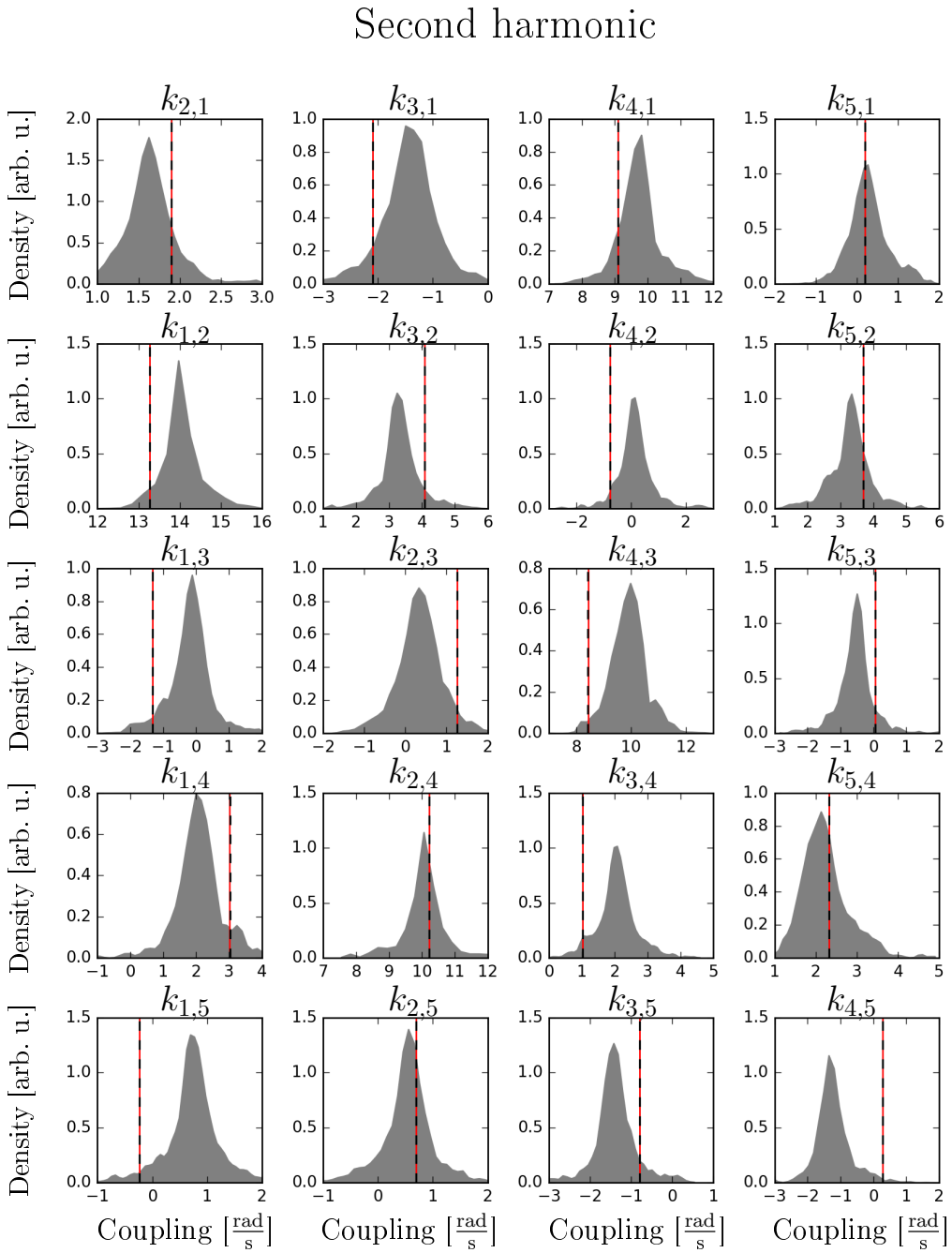


Figure B.10: The probability density for coupling scaling values related to the second harmonic obtained for the 3$^{\text{rd}}$ order KurSL in the comparison experiment. All coupling values are in rad/s units. Red vertical lines marks the global optima found for respective parameters. Values used to generate the input signal are indicated by a black dashed line.

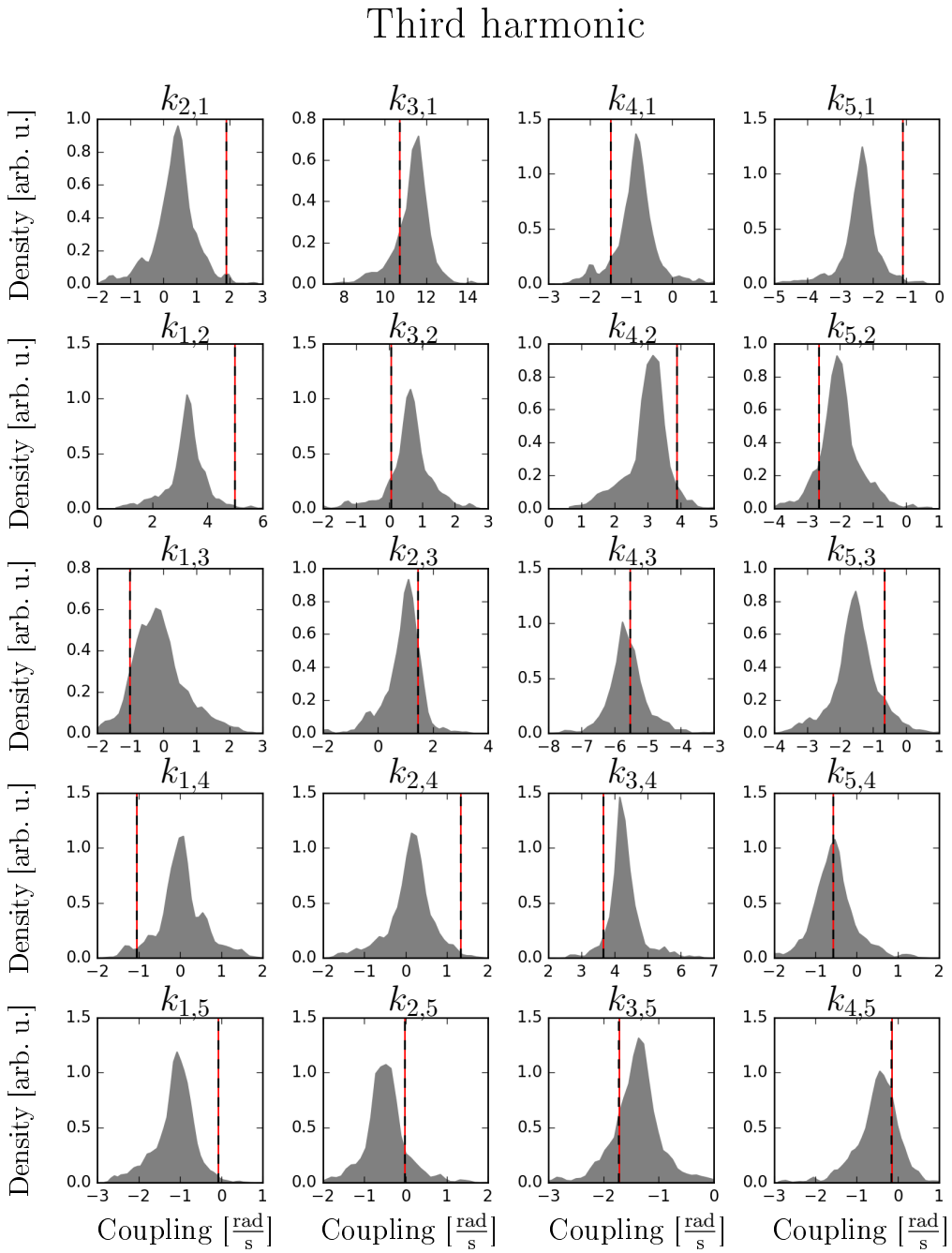


Figure B.11: The probability density for coupling scaling values related to the third harmonic obtained for the 3$^{\text{rd}}$ order KurSL in the comparison experiment. All coupling values are in rad/s units. Red vertical lines marks the global optima found for respective parameters. Values used to generate the input signal are indicated by a black dashed line.

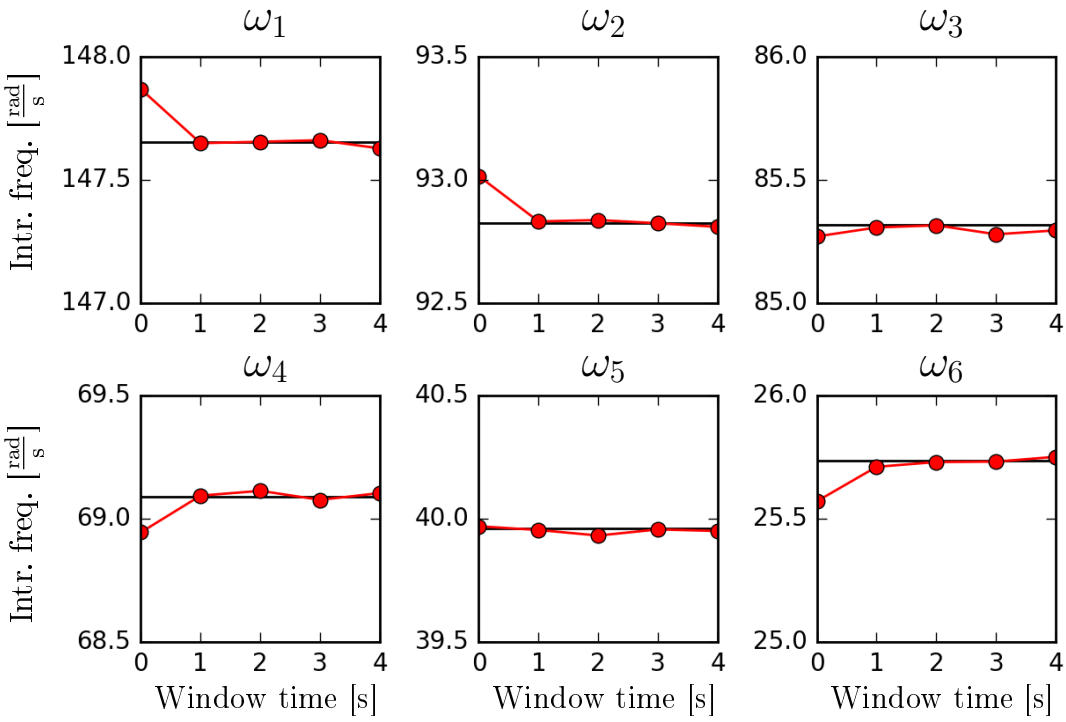


Figure B.12: Evolution of intrinsic frequencies for oscillators as indicated in the title. Values on x-axis denote initial time value for the respective segment, whereas y-axis holds range for intrinsic frequency values.

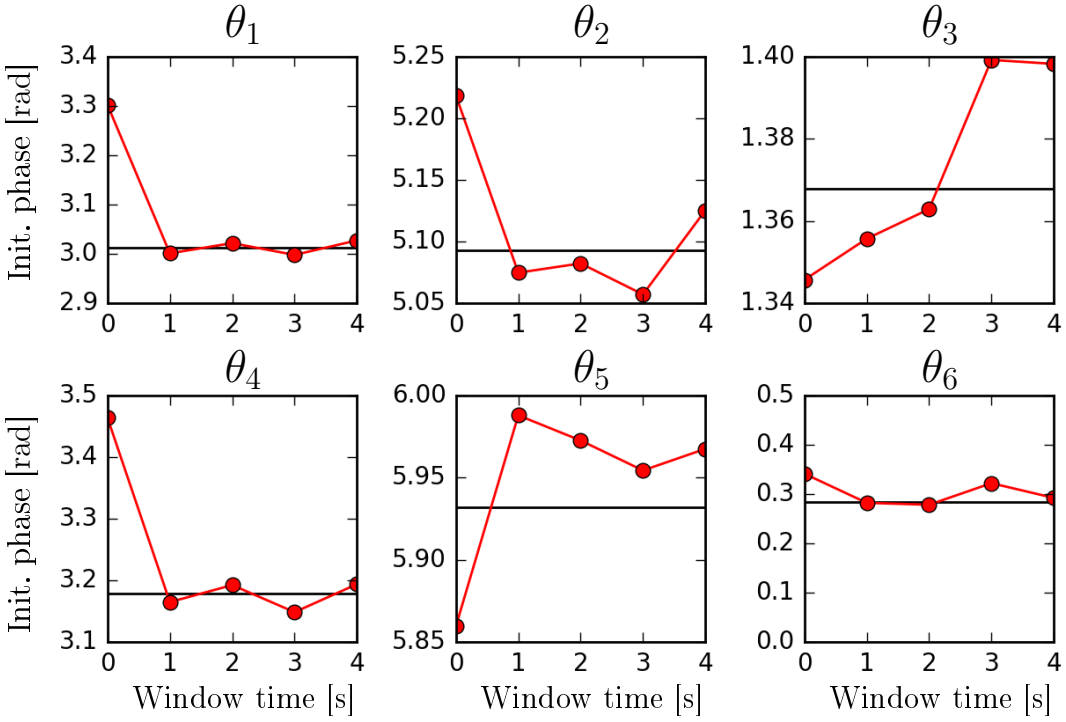


Figure B.13: Evolution of initial phases for oscillators as indicated in the title. Values on x-axis denote initial time value for the respective segment, whereas y-axis holds range for phase values.

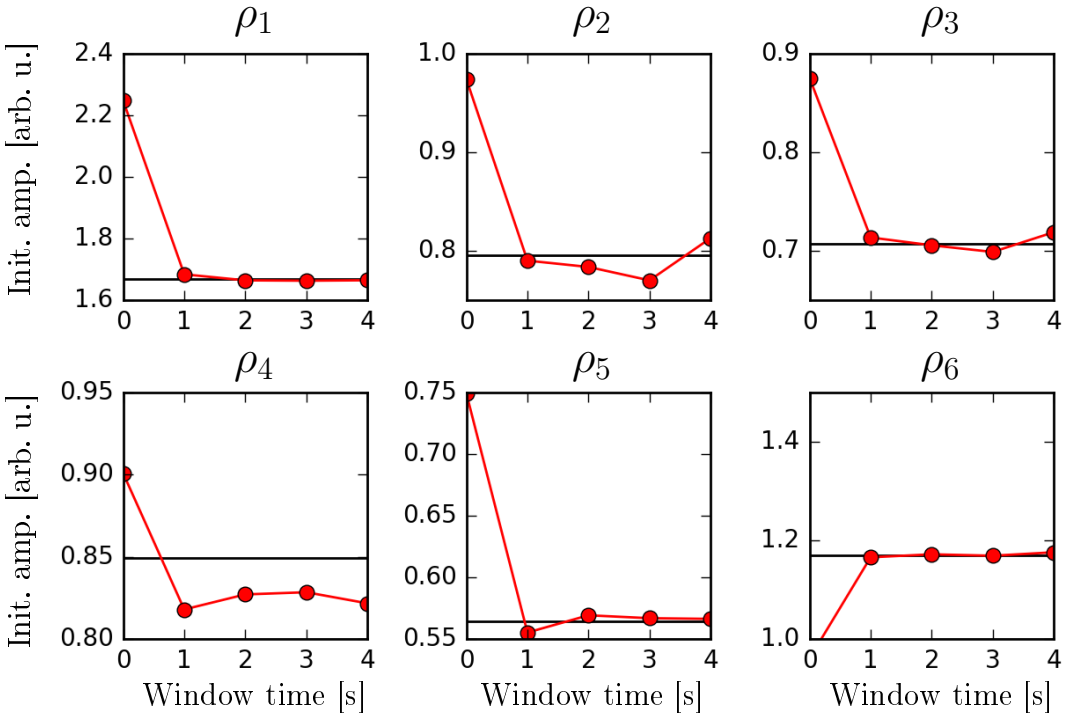


Figure B.14: Evolution of amplitudes for oscillators as indicated in the title. Values on x-axis denote initial time value for the respective segment, whereas y-axis holds range for amplitude values.

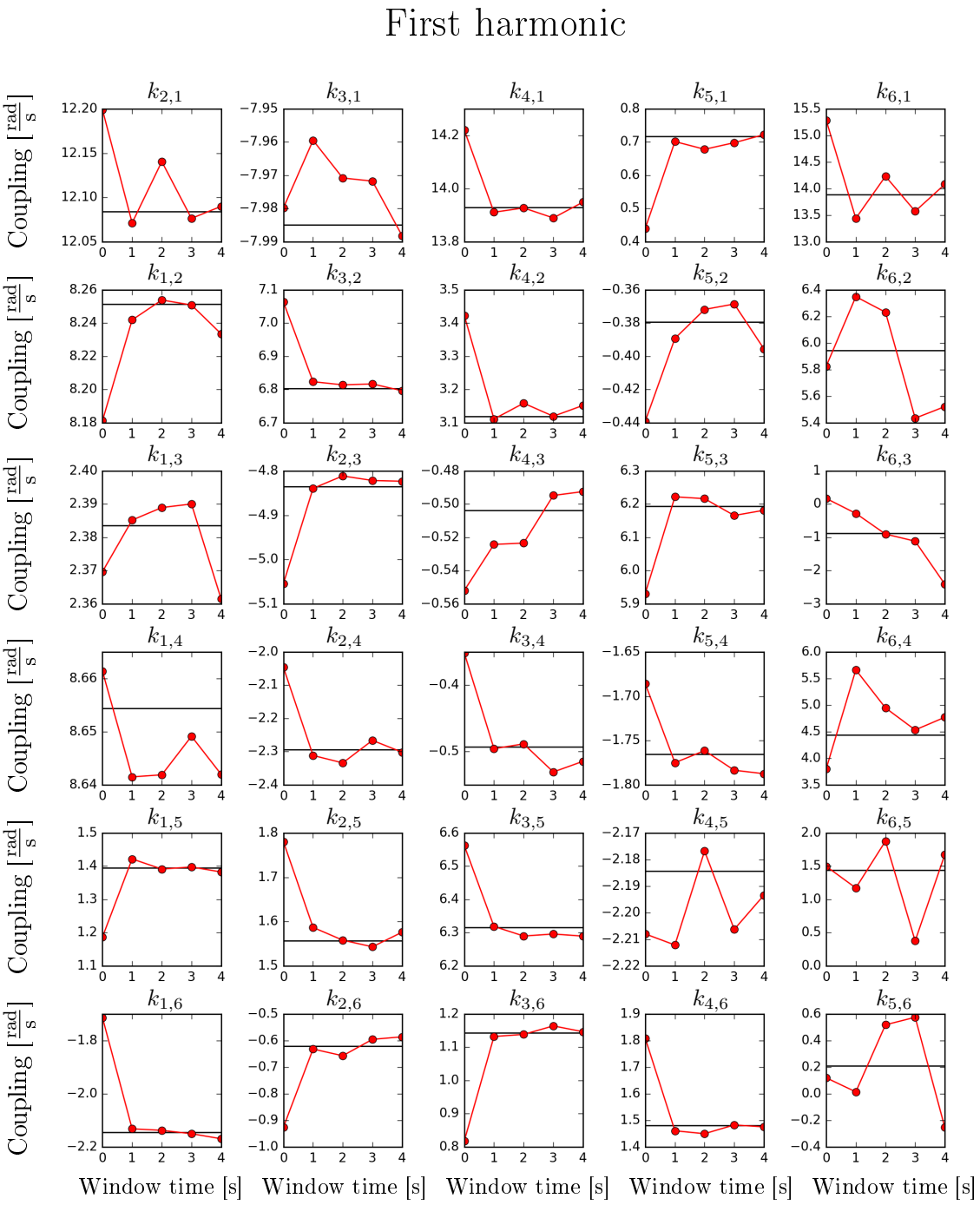


Figure B.15: Evolution of the coupling strengths scaling the first harmonics between oscillators indicated in the title. Values on x-axis denote initial time value for the respective segment, whereas y-axis holds range for coupling values.

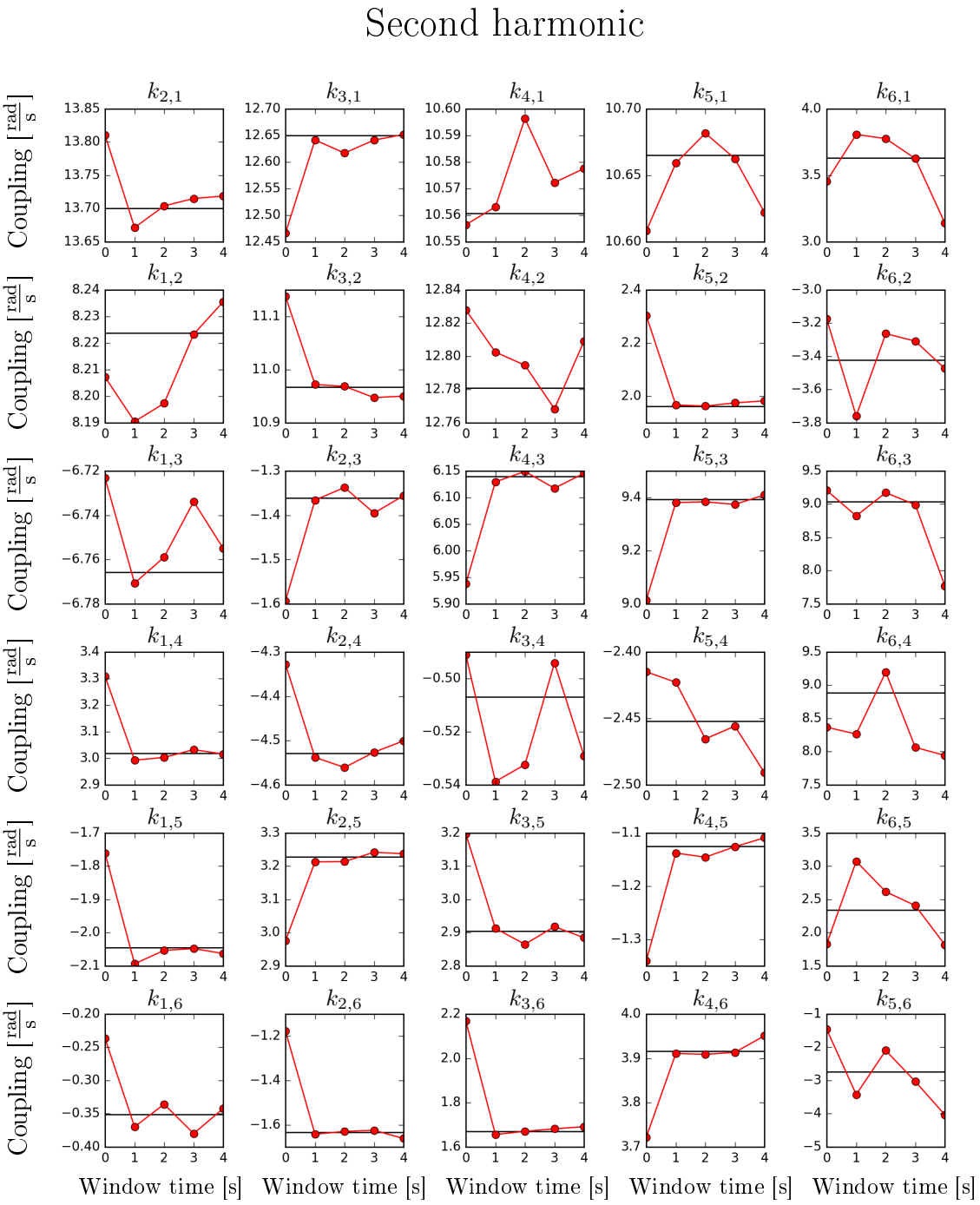


Figure B.16: Evolution of the coupling strengths scaling the second harmonics between oscillators indicated in the title. Values on x-axis denote initial time value for the respective segment, whereas y-axis holds range for coupling values.

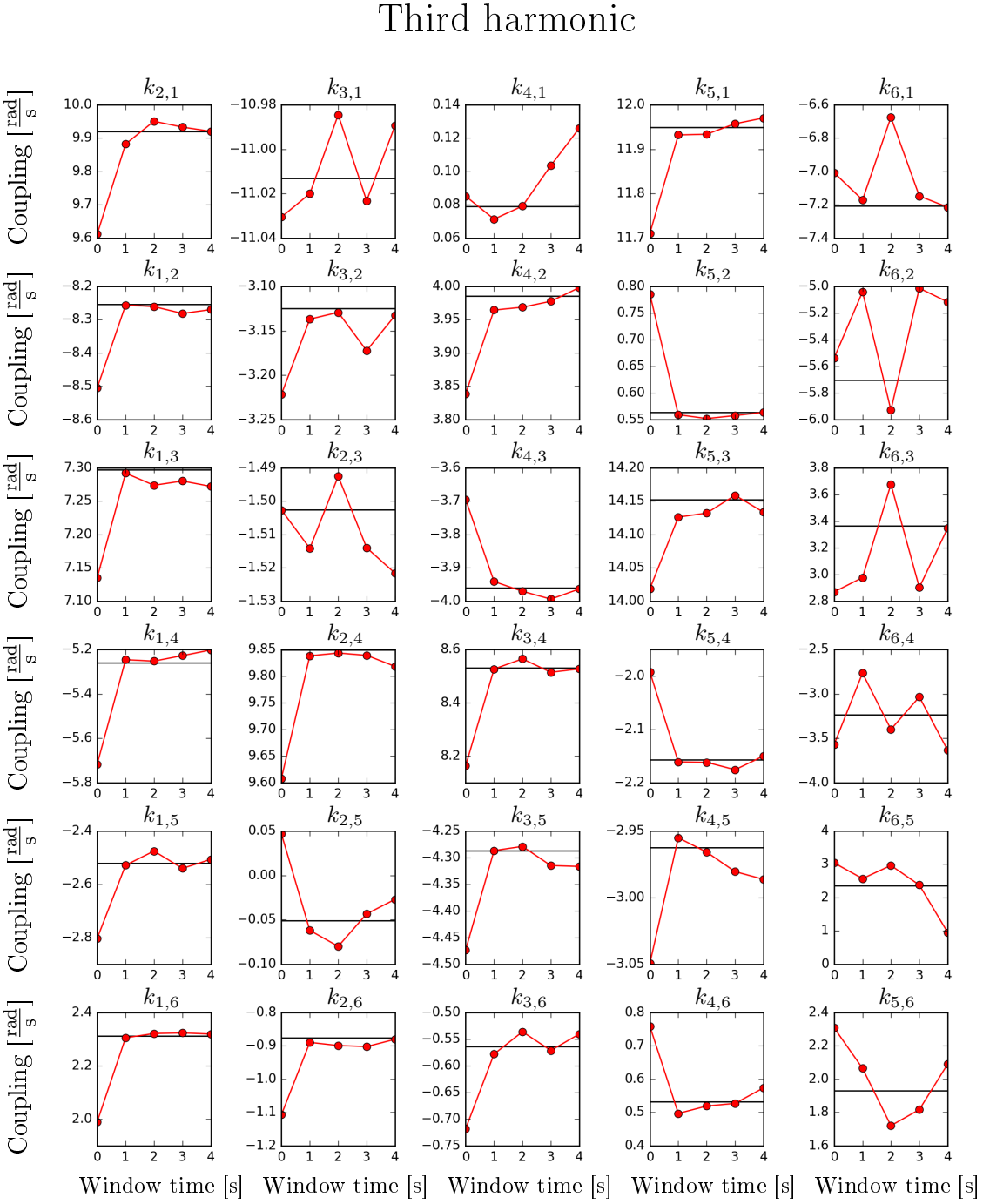


Figure B.17: Evolution of the coupling strengths scaling the third harmonics between oscillators indicated in the title. Values on x-axis denote initial time value for the respective segment, whereas y-axis holds range for coupling values.

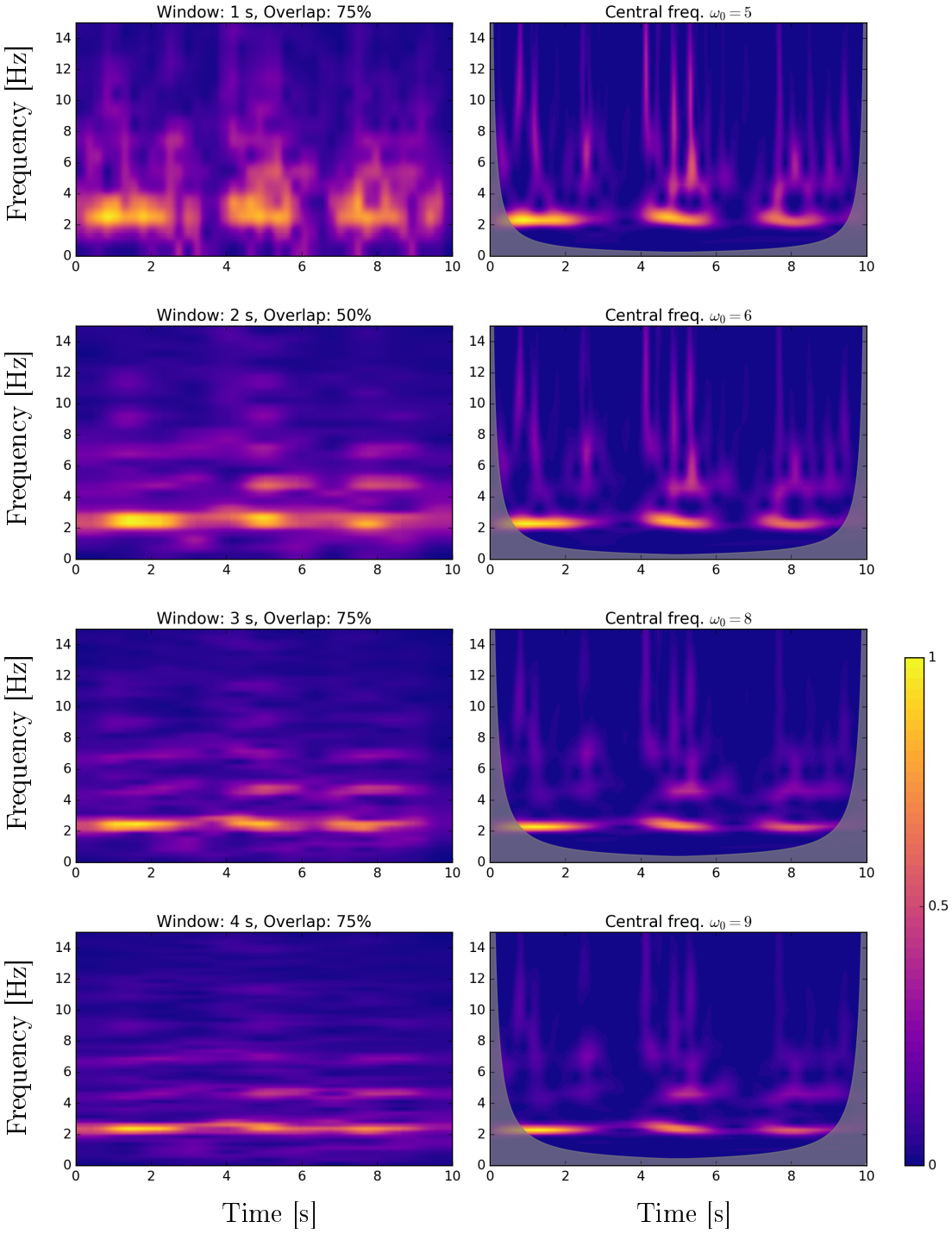


Figure B.18: Time-frequency representations computed for EEG signal used in stationary experiment (Sec. 7.3.1). The left column presents spectrogram representations from the top computed Hann window function using 1 s window with 75% overlap, 2 s window with 50% overlap, 3 s window with 75% overlap and 4 s window with 75% overlap. For the scaleogram which is presented in the right column all representations were obtained using Morlet wavelet. Each figure was obtained with different central frequency which from the top are $\omega_0 = 5$, $\omega_0 = 6$, $\omega_0 = 8$ and $\omega_0 = 9$. Values in all figures were scaled such that the maximum for any representation is one and the progression bar is displayed in the bottom right corner.

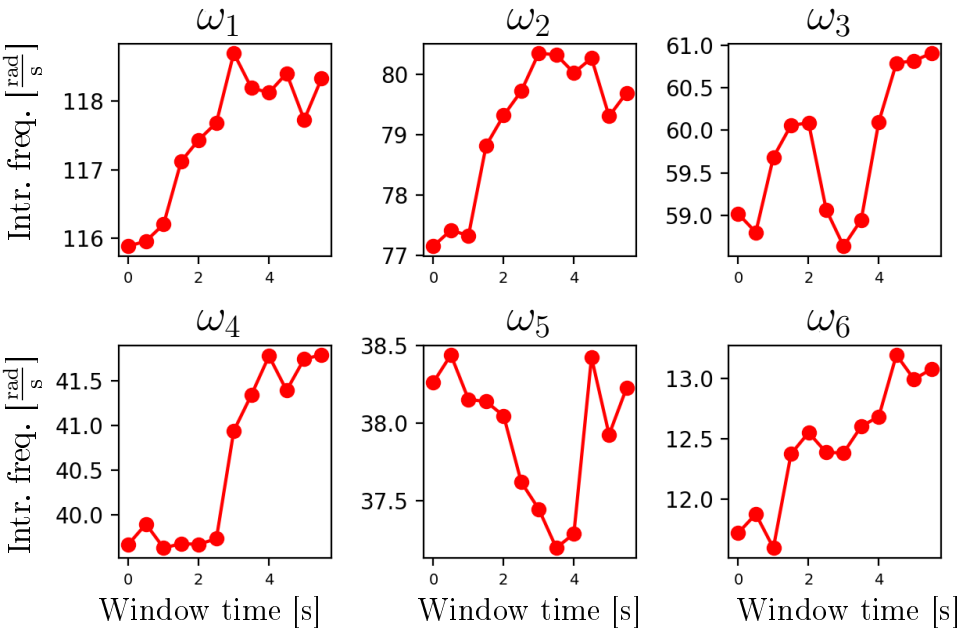


Figure B.19: Evolution of intrinsic frequencies for oscillators as indicated in the title. Values on x axis denote initial time value for respective segment, whereas y axis holds range for intrinsic frequency values.

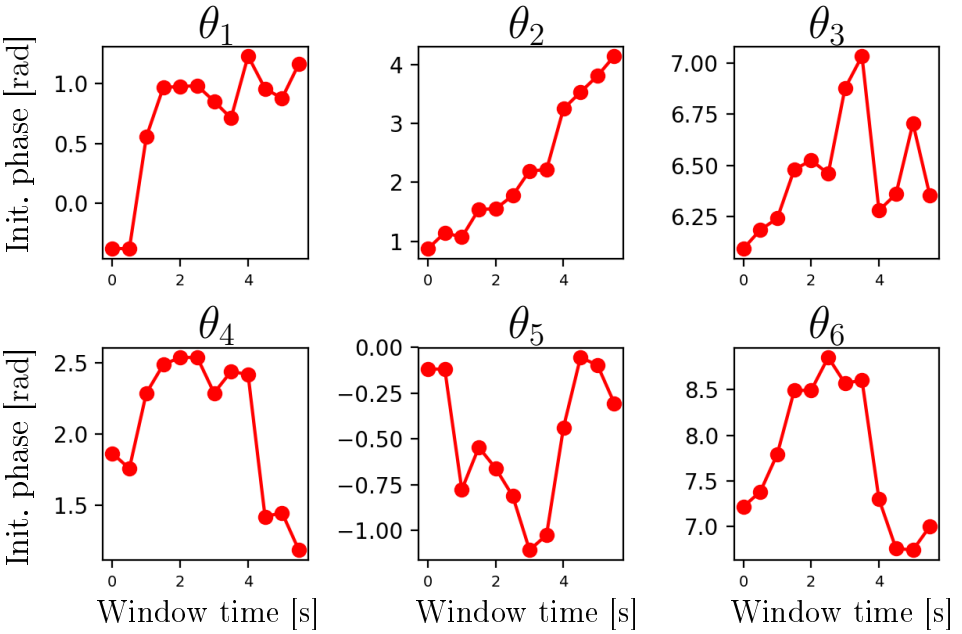


Figure B.20: Evolution of initial phases for oscillators as indicated in the title. Values on x axis denote initial time value for respective segment, whereas y axis holds range for phase values.

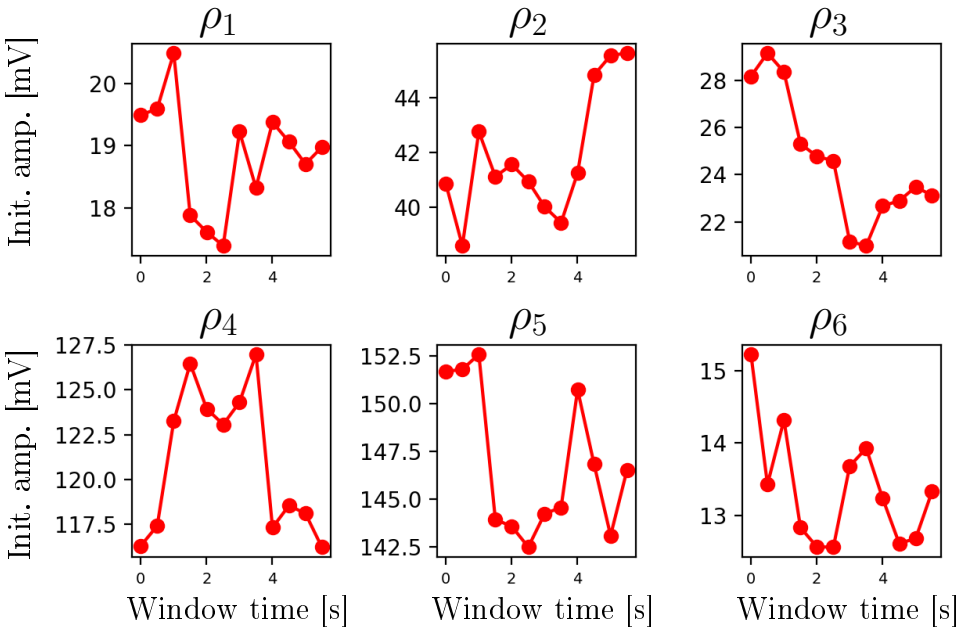


Figure B.21: Evolution of amplitudes for oscillators as indicated in the title. Values on x axis denote initial time value for respective segment, whereas y axis holds range for amplitude values in millivolts.

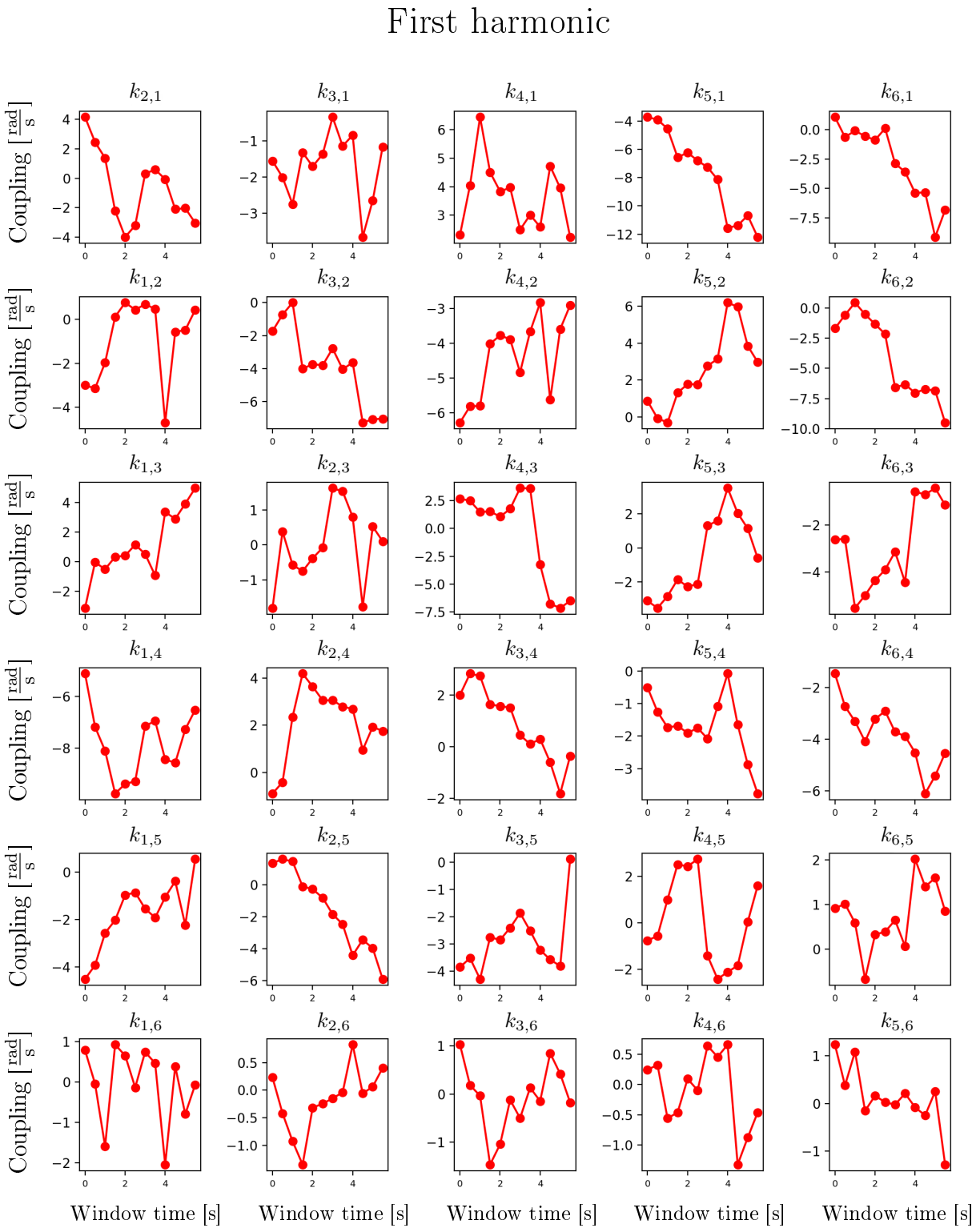


Figure B.22: Evolution of the coupling strengths scaling the first harmonics between oscillators indicated in the title. Values on x axis denote initial time value for respective segment, whereas y axis holds range for coupling values.

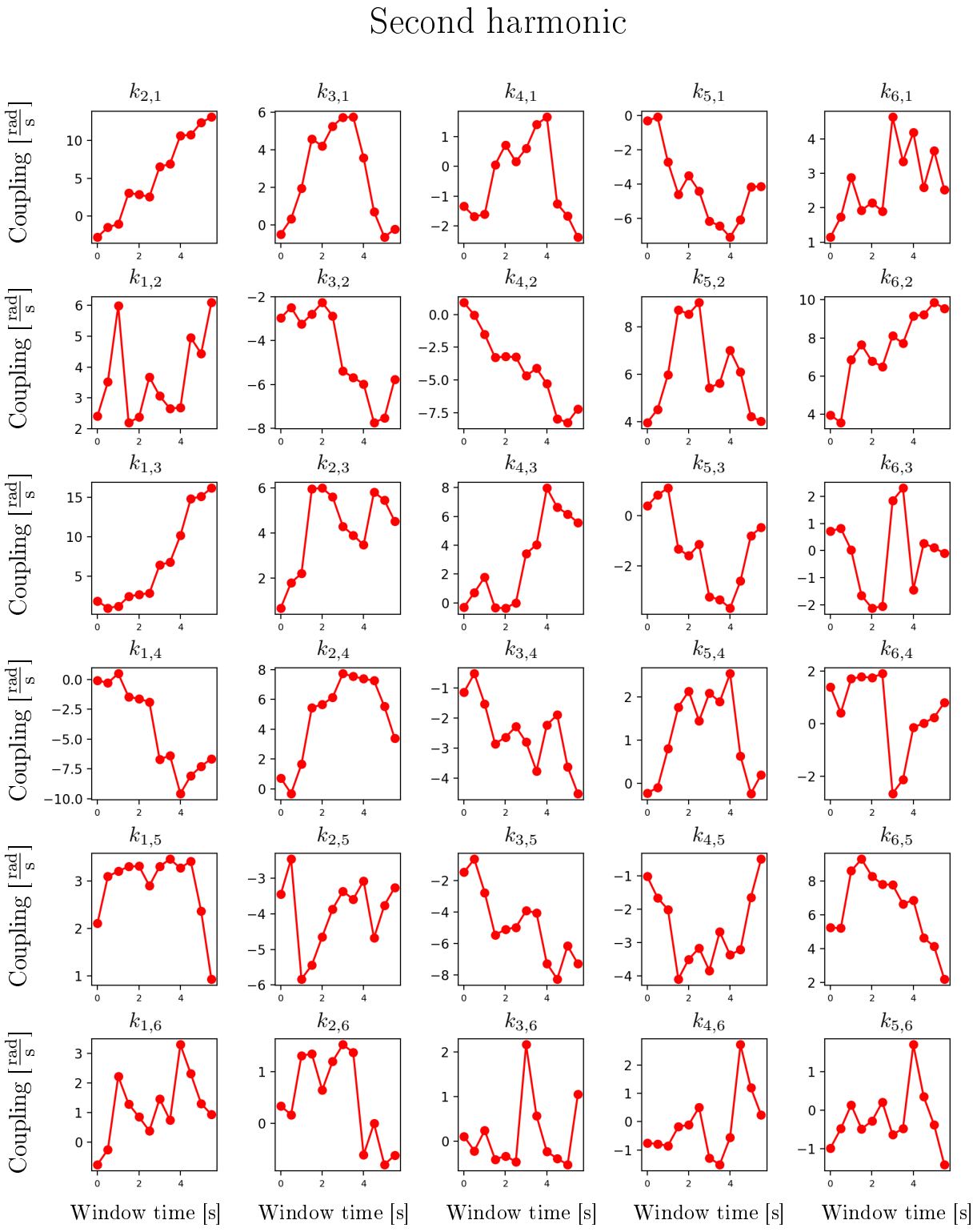


Figure B.23: Evolution of the coupling strengths scaling the second harmonics between oscillators indicated in the title. Values on x axis denote initial time value for respective segment, whereas y axis holds range for coupling values, whereas y axis holds range for coupling values.

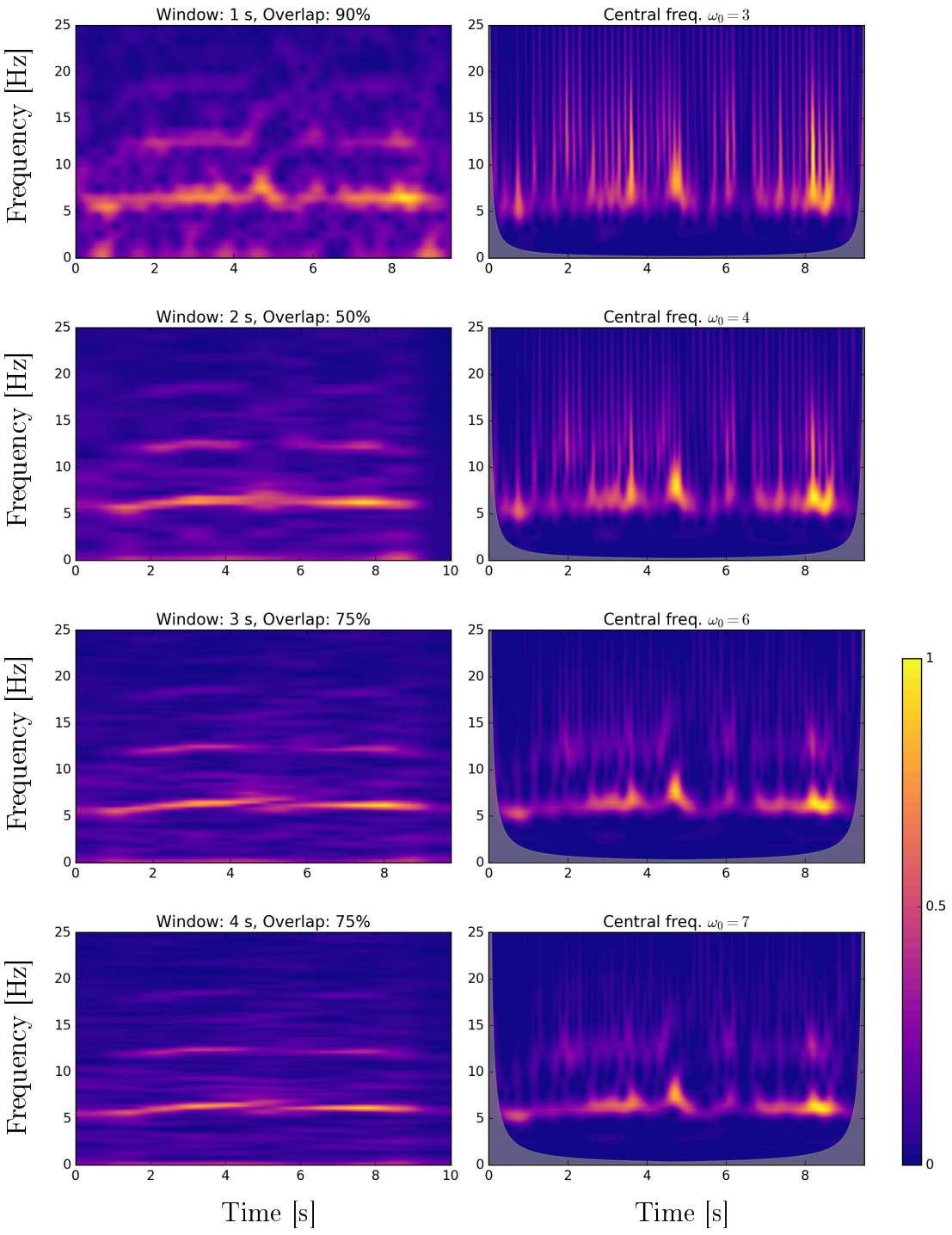


Figure B.24: Time-frequency representations computed for EEG signal used in the experiment with dynamic analysis (Sec. 7.3.2). The left column presents spectrogram representations from the top computed Hann window function using 1 s window with 90% overlap, 2 s window with 50% overlap, 3 s window with 75% overlap and 4 s window with 75% overlap. For the scaleogram which is presented in the right column all representations were obtained using Morlet wavelet. Each figure was obtained with different central frequency which from the top are $\omega_0 = 3$, $\omega_0 = 4$, $\omega_0 = 6$ and $\omega_0 = 7$. Values in all figures were scaled such that the maximum for any representation is one and the progression bar is displayed in the bottom right corner.