# Spectral Analysis for Signal Detection and Classification

## Reducing Variance and Extracting Features

Reinhold, Isabella

2021

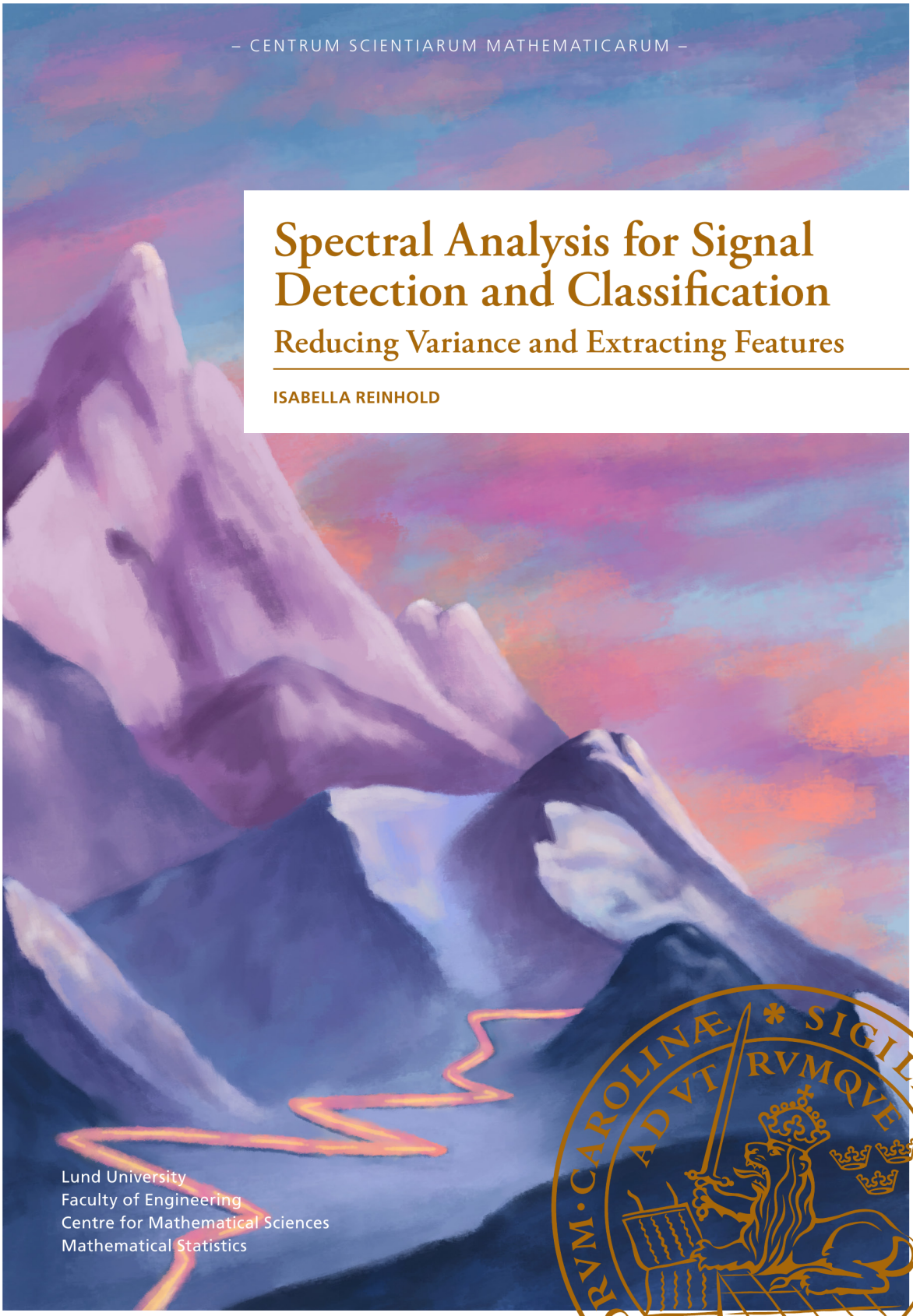
*Document Version:*
Publisher's PDF, also known as Version of record

– CENTRUM SCIENTIARUM MATHEMATICARUM –

# Spectral Analysis for Signal Detection and Classification

## Reducing Variance and Extracting Features

ISABELLA REINHOLD



Lund University
Faculty of Engineering
Centre for Mathematical Sciences
Mathematical Statistics

# Spectral Analysis for Signal Detection and Classification

Reducing Variance and Extracting Features

by

ISABELLA REINHOLD



Thesis for the degree of Doctor of Philosophy

by due permission of the Faculty of Engineering at Lund University, to be publicly defended on Wednesday, the 19th of May 2021, at 13:15, in lecture hall MH:Riesz, Sölvegatan 18, Lund.

Faculty opponent: Prof. Selin Aviyente,
Department of Electrical and Computer Engineering,
Michigan State University, USA.

Organization
**LUND UNIVERSITY**

Centre for Mathematical Sciences
Mathematical Statistics
Box 118
SE-221 00 Lund, Sweden

Document name
**DOCTORAL DISSERTATION**

Date of disputation
2021-05-19

Sponsoring organization

Author(s)
Isabella Reinhold

Title and subtitle
Spectral Analysis for Signal Detection and Classification: Reducing Variance and Extracting Features

Abstract
Spectral analysis encompasses several powerful signal processing methods. The papers in this thesis present methods for finding good spectral representations, and methods both for stationary and non-stationary signals are considered. Stationary methods can be used for real-time evaluation, analysing shorter segments of an incoming signal, while non-stationary methods can be used to analyse the instantaneous frequencies of fully recorded signals. All the presented methods aim to produce spectral representations that have high resolution and are easy to interpret. Such representations allow for detection of individual signal components in multi-component signals, as well as separation of close signal components. This makes feature extraction in the spectral representation possible, relevant features include the frequency or instantaneous frequency of components, the number of components in the signal, and the time duration of the components. Two methods that extract some of these features automatically for two types of signals are presented in this thesis. One adapted to signals with two longer duration frequency modulated components that detects the instantaneous frequencies and cross-terms in the Wigner-Ville distribution, the other for signals with an unknown number of short duration oscillations that detects the instantaneous frequencies in a reassigned spectrogram. This thesis also presents two multitaper methods that reduce the influence of noise on the spectral representations. One is designed for stationary signals and the other for non-stationary signals with multiple short duration oscillations. Applications for the methods presented in this thesis include several within medicine, e.g. diagnosis from analysis of heart rate variability, improved ultrasound resolution, and interpretation of brain activity from the electroencephalogram.

Key words
Instantaneous frequency, Multitaper, Signal resolution, Reassignment method, Time-frequency analysis

Classification system and/or index terms (if any)

Supplementary bibliographical information

Language
English

ISSN and key title
1404-0034, Doctoral Theses in Mathematical Sciences 2021:3

ISBN
978-91-7895-803-0

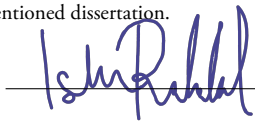
Recipient's notes

Number of pages
190

ISRN
LUTFMS-1049-2021

Security classification



Signature  Date 2021-04-07

# Spectral Analysis for Signal Detection and Classification

Reducing Variance and Extracting Features

ISABELLA REINHOLD

LUND UNIVERSITY

LTH

FACULTY OF ENGINEERING

Centre for Mathematical Sciences
Mathematical Statistics

Cover image by Isabella Reinhold



Mathematical Statistics
Centre for Mathematical Sciences
Lund University
Box 118
SE-221 00 Lund
Sweden

http://www.maths.lu.se/

Doctoral Theses in Mathematical Sciences 2021:3
ISSN: 1404-0034
ISBN: 978-91-7895-803-0 (print)
ISBN: 978-91-7895-804-7 (pdf)
ISRN: LUTFMS-1049-2021

Printed in Sweden by Media-Tryck, Lund 2021

*Whatever is going on is as usual.*

— Margaret Atwood, The Handmaid's tale

# Contents

## Acknowledgements

I want to thank my supervisor Maria Sandsten, who has been supportive and encouraging throughout the whole of my PhD studies. It has truly been so much fun working together, and I have learned a lot about research, teaching, and the benefits of mountaineering. I would also like to thank my co-supervisors, Josefin Starkhammar, for introducing me to the interesting field of dolphin echolocation and ultrasound, and Andreas Jakobsson, for making conferences more fun. A special thanks as well to my colleagues and co-authors, Johan Brynolfsson and Rachele Anderson, who always impress me, and to my colleague Maria Juhlin on whom I can always rely on for support and interesting insights. Finally, I want to thank the amazing administration staff, James Hakim, Susann Nordqvist, Denise Bokander, Maria Lövgren, and Natasa Olsson, who are always helpful and make the workplace warm and inviting.

Tack till min farmor som gjorde det roligt att arbeta med matematikläxorna innan fotbollsträningen. Tack till mina föräldrar som alltid uppmuntrat mig att ställa frågor. Tack till mina vänner, speciellt Alex, Meng, och Åse, som alltid får mig att skratta och alltid finns där. Och slutligen, tack till min man som inte bara stått ut med mig under forskarstudierna men också gjort livet lite enklare och mycket roligare.

*Lund, 2021* *Isabella Reinhold*

## Abstract

Spectral analysis encompasses several powerful signal processing methods. The papers in this thesis present methods for finding good spectral representations, and methods both for stationary and non-stationary signals are considered. Stationary methods can be used for real-time evaluation, analysing shorter segments of an incoming signal, while non-stationary methods can be used to analyse the instantaneous frequencies of fully recorded signals. All the presented methods aim to produce spectral representations that have high resolution and are easy to interpret. Such representations allow for detection of individual signal components in multi-component signals, as well as separation of close signal components. This makes feature extraction in the spectral representation possible, relevant features include the frequency or instantaneous frequency of components, the number of components in the signal, and the time duration of the components. Two methods that extract some of these features automatically for two types of signals are presented in this thesis. One adapted to signals with two longer duration frequency modulated components that detects the instantaneous frequencies and cross-terms in the Wigner-Ville distribution, the other for signals with an unknown number of short duration oscillations that detects the instantaneous frequencies in a reassigned spectrogram. This thesis also presents two multitaper methods that reduce the influence of noise on the spectral representations. One is designed for stationary signals and the other for non-stationary signals with multiple short duration oscillations. Applications for the methods presented in this thesis include several within medicine, e.g. diagnosis from analysis of heart rate variability, improved ultrasound resolution, and interpretation of brain activity from the electroencephalogram.



**Keywords:** Instantaneous frequency, Multitaper, Signal resolution, Reassignment method, Time-frequency analysis

## Populärvetenskaplig sammanfattning

Att vi i en nära framtid kommer kunna styra tekniska prylar med tankekraft och använda våra mobiler för att själva ställa medicinska diagnoser känns kanske mest som en dröm, men med nya metoder för signalbehandling är det en möjlighet. Vi måste komma ihåg att vi idag är vana vid röststyrning av våra mobiltelefoner och att robotdammsugare städar i våra hem, något som verkade vara avlägsen science fiktion för bara några årtionden sedan. Dagens vanliga mobiltelefoner är flera miljoner gånger snabbare än de mest avancerade datorerna på 1960-talet och denna tekniska revolution har utökat möjligheterna för signalbehandling.

Min forskning fokuserar på att få information från signaler med hjälp av spektralanalys. Det är enkelt att tänka på ljud, som musik och fågelsång men också puttrandet från kaffebryggaren, som signaler, men mycket mer kan klassas som signaler. Allt som varierar över tid eller över ett avstånd kan anses vara en signal, det betyder att till exempel mätningar av temperatur och hjärtslag är signaler. Dessa är också bra exempel på att signaler ofta är stokastiska. Alltså att även om det finns en struktur som signalen följer, som att dagens temperatur liknar gårdagens och att hjärtslag generellt ser likadana ut för alla människor, så finns det en slumpmässighet. Dagens temperatur kan inte förutspås exakt med gårdagens och hjärtslag kan till exempel ta olika lång tid och vara olika starka. En av förutsättningarna i min forskning är alltså att metoderna ska fungera trots att signalerna är stokastiska.

Hur en signal varierar kan beskrivas med dess frekvenser. Spektralanalys är den del inom signalbehandling som med olika metoder undersöker vilka frekvenser signaler innehåller, och ofta hur frekvenserna varierar med tid eller avstånd. Att på ett säkert sätt kunna ta fram en signals frekvenser är inte alltid lätt, men det kan ge mycket information om signalen och kanske viktigast, information om den eller det som orsakat signalen. Frekvenserna i en fågelsång kan avslöja vilken fågelart som sjunger, vissa ändringar i hjärtrytmen indikerar sjukdomstillstånd,

och en förändring i kaffebryggarens normala puttrande kan berätta att apparaten är på väg att gå sönder.

En signal är ofta inte strikt bara en signal, utan en summa av flera olika komponenter. Tänk själv på hur mycket ljud vi hör hela tiden, när vi lyssnar på något specifikt kan vi också höra trafik, maskinsurr, djur, andra människor eller något annat i bakgrunden, vi kallar sådana störningar för brus. Det är därför viktigt att kunna detektera och klassificera olika komponenter i en signal. I den här avhandlingen presenteras två metoder som automatiskt detekterar signalkomponenter. Metoderna är anpassade för olika typer av signaler och situationer. Den ena används för signaler med två komponenter i en brusig omgivning och detekterar komponenterna för att sedan ge en så optimal visualisering av signalernas frekvenser som möjligt. Den andra hittar alla korta, plötsliga signalkomponenter i en längre signal med okänt många komponenter i en brusig omgivning, vilket används för att separera komponenter som uppmäts nära i tid.

Då störningar i signaler är ett stort problem i många applikationer för spektralanalys, är det viktigt med metoder som minskar störningarnas inverkan. Den här avhandlingen presenterar två metoder som gör just det. Den ena är designad för signaler vars frekvenser inte varierar över tid eller plats och kombinerar bra brusreducering med en effektiv implementering. Den andra metoden fokuserar på att minska brusets påverkan för signaler som har väldigt mycket brus och flera korta, plötsliga signalkomponenter.

Dessa metoder kräver vissa strukturer på signalerna för att kunna användas, men eftersom dessa olika strukturer är vanligt förekommande kan metoderna användas inom flera applikationer. I avhandlingen finns exempel på flera medicinska tillämpningar, bland annat en som analyserar variationer i hjärtrytm. Variationer i ens hjärtrytm är normalt och det kan vara svårt att identifiera dåliga variationer, samt exakt vad de betyder. En annan tillämpning ligger inom analys av hjärnsignaler. Vi mäter gärna hjärnsignaler utanpå huvudet, vilket medför att signalerna är väldigt brusiga och svårtydda. Men det är just med bättre metoder för att analysera de här signalerna som det blir möjligt att identifiera sjukdomar både tidigt och med god säkerhet, samt att tyda avsikter från hjärnsignaler som kommer ge oss möjligheten att styra saker med tankekraft.

## Popular summary

It might seem like wishful thinking that we in the near future will be able to control technical devices with our minds, or that we will be able to make reliable diagnoses using our own mobile phones. With new methods in signal processing this will however soon be our reality. We are already used to voice commands on our mobile phones and robot vacuum cleaners in our homes, which would have seemed like pure science fiction only a few decades ago. Today's standard mobile phone is several million times faster than the most advanced computers of the 1960's, and this technological revolution has greatly expanded the possibilities for signal processing.

My research is focused on gaining information from signals using spectral analysis. It is easy to think of sounds, like music and bird song but also the rumbling and gurgling sounds of a kettle or coffee maker, as signals. However, everything that varies over time or space can be defined as a signal. So, for instance a series of temperature measurements or heartbeats, are signals. These are also good examples of stochastic signals, which means that they have some structure but also random properties. Today's temperature will be similar to yesterday's but not the same, and we can easily recognise a heartbeat, but a heartbeat will vary in length and strength. In my field of research, we need good methods to analyse signals even when they are stochastic, and this is where spectral analysis comes in.

How a signal varies can be described by its frequencies. Spectral analysis is a field within signal processing focused on finding what frequencies signals have, and typically how these frequencies vary with time or space. It is not always easy to accurately extract a signal's frequencies, but they can provide much information about the signal, and perhaps most importantly, about the person, animal or object that made the signal. The frequencies in bird songs can tell us what type of bird is singing, certain changes in the heart rate can indicate illness, and a change in the sound from your kettle or coffee maker can tell if it is about to break.

A signal is often not just one signal, but a sum of several components. Consider the everyday background sounds when you listen for something specific but also hear traffic, machines humming, animals, people, or something else. We call these disturbances noise. It is therefore important to be able to detect and classify different components in signals, separating the noise from the components that we are interested in. This thesis presents two methods for automatic component detection, and the methods are adapted for different types of signals and situations. The first is used on signals with two components in a noisy environment, it detects the two components and then calculates an optimal visual representation of the frequencies of those components. The second method finds short, sudden components in a longer noisy signal, where the number of components initially is unknown. This method is used to separate components that have been measured close together in time.

Disturbances in signals is a large problem for many applications of spectral analysis, so methods to lessen the effects of noise are very important. The first method is designed for signals with frequencies that do not vary over the measurement time or space, and it combines good noise reduction capabilities with an efficient implementation. The second method focuses on minimising the effect of noise when there is much of it, in signals with several short and sudden components.

All these methods need the signals to have certain qualities to work well, but these qualities are general and common, which means that the methods can be used in many applications. This thesis includes examples of several medical applications, including one that analyses variation in heart rate. Some variation in heart rate is normal and it can be difficult to detect bad variations and find out what they are indicators of. Another application is within analysis of brain signals. We typically prefer to measure brain signals on the outside the head, but this also makes the signals noisy and hard to interpret. With methods that give good and reliable analysis of these signals, we will be able to detect diseases at an earlier stage, and we will be able to interpret intentions from the brain signals giving us the possibility to control technical devices with our minds.

## List of publications

This thesis is based on the following publications:

**A** Isabella Reinhold and Maria Sandsten, "Efficient Thomson Spectral Estimator with Time-Shifted Windows", *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 8236–8240, 2019. doi: 10.1109/ICASSP.2019.8683588

**B** Isabella Reinhold and Maria Sandsten, "Multitaper Reassignment for Oscillating Transients with Gaussian Envelopes", manuscript.

**C** Isabella Reinhold and Maria Sandsten, "Optimal Time-Frequency Distributions using a Novel Signal Adaptive Method for Automatic Component Detection", *Signal Processing*, vol. 133, pp. 250–259, 2017. doi: 10.1016/j.sigpro.2016.11.028

**D** Isabella Reinhold, Maria Sandsten, and Josefin Starkhammar, "Objective Detection and Time-Frequency Localization of Components within Transient Signals", *J. Acoust. Soc. Am.*, vol. 143, no. 4, pp. 2368–2378, 2018. doi: 10.1121/1.5032215

Publications not included in this thesis:

I Josefin Starkhammar, Isabella Reinhold, Tobias Erlöv, and Maria Sandsten, "Scaled reassigned spectrograms applied to linear transducer signals", under revision for *JASA Express Letters*.

II Johan Brynolfsson, Isabella Reinhold, and Maria Sandsten, "A time-frequency-shift invariant parameter estimator for oscillating transient functions using the matched window reassignment", *Signal Processing*, vol. 183, 2021.

III Maria Sandsten, Isabella Reinhold, and Rachele Anderson, "A multitaper reassigned spectrogram for increased time-frequency localization precision", *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 5310–5314, 2020.

IV Maria Sandsten, Rachele Anderson, Isabella Reinhold, and Johan Brynolfsson, "The matched reassigned cross-spectrogram for phase estimation", *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 5960–5964, 2020.

V Johan Brynolfsson, Isabella Reinhold, Josefin Starkhammar, and Maria Sandsten, "The matched reassignment applied to echolocation data", *ICASSP 2019 -2019 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 8236–8240, 2019.

VI Josefin Starkhammar, Isabella Reinhold, Patrick W. Moore, Dorian S. Houser, and Maria Sandsten, "Detailed analysis of two detected overlaying transient components within the echolocation beam of a bottlenose dolphin (Tursiops truncatus)", *J. Acoust. Soc. Am.*, vol. 145, no. 4, pp. 138–2148, 2019.

VII Maria Sandsten, Johan Brynolfsson, and Isabella Reinhold, "The matched window reassignment", *26th European Signal Processing Conference (EUSIPCO)*, pp. 2340–2344, 2018.

VIII Maria Sandsten, Isabella Reinhold, and Josefin Starkhammar, "Automatic time-frequency analysis of echolocation signals using the matched Gaussian multitaper spectrogram", *Interspeech 2017*, pp. 3048–3052, 2017.

IX Isabella Reinhold, Kenneth Batstone, Isabel M. Gallardo González, Andrea Troian, and Rixin Yu, "Quality of academic writing for engineering students at Lund University", *The 2nd European conference for the scholarship of teaching and learning*, pp. 208–214, 2017.

X Isabella Reinhold, Josefin Starkhammar, and Maria Sandsten, "The scaled reassigned spectrogram adapted for detection and localisation of transient signals", *25th European Signal Processing Conference (EUSIPCO)*, pp. 937–941, 2017.

xi Josefin Starkhammar, Isabella Reinhold, Patrick Moore, Dorian Houser, and Maria Sandsten, "Intra-click time-frequency patterns across the echolocation beam of a beluga whale", *J. Acoust. Soc. Am.*, vol. 140, no. 4, pp. 3239, 2016.

xii Isabella Reinhold, and Maria Sandsten, "A novel Doppler penalty function for the multitaper Wigner-Ville distribution", *11th IMA International Conference on Mathematics in Signal Processing*, 2016.

## Abbreviations and symbols

| | |
|---|---|
| 1D, 2D, 3D | one-dimensional, two-dimensional, three-dimensional |
| AM | amplitude modulated |
| FM | frequency modulated |
| SNR | signal-to-noise ratio |
| TF | time-frequency |
| | |
| $i$ | imaginary unit |
| $f$ | frequency |
| $\phi$ | phase |
| $t$ | time (continuous) |
| $n$ | time (discrete) |
| $\tau$ | time-lag |
| $\mathbb{C}$ | complex numbers |
| $\mathbb{N}$ | natural numbers without zero |
| $\mathbb{R}$ | real numbers |
| $\mathbb{Z}$ | integers |
| $\mathcal{N}(\mu, \sigma^2)$ | normal distribution with expected value $\mu$ and variance $\sigma^2$ |
| $\mathcal{U}(a, b)$ | uniform distribution with lower and upper bounds $a$ and $b$ |
| $\int$ | integrals without indicated limits go from $-\infty$ to $\infty$ |
| $x^*(t)$ | complex conjugate of $x(t)$ |
| $f(t) * g(t)$ | convolution between $f(t)$ and $g(t)$ |
| E[X] | expected value of X |
| Var[X] | variance of X |
| Cov[X] | covariance of X |
| $\hat{\mathrm{X}}$ | estimate of X |
| $\Re(x)$ | real part of $x$ |
| $\Im(x)$ | imaginary part of $x$ |

# Introduction

# Introduction to Spectral Analysis and Signal Processing

*The arts and sciences are avatars of human creativity.*
— Mae Jemison

The focus of this thesis is spectral analysis of signals based on mathematical models and statistical analysis. The original research is presented in four papers, covering different methods and signals. While the papers consider different signals and applications, they all describe methods aimed to extract information from signals from either a pure frequency or a joint time-frequency representation. There is also an overarching focus on methods that have easily interpreted results, this can be intuitive visual representations or the automatic detection of key features. This first chapter provides a background to the concepts developed further by the papers.

## 1 The idea of spectral analysis

Spectral analysis allows us to study the frequency content of signals. While the frequency or possibly several frequencies of a signal can be observed as the period of oscillations in the *time domain* of signals (if indeed the signal has oscillations), a transformation to the *frequency domain* makes analysis easier. Spectral analysis of signals not dependent on time, but instead spatial variables, is also possible but not a focus in this thesis. Methods for spectral analysis started emerging during the first half of the 1800s, most notable is perhaps the Fourier transform

$$X(f) = \int x(t)e^{-i2\pi ft}dt, \tag{1}$$

which is a linear transform, with the inverse

$$x(t) = \int X(f) e^{i2\pi f t} df. \tag{2}$$

The Fourier transform holds for almost all real- and complex-valued functions $x(t)$, even if they are not continuous in $t$ or if $\int |x(t)| dt$ is not bounded. The exact definition for a set of functions for which the Fourier transform holds and accompanying proofs are outside the scope of this thesis, it is however important to note that the Fourier transform holds also for functions that are not Riemann or Lebesgue integrable, as such functions will be handled in this thesis.

The theoretical concept of the spectral density

$$S_{\mathrm{X}}(f) = \int r_{\mathrm{X}}(\tau) e^{-i2\pi f \tau} d\tau, \tag{3}$$

is the Fourier transform of the theoretical covariance function, $r_{\mathrm{X}}(\tau)$, of a (wide-sense) stationary stochastic process, $\mathrm{X}(t)$. For now we will be content to know that a stochastic process is a random function for which the fixed outcomes, or realisations, have some characteristics in common. For stationary processes these characteristics are also unchanging over time.

Two signals with interesting spectral densities and covariance functions are the random harmonic oscillation with a deterministic frequency, and the white noise process. As stochastic processes these are well defined by their corresponding covariance function and spectral density, which is demonstrated in Figure 1. Two realisations of each process, Figures 1(a)–(b), might hint at similarities between the realisations, but it is much easier to see these similarities in the covariance function and spectral density, Figure 1(c)–(f).

The covariance function describes dependencies within a process and in the stationary case it is a function of time lag, i.e. differences in time within the process. An intuitive understanding of this can be gained by studying the difference between the covariance function of a random harmonic oscillation and a white noise process. After observing a harmonic oscillation for a short time it is easy to know how it will progress, it will continue to oscillate with a given period, $1/f_0$, this is also what the covariance function, Figure 1(c), tells us. Similar observation of white noise will not make it easier to predict how the realisation will progress, the covariance function, Figure 1(d), is thus zero except for the variance that indicates how large variations can be expected.

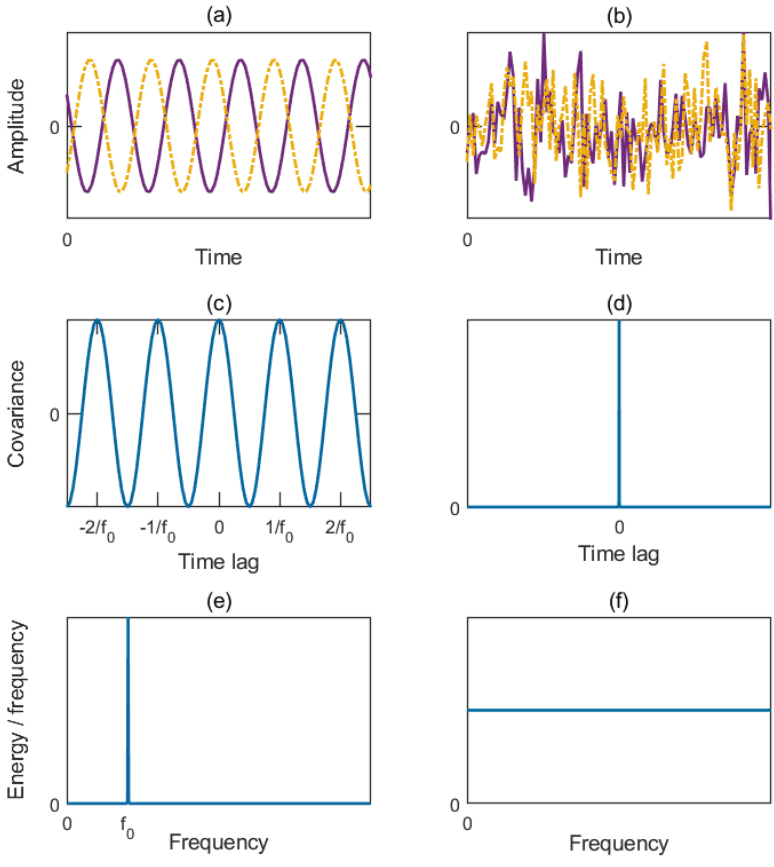



**Figure 1:** (a) Two realisations of a harmonic sinusoidal oscillation, (b) two realisations of white noise, (c) the covariance function of a random harmonic oscillation, (d) the covariance function of white noise, (e) the spectral density a random harmonic oscillation, (f) the spectral density of white noise.

The spectral density, being the Fourier transformation of the covariance function, also relates to the dependencies within a process, it describes how much of each possible frequency a process has. The spectral density of the random harmonic oscillation, Figure 1(e), is zero except for at the frequency of the oscillation. The lack of dependencies of the white noise is translated into a spectral density that has equal power for all frequencies, Figure 1(f), as white noise can change both very slow and very fast, and everything in between.

It is important to note that the spectral density and covariance function are theoretical concepts that typically cannot be obtained from one finite realisation. Likewise, it is usually not possible to recreate a specific realisation from the spectral density or covariance function of the stochastic process, because of the random elements. However, both the spectral density and covariance function describe characteristics that are useful to know about a signal, and thus we are interested in estimating them. In this thesis, the spectral density is of most interest. For a stationary signal, a spectral density estimate, or frequency spectrum, can be defined

$$\hat{S}_{\mathrm{X}}(f) = S_x^h(f) = \left| \int x(t)h^*(t)e^{-i2\pi ft}dt \right|^2, \tag{4}$$

where $x(t)$ is the signal and $h(t)$ is a window function. The choice of window is important but for now the simple rectangular window, $h(t) = a_h$, will be used. If we assume that the signal is sampled, i.e. discrete time increments are used, and that the window is unit energy, then this spectral density estimate is the periodogram.

To obtain the frequency spectrum, the squared Fourier transform of the signal is used. It would be possible to examine only the Fourier transform of the signal, however it will be complex-valued, have negative values and contain information about phase, which would not make for an intuitive spectral density estimate. By using the squared absolute value of the Fourier transform, the estimate will show what frequencies, and how much of them, that particular signal has. The spectrum is thus a *frequency representation* of the signal.

If a process is non-stationary, its covariance function will vary with time and must be described by two variables. We call this the instantaneous correlation function $r_{\mathrm{X}}(t, \tau)$, the Fourier transform of this and thus the non-stationary equivalent of the spectral density might be called the Wigner spectrum. This Wigner spectrum has some properties that most likely deviates from our expectations coming from the frequency spectrum, this will be discussed in Section 4.4 and instead we will introduce the short-time Fourier transform of a signal

$$F_x^h(t, f) = \int x(s)h^*(s - t)e^{-i2\pi fs}ds. \tag{5}$$

This makes it possible to estimate several frequency spectra by dividing the signal into shorter segments that are determined by the window function. The squared

absolute value of the short-time Fourier transform gives the *spectrogram*

$$S_x^h(t,f) = \left| \int x(s)h^*(s-t)e^{-i2\pi fs}ds \right|^2. \tag{6}$$

The choice of window function is again important, and it is often best not to choose a rectangular window for the spectrogram. Not only the shape of the window function is important now however, it is also important to chose a suitable length. This will be discussed in Section 4.3, and for now we will assume that a suitable window function is used.

Much like the frequency spectrum, the spectrogram will describe what frequencies the signal has, but it will also be possible to see how these change with time as the spectrogram is a function of both time and frequency. This type of joint *time-frequency (TF) representation* is very intuitive as it will be real-valued and positive, it can be interpreted as describing the amount of each frequency at any given time of the signal (if such an interpretation can be accepted). While the short-time Fourier transform, like the Fourier transform, is injective, the spectrogram is not and some information about the signal is lost. It should be noted that several invertible transforms used for spectral analysis do exist, including wavelet based methods [1–3], the empirical mode decomposition [4], and analysis based on the Fourier transform [5, 6].

To demonstrate the performance of the frequency spectrum and the spectrogram, let us again consider a harmonic oscillation

$$x_1(t) = a_1 \cos\left(2\pi f_0 t\right), \ 0 \leq t < 250, \tag{7}$$

with some amplitude $a_1$ and frequency $f_0 = 0.1$, that has been observed during $0 \leq t < 250$. The amplitude and frequency of the oscillations are easy to observe even in its time representation, shown in Figure 2(a). If we multiply this oscillation with an envelope function we can get an *oscillating transient*

$$x_2(t) = a_2 \cos\left(2\pi f_0 t\right) e^{-\frac{(t-t_0)^2}{2\sigma^2}}, \ 0 \leq t < 250, \tag{8}$$

with some amplitude $a_2 \neq a_1$ and a time centre of the envelope $t_0 = 125$. The frequency of the oscillation is the same, $f_0 = 0.1$, and the observation time is also unchanged. The envelope has two parameters, $t_0$ is the time where the envelope

has its centre and $\sigma$ determines the scaling, i.e. length of the envelope. This non-stationary signal is shown in Figure 2(b).

The frequency spectrum of the harmonic oscillation, Figure 2(c), deviates from the spectral density of such a process in two ways. First, the peak is not perfectly narrow, instead it has some slight width and some smaller peaks beside the large peak. This is due to the signal being observed for a finite amount of time, thus there will be some uncertainty in the frequency information. Another interpretation is that we measure the frequencies of the signal starting and stopping to oscillate. Secondly, the signal is described to have finite energy, seen as the peak almost reaching 1. This is only true if the signal exists for a finite amount of time and is not assumed to continue to oscillate indefinitely, otherwise it would have infinite energy. Because of this, it is sometimes preferred to talk about power instead of energy, but in this thesis all simulated and measured signals are assumed to have finite energy.

The transient will not have a theoretical spectral density, as it is a non-stationary signal. Its frequency representation in Figure 2(d) shows a bell shape centred at the frequency of the oscillation. The bell shape comes from the frequency representation of the envelope function, and the centre frequency is due to the harmonic oscillation. This representation might be desirable, but in some applications it might be more interesting for a frequency spectrum to show only one peak at the oscillation frequency, and then the width of the bell shaped peak can be interpreted as uncertainty.

The spectrograms of the harmonic oscillation and transient, their 2D projections visualised in Figures 2(e)–(f) respectively, shows the distribution of signal energy over time and frequency. In the figures, light yellow represents a high energy concentration and dark blue a low concentration. The spectrogram of the harmonic oscillation shows uncertainty in frequency much like the frequency spectrum, as the energy is not only located exactly along $f_0 = 0.1$. The spectrogram also shows some uncertainty in time, seen as the low energy concentration at the start and end. The spectrogram of the transient is a circle with increasing energy concentration towards its centre, this is just the 3D bell shape as seen from above (compared to the side view in the frequency spectrum). The centre of the circle is the intersection of the time centre of the envelope $t_0 = 125$ and the frequency of the oscillation $f_0 = 0.1$.

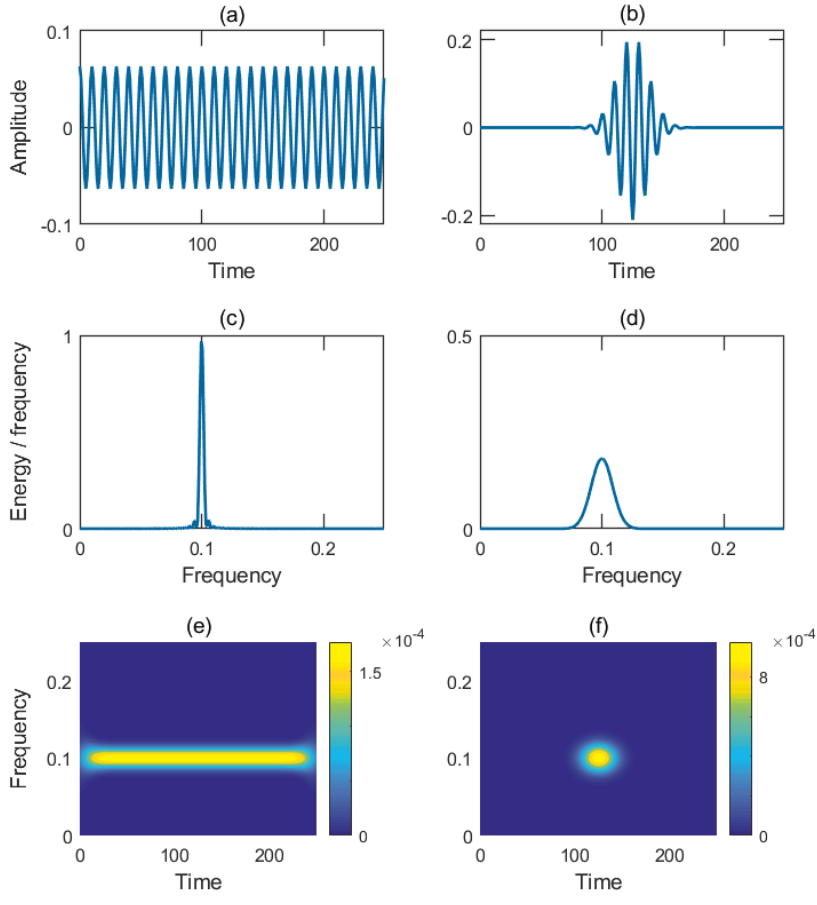



**Figure 2:** (a) The harmonic oscillation in Eq. (7), (b) the transient in Eq. (8), (c) frequency spectrum of the harmonic oscillation in Eq. (7), (d) frequency spectrum of the transient in Eq. (8), (e) spectrogram of the harmonic oscillation in Eq. (7), (f) spectrogram of the transient in Eq. (8).

Spectral analysis can thus be used as a tool for extracting interesting information that is not easily found in the time domain. While it is easy to find good frequency and TF representations of the harmonic oscillation and the transient, these simple examples have also highlighted some of the difficulties with spectral analysis. We can see that even for the most simple signal, a harmonic oscillation, the spectral density estimate is not without bias. When the signal instead is a

sum of several harmonic oscillations and includes some disturbance, the presented frequency spectrum might not be a good enough spectral density estimate. The spectrogram is generally regarded as a good TF representation, it is fast to compute and easy to interpret, but its greatest weakness is the relatively high uncertainty in time and frequency. For a signal with two transients that have similar oscillation frequencies and time centres, the spectrogram will not show two high energy circles but instead one single, more irregular, high energy shape. Because different applications have very different needs, there exist many different frequency and TF representations, with current research still improving on these methods to find representations that are faster, work for complex signals structures, and are more robust to disturbances.

## 2 Stochastic signals

The definition of a signal can be ambiguous, while we made the distinction between the stochastic process and its realisations in the previous section, we will let a signal be either a theoretical model or measured/simulated data. It is sometimes customary to separate signal from noise, defining a signal as some ordered waveform and noise as something disordered, this thesis will not make such a separation. Instead, ordered waveforms within a possibly noisy signal will be referred to as signal components, three such ordered waveforms of special interest are the stationary oscillation

$$x_{osc}(t) = e^{i2\pi f_k t}, \quad t, f_k \in \mathbb{R}, \tag{9}$$

the non-stationary FM chirp

$$x_{chirp}(t) = e^{i2\pi f_k(t) t}, \quad t, f_k(t) \in \mathbb{R}, \tag{10}$$

and the non-stationary oscillating transient

$$x_G(t) = e^{-\frac{(t-t_k)^2}{2\sigma^2}} e^{i2\pi f_k t}, \quad t, t_k, f_k \in \mathbb{R},\ \sigma \in \mathbb{R}_{>0}. \tag{11}$$

These signals, being complex-valued, defined for all real values of time and frequency and expressed with simple mathematical expressions, are obviously theoretical models. Linear combinations of these components and noise are assumed to adequately model signals that we can measure in the real world, and

simulations from these models are used to assess the performance of proposed spectral analysis methods. While the signals Eqs. (9)–(11) are expressed in their complex-valued form, their real-valued counterparts are also of interest.

The spectral analysis methods presented in this thesis will all require some assumptions on the signals for the respective methods to work well. Such assumptions can be stationarity or the signal being a linear combination of only one or two of the signal component types Eqs. (9)–(11). The methods will not assume specific frequencies, $f_k$ and $f_k(t)$, or time centres, $t_k$, but these will be kept deterministic when simulating signals, as this gives the ability to construct difficult situations for which the methods can be evaluated. The stochastic parameters are instead the phase shift of an oscillation

$$e^{i2\pi\phi_k}, \quad \phi_k \in \mathcal{U}(0,1), \tag{12}$$

and the added noise, usually white Gaussian noise, which is best described in discrete time

$$\epsilon_w(n) \in \mathcal{N}(0,\sigma_w^2), \quad n \in \mathbb{Z}, \ \sigma_w \in \mathbb{R}_{>0}. \tag{13}$$

Complex-valued white Gaussian noise is constructed by simulating from the normal distribution and multiplying with the imaginary unit. Letting the phase shift and noise be stochastic parameters allows us to test if the spectral analysis methods are robust to different phase shifts and noise realisations. As the phase is not shown in the frequency or TF representations and the noise is random disturbance, it is more difficult to do controlled evaluations of their effect on the spectral analysis.

## 3 Fundamental ideas and limitations of spectral analysis

There are numerous frequency and joint TF representations, all designed for different signals and applications. However, there are some fundamental ideas and limitations that are interesting to consider for a general representation. To define some of these properties we let $x(t)$ be the signal in time domain and $X(f)$ the signal in frequency domain, i.e. the Fourier transform of the signal, according to Eqs. (1)–(2). The ideal frequency representation is denoted $S_x(f)$, interpreted as the energy per unit frequency at frequency $f$, and the ideal TF representation is denoted $P_x(t,f)$, interpreted as the energy per unit time and frequency at time $t$ and frequency $f$.

### 3.1 Shifts, scaling and support

To enable analysis of a wide variety of signals, we want changes in parameters of the signal to result in changes that correspond to those parameters in the frequency and TF representations. For the Fourier transform of a signal it holds that

$$\begin{aligned} X(f-f_0) &= \int x(t)e^{i2\pi f_0 t}e^{-i2\pi f t}dt, \\ x(t-t_0) &= \int X(f)e^{-i2\pi t_0 f}e^{i2\pi f t}df. \end{aligned} \tag{14}$$

For a frequency representation, such time and frequency shift invariance gives

$$X(f) \to X(f-f_0) \Longrightarrow S_x(f) \to S_x(f-f_0), \tag{15}$$

and for a TF representation

$$\begin{aligned} x(t) \to x(t-t_0) &\Longrightarrow P_x(t,f) \to P_x(t-t_0,f), \\ X(f) \to X(f-f_0) &\Longrightarrow P_x(t,f) \to P_x(t,f-f_0), \\ x(t) \to e^{i2\pi f_0 t}x(t-t_0) &\Longrightarrow P_x(t,f) \to P_x(t-t_0,f-f_0). \end{aligned} \tag{16}$$

Thus, the shape of the energy distribution in the representations should ideally not be changed when a signal is shifted in time and/or frequency.

Linear scaling of a signal also scales the Fourier transform

$$x(t) \to x(\alpha t) \Longrightarrow X(f) \to \frac{1}{\alpha}X(f/\alpha), \quad \alpha \in \mathbb{R}_{>0}, \tag{17}$$

i.e. a signal that is compressed in time will have a Fourier transform that is expanded, and vice versa. Thus we want a frequency representation to satisfy

$$x(t) \to x(\alpha t) \Longrightarrow S_x(f) \to \beta S_x(f/\alpha), \tag{18}$$

and a TF representation to satisfy

$$x(t) \to x(\alpha t) \Longrightarrow P_x(t,f) \to \beta P_x(\alpha t, f/\alpha), \tag{19}$$

where the factor $\beta$ could be $1/\alpha^2$ but will depend on representation. The more interesting property is that the linear scaling in time and frequency shows in the representations.

Considering a signal that has energy in some time duration and frequency bandwidth, it seems natural that we want the frequency and TF representations to be zero outside that time duration and frequency bandwidth. If we let the time duration be $(t_1, t_2)$ and frequency bandwidth be $(f_1, f_2)$ then we define weak finite support

$$X(f) = 0, \forall f \notin (f_1, f_2) \Longrightarrow S_x(f) = 0, \forall f \notin (f_1, f_2), \tag{20}$$

for a frequency representation and

$$\begin{aligned} x(t) = 0, \forall t \notin (t_1, t_2) &\Longrightarrow P_x(t, f) = 0, \forall t \notin (t_1, t_2), \\ X(f) = 0, \forall f \notin (f_1, f_2) &\Longrightarrow P_x(t, f) = 0, \forall f \notin (f_1, f_2)., \end{aligned} \tag{21}$$

for a TF representation. If we want the representations to be zero when the signal is zero, then we must define strong finite support. Strong finite support is fulfilled for a frequency representation if it for any given $f_1$ holds that

$$X(f_1) = 0 \Longrightarrow S_x(f_1) = 0, \tag{22}$$

and for a TF representation if it for any given $t_1$ and $f_1$ holds that

$$\begin{aligned} x(t_1) = 0 &\Longrightarrow P_x(t_1, f) = 0, \\ X(f_1) = 0 &\Longrightarrow P_x(t, f_1) = 0. \end{aligned} \tag{23}$$

Strong finite support implies weak finite support, but not the other way around. Many popular TF representations do not have strong finite support because it requires other sacrifices, it is however helpful when analysing multi-component signals, as strong finite support makes it easier to identify individual components.

## 3.2 Total energy and marginals

The total energy of a signal is

$$E_{tot} = \int |x(t)|^2 dt = \int |X(f)|^2 df, \tag{24}$$

which is known as Parseval's relation. Thus, $|x(t)|^2$ and $|X(f)|^2$ can be interpreted as the energy per unit time at time $t$ and the energy per unit frequency at frequency $f$ respectively, making them time and frequency

representations of the signal. Assuming that $P_x(t,f)$ indeed is a distribution of the signal energy over time and frequency, this gives the idea that it is possible to get two so called marginals

$$\begin{aligned} P_x(t) &= \int P_x(t,f)df, \\ P_x(f) &= \int P_x(t,f)dt. \end{aligned} \tag{25}$$

We chose to make a distinction between the notation of $P_x(f)$ and $S_x(f)$ because typically a frequency representation is not calculated from the marginal of a TF representation.

Ideally the marginals should then accurately describe the energy per unit time and per unit frequency respectively. This would mean that

$$\begin{aligned} P_x(t) &= \int P_x(t,f)df \quad &= |x(t)|^2, \\ P_x(f) &= \int P_x(t,f)dt \quad &= |X(f)|^2. \end{aligned} \tag{26}$$

A TF representation for which this holds is said to fulfil the marginal conditions. However, it is possible for a TF representation to not fulfil this but still satisfy the total energy requirement

$$E_{tot} = \iint P_x(t,f)dfdt. \tag{27}$$

There exist no marginal conditions for frequency representation, but the total energy requirement would be

$$E_{tot} = \int S_x(f)df. \tag{28}$$

It is worth noting that many popular TF representations do not fulfil the marginal conditions or the total energy requirement, and while it is important to be aware of this, they can still be good analysis methods for non-stationary and multi-component signals. It is also possible to calculate the marginals for all TF representations, but the marginals might not have the expected properties.

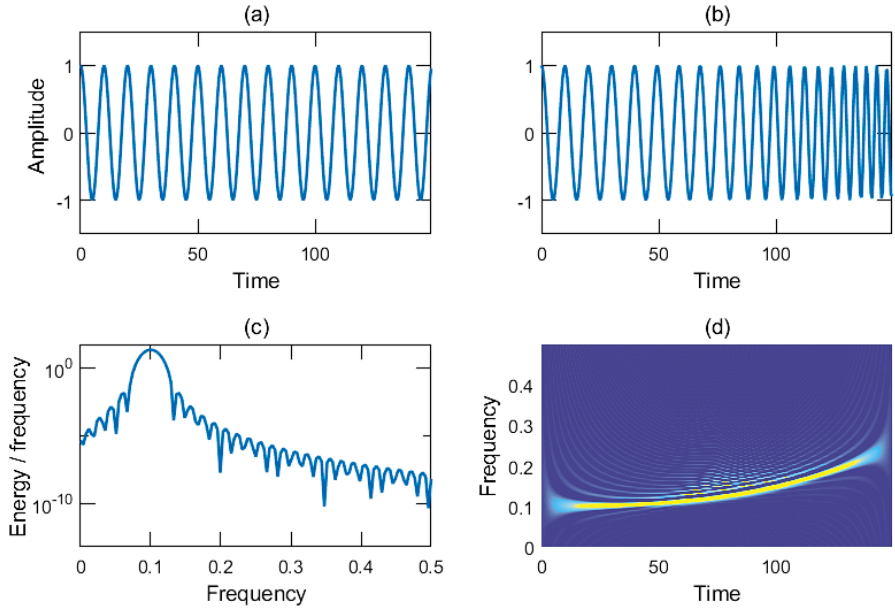



**Figure 3:** (a) A harmonic oscillation, (b) a chirp, (c) a frequency representation of the harmonic oscillation with dB scale on the y-axis, (d) a TF representation of the chirp.

## 3.3 Bias and uncertainty

A signal cannot have both finite duration and limited bandwidth. Thus, no simulated or measured signals will be band-limited, and this affects any frequency and TF representation. Figure 3 gives examples of this effect for a stationary harmonic oscillation with frequency 0.1, shown in (a), and a non-stationary chirp with frequencies between 0.1 and 0.25, shown in (b). The frequency representation of the harmonic oscillation in Figure 3(c) shows a peak or main lobe at frequency 0.1, but also multiple smaller oscillations or side lobes for the other frequencies. The plot is shown in dB scale on the y-axis to emphasise the side lobes, their amplitudes are very small, but the frequency representation is never zero for this signal. The same type of side lobes can be seen in the TF representation of the chirp in Figure 3(d) as the smaller ridges in yellow and light blue. The effect of the side lobes is not seen as clearly in this plot because the amplitude is scaled linearly.

The side lobes are also referred to as spectral leakage, referring to how signal energy leaks from the frequency of an oscillation to all other frequencies. How large this bias will be is decided by the choice of method to calculate the frequency or TF representation. However, assuming that the duration of the signal is kept constant and no additional information is used, decreasing the amplitude of the side lobes will increase the width of the main lobe, and thus lower the resolution of the representation. Low resolution is a problem especially for multi-component signals, as the main lobes of close components will merge together. High spectral leakage is however also a problem for multi-component signals, as components with smaller amplitudes will not be visible. Therefore, resolution and spectral leakage always need to be balanced.

For stationary signals, the width of the main lobe can also be decreased by having a longer signal. This is because increasing the time duration of a signal with constant frequencies is the same as increasing the energy spread in time domain, and if the energy spread increases in the time domain it decreases in the frequency domain. This relationship is perhaps most interesting for non-stationary signals, which can have components that are shorter than the signal.

The relationship between the energy spread in the time respective frequency domain comes from that $x(t)$ and $X(f)$ are Fourier transform pairs, and is visualised for some Gaussian functions in Figure 4. The total energy of a signal is given by Eq. (24), and thus $|x(t)|^2$ and $|X(f)|^2$ are distributions of the signal energy. Considering $|x(t)|^2$ and $|X(f)|^2$ as distributions means that we can calculate the variances

$$\begin{aligned} \Delta t^2 &= \int \left(t - \mu_t\right)^2 |x(t)|^2 dt, \\ \Delta f^2 &= \int \left(f - \mu_f\right)^2 |X(f)|^2 df, \end{aligned} \tag{29}$$

where $\mu_t$ and $\mu_f$ are the expected values. It is now interesting to consider the Heisenberg uncertainty principle for energy and time

$$\Delta E \Delta t \geq \frac{\hbar}{2}, \tag{30}$$

where $\Delta E$ is the uncertainty in energy and $\hbar$ is the reduced Planck's constant. Then, from the relationship of the energy of a photon and the frequency of that light

$$\Delta E = 2\pi\hbar\Delta f, \tag{31}$$

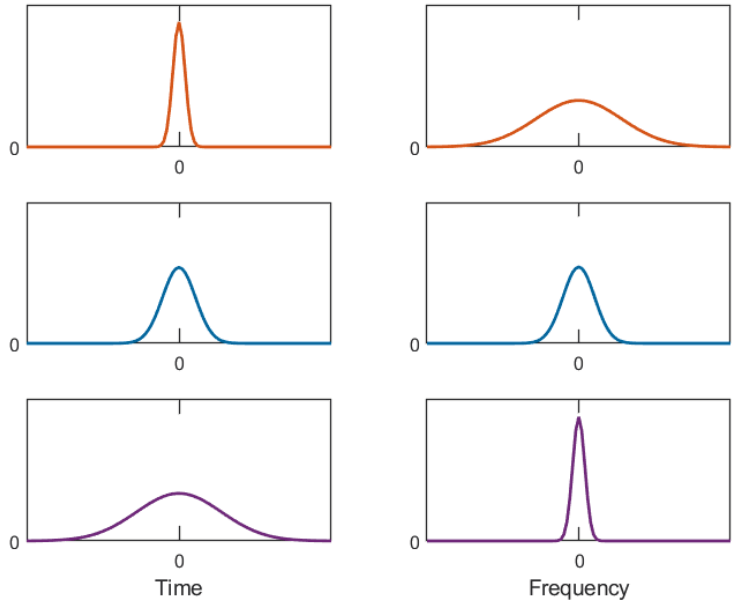



**Figure 4:** Illustration of the energy spread in the time and frequency domains, using three Gaussian functions and their Fourier transforms, which are also Gaussian functions.

we get the uncertainty principle for time and frequency

$$\Delta t \Delta f \geq \frac{1}{4\pi}, \tag{32}$$

which holds when using the Fourier transform. The uncertainty principle can also be calculated using the Cauchy–Schwarz inequality [5]. According to the uncertainty principle, both standard deviations cannot be arbitrarily small, and this limits the TF resolution. The lower bound of the inequality is reached by unit energy Gaussian waveforms

$$x_G(t) = \frac{1}{\sqrt{\sigma\sqrt{\pi}}} e^{-\frac{(t-t_0)^2}{2\sigma^2}} e^{i2\pi(f_0 t+\phi)}, \tag{33}$$

indicating that they have an optimal energy spread in time and frequency, which made Gabor call this waveform the elementary signal [5].

### 3.4 The instantaneous frequency

A complex-valued signal with one component can be written

$$x(t) = A(t)e^{i\phi(t)}, \tag{34}$$

where $A(t)$ is the amplitude function and $\phi(t)$ the phase function. If we let $\phi(t) = 2\pi f(t)t$, the signal is a general AM-FM chirp, from which all the signals described by Eqs. (9)–(11) can be formed. The frequency of this chirp will, barring a few exceptions, vary over time, it is thus interesting to know the frequency at any given time. That is, we want to know the instantaneous frequency (IF).

To derive a mathematical expression of the IF, we consider the expected value of the frequency for the chirp in Eq. (34)

$$\begin{aligned} \mathrm{E}[f] = \int f\,|X(f)|^2 df &= \iiint f\; x^*(t)x(s)\; e^{i2\pi(t-s)f}\, df\, ds\, dt \\ &= \frac{1}{i2\pi} \iiint x^*(t)x(s) \frac{\partial}{\partial t} e^{i2\pi(t-s)f}\, df\, ds\, dt. \end{aligned} \tag{35}$$

Now, since

$$\delta(t-s) = \int e^{i2\pi(t-s)f}\, df, \tag{36}$$

where $\delta(t)$ is the Dirac impulse, we can simplify

$$\mathrm{E}[f] = \frac{1}{i2\pi} \iint x^*(t)x(s) \frac{\partial}{\partial t} \delta(t-s)\, ds\, dt. \tag{37}$$

To further simplify, we can use

$$\int x(s) \frac{\partial}{\partial t} \delta(t-s)\; ds = \frac{d}{dt} x(t), \tag{38}$$

and thus, using Eq. (34)

$$\begin{aligned} \mathrm{E}[f] = \frac{1}{i2\pi} \int x^*(t) \frac{d}{dt} x(t)\, dt &= \frac{1}{2\pi} \int x^*(t) \left( \phi'(t) - i\frac{A'(t)}{A(t)} \right) x(t)\, dt \\ &= \frac{1}{2\pi} \int |x(t)|^2 \left( \phi'(t) - i\frac{A'(t)}{A(t)} \right) dt. \end{aligned} \tag{39}$$

If we assume that $f(t)$ is real-valued, and we should, it is easy to see that the second term has to be zero, as that term is imaginary. This means that

$$\mathrm{E}[f] = \frac{1}{2\pi} \int \phi'(t) |x(t)|^2 dt. \tag{40}$$

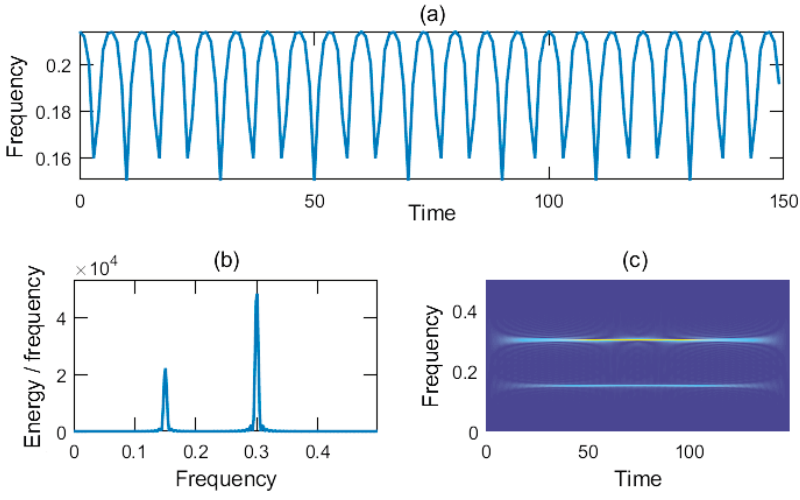



**Figure 5:** Illustration of the IF of the two-component signal Eq. (42), (a) the IF calculated from the derivative of the phase function, (b) and (c) intuitive representations of the IF.

Thus, the expected value of the frequency is given by integrating, over time, the product of the derivative of the phase function and the energy density. The derivative of the phase function must therefore be the instantaneous value of the frequency, i.e. the IF is

$$f(t) = \frac{\phi'(t)}{2\pi}. \tag{41}$$

However, this result only holds for complex-valued signals with one component. To fix one of these problems we can define a complex-valued signal for every real-valued signal, this is done by utilising a property in the frequency domain of complex-valued signals, which will be explained further in Section 4.4. It is more challenging to calculate the IF of multi-component signals. Figure 5 illustrates this for the simple two-component signal

$$x_2(t) = e^{i2\pi 0.15t} + 1.5e^{i2\pi 0.3t}. \tag{42}$$

Figure 5(a) shows the derivative of the phase function of the signal, and it does not show the two frequencies that we expect, instead the IF oscillates between 0.15 and 0.3. According to the intuitive understanding of the IF, Figure 5(b) and (c) would be much better representations, and thus a frequency respective TF representation that we would like to obtain for the signal.

# 4 Well-known spectral representations

In order to calculate frequency and TF representations, one needs to consider sampled signals. That is signals that are not a function of the continuous time $t \in \mathbb{R}$, but the discrete time $n \in \mathbb{Z}$. When sampling a signal, the sampling frequency, $f_s$, becomes important as it defines the range of frequencies the sampled signal can have. However, in this thesis it will always be assumed that the sampling frequency is sufficiently high and if nothing else is stated it is assumed that $f_s = 1$, so that $-1/2 < f \leq 1/2$.

Because of different traditions some spectral representations are defined in discrete time and others in continuous time, this section will therefore mix continuous and discrete time. Any application of the discussed methods of course requires a discrete implementation, which typically involves the discrete or fast Fourier transform. The discrete Fourier transform is defined

$$X(f) = \sum_{n=0}^{N-1} x\left(\frac{n}{f_s}\right) e^{-i2\pi f \frac{n}{f_s}}, \quad f = \frac{kf_s}{N},\ k = 0, 1, \ldots, N-1, \tag{43}$$

and the fast Fourier transform is an implementation that allows for faster calculations, $\mathcal{O}(n \log n)$. With this formulation only positive frequencies, $0 \leq f < f_s$, are obtained, however for $f_s/2 < f < f_s$ we get the same values as we would have got for $-f_s/2 < f < 0$. Thus, the negative frequencies can be obtained by appropriately shifting $X(f)$ for the higher values of $f$.

## 4.1 The periodogram

The periodogram was given its name by Arthur Schuster already in 1898 [7], and it has been frequently used since the modern invention of the fast Fourier transform by James W. Cooley and John W. Tukey in 1965 [8]. The windowed periodogram is defined as

$$S_P(f) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x(n) h(n) e^{-i2\pi f n} \right|^2, \tag{44}$$

thus from the discrete-time Fourier transform for a finite number of samples, $N$, with the signal, $x(n)$, and a window function, $h(n)$. Classically $h(n)$ is the rectangular window, otherwise $S_P(f)$ is also called the modified periodogram. The choice of window function affects the resolution and spectral leakage of the

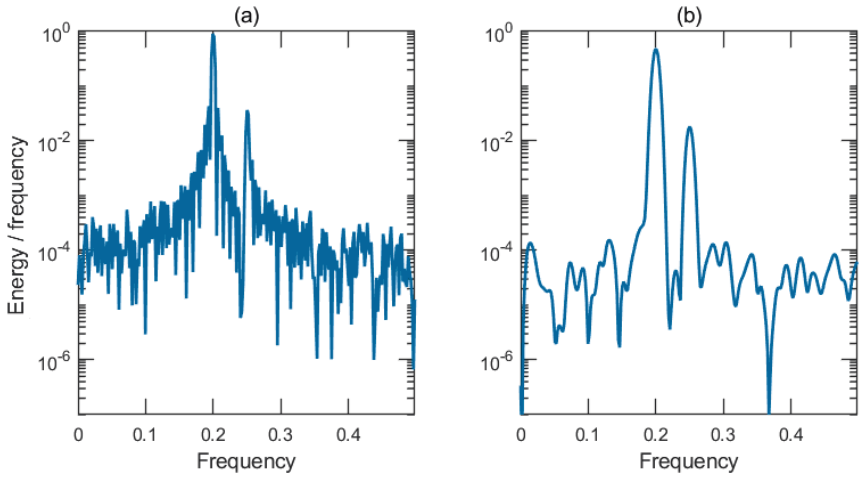



**Figure 6:** Illustration of the periodogram of a two-component signal with white Gaussian noise, (a) calculated with a rectangular window, (b) calculated with a Slepian window.

periodogram, which can be seen for two examples, the rectangular window and a Slepian window, in Figure 6. The rectangular window in Figure 6(a) gives narrow main lobes for the components, but large spectral leakage is seen in the high energy around the two peaks, this contrast with the Slepian window in Figure 6(b) where the main lobes are wider but the spectral leakage is considerably lower. This trade-off between resolution and spectral leakage is not unique for the periodogram and makes the choice of window function for the Fourier transform very important.

If we assume that $x(n)$ is stationary, real-valued and has a zero mean, then if a rectangular window is used, we can reformulate the periodogram

$$\begin{aligned} S_P(f) &= \frac{1}{N}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} x(n)x(m)e^{-i2\pi f(m-n)} \\ &= \sum_{\tau=-(N-1)}^{N-1} \frac{1}{N}\sum_{n=0}^{N-1-|\tau|} x(n)x(n+|\tau|)e^{-i2\pi f\tau} \\ &= \sum_{\tau=-(N-1)}^{N-1} \hat{r}_{\mathrm{X}}(\tau)e^{-i2\pi f\tau}, \end{aligned} \tag{45}$$

where $\hat{r}_{\mathrm{X}}(\tau)$ is an estimate of the covariance function of the process from which $x(n)$ is a realisation. Thus, the expected value

$$\begin{aligned} \mathrm{E}[S_P(f)] &= \sum_{\tau=-(N-1)}^{N-1} \mathrm{E}[\hat{r}_{\mathrm{X}}(\tau)] e^{-i2\pi f\tau} \\ &\to \sum_{\tau=-\infty}^{\infty} r_{\mathrm{X}}(\tau) e^{-i2\pi f\tau} = S_{\mathrm{X}}(f), \text{ if } N \to \infty, \end{aligned} \tag{46}$$

where $S_{\mathrm{X}}(f)$ is the spectral density of the process from which $x(n)$ is a realisation. This is a great result, however unfortunately very long signals are usually not at hand, and therefore it is interesting to also consider the variance of the periodogram. If we let $x(n)$ be white Gaussian noise, then it can be shown that

$$\mathrm{Var}[S_P(f)] \approx \sigma_w^4, \tag{47}$$

where $\sigma_w^2$ is the variance of the noise [9]. This shows not only that the variance is large, but also that it does not decrease if $N \to \infty$, which means that the periodogram is not a consistent estimator of the spectral density. Therefore, the periodogram should only be used for signals with high SNR and preferably only signals with many samples. If the SNR is low, other methods, more robust to noise, should be considered.

### 4.2 The Welch method

One way to reduce the variance of the periodogram is to average several periodograms, which was first suggested by Bartlett in 1948 [10]. However, in order to average spectral estimates in this way, the signal has to be divided into shorter segments, which reduces the resolution of the final estimate. This result follows from the same reasoning leading to the uncertainty principle in Eq. (32), though for time-limited and discrete signals. To mitigate this loss in resolution, Peter D. Welch proposed the weighted overlap segmented averaging algorithm in 1967 [11], which is now commonly referred to as the Welch method.

If we let the signal $x(n)$ be divided into segments that are $N_b$ samples long, and let $N_b - L$ be the overlap in samples, then the Welch method can be defined

$$S_W(f) = \frac{1}{K} \sum_{k=1}^{K} \left| \sum_{n=0}^{N_b-1} x(n + (k-1)L) h(n) e^{-i2\pi f n} \right|^2, \tag{48}$$

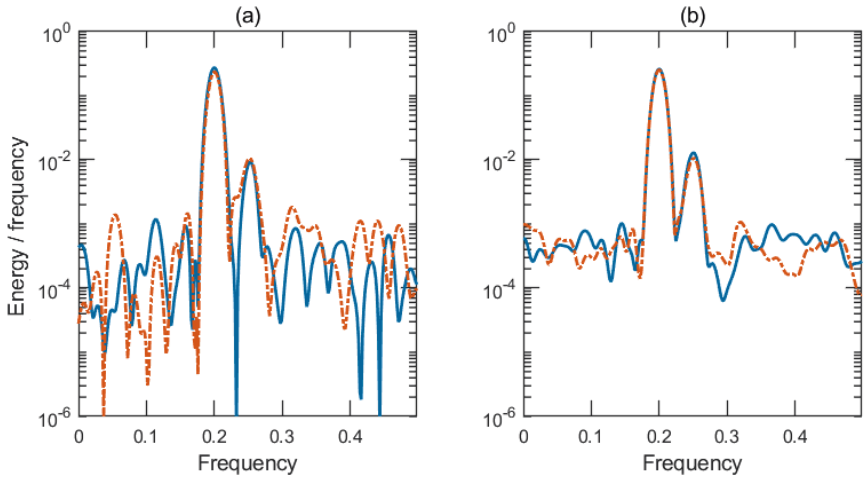



**Figure 7:** Illustration the Welch method of two signals with the same two oscillating components but different realisations of white Gaussian noise, (a) no overlap and $K = 3$, (b) 50% overlap and $K = 5$. All estimations are calculated using a Slepian window of length $N_b = N/3$.

where $K$ is the number of windows. If we assume that $x(n)$ is white Gaussian noise with variance $\sigma_w^2$, then the variance of the Welch method can be calculated

$$\mathrm{Var}[S_W(f)] = \frac{1}{K^2} \sum_{j=1}^{K} \sum_{k=1}^{K} \mathrm{Cov}[S_j(f), S_k(f)], \tag{49}$$

where $S_j(f)$ and $S_k(f)$ are the (modified) periodograms calculated from the $N_b$ long segments of $x(n)$. If $S_j(f)$ and $S_k(f)$ are uncorrelated when $j \neq k$, then

$$\mathrm{Var}[S_W(f)] = \frac{1}{K^2} \sum_{k=1}^{K} \mathrm{Var}[S_k(f)] \approx \frac{\sigma_w^4}{K}, \tag{50}$$

and thus, compared to the periodogram, the variance is approximately reduced by a factor equal to the number of windows. Figure 7 shows the Welch spectral estimate for two signals with the same components but different white Gaussian noise, one in blue and the other in orange. The two estimations in Figure 7(b), where $K = 5$, are more similar compared to the estimations in Figure 7(a), where $K = 3$, illustrating how the increase in $K$ lowers the variance introduced by the noise.

Because of the nature of white Gaussian noise, if there is no overlap between the signal segments, the spectral estimates, $S_j(f)$ and $S_k(f)$ $j \neq k$, will be uncorrelated. However, if the segments overlap, then it cannot be certain that the estimates are uncorrelated. Even so, we want sufficient overlap to reduce the loss of resolution. If the length of the segments are shortened from $N$ to $N_b = N/M$, and the same window function is used, then the main lobe width increases with at factor $M$. It is therefore important to balance the trade-off in variance reduction and resolution loss. Figure 7 shows that if we allow for 50% overlap, compared to no overlap, then we can increase the number of windows significantly and reduce variance, but the resolution is kept the same. The optimal overlap will depend on the window function, as the window function will determine the weight of each sample in $x(n)$ and thus affect the covariance between the spectral estimates of overlapping sections.

### 4.3 The spectrogram

One of the most well-known and widely used TF representations is the spectrogram, though the name might differ in some fields, e.g. the sonogram, spectrograph or waterfall plot. In 1946, Dennis Gabor argued for the necessity in mathematics to describe signals as a joint function of time and frequency, instead of only considering frequency in an infinite interval of time like the Fourier transform. Through this reasoning that frequencies must be allowed to change with time, not only in reality but also in mathematics, he presented the Gabor transform [5], which today is considered a special case of the short-time Fourier transform (STFT) [12]. The STFT uses a window function $h(t)$ to divide the signal $x(t)$ (real- or complex-valued) into smaller segments, for which the Fourier transform is calculated

$$F_x^h(t,f) = \int x(s)h^*(s-t)e^{-i2\pi fs}ds. \tag{51}$$

This is not unlike the approach of the Welch method, however instead of averaging Fourier transforms they should be thought of as stacked to create a time dimension. However, unlike the Welch method, the overlap is generally very large. In this thesis a sliding Fourier transform is always used, i.e. maximum overlap is used and there is no downsampling in time. The spectrogram is defined as

$$S_x^h(t,f) = \left|F_x^h(t,f)\right|^2, \tag{52}$$

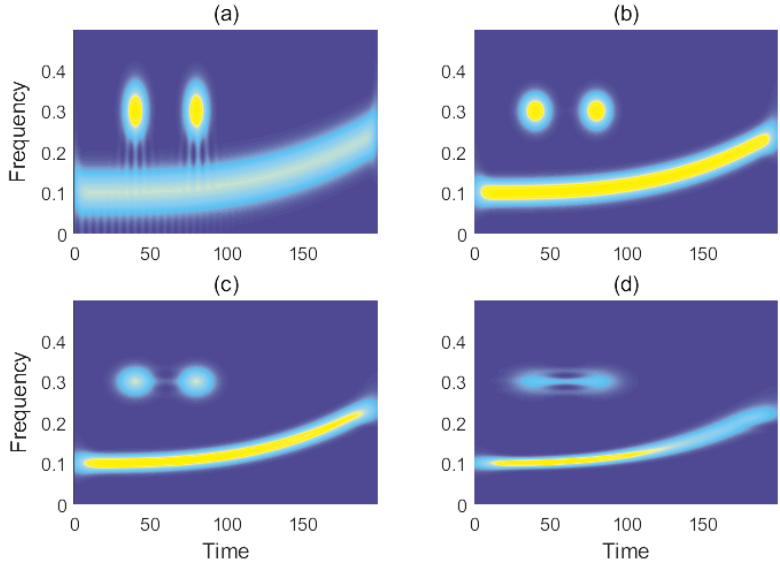



**Figure 8:** The spectrogram of a multi-component signal, with increasing length of the window function so that (a) has the shortest window and (d) the longest.

which makes it a non-negative distribution of the signal energy. For the Gabor transform, a Gaussian window function is used, which is why the spectrogram with a Gaussian window is also called the Gabor spectrogram.

The spectrogram has two major advantages. First, it is very easy to interpret, there are even people who can "read" the spectrogram of speech [13], and the spectrogram of audio signals is closely linked to how the cochlea of the inner ear of humans encode audio [14]. The spectrogram of multi-component signals has little interaction, i.e. artefacts, between signal components, thus almost achieving strong finite support. Second, it has a computationally efficient implementation due to the fast Fourier transform.

The main drawback of the spectrogram is the resolution. There is a trade-off between the resolution in time and frequency, which is determined by the length of the window function. This is demonstrated by Figure 8, where the window length is increased from (a) to (d). The shortest window in (a) gives very poor frequency resolution, in (b) the window length is optimal for the two transient components, the change of the frequencies in the chirp is perhaps best seen in (c)

but the transients are smeared in time, and in (d) the time resolution is very poor. Similar to the previously discussed frequency representations, the shape of the window function also affect the width of the main lobe and the height of the side lobes. The choice of window function can therefore determine if close components are resolved and if weak components are hidden by the spectral leakage of stronger components.

### 4.4 The Wigner distributions

In 1932, Eugene Wigner presented what we today call the Wigner distribution (WD). Originally it was presented in the field of quantum mechanics as a joint distribution of time and momentum [15]. This was not the only joint distribution developed around this time, but since the 1980's it has been the most influential in spectral analysis and signal processing [16–19]. The WD relates to the instantaneous correlation function $r_{\mathrm{X}}(t,\tau)$ of a non-stationary processes $\mathrm{X}(t)$

$$r_{\mathrm{X}}(t,\tau) = \mathrm{E}\left[\mathrm{X}\left(t+\tfrac{\tau}{2}\right)\mathrm{X}^{*}\left(t-\tfrac{\tau}{2}\right)\right]. \tag{53}$$

From the realisation $x(t)$ of that process, we can get an estimate of $r_{\mathrm{X}}(t,\tau)$

$$K_x(t,\tau) = x\left(t+\tfrac{\tau}{2}\right)x^{*}\left(t-\tfrac{\tau}{2}\right), \tag{54}$$

and the WD is the calculated as

$$W_x(t,f) = \int K_x(t,\tau)e^{-i2\pi f\tau}d\tau. \tag{55}$$

Thus, the WD is the Fourier transform of the estimated instantaneous correlation function, linking it to the spectral density estimates of stationary stochastic processes. For most mono-component signals with no noise, AM-FM chirps where the FM is linear, the WD gives exactly the IF of the signal, and the TF concentration is good. However, for multi-component signals, $K_x(t,\tau)$ will include cross-correlations between signal components, resulting in interference, called artefacts or cross-terms, between signal components in the WD. The WD is also cumbersome to compute as $K_x(t,\tau)$ is a 2D function.

The cross-terms are especially troublesome for real-valued signals, as their Fourier transform includes negative frequencies and the cross-terms will appear between the positive and negative frequency components. This can be seen in Figure 9(a)

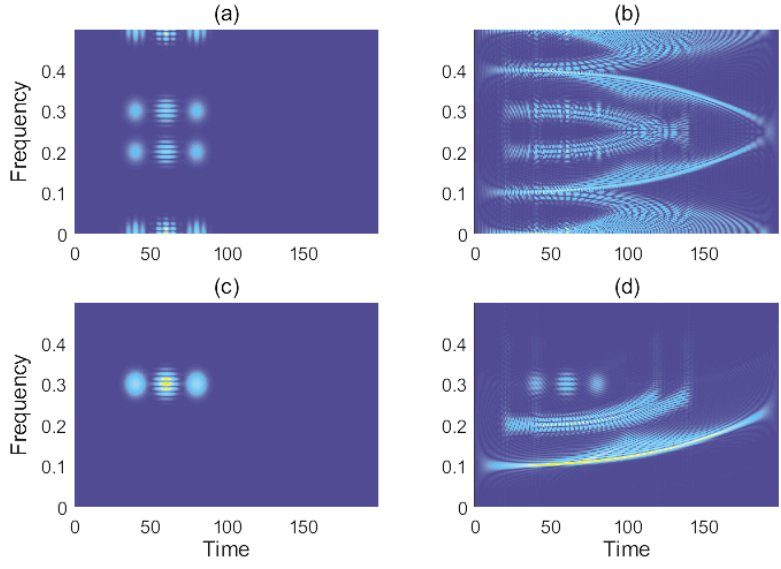



**Figure 9:** The WD and WVD of two multi-component signals, a two-component and a three-component, (a) the WD of the two-component signal (b) the WD of the three-component signal, (c) the WVD of the two-component signal, (d) the WVD of the three-component signal.

and (b). This problem can be solved by using complex-valued signals, which do not have negative frequency components. The version of WD that requires a complex-valued signal is called the Wigner-Ville distribution (WVD), and it was Jean-André Ville who first introduced the WD in the context of signal processing in 1948 [20]. The improvement gained from using complex-valued signals can be seen in Figure 9(c) and (d).

The three-component signal in Figure 9 is the same as in Figure 8, and the two-component signal has the same transient components as that signal. Comparing the WD, WVD and spectrogram, it is clear that the WD and WVD have much better TF resolution, the energy localisation around the components is very high. However, the WVD and WD especially are not easy to interpret. The cross-terms, located midway between all signal components, can have twice the amplitude of the signal components and while it is easy to determine the cross-term in Figure 9(c), it is much harder when the number of components increase, as seen in Figure 9(d).

In order for the WVD to be useful in application, there needs to be a way to make a complex-valued signal that corresponds to a real-valued signal. Therefore, we define the analytical signal

$$z(t) = x_r(t) + x_i(t), \tag{56}$$

where $x_r(t)$ is the real-valued signal, and $x_i(t)$ is an imaginary part that needs to be defined. The Fourier transform of the real-valued signal is symmetric so that

$$X_r(f) = X_r(-f), \tag{57}$$

and thus the real-valued signal contains more information of the frequency than is needed. We can then choose to subtract that information so that

$$Z(f) = 0, \quad f < 0, \tag{58}$$

which gives

$$z(t) = 2\int_0^\infty X(f)e^{i2\pi ft}df = x(t) + i\left(\frac{1}{\pi t} * x(t)\right). \tag{59}$$

### 4.5 Smoothed Wigner-Ville distributions

The WVD is sometimes difficult to interpret, but the TF concentration is very good, it is time and frequency shift invariant, and fulfils the marginal conditions. It is therefore not surprising that many methods have been developed with the aim to suppress the cross-terms, thus improving the readability, while trying to maintain good TF concentration. In 1966, Leon Cohen set up a framework for such TF representations, known as Cohen's class, the quadratic class or the bilinear class. Cohen also showed that some already existing representations, such as the Page and Margenau-Hill distributions [21, 22], could be described in this framework [23].

The quadratic class can be described by first defining a kernel

$$\Psi(t,f) = \iint \psi(\nu,\tau)e^{-i2\pi(f\tau-\nu t)}d\tau d\nu = \int G(t,\tau)e^{-i2\pi f\tau}d\tau, \tag{60}$$

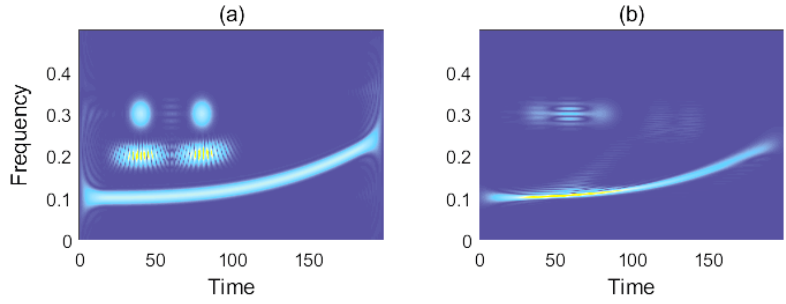



**Figure 10:** Two smoothed WVD of a multi-component signal, (a) smoothing only in frequency (b) smoothing only in time.

and then filtering the WVD with that kernel

$$\begin{aligned} W_x^Q(t,f) &= W_x(t,f) \underset{t}{*}\underset{f}{*} \Upsilon(t,f) \\ &= \iiint K_x(s,\tau)\psi(\nu,\tau)e^{-i2\pi(\nu s+f\tau-\nu t)}dsd\tau d\nu \\ &= \iint K_x(s,\tau)G(t-s,\tau)e^{-i2\pi f\tau}dsd\tau. \end{aligned} \tag{61}$$

The kernels $\Upsilon(t,f)$, $\psi(\nu,\tau)$, and $G(t,\tau)$, all give equal quadratic distributions, they are just defined on different domains, the time-frequency, ambiguity, and time-lag domain respectively. There is also a Doppler-frequency domain, and all four domains are connected with the Fourier transform. Filtering with a kernel causes smoothing in time and/or frequency, depending on the design of the kernel, which is why the resulting TF representations can be thought of as smoothed WVDs.

It is possible to design kernels to have certain properties and be adapted to specific types of signals. This area of research, to improve on the flaws of the WVD, was very popular during the 1980's and 1990's, and there are now many well-known quadratic distributions. These include the Choi-Williams [24], Rihaczek [25], Zhao-Atlas-Marks [26], signal adaptive kernels [27, 28], and many more [29, 30]. Most kernels aim at achieving good energy localisation of signal components and suppression of cross-terms, but there is a trade-off between the two, and kernels can be designed to keep some cross-term to gain better resolution. This is demonstrated in Figure 10 where the kernel in (a)

smooths in frequency, only removing cross-terms that are between components along the time axis, and the kernel in (b) smooths in time, only removing cross-terms that are between components along the frequency axis. In Figure 10(a) the time resolution is high, the transients are well separated, but there is a loss in frequency resolution, i.e. the components are wide in frequency. For Figure 10(b) the frequency resolution is high, the chirp is narrow when the frequency is only slowly increasing, but the time resolution is reduced and the transients are drawn together. A kernel that smooths only in frequency is called Doppler-independent, and a kernel that smooths only in time is called lag-independent.

The spectrogram is also part of the quadratic class, which can be shown if we rewrite the spectrogram

$$\begin{aligned} S_x^h(t,f) &= \left| \int x(s)h^*(s-t)e^{-i2\pi fs}ds \right|^2 \\ &= \iint x(s)h^*(s-t)x^*(u)h(u-t)e^{-i2\pi f(s-u)}dsdu. \end{aligned} \tag{62}$$

If we let $s = w + \tau/2$ and $u = w - \tau/2$, we can write

$$S_x^h(t,f) = \iint K_x(w,\tau)h^*\left(w+\tfrac{\tau}{2}-t\right)h\left(w-\tfrac{\tau}{2}-t\right)e^{-i2\pi f\tau}dwd\tau \tag{63}$$

and thus the time-lag kernel for the spectrogram is

$$G(t,\tau) = h^*(-t+\tfrac{\tau}{2})h(-t-\tfrac{\tau}{2}). \tag{64}$$

## 5 The reassignment method

Accurately estimating the IF of non-stationary multi-component signals is challenging, the WVD has good resolution but can be very difficult to interpret, the spectrogram is easy to interpret but the resolution is not great. In 1976 Kodera, de Villedary and Gendrin presented a technique to improve the energy localisation of the spectrogram [31]. However, this technique failed to gain interest until Francois Auger and Patrick Flandrin reintroduced it as the reassignment method in 1995 [32]. This method uses the phase information, present in the Fourier transform of a signal but discarded when calculating the

spectrogram, to move signal energy closer to the IFs of a signal. The resulting TF representation is called the reassigned spectrogram.

The reassigned spectrogram of a signal $x(t)$ is defined

$$RS_x(t,f) = \iint S_x(s,\xi)\partial\left(t-\hat{t}(s,\xi), f-\hat{f}(s,\xi)\right) dsd\xi, \tag{65}$$

where $S_x(t,f)$ is the spectrogram of the signal, $\hat{t}(t,f)$ and $\hat{f}(t,f)$ the reassignment coordinates, and $\partial(t,f)$ the 2D Dirac impulse

$$\iint f(t,f)\partial(t-t_0, f-f_0)dtdf = f(t_0,f_0). \tag{66}$$

To calculate the reassigned spectrogram, we first need to calculate the spectrogram and the reassignment coordinates. When they are known, the instantaneous signal energy for every location $(t,f)$ in $S_x(t,f)$ is mapped to new locations $(\hat{t}(t,f),\hat{f}(t,f))$. If the reassignment coordinates have been successfully calculated, the new locations will be the exact IFs of the signal. If many locations $(t,f)$ map to the same $(\hat{t}(t,f),\hat{f}(t,f))$, the energy is added.

The traditional reassignment coordinates from [32] are calculated for a general smoothed WVD

$$\begin{aligned}
\hat{t}_x(t,f) &= t - \frac{\iint s\,\Upsilon(s,\xi)W_x(t-s,f-\xi)dsd\xi}{\iint \Upsilon(s,\xi)W_x(t-s,f-\xi)dsd\xi},\\
\hat{f}_x(t,f) &= f - \frac{1}{2\pi}\frac{\iint \xi\,\Upsilon(s,\xi)W_x(t-s,f-\xi)dsd\xi}{\iint \Upsilon(s,\xi)W_x(t-s,f-\xi)dsd\xi},
\end{aligned} \tag{67}$$

and for the spectrogram these can be calculated with the more efficient

$$\begin{aligned}
\hat{t}(t,f) &= t + \Re\left(\frac{F_x^{th}(t,f)}{F_x^{h}(t,f)}\right),\\
\hat{f}(t,f) &= f - \frac{1}{2\pi}\Im\left(\frac{F_x^{\frac{dh}{dt}}(t,f)}{F_x^{h}(t,f)}\right).
\end{aligned} \tag{68}$$

Here $F^h(t,f)$, $F^{th}(t,f)$, and $F^{\frac{dh}{dt}}(t,f)$ are the STFTs of the signal $x(t)$ and the windows $h(t)$, $t\cdot h(t)$, and $dh(t)/dt$ respectively. This traditional reassignment

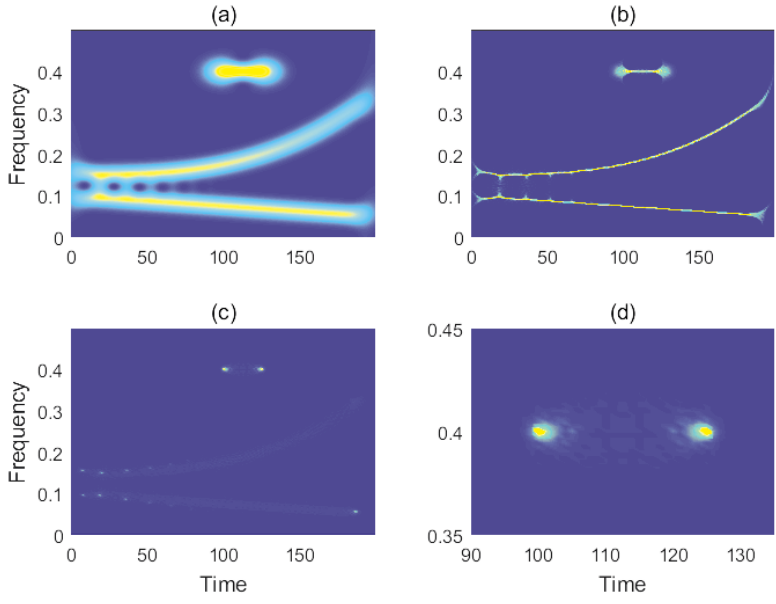



**Figure 11:** Illustration of the reassignment method for a signal with four components, (a) the spectrogram with a matched window, (b) the traditional reassigned spectrogram, (c) the scaled reassigned spectrogram with a matched window, (d) closer look at the transient components in the scaled reassigned spectrogram with a matched window.

works very well for longer duration chirps, especially if the FM is linear, however it does not work well for short duration chirps, that is transients. This is demonstrated in Figures 11(a) and (b), where the signal has two close transient components and two long chirps, one with linear FM and the other with non-linear FM. The traditional reassigned spectrogram in Figure 11(b) improves the energy localisation of the longer duration chirps, compared to the spectrogram in (a), and even though there is some interaction between the two chirps the reassigned spectrogram gives good estimates of their IFs. However, the transients are not resolved by the spectrogram or the traditional reassigned spectrogram. The traditional reassigned spectrogram instead reassigns the energy to a line going through the IFs of the transients, which is consistent with this method assuming longer duration signal components.

The reassignment method was adapted for transients that have a Gaussian envelope by Maria Sandsten and Johan Brynolfsson in 2015 [33], and extended to transients with any envelope in 2018 by Sandsten, Brynolfsson and Reinhold [34]. This method requires that the spectrogram is calculated with a window that matches the envelope of the transient, i.e. the window function must have the same shape and length as the envelope. The reassignment coordinates are also scaled

$$\begin{aligned} \hat{t}(t,f) &= t + c_t \Re\left(\frac{F^{th}(t,f)}{F^{h}(t,f)}\right), \\ \hat{f}(t,f) &= f - \frac{1}{2\pi} c_f \Im\left(\frac{F^{\frac{dh}{dt}}(t,f)}{F^{h}(t,f)}\right), \end{aligned} \tag{69}$$

where $c_t = c_f = 2$ gives reassignment to the IF of a transient when the window is matched. If the window does not match the envelope the energy will be more scattered around the IF, reducing how much the reassignment improves on the spectrogram.

The scaled matched window reassignment method works well for transients, but not long duration chirps, which is demonstrated by Figures 11(c) and (d). The scaled reassignment using a matched window in (c) can resolve the two close transients that are merged together in the spectrogram, and (d) provides a closer look at the transient components. Almost all energy from the transients is reassigned to the IF of each transient, i.e. the time centre of the envelope and the frequency of the oscillation remembered from Eq. (11). The long duration chirps are almost not visible in Figure 11(c), their energy has been reassigned to several points along the IFs of the chirps. The energy is however much more scattered compared to Figure 11(b) and the energy is mostly concentrated around the start and endpoints of the long chirps.

Given that the transient envelope needs to be known, the scaled matched window reassignment should be considered a semi-parametric method. The spectrogram, WVD, and traditional reassignment are all non-parametric methods. However, the spectrogram used in Figure 11(a) is calculated with the window that matches the transient envelopes, this choice of window function greatly affects the spectrogram's ability to resolve the transient components, and the matched window is optimal. If the matched window is not know it can be estimated using the properties of the scaled reassigned spectrogram [35].

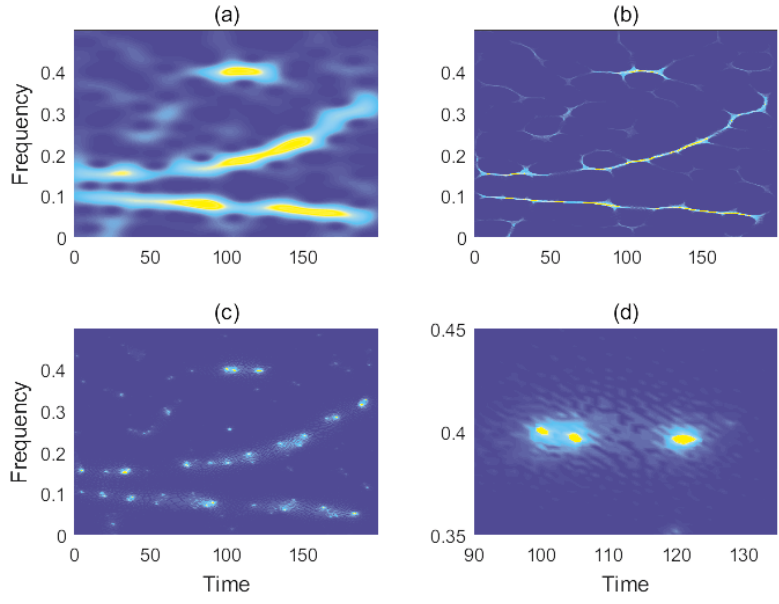



**Figure 12:** Illustration of the reassignment method for a signal with four components and white Gaussian noise, (a) the spectrogram with a matched window, (b) the traditional reassigned spectrogram, (c) the scaled reassigned spectrogram with a matched window, (d) closer look at the transient components in the scaled reassigned spectrogram with a matched window.

There are many advantages to using the traditional and scaled matched window reassignments as they greatly improve the energy localisation around the IFs of signal components. For their respective signal types their representations almost have strong finite support, which makes them easy to interpret. Because the reassignment coordinates can be calculated using only two additional STFTs compared to the spectrogram, they also have a fast implementation. The major disadvantage of the reassignment methods, is that they only have a consistent performance for moderately high SNR. The reassignment coordinates are noise sensitive. The signal used for Figure 11 has no noise.

Figure 12 shows the spectrogram, the traditional reassigned spectrogram, and the scaled matched window reassigned spectrogram for a signal with noise and relatively low SNR. The signal has the same signal components as the one used

for Figure 11. The performance of the reassignment methods deteriorates significantly with the added noise as seen in Figures 12(b), (c) and (d). Signal energy is still reassigned to locations around the IFs of the signal, but the reassignment is not as accurate and the representations are harder to interpret. While the scaled reassigned spectrogram with a matched window resolves the two transients in Figures 12(c) and (d), they are hard to identify as (the only) two transient components. In this thesis Paper D presents a method that automatically identifies transient components from the scaled matched window reassigned spectrogram, when the signals have an unknown number of (only) transients components and noise. Paper B presents a method to reduce the effects of noise on the scaled reassigned spectrogram.

There exist other methods related to the two presented reassignment techniques. The synchrosqueezing transform introduced first in the wavelet-framework [3], then related to the empirical mode decomposition [4, 36], and later defined in a Fourier-framework [6, 37]. This transform reassigns signal energy but only in the frequency dimension, it is designed for signals with longer duration components, like the traditional reassigned spectrogram. The Levenberg-Marquardt reassignment is based on the traditional reassigned spectrogram and the Levenberg-Marquardt algorithm [38–40] and has a recursive implementation [41]. This method allows the user to choose if they want weak or strong energy localisation, where strong localisations gives the normal reassignment results.

## 6 Multitaper spectral analysis

When averaging spectral representations with the aim to reduce the variance of the final representation, the requirement is that the initial spectral representations are almost uncorrelated. With the Welch method, this is achieved by using different segments of the signal when calculating the spectral estimates, and with a clever choice of window some overlap between the segments is possible. Overlap between signal segments allows for longer segments and thus the possibility for better resolution in frequency, while keeping the variance reduction almost constant. It would therefore be desirable to have 100% overlap if the spectral representations could still be made almost uncorrelated.

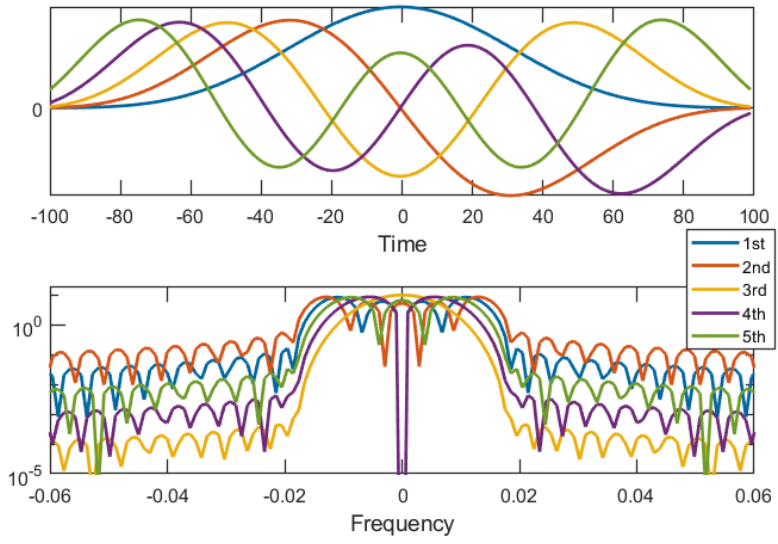



**Figure 13:** Illustration of the first five DPSS, $B = 0.035$ and $N = 200$, in time and frequency domain.

### 6.1 The Thomson multitaper method and the DPSS

David Thomson presented the idea of using multiple windows, or multitapers, to get several approximately uncorrelated spectral estimates from the same stationary signal in 1982 [42]. Today, we call this the Thomson multitaper method or the Thomson spectral estimate, and it is calculated according to

$$S_T(f) = \frac{1}{K}\sum_{k=1}^{K} S_k(f) = \frac{1}{K}\sum_{k=1}^{K}\left|\sum_{n=0}^{N-1} x(n)h_k(n)e^{-i2\pi fn}\right|^2, \tag{70}$$

where $K$ is the number of windows, $N$ is the number of samples in the stationary signal $x(n)$, and $h_k, (n)$ is the $k$th window function. Thomson suggested that the spectral estimates $S_k(f)$ are different enough to the reduce the variance of $S_T(f)$, if the $S_k(f)$ are uncorrelated for white noise in a frequency band $B$.

The window functions that Thomson proposed to use for his method are the discrete prolate spheroidal sequences (DPSS), also called the Slepian functions. These functions were proposed and thoroughly investigated by Slepian, Pollak, and Landau in the 1960s and 1970s [43–47]. The DPSS are orthogonal, and the

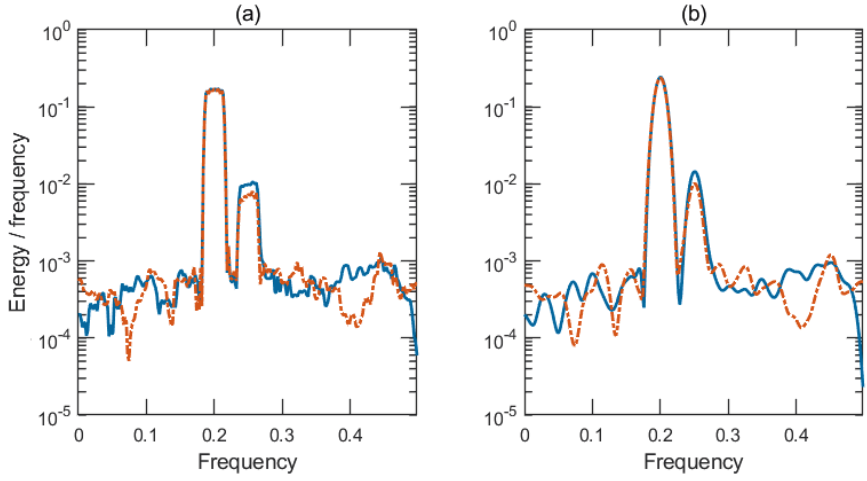



**Figure 14:** Illustration of the resolution gained by the Thomson multitaper method of two signals with the same two oscillating components but different realisations of white Gaussian noise, (a) the Thomson multitaper method, DPSS obtained according to Eq. (71) with $K = 5$ and $N = 200$, (b) the Welch method with 5 windows and 50% overlap, DPSS obtained according to Eq. (71) with $K = 1$ and $N = 66$.

time-limited windows that are most localised in the frequency domain. The width of the spectral main lobe, $B$, and duration of the window, $N$, decide how many windows, $K$, that should be used for the Thomson multitaper method. Thus, there is still a trade-off between the duration of the signal, the resolution in frequency, and the variance reduction

$$K \approx NB - 2. \tag{71}$$

The main lobe width is the same frequency band for which $S_k(f)$ are uncorrelated for white noise. The first five DPSS, $B = 0.035$ and $N = 200$, are illustrated in Figure 13, where the characteristic square main lobes can be noted. Both the time and spectral DPSS windows are orthogonal [47], thus all window functions are different both in time and frequency domain, which results in most DPSS not looking like what we typically expect from a window function.

Figure 14(a) shows the Thomson spectral estimate of a two-component signal with white Gaussian noise using the DPSS in Figure 13. Comparing Figure 14(a) with the Welch spectral estimate in Figure 14(b), both with approximately the same

variance reduction, we can see the gain in frequency resolution from the Thomson multitaper method. The downside of the Thomson method, compared to the Welch method, is that the multiple DPSS windows needs to be calculated and stored, the Welch method only requires one window and a shift parameter to be calculated and stored, which makes it more efficient. The Welch method also has a more natural connection to real-time applications, since the whole signal does not need to be measured before analysis starts. Paper A in this thesis presents a spectral estimator for stationary signals that combine some of the ideas from both the Welch and Thomson spectral estimator.

## 6.2 Time-frequency multitaper methods and the Hermite functions

Inspired by the Thomson multitaper method, multitaper TF methods began to emerge in the 1990's and several methods have been proposed since then [48–57]. The most common approach of is to calculate the average or weighted sum of spectrograms, the multitaper spectrogram

$$S_{MT}(t,f) = \sum_{k=1}^{K} \alpha_k S_k(t,f) = \sum_{k=1}^{K} \alpha_k \left| \int x(s) h_k^*(s-t) e^{-i2\pi f s} ds \right|^2, \qquad (72)$$

where the window functions $h_k(t)$ are orthogonal, and $\alpha_k$ are weights. The connection between the spectrogram and the smoothed WVD should now be remembered. By decomposing a kernel, Eq. (60), into its eigenfunctions it is possible to decompose a smoothed WVD into a multitaper spectrogram. This decomposition approach can be used to achieve more efficient calculations of smoothed WVD by using the eigenvectors as window functions for the multitaper spectrogram [58, 59].

Other sets of window functions can be found by considering the region in the TF plane where we want high localisation. For a square area $(-T,T) \times (-B,B)$, where $T$ is the time duration and $B$ the frequency bandwidth, the DPSS are good localisers, being eigenvectors of the operator defining the square area. However, the DPSS inherently treat the time and frequency domain separately and they depend on both $T$ and $B$. This makes the DPSS difficult to calculate and generally not considered optimal for the joint TF domain. The Hermite functions are eigenvectors of the operator defining the circular region $\{(t,f) \in \mathbb{R}^2; t^2 + f^2 \leq R^2\}$. They do not depend on the size of the region, i.e.

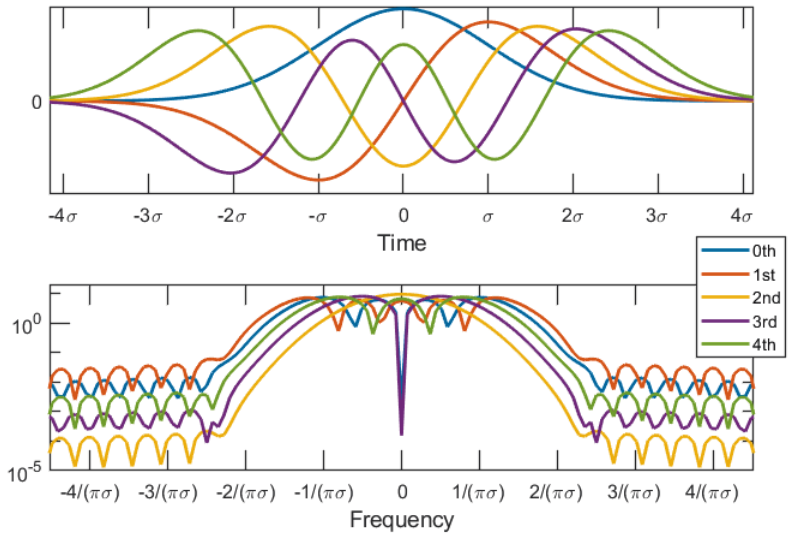



**Figure 15:** Illustration of the first five Hermite functions in the time and frequency domain.

the radius $R$, treat the time and frequency domain jointly, and are easy to calculate [60, 61]. Because of their easy calculation and good localisation properties, many TF multitaper methods use the Hermite functions [49–55].

The Hermite functions, also called Hermite-Gaussian functions, are calculated form the physicist's Hermite polynomials

$$H_k(t) = (-1)^k e^{t^2} \frac{d^k}{dt^k} e^{-t^2}, \; t \in \mathbb{R}, \; k \in \mathbb{Z}_{\geq 0}, \tag{73}$$

and the square root of their weight function

$$w_{1/2}(t) = e^{-\frac{t^2}{2}}, \; t \in \mathbb{R}. \tag{74}$$

By allowing a scaling in the duration of the polynomials and weight function we can expand from a circular region to a more general elliptical shape. The orthogonal Hermite functions are then

$$h_k(t) = H_k\left(\frac{t}{\sigma}\right) e^{-\frac{t^2}{2\sigma^2}}, \; t \in \mathbb{R}, \; k \in \mathbb{Z}_{\geq 0}, \; \sigma \in \mathbb{R}_{\geq 0}. \tag{75}$$

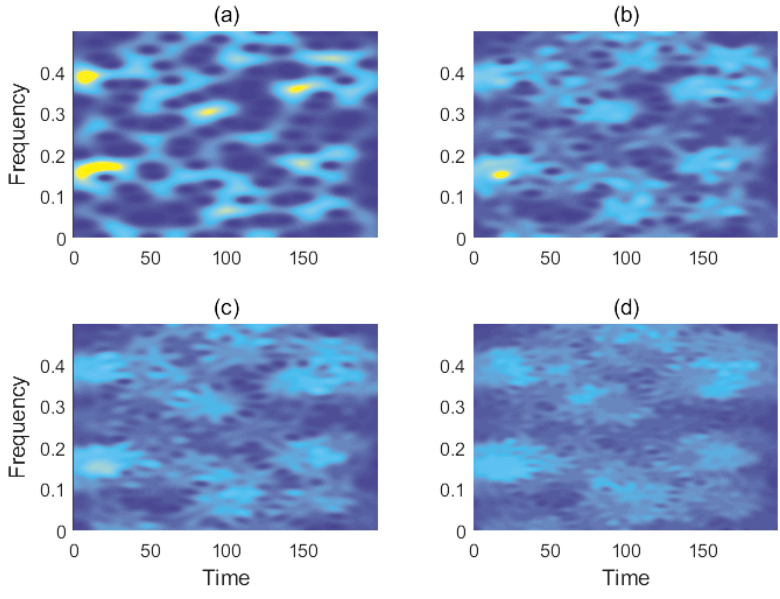



**Figure 16:** Illustration of how the multitaper spectrogram with Hermite window functions goes towards a constant spectral estimate for simulated white noise, when the number of windows increase, (a) $K = 1$, (b) $K = 2$, (c) $K = 3$, (d) $K = 5$.

The first five scaled Hermite functions are shown in Figure 15, and note that the 0th order window is the Gaussian function, which we should remember gives the optimal energy spread in time and frequency. The time representation of the Hermite functions expands with increased $\sigma$, while the main lobe in the frequency representation becomes narrower.

Just as with the DPSS, most of these windows do not look like what we expect from window functions, but this is necessary for them to give almost uncorrelated spectral estimates of white noise. Figure 16 shows the spectrogram and three multitaper spectrograms of simulated white noise, all the spectrograms use Hermite window functions. White noise is a stationary process and should have a constant spectral density, therefore we expect the multitaper spectrogram of simulated white noise to go towards a constant level when we increase the number of windows. We can see that this is achieved when using the Hermite window functions in Figure 16.

# 7 Measuring performance

The different frequency and TF representations discussed in Sections 4–6 all have different advantages and disadvantages. So, how do we compare them, measure their performance, and choose which one to use? One way is to go back to the fundamental ideas for spectral analysis presented in Section 3, but it is also possible that some of these theoretical ideas are unimportant in some applications.

The WVD fulfils many of the presented ideas in Section 3. It is time and frequency shift invariant, scale invariant, it has weak finite support, and satisfies the marginal and energy conditions. For mono-component signals the WVD also has a very high energy localisation. But the WVD has a high variance, it is greatly affected by noise, it is cumbersome to compute, and the cross-terms for multi-component signals affects readability. The spectrogram is time and frequency shift invariant, scale invariant, if the window function is unit energy it fulfils the energy condition, but not generally the marginals, nor does it technically have finite support. However, the spectrogram is fast to compute, has low variance and is easy to interpret even for many multi-component signals. Which of these qualities is most important, and what trade-offs are best, depend on the application. There is thus a need for different performance measures.

It is common to evaluate different representations with the help of simulated signals. With simulated signals it is possible to know and control the IFs of the signal components, the number of components, the type of noise, the SNR, and any other features of interest. If the signal components are easily identified in the representation, the bias and variance of any relevant estimate can be calculated. It is possible to see how noise robust the representation is by changing the SNR, and the resolution can be evaluated by moving signal components closer together. However, this type of evaluation can only be done on simulated or measured signals where the features are known beforehand, and it requires some method for extracting the relevant estimates.

There are some ways to measure performance even if the signal features of interest are unknown. One approach is to measure the energy concentration in the representation. The Rényi entropy, named after Alfréd Rényi [62], is a generalised measure of entropy. In information theory, entropy is a measure of uncertainty in a random variable, and in the context of TF representations, it can be used to measure how localised the energy is [63–65]. If energy is localised around the IFs

in a signal it is easier to extract signal features. The Rényi entropy gives a low value if the energy is localised, and is calculated according to

$$RE_\alpha = \frac{1}{1-\alpha} \log_2 \iint \left( \frac{P(t,f)}{\int\int P(t,f)dt\,df} \right)^\alpha dt\,df, \tag{76}$$

where $P(t,f)$ is the TF representation, and $\alpha = 3$ is an appropriate choice [64]. It is also possible to calculate the Rényi entropy locally in a selected area of the TF representation. This is usually preferable because the Rényi entropy favours peaky representations, which means that if a signal contains both strong and weak components, a low Rényi entropy calculated over the whole representation might only optimise localisation for the strong component(s).

The Rényi entropy is generally not a good measure for signals with close or overlapping components, since it will be minimised when the resolution is low, and the components form one peak or ridge. For such signals, a resolution measure might be more appropriate. Considering TF representations and starting with the WVD of a signal with at least two components, there are a few things of interest to achieve good resolution. The cross-term(s) need(s) to be sufficiently suppressed, the main lobes of the signal components need to be narrow, and the side lobes low. It might not however be very important that the representation fulfils the marginal conditions, the energy condition, or that it has finite support. One such resolution measure is presented by Boashash and Sucic in [66], the measure balances the importance of suppressing cross-terms, suppressing side lobes, and having high and narrow main lobes. This also means that the signal components, cross-terms, and side lobes need to be identified before the measure is used. Paper C presents a method that automatically identifies the signal components, so that the resolution measure can be calculated.

## 8 Applications

Despite that the Fourier transform calculates frequencies from an infinite time interval, and thus the mathematical interpretation must be that frequencies are time independent, most signals are by nature non-stationary. This contradiction was understood even in the early days of Fourier analysis, and captured wonderfully

by John Carson commenting on the (frequency) spectrum of a FM chirp in 1922 [67]:

> *The foregong solutions, tho unquestionably mathematically correct, are somewhat difficult to reconcile with our physical intuitions, and our physical concepts of such "variable-frequency" mechanisms as, for example, the siren.*

Many signals are however almost stationary if the duration is sufficient short, e.g. human speech is assumed to be stationary in 20 ms intervals. This means that there are still many applications for stationary spectral methods, like the one presented in Paper A. Short-term stationarity in signals means that stationary spectral methods can be used for effective real-time applications. Such applications can be in digital communication [68], or continuous monitoring, e.g. in medicine, of seismic data, or machine and structural health [69–71].

The multitaper reassignment method in Paper B aims to reduce the variance of TF representations of noisy signals while having a high energy localisation at the IFs of transient components. The paper incudes an example of the method's performance on electroencephalography (EEG) data, which is very noisy when recorded non-invasively from electrodes on the scalp. The EEG is a very popular method for measuring the electric activity of the brain, and it offers a high resolution in time [72, 73]. Recordings of EEG are used for medical diagnosis, e.g. epilepsy, stroke, and depression, but can also be used to decode mental state information [74]. Decoding of mental states and brain mapping involve finding what areas of the brain that are involved in certain tasks and understanding the brain activity related to that task, e.g. the task can be wanting to lift ones arm or remembering an image. This is a difficult and time consuming problem however, as the brain activity of the same task can be very different for any two people. But with successful decoding of mental states, it is possible to construct a brain-computer interface (BCI), i.e. a direct communication pathway between a brain and an external device. BCIs are mainly used to assist people with severe motor impairment, but can also be used for smart hearing aids, gaming devices, and medical diagnosis [75].

Paper C presents a feature extraction method that is used on heart rate variability (HRV) data. The HRV is obtained from the electrocardiogram (ECG), which records the electrical activity of the heart using electrodes on the skin. The ECG

will show the heart beats and by measuring the time between every two consecutive heart beats, we get the HRV. The time between heart beats will vary, so the HRV will not be a constant signal, but will vary with time. The HRV has two main components, the low frequency oscillations, which is associated with changes in blood pressure, and the respiratory sinus arrhythmia, which is in a higher frequency band and relates to the breathing. The method presented in Paper C identifies these two components in the TF representation. The low frequency component is fairly stationary, but the frequency of the respiratory sinus arrhythmia will change with the breathing, it will follow the breathing rhythm as well as typically increase during inhales and decrease during exhales [76]. The HRV, and mainly the respiratory component, is a good indicator of compromised health and unhealthy stress [76–78].

Estimating very short transients in the TF domain is more challenging than estimating longer signals, simply because the short duration implies less information about the component and more uncertainty in the frequencies obtained from the Fourier transform. Extracting features from short transients are of interest in a number of fields, including seizure detection, seismic wave detection, and vibration signal characterisation [79–81], but also ultrasound and biosonar analysis. In Paper D, the scaled reassigned spectrogram is used so that the TF representations have high energy localisation around transient components, and a method for detecting and classifying individual transients, in the presence of noise, is presented. The method is tested on biosonar signals and ultrasound pulse echoes both in Paper D, and more extensively on biosonar in [82] and on ultrasound in an ongoing project led by Josefin Starkhammar.

In biosonar analysis, recent research has shown that the echolocation clicks used by beluga whales (*Delphinapterus leucas*) and bottlenose dolphins (*Tursiops truncatus*) contain two close transients. This was first seen in off-axis angles where the transients become more separated in time [83, 84], but with the high resolution scaled reassigned spectrogram and the automatic detection and identification method of Paper D, analysis on a large number of clicks from bottlenose dolphins shows two transients also in on-axis recordings. These results are used to further understand how dolphins produce their echolocation clicks and how their acoustic field is affected by the structures in the head of dolphins [82].

The ability of ultrasound machines to detect small objects depend largely on the length of the pulses the machine, or transducer, sends out. A short pulse means that the recorded pulse echoes from reflections on the edges of small objects will overlap less in time compared to when a longer pulse is used [85]. It does however require more expensive hardware to produce the short pulses. Therefore, software improvements that can resolve heavily overlapping pulse echoes are of interest, and the combination of the scaled reassigned spectrogram and the transient detection method from Paper D are well suited for this. Ultrasound has many medical uses, it is non-invasive and fast to use, many will probably know of ultrasound images of fetuses and unborn children, but it can also be used to get images of other parts of the body, e.g. blood-vessels where clots can be detected, and it is used to guide the administration of regional anaesthesia and pain relief [86].

## References

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

[2] A. Grossmann and J. Morlet, "Decomposition of hardy functions into square integrable wavelets of constant shape," *SIAM Journal on Mathematical Analysis*, vol. 15, pp. 723–736, 1984.

[3] I. Daubechies and S. Maes, "A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models," in *Wavelets in Medicine and Biology*, A. Aldroubi and M. Unser, Eds. CRC Press, 1996, pp. 527–546.

[4] 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," *Proc. R. Soc. Lond. A.*, vol. 454, pp. 903–995, 1998.

[5] D. Gabor, "Theory of communications," *Journal of the IEE*, vol. 93, pp. 429 – 460, 1946.

[6] F. Auger, P. Flandrin, L. Yu-Ting, S. McLaughlin, S. Meignen, T. Oberlin, and W. Hau-Tieng, "Time-frequency reassignment and synchrosqueezing:

an overview." *IEEE Signal Processing Magazine*, vol. 30, no. 6, pp. 32 – 41, 2013.

[7] A. Schuster, "On the investigation of hidden periodicities with application to a supposed 26-day period of meteorological phenomena," *Terr. Magnet.*, vol. 3, pp. 13–41, 1898.

[8] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of Fourier series," *Math. Comput.*, vol. 19, pp. 297–301, 1965.

[9] G. Lindgren, H. Rootzen, and M. Sandsten, *Stationary Stochastic Processes for Scientists and Engineers*. CRC Press, 2014.

[10] M. Bartlett, "Smoothing periodograms from time-series with continuous spectra," *Nature.*, vol. 161, p. 686–687, 1948.

[11] P. D. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. on Audio Electroacoustics*, vol. AU-15, no. 2, pp. 70–73, June 1967.

[12] M. J. Bastiaans, "Gabor's expansion of a signal into Gaussian elementary signals," *Proceedings of the IEEE*, vol. 68, no. 4, pp. 538–539, April 1980.

[13] V. Zue and R. Cole, "Experiments on spectrogram reading," in *ICASSP '79. IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, 1979, pp. 116–119.

[14] M. Svirsky, "Cochlear implants and electronic hearing," *Physics Today*, vol. 70, no. 8, pp. 52 – 58, 2017.

[15] E. P. Wigner, "On the quantum correction for thermodynamic equilibrium," *Physics Review*, vol. 40, pp. 749–759, June 1932.

[16] T. A. C. M. Claasen and W. F. C. Mecklenbräuker, "The Wigner distribution: A tool for time frequency signal analysis. I. continuous time signals, II. discrete time signals, III. relations with other time-frequency signal transformations," *Philips J.Res.*, vol. 35, pp. 217–250, 276–300, 372–383, 1980.

[17] P. Flandrin and B. Escudie, "Time and frequency representation of finite energy signals: A physical property as a result of an Hilbertian condition," *Signal Processing*, vol. 2, no. 2, pp. 93–100, 1980.

[18] W. Martin, "Time-frequency analysis of random signals," in *ICASSP '82. IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 7, 1982, pp. 1325–1328.

[19] W. Martin and P. Flandrin, "Wigner-Ville spectral analysis of nonstationary processes," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 6, pp. 1461–1470, 1985.

[20] J. Ville, "Theorie et applications de la notion de signal analytique," *Cables et Transmissions*, vol. 2A, pp. 61–74, 1948.

[21] C. H. Page, "Instantaneous power spectra," *J. of Applied Physics*, vol. 23, pp. 103–106, January 1952.

[22] H. Margenau and R. N. Hill, "Correlation between measurements in quantum theory," *Progress of Theoretical Physics*, vol. 26, pp. 722–738, 1961.

[23] L. Cohen, "Generalized phase-space distribution functions," *Jour. Math. Phys.*, vol. 7, pp. 781–786, 1966.

[24] H. I. Choi and W. J. Williams, "Improved time-frequency representation of multi-component signals using exponential kernels," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 37, pp. 862–871, June 1989.

[25] A. W. Rihaczek, "Signal energy distribution in time and frequency," *IEEE Trans. Information Theory*, vol. 14, pp. 369–374, May 1968.

[26] Y. Zhao, L. E. Atlas, and R. J. Marks, "The use of cone-shaped kernels for generalized time-frequency representations of nonstationary signals," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 38, pp. 1084–1091, July 1990.

[27] R. Liao, C. Guo, K. Wang, Z. Zuo, and A. Zhuang, "Adaptive optimal kernel time–frequency representation technique for partial discharge ultra-high-frequency signals classification," *Electric Power Components and Systems*, vol. 43, no. 4, pp. 449–460, 2015.

[28] R. G. Baraniuk and D. L. Jones, "A signal-dependent time-frequency representation: Optimal kernel design," *IEEE Transactions on Signal Processing*, vol. 41, no. 4, pp. 1589 – 1602, 1993.

[29] L. Cohen, *Time-Frequency Analysis*, ser. Signal Processing Series. Upper Saddle River, NJ, USA: Prentice-Hall, 1995.

[30] B. Boashash, "Part I: Introduction to the concepts of TFSAP," in *Time Frequency Signal Analysis and Processing - A Comprehensive Reference*, B. Boashash, Ed. Oxford: Elsevier Ltd, 1994.

[31] K. Kodera, C. de Villedary, and R. Gendrin, "A new method for the numerical analysis of nonstationary signals," *Physics of the Earth & Planetary Interiors*, vol. 12, pp. 142–150, 1976.

[32] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. on Signal Processing*, vol. 43, pp. 1068–1089, May 1995.

[33] M. Hansson-Sandsten and J. Brynolfsson, "The scaled reassigned spectrogram with perfect localization for estimation of Gaussian functions," *IEEE Signal Processing Letters*, vol. 22, no. 1, pp. 100–104, January 2015.

[34] M. Sandsten, J. Brynolfsson, and I. Reinhold, "The matched window reassignment," in *26th European Signal Processing Conference, EUSIPCO 2018*. European Association for Signal Processing (EURASIP), 2018, pp. 2340–2344.

[35] J. Brynolfsson, I. Reinhold, and M. Sandsten, "A time-frequency-shift invariant parameter estimator for oscillating transient functions using the matched window reassignment," *Signal Processing*, vol. 183, p. 107913, 2021.

[36] I. Daubechies, J. Lu, and H. T. Wu, "Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool," *Applied and Computational Harmonic Analysis*, vol. 30, no. 1, pp. 243–261, 2011.

[37] T. Oberlin, S. Meignen, and V. Perrier, "The Fourier-based synchrosqueezing transform," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 315–319.

[38] F. Auger, E. Chassande-Mottin, and P. Flandrin, "Making reassignment adjustable: The Levenberg-Marquardt approach," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2012, pp. 3889–3892.

[39] K. Levenberg, "A method for the solution of certain non-linear problems in least squares," *Quarterly of Applied Mathematics*, vol. 2, no. 2, p. 164–168, 1944.

[40] D. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *SIAM Journal on Applied Mathematics*, vol. 11, no. 2, p. 431–441, 1963.

[41] D. Fourer, F. Auger, and P. Flandrin, "Recursive versions of the Levenberg-Marquardt reassigned spectrogram and of the synchrosqueezed STFT," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2016, pp. 4880–4884.

[42] D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proc. of the IEEE*, vol. 70, no. 9, pp. 1055–1096, Sept 1982.

[43] D. Slepian and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty – I," *The Bell System Technical Journal*, vol. 40, no. 1, pp. 43–63, 1961.

[44] H. J. Landau and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty – II," *The Bell System Technical Journal*, vol. 40, no. 1, pp. 65–84, 1961.

[45] ——, "Prolate spheroidal wave functions, Fourier analysis and uncertainty – III: The dimension of the space of essentially time- and band-limited signals," *The Bell System Technical Journal*, vol. 41, no. 4, pp. 1295–1336, 1962.

[46] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis and uncertainty – IV: Extensions to many dimensions; generalized prolate spheroidal functions," *The Bell System Technical Journal*, vol. 43, no. 6, pp. 3009–3057, 1964.

[47] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis and uncertainty-V: The discrete case," *Bell System Journal*, vol. 57, no. 5, pp. 1371–1430, May-June 1978.

[48] G. Fraser and B. Boashash, "Multiple window spectrogram and time-frequency distributions," in *Proceedings of ICASSP '94. IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. IV, 1994, pp. 293–296.

[49] M. Bayram and R. G. Baraniuk, "Multiple window time-frequency analysis," in *Proceedings of Third International Symposium on Time-Frequency and Time-Scale Analysis (TFTS-96)*, 1996, pp. 173–176.

[50] Y. Xu, S. Haykin, and R. J. Racine, "Multiple window time-frequency distribution and coherence of EEG using Slepian sequences and Hermite functions," *IEEE Transactions on Biomedical Engineering*, vol. 46, no. 7, pp. 861–866, July 1999.

[51] F. Cakrak and P. J. Loughlin, "Multiwindow time-varying spectrum with instantaneous bandwidth and frequency constraints," *IEEE Transactions on Signal Processing*, vol. 49, no. 8, pp. 1656–1666, 2001.

[52] S. Aviyente and W. J. Williams, "Multitaper marginal time–frequency distributions," *Signal Processing*, vol. 86, no. 2, pp. 279–295, 2006.

[53] J. Xiao and P. Flandrin, "Multitaper time-frequency reassignment for non-stationary spectrum estimation and chirp enhancement," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2851–2860, 2007.

[54] I. Orović, S. Stanković, T. Thayaparan, and L. Stanković, "Multiwindow S-method for instantaneous frequency estimation and its application in radar signal analysis," *IET Signal Processing*, vol. 4, no. 4, pp. 363–370, 2010.

[55] M. Hansson-Sandsten, "Optimal multitaper Wigner spectrum estimation of a class of locally stationary processes using Hermite functions," *EURASIP J. Adv. Signal Process.*, p. Article ID 980805, 2011.

[56] ——, "Multitaper Wigner and Choi–Williams distributions with predetermined Doppler–lag bandwidth and sidelobe suppression," *Signal Processing*, vol. 91, no. 6, pp. 1457–1465, 2011.

[57] K. Czarnecki, D. Fourer, F. Auger, and M. Rojewski, "A fast time-frequency multi-window analysis using a tuning directional kernel," *Signal Processing*, vol. 147, pp. 110–119, 2018.

[58] G. S. Cunningham and W. J. Williams, "Kernel decomposition of time-frequency distributions," *IEEE Transactions on Signal Processing*, vol. 42, no. 6, pp. 1425–1442, 1994.

[59] ——, "Fast implementations of generalized discrete time-frequency distributions," *IEEE Transactions on Signal Processing*, vol. 42, no. 6, pp. 1496–1508, 1994.

[60] I. Daubechies, "Time-frequency localization operators: a geometric phase space approach," *IEEE Transactions on Information Theory*, vol. 34, no. 4, pp. 605–612, 1988.

[61] R. G. Shenoy and T. W. Parks, "The Weyl correspondence and time-frequency analysis," *IEEE Transactions on Signal Processing*, vol. 42, no. 2, pp. 318–331, 1994.

[62] A. Rényi, "On measures of entropy and information," in *Proc. 4th Berkeley Symp. Mathematics of Statistics and Probability*, vol. 1, 1961, pp. 547–561.

[63] W. J. Williams, M. L. Brown, and A. O. H. III, "Uncertainty, information, and time-frequency distributions," in *Advanced Signal Processing Algorithms, Architectures, and Implementations II*, F. T. Luk, Ed., vol. 1566, International Society for Optics and Photonics. SPIE, 1991, pp. 144 – 156.

[64] R. G. Baraniuk, P. Flandrin, A. J. E. M. Janssen, and O. J. J. Michel, "Measuring time-frequency information content using Rényi entropies," *IEEE Trans. on Information Theory*, vol. 47, no. 4, pp. 1391–1409, May 2001.

[65] N. Saulig, I. Orović, and V. Sucic, "Optimization of quadratic time-frequency distributions using the local Rényi entropy information," *Signal Processing*, vol. 129, pp. 17–24, 2016.

[66] B. Boashash and V. Sucic, "Resolution measure criteria for the objective assessment of the performance of quadratic time-frequency distributions," *Signal Processing*, vol. 51, pp. 1253–1263, 2003.

[67] J. R. Carson, "Notes on the theory of modulation," *Proceedings of the Institute of Radio Engineers*, vol. 10, no. 1, pp. 57–64, 1922.

[68] Chen-Mou Cheng, H. T. Kung, and Koan-Sin Tan, "Use of spectral analysis in defense against dos attacks," in *Global Telecommunications Conference, 2002. GLOBECOM '02. IEEE*, vol. 3, 2002, pp. 2143–2148.

[69] E. M. Schumacher, A. S. Westvik, P. G. Larsson, R. Lindemann, J. Westvik, and T. A. Stiris, "Feasibility of long-term continuous EEG monitoring during the first days of life in preterm infants: An automated quantification of the EEG activity," *Pediatr Res*, vol. 69, pp. 413–417, 2011.

[70] R. Carniel, "Characterization of volcanic regimes and identification of significant transitions using geophysical data: a review," *Bull Volcanol*, vol. 76, p. 848, 2014.

[71] Y. Kaya and E. Safak, "Real-time analysis and interpretation of continuous data from structural health monitoring (SHM) systems," *Bull Earthquake Eng*, vol. 13, p. 917–934, 2015.

[72] J. Jiang, I. Bramão, A. Khazenzon, S.-F. Wang, M. Johansson, and A. D. Wagner, "Temporal dynamics of memory-guided cognitive control and generalization of control via overlapping associative memories," *Journal of Neuroscience*, vol. 40, no. 11, pp. 2343–2356, 2020.

[73] C. Kerrén, I. Bramão, R. Hellerstedt, and M. Johansson, "Strategic retrieval prevents memory interference: The temporal dynamics of retrieval orientation," *Neuropsychologia*, vol. 154, p. 107776, 2021.

[74] B. He, L. Astolfi, P. A. Valdés-Sosa, D. Marinazzo, S. O. Palva, C.-G. Bénar, C. M. Michel, and T. Koenig, "Electrophysiological brain connectivity: Theory and implementation," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 7, pp. 2115 – 2137, 2019.

[75] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger, "A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update," *J Neural Eng.*, vol. 15, no. 3, p. 031005, 2018.

[76] G. Billman, "Heart rate variability – a historical perspective," *Frontiers in Physiology*, vol. 2, p. 86, 2011.

[77] M. T. Verklan and N. S. Padhye, "Spectral analysis of heart rate variability: An emerging tool for assessing stability during transition to extrauterine life," *Journal of Obstetric, Gynecologic & Neonatal Nursing*, vol. 33, no. 2, pp. 256–265, 2004.

[78] A. Hernando, J. Lázaro, E. Gil, A. Arza, J. M. Garzón, R. López-Antón, C. de la Cámara, P. Laguna, J. Aguiló, and R. Bailón, "Inclusion of respiratory frequency information in heart rate variability analysis for stress assessment," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 4, pp. 1016–1025, 2016.

[79] H. Kalbkhani and M. G. Shayesteh, "Stockwell transform for epileptic seizure detection from EEG signals," *Biomedical Signal Processing and Control*, vol. 38, pp. 108–118, 2017.

[80] J. Pons-Llinares, M. Riera-Guasp, J. A. Antonino-Daviu, and T. G. Habetler, "Pursuing optimal electric machines transient diagnosis: The adaptive slope transform." *Mechanical Systems and Signal Processing*, vol. 80, pp. 553 – 569, 2016.

[81] Q. He and X. Ding, "Sparse representation based on local time-frequency template matching for bearing transient fault feature extraction." *Journal of Sound and Vibration*, vol. 370, pp. 424 – 443, 2016.

[82] J. Starkhammar, I. Reinhold, P. Moore, D. Houser, and M. Sandsten, "Detailed analysis of two detected overlaying transient components within the echolocation beam of a bottlenose dolphin (Tursiops truncatus)," *J. Acoust. Soc. Am.*, vol. 145, no. 4, pp. 2138–2148, 2019.

[83] M. Lammers and M. Castellote, "The beluga whale produces two pulses to form its sonar signal," *Biol.Lett.*, vol. 5, pp. 297–301, 2009.

[84] W. Au, B. Branstetter, P. Moore, and J. Finneran, "Dolphin biosonar signals measured at extreme off-axis angles: Insights to sound propagation in the head," *J. Acoust. Soc. Am.*, vol. 132, no. 2, pp. 1199–1206, 2012.

[85] A. Ng and J. Swanevelder, "Resolution in ultrasound imaging," *Continuing Education in Anaesthesia Critical Care & Pain*, vol. 11, no. 5, pp. 186–192, 2011.

[86] P. Marhofer, W. Harrop-Griffiths, H. Willschke, and L. Kirchmair, "Fifteen years of ultrasound guidance in regional anaesthesia: Part 2—recent developments in block techniques," *British Journal of Anaesthesia*, vol. 104, no. 6, pp. 673–683, 2010.

# Scientific Publications

# Scientific Publications

The first two papers in this thesis present methods to reduce the variance of spectral representations, the first for generic stationary signals using a time-shifted optimal window, and the second for the reassigned spectrogram of transients. The last two papers present automatic methods for detecting signal components, the first adapted for the Wigner-Ville distribution of long duration FM signals, and the second adapted for the reassigned spectrogram of transients. This chapter provides a brief overview of these papers and their presented key concepts, followed by a list detailing my contributions as author.

**Paper A:** Efficient Thomson Spectral Estimator with Time-Shifted Windows

This paper presents a spectral estimation method that is a hybrid between the Welch method and the Thomson multitaper method. By defining a bandwidth for the main lobe of the window and a bias constraint, a single window function is calculated as a weighted sum of the discrete prolate spheroidal sequences. This single window is shifted in time to create a set of windows, and both the weights for the sum and the time shifts are optimised to minimise the variance of white noise. The paper presents a 5-step procedure for the window estimation. This approach combines the computational efficiency of the Welch estimator and the ability to explicitly define a frequency resolution, which is associated with the Thomson method. The minimisation problem is not convex and search for a global minimum can be costly, therefore the paper also presents a convex approximation of the minimisation problem. The spectral estimators for the convex and non-convex minimisation are evaluated and compared to the Welch and Thomson spectral estimators.

**Paper B**: Multitaper Reassignment for Oscillating Transients with Gaussian Envelopes

This paper presents a multitaper reassigned spectrogram for transients with a Gaussian envelope. This method makes the reassignment more robust to noise, introducing the ability to balance the resolution and variance of the scaled reassigned spectrogram. The Hermite functions are used as window functions, and the paper presents new reassignment coordinates that reassign all signal energy of one transient to its instantaneous frequency for all the window functions. These new reassignment coordinates are then combined, rather than averaging the reassigned spectrograms, to produce a multitaper reassigned spectrogram. The correlation between the reassignment coordinates for white noise and different windows is evaluated, to ensure that the multitaper method reduces the variance. The method is evaluated for simulated signals with two transient components that are either separated or overlapping, and the signals are disturbed by either white Gaussian noise or pink noise.

**Paper C**: Optimal Time-Frequency Distributions using a Novel Signal Adaptive Method for Automatic Component Detection

This paper presents a method for automatic detection of the signal components in the Wigner-Ville distributions of signals with two long duration chirps. The cross-term, side lobes and other features are also identified with this method. The novel automatic detection improves on a previously existing method, and it is used with a normalised instantaneous resolution performance measure to optimise the modified B-distribution. The modified B-distribution is a smoothed Wigner-Ville distribution. The motivation for a new detection method is to improve the ability to detect components when they have unequal amplitudes and when one or both have non-linear FM. The performance of the novel automatic detection method is evaluated for different parameter choices in the two-component signal and for different levels of smoothing of the Wigner-Ville distribution.

**Paper D**: Objective Detection and Time-Frequency Localization of Components within Transient Signals

This paper presents a method that automatically detects transient components in the scaled reassigned spectrogram of signals with an unknown number of

transients. The scaled reassigned spectrogram is used with a matched window, which resolves close transient components, and the proposed detection algorithm detects the individual components and provides their instantaneous frequencies. The novel automatic detection method also estimates how many transient components a signal has. Evaluation is mainly done on simulated signals with white Gaussian noise and varying SNR to determine the resolution and if the method is noise robust. But the evaluation also includes recorded data from ultrasound pulse-echoes and marine biosonar.

## Author contributions

**Paper A**: Efficient Thomson Spectral Estimator with Time-Shifted Windows

I calculated the minimization problem and defined the convex approximation. I developed the procedure and evaluated both methods. We collaborated on the writing.

**Paper B**: Multitaper Reassignment for Oscillating Transients with Gaussian Envelopes

It was a joint idea. I derived the new reassignment coordinates, and we collaborated on the method for combining the reassignment coordinates. I did the simulations and evaluation. We collaborated on the writing.

**Paper C**: Optimal Time-Frequency Distributions using a Novel Signal Adaptive Method for Automatic Component Detection

I had the idea for the automatic detection algorithm, developed the algorithm and did the evaluation. We collaborated on the writing.

**Paper D**: Objective Detection and Time-Frequency Localization of Components within Transient Signals

It was a joint idea. I developed the detection algorithm and did the evaluation on the simulated signals, excluding the noise robust testing. I helped in collecting the ultrasound data and did the evaluation on that data. We collaborated on the writing.

# Paper A

**Paper A**

# Efficient Thomson Spectral Estimator with Time-Shifted Windows

Isabella Reinhold and Maria Sandsten

*Mathematical Statistics, Centre for Mathematical Sciences, Lund University, Sweden.*

**Abstract**

In this paper, optimal spectral analysis window shapes, using weighted discrete prolate spheroidal sequences as basis functions, are proposed. These windows are not typically positive or even. The windows are time-shifted, combining the computational efficiency of the Welch method and the appealing property of predefined frequency resolution of the Thomson spectral estimator. The parameters of the optimal windows are found by minimising the resulting spectral covariances and optimising the window overlap, for the predetermined frequency resolution and number of windows. The windows are found to have low side lobes, giving small spectral leakage, and the final spectral estimate gives close to optimal variance reduction, i.e. the covariance between different sub-spectra is very small.

**Keywords:** DPSS, Slepian functions, Spectral leakage, Variance, Welch method

## 1 Introduction

The Welch method [1] is well known and used as spectral estimator in many different applications [2–4]. The spectral estimate has low variance as the time-shifted windows result in almost orthogonal sub-spectra that are averaged for the final estimate. Partly overlapping and smooth windows, such as the commonly used Hanning window, are more beneficial than non-overlapping rectangular windows [5, 6]. From frequency resolution and leakage viewpoint, different window shapes have been thoroughly investigated and compared [7, 8] and for a predefined frequency resolution, the discrete prolate spheroidal sequences (DPSS) are the most optimal from leakage viewpoint [9]. Multitaper estimators [10], which uses all data samples (100% overlap) for all the windowed periodograms, are also popular choices for spectral estimation. The properties of the different windows give uncorrelated periodograms and thereby reduced variance.

Optimising variance, resolution and leakage is often of great interest, where window shapes with certain properties are chosen followed by optimisation of the overlap and minimising the overall variance [11, 12]. It has been shown that the Thomson multitaper estimator, based on the DPSS, outperforms the Welch method in terms of bias and variance [13]. However, the Welch method is more efficient in real-time applications, with less computations and less memory allocation, as the windowed sequences require shorter discrete Fourier transforms and less storage [11]. The appealing Thomson multitaper property of predefined resolution was used to optimise a time-shifted window shape of the Welch estimator in [14], but the resulting windows did not fulfil the property of window orthogonality and well suppressed side lobes.

This paper proposes using the DPSS as basis functions to estimate an optimal analysis window shape for a predefined frequency resolution and a fixed number of time-shifted windows. The window shape is optimal in the sense that it minimises the variance of white Gaussian noise, under the constraint of well suppressed side lobes, i.e. low spectral leakage. A procedure to find the optimal overlap and corresponding window shapes is also proposed. The windows are not restricted to be positive or even.

## 2 Thomson estimator with Welch window structure

Given the discrete-time zero-mean stationary stochastic process, $x(n)$, with spectral density, $S_x(f)$, the spectrum can be estimated from the $N$ samples $\mathbf{x} = [x(0) \ \ldots \ x(N-1)]^T$, ($T$ denotes transpose), using the Welch method

$$\hat{S}_x(f) = \frac{1}{K} \sum_{k=1}^{K} \left| \sum_{n=0}^{N_g-1} x(n + (k-1)L) g(n) e^{-i2\pi f n} \right|^2, \tag{1}$$

where $K$ is the number of windows, $L$ is the time-shift in samples, $\mathbf{g} = [g(0) \ \ldots \ g(N_g-1)]^T$ is the window function and the length of the window is the largest integer $N_g \leq N - L(K-1)$. Another way of estimating the spectrum is instead to consider the set of time-shifted windows

$$\mathbf{h}_k = [\underbrace{0 \ \ldots \ 0}_{(k-1)L} \ \mathbf{g} \ \overbrace{0 \ \ldots \ 0}^{N-((k-1)L+N_g)}]^T, \quad k = 1 \ \ldots \ K. \tag{2}$$

These time-shifted windows make it possible to formulate a Thomson estimator

$$\hat{S}_x(f) = \frac{1}{K} \sum_{k=1}^{K} \left| \sum_{n=0}^{N-1} x(n) h_k(n) e^{-i2\pi f n} \right|^2. \tag{3}$$

This paper proposes that for some given frequency resolution, indicated by $B$, set a number of overlapping windows $K$, appropriate for the application, and then optimise the shape of the single window function $\mathbf{g}$ and its length $N_g$ to reduce the (co)variance and leakage.

## 3 An optimal time-shifted window

The approach in this paper is to combine a set of $M$ basis functions $\mathbf{q}_m = [q_m(0) \ \ldots \ q_m(N_g-1)]^T$, $m = 1 \ldots M$, to express the single window

$$\mathbf{g} = \sum_{m=1}^{M} \alpha_m \mathbf{q}_m = \mathbf{Q}\boldsymbol{\alpha}, \tag{4}$$

where $\boldsymbol{\alpha} = [\alpha_1 \ \alpha_2 \ \ldots \ \alpha_M]^T$, is the scaling vector and $\mathbf{Q} = [\mathbf{q}_1 \ \mathbf{q}_2 \ \ldots \ \mathbf{q}_M]$ is the matrix including the basis functions as column vectors.

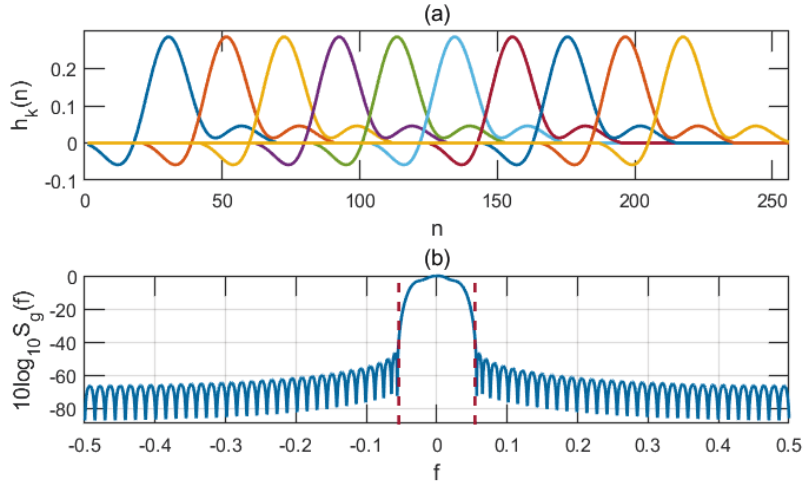



**Figure 1:** Optimal time-shifted windows $\mathbf{h}_k$ for the frequency band $B = 0.108$, when $N = 256$ and $K = 10$; (a) time-shifted windows using $M = 4$ DPSS basis functions; (b) window spectrum of each single window $\mathbf{g}$.

The DPSS [9] are used as basis functions since they are all orthogonal and localised to have maximum power inside a predefined frequency band. The spectral leakage is determined by the corresponding eigenvalue, thus making it easy to define a tolerated leakage $\epsilon_T = 1 - \lambda_i$ and at the same time determine the number of basis functions $M$, i.e. all eigenvectors with corresponding $\lambda_m \geq 1 - \epsilon_T,\ m = 1 \ldots M$.

The resulting shapes of the windows $\mathbf{g}$ are typically not positive or even. Figure 1(a) shows an example of this when $B = 0.108$, $N = 256$ and $K = 10$. The windows are a sum of $M = 4$ DPSS, the tolerated leakage is $\epsilon_T = 0.0005$, the overlap 70% and $\boldsymbol{\alpha} = [0.87\ 0.28\ -0.17\ -0.36]^T$. Figure 1(b) shows the spectrum of the window $\mathbf{g}$. The dashed vertical lines marks the frequency band $|f| = B/2 = 0.054$, and it can be noted that for $|f| > B/2$ the spectrum is very low, the first side lobe is only -47 dB high.

An optimal spectral estimator should minimise the variance of a white Gaussian noise process, $\mathcal{N}(0,1)$,

$$\text{Var}\ \hat{S}_w(f) = \frac{1}{K^2} \sum_{k_1=1}^{K} \sum_{k_2=1}^{K} |\mathbf{h}_{k_1}^T \mathbf{h}_{k_2}|^2, \tag{5}$$

under the constraint of zero bias,

$$\mathrm{E}[\hat{S}_w(f)] = \frac{1}{K}\sum_{k=1}^{K} \mathbf{h}_k^T \mathbf{h}_k = 1, \tag{6}$$

[10]. As the DPSS windows are orthonormal, i.e., $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix, the zero bias constraint is simplified to

$$\mathrm{E}[\hat{S}_w(f)] = \mathbf{g}^T\mathbf{g} = \boldsymbol{\alpha}^T\mathbf{Q}^T\mathbf{Q}\boldsymbol{\alpha} = \boldsymbol{\alpha}^T\boldsymbol{\alpha} = 1. \tag{7}$$

Thus the minimisation problem that needs to be solved is

$$\arg\min_{|\boldsymbol{\alpha}|^2=1} \mathrm{Var}\ \hat{S}_w(f), \tag{8}$$

which will give the scaling vector $\boldsymbol{\alpha}$ to construct the single window function $\mathbf{g}$. However, the shape of $\mathbf{g}$ is dependent on the window length, and thus the window overlap, a procedure to find both the optimal window shape and overlap is presented next.

### 3.1 Procedure

This procedure finds the optimal shape of the window $\mathbf{g}$ and the optimal overlap for the set of time-shifted windows $\mathbf{h}_k$, assuming there is a predetermined frequency band $B$ and number of windows $K$.

1. Decide a range of overlaps to be evaluated.
2. Find the number of DPSS basis functions which fulfil $\lambda_m \geq 1 - 0.0005,\ m = 1 \ldots M$.
3. Solve the minimisation problem Eq. (8) for all considered overlaps and calculate the minimum variances given by the obtained $\boldsymbol{\alpha}$s.
4. Find the smallest minimum variance $\eta$. Consider all, though possible just one, of the $\boldsymbol{\alpha}$ that give variances smaller than $\eta + 0.0005^2$, choose among them the $\boldsymbol{\alpha}$ that corresponds to the smallest overlap.
5. From the chosen $\boldsymbol{\alpha}$ and corresponding overlap, construct the time-shifted windows $\mathbf{h}_k$.

The tolerated leakage

$$\epsilon_T = 1 - P_B = 1 - \int_{-B/2}^{B/2} S_b(f)df \leq 0.0005, \tag{9}$$

used in step 2, is chosen as the first sidelobe of the Hanning window spectrum is found slightly below -30 dB. In step 4 a deviation from the smallest minimum variance is allowed to balance the variance minimisation and the increased computational complexity of larger overlaps.

## 4 Variance minimisation

When solving the minimisation problem Eq. (8) it is of interest to consider the overlaps of the time-shifted windows. Lets define the lower and upper parts of the window vector $\mathbf{g}$

$$\mathbf{g}_{l_k} = [g(kL) \dots g(N_g - 1)]^T, \tag{10}$$

$$\mathbf{g}_{u_k} = [g(0) \dots g(N_g - 1 - kL)]^T. \tag{11}$$

They can be used to rewrite the expression Eq. (5) into

$$\begin{aligned} \min \operatorname{Var} \hat{S}_w(f) &= \min \frac{1}{K} |\mathbf{g}^T \mathbf{g}|^2 + \sum_{k=1}^{K_C} w_k |\mathbf{g}_{l_k}^T \mathbf{g}_{u_k}|^2 \\ &= \min \operatorname{Var}_0 \hat{S}_w(f) + \operatorname{Cov} \hat{S}_w(f), \end{aligned} \tag{12}$$

where the number of different overlaps is $K_C < N_g/L < K$ and $w_k = 2\frac{(K-k)}{K^2}$. However, the first term $\operatorname{Var}_0$ is always $1/K$ as $|\mathbf{g}^T \mathbf{g}|^2 = 1$, therefore only the covariance terms needs to be minimised. Using that

$$\mathbf{g}_{l_k} = \mathbf{Q}_{l_k} \boldsymbol{\alpha}, \tag{13}$$

$$\mathbf{g}_{u_k} = \mathbf{Q}_{u_k} \boldsymbol{\alpha}, \tag{14}$$

where $\mathbf{Q}_{l_k}$ and $\mathbf{Q}_{u_k}$ are the corresponding lower and upper part of the DPSS basis functions matrix $\mathbf{Q}$ respectively, the minimisation problem Eq. (8) can then be reformulated as

$$\arg\min_{|\boldsymbol{\alpha}|^2=1} \operatorname{Var} \hat{S}_w(f) = \arg\min_{|\boldsymbol{\alpha}|^2=1} \sum_{k=1}^{K_C} w_k |\boldsymbol{\alpha}^T \mathbf{Q}_{l_k}^T \mathbf{Q}_{u_k} \boldsymbol{\alpha}|^2. \tag{15}$$

The DPSS are either even or odd sequences, $q_m(N_g-1-n) = (-1)^{(m-1)} q_m(n)$ [9], giving the following relationship

$$\mathbf{Q}_{l_k} = \mathbf{I}_p \mathbf{Q}_{u_k} \mathbf{I}_{cs}, \tag{16}$$

where $\mathbf{I}_{cs}$ is the $(M \times M)$ diagonal signature matrix with elements $(-1)^{(m-1)}$, and $\mathbf{I}_p$ is the $(N_g - 1 - |kL| \times N_g - 1 - |kL|)$ diagonal exchange matrix. Thus the product

$$\mathbf{A}_k = \mathbf{Q}_{l_k}^T \mathbf{Q}_{u_k} = \mathbf{I}_{cs} \mathbf{Q}_{u_k}^T \mathbf{I}_p \mathbf{Q}_{u_k} \tag{17}$$

is a $(M \times M)$ non-symmetric matrix. The minimisation problem can be solved using iterative optimisation methods for non-linear problems, but since the problem is non-convex, these methods can be costly. However, both $M$ and $K_C$ are assumed to always be rather small, which still make these optimisation methods a valid choice. The time-shifted windows resulting from solving Eq. (15) will be referred to as the iterative optimal time-shifted windows (I-OTSW).

### 4.1 Approximation

In order to achieve more computational efficiency a simplification of the minimisation problem Eq. (15) with an analytical solution is also proposed. The time-shifted windows resulting from solving this problem will be called the approximative optimal time-shifted windows (A-OTSW).

The original minimisation problem is a sum of non-negative numbers

$$\begin{aligned} \min \operatorname{Var} \hat{S}_w(f) &= \arg\min_{|\boldsymbol{\alpha}|^2=1} \sum_{k=1}^{K_c} w_k |\boldsymbol{\alpha}^T \mathbf{A}_k \boldsymbol{\alpha}|^2 \\ &= \arg\min_{|\boldsymbol{\alpha}|^2=1} \sum_{k=1}^{K_c} w_k \boldsymbol{\alpha}^T \mathbf{A}_k \boldsymbol{\alpha} \boldsymbol{\alpha}^T \mathbf{A}_k^T \boldsymbol{\alpha}, \end{aligned} \tag{18}$$

and a reasonable simplification is to instead consider the much simpler, although similar problem

$$\arg\min_{|\boldsymbol{\alpha}|^2=1} \sum_{k=1}^{K_c} w_k \boldsymbol{\alpha}^T \mathbf{A}_k \mathbf{A}_k^T \boldsymbol{\alpha}. \tag{19}$$

This is also a sum of non-negative numbers, since the matrix product $\mathbf{A}_k \mathbf{A}_k^T$ is Hermitian and positive semi-definite. The solution to this new problem is found

by finding the right-singular vector of the $(K_c M \times M)$ block matrix

$$\begin{bmatrix} w_1 \mathbf{A}_1 \mathbf{A}_1^T \\ \vdots \\ w_{K_c} \mathbf{A}_{K_c} \mathbf{A}_{K_c}^T \end{bmatrix}, \tag{20}$$

corresponding to the smallest singular value.

At first glance it might seem tempting to first minimise the problem in Eq. (19) to get a solution $\boldsymbol{\alpha}_0$, and then use that to solve $\arg\min_{|\boldsymbol{\alpha}|^2=1} \sum_{k=1}^{K_c} w_k \boldsymbol{\alpha}^T \mathbf{A}_k \boldsymbol{\alpha}_0 \boldsymbol{\alpha}_0^T \mathbf{A}_k^T \boldsymbol{\alpha}$. However, since $\boldsymbol{\alpha}_0$ is a singular vector of the block matrix Eq. (20), the rank of the new block matrix, with rows $w_k \mathbf{A}_k \boldsymbol{\alpha}_0 \boldsymbol{\alpha}_0^T \mathbf{A}_k^T$, is reduced and the solution will always be $\boldsymbol{\alpha} = \alpha_1 = 1$.

## 5 Evaluation

The proposed I-OTSW and A-OTSW estimators are evaluated with the usual Hanning window (Welch) and the first DPSS window (STSW) as Welch estimators and also to the Thomson multitaper estimator (Thomson). All windows are optimised or normalised to fulfil the zero bias condition Eq. (7). The evaluation is done for $N = 256$ and a range of different frequency bands according to Table 1 and overlap between 30% - 75%, for the I-OTSW, A-OTSW and STSW. The Nelder-Mead simplex method of Matlab (fminsearch) with a set of random initial values is used for the I-OTSW estimator. The iterative search is repeated for reliable convergence. The Welch estimator is used with 50% overlap and the Thomson estimator with $NB - 3$ multitapers [10].

**Table 1:** Range of frequency bands and corresponding number of windows for the I-OTSW and A-OTSW estimators.

| $B$ | $K$ |
|---|---|
| [0.040, 0.052] | 4 |
| [0.056, 0.072] | 6 |
| [0.076, 0.088] | 8 |
| [0.092, 0.108] | 10 |

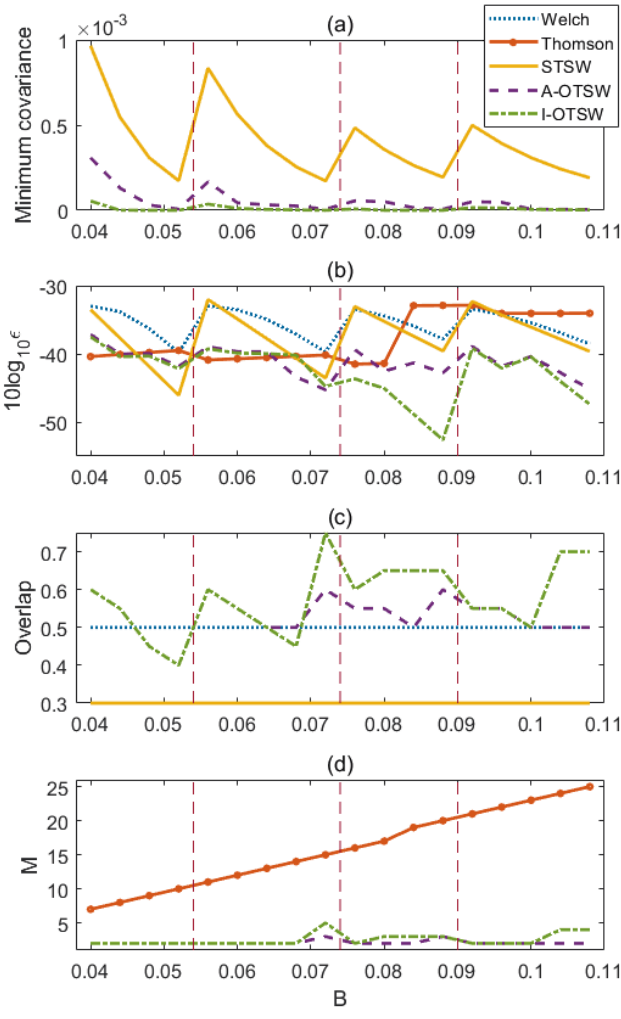



**Figure 2:** Evaluation results for $N = 256$, different frequency bands $B$ and corresponding $K$ according to Table 1; (a) minimum covariances; (b) spectral leakage in dB; (c) optimal overlap; (d) number of DPSS basis functions.

The results of the evaluation are shown in Figure 2, (a) shows the minimum covariances for the STSW, I-OTSW and A-OTSW estimators, which shows how close to the smallest possible variance $1/K$ the methods reach. The covariances for the Welch estimator are much larger than for the other methods,

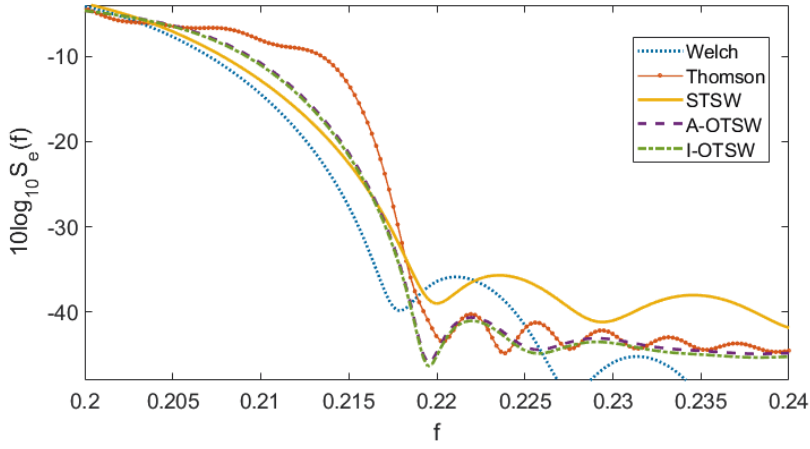



**Figure 3:** Example of spectral estimates of low-pass filtered white Gaussian noise with cut-off frequency $f = 0.2$.

around 0.01 for all $K$. The Thomson estimator has covariance zero, however the minimum covariances for the I-OTSW and A-OTSW estimators are consistently very low and close to zero. Figure 2(b) shows that the leakage is low for both the I-OTSW and A-OTSW estimators, often under -40 dB, and (c) shows that the overlap is mostly 55% for the A-OTSW and slightly higher for the I-OTSW. This means that most often $K_C = 2$ and $K_C \leq 4$, indicating low complexity computations. In Figure 2(d) it can be seen that for the I-OTSW and A-OTSW estimators most often $M = 2, 3$, also indicating low complexity calculations, this is compared to the Thomson estimator that uses $7 - 25$ basis functions.

### 5.1 Low-pass filtered noise example

Low-pass filtered white Gaussian noise, $\mathcal{N}(0, 1)$, is used to visualise the performance of the estimators on spectra with large dynamics. The spectral estimations of a filtered noise realisation, cut-off frequency $f = 0.2$ and $N = 256$, are shown in Figure 3. It can be seen that, the Welch estimator has the most narrow main lobe, with the I-OTSW, A-OTSW and STSW estimators almost as narrow. However, the I-OTSW, A-OTSW and Thomson estimators have the lowest first side lobes. In this example $B = 0.04$, $K = 4$ and the optimal overlaps presented in Figure 2(c) are used.

## 6 Conclusion

Two methods, I-OTSW and A-OTSW, for finding the optimal shape of a time-shifted window have been presented. The window function, typically not positive or even, is a weighted sum of the DPSS, where the number of basis functions is determined by a tolerated leakage and the weights are optimised to minimise the variance of a white Gaussian noise process. Using a novel procedure to determine the optimal overlap of the time-shifted windows, the resulting optimal spectral estimators are shown to give close to optimal variance reduction and low leakage, (side lobes -40 dB). Both methods outperform the usual Welch method, using common windows, with regards to variance and leakage. Since the methods use few DPSS, low overlap and time-shifted windows, they are more computationally efficient than the Thomson estimator. The methods however still have the appealing quality of a predetermined frequency resolution, usually associated with the Thomson estimator. It is also shown that the two novel methods perform very similar to each other, making the simplified A-OTSW estimator highly beneficial.

## References

[1] P. D. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. on Audio Electroacoustics*, vol. AU-15, no. 2, pp. 70–73, June 1967.

[2] B.Babadi and E. N. Brown, "A review of multitaper spectral analysis," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 5, pp. 1555–1564, 2014.

[3] K. Barbé, R. Pintelon, and J. Schoukens, "Welch method revisited: Nonparametric power spectrum estimation via circular overlap," *IEEE Transations on Signal Processing*, vol. 58, no. 2, pp. 553–565, Sept 2010.

[4] F. Attivissimo, M. Savino, and A. Trotta, "Power spectral density estimation via overlapping nonlinear averaging.," *IEEE Transactions on Instrumentation and Measurement*, vol. 50, no. 5, pp. 1418 – 1424, 2001.

[5] G. C. Carter and A. H. Nuttall, "On the weighted overlapped segment averaging method for power spectral estimation," *Proc. of the IEEE*, vol. 68, no. 10, pp. 1352–1354, Oct 1980.

[6] A. H. Nuttall and G. C. Carter, "A generalized framework for power spectral estimation," *IEEE Trans. on ASSP*, vol. ASSP-28, no. 3, pp. 334–335, June 1980.

[7] A. H. Nuttall, "Some windows with very good sidelobe behavior," *IEEE Trans. on ASSP*, vol. ASSP-29, no. 1, pp. 84–91, Feb 1981.

[8] F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proc. of the IEEE*, vol. 66, no. 1, pp. 51–83, Jan 1978.

[9] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis and uncertainty-V: The discrete case," *Bell System Journal*, vol. 57, no. 5, pp. 1371–1430, May-June 1978.

[10] D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proc. of the IEEE*, vol. 70, no. 9, pp. 1055–1096, Sept 1982.

[11] J. Antoni and J. Schoukens, "Optimal settings for measuring frequency response functions with weighted overlapped segment averaging," *IEEE Trans. on Instrumentation and Measurement*, vol. 58, no. 9, pp. 3276–3287, Sept 2009.

[12] P. A. Treverso P. Magnone and C. Fiegna, "Experimental technique for the performance evaluation and optimization of 1/f noise spectrum investigation in electron devices," *Measurement*, vol. 98, pp. 421–428, 2017.

[13] T. P. Bronez, "On the performance advantage of multitaper spectral analysis," *IEEE Transactions on Signal Processing*, vol. 40, no. 12, pp. 2941–2946, Dec 1992.

[14] M. Hansson-Sandsten, "A Welch method approximation of the Thomson multitaper spectrum estimator," in *European Signal Processing Conference (EUSIPCO)*, Bucharest, Romania, 2012.

# Paper B

**Paper B**

# Multitaper Reassignment for Oscillating Transients with Gaussian Envelopes

Isabella Reinhold and Maria Sandsten

*Mathematical Statistics, Centre for Mathematical Sciences, Lund University, Sweden.*

**This manuscript is not included in the electronic copy of this thesis.**

**Keywords:** Hermite functions, Instantaneous frequency estimation, Reassignment method, Signal detection, Time-frequency analysis

# Paper C

**Paper C**

# Optimal Time-Frequency Distributions using a Novel Signal Adaptive Method for Automatic Component Detection

Isabella Reinhold and Maria Sandsten

*Mathematical Statistics, Centre for Mathematical Sciences, Lund University, Sweden.*

**Abstract**

Finding objective methods for assessing the performance of time-frequency distributions (TFD) of measured multi-component signals is not trivial. An optimal TFD should have well resolved signal components (auto-terms) and well suppressed cross-terms. This paper presents a novel signal adaptive method, which is shown to have better performance than the existing method, of automatically detecting the signal components for TFD time instants of two-component signals. The method can be used together with a performance measure to receive automatic and objective performance measures for different TFDs, which allows for an optimal TFD to be obtained. The new method is especially useful for signals including auto-terms of unequal amplitudes and non-linear frequency modulation. The method is evaluated and compared to the existing method, for finding the optimal parameters of the modified B-distribution. The performance is also shown for an example set of Heart Rate Variability (HRV) signals.

**Keywords:** Detection, Heart Rate Variability, Multi-component signal, Performance measure, Time-frequency

## 1 Introduction

There are many types of non-stationary signals, most of which are multi-component. These signals need to be visualised in time and frequency simultaneously to characterise their time-varying nature. To do this the distribution of the signal energy over the time-frequency plane, i.e. the time-frequency distribution (TFD), can be studied.

The Wigner-Ville distribution (WVD) is a common TFD. For mono-component, linear frequency modulated (FM) signals the WVD gives exactly the instantaneous frequency (IF) making it the optimal TFD for such signals. The problem with the WVD occurs when dealing with multi-component signals or signals disturbed by noise. For such a signal the WVD is not always zero when the signal has no power for a given time-frequency instant. These contributions are called cross-terms and can have twice the amplitude of the signal components. This makes it difficult to distinguish the actual signal components, also called auto-terms, from the cross-terms [1].

There exist many TFDs which aim to suppress cross-terms by means of filtering the WVD with a kernel, such as Choi-Williams, Zhao-Atlas-Marks [1] and modified B-distribution [2]. However suppression of the cross-terms can also result in loss of resolution of the signal components. Finding good representations of multi-component signals is a complex problem and is still a large field of research [3, 4]. When looking at different TFDs for multi-component signals it might be possible to say that some plots look cleaner and thus better. However, assessing the performance based only on this visual comparison is very subjective and finding the optimal parameter for a specific kernel would be very tiresome if not impossible. Not many methods exist, for assessing which TFD is the best for a given signal, especially when dealing with measured signals.

A quantitative performance measure for TFDs of two-component signals, called normalised instantaneous resolution (NIR) performance measure, was presented in [5]. The NIR performance measure makes it possible to compare different TFDs and optimise kernel parameters which control the tradeoff between signal component resolution and cross-term suppression. The NIR performance measure can be used for simulated as well as measured signals and was recently used in [6, 7] to find optimal TFDs for different multi-component signals. However, the measure relies on parameters connected to correct detection of the signal components for each time instant of the TFD, and the method used for

automatic detection of auto-terms is the one presented in [8]. One restriction of this method is the requirement that the amplitudes of the two signal components are (approximately) equal, which is an assumption that limits the use of the method. The method also fails when signal components are close to each other or has components with non-linear FM law [9], which is a well known restriction of many methods [3]. These restrictions in the detection method narrows the use of the NIR performance measure as the choice of analysed kernel parameters needs to be made with care. This limits the use of the performance measure for automatic optimisation of signal adaptive kernels, compared to other methods such as [10]. A large number of methods for identification of signal components exist, e.g. [11–13], who require that cross-terms are well suppressed and locates the maximum peaks as signal components. Other methods such as [14] which uses a method called non-linear squeezing time-frequency transform exist as well. However, these methods require already optimised or semi-optimised TFDs or are more complex and computationally heavy.

This paper presents a new signal adaptive method for automatically detecting signal components in two-component signals which outperforms the method in [8]. The new method is not limited by requiring that the signal components have equal amplitudes. Additionally, the method succeeds in detecting components with non-linear FM laws. Further, the new algorithm overcomes one of the main drawbacks when the estimated parameters are used in the NIR performance measure, as it successfully identifies auto-terms for a larger interval of kernel parameters, allowing for a more objective kernel optimisation. It is also feasible that for two-component signals the new automatic detection method can be used to find the direction of the auto-terms which is used to create an adaptive directional kernel [15, 16]. This kernel smooths at each point in the time-frequency domain based on the direction of the energy distribution of the signal.

To illustrate the use of the new signal adaptive automatic detection method together with the NIR performance measure, this paper shows how the optimal kernel parameters for the modified B-distribution [2] of an example set of Heart Rate Variability (HRV) signals with a non-linear component can be obtained. HRV, which is the variation of inter-heartbeat intervals, is measured non-invasively using ECG. It provides information on the autonomic regulation of the cardiovascular system. This means that the HRV is a sensitive indicator of compromised health [17, 18]. The HRV has a non-stationary nature, however

only recently methods which do not assume stationarity have been evaluated for HRV [19, 20]. It is common to study HRV during treadmill running [21, 22], making the need for methods of studying HRV in time and frequency concurrently even more important.

The paper is organised as follows. Section 2 provides an introduction to the basics of time-frequency analysis. Section 3 shortly presents the NIR performance measure which will be used and details the new signal adaptive method for automatic detection of the signal components. In Section 4 the performance of the new automatic detection method is evaluated and compared to the performance of the method in [8]. The basis for the evaluation is simulated signals and an example set of HRV signals. The optimal modified B-distributions of the example HRV signals are presented in Section 5. Sections 6 and 7 finish the paper with discussion and conclusions.

## 2 Time-frequency methods

The Wigner-Ville distribution (WVD),

$$W_z(t,f) = \int_{-\infty}^{\infty} z\left(t+\tfrac{\tau}{2}\right) z^*\left(t-\tfrac{\tau}{2}\right) e^{-i2\pi f\tau} d\tau, \tag{1}$$

where $^*$ represents the complex conjugate, is a TFD defined using an analytic signal, $z(t)$. The analytic signal is defined such that $Z(f) = 0$ if $f < 0$, where $Z(f) = \mathcal{F}\{z(t)\}$, is the Fourier transform of the signal. The quadratic class of TFDs, a subclass of TFDs where the signal kernel is of quadratic form, can be written as

$$\rho_z(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(t-u,\tau) z\left(u+\tfrac{\tau}{2}\right) z^*\left(u-\tfrac{\tau}{2}\right) e^{-i2\pi f\tau} du d\tau, \tag{2}$$

where the time-lag kernel $G(t,\tau)$ is specific for each different quadratic TFD. The convolution of the kernel in Eq. (2) is (in most cases) equal to a 2D filtering of the TFD and is used to suppress cross-terms. The design of the kernels is usually done in the ambiguity (doppler-lag) domain, where auto- and cross-terms are more easily differentiable [1].

### 2.1 Separable and lag-independent kernels

One simple, yet useful, class of kernels is the separable kernels. With the separable kernel the TFD can be written

$$\rho_z(t,f) = g_1(t) \underset{t}{*} W_z(t,f) \underset{f}{*} G_2(f). \tag{3}$$

The convolutions in time and frequency can now be made in either order which simplifies the calculations. It also means that the design of the kernel will be greatly simplified, the 2D filtering operation is replaced by two consecutive 1D filtering operations. A special case of the separable kernel is the lag-independent (LID) kernel. It is obtained by setting

$$G_2(f) = \delta(f), \tag{4}$$

which means that the kernel only will depend on time $t$. The calculations for the TFD then only require one convolution, in the time direction only

$$\rho_z(t,f) = g_1(t) \underset{t}{*} W_z(t,f). \tag{5}$$

Since the LID kernel only applies one 1D filtering, the resulting TFD will be smoothed in the time direction only. This property makes the LID kernel suitable for slowly varying frequency modulated signals or other signals where cross-terms exist mainly some frequency distance from the auto-terms and single auto-terms do not vary much in frequency. LID-TFDs have been shown to have better performance in characterising HRV signals, compared to other time-frequency methods [23]. The LID kernel can have different distributions, one is the modified B-distribution (MBD), which has been shown to be suitable for HRV signals [24]. The MBD kernel is defined as

$$g_{\mathrm{MBD}}(t) = \frac{\cosh^{-2\beta}(t)}{\int_{-\infty}^{\infty} \cosh^{-2\beta}(\xi) d\xi}, \tag{6}$$

where $\beta$ is the scaling parameter which determines the trade-off between resolution of signal components and cross-term suppression. The MBD, designed specifically for multi-component IF estimation, is almost cross-term free and has high resolution of signal components in the time-frequency plane [2].

## 3 Performance measure and a novel signal adaptive method for automatic detection of auto-terms

The NIR performance measure, which combines the concepts of high energy concentration around the IF laws and clearly resolved signal components is presented in [5]. The measure doesn't take into account some properties usually demanded for TFDs, which impose strict constraints on the TFD design, such as satisfying the marginals [1]. Instead it focuses on resolution of signal components and suppression of cross-terms and sidelobes, which are important for practical use. The measure is defined as

$$P(t) = 1 - \frac{1}{3}\left(\frac{A_S(t)}{A_M(t)} + \frac{1}{2}\frac{A_X(t)}{A_M(t)} + (1 - D(t))\right), \quad 0 \leq P(t) \leq 1, \tag{7}$$

where $A_S(t)$ is the average absolute amplitude of the largest sidelobes, $A_M(t)$ the average amplitude of the auto-terms (mainlobes), $A_X(t)$ the absolute amplitude of the cross-term and $D(t)$ a measure of the separation of the signal components' mainlobes. It is defined as

$$D(t) = \frac{\left(f_2(t) - \frac{V_2(t)}{2}\right) - \left(f_1(t) - \frac{V_1(t)}{2}\right)}{f_2(t) - f_1(t)}, \tag{8}$$

where $f_1(t)$ and $f_2(t)$ are the centres of the mainlobes and $V_1(t)$ and $V_2(t)$ are the instantaneous bandwidths of the auto-terms, calculated at $\sqrt{2}/2$ of the height of the mainlobe.

For this measure a value close to 1 is a good performance. The performance measure is calculated for a time instant (slice) of the TFD. If $P(t)$ is calculated for several time instants, an estimate of the performance for the whole TFD can be formed [5]. This measure works well for signals with both linear and non-linear FM components [8, 9]. The only restriction is that the signal should have only two components where the performance measure is calculated.

### 3.1 A novel signal adaptive method for automatic detection of auto-terms

In order to use the resolution performance measure on signals there is a need for a signal adaptive method which automatically detects auto-, cross-terms and sidelobes for a TFD time slice. Such a method for two-component signals is

proposed by Sucic et al. [8]. However, the difficulty lies within detecting the auto-terms and a restriction is the assumption that the signal components have equal amplitudes. The algorithm for Sucic's automatic detection of auto-terms (ADAT) follows these steps:

1. Normalise the time slice such that the absolute maximum is equal to 1.
2. Determine the three largest maxima (peaks) of the slice.
3. The cross-term is located between the auto-terms, so initially set the middle peak to be the cross-term and the remaining as auto-terms.
4. Make sure that the ratio between the amplitudes of remaining two peaks is close to 1, and that the peak chosen as the cross-terms is close to the middle point between the centres of the other two peaks. This checks whether the assumption in the previous step is correct. If not, select the two largest peaks of the slice as the auto-terms.

This method is simple and does in many cases successfully identify the auto-terms. However, the requirement that the amplitudes of the two signal components are (approximately) equal limits its use. Another drawback is that the method has a degraded performance for signals containing components with non-linear frequency modulated (FM) law [9]. The novel method presented here does not require the signal component amplitudes to be equal, instead it relies on sidelobes and noise peaks of restricted amplitudes. The steps of the novel Reinhold's ADAT algorithm are:

1. Normalise the time slice such that the absolute maximum is equal to 1.
2. Determine an amplitude threshold, $\lambda$, for the auto-terms.
3. Determine between which frequencies all peaks above $\lambda$ are located. This is the estimated frequency distance between the auto-terms, $\widehat{\Delta f_a}$. Set $\delta \approx \widehat{\Delta f_a}/2$ as the minimum allowed frequency distance between the auto- and cross-terms.
4. Identify the largest peaks above the threshold $\lambda$, which are separated with at least $\delta$. These are the only peaks which will be considered when identifying auto-terms. Select the peaks furthest away from each other as the auto-terms.

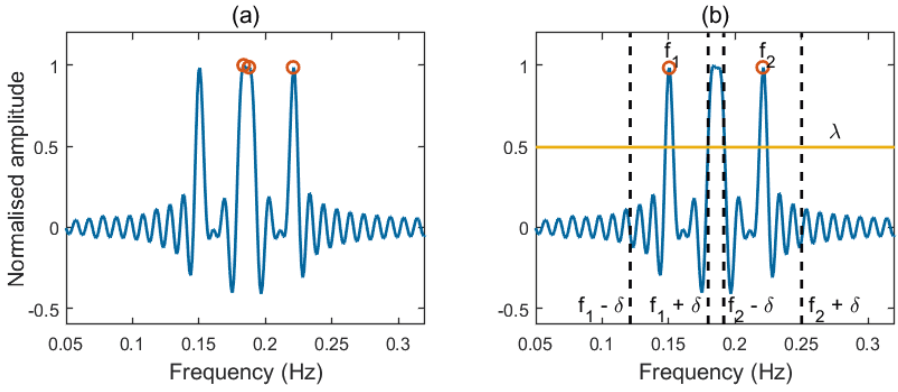



**Figure 1:** Time slice of the WVD of a two-component linear FM signal, with IFs 0.15 and 0.22; (a) Potential auto-terms detected with Sucic's ADAT, marked with circles; (b) Detected auto-terms with Reinhold's ADAT, the auto-terms are marked by circles and the labels $f_1$ and $f_2$ respectively. The horizontal line shows the threshold $\lambda$ and the dashed lines are $\eth$ away from each identified auto-term.

The minimum distance $\eth$ is set as approximately half the estimated frequency distance between the auto-terms since theoretically within $\Delta f_a$ there should be three peaks, the two auto-terms and the cross-term. It is reasonable to choose it as $\eth = \widehat{\Delta f_a}/2 - \epsilon$, where $\epsilon$ is a small error tolerance. This allows for some error in the estimation of $\Delta f_a$ and small deviations of the placement of the cross-term.

The selection of the parameter $\lambda$ could also be made automatically and should be allowed to vary with each time slice for optimal results. This paper proposes to let $\lambda = cA_2$, where $A_2$ is the amplitude of the second largest peak and $c$ is some scale factor, $0 < c < 1$. This means that $\lambda$ always will relate to the amplitudes of the signal content. For simulations in this paper $\lambda = 0.5A_2$, if nothing else is stated. Other choices of $c$ can be used and Section 4.1 will evaluate how robust the novel ADAT is to different $\lambda$s.

### 3.2 Motivation for new automatic detection method

This section will motivate the need for a new ADAT by studying two example time slices from TFDs of two-component signals. The examples demonstrate situations when Sucic's ADAT fails to correctly detect the auto-terms, whereas Reinhold's ADAT is successful. The first example is a time slice of a WVD of a

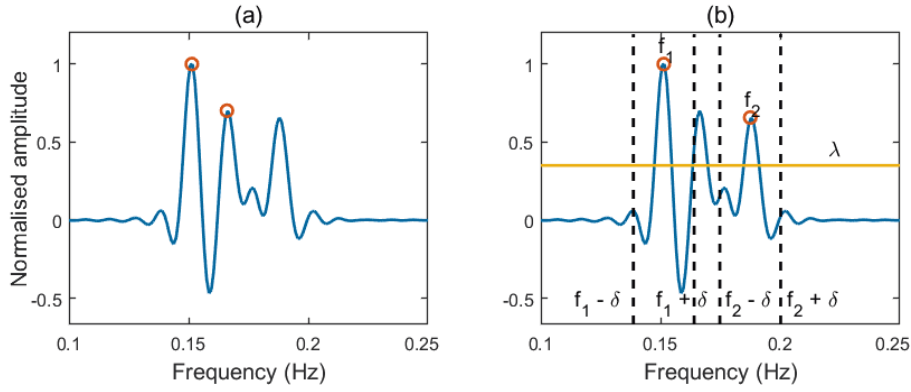



**Figure 2:** Time slice of a LID-TFD of a two-component non-linear FM signal, with IFs 0.15 and 0.19; (a) Auto-terms detected by Sucic's ADAT, marked by circles; (b) Auto-terms detected by Reinhold's ADAT, marked by circles and the labels $f_1$ and $f_2$ respectively. The horizontal line shows the threshold $\lambda$ and the dashed lines are $\delta$ away from each identified auto-term.

signal with components of equal amplitude. Sucic's ADAT initially identifies the three largest peaks, if these peaks are close in amplitude, two of them are identified as auto-terms. Figure 1(a) shows the identified three peaks. The ratio between any two of these peaks is close to one and thus two of the three peaks are identified as auto-terms, which two depend on implementation choices for the algorithm. However since the first auto-term is not among these peaks Sucic's ADAT fails to correctly identify the auto-terms.

In this example only four peaks are above the threshold $\lambda$, marked by a horizontal line in Figure 1(b). The peaks of maximum distance are initially identified as the auto-terms, which gives an estimate of $\Delta f_a$. Reinhold's ADAT will then detect three peaks above $\lambda$ which are separated by at least $\delta$, the auto-terms and one of the dual peaks of the cross-term. The distance between the dual peaks of the cross-term is (much) smaller than $\delta$, hence only one of the peaks are detected. Of the three detected peaks, the two at maximum distance are finally identified as the auto-terms. These two are the actual auto-terms and they are marked in the figure by the labels $f_1$ and $f_2$. The figure also shows, in dashed lines, $f_1 \pm \delta$ and $f_2 \pm \delta$.

The second example is a time slice from a LID-TFD of a signal with non-linear FM law for one component, where the signal components have equal amplitude. However, for some time slices of the TFD, there will be poor energy concentration

around the non-linear component's IF law, which will result in differences in the amplitudes of the auto-terms. According to Sucic's ADAT, after identification of the three largest peaks, if the ratio between the two outer peaks is not close to one, the two largest peaks are chosen as auto-terms. Figure 2(a) shows how these steps identifies the wrong peaks as auto-terms.

Reinhold's ADAT correctly identifies the auto-terms of this time slice, which is shown in Figure 2(b). As seen in the figure, only three peaks are above the threshold $\lambda$, the outermost are the auto-terms and those are identified as auto-terms. Figure 2(b) also shows $f_1 \pm \delta$ and $f_2 \pm \delta$.

### 3.3 Signals with more than two components

The purpose of the NIR performance measure is to resolve two components which are close in frequency, however it can still be interesting to resolve close components in a signal with more than two components. There are two kinds of signals which are particularly interesting to consider, both has three components, however one has only two components present in the signal at any given time. For such a signal there is no theoretical problem using the ADAT algorithms, since the signal's TFD is analysed at each time instant. Therefore it does not matter how long time duration signal components have or how many signal components the signal has, as long as there are at maximum two for any given time instant.

An example of the other type of three component signal is defined by

$$\begin{aligned} s(n) = &\cos\left(2\pi\left(0.15 + 0.04\left(\tfrac{n}{256}\right)\right)n\right) + \cos\left(2\pi\left(0.24 - 0.04\left(\tfrac{n}{256}\right)\right)n\right) \\ &+ \cos(2\pi 0.3n), \quad 0 < n \leq 256, \end{aligned} \tag{9}$$

and shown in Figure 3, it has three components which are all present at the same time. This signal presents a problem for both ADAT methods, since they are designed to find only two auto-terms separated by some frequency distance in each time instant, and both methods usually fail to identify the desired auto-terms for such a signal.

However the NIR performance measure and ADAT algorithms can still be used for such signals if only the possibly known frequency band containing the two close components are considered. The ADAT methods will then only be applied to that frequency band and this requires omitting parts of the TFD beforehand.

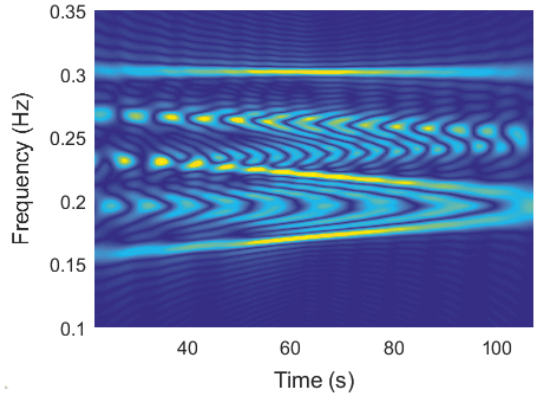



**Figure 3:** MBD of the signal Eq. (9), which is a three component signal with two close components and one further away.

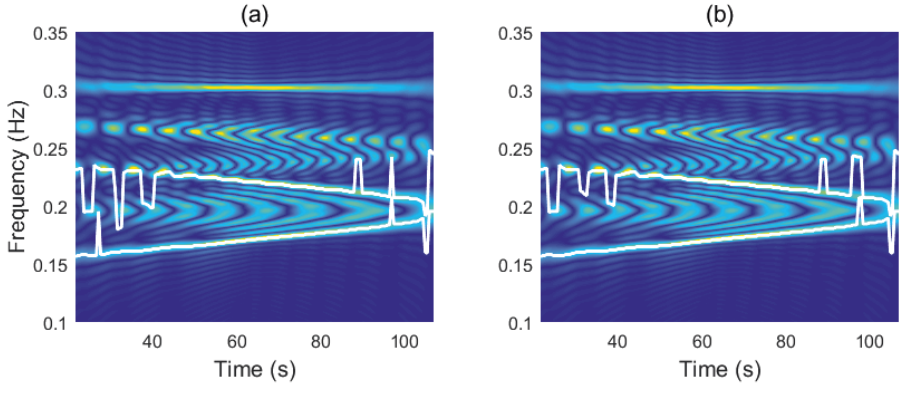



**Figure 4:** Detection of auto-terms in the example three-component signal Eq. (9), when only considering the frequency band 0 - 0.25 in the ADAT algorithms; (a) Sucic's ADAT; (b) Reinhold's ADAT.

In this example the NIR performance is calculated in the frequency band 0 - 0.25. Figures 4(a) and (b) shows the detected auto-terms on top of the TFDs for Sucic's and Reinhold's ADAT respectively. The methods perform similarly and identifies the correct auto-terms for most time instants. The added disturbance

of more cross-terms can decrease the performance of both ADAT algorithms and some initial filtering with an appropriate kernel might be needed to suppress cross-terms. An indication that filtering by a kernel is needed is if the auto-terms are hard to distinguish by eye. The suggestion is then to apply a wide kernel to get a TFD which closely resembles the WVD, keeping the resolution of the auto-terms, but with cross-terms slightly suppressed. In this example a wide MBD-kernel ($\beta = 0.5$) is used, the performance of both methods increase if the kernel is more narrow. If the kernel is instead made more wide, the middle auto-term will be very hard to distinguish from the cross-term which intercepts it.

## 4 Detection of auto-terms

To compare the performance of Sucic's and Reinhold's ADAT, the rate of detection for different two-component signals and TFDs have been studied. The rate of detection is measured by checking if the detected auto-terms are close to the signal component IFs in each time slice of the TFD. The detection for each time instant can be either successful or unsuccessful. If the frequencies of both the detected auto-terms vary no more than $\Delta f_a/4$ from the respective signal component IF, the detection is called successful. Every successful detection yields a value 1 and every unsuccessful a 0. The detection results for each time instant of the TFD is then added together and the sum is divided by the number of time instants, this gives a rate of detection in the interval $[0, 1]$ for the whole signal.

In this section the rate of detection will be examined for signals with additive Gaussian white noise, with signal-to-noise ratio (SNR) 5 dB. The rate of detection for a given signal will vary with different noise simulations. Thus to give an accurate description of the rate of detection, 500 different noise simulations will be used to find the average rate of detection for a given signal and kernel parameter. The lower bound of the one sided confidence interval with 5% significance is also presented.

When computing the TFD of a (finite) signal, there will be some effects around the edges, in time and frequency. In this section, the middle 2/3 time slices of the TFD will be evaluated when calculating the rate of detection, where the initial and end time slices are ignored. The simulated signals are 256 samples, which gives $256 \cdot 2/3 = 172$ evaluated time instants for each TFD.

**Table 1:** Setup parameters for calculation of rate of detection. Parameters $a_0$, $f_0$, $f_I$ and $k$ are the variable amplitude, starting frequency, frequency increase and factor for the signal in Eq. (10). $\beta$ is the scaling parameter for the MBD kernel Eq. (6).

| Setup | $a_0$ | $f_0$ | $f_I$ | $k$ | $\beta$ |
|---|---|---|---|---|---|
| 1 | $[0.6, 1.4]$ | 0.17 | 0.06 | 1 | 0.30 |
| 2 | 1.0 | $[0.16, 0.21]$ | 0.04 | 1 | 0.50 |
| 3 | 1.0 | 0.19 | 0.07 | 3 | $[0.08, 0.20]$ |

Three different setups of TFDs will be evaluated, the parameters for these are shown in Table 1. The parameters refer to the general signal

$$s(n) = a_0 \cos(2\pi 0.15n) + \cos\left(2\pi\left(f_0 + f_I\left(\tfrac{n}{256}\right)^k\right)n\right) + e(n), \quad 0 < n \leq 256, \tag{10}$$

where $e(n)$ is stationary Gaussian white noise, and to the kernel parameter, $\beta$, of the MBD kernel in Eq. (6).

The first setup calculates the rate of detection when the components have linear FM laws. The amplitude of one of the components is varied, in accordance to Table 1. The kernel parameter is chosen so that the cross-term and noise peaks are slightly suppressed, whilst the signal components should be relatively unaffected by the filtering. Figures 5(a) and 5(b) show two examples of the evaluated MBDs of signals with the smallest ($a_0 = 0.6$) and largest ($a_0 = 1.4$) component amplitudes. The results for Sucic's and Renhold's ADAT are presented in Figure 6(a) and it can be seen that Reinhold's ADAT performs better than Sucic's for all values of $a_0$. In fact the lower bound of Reinhold's ADAT is in all cases higher than the average detection rate for Sucic's.

Sucic's ADAT performs poorly for low and high $a_0$, i.e. when the difference in amplitude of auto-terms is significant, which is in accordance of the results presented in [9]. Especially notable is, that the rate of detection is as low as around 0.50 for $a_0 = 0.6$, this would make any evaluation of the performance using Eq. (7) very unreliable. Reinhold's ADAT however has a detection rate of around 0.75 for the same amplitude, which although not perfect is considerably better.

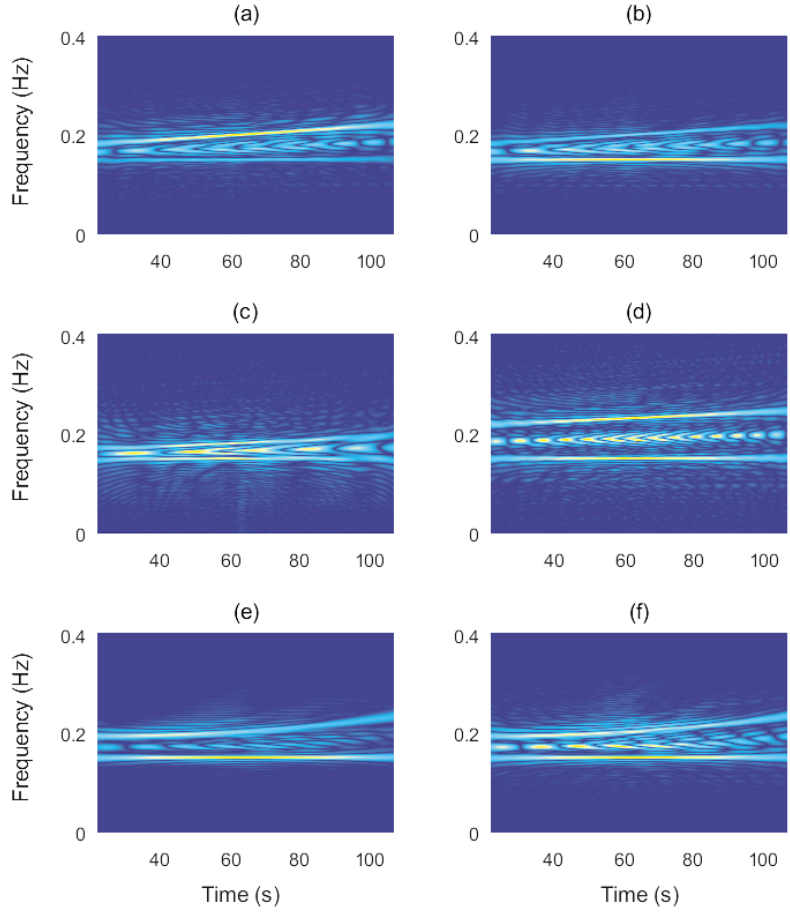



**Figure 5:** Some MBDs of the signal in Eq. (10) with parameters according to the setups in Table 1 evaluated when calculating the rate of detection. The figures show the part of the distribution which is assessed; (a) Setup 1 with $a_0 = 0.6$; (b) Setup 1 with $a_0 = 1.4$; (c) Setup 2 with $f_0 = 0.16$; (d) Setup 2 with $f_0 = 0.21$; (e) Setup 3 with $\beta = 0.08$; (f) Setup 3 with $\beta = 0.20$.

The second setup, Table 1, varies the frequency distance of the two components of the signal. The kernel parameter for this setup is chosen large so that much of the cross-term and noise remains, making the auto-term detection difficult. This setup will thus show how the two methods perform for quite challenging

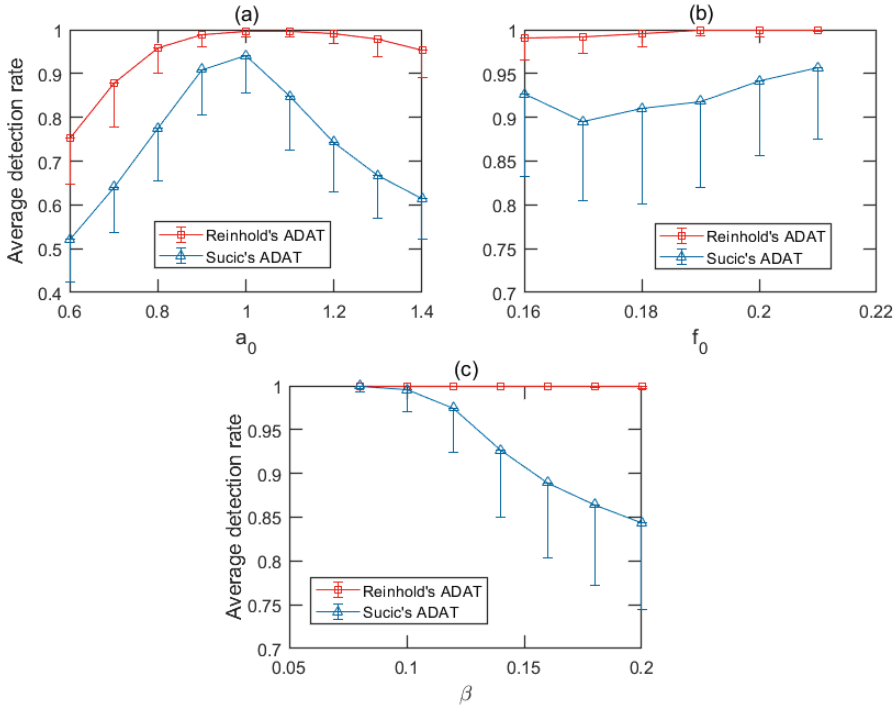



**Figure 6:** The average rate of detection for the ADAT algorithms according to the three setups in Table 1. The lower bound is a one sided confidence interval with 5% significance. The results are obtained from 500 simulations with different realisations of stationary Gaussian white noise with SNR 5 dB; (a) Setup 1; (b) Setup 2; (c) Setup 3.

TFDs. Figures 5(c) and 5(d) show the MBD when the signal components are closest together ($f_0 = 0.16$) and furthest apart ($f_0 = 0.21$). In Figure 6(b) it can be seen that both ADAT methods have the highest rate of detection when the signal components are furthest apart and the average detection rate is rather high. However Reinhold's ADAT still outperforms Sucic's, again the lower bound for Reinhold's ADAT is higher than the average value for Sucic's ADAT.

The third setup, Table 1, uses the same signal and instead varies the scaling parameter $\beta$ of the MBD kernel, making this test different from the other two. The signal has one component with a non-linear FM law. A MBD of the signal with the smallest scaling parameter ($\beta = 0.08$) is shown in Figure 5(e), in this TFD much of the noise and cross-term have been suppressed. Figure 5(f) shows

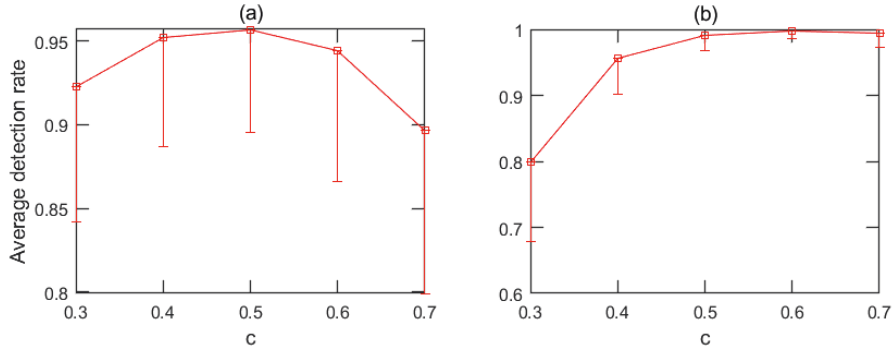



**Figure 7:** The average rate of detection for Reinhold's ADAT according to the setups in Table 1. The lower bound is a one sided confidence interval with 5% significance. The results are obtained from 500 simulations with different realisations of stationary Gaussian white noise with SNR 5 dB; (a) Setup 1 with $a_0 = 0.8$; (b) Setup 2 with $f_0 = 0.16$.

a MBD with the largest scaling parameter ($\beta = 0.20$), this TFD has some noise peaks and quite high cross-terms. Figure 6(c) shows the resulting average rates of detection for Sucic's and Reinhold's ADAT. It can be seen that the performance of Sucic's ADAT decreases when $\beta$ increases, however for Reinhold's the average rate of detection is 1 for all evaluated $\beta$. This suggests that Reinhold's ADAT is robust to different levels of filtering of the TFD.

### 4.1 Robustness to choices of amplitude threshold

The amplitude threshold used for Reinhold's ADAT in this paper is $\lambda = 0.5A_2$, where $A_2$ is the amplitude of the second largest peak. This section evaluates how robust Reinhold's ADAT is to other choices of $\lambda$. Different thresholds are tested by letting $\lambda = cA_2$, $0 < c < 1$, and letting the scale factor $c$ vary.

To evaluate the robustness, the average rate of detection is calculated for 500 simulations of the signal Eq. (10), using the first setup with $a_0 = 0.8$ and the second setup with $f_0 = 0.16$, with Gaussian white noise, SNR 5 dB. The results for the signal with different amplitudes of the signal components are shown in Figure 7(a), all average detection rates are around 0.9 or higher, which can be compared to the average detection rate for the same setup using Sucic's ADAT which is under 0.8.

The results for the signal with close components and large kernel parameter are shown in Figure 7(b). It can be seen that choosing $c = 0.3$ gives the worst performance for this signal, which is not surprising since there are much high-amplitude disturbance in the signal and a low $\lambda$ would allow such peaks to be identified as auto-terms. The average rate of detection for Sucic's ADAT of the same signal is almost 0.95, which is higher than for Reinhold's when $c = 0.3$, however for other choices of $c$ the performance of Reinhold's ADAT is equivalent or superior.

## 4.2 Detection of auto-terms on real HRV data examples

The strength of the NIR performance measure Eq. (7) is that it can be used to asses the performance of different TFDs of measured signals [5]. However, for the performance measure to be as accurate as possible, the signal adaptive method for automatic detection of the auto-terms need to detect the correct IFs of the signal for as many time instants of the TFD as possible. When using measured signals, the signal IFs are unknown, which makes the rate of detection more difficult to calculate. This section will instead show the detected auto-terms on top of the TFDs.

The signals in this section is the Heart Rate Variability (HRV) signals from adult humans which have been asked to breathe with the same frequency as a metronome. The frequency of the metronome was increased non-linearly over time, and thus the breathing frequency is increased non-linearly with time. This gives a HRV signal with two components, one assumed stationary low frequency (LF) and and one high frequency (HF) with non-linear FM law, approximately following the breathing frequency.

Figure 8 shows the MBDs, $\beta = 0.08$, of four HRV signals and the detected auto-terms using Sucic's and Reinhold's ADAT. Figures 8(a), 8(c), 8(e) and 8(g) show the detected auto-terms using Sucic's ADAT. Figures 8(b), 8(d), 8(f) and 8(h) show the detected auto-terms using Reinhold's ADAT.

As seen in Figures 8(a) and 8(b) both methods fail to identify the auto-terms when the stationary signal component is corrupted by much noise, around $t \in [80, 120]$. The noise peaks in this region have high amplitudes and are close to the signal component in frequency, so identification is expected to be hard. However Reinhold's ADAT performs over all much better.

In Figures 8(c) and 8(d) it can be seen that Sucic's ADAT fails much more than Reinhold's. This is because the stationary signal components has a low amplitude compared to the non-stationary component. Figure 8(e) shows that Sucic's ADAT detects the incorrect peaks as auto-terms when the (almost) stationary signal component is noisy, at $t \in [100, 140]$. Reinhold's ADAT however identifies the correct peaks as auto-terms, see Figure 8(f), this is because the noise peaks around the (almost) stationary signal component are close in frequency to the actual IF of the component, the least distance allowed between peaks, $\delta$, is large enough to avoid these peaks being identified as auto-terms.

The non-stationary signal component in Figures 8(g) and 8(h) seems to have strong inner artifacts [1] at $t \in [30, 80]$, i.e. peaks due to the non-linear frequency increase. This makes detection hard and both methods fail sometimes, however over all the performance of Reinhold's ADAT is much higher.

## 5 Optimal parameter estimation of kernels for HRV signals

A method for finding the optimal TFD for a given multi-component signal is presented in [5]. The basic steps are to define a set of criteria for comparison of TFDs, then define a quantitative measure for evaluating TFD performance based on these criteria. This quantitative measure can be the NIR performance measure Eq. (7). After choosing a measure the TFDs should be optimised to match the comparison criteria as close as possible. When looking at the MBD this means that one chooses an initial value of the kernel parameter $\beta$ and calculates the MBD for this. Then the performance is calculated for each time instant within some time interval of interest. The average of all instantaneous measures is the interval performance measure of the MBD for the given $\beta$. This procedure is repeated for an interval of $\beta$ with an appropriate length of the increments. The optimal kernel parameter $\beta$ is the one which gives the best interval performance measure. Other TFDs could be optimised in the same manner by comparing the NIR interval performance measures while varying one or several parameters connected to the relevant TFD.

When the optimal parameters has been found for several different TFDs, the TFDs can be compared. The TFD with the maximum interval performance measure is the optimal TFD for the given signal. In this section the optimal MBDs, obtained by the above described method using the NIR performance

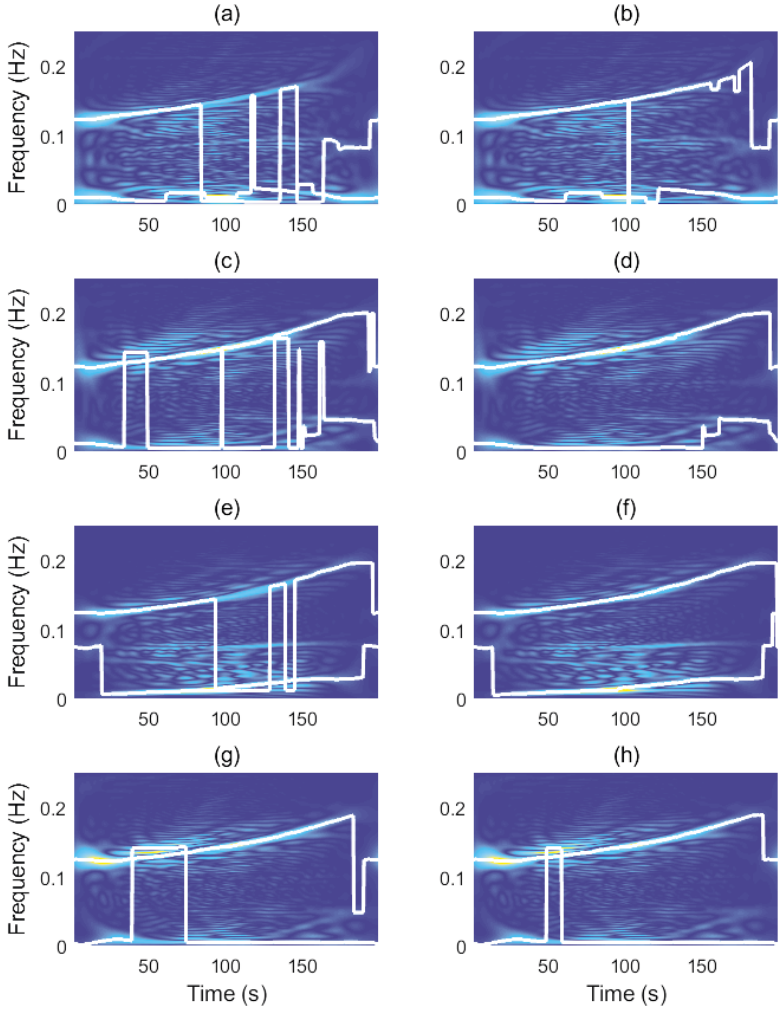



**Figure 8:** Detection of auto-terms in measured HRV using Sucic's and Reinhold's ADAT; (a) subject 1, Sucic's ADAT; (b) subject 1, Reinhold's ADAT; (c) subject 2, Sucic's ADAT; (d) subject 2, Reinhold's ADAT; (e) subject 3, Sucic's ADAT; (f) subject 3, Reinhold's ADAT; (g) subject 4, Sucic's ADAT; (h) subject 4, Reinhold's ADAT.

measure, will be found for the four example HRV signals presented in the previous section. The choice to optimise the MBD for the HRV signals is because this TFD has been shown suitable for HRV signals [24].

**Table 2:** The evaluated time intervals when finding the optimal MBDs for the HRV signals of four subjects, together with the resulting optimal kernel parameters and the interval performance measures.

| Subject | Time interval | Optimal parameter | Interval performance measure |
|---|---|---|---|
| 1 | $(0, 85]$ | $\beta = 0.053$ | 0.7860 |
| 2 | $(0, 120]$ | $\beta = 0.056$ | 0.8193 |
| 3 | $(25, 95]$ | $\beta = 0.047$ | 0.8035 |
| 4 | $(70, 170]$ | $\beta = 0.061$ | 0.8190 |

The interval performance measure is calculated for different time intervals for the four HRV signals using the measure in Eq. (7). This gives an interval performance measure $[0, 1]$, where 1 is optimal performance. The time intervals are chosen such that Reinhold's ADAT detects the correct auto-terms for each time instant for all the evaluated $\beta$. The time intervals for each subject is shown in Table 2, which also shows the resulting optimal parameters and interval performance measure. The optimal MBDs are shown in Figure 9.

Longer time intervals can be used when calculating the interval performance measure with Reinhold's ADAT compared to Sucic's, since it correctly identifies the auto-terms for more time instants and longer compact time intervals of the HRV signals. The resulting performance measure will thus more accurately describe the performance of the TFD, thus giving a more correct estimate of which parameter and corresponding TFD is the optimal.

## 6 Discussion

The novel Reinhold's ADAT presented in this paper relies less on the amplitudes of the auto-terms being equal compared to Sucic's ADAT presented in [8]. For Reinhold's ADAT to succeed in each TFD slice, $\lambda$ and $\delta$ need to fulfil the following:

- Outer peaks which have a distance larger than $\delta$ to their closest auto-term are smaller than $\lambda$.
- Peaks with less distance than $\delta$ to their closest auto-term have smaller amplitudes than that auto-term.

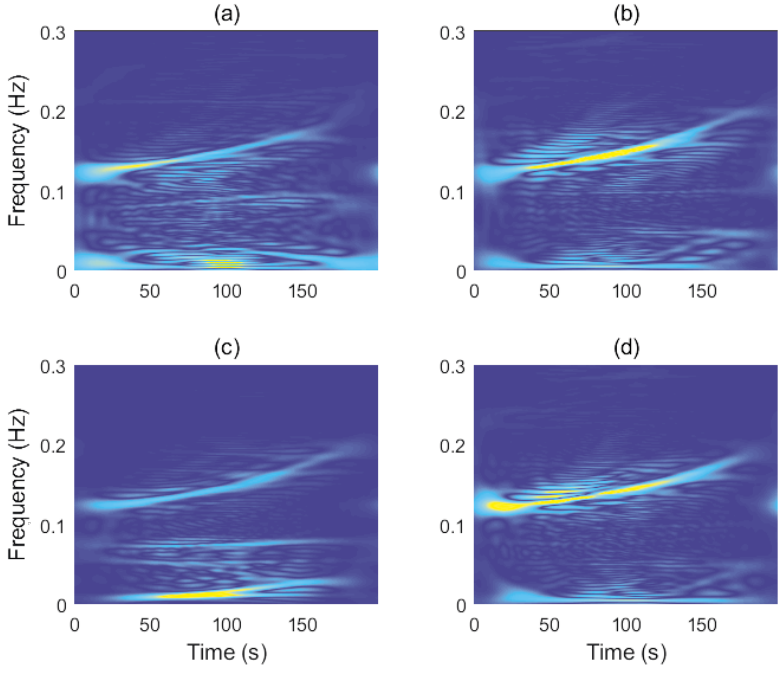



**Figure 9:** Optimal MBD for four examples of HRV signals; (a) subject 1; (b) subject 2; (c) subject 3; (d) subject 4.

- If the cross-term is the largest peak, the parameter $\delta$ is smaller than the actual distance between the cross-term and auto-terms.

It is therefore important that $\lambda$ and $\delta$ depend on the examined TFD time slice. It is reasonable to choose $\delta = \widehat{\Delta f_a}/2 - \epsilon$, where $\epsilon$ is a small error tolerance as discussed in Section 3.1. This paper suggests choosing $\lambda = cA_2$ and Section 4.1 shows that Reinhold's ADAT is robust for such $\lambda$s. The scale factor $c$ can be adapted to increase performance if there exist some knowledge of the signal, such as the relative amplitudes of the signal components or the abundance of noise peaks, however $c = 0.5$ is shown to give a good over all performance.

Reinhold's ADAT, as well as Sucic's ADAT, is designed to find two auto-terms some frequency distance apart in each time instant of a TFD. If a signal has more than two components, Reinhold's ADAT can be used with the NIR performance

measure to get good resolution between two signal components which are close. As discussed in Section 3.3, this requires cutting away the TFD which is outside the relevant frequency bandwidth, i.e. where the two close components are. This should be done before applying the ADAT algorithm and can be done manually, however an automatic and adaptive method to select the bandwidth is suggested for further research.

## 7 Conclusion

This paper presents a novel signal adaptive method for automatic detection of auto-terms in time slices of TFDs for two-components signals. This method performs better than the existing method for several types of signals and is less dependent on signal components to have equal amplitudes. Since the new method is shown to be more robust to the choice of kernel parameter, a larger range of kernel parameters can be tested, lowering the risk of erroneous conclusions to be drawn of the optimal TFD. This novel detection method can successfully be used together with a performance measure for TFDs to find optimal TFDs.

## Acknowledgements

Thanks to Associate Professor Peter Jönsson, Kristianstad University, for sharing HRV data for the real data examples. Ethical approval for the study was given by the ethical committee of Lund University, Dnr 2010/22.

## References

[1] B. Boashash, "Part I: Introduction to the concepts of TFSAP," in *Time Frequency Signal Analysis and Processing - A Comprehensive Reference*, Boualem Boashash, Ed. Elsevier Ltd, Oxford, 1994.

[2] Z. M. Hussain and B. Boashash, "Multi-component IF estimation," in *Statistical Signal and Array Processing*, August 2000, pp. 559 – 563.

[3] L. Stanković, I. Djurović, S. Stanković, M. Simeunović, S. Djukanović, and M. Daković, "Instantaneous frequency in time-frequency analysis:

Enhanced concepts and performance of estimation algorithms," *Digital Signal Processing*, vol. 35, pp. 1–13, December 2014.

[4] F. Auger, P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, and H.-T. Wu, "Time-frequency reassignment and synchrosqueezing," *IEEE Signal Processing Magazine*, pp. 32–41, November 2013.

[5] B. Boashash and V. Sucic, "Resolution measure criteria for the objective assessment of the performance of quadratic time-frequency distributions," *Signal Processing*, vol. 51, pp. 1253–1263, 2003.

[6] N. Saulig, I. Orović, and V. Sucic, "Optimization of quadratic time-frequency distributions using the local Rényi entropy information," *Signal Processing*, vol. 129, pp. 17–24, 2016.

[7] M. Abed, A. Belouchrani, and B. Boashash, "Time-frequency distributions based on compact support kernels: Properties and performance evaluation," *Signal Processing*, vol. 60, no. 6, pp. 2814–2827, June 2012.

[8] V. Sucic and B. Boashash, "Parameter selection for optimising time-frequency distributions and measurements of time-frequency characteristics of non-stationary signals," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2001, vol. 6, pp. 3557–3560.

[9] V. Sucic and B. Boashash, "Optimisation algorithm for selecting quadratic time-frequency distributions: Performance results and calibration," *Signal Processing and its Applications*, vol. 1, pp. 331–334, 2001.

[10] D. Malanar, V. Sucic, and J. M. O'Toole, "Automatic quality assesment and optimisation of quadratic time-frequency representation," *Electronic letters*, vol. 51, no. 13, pp. 1029–1031, June 2015.

[11] J. Lerga, V. Sucic, and B. Boashash, "An efficient algorithm for instantaneous frequency estimation of nonstationary multicomponent signals in low SNR," *EURASIP Journal on Advances in Signal Processing*, pp. 1–16, January 2011.

[12] B. Barkat and K. Abed-Meraim, "Algorithms for blind components separation and extraction from the time-frequency distribution of their mixture," *EURASIP Journal on Applied Signal Processing*, vol. 13, pp. 2025–2033, 2004.

[13] L. A. Cirillo and M. G. Amin, "Auto-term detection using time-frequency array processing," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, April 2003, vol. 6, pp. VI–465–VI–468.

[14] S. Wang, X. Chen, Y. Wang, G. Cai, B. Ding, and X. Zhang, "Nonlinear squeezing time-frequency transform for weak signal detection," *Signal Processing*, vol. 113, pp. 195–210, August 2015.

[15] N. Ali Khan and B. Boashash, "Multi-component instantaneous frequency estimation using locally adaptive directional time freqency distributions," *International Journal of Adaptive Control and Signal Processing*, pp. 1099–1115, July 2015.

[16] N. A. Khan and M. Sandsten, "Time-frequency image enhancement based on interference suppression in Wigner-Ville distribution," *Signal Processing*, vol. 127, pp. 80–85, 2016.

[17] G. G. Berntson, J. T. Bigger, JR., D. L. Eckberg, P. Grossman, P. G. Kaufmann, M. Malik, H. N. Nagaraja, S. W. Porges, J. P. Saul, Peter H. Stone, and M. W. Van Der Molen, "Heart rate variability: Origins, methods, and interpretive caveats," *Psychophysiology*, vol. 34, pp. 623–648, 1997.

[18] M. T. Verklan and N. S. Padhye, "Spectral analysis of heart rate variability: An emerging tool for assessing stability during transition to extrauterine life," *Journal of Obstetric, Gynecologic & Neonatal Nursing*, vol. 33, no. 2, pp. 256–265, 2004.

[19] K. M. Gates, L. M. Gatzke-Kopp, M. Sandsten, and A. Y. Blandon, "Estimating time-varying RSA to examine psychophysiological linkage of martial dyads," *Psychophysiology*, vol. 52, no. 8, pp. 1059–1065, August 2015.

[20] V. Maganin, T. Bassani, V. Bari, M. Turiel, R. Maestri, G. D. Pinna, and A. Porta, "Non-stationarities significantly distort short-term spectral, symbolic and entropy heart rate variability indices," *Physiological Measurement*, vol. 32, pp. 1775–1786, 2011.

[21] R. Bailón, N. Garatachea, I. de la Iglesia, J. A. Casajús, and P. Laguna, "Influence of running stride frequency in heart rate variability analysis during treadmill exercise testing," *Biomedical Engineering*, vol. 60, no. 7, pp. 1796–1805, July 2013.

[22] S.V. Bozhokin and I.B. Suslova, "Analysis of non-stationary HRV as a frequency modulated signal by double continuous wavelet transformation method," *Biomedical Signal Processing and Control*, vol. 10, pp. 34–40, March 2014.

[23] S. Dong, G. Azemi, and B. Boashash, "Improved characterization of HRV signals based on instantaneous frequency features estimated from quadratic time-frequency distributions with data-adapted kernels," *Biomedical Signal Processing and Control*, vol. 10, pp. 153–165, December 2013.

[24] M. B. Malarvili, L. Rankine, M. Mesbah, and B. Boashash, "Time-frequency based Renyi entropy of heart rate variability for newborn seizure detection," in *International Symposium on Signal Processing and Its Applications*, February 2007, pp. 1–4.

# Paper D

**Paper D**

# Objective Detection and Time-Frequency Localization of Components within Transient Signals

Isabella Reinhold[1], Maria Sandsten[1], and Josefin Starkhammar[2]

[1] *Mathematical Statistics, Centre for Mathematical Sciences, Lund University, Sweden.*

[2] *Department of Biomedical Engineering, Lund University, Sweden.*

**Abstract**

An automatic component detection method for overlapping transient pulses in multi-component signals is presented and evaluated. The recently proposed scaled reassignment technique is shown to have the best achievable resolution for closely located Gaussian shaped transient pulses, even in heavy disruptive noise. As a result, the method automatically detects and counts the number of transients, giving the center times and center frequencies of all components with considerable accuracy. The presented method shows great potential for applications in several acoustic research fields, where coinciding Gaussian shaped transients are analysed. The performance is tested on measured data from a laboratory pulse-echo set-up and from a dolphin echolocation signal measured simultaneously at two different locations in the echolocation beam. Since the method requires little user input, it should be easily employed in a variety of research projects.

**Keywords:** Component detection, Pulse-echo, Reassignment, Time-frequency analysis

## 1 Introduction

The time-frequency (TF) characterization of transient signals is of interest in many different research areas, such as ultrasonic and marine biosonar signal analysis as well as machine fault diagnosis and biomedical signal processing. In these fields, the measurement signal is often of short duration, includes several closely spaced or even coinciding components, and can be heavily disturbed by noise [1–5]. Methods which are tailored to signals of this type are scarce and not conclusively investigated, in comparison to methods for longer signals. This paper thoroughly investigates a TF representation, optimal for transient signals, and presents an automatic method for counting and characterizing the individual components, in terms of TF localization, within a signal.

Transient signals are by nature harder to characterize, compared to longer signals. The short, sometimes extremely short, duration of the pulses implies that the uncertainty in frequency is high according to the Heisenberg uncertainty principle [6]. Still, these signals are essential in fields such as ultrasonic analysis where pulse reflections are located closely in the TF domain, and for the ultrasonic biosonar analysis of several toothed whale species, where the broadband signals (30–60 kHz) are only a few periods long [7, 8]. The sonar beam of toothed whales contains signal components from various acoustic propagation pathways inside the animals' forehead. It is thought that these pathways can be altered by the animal by for instance varying the grade of inflation of airsacks around the sound generation mechanism or by altering the shape of the melon, an anatomic structure thought to function as an acoustic lens for the echolocation beam. However, to what extent the signal can be controlled by the animal and what specific functions it serves, are not yet fully understood [9,10].

Methods in the quadratic class often aim for reaching the best possible TF concentration, where one of the most known is the Choi-Williams distribution (CWD) [6, 11], often applied for TF estimation of biomedical signals. For broadband excitations where the multiple components appear very close in the TF domain, the signals are increasingly difficult to analyze [12]. The TF representations employed in previous studies of broadband echolocation signals can often be used in the off-axis part of the echolocation beam where the components are more separated in time. Along the beam axis, the time and frequency information of possible individual components, are still unknown since all previously used methods are unable to resolve these

signal components [2], and the topic is currently a research field of great importance [1,13].

A technique to improve the localization of single TF components and enhance the readability of the TF representation of multi-component signals was introduced by [14] and later reintroduced by [15]. The method reassigns signal energy to the center of gravity, giving higher energy concentration at the instantaneous frequencies of the signal. A similar method, the synchrosqueezing transform by [16], related to the empirical mode decomposition [17], reassigns all energy in frequency at a certain time point. However these methods are designed for longer chirps and constant frequency signals, and are based on the assumption of a linear frequency modulation, essentially of infinite length, and are hence not accurately applicable to transients signals. Methods exist that convert the possible non-linear instantaneous frequency into a linear one and in [18] a nonlinear squeezing transform, especially designed for weak signal detection, is proposed. There exist many other methods for localizing and counting signal components, e.g. [19–23], however these methods are also developed for longer lasting signals.

Short transient signals can often be assumed to have a Gaussian like shape in time, and modern algorithms that resolve the parameters of a Gaussian shaped function in time have been described by [24] and [25]. However for components that overlap heavily in time, TF based methods, such as Gabor and wavelet based algorithms have been applied to a larger extent, for which the main aim is to find the analysis window achieving the best TF resolution. Similarly the signal adaptive Stockwell transform estimates the width of a Gaussian window function using a concentration criterion [26]. The Gabor and Stockwell transforms are widely used and adapted in many fields of research, e.g. estimating the direction of arrival [27], automatically adapting the TF resolution of transients [28], detecting epileptic seizures [29] or double-talk in acoustic echo cancellation [30].

A method tailored to very short transients, which goes beyond the lower bound of the Gabor transform and the reassignment of longer lasting transients, is presented in [31]. This reassignment technique, the scaled reassigned spectrogram, finds the TF centers of individual signal components by utilizing that many transients have a Gaussian like shape in time. The scaled reassigned spectrogram is a high TF resolution method and is therefore well suited for detection and localization of transient signal components, also when they are closely located in the TF domain.

In literature numerous of peak detection algorithms for specific applications can be found. Closest to our representations are image peak detection algorithms and image thresholding methods, see overview in [32]. One of the more popular image peak detection algorithms is presented in [33], which is developed for histogram-based image data reduction. Other approaches use the wavelet transform [34]. However, to our knowledge there exist no researched methods for finding and resolving the localization of time-frequency peaks.

In this paper we present a novel method for objective detection, counting and TF localization of components within transient signals. We also present a thorough evaluation of the novel method and the resolution of the suggested TF representation. Our method is unique in that it is developed for short transient signals, it exploits the high resolution of the scaled reassigned spectrogram and can be adapted to signals with heavy disruptive noise. The paper offers an comprehensive evaluation of the method on simulated signals, and shows results for measured ultrasound pulse-echoes and marine biosonar signals. The results are of importance especially to the acoustic research community.

## 2 The ReSTS

In order to detect and localize individual transient pulses in a multi-component signal, there is a need for a TF representation with appropriate resolution. Figure 1 shows the time signal and a joint TF representation, the spectrogram, of three different signals with decreasing time distance between the component TF centers. For the first two signals, Figure 1(a) the spectrogram has adequate resolution and separates the components, but for the second case, Figure 1(b), the overlap in time is too large and as a result the spectrogram does not fully resolve the two components. The spectrogram is computed using a Gaussian window of the same length as the signal component and with step size one sample. The matched window, with equal shape and length of the signal, gives the optimal time-frequency concentration of the spectrogram.

However the components of the last example can be resolved using the reassigned spectrogram for transient signals (ReSTS), proposed in [31] as the scaled reassignment technique which is an adaptation of the reassigned spectrogram [14, 15]. The ReSTS relies on the assumption of a Gaussian shaped transient signal and can in theory give perfect TF localization if the time- and frequency distances of several close transient components are large enough. The

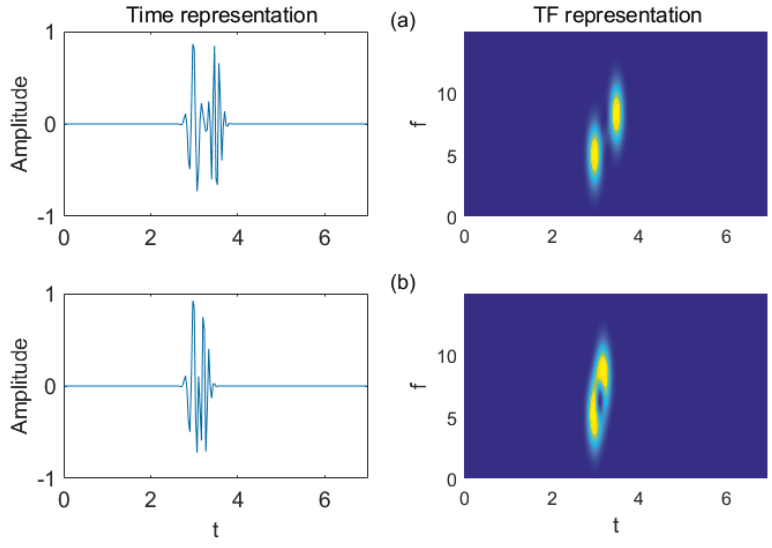



**Figure 1:** Illustration of a transient signal with two Gaussian pulses, each of approximate length 0.6 with varying TF centers $(t_0, f_0)$. The spectrogram is computed using a Gaussian window of the same length as the signal component and with step size one sample; (a) pulses overlapping somewhat in time but clearly separated in frequency, $(t_1, t_2) = (3, 3.5)$ and $(f_1, f_2) = (5, 8.3)$; (b) pulses overlapping so much in time that the frequency separation is not clear, $(t_1, t_2) = (3, 3.2)$ and $(f_1, f_2) = (5, 8.3)$.

ReSTS is obtained by first calculating the spectrogram of a signal $x(t)$ using a desired time window $h(t)$

$$S_x^h(t,\omega) = \left|F_x^h(t,\omega)\right|^2 = \left|\int x(s)h^*(s-t)e^{-i\omega s}ds\right|^2, \tag{1}$$

where $*$ denotes complex conjugate, $\omega = 2\pi f$ and the integral runs from $-\infty$ to $\infty$. The signal energy is then reassigned by introducing the reassignment coordinates $\hat{t}_x(t,\omega)$ and $\hat{\omega}_x(t,\omega)$ and the two-dimensional Dirac impulse, $\iint f(t,\omega)\delta(t-t_0,\omega-\omega_0)dtd\omega = f(t_0,\omega_0)$. The ReSTS is then defined as

$$ReS_x^h(t,\omega) = \iint S_x^h(s,\xi)\delta\left(t-\hat{t}_x(s,\xi),\omega-\hat{\omega}_x(s,\xi)\right)dsd\xi, \tag{2}$$

and thus maps signal energy from a point $(t_0,\omega_0)$ to the point $(\hat{t}_x(t_0,\omega_0),\hat{\omega}_x(t_0,\omega_0))$ in the spectrogram. The reassignment coordinates need to

be calculated for each selection of signal and time window according to

$$\begin{aligned}\hat{t}_x(t,\omega) &= t + c_t \Re\left(\frac{F_x^{th}(t,\omega)}{F_x^{h}(t,\omega)}\right),\\ \hat{\omega}_x(t,\omega) &= \omega - c_\omega \Im\left(\frac{F_x^{dh/dt}(t,\omega)}{F_x^{h}(t,\omega)}\right),\end{aligned} \tag{3}$$

where $\Re$ and $\Im$ represents the real and imaginary parts and $F_x^h$, $F_x^{th}$ and $F_x^{dh/dt}$ are the short-time Fourier transforms (STFTs) with different time windows. The included scaling factors $c_t$ and $c_\omega$ makes the ReSTS adaptable to transient signals, and separates it from the normal reassigned spectrogram that have $c_t = c_\omega = 1$ [15, 31, 35].

Transient signals are often assumed to be Gaussian shaped in time, it is thus interesting to consider the unit energy Gaussian function

$$x_G(t) = \sigma^{-1/2}\pi^{-1/4}e^{-\frac{t^2}{2\sigma^2}}, \tag{4}$$

and multi-component signals constructed by time, frequency and phase shifted Gaussian shaped pulses

$$x(t) = \sum_{k=1}^{K} a_k x_G(t - t_k)e^{i2\pi f_k t}e^{i2\pi\phi_k}, \tag{5}$$

where $a_k$ is the amplitude, $t_k$ and $f_k = \omega_k/2\pi$ are the time and frequency centers and $\phi_k \in [0\ 1)$ the phase shift.

Hansson-Sandsten and Brynolfsson [31] calculated the reassignment coordinates for a Gaussian signal with time-frequency center at the origin, and a matching Gaussian time window

$$\begin{aligned}\hat{t}_{x_G}(t,\omega) &= t - c_t\frac{t}{2},\\ \hat{\omega}_{x_G}(t,\omega) &= \omega - c_\omega\frac{\omega}{2}.\end{aligned} \tag{6}$$

Perfect TF localization is then achieved when $c_t = c_\omega = 2$, thus giving the reassignment coordinates $(\hat{t}_{x_G}(t,\omega), \hat{\omega}_{x_G}(t,\omega)) = (0,0)$, i.e. the correct TF center. This result can easily be expanded to Gaussian signals with other time-frequency centers since the spectrogram obeys time-frequency

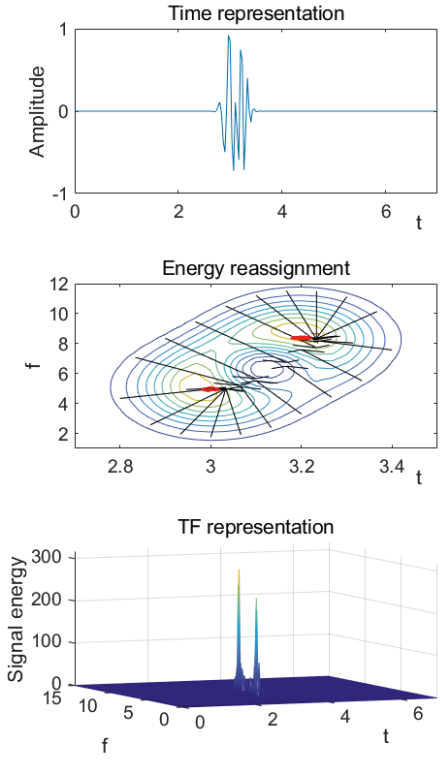



**Figure 2:** Illustration of the reassignment of signal energy and the resulting TF representation, adapted for transient signals, for the signal in Figure 1. The spectrogram is computed using a Gaussian window of the same length as the signal component and with step size one sample. The signal has two transient pulses, each of approximate length 0.6, which overlap heavily in time, $(t_1, t_2) = (3, 3.2)$ and $(f_1, f_2) = (5, 8.3)$, however when the signal energy is reassigned to the center of mass, the ReSTS shows two clear peaks at the correct TF centers.

shift-invariance and due to the linearity of the Fourier transform, the reassignment coordinates are linear [35].

For multi-component signals [Eq. (5)], there will be some interaction between the components after reassignment. However it is still possible for the ReSTS to show clearly separated components, Figure 2, to be compared with the spectrogram in Figure 1(c). The figure illustrates the reassignment of the signal energy to the new

TF coordinates, according to Eq. (6) with $c_t = c_\omega = 2$, shown by the arrows in the spectrogram contour plot to the TF centers of the two pulses. The resulting ReSTS has two clear peaks at the correct TF centers of the pulses and very little scattered signal energy.

## 3 Automatic component detection algorithm

This section proposes an algorithm that automatically counts the number of signal components in the ReSTS. The algorithm will enable users to automatically find the number of transient pulses in a signal and the individual TF locations of the pulses. It will thus be possible to objectively and efficiently analyze transient signals.

The ReSTS consists of high energy peaks in a low energy surrounding. Both signal components and noise will form peaks after the reassignment, however peaks will have higher amplitude, indicative of higher signal energy, if they are the result of signal components compared to peaks originating from noise. The algorithm therefore assumes that any local maxima is a peak, either from a signal components or from noise, and that the amplitudes of the peaks will differ between signal components and noise.

A pseudo code of the proposed algorithm is presented. It uses the discrete time and discrete frequency ReSTS matrix, denoted $ReS$, as TF representation. The user decides a maximum number of components for the signal, $K_{max}$, the guess can be much larger than the expected number of components without compromising the performance, however very large numbers would increase the computational time. The user also sets an area around a local maximum, $2\delta_t$ wide in time and $2\delta_f$ wide in frequency, that will be assumed not to include more than one local maximum. The choices of $\delta_t$ and $\delta_f$ depend on the resolution of the ReSTS which will be evaluated in Section 4.1.

The constant $0 < \rho \leq 1$ is set by the user and allows the algorithm to be used for signals with relatively low SNR. Depending on the SNR, all noise peak amplitudes could be low and relatively constant or some noise peak amplitudes could be rather high and close in amplitude to the signal peaks. A large $\rho$ assumes high SNR, where signal and noise peak amplitudes are clearly separated, a small $\rho$ assumes low SNR and allows the algorithm to find relatively smaller signal peaks.

The output from the algorithm is the time locations $T_1, T_2, \ldots, T_K$ and the corresponding frequency locations, and $F_1, F_2, \ldots, F_K$, of the estimated signal peaks.

**Algorithm 1** Pseudo code

**Input:** $ReS$, $K_{max}$, $\delta_t$, $\delta_f$, $\rho$
**Output:** $T_1, T_2, \ldots, T_K, F_1, F_2, \ldots, F_K$

$ReS$ : *The ReSTS matrix*
$K_{max}$ : *Initial guess of maximum number of signal components*
$\delta_t$ : *Smallest time separation*
$\delta_f$ : *Smallest frequency separation*
$\rho$ : *Normalizing constant for the noise amplitudes*
$T_1, T_2, \ldots, T_K$ : *Time centers of the signal components*
$F_1, F_2, \ldots, F_K$ : *Frequency centers of the signal components*

$N = 3K_{max}$
**for** i=1:N **do**
    find coordinates of maximum in $ReS$, $(T(i), F(i))$
    save maximum amplitude in vector, $A(i)$
    define the rectangle area, $B_{TF}$, with center $(T(i), F(i))$ and area $2\delta_t \cdot 2\delta_f$
    set the area $B_{TF}$ in $ReS$ to 0
**end for**
$K = K_{max} + 1$
**repeat**
    $K = K - 1$
    $\Delta_n = \rho\,(A(K+1) - A(N))$
    $\Delta_s = A(1) - A(K)$
    $\Delta = A(K) - A(K+1)$
    **if** $\Delta > \Delta_s$ **then**
        peak $K$ is a signal component
    **else if** $\Delta < \Delta_n$ **then**
        peak $K$ is a noise component
    **else**
        **if** $\Delta/\Delta_n > \Delta_s/\Delta$ **then**
            peak $K$ is a signal component
        **else**
            peak $K$ is a noise component
        **end if**
    **end if**
**until** peak K is a signal component

The algorithm runs as described below:

- $N = 3K_{max}$ peaks are located to always include an adequate number of noise peaks.
- The maximum peaks are estimated sequentially and the area of size $2\delta_t \cdot 2\delta_f$ around the current maximum peak is set to zero. Then the next maximum peak can be located. The identified local maxima of the ReSTS can be assumed to be peaks because of the properties of the ReSTS, and due to the structure of the algorithm, the $N$ identified peaks are sorted with descending order of amplitude in all relevant vectors.
- The number of assumed signal peaks among the $N$ maximum peaks is set to $K = K_{max} + 1$ and is then decreased with one for each iteration the repeat loop. $\Delta_n$ is the difference in amplitude of all known noise peaks, normalized with the constant $\rho$. $\Delta_s$ is the difference in amplitude of all the possible signal peaks. $\Delta$ is the difference in amplitude of the smallest possible signal peak and largest known noise peak.
- In each iteration of the repeat loop (line 9 – 25), it is determined if peak $K$ is a signal or noise component. If peak $K$ is determined to be from a signal component, all peaks with larger amplitude are assumed to also be signal peaks and the algorithm has finished its search for signal peaks.
- Peak $K$ will be determined to be from a signal or noise component depending on how much its amplitude deviates from the amplitude slope created by the known noise peaks.
- If the test $\Delta > \Delta_s$ (line 14) is passed, peak $K$ should be a clear signal component as its difference in amplitude to the largest noise peak is larger than its difference in amplitude to the largest identified peak.
- If the test $\Delta < \Delta_n$ (line 16) is passed, peak $K$ should be a noise peak as $\Delta$ will be relatively small.
- Note that both $\Delta > \Delta_s$ and $\Delta < \Delta_n$ can be true, then peak $K$ is assumed to be from a signal component, however the SNR is probably very low for such a signal.
- If none of the test $\Delta > \Delta_s$ and $\Delta < \Delta_n$ are true, then both $\Delta/\Delta_n \geq 1$ and $\Delta_s/\Delta \geq 1$. The test $\Delta/\Delta_n > \Delta_s/\Delta$ (line 19) is true if $\Delta_s/\Delta$ is closer

to 1 compared to $\Delta/\Delta_n$. This means that the amplitude of peak $K$ deviates from the amplitude slope of the noise peaks, and peak $K$ is assumed to be a signal component.

## 4 Resolution of the ReSTS

The proposed automatic component detection algorithm is designed to be used with the ReSTS, which means that the TF resolution of the ReSTS is of importance. In theory the ReSTS with Gaussian window can give perfect TF localization to Gaussian signal components, however the resolution of components in signals disrupted by noise needs to be evaluated in order for the proposed algorithm to be usable in practice.

We consider the signal $x(t)$ in Eq. (5) that is a linear combination of Gaussian pulses [Eq. (4)] and add white Gaussian noise. For such a signal with two components, that have the same frequency centers and amplitudes, but different time centers, the components can be moved closer together in time to examine when different TF distributions no longer can resolve the components.

In this section the simulated signals have centre frequency 5 MHz, sampling frequency 100 MHz and the scaling of the Gaussian pulses [Eq. (4)], $\sigma = 0.5$ $\mu$s, which gives an approximate signal length of 1.2 $\mu$s, full width at half maximum, or approximately 5 periods. The evaluated time distances range from to 0.5 $\mu$s to 2.0 $\mu$s, this means that the signal components will heavily overlap for some test signals. White noise is additively disturbing the signals, with SNR = 5 dB, defined as averaged total signal energy to variance of the noise.

Figure 3 shows a realization of the signal when $(t_1, t_2) = (4.00, 5.50)$ $\mu$s and it can be seen that for this noisy signal, the overlap of the components is noticeable in time, Figure 3(a), as well as in the spectrogram, Figure 3(b) and in the CWD [11], Figure 3(c). However, in the ReSTS, Figure 3(d), the components are clearly separated, showing the TF centers as clear peaks. Important to note is that the scaling parameter for the CWD, $\alpha = 0.2$, is evaluated and chosen so that it balances the suppression of interference and loss of resolution.

The three different TF distributions, the spectrogram and the ReSTS with a matched Gaussian window, $\sigma = 0.50$ $\mu$s, and the CWD, are evaluated by simulating 1000 realizations of each test signal with different time distances between components. Each simulation has different noise disturbance

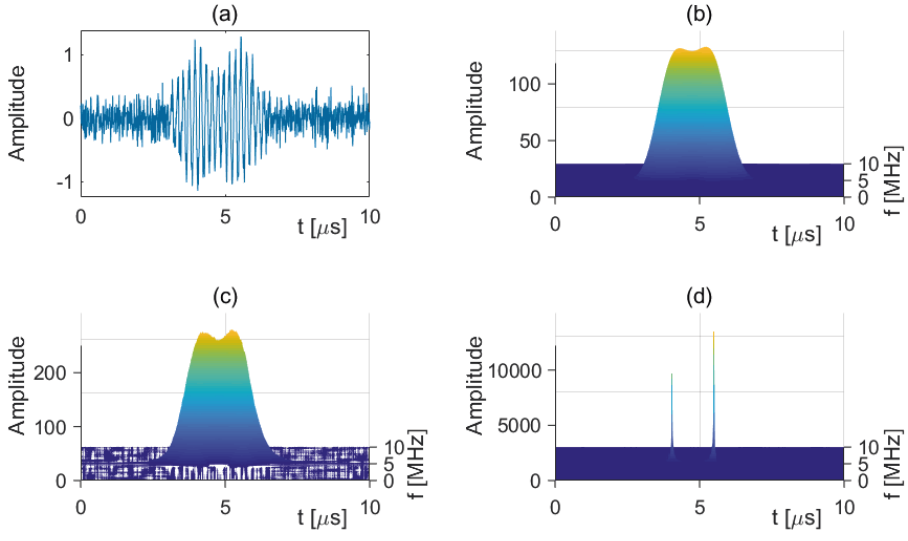



**Figure 3:** Illustrations for a realization of the simulated signal of two Gaussian pulses with scaling $\sigma = 0.5$ $\mu$s, centre frequencies $(f_1, f_2) = 5$ MHz of both components, centre times $(t_1, t_2) = (4.00, 5.50)$ $\mu$s, sampled with frequency 100 MHz. A disturbance of white Gaussian noise with SNR = 5 dB is added to the signal; (a) the time signal; (b) the spectrogram; (c) the CWD; (d) the ReSTS.

realizations and phase shifts for both signal components. The TF distributions are calculated for each realization and the maximum in an area around the true TF center for both components are extracted to calculate the mean estimated TF centers and mean peak amplitudes. These results are shown in Tables 1 and 2, where the standard deviation (SD) of the estimated time and frequency centers and coefficient of variation (CV) of the amplitudes are calculated to show the spread of the estimates and peak amplitudes. An evaluation of the required number of realizations was carried out and confirms stability of the estimates at 1000 realizations, as shown in Figure 4.

It can be seen that all three methods separate the signal components when time distance between the components is 2.00 $\mu$s, even though the CWD gives some deviations in the time centers, see Table 1. When the time distance is 1.50 $\mu$s the spectrogram and the CWD are not able to give the correct time centers, while the ReSTS gives a good estimate mean and low standard deviation. This means that

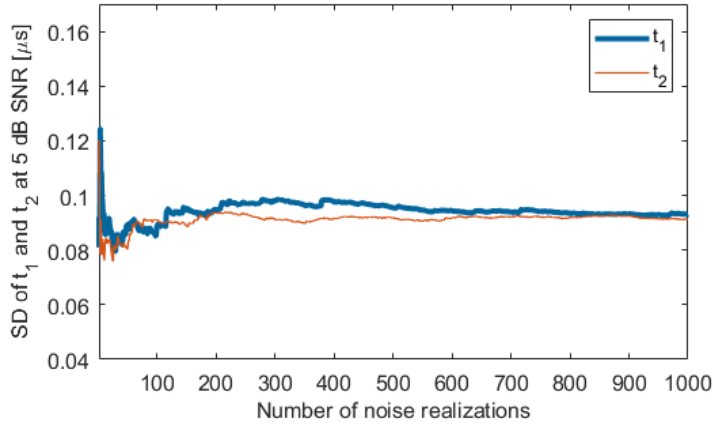



**Figure 4:** The SD of detected centre times over number of realization of the simulated signal of two Gaussian pulses with $(t_1, t_2) = (4.00, 5.00)$ $\mu$s and $(f_1, f_2) = 5$ MHz.

**Table 1:** Mean and standard deviation of the estimated time centers from 1000 simulations of the signal [Eq. (5)] with white Gaussian noise, SNR = 5 dB and random phase for each pulse.

| | | Spectrogram | | | | CWD | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Mean [$\mu$s] | | SD [$\mu$s] | | Mean [$\mu$s] | | SD [$\mu$s] | |
| Test | True $(t_1, t_2)$ | $\hat{t}_1$ | $\hat{t}_2$ | $\hat{t}_1$ | $\hat{t}_2$ | $\hat{t}_1$ | $\hat{t}_2$ | $\hat{t}_1$ | $\hat{t}_2$ |
| 1 | (4.00, 6.00) $\mu$s | **4.00** | **5.99** | **0.03** | **0.03** | **4.07** | **5.93** | **0.05** | **0.06** |
| 2 | (4.00, 5.50) $\mu$s | 4.14 | 5.36 | 0.19 | 0.19 | 4.12 | 5.38 | 0.11 | 0.11 |
| 3 | (4.00, 5.00) $\mu$s | 4.30 | 4.69 | 0.27 | 0.27 | 4.22 | 4.77 | 0.24 | 0.24 |
| 4 | (4.00, 4.50) $\mu$s | 4.18 | 4.32 | 0.16 | 0.16 | 4.12 | 4.37 | 0.17 | 0.17 |

| | | ReSTS | | | |
|---|---|---|---|---|---|
| | | Mean [$\mu$s] | | SD [$\mu$s] | |
| Test | True $(t_1, t_2)$ | $\hat{t}_1$ | $\hat{t}_2$ | $\hat{t}_1$ | $\hat{t}_2$ |
| 1 | (4.00, 6.00) $\mu$s | 4.00 | 6.00 | 0.02 | 0.02 |
| 2 | (4.00, 5.50) $\mu$s | 4.00 | 5.49 | 0.04 | 0.04 |
| 3 | (4.00, 5.00) $\mu$s | **4.03** | **4.97** | **0.09** | **0.09** |
| 4 | (4.00, 4.50) $\mu$s | 4.12 | 4.37 | 0.14 | 0.14 |

**Table 2:** Mean and coefficient of variation of the peak amplitudes from 1000 simulations of the signal [Eq. (5)] with white Gaussian noise, SNR = 5 dB and random phase for each pulse.

| | Spectrogram | | | | CWD | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | | CV | | Mean | | CV | |
| Test | $A_1$ | $A_2$ | $A_1$ | $A_2$ | $A_1$ | $A_2$ | $A_1$ | $A_2$ |
| 1 | **88.9** | **88.9** | **0.05** | **0.04** | **189** | **188** | **0.12** | **0.12** |
| 2 | 96.4 | 96.4 | 0.16 | 0.17 | 200 | 200 | 0.21 | 0.21 |
| 3 | 135 | 135 | 0.44 | 0.44 | 230 | 230 | 0.42 | 0.42 |
| 4 | 165 | 165 | 0.63 | 0.63 | 262 | 262 | 0.63 | 0.63 |

| | ReSTS | | | |
|---|---|---|---|---|
| | Mean | | CV | |
| Test | $A_1$ | $A_2$ | $A_1$ | $A_2$ |
| 1 | 30800 | 30600 | 0.35 | 0.34 |
| 2 | 10500 | 10600 | 0.31 | 0.31 |
| 3 | **3840** | **3840** | **0.48** | **0.45** |
| 4 | 12200 | 12900 | 1.12 | 1.14 |

for this type of signal with low SNR and random phase shifts, the spectrogram does not succeed to resolve the Gaussian pulses that are $2\sigma = 1.0$ $\mu$s apart, which is the theoretical minimum time separation needed between two Gaussian pulses for the peaks to be resolved [36]. It is notable that the ReSTS goes beyond the resolution limit of the spectrogram, using the a priori information of a Gaussian shaped transient pulse, and we see that the estimated time centers for the ReSTS are still very good when $(t_1, t_2) = (4.00, 5.00)$ $\mu$s. It is not until the time distance is 0.5 $\mu$s that the estimates for the ReSTS becomes unreliable.

The signal peak amplitudes of the ReSTS are much higher compared to the other TF distributions, Table 2. This clearly shows that the signal energy is more localized to the TF centers of each component in the ReSTS compared to both the spectrogram and CWD. The peak amplitudes in the spectrogram and CWD increase when the component separation decrease, and the energy from the components combine. The amplitudes in the ReSTS first decrease because reassignment to the correct mass centers becomes more difficult for closer components,

and some energy is reassigned to positions in between the true TF centers. For the smallest time distance, the ReSTS is unreliable and sometimes only gives one strong peak, resulting in high mean amplitude but also very high coefficient of variation.

The estimated frequency centers are not presented, as all methods have good estimations for these signals. This is expected, since the signal components have the same frequency, and thus any smoothing of the TF distributions will be mainly in the time domain.

### 4.1 Guidance in parameter choices of $\delta_t$ and $\delta_f$

The proposed automatic component detection algorithm requires the user to define the parameters $\delta_t$ and $\delta_f$, which are signal dependent. Finding theoretical values for these parameters is an arduous task because of the interaction between close signal components and noise in the ReSTS. Values can instead be found experimentally and be translated to a general signal.

To achieve the best performance of the proposed automatic component detection algorithm together with the ReSTS, the parameters $\delta_t$ and $\delta_f$ should be defined equal to the time and frequency resolution of the ReSTS. The algorithm will then be able to both resolve, and thus accurately detect, any two components which are separated by at least $\delta_t$ and $\delta_f$. We are also interested in results that can be applied to real, measured signals, therefore this section will experimentally find the resolution of the ReSTS for noisy, transient signals.

According to the Heisenberg inequality, even the most optimal resolution need to fulfill, $\delta_t \delta_f \geq 1/(4\pi)$, where $\delta_t$ is the uncertainty in time, i.e. the length of the pulse, and $\delta_f$ is the uncertainty in frequency. Two Gaussian pulses of equal amplitude and time length, are separable if

$$\delta_t[s] = 2\sigma[s], \quad \delta_f[s^{-1}] = 1/(\pi\sigma[s]), \tag{7}$$

where $\sigma$ is the scaling parameter of the Gaussian [36]. This means that these distances correspond to the best resolution the spectrogram can possibly achieve, without reassignment.

Defining the signals as in Eq. (5) with two Gaussian pulse components and changing the TF center of one component, the minimum time and frequency distances needed to resolve two components for the ReSTS can be determined.

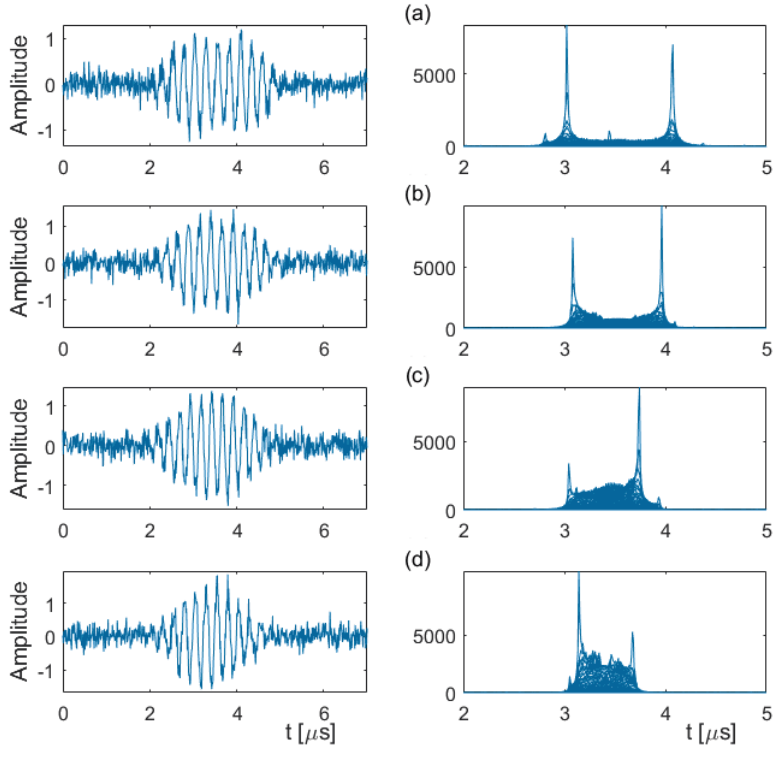



**Figure 5:** Realizations of the test signals [Eq. (5)] with two Gaussian pulses and white Gaussian noise, SNR = 5 dB, where $f_1 = f_2 = 4.00$ MHz and random phases of the components. The time signal is shown to the left and the ReSTS is viewed to the right so that the amplitude and time axis can be seen; (a) signal with $(t_1, t_2) = (3.00, 4.10)$ $\mu$s; (b) signal with $(t_1, t_2) = (3.00, 4.00)$ $\mu$s; (c) signal with $(t_1, t_2) = (3.00, 3.90)$ $\mu$s; (d) signal with $(t_1, t_2) = (3.00, 3.80)$ $\mu$s.

For the used test signals the Gaussian pulses have $\sigma = 0.50$ $\mu$s, the sampling frequency is 100 MHz and 0.01 $\mu$s corresponds to 1 sample. The simulated signals are disturbed by white Gaussian noise, SNR = 5 dB, which will give $\delta_t$ and $\delta_f$ that can be used when applying the automatic component detection method on measured data, possibly also with severe disrupting noise.

The minimum required time separation is evaluated by keeping the frequency of the two signal components constant and decreasing the time distance between

**Table 3:** Mean and standard deviation of the estimated time centers and the mean and coefficient variation of the peak amplitudes from 1000 simulations of the signal [Eq. (5)] with $f_1 = f_2 = 4.00$ MHz, white Gaussian noise, SNR = 5 dB and random phase for each pulse.

| | Mean [$\mu$s] | | SD [$\mu$s] | | Mean | | CV | |
|---|---|---|---|---|---|---|---|---|
| True $(t_1, t_2)$ | $\hat{t}_1$ | $\hat{t}_2$ | $\hat{t}_1$ | $\hat{t}_2$ | $A_1$ | $A_2$ | $A_1$ | $A_2$ |
| (3.00, 4.10) $\mu$s | 3.02 | 4.08 | 0.08 | 0.08 | 6630 | 6630 | 0.42 | 0.43 |
| (3.00, 4.00) $\mu$s | **3.03** | **3.97** | **0.09** | **0.09** | **5910** | **5790** | **0.45** | **0.47** |
| (3.00, 3.90) $\mu$s | 3.04 | 3.85 | 0.11 | 0.11 | 5180 | 5160 | 0.53 | 0.56 |
| (3.00, 3.80) $\mu$s | 3.07 | 3.72 | 0.13 | 0.13 | 5390 | 5490 | 0.67 | 0.67 |

the TF centers of the components. Figure 5 shows realizations of the signals with decreasing time separation of the TF centers, it shows the time signal (left) and the ReSTS from the viewpoint so that only the time axis is seen (right). The figure shows that the signal energy becomes more scattered, as it reassigns to locations between the components, when the time distance decreases.

Table 3 shows the results of 1000 simulations of the test signals, with random noise and phase shift of the components for each simulation. The table only shows the estimated time centers and peak amplitudes, since the estimated frequency centers are consistently reliable for all time distances. It can be seen that the mean estimates differ at most 6 samples from the true positions, however when $t_2 = 3.90$ $\mu$s the mean of $\hat{t}_2$ is closer to 3.80 $\mu$s than the true value. Also when the time distance is less than 1.0 $\mu$s the coefficient variations of the peak amplitudes are more than 50%, which indicates that the signal energy can be rather scattered, as can be seen in Figure 5(c) and (d). It can therefore be safer to set $\delta_t$ corresponding to 1.0 $\mu$s even though the estimated time centers for the smaller time distances are close to the truth. For the chosen signals of this experiment, 1.0 $\mu$s corresponds to $2\sigma$, which then can be applied to a general Gaussian pulse, multi-component, transient signal.

The required minimum frequency distance between two components is examined by keeping the time centers of the two signal components constant and decreasing the frequency distance between the TF centers of the components. Again, $\sigma = 0.50$ $\mu$s and the sampling frequency is 100 MHz. For

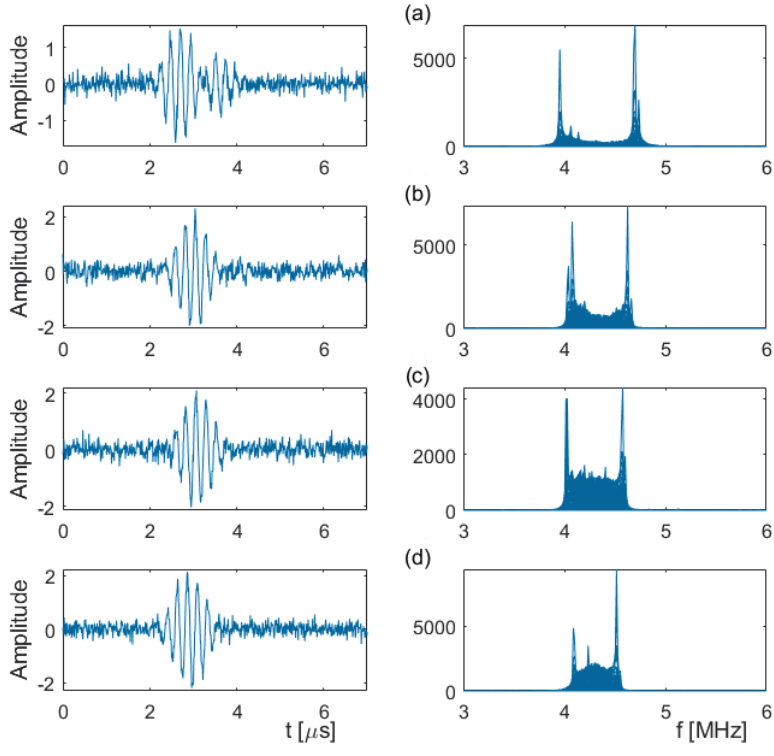



**Figure 6:** Realizations of the test signals [Eq. (5)] with two Gaussian pulses and white Gaussian noise, SNR = 5 dB, where $t_1 = t_2 = 3.00$ $\mu$s and random phase of the components. The time signal is shown to the left and the ReSTS is viewed so that the amplitude and frequency axis (right) can be seen; (a) signal with $(f_1, f_2) = (4.00, 4.70)$ MHz; (b) signal with $(f_1, f_2) = (4.00, 4.65)$ MHz; (c) signal with $(f_1, f_2) = (4.00, 4.60)$ MHz; (d) signal with $(f_1, f_2) = (4.00, 4.55)$ MHz.

the calculated ReSTS matrices, the distance between two frequency values is 0.012 MHz.

Figure 6 shows realizations of the simulated signals for each of the four chosen frequency separations. The figure shows the time signal and the ReSTS, first the time signal (left) and then ReSTS from the viewpoint so that only the frequency axis is seen (right). It can be noted that the energy seem more scattered for smaller frequency distances.

**Table 4:** Mean and standard deviation of the estimated frequency centers and the mean and coefficient variation of the peak amplitudes from 1000 simulations of the signal [Eq. (5)] with $t_1 = t_2 = 3.00$, white Gaussian noise, SNR 5 = dB and random phase for each pulse.

| | Mean [MHz] | | SD [MHz] | | Mean | | CV | |
|---|---|---|---|---|---|---|---|---|
| True $(f_1, f_2)$ | $\hat{f}_1$ | $\hat{f}_2$ | $\hat{f}_1$ | $\hat{f}_2$ | $A_1$ | $A_2$ | $A_1$ | $A_2$ |
| (4.00, 4.70) MHz | 4.02 | 4.69 | 0.05 | 0.05 | 6670 | 6800 | 0.39 | 0.41 |
| (4.00, 4.65) MHz | **4.02** | **4.64** | **0.06** | **0.06** | **5860** | **5880** | **0.43** | **0.44** |
| (4.00, 4.60) MHz | **4.03** | **4.58** | **0.07** | **0.07** | **5360** | **5420** | **0.48** | **0.51** |
| (4.00, 4.55) MHz | 4.04 | 4.52 | 0.07 | 0.08 | 5390 | 5290 | 0.57 | 0.58 |

Table 4 shows the estimated frequency centers and peak amplitudes from 1000 simulations of the test signals, where the noise and phase shifts of the components are random for each simulation. The estimated time centers were all consistent and close to the true value $t_1 = t_2 = 3.00$ $\mu$s. It can be seen that the estimated frequency centers over all are reliable, the mean of the estimates deviates at most 4 samples from the true frequency centers. The standard deviations for the estimated frequency centers are also rather low, however for $f_2 = 4.55$ MHz the mean of $\hat{f}_2$ is almost 4.50 MHz and already for $f_2 = 4.60$ the coefficient variation is 50% for one of the peaks. Thus, a reasonable choice for $\delta_f$ would be somewhere between $0.60 - 0.65$ MHz for this signal and to translate this to a general Gaussian pulse, multi-component, transient signal, a reasonable choice is $1/(\pi\sigma)$, for this signal that is approximately 0.64 MHz.

This evaluation is done with rather low SNR, and recommends to chose $\delta_t$ and $\delta_f$ according to Eq. (7). This applies for a general signal which can be considered to consist of multiple, transient Gaussian pulses and is disrupted by noise. This means that the proposed automatic component detection algorithm used with the ReSTS can resolve transient signal components, which have the smallest time and frequency distance required for two Gaussian pulses to be separable. The parameters depend on the length of the Gaussian pulses in the signal and have an inverse relation to each other, meaning that shorter pulses give a smaller uncertainty in time but a larger uncertainty in frequency.

**Table 5:** Time and frequency centers and amplitudes for the transient, multi-component Gaussian signal [Eq. (5)], used for evaluating the performance of the proposed automatic component detection algorithm.

| $k$ | $t_k$ [$\mu$s] | $f_k$ [MHz] | $a_k$ |
|---|---|---|---|
| 1 | 3.0 | 2.0 | 1.0 |
| 2 | 5.0 | 5.0 | 0.6 |
| 3 | 3.0 | 8.0 | 0.8 |
| 4 | 7.0 | 2.0 | 0.6 |
| 5 | 7.0 | 8.0 | 0.4 |

## 5 Performance of the automatic component detection algorithm

Using the parameter choices found in the previous section, $\delta_t = 2\sigma$ and $\delta_f = 1/(\pi\sigma)$, the performance of the proposes automatic component detection algorithm can now be evaluated. This is done using test signals [Eq. (5)] with the different number of components, TF centers and amplitudes given in Table 5. Two multi-component signals are used, one with two and one with five components. The first two presented in Table 5 forms the two-component signal and the five-component signal includes all five.

The test signals are evaluated with SNR = 5 dB and SNR = 15 dB, where 1000 simulations are done for each signal and SNR, with different noise realizations (white Gaussian noise) and random phase shifts for the signal components. The sampling frequency is 100 MHz and the scaling of the Gaussian pulses $\sigma = 0.5$ $\mu$s. The automatic algorithm extracts $K_{max}$ peaks of the ReSTS and determines which peaks that are signal components.

For the evaluation a detection rate is calculated, giving a value 0 – 1. If the algorithm finds the correct number of signal components, which all are close to the true TF centers, then the detection is considered correct (1). In all other cases the detection is incorrect (0). The correct and incorrect detections from the 1000 simulations are averaged to get the detection rate. Thus 1 means that detection was correct for all 1000 simulations and 0 that detection was incorrect for all simulations.

**Table 6:** Detection rate of the automatic component detection algorithm for multi-component, transient signals [Eq. (5)], disturbed by white noise and with parameters according to Table 5. The detection rates are obtained from 1000 simulations with different noise realizations and where the signal components have random phase shifts. If all detections for all 1000 simulations are correct, the rate is 1.

| | | Detection rate for | |
|---|---|---|---|
| SNR [dB] | $\rho$ | 2 components | 5 components |
| 15 | 2/3 | 1 | 1 |
| 5 | 2/3 | 1 | 0.94 |
| 5 | 1/3 | 1 | 0.98 |

Table 6 shows the resulting detection rates for six tests. For all tests with two signal components, the maximum number of peaks parameter is chosen to $K_{max} = 5$ and when the signal has five components $K_{max} = 8$. An estimated TF center is considered close enough to the true TF center if the time and frequency difference is less than $\delta_t = 2\sigma = 1\ \mu$s and $\delta_f = 1/(\pi\sigma) = 0.65$ MHz respectively.

It can be seen in Table 6 that for the signals with two components, the proposed algorithm always detects the correct signal components. For the signals with five components, the detection rate is 0.95 when the SNR is low, however the result can be improved to 0.98 by lowering the normalization constant $\rho$. When the SNR is higher, the algorithm correctly identifies all components even for the five components signals.

The algorithm's ability to cope with noise is exemplified in Figure 7 where 1000 realizations of each SNR level were simulated for the same two component signal used in Table 6, which parameters are found in Table 5 and $\rho$=2/3. The performance of the algorithm is immaculate down to 0 dB SNR, it still produces reliable detections in the range of 0 to −5 dB, and the performance deteriorates rather linearly from −5 dB to −12.5 dB. Considering the fact that a measured signal with −5 dB SNR would be regarded as a low quality signal in most applications, the algorithm can be described as rather robust to noise and insensitive to noise over 0 dB.

Since there, to our knowledge, exist no other researched methods, the performance of our algorithm can not be compared to other known methods.

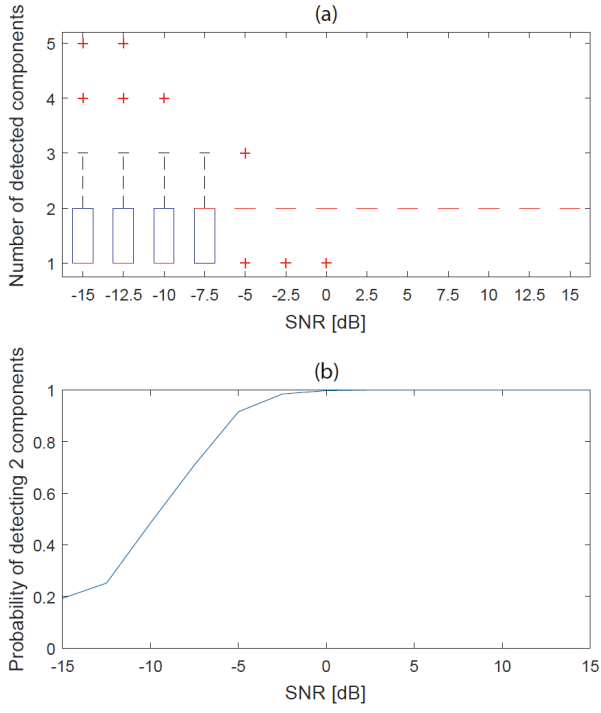



**Figure 7:** (a) Boxplot showing how the number of detected signal components for a two component signal vary over different levels of SNR. (b) Probability of detecting two components in a signal containing two components for different levels of SNR. Both plots are generated from 1000 realizations.

However when considering the ReSTS matrix, two other, perhaps simpler, approaches seem natural. The first approach is to calculate some threshold for the noise peak amplitudes. The distribution of the (scaled reassigned) noise peaks is unknown, however a threshold based on the Gaussian distribution might be reasonable, or the universal threshold for noise reduction using discrete wavelet transform [37]. The second approach is to look at the peaks sequentially from the largest peak and continue until the peak amplitudes drop and then level out, the peaks after the drop will then be assumed to be noise peaks.

We have implemented such schemes and evaluated the detection rate for the same test signals used previously. The wavelet universal threshold has a detection rate from 0.90 to 0.94 for the different cases. The Gaussian threshold only has detec-

tion rates around 0.27 to 0.29, which is too low to be useful. The level out approach has a detection from 0.58 to 0.62 for the cases with two components and 0.53 to 0.55 for the cases with five components.

## 6 Examples on measured data

This section shows how the proposed automatic component detection algorithm performs on real, measured signals, from two acoustic fields. Before using the method on measured data, the time window length of the ReSTS needs to be decided. The time window should have the same length as, e.g. time duration of, the transient signal components. If the time duration is not known, an appropriate length of the time window can be determined by evaluating the local Rényi entropy of the ReSTS for different lengths of the time window [31]. When an appropriate length is used, the energy concentration will be high and accordingly the Rényi entropy small. It is not essential that the length of the time window exactly matches the duration of the transient signal components, the ReSTS is stable for different window lengths.

### 6.1 Ultrasound pulse-echo measurements

The automatic component detection algorithm was tested on real measurements from a simple pulse-echo measurement set-up in water. A 2.1 MHz in-house built transducer functioned as both sender and receiver. An ultrasonic pulse was generated using a Panametrix Pulse/receiver Model 5072PR device (Panametrics-NDT, Waltham, MA) and was measured at 100 MHz sampling rate with a Textronix TDS 2002C oscilloscope (Techtronix, Inc., Beaverton, OR). The reflective object was a Plexiglas phantom with the shape of a solid stairway with step sizes ranging from 2 mm to 0.25 mm. A sketch of the phantom can be seen in Figure 8.

Measurements were taken in each transition between two steps of the Plexiglas phantom. This resulted in a total of four measured time signals containing two echo components, one from each of the two adjacent steps. These components thus had different relative time delays (two way travel times). The time plots in Figure 9 show the measured reflections and the red circles indicate where our method detected a signal component. The time separations of the detected components correspond to an estimation of the step sizes of the Plexiglas

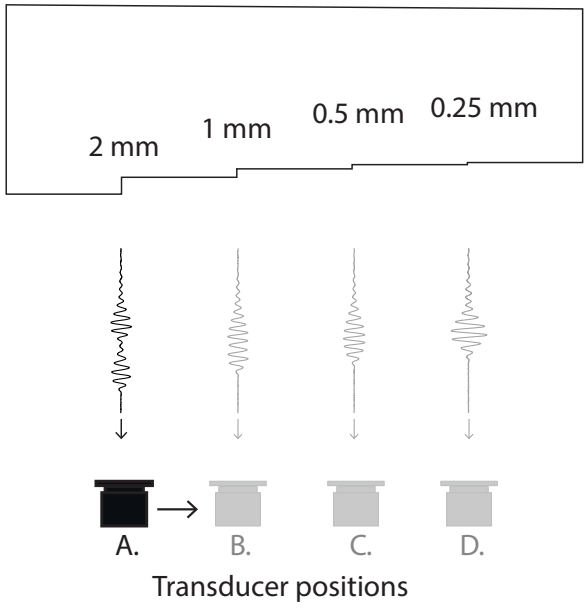



**Figure 8:** Set-up for the pulse-echo measurement of two surfaces, separated by 2, 1, 0.5 or 0.25 mm (positions A–D).

phantom, which are shown in the figure and should be compared with the step sizes in Figure 8.

The simple piezoceramic transducer with an approximate pulse length of 1 $\mu$s, full width at half maximum, can resolve surfaces separated by at least 1.5 mm. Thus only for case A, exemplified in Figure 9(a), it is possible to detect the two echoes visually. Our algorithm can for this transducer accurately estimate distances between two surfaces if they are larger or equal to 0.67 mm. This means that for cases A and B, Figure 9(a) and (b), the algorithm correctly identifies the TF centers of the two pulses. For case C shown in Figure 9(c), the pulses act more like one long pulse, not suited for the ReSTS, which is adapted for short signals. Still, two components are detected, although the distance is not correctly identified. For case D, Figure 9(d), only one component is detected, which is expected.

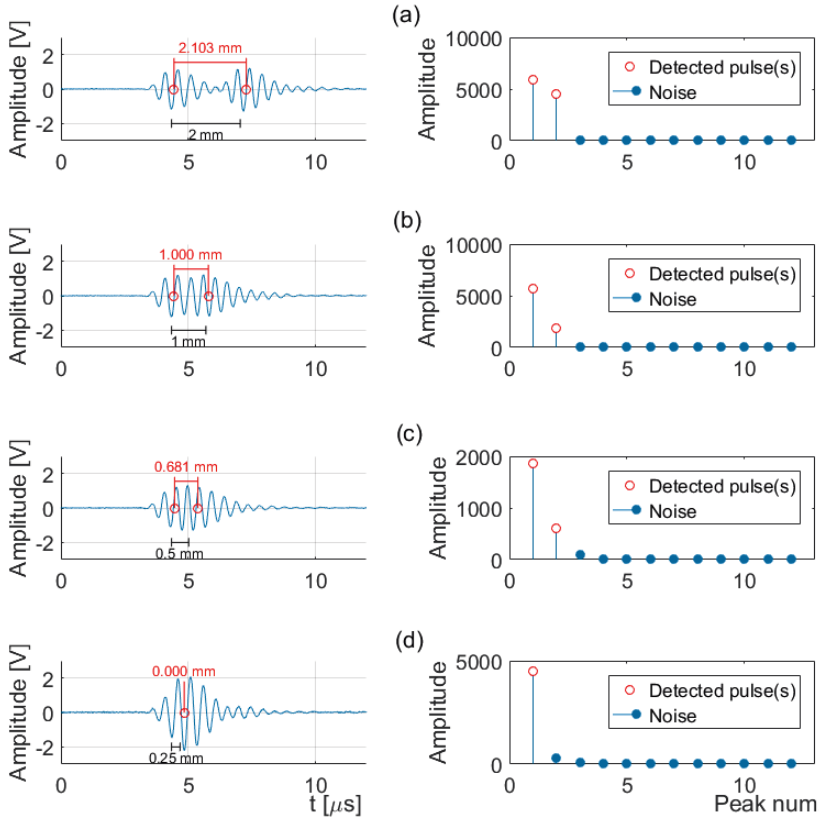



**Figure 9:** Time signals and peak detections from ReSTS, of pulse-echo measurements of two surfaces that are separated by 2, 1, 0.5 or 0.25 mm (a) – (d). The red circles indicate the detected time centers, obtained by the automatic detection method from the ReSTS. The time separation between the pulse centers correspond to an estimation of the step sizes of the Plexiglas phantom, Figure 8. The estimates are shown in red (above), the true step sizes are shown in black (under).

For our algorithm, the parameters were set to $\delta_t = 2\sigma = 0.9$ $\mu$s, $\delta_f = 1/(\pi\sigma) = 0.7$ MHz, $\rho = 1$ and $K_{max} = 4$. The length of the signal, and thus $\sigma$, could be estimated studying a single pulse using full width at half maximum. It is important to note that while we choose to use $\sigma = 0.45$ $\mu$s, equivalent results were obtained for $\sigma \in [0.38\ 0.65]$ $\mu$s. Thus for these signals our method is robust to choices of $\sigma$.

### 6.2 Marine biosonar signals

The usefulness of the developed automatic component detection algorithm is further exemplified by applying it to recorded dolphin echolocation signals. Recordings were made from different parts of the echolocation beam main lobe. Details regarding how these recordings were made are explained by [10].

Figure 10 left hand side, shows recordings from a bimodal transient signal, expected to contain more than one component. Figure 10 right hand side, shows recordings from a unimodal transient signal, expected to contain only one component. The algorithm outcome in Figure 10 (left) shows that it detects two signal components separated in both time and frequency, while in Figure 10 (right) only one component is detected.

For the ReSTS and the automatic component detection algorithm, $\sigma = 6.4$ $\mu$s, $K_{max} = 3$ and $\rho = 2/3$ are used. This means that the resolution in time is $\delta_t = 2\sigma = 13$ $\mu$s and in frequency $\delta_f = 1/(\pi\sigma) = 53$ kHz. However similar results are obtained for $\sigma \in [5\ 8]$ $\mu$s, so our method is robust to choices of $\sigma$ for these signals.

From a biosonar perspective, it is interesting to compare the signal component time and frequency centers in Figure 10. Although such a comparison lies outside the scope of this paper, it brings new information and insights to how the different parts of the echolocation beam of bottle nose dolphins (*Tursiops truncatus*) are generated in terms of suggested internal frequency filters, acoustic reflection pathways and possibly multiple echolocation sources [9, 10, 38, 39].

## 7 Conclusion

In this paper an automatic component detection method for short, multi-component transient signals has been proposed. The method combines a novel detection algorithm with a high-resolution TF representation adapted for short transient signals, ReSTS. As a result, the method automatically counts, the beforehand unknown, number of transient components and estimates the TF centers of individual components with great precision. The results are also easily visualized by the ReSTS, showing the TF centers of individual components, marked by the automatic algorithm.

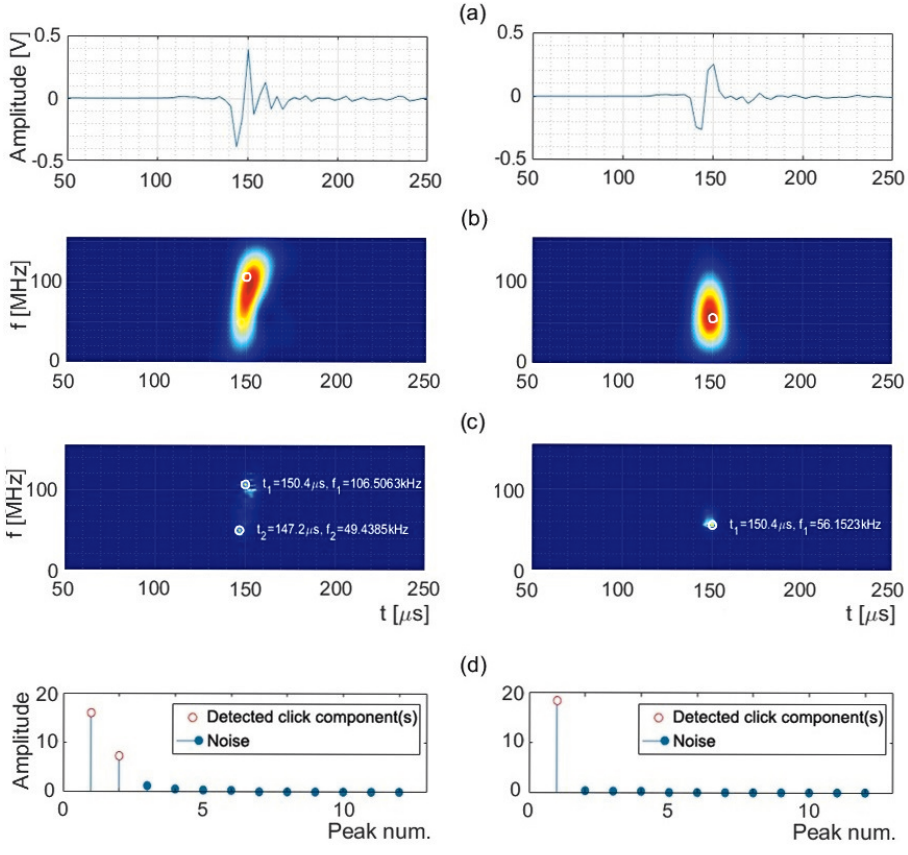



**Figure 10:** Two dolphin echolocation signals, to the left a bimodal, to the right a unimodal; (a) time signal; (b) spectrogram; (c) ReSTS representation and detected signal TF centers; (d) pulse detections.

Our proposed automatic component detection algorithm and TF representation are useful for severely coinciding Gaussian shaped transients, where other comparable methods fail to resolve the components. The resolution of the suggested TF representation is shown to be the best achievable for Gaussian shaped pulses. In addition, the calculations of time and frequency centers of the components are robust to noise. The method is easy to use since the algorithm requires very little user input. The paper provides guidelines on how to choose the input parameters, although the method is quite insensitive to parameter choices.

This new method should be useful in several acoustic research fields, not the least sonar and biosonar applications, or other fields where coinciding Gaussian shaped transients are analyzed. In this paper the method shows promising results on measured data, both from a laboratory pulse-echo set-up and a dolphin echolocation signal measured simultaneously at two different locations in the echolocation beam. The method resolves heavily overlapping pulses from the pulse-echo signals and automatically detects the expected number of components in the bimodal and unimodal echolocation signal.

## Acknowledgments

The authors would like to thank the Navy Marine Mammal Program and Dorian S. Houser, Patrick W. Moore and Lois Talmadge, National Marine Mammal foundation, USA, for their generous contribution of dolphin echolocation data. The research reported in this paper was supported by the Swedish Research Council.

## References

[1] C. Wei, W. W. L. Au, D. R. Ketten, Z. Song, and Y. Zhang, "Biosonar signal propagation in the harbor porpoise's (Phocoena phocoena) head: The role of various structures in the formation of the vertical beam.," *Journal of the Acoustical Society of America*, vol. 141, no. 6, pp. 4179, 2017.

[2] C. Capus, Y. Pailhas, K. Brown, D.M. Lane, P. Moore, and D. Houser, "Bio-inspired wideband sonar signals based on observations of the bottlenose dolphin (Tursiops truncatus)," *J. Acoust. Soc. Am.*, vol. 121, no. 1, pp. 594–604, 2007.

[3] J. Pons-Llinares, M. Riera-Guasp, J. A. Antonino-Daviu, and T. G. Habetler, "Pursuing optimal electric machines transient diagnosis: The adaptive slope transform.," *Mechanical Systems and Signal Processing*, vol. 80, pp. 553 – 569, 2016.

[4] Q. He and X. Ding, "Sparse representation based on local time-frequency template matching for bearing transient fault feature extraction.," *Journal of Sound and Vibration*, vol. 370, pp. 424 – 443, 2016.

[5] H. Liu, W. Huang, S. Wang, and Z. Zhu, "Adaptive spectral kurtosis filtering based on morlet wavelet and its application for signal transients detection.," *Signal Processing*, vol. 96, no. Part A, pp. 118 – 124, 2014.

[6] L. Cohen, *Time-Frequency Analysis*, Signal Processing Series. Prentice-Hall, Upper Saddle River, NJ, USA, 1995.

[7] W. W. L. Au, "Echolocation signals of the atlantic bottlenose dolphin (tursiops truncatus) in open waters," in *Animal Sonar Systems*, René-Guy Busnel and James F. Fish, Eds., pp. 251–282. Springer US, Boston, MA, 1980.

[8] W. W. L. Au, *The Sonar of Dolphins*, Springer-Verlag, 1993.

[9] W. W. L. Au, B. Branstetter, P.W. Moore, and J.J. Finneran, "Dolphin biosonar signals measured at extreme off-axis angles: Insights to sound propagation in the head," *J. Acoust. Soc. Am.*, vol. 132, no. 2, pp. 1199–1206, 2012.

[10] J. Starkhammar, P.W.B. Moore, L. Talmadge, and D.S. Houser, "Frequency-dependent variation in the two-dimensional beam pattern of an echolocating dolphin," *Biol. Lett.*, vol. 7, no. 6, pp. 836–839, 2011.

[11] H.I. Choi and W. J. Williams, "Improved time-frequency representation of multi-component signals using exponential kernels," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 37, pp. 862–871, June 1989.

[12] S. Wang, S. Huang, Q. Wang, Y. Zhang, and W. Zhao, "Mode identification of broadband Lamb wave signal with squeezed wavelet transform.," *Applied Acoustics*, vol. 125, pp. 91 – 101, 2017.

[13] J. J. Finneran, B. K. Branstetter, D. S. Houser, P. W. Moore, J. Mulsow, C. Martin, S. Perisho "High-resolution measurement of a bottlenose dolphin's (tursiops truncatus) biosonar transmission beam pattern in the horizontal plane," *J. Acoust. Soc. Am.*, vol. 136, no. 4, pp. 2025–2038, 2014.

[14] K. Kodera, C. de Villedary, and R. Gendrin, "A new method for the numerical analysis of nonstationary signals," *Physics of the Earth & Planetary Interiors*, vol. 12, pp. 142–150, 1976.

[15] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. on Signal Processing*, vol. 43, pp. 1068–1089, May 1995.

[16] I. Daubechies, J. Lu, and H.-T. Wu, "Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool.," *Applied & Computational Harmonic Analysis*, vol. 30, no. 2, pp. 243 – 261, 2011.

[17] 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, RG2006, 2008.

[18] S. Wang, X. Chen, Y. Wang, G. Cai, B. Ding, and X. Zhang, "Nonlinear squeezing time-frequency transform for weak signal detection," *Signal Processing*, vol. 113, pp. 195–210, August 2015.

[19] I. Reinhold and M. Sandsten, "Optimal time-frequency distributions using a novel signal adaptive method for automatic component detection," *Signal Processing*, vol. 133, pp. 250 – 259, 2017.

[20] E. Sejdić, I. Djurović, J. Jiang, "Time-frequency feature representation using energy concentration: An overview of recent advances," *Digital Signal Processing*, vol. 19, pp. 153 – 183, 2009.

[21] J. Lerga, V. Sucic, and B. Boashash, "An efficient algorithm for instantaneous frequency estimation of nonstationary multicomponent signals in low SNR," *EURASIP Journal on Advances in Signal Processing*, pp. 1–16, January 2011.

[22] B. Barkat and K. Abed-Meraim, "Algorithms for blind components separation and extraction from the time-frequency distribution of their mixture," *EURASIP Journal on Applied Signal Processing*, vol. 13, pp. 2025–2033, 2004.

[23] L. A. Cirillo and M. G. Amin, "Auto-term detection using time-frequency array processing," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 6, pp. 465–468, 2003.

[24] H. Guo, "A simple algorithm for fitting a Gaussian function [dsp tips and tricks]," *IEEE Signal Processing Magazine*, vol. 28, no. 5, pp. 134–137, Sept 2011.

[25] E. Kheirati Roonizi, "A new algorithm for fitting a Gaussian function riding on the polynomial background.," *IEEE Signal Processing Letters*, vol. 20, no. 11, pp. 1062 – 1065, 2013.

[26] S.C. Pei and S.G. Huang, "STFT with adaptive window width based on the chirp rate," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4065 – 4080, 2012.

[27] S. Ghofrani, "Matching pursuit decomposition for high-resolution direction of arrival.," *Multidimensional Systems and Signal Processing*, vol. 26, no. 3, pp. 693 – 716, 2015.

[28] P. Balazs, M. Dörfler, F. Jaillet, N. Holighaus, and G. Velasco, "Theory, implementation and applications of nonstationary Gabor frames.," *Journal of Computational and Applied Mathematics*, vol. 236, pp. 1481 – 1496, 2011.

[29] H. Kalbkhani and M. G. Shayesteh, "Stockwell transform for epileptic seizure detection from EEG signals.," *Biomedical Signal Processing and Control*, vol. 38, pp. 108 – 118, 2017.

[30] M. Hamidia and A. Amrouche, "A new robust double-talk detector based on the Stockwell transform for acoustic echo cancellation.," *Digital Signal Processing*, vol. 60, pp. 99 – 112, 2017.

[31] M. Hansson-Sandsten and J. Brynolfsson, "The scaled reassigned spectrogram with perfect localization for estimation of Gaussian functions," *IEEE Signal Processing Letters*, vol. 22, no. 1, pp. 100–104, January 2015.

[32] M. Sezgin, B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *Journal of Electronic Imaging*, vol. 13, no. 1, pp. 146, 2004.

[33] M. I. Sezan, "A peak detection algorithm and its application to histogram-based image data reduction," *Computer Vision, Graphics and Image Processing*, vol. 49, pp. 36 – 51, 1990.

[34] J. Olivo, "Automatic threshold selection using the wavelet transform," *CVGIP: Graphical Models and Image Processing*, vol. 56, no. 3, pp. 205, 1994.

[35] E. Chassande-Mottin, F. Auger, and P. Flandrin, "Reassignment," in *Time-Frequency Analysis*, F. Hlawatsch and F. Auger, Eds., chapter 9, pp. 249–277. U.K.: ISTE, London, 2008.

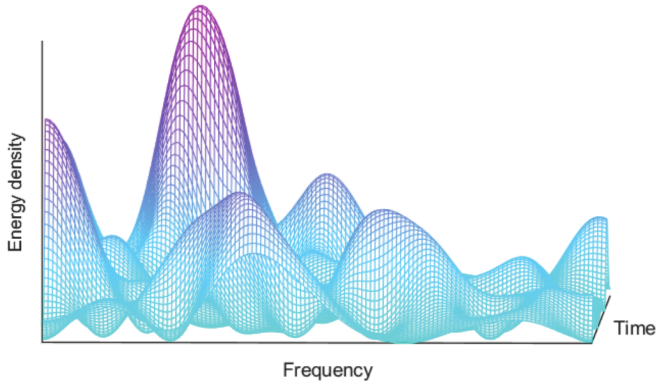
[36] H. Holzmann and S. Vollmer, "A likelihood ratio test for bimodality in two-component mixtures with application to regional income distribution in the EU," *AStA Advances in Statistical Analysis*, vol. 92, no. 1, pp. 57–69, Feb 2008.

[37] D. B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis*, Cambridge University Press, Cambridge, United Kingdom, July 2000.

[38] T. W. Cranford, W. R. Elsberry, W. G. Van Bonn, J. A. Jeffress, M. S. Chaplin, D. J. Blackwood, D. A. Carder, T. Kamolnick, M. A. Todd, and S. H. Ridgway, "Observation and analysis of sonar signal generation in the bottlenose dolphin (tursiops truncatus): Evidence for two sonar sources," *Journal of Experimental Marine Biology and Ecology*, vol. 407, no. 1, pp. 81 – 96, 2011.

[39] P. T. Madsen, M. Lammers, D. Wisniewska, and K. Beedholm, "Nasal sound production in echolocating delphinids (tursiops truncatus and pseudorca crassidens) is dynamic, but unilateral: clicking on the right side and whistling on the left side," *Journal of Experimental Biology*, vol. 216, no. 21, pp. 4091–4102, 2013.

Printed by Media-Tryck, Lund 2021 NORDIC SWAN ECOLABEL 3041 0903




It might seem like wishful thinking that we in the near future will be able to control technical devices with our minds, or that we will be able to make reliable diagnoses using our own mobile phones. With new methods in signal processing and spectral analysis, this will however soon be our reality. We are already used to voice commands on our mobile phones and robot vacuum cleaners in our homes, which would have seemed like pure science fiction only a few decades ago. Today's standard mobile phone is several million times faster than the most advanced computers of the 1960's, and this technological revolution has greatly expanded the possibilities for spectral analysis. The research in this book focuses on the development of spectral methods for extraction of useful information from a variety of signals, including heart rate (ECG), brain activity (EEG), and dolphin echolocation.

LUND UNIVERSITY

Doctoral Theses in Mathematical Sciences 2021:3
ISBN 978-91-7895-803-0
ISSN 1404-0034