High resolution time-frequency representations

Reinhold, Isabella

2017

Document Version:
Annan version

High resolution time-frequency representations

by

Isabella Reinhold

LUND
UNIVERSITY

Thesis for the degree of Licentiate of Engineering
To be presented, with the permission of the Faculty of Engineering at Lund University, for public criticism in room MH:309A, Mattehuset, Sölvegatan 18, Lund, on Tuesday, the 28th of November 2017 at 13:15 (1.15pm).

Faculty opponent: Prof. Sven Nordebo,
Department of Physics and Electrical Engineering, Linnaeus University, Sweden.

Organization
LUND UNIVERSITY
Centre for Mathematical Sciences
Mathematical Statistics
Box 118
SE–221 00 Lund, Sweden

Document name
LICENTIATE THESIS
Date of disputation
2017-11-28
Sponsoring organization

Author(s)
Isabella Reinhold

Title and subtitle
High resolution time-frequency representations

Abstract
Non-stationary signals are very common in nature, e.g. sound waves such as human speech, bird song and music. It is usually meaningful to describe a signal in terms of time and frequency. Methods for doing so exist and are well defined. From the time representation it is possible to see the oscillations or waves of the signal and if the signal changes over time. From the frequency representation, obtained from the Fourier transform, the frequency decomposition of the signal can be seen, i.e. which frequencies the signal contains.

However the time and frequency representations are not unique for any given signal, i.e. the transformation from time to frequency is not injective. It is therefore, especially for non-stationary and multi-component signals, important to study a joint time-frequency (TF) representation, which shows how the frequency content of the signal varies with time. This is done in the field of time-frequency analysis, which is the topic of this thesis.

There exist many different joint TF representations for any given signal and choosing an appropriate representation is most often not straight forward. Unfortunately there exist no optimal TF representation for all signals and finding good representations, especially for multi-component signals is a complex problem.

In this thesis, methods for obtaining good TF representations, for two types of non-stationary and multi-component signals, and for extracting meaningful information from these representations, are developed. The two types of signals are long, frequency modulated signals and short, transient signals. Even though the types of signals are very different and require very different TF representations, the aim is to resolve components that are close in time, frequency or both. This requires TF representations with high resolution.

For the long, frequency modulated signals, a signal adaptive method, which enables automatic comparison between different TF representations, is proposed. For the short, transient signals, a method which finds the TF centres of transient pulses and counts the number of pulses in a signal is presented. An approach for determining the (time) shape of transient pulses is also given.

Key words
time-frequency analysis, non-stationary signals, multi-component signals, IF estimation, reassignment

Classification system and/or index terms (if any)

Supplementary bibliographical information

Language
English

Licentiate Theses in Mathematical Sciences
2017:2

ISSN
1404-028X

Recipient's notes

Number of pages
134

ISRN
LUTFMS-2019-2017

Security classification

High resolution time-frequency representations

Isabella Reinhold

Lund University

Faculty of Engineering
Centre for Mathematical Sciences
Mathematical Statistics

Public defence
November 28th, 2017, 13:15 in MH:309A, Centre for Mathematical Sciences, Lund University.
Sölvegatan 18, 223 62 Lund, Sweden.

Advisors
Prof. Maria Sandsten,
Mathematical Statistics, Centre for Mathematical Sciences, Lund University, Sweden.
Assoc. Prof. Josefin Starkhammar,
Department of Biomedical Engineering, Lund University, Sweden.

Faculty opponent
Prof. Sven Nordebo,
Department of Physics and Electrical Engineering,
Linnaeus University, Sweden.

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

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

Licentiate Theses in Mathematical Sciences 2017:2
ISSN: 1404-028X
ISRN: LUTFMS-2019-2017

Printed in Sweden by Media-Tryck, Lund University, Lund 2017

To everyone who has given me free coffee

Contents

List of publications

This thesis is based on the following publications:

A **Optimal time–frequency distributions using a novel signal adaptive method for automatic component detection**

Isabella Reinhold, Maria Sandsten.
Signal Processing, Vol. 133, 250-259, 2017.

B **Objective detection and time-frequency localization of components within transient signals**

Isabella Reinhold, Maria Sandsten, Josefin Starkhammar.
Submitted to J. Acoust. Soc. Am.

C **The scaled reassigned spectrogram adapted for detection and localisation of transient signals**

Isabella Reinhold, Josefin Starkhammar, Maria Sandsten.
25th European Signal Processing Conference (EUSIPCO), 937-941, 2017.

Publications not included in this thesis:

1 **Quality of academic writing for engineering students at Lund University**

Isabella Reinhold, Kenneth Batstone, Isabel M. Gallardo González, Andrea Troian, Rixin Yu.
The 2nd EuroSoTL conference, 2017.

2 **Automatic time-frequency analysis of echolocation signals using the matched Gaussian multitaper spectrogram**

Maria Sandsten, Isabella Reinhold, Josefin Starkhammar.
Interspeech 2017, 3048-3052, 2017.

3 **A novel Doppler penalty function for the multitaper Wigner-Ville distribution**

Isabella Reinhold, Maria Sandsten.
11th IMA International Conference on Mathematics in Signal Processing, 2016.

4 **Intra-click time-frequency patterns across the echolocation beam of a beluga whale**

Josefin Starkhammar, Isabella Reinhold, Patrick Moore, Dorian Houser, Maria Sandsten.
J. Acoust. Soc. Am., Vol. 140 (4), 3239, 2016.

Acknowledgements

I wish to express my sincere thanks to my supervisor Maria Sandsten for her support and guidance during these two years. I would also like to thank my assistant supervisor Josefin Starkhammar, who have introduced interesting applications for my work. A heartfelt thanks also to the present and previous administrative and technical staff at the mathematical statistics department for their help in practical matters. Thanks also to Rachele, who has shared the stress of simultaneously teaching and writing a thesis. Lastly, thanks to my parents Jeanette and Kent, my wonderful boyfriend Oskar, my brother Viktor, my first roommate Åse and all my friends for always being there in times of need.

Isabella Reinhold

Abstract

Non-stationary signals are very common in nature, e.g. sound waves such as human speech, bird song and music. It is usually meaningful to describe a signal in terms of time and frequency. Methods for doing so exist and are well defined. From the time representation it is possible to see the oscillations or waves of the signal and if the signal changes over time. From the frequency representation, obtained from the Fourier transform, the frequency decomposition of the signal can be seen, i.e. which frequencies the signal contains.

However the time and frequency representations are not unique for any given signal, i.e. the transformation from time to frequency is not injective. It is therefore, especially for non-stationary and multi-component signals, important to study a joint time-frequency (TF) representation, which shows how the frequency content of the signal varies with time. This is done in the field of time-frequency analysis, which is the topic of this thesis.

There exist many different joint TF representations for any given signal and choosing an appropriate representation is most often not straight forward. Unfortunately there exist no optimal TF representation for all signals and finding good representations, especially for multi-component signals is a complex problem.

In this thesis, methods for obtaining good TF representations, for two types of non-stationary and multi-component signals, and for extracting meaningful information from these representations, are developed. The two types of signals are long, frequency modulated signals and short, transient signals. Even though the types of signals are very different and require very different TF representations, the aim is to resolve components that are close in time, frequency or both. This requires TF representations with high resolution.

For the long, frequency modulated signals, a signal adaptive method, which enables automatic comparison between different TF representations, is proposed. For the short, transient signals, a method which finds the TF centres of transient pulses and counts the number of pulses in a signal is presented. An approach for determining the (time) shape of transient pulses is also given.

Keywords: time-frequency analysis, non-stationary signals, multi-component signals, IF estimation, reassignment

Popular summary in English

Non-stationary signals are very common in nature, e.g. sound waves such as human speech, bird song and music. In the fields of mathematical statistics and signal processing, it is interesting to study how these signals look and behave in time and frequency. The time representation shows when something happens in the signal. An example can be a recording of a person speaking, from the time representation it is possible to see when the person starts to talk and when it stops. The frequency representation shows which frequencies are present in the signal, but not when. The frequency decides how something sounds, e.g. a person can have a deep voice and another a shrill voice, they then speak with different frequencies even if they say the same things.

This thesis is in the field of time-frequency analysis, which means that it aims to study signals in time and frequency at the same time. If the joint time-frequency representation is studied, it can be seen when something happens in the signal and what happens at that time, i.e. what frequencies appear at any given time in a signal.

The start of frequency analysis of signals can be traced to the beginning of the 1700s with Sir Isaac Newton (1642-1727), and the modern frequency analysis of measured signals was invented in the late 1800s. The discoveries that made joint time-frequency analysis possible were done in the mid-1900s and since then it has been a field of interest for many mathematicians and other scientists working with signal processing.

It is perhaps a hint, since the field has existed for almost a hundred years, that representing a signal in time and frequency at the same time is not simple or straight forward. There exist no optimal representation for all non-stationary signals, or even for specific types of non-stationary signals. It is therefore possible to still make contributions to this interesting and complex field.

In this thesis a few of the applications of time-frequency analysis are detailed. The results in paper A are used to better characterise measured heart rate variability signals. These signals are the variation of inter-heartbeat intervals, which are measured non-invasively using ECG, and can be used to see if a person suffers from some health issues. Improved characterisation of these signals can result in better medical care, especially for very small babies.

The results in papers B and C are used to study the echolocation signals of dolphins and with the methods presented, especially in paper B, it is possible to get a more accurate picture of how the echolocation signals behave. This is useful information for understanding how the signal is generated by the dolphin.

The method in paper B, can also be used for improved characterisation of similar signals, i.e. signal that have very, very short duration, which are common in fields such as ultrasonic signal analysis machine fault diagnosis and biomedical signal processing. Essentially fields that focus on ”looking” at something, e.g. a machine or human, using signals and recordings of the reflections, instead of opening that something up and actually looking at it.

Populärvetenskaplig sammanfattning på svenska

Icke-stationära signaler är väldigt vanliga i naturen, t.ex. ljudvågor så som mänskligt tal, fågelsång och musik. Det är inom fälten matematisk statistik och signalbehandling intressant att studera hur dessa signaler beter sig i tid och frekvens. Tidsrepresentationen visar när något händer i en signal. Ett exempel kan vara en inspelning av en person som pratar, från tidsrepresentationen är det då möjligt att se när personen börjar och slutar prata. Frekvensrepresentationen visar vilka frekvenser som finns i en signal, men inte när de finns. Frekvensen bestämmer hur något låter, t.ex. kan en viss person ha en djup röst medan en annan har en gäll röst, de personerna pratar med olika frekvenser, även om de säger samma sak.

Denna avhandling ligger inom fältet för tids-frekvensanalys, vilket betyder att målet är att studera signaler i tid och frekvens samtidigt. Om den gemensamma tids-frekvensrepresentationen studeras, så är det möjligt att se när något händer i en signal och vad det är som händer vid den tiden, alltså vilka frekvenser som finns vid en given tidpunkt i signalen.

Frekvensanalysens uppkomst kan spåras till 1700-talets början och Sir Isaac Newton (1642-1727), medan den moderna frekvensanalysen med uppmätta signaler tog vid på sent 1800-tal. Upptäckterna som banade väg för en gemensam tids-frekvensanalys gjordes under mitten av 1900-talet och sedan dess har fältet varit av intresse för många matematiker och andra forskare som arbetar med signalbehandling. Att fältet har existerat i nästan hundra år kan vara en ledtråd om att det inte är ett helt trivialt att uttrycka en signal i tid och frekvens samtidigt. Det existerar ingen optimal representation för alla icke-stationära signaler, eller ens för specifika typer av icke-stationära signaler. Därför är det fortfarande möjligt att bidra till detta intressanta och komplexa fält.

Några tillämpningar för tids-frekvensanalys presenteras i den här avhandlingen. Resultaten i artikel A används till att bättre beskriva uppmätta HRV-signaler, som mäter variationer mellan hjärtslag. Dessa signaler mäts med EKG och kan visa om en person lider av vissa åkommor. En förbättrad beskrivning av dessa signaler kan innebära bättre medicinsk vård, speciellt för spädbarn. Resultaten i artiklarna B och C används för att studera ekolokaliseringssignaler från delfiner, och med metoden som presenteras i artikel B, så är det möjligt att få en tydligare bild av

hur ekolokaliseringssignaler beter sig. Detta är viktigt för att förstå hur delfinerna generar sina ekolokaliseringssignaler.

Metoden i artikel B kan också användas för att på ett bättre sätt beskriva liknande signaler, alltså signaler som är väldigt, väldigt korta. Dessa signaler är vanliga i fält som berör ultraljudsanalys, diagnos av fel i maskiner och biomedicinsk signalbehandling. Med andra ord fält som ”tittar” på objekt, t.ex. maskiner eller människor, med hjälp av signaler och de reflektioner signalerna ger upphov till, istället för att faktiskt öppna upp och titta i objektet.

Summary of notations

In this thesis the following notations are used:

Integrals without limits if no limits are given for an integral, they are $-\infty$, ∞

$$\int \equiv \int_{-\infty}^{\infty}$$

Fourier transform pairs if a signal is denoted $s(t)$, its Fourier transform is denoted $S(f)$

$$S(f) \equiv \int s(t) e^{-i2\pi ft} dt, \quad s(t) \equiv \int S(f) e^{i2\pi ft} df$$

Imaginary unit if nothing else is stated, i denotes the imaginary unit

$$i^2 \equiv -1$$

Complex conjugate for any function denoted $f(t)$, its complex conjugate is denoted $f^*(t)$

$$f(t) = a(t) + ib(t) \Longleftrightarrow f^*(t) = a(t) - ib(t)$$

Convolution for any two functions $f(t)$, $g(t)$, their convolution is denoted $f(t) * g(t)$

$$f(t) * g(t) \equiv \int f(\tau) g(t - \tau) d\tau$$

High resolution time-frequency representations

1 The need for time-frequency analysis

It is generally meaningful to describe a signal in terms of time and frequency. Methods for doing so exist and are well defined. From the time representation it is possible to see the oscillations or waves of the signal and if the signal changes over time. From the frequency representation, invented by Fourier (1807) and developed by Bunsen and Kirchhoff (around 1865), the frequency decomposition of the signal can be seen, i.e. which frequencies the signal contains.

A signal can have a constant frequency for its whole duration or the frequency could change over time, if the frequency changes with time the signal is called non-stationary. These signals are very common in nature, for example sound waves such as human speech, bird song and music are all signals which usually vary in frequency. Non-stationary signals can take any shape or form and have very complex natures, while there of course also exist fairly simple signals.

Figures 1 illustrates two such simple signals, a chirp with linearly increasing frequency

$$s(t) = a_1 \cos\left(2\pi\left(f_0 + f_I t\right) t\right), \; t \geq 0, \tag{1}$$

and a sinc

$$s(t) = \frac{a_2 \sin(\pi t)}{\pi t}, \; t \geq 0, \tag{2}$$

where a_k is a constant amplitude, f_0 a constant frequency and f_I a constant frequency increase. The time representations of the chirp and sinc are shown in

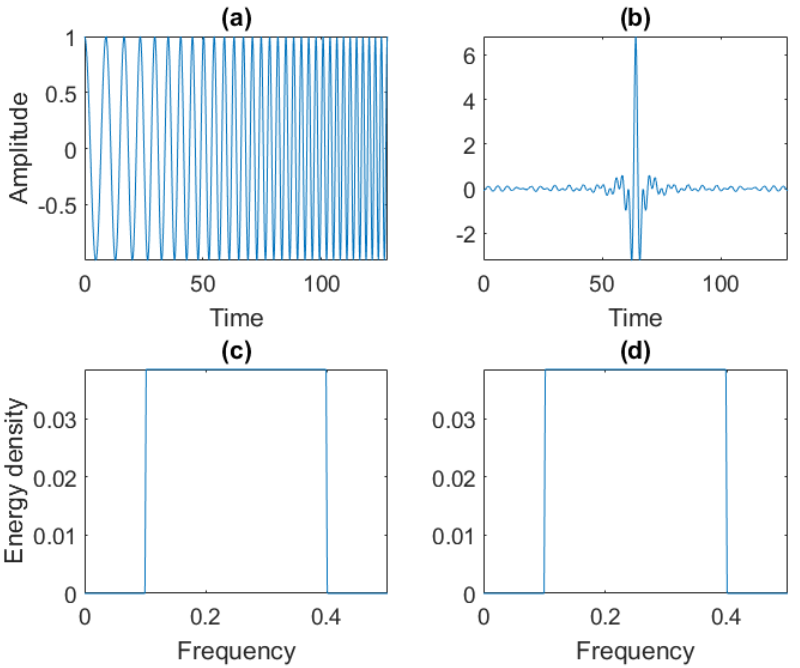

Figure 1: Separate time and frequency representations of two signals; (a) time-domain representation of a chirp; (b) time-domain of a sinc; (c) frequency-domain representation of the chirp; (d) frequency-domain representation of the sinc.

Figure 1 (a) and (b) respectively, they are clearly different as the chirp has one specific frequency at any given time, while the sinc has many frequencies at any given time. Especially for the sinc it will be hard, even impossible, to estimate the frequency content by only studying the time representation, it is therefore important to look at the frequency representations of the signals, shown in Figure 1 (c) and (d). It can be seen that the signals have the same frequency content and the only difference between these signals is when in time certain frequencies occur.

From these two examples it is also apparent that the frequency representation is not unique for any given signal, i.e. the transformation from time to frequency is not injective. It is therefore important to study both the time and frequency

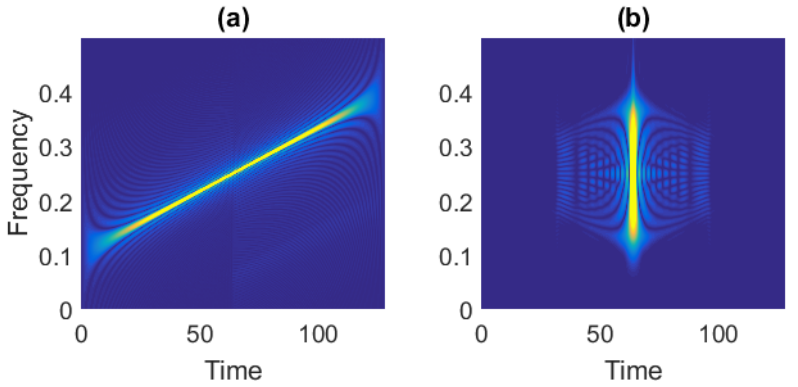

Figure 2: Joint time-frequency representations of two signals; (a) representation of a chirp; (b) representation of a sinc. Light yellow represents high energy density and dark blue low.

representations of a signal. However sometimes this is not enough, it is still not clear when in time certain frequencies occur. To know this, a joint time-frequency (TF) representation would have to be studied.

Figure 2 (a) shows a joint TF representation of the chirp and (b) of the sinc. From these representations, the frequency at any given instant in time can be known. A TF representation shows the distribution of signal energy for its time duration and frequency bandwidth, in the figure, light yellow represents a high energy concentration and dark blue a low concentration. In Figure 2 (a), the chirp, it can be seen that the frequency increases linearly with time and that the signal has the same signal energy for the whole signal duration. In Figure 2 (b), the sinc, most of the signal energy is concentrated at the midpoint of the duration and all frequencies within the signal occurs at that one time instant.

There exist many different joint TF representations for any given signal and choosing an appropriate representation is most often not straight forward. Some of the challenges can be seen in Figure 2. The linear chirp has the same signal energy for its whole duration, however it can be seen i Figure 2 (a) that the energy seem to be the highest midway through the signal. This is due to that the signal duration is not infinite. There is also an uncertainty in time and frequency, illustrated by the width of the high energy lines in the TF representations, even though the uncertainty is not huge for these simple signals. Smaller oscillations

outside the main, high energy line can also be seen, these so called side lobes are also a result of the finite signal length. When a signal has non-linear behaviour or discontinuities in time and/or frequency, or is disturbed by noise, more problems arise.

Non-stationary and multi-component signals, i.e. signals that vary in frequency over time and has discontinuities in time and/or frequency, can be modelled in different ways. It is natural to imagine a real valued signal

$$s(t) = A(t)\cos\left(\varphi(t)\right) = A(t)\cos\left(2\pi f(t)t\right), \quad t \geq 0, \tag{3}$$

where $A(t)$ is an amplitude function and $\varphi(t) = 2\pi f(t)t$ denotes the different frequencies present in the signal. A complex valued signal, on the other hand is less intuitive, while still useful. The complex signal can be separated into an amplitude function and a phase function according to

$$s(t) = A(t)e^{i\varphi(t)} = A(t)e^{i2\pi f(t)t}, \quad t \geq 0. \tag{4}$$

It is also possible to model a non-stationary and multi-component signal as a sum of time, frequency and phase shifted functions

$$s(t) = \sum_{k=1}^{K} a_k x_k(t - t_k)e^{i2\pi f_k t}e^{i2\pi\varphi_k}, \quad t \geq 0, \tag{5}$$

where a_k are amplitudes, t_k and f_k are time and frequency centres, φ_k phase shifts and $x_k(t)$ some appropriate functions, perhaps the Hermite functions. The sum is in this case a complex valued signal, but can be made real valued by simply taking the real value of the whole signal or individual signal components. Regardless of how the signal is modelled or looks, the goal of TF analysis is to accurately represent it in the TF plane.

Unfortunately there exist no optimal TF representation for all signals. Finding good representations, especially for multi-component signals, is a complex problem and still a large field of research [1, 2, 3, 4, 5]. Good TF representations are also useful in many applied fields, e.g. in biomedical fields accurate TF representations can result in better medical diagnoses using non-invasive procedures [6, 7], faults in machines can be detected and diagnosed [8, 9] and the acoustic pathway of echolocation dolphins can be understood [10, 11, 12].

2 Fundamental ideas for joint time-frequency representations

A joint TF representation could be a two dimensional density, that is a measure, $P(t,f)$, of the amount of something per unit t and per unit f at the point (t,f). The normalised total amount of the something is

$$\int\int P(t,f)dtdf = 1. \tag{6}$$

In the area of TF analysis the terms density and distribution are used interchangeably and a two dimensional TF density is usually called a TF distribution (TFD). The density $P(t,f)$ is by definition always non-negative.

The something that $P(t,f)$ measures, is the intensity of the signal at the point (t,f). The total energy of the signal, i.e. the amount of energy required to produce the signal, is calculated by integrating over the whole time and all frequencies. If the total energy is normalised according to equation (6), then $P(t,f)$ is the fraction of energy at that time and frequency.

For a signal $s(t)$ its Fourier transform $t \rightarrow f$ can be called $S(f)$. Then $|s(t)|^2$ is the intensity per unit time at time instant t and $|S(f)|^2$ is the intensity per unit frequency at frequency instant f. For a TFD it is not always true that the total energy of the distribution equals the total energy of the signal, i.e. that the following equalities hold

$$E_{tot} = \int\int P(t,f)dtdf = \int |s(t)|^2 dt = \int |S(f)|^2 df, \tag{7}$$

where E_{tot} is the total energy of the signal. If the above equalities hold however, the TFD is said to fulfil the total energy requirement. This requirement is weak and many TFDs which do not fulfil the total energy requirement still give a good representation of the TF structure of the signal.

2.1 Desired characteristics

Some fundamental properties are desired for a TFD that stem from the nature of signals. Since a signal easily can be translated in time or frequency, it is desired

for a TFD to be time and frequency shift invariant so that

$$\begin{aligned} s(t) \rightarrow s(t-t_0) &\Longrightarrow P(t,f) \rightarrow P(t-t_0,f), \\ S(f) \rightarrow S(f-f_0) &\Longrightarrow P(t,f) \rightarrow P(t,f-f_0), \\ s(t) \rightarrow e^{i2\pi f_0 t} s(t-t_0) &\Longrightarrow P(t,f) \rightarrow P(t-t_0,f-f_0). \end{aligned} \tag{8}$$

Signals can also be scaled and as the Fourier transform of a scaled signal is

$$s(t) \rightarrow s(at) \Longrightarrow S(f) \rightarrow \frac{1}{a} S(f/a), \quad a > 0, \tag{9}$$

a signal which is compressed in time will have a Fourier transform which is expanded. The TFD should fulfil

$$s(t) \rightarrow s(at) \Longrightarrow P(t,f) \rightarrow P(at,f/a), \tag{10}$$

to have the same properties as the time representation and Fourier transform of the signal.

For the TFD to be a good representation of the signal, it is desired that it at least has weak finite support, i.e. so that

$$\begin{aligned} s(t) = 0, \text{ for } t \text{ outside } (t_1,t_2) &\Longrightarrow P(t,f) = 0, \text{ for } t \text{ outside } (t_1,t_2), \\ S(f) = 0, \text{ for } f \text{ outside } (f_1,f_2) &\Longrightarrow P(t,f) = 0, \text{ for } f \text{ outside } (f_1,f_2). \end{aligned} \tag{11}$$

A TFD with weak finite support is thus zero outside the time duration and frequency bandwidth of a signal. Strong finite support is also desired, although many TFDs do not fulfil this requirement. Strong finite support is fulfilled if

$$\begin{aligned} s(t_1) = 0 &\Longrightarrow P(t_1,f) = 0, \\ S(f_1) = 0 &\Longrightarrow P(t,f_1) = 0. \end{aligned} \tag{12}$$

In order to have strong finite support, the TFD then needs to be zero for all time and frequency instants that do not exist in the signal. This requirement is relevant for multi-component signals, which has gaps in time and/or frequency, i.e. areas $s(t) = 0$ or $S(f) = 0$, surrounded by areas where the signal energy is not zero.

2.2 Marginals

It is possible to get the one dimensional densities $P(t)$ and $P(f)$ from a two dimensional density $P(t,f)$. The density $P(t)$ describe the density of energy per unit t irrespective of f and similarly for $P(f)$. The one dimensional densities are obtained by integrating over the other variable

$$P(t) = \int P(t,f)df, \quad P(f) = \int P(t,f)dt. \tag{13}$$

These densities are called marginals, however the above equalities are not satisfied for all TFDs. A TFD is said to fulfil the marginals if it fulfils the equalities.

For a TFD that fulfils the marginals, the integral over all frequencies of the density equals the intensity per unit time at time instant t. The integral over the whole time similarly equals the intensity per unit frequency at frequency instant f. This can be expressed

$$\int P(t,f)df = |s(t)|^2, \quad \int P(t,f)dt = |S(f)|^2. \tag{14}$$

If a TFD fulfils the marginals, it also fulfils the total energy requirement, however the converse is not true.

2.3 Uncertainty principle

An essential fact in signal processing is that a signal can not both have finite duration and limited bandwidth. This is the reason for the smaller oscillations, side lobes, in the TFDs of Figure 2. The signals used for the TFD calculations are obviously finite in time, since they are sampled, which means that they can not be limited in frequency, even if that perhaps was intended, resulting in relatively small, but infinite, oscillations in both time and frequency.

It is possible to calculate the time and frequency standard deviations for a signal

$$\begin{aligned} T^2 &= \int (t-\bar{t})^2 \, |s(t)|^2 dt, \\ B^2 &= \int \left(f-\bar{f}\right)^2 |S(f)|^2 df, \end{aligned} \tag{15}$$

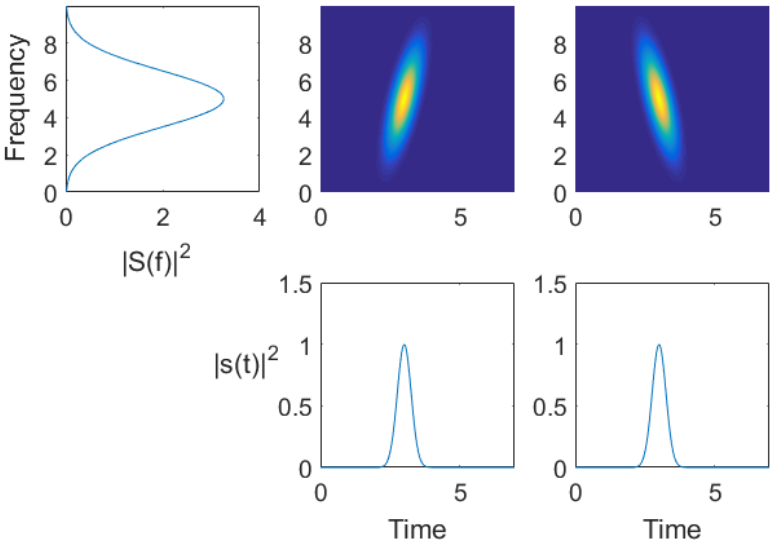

Figure 3: Two different signals with equal time and frequency marginals.

where $\bar{t}$ and $\bar{f}$ are the mean time and frequency. The uncertainty principle is then defined by

$$TB \geq \frac{1}{4\pi}, \tag{16}$$

and this applies to all signals [13]. A signal can thus not be constructed to have both standard deviations, T and B, arbitrarily small. The lower bound can be reached in theory using a Gaussian signal since it has the optimal concentration in time and frequency [14].

The uncertainty principle only depends on the marginals, $|s(t)|^2$ and $|S(f)|^2$, and not the whole density $P(t,f)$. This means that even if there are restrictions on the smallness of T and B, there is possibility for infinite variations of signals, since the marginals will not show how time and frequency are correlated. Thus two different signals might have the same marginals, a simple example is shown in Figure 3.

For a TFD the standard deviations are obtained by

$$\begin{aligned}\sigma_t^2 &= \int\int (t-\bar{t})^2\, P(t,f)dt = \int (t-\bar{t})^2\, P(t)dt,\\ \sigma_f^2 &= \int\int \left(f-\bar{f}\right)^2 P(t,f)df = \int \left(f-\bar{f}\right)^2 P(f)df.\end{aligned} \tag{17}$$

This means that the correct uncertainty principle is obtained if the TFD fulfils the marginals.

2.4 Instantaneous frequency

With the complex valued signal it is possible to define an operator $\mathcal{W}$ such that

$$\mathcal{W}s(t) = \mathcal{W}A(t)e^{i\varphi(t)} = \frac{1}{i}\frac{d}{dt}A(t)e^{i\varphi(t)} = \left(\varphi'(t) - i\frac{A'(t)}{A(t)}\right)s(t). \tag{18}$$

This can be used to calculate the mean frequency

$$\begin{aligned}\bar{f} = \int f\,|S(f)|^2 df &= \int\int\int f\; s^*(t)s(\tau)\; e^{i2\pi(t-\tau)f}\, df\; d\tau\, dt\\ &= \frac{1}{i2\pi}\int\int\int s^*(t)s(\tau)\frac{\partial}{\partial t}e^{i2\pi(t-\tau)f}\, df\; d\tau\, dt,\end{aligned} \tag{19}$$

for the equality it was used that

$$\frac{\partial}{\partial t}e^{i2\pi(t-\tau)f} = i2\pi f\; e^{i2\pi(t-\tau)f}, \tag{20}$$

and since

$$\delta(t-\tau) = \int e^{i2\pi(t-\tau)f}\, df, \tag{21}$$

the mean frequency is

$$\bar{f} = \frac{1}{i2\pi}\int\int s^*(t)s(\tau)\frac{\partial}{\partial t}\delta(t-\tau)d\tau dt. \tag{22}$$

Now it can be utilised that

$$\int s(\tau)\frac{\partial}{\partial t}\delta(t-\tau)\, d\tau = \frac{d}{dt}s(t), \tag{23}$$

which gives

$$\begin{aligned}\bar{f} &= \frac{1}{2\pi}\int s^*(t)\frac{1}{i}\frac{d}{dt}s(t)dt \\ &= \frac{1}{2\pi}\int s^*(t)\left(\varphi'(t) - i\frac{A'(t)}{A(t)}\right)s(t)dt \\ &= \frac{1}{2\pi}\int |s(t)|^2\left(\varphi'(t) - i\frac{A'(t)}{A(t)}\right)dt,\end{aligned} \tag{24}$$

here the second term is zero, easily realised since it is imaginary and the mean, $\bar{f}$, is real valued. This means that

$$\bar{f} = \frac{1}{2\pi}\int \varphi'(t)|s(t)|^2 dt, \tag{25}$$

which says that the average frequency is given by integrating the derivative of the phase with the density over all time. The derivative of the phase must then be the instantaneous value of the frequency, it is therefore appropriate to call

$$f_k(t) = \frac{\varphi'(t)}{2\pi}, \tag{26}$$

the frequency at each time or, the instantaneous frequency (IF).

The conceptual idea of IF is rather intuitive, as it would be the frequency of a signal at a given time instant. However the mathematical description and understanding of IF is not equally simple, e.g. for a general real valued signal the IF would be zero, since it does not have the phase signal $e^{i\varphi(t)}$. This absurd result can be rectified by defining a complex valued signal that corresponds to a given real valued signal, this is done in the next section. The mathematical definition of the IF also gives paradoxical results for multi-component, complex valued signals, and the IF can sometimes extend outside the bandwidth of the signal. Nevertheless the term IF is widely used in signal processing when estimating the frequencies in signals for any given time instant.

2.5 The analytic signal

In TF analysis it is useful to define a complex valued signal that corresponds to a given real valued signal. Such a complex valued signal would be

$$z(t) = s_r(t) + s_i(t), \tag{27}$$

where $s_r(t)$ is the real valued signal and $s_i(t)$ is an imaginary part. A property of any real valued signal can be utilised for a good choice of $s_i(t)$. The Fourier transform of a real valued signal is symmetric, $S_r(f) = S_r(-f)$, thus the real valued signal contains more information of the frequency than is needed. If the imaginary part is chosen so that

$$Z(f) = 0, \quad f < 0, \tag{28}$$

then

$$z(t) = 2\int_0^\infty S(f)e^{i2\pi ft}df = s(t) + i\left(\frac{1}{\pi t} * s(t)\right), \tag{29}$$

which is called the analytic signal. It is very useful when calculating some of the most common TFDs and can also be obtained from a general complex valued signal using equation (29).

3 Spectrogram and quadratic time-frequency distributions

One of the most common TFDs is the spectrogram, which is obtained from the short-time Fourier transform (STFT). The idea of the STFT is simple, it breaks the signal into small time segments and Fourier transforms those segments. This means that the STFT is defined as

$$F_s^h(t,f) = \int s(\tau)h^*(\tau - t)e^{-i2\pi f\tau}d\tau, \tag{30}$$

where $s(t)$ could be a real or complex valued signal and $h(t)$ is a time window, centred at time t. The length of the window, decides the length of the time segments used for the Fourier analysis. Usually an even, positive, unit energy time window, centred around zero, is used.

The spectrogram is defined as

$$S_s^h(t,f) = \left|F_s^h(t,f)\right|^2, \tag{31}$$

which makes it a non-negative measure of the intensity of a signal at the point (t,f). There are many advantages to using the spectrogram, it is time and frequency shift invariant, easy to implement and fast to use, it is also easily related to the periodogram. The spectrogram also has little interaction, i.e. artefacts or

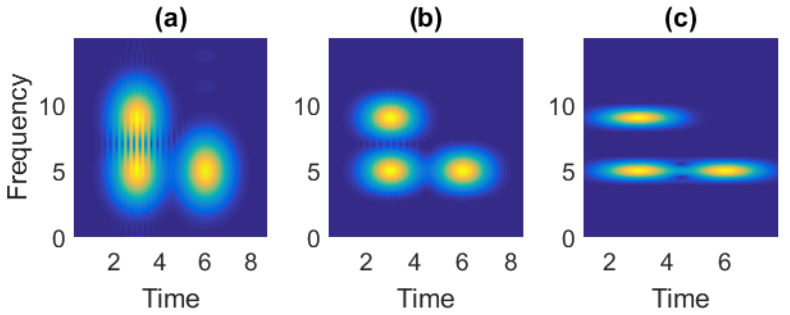

Figure 4: The spectrogram of a multi-component signal, with different time window lengths; (a) too short window length; (b) appropriate window length; (c) too long window length.

cross-terms, between signal components in multi-component signals, therefore it is mostly zero when the signal has no frequency contribution at a given time.

However the spectrogram does not fulfil the marginals, and its main problem is the trade-off between resolution in either time or frequency. The length of the time window, for the STFT, decides this trade-off, which can be seen in Figure 4. Other popular TFDs have better TF resolution compared to the spectrogram, and the spectrogram will not be able to resolve signal components that are close in time and frequency.

Another frequently used TFD is the Wigner-Ville distribution (WVD), which is defined using the analytic signal. If the analytic signal is not used, the Wigner distribution is obtained instead, which is much harder to interpret for multi-component signals. The WVD is defined as

$$W_z(t,f) = \int z\left(t+\tfrac{\tau}{2}\right) z^*\left(t-\tfrac{\tau}{2}\right) e^{-i2\pi f\tau} d\tau, \tag{32}$$

where it can be noted that $K_z(t,\tau) = z\left(t+\frac{\tau}{2}\right) z^*\left(t-\frac{\tau}{2}\right)$ is an estimate of the auto-correlation function of the signal, called the instantaneous auto-correlation function (IAF). This results in the WVD giving exactly the IF for mono-component signals, and when compared to other TFDs it achieves the best energy concentration around the signal IF law [15]. The WVD is time and frequency invariant and fulfils the marginals, thus also the total energy requirement.

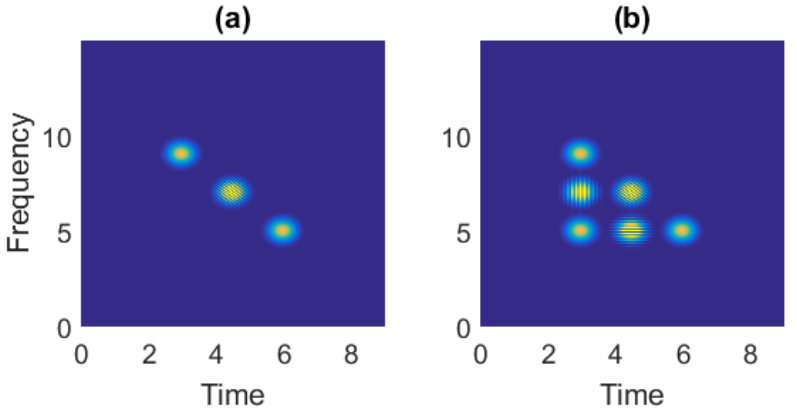

Figure 5: The WVD of two multi-component signals; (a) signal with two components and one resulting cross-term; (b) signal with three components and three resulting cross-terms.

The problem with the WVD occurs when dealing with multi-component signals or signals with noise. For such signals the WVD is not always zero when the signal has no frequency contribution at a given time, this can be seen in Figure 5. When the signal only has two components, Figure 5 (a), it is still possible to distinguish the two pulses, even though there is a cross-term midway between them which has twice the magnitude. In Figure 5 (b) the signal has three components, and the interpretation of the WVD is harder, since there now are three cross-terms. However the figure also shows that the energy concentration of the WVD is much better compared to the spectrogram, the signal in (b) is the same as in Figure 4.

3.1 Relationship between the spectrogram and the Wigner-Ville distribution

The spectrogram and the WVD are both so called quadratic TFDs, i.e. they are quadratic in the signal. This is the reason for the appearance of cross-terms, an interference midway between any two signal components in the TFD. Cross-terms can have twice the amplitude of the signal components, which can make interpretation of the TFD very hard. The presence of noise in a signal, and the additional cross-terms from noise components, can easily make it impossible to interpret the TFD.

The appearance of cross-terms is much more prevalent in the WVD compared to the spectrogram, even though both are quadratic. This is because the spectrogram can be seen as a smoothed version of the WVD, the smoothing removes most of the cross-terms, however it is also the reason for the loss in resolution or energy concentration, which can be seen when comparing Figures 4 and 5.

With the introduction of a time-lag kernel $G(t, \tau)$ it is possible to write both the WVD and the spectrogram as Fourier transforms of the convolution (in time) of the kernel and the IAF

$$G(t, \tau) \underset{t}{*} K_z(t, \tau). \tag{33}$$

Since the WVD is calculated directly from the Fourier transform of $K_z(t, \tau)$, clearly the kernel is $G(t, \tau) = \delta(t)$.

Finding the kernel for the spectrogram requires more calculations. Using the definition of the spectrogram with an analytic signal $z(t)$ it is possible to write

$$\begin{aligned} S_z^h(t,f) &= \left| \int z(t_1) h^*(t - t_1) e^{-i2\pi f t_1} dt_1 \right|^2 \\ &= \int \int z(t_1) h^*(t - t_1) z^*(t_2) h(t - t_2) e^{-i2\pi f (t_1 - t_2)} dt_1 dt_2. \end{aligned} \tag{34}$$

With the variable changes $t_1 = u + \frac{\tau}{2}$ and $t_2 = u - \frac{\tau}{2}$ the spectrogram is

$$\begin{aligned} S_z^h(t,f) &= \int \int h(t - u + \tfrac{\tau}{2}) h^*(t - u - \tfrac{\tau}{2}) z(u + \tfrac{\tau}{2}) z^*(u - \tfrac{\tau}{2}) e^{-i2\pi f \tau} du d\tau \\ &= \int \left(h(t + \tfrac{\tau}{2}) h^*(t - \tfrac{\tau}{2}) \right) \underset{t}{*} \left(z(t + \tfrac{\tau}{2}) z^*(t - \tfrac{\tau}{2}) \right) e^{-i2\pi f \tau} d\tau, \end{aligned} \tag{35}$$

the equality uses the evenness of the window function $h(t)$. This means that the time-lag kernel for the spectrogram depend on the time window used for the spectrogram, $G(t, \tau) = h(t + \frac{\tau}{2}) h^*(t - \frac{\tau}{2})$.

It is possible to design kernels with specific types of signals in mind. The kernels are usually designed to suppress cross-terms and noise, while keeping most of the resolution of the signal components. There is always a trade-off between suppressing interference and maintaining resolution. The kernels can be designed so that the TFD fulfil the marginals, like the WVD, or not, like the spectrogram. As there are many types of signals, there are many types of TFDs which smooths the WVD

via the use of a kernel. Probably the most common is the Choi-Williams distribution (CWD) [16], there also exist signal adaptive or optimal kernels [17, 18] and many others [13, 15].

Outside of the large class of quadratic distributions, with adapted kernels for desired TD characteristics, there exist other well used TF methods, such as the Gabor expansion [14], which was related to the STFT and spectrogram by Bastiaans [19], and wavelet based algorithms, developed by e.g. Haar (early 1900s), Zweig, Morlet, Grossmann and Daubechies (1970-1980). For these methods the aim is to find the analysis window achieving the best TF resolution. In a similar way, the Stockwell transform estimates the width of a Gaussian window function using a concentration criterion [20].

4 Resolution and localisation

Real, measured signals are always of finite duration and disrupted by noise. This makes TF characterisation of real, measured signals a difficult task. Even if a TFD has strong finite support, noise and the finite signal duration will be a problem for the interpretation of the density. This means that in practice, there has to be compromises between resolution and localisation of signal components, i.e. how dense the signal energy is close to the true frequencies for all time instants, and the level of suppression on undesired energy dense areas in the TFD. The theoretical bases of desired characteristics for a TFD might not be enough for easy interpretation of TFDs for measured signals. In some applied fields of research it might not be important that the TFD fulfils the marginals, but instead something else might be desired.

4.1 Defining good resolution

Studying different TFDs, for examples the TFDs in Figure 6, which shows the spectrogram (a), the WVD (b), the spectrogram with a matching window (c) and (d) the CWD, it is perhaps possible to say that one looks cleaner and thus better than the others. In Figure 6 it seems obvious that the spectrogram with matched window, i.e. the spectrogram with a time window that matches the signal components in both shape and duration, is better compared to the spectrogram without the matching window. However it is harder to determine if the CWD

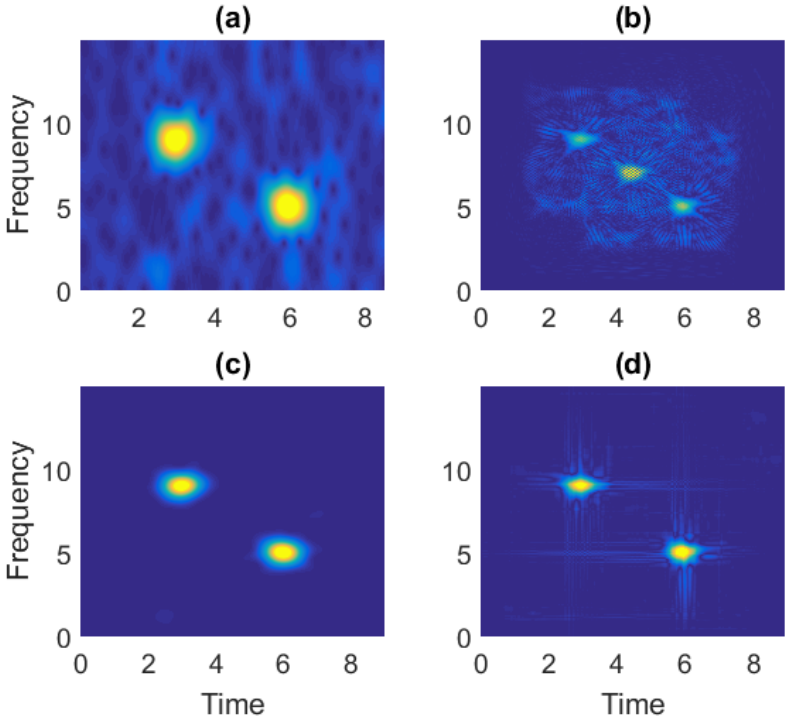

Figure 6: Four different TFDs of the same multi-component signal, disturbed by noise; (a) spectrogram; (b) Wigner-Ville distribution; (c) spectrogram with matching window; (d) Choi-Williams distribution.

is better than the spectrogram with matching window. Even though the noise and cross-terms make interpretation of the WVD hard, it has the most localised signal components, which might be the most important characteristic in certain situations.

In reality most would probably agree that either the spectrogram with matched window (c) or the CWD (d) it the best TF representation for the signal. It would however be nice if this comparison could be done (semi) automatically or without a subjective eye. There exist some methods for this.

If the analysed signal is known beforehand, the optimal TFD can be calculated and the other TFDs can be compared, by least squares, or another measure, to find the deviation from the optimal TFD. The TFD with smallest deviation, or

highest likeness, would then be the best. For known signals it is also possible to compare different TFDs by calculating the TFDs for many realisation of the known signal, the robustness to random noise, phase or other important variables could be measured. This can be done by studying if the high energy areas in the TFDs correspond to signal content. Results from these types of evaluations can be expanded to signals similar to the known one.

For a unknown signal it is possible to use the Rényi entropy [21]. It measures the energy concentration, when the concentration is high, the Rényi entropy will be small. However it is only appropriate to use on single signal components, otherwise the measure might be smallest when the resolution is very low and two (or more) components have leaked into each other. It is possible to use the Rényi entropy locally.

It is also possible to measure TFD performance for unknown signals by a quantitative performance measure presented by Boashash and Sucic [22]. The performance measure aims to resolve two close components and requires some method to find the components in each time or frequency slice of the TFD. It has been used om both simulated and measured signals [2, 23].

4.2 Reassignment

A technique to improve the localisation of single TF components and enhance the readability of the spectrogram of multi-component signals was introduced by Kodera *et al.* [24] and later reintroduced by Auger and Flandrin [25]. The method reassigns signal energy to the centre of gravity, giving higher energy concentration at the instantaneous frequencies of the signal. A similar method, the synchrosqueezing transform by Daubechies *et al.* [26], related to the empirical mode decomposition [27], reassigns all energy in frequency at a certain time point.

The reassigned spectrogram is defined by

$$ReS_s^h(t,\omega) = \int\int S_s^h(\tau,\xi)\delta\left(t-\hat{t}_s(\tau,\xi),\omega-\hat{\omega}_s(\tau,\xi)\right)d\tau d\xi, \tag{36}$$

where the two-dimensional Dirac impulse is defined as

$$\int\int f(t,\omega)\delta(t-t_0,\omega-\omega_0)dtd\omega = f(t_0,\omega_0). \tag{37}$$

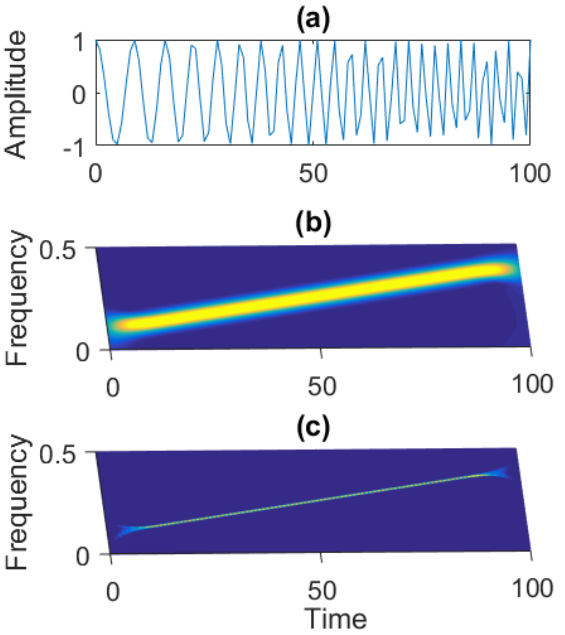

Figure 7: Illustration of reassignment for a linear chirp; (a) time representation; (b) spectrogram; (c) reassigned spectrogram.

The reassignment thus maps signal energy from a point (t_0, ω_0) to the point $(\hat{t}_x(t_0, \omega_0), \hat{\omega}_x(t_0, \omega_0))$ in the spectrogram. The new coordinates can be calculated using the phase of the STFT

$$\begin{aligned} \hat{t}_s(t, \omega) &= \frac{t}{2} - \frac{\partial}{\partial \omega}\varphi(t, \omega), \\ \hat{\omega}_s(t, \omega) &= \frac{\omega}{2} + \frac{\partial}{\partial t}\varphi(t, \omega). \end{aligned} \tag{38}$$

The reassigned spectrogram has the same smoothing as the spectrogram, given by the time window, which suppresses oscillating interference and widens the signal components. Then the reassignment compress the signal components to be more localised, by moving signal energy that is left after the smoothing. Figure 7 illustrates the reassignment of a linear chirp. The spectrogram, Figure 7 (b), of the chirp, smooths the TFD, suppressing the side lobes but also making the IF estima-

tion less accurate, thus widening the high density area in the TFD. The reassigned spectrogram, Figure 7 (c), has the same smoothing as the spectrogram, with no visible side lobes, however the energy is moved to be more localised around the IF, making the high density area more narrow.

The reassignment method, as well as the synchrosqueezing transform, works well for longer chirps and constant frequency signals. They are based on the assumption of a linear frequency modulation, essentially of infinite length. This means that the methods do not work for signals with short duration, i.e. transient signals.

The reassignment method has been adapted for transient signals by Hansson-Sandsten and Brynolfsson [3]. The adaptation allows the length of the reassignment to be scaled and not fixed to one, as for the traditional reassignment spectrogram. By allowing the signal energy to move further when reassigned, it is possible to get good resolution and localisation for very short signals. The reassignment operators for this adapted reassignment are

$$\begin{aligned} \hat{t}_s(t,\omega) &= \frac{t}{2} - c_t \frac{\partial}{\partial \omega}\varphi(t,\omega), \\ \hat{\omega}_s(t,\omega) &= \frac{\omega}{2} + c_\omega \frac{\partial}{\partial t}\varphi(t,\omega), \end{aligned} \tag{39}$$

where c_t and c_ω are the scaling factors. The reassignment coordinates has to be calculated for every set of signal and window type. Hansson-Sandsten and Brynolfsson [3] presents the coordinates when the signal and window are both Gaussian shaped. The calculations show that for a single component signal with no noise, all energy is then reassigned to the TF centre.

Figure 8 illustrates the normal and adapted reassignment for a signal with two Gaussian shaped, transient components. While the spectrogram, Figure 8 (b), resolves the two components, the smoothing of the window that suppress the side lobes also smooths the energy of the signal components. The reassigned spectrogram, Figure 8 (c), should localise the signal energy, however since it assumes linear frequency modulation, the result is not as desired and the energy is scattered. The reassigned spectrogram for transient signals, Figure 8 (d), obtains good energy localisation around the IF of the transient components.

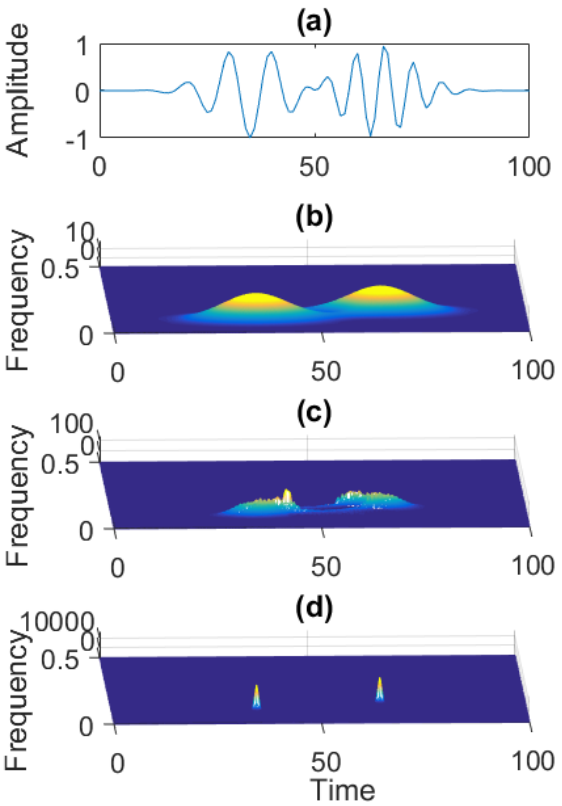

Figure 8: Illustration of reassignment for a transient signal; (a) time representation; (b) spectrogram; (c) reassigned spectrogram; (d) reassigned spectrogram for transient signals.

In order to achieve faster computations of the reassignment coordinates a different formulation can be used

$$\begin{aligned} \hat{t}_x(t,\omega) &= t + c_t \Re\left(\frac{F_x^{th}(t,\omega)}{F_x^{h}(t,\omega)}\right), \\ \hat{\omega}_x(t,\omega) &= \omega - c_\omega \Im\left(\frac{F_x^{dh/dt}(t,\omega)}{F_x^{h}(t,\omega)}\right), \end{aligned} \tag{40}$$

where $\mathfrak{R}$ and $\mathfrak{I}$ represents the real and imaginary parts and F_x^h, F_x^{th} and $F_x^{dh/dt}$ are the STFTs with different time windows. The formulation, without the scaling factors, is presented in [25].

An other adaptation of the reassigned spectrogram was presented by Auger *et al.* [28], it uses the Levenberg-Marquardt algorithm [29, 30], to allow for a weak or strong reassignment. The reassignments coordinates are derived from the second-order partial derivatives of the phase of the STFT, and it is possible to get good localisation of long or short signal by adjusting a dampening parameter. The method requires the additional calculation of one more STFT with the time window $t^2h(t)$. A recursive implementation of the Levenberg-Marquardt reassignment has also been proposed recently [1].

5 Main results of the research papers

All three papers in this thesis concern the TF analysis of non-stationary and multi-component signals. The aim of the papers involve finding high resolution TF representations of different signals. Paper A and B have more practical applications and paper C is more theoretical. This section outlines the papers.

Paper A: Optimal time–frequency distributions using a novel signal adaptive method for automatic component detection

Paper A presents a signal adaptive method for automatically detecting signal components in two-component signals. The method is based on, and outperforms, a method presented by Sucic and Boashash [31]. In the paper the new method is used with a quantitative performance measure presented by Boashash and Sucic [22] with the aim of finding the optimal, high resolution TFD for long duration, frequency modulated (FM) signals with two close components.

The proposed method is shown to accurately detect signal components for different types of two component FM signals. The evaluation is done on simulated signals with white Gaussian noise, and the method is found to be insensitive to amplitude differences of the components, frequency distance between the components and smoothing of the TFD. Since the method is robust to the level of smoothing of the TFD, a large range of TFDs can be tested, lowering the risk of erroneous conclusions to be drawn of the optimal TFD.

To illustrate the use of the method, with the performance measure, the paper shows how the optimal kernel parameters for the modified B-distribution [32] 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 and is a sensitive indicator of compromised health [7, 33]

Paper B: Objective detection and time-frequency localization of components within transient signals

Paper B presents a method that automatically detects and counts transient signal components in a signal with an unknown number of components. The paper also thoroughly investigates the reassigned spectrogram for transient signals [3], which is used in the detection method. The TFD is found to have good TF resolution even for noisy signals, i.e. it resolves heavily overlapping components. The method is unique in that it is developed for short transient signals, other component detection methods are designed for longer signals.

The aim of the automatic detection method is to give the TF centres of all components within a multi-component, transient signal. The evaluation of the method shows that the estimated TF centres have good accuracy and that the correct number of components almost always are detected for a variety of transient signals with white Gaussian noise. For the evaluation, the transient components are Gaussian shaped, which is an usual assumption for transients.

Promising results are shown when the method 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. This is of great interest since transient signals are common in several fields, e.g. ultrasonic and marine biosonar signal analysis, machine fault diagnosis and biomedical signal processing [8, 9, 34, 35, 36].

Paper C: The scaled reassigned spectrogram adapted for detection and localisation of transient signals

Paper C builds on the reassigned spectrogram for transient signals [3] and presents the reassignment coordinates for when the signal and time window of the spectrogram are combinations of the first and second Hermite functions. The aim of the paper is to show that it is possible to get perfect localisation, i.e. that all signal energy is reassigned to the TF centre of the signal, not only for signals that are Gaussian shaped.

It has previously been shown that for a transient, Gaussian (first Hermite function) shaped signal it is possible to get perfect localisation after reassignment if a Gaussian shaped window is used for the spectrogram [3]. This paper shows

that perfect localisation is also possible for a signal with the shape of the second Hermite function, if the time window used for the spectrogram is also a second Hermite function. It is also shown that perfect localisation is not possible to achieve if matching a Gaussian signal with a second Hermite window, or a second Hermite signal with a Gaussian window, then instead the signal energy is scattered in ellipses. The Hermite functions are analysed since they can be combined linearly to model transient signals [37, 38, 39, 40].

Perfect localisation is only possible for single component signals without noise, however the signal energy for multi-component signals with noise is still localised after reassignment if the signal and window shapes match. When the signal and window shapes do not match, the signal energy will appear more scattered. This can be used to detect the shape of and localise the TF centres of individual transient components in a non-stationary signal.

The results from simulated multi-component signals with white Gaussian noise show that close components can be resolved and correctly identified as being either Gaussian or second Hermite shaped. Good results are obtained even for signals heavily disturbed by noise. For a measured dolphin echolocation signal, it appears that two Gaussian-like components can be detected.

References

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

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

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

[15] 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.

[16] 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.

[17] R. Liao, C. Guo, Ke 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.

[18] 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.

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

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] N. E. Huang and Z. Wu, "A review on Hilbert-Huang transform: Method and its applications to geophysical studies," *Reviews of Geophysics*, vol. 46, no. 2, pp. n/a–n/a, 2008, RG2006.

[28] 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.

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

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

[31] 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.

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

[33] 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, P. H. Stone, and M. W. Van Der Molen, "Heart rate variability: Origins, methods, and interpretive caveats," *Psychophysiology*, vol. 34, pp. 623–648.

[34] 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.

[35] H. Liu, Weiguo 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.

[36] 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.

[37] R. Ma, Z. Huang L. Shi, and Y. Zhou, "EMP signal reconstruction using associated-Hermite orthogonal functions," *IEEE Trans. on Electromagnetic Compatibility*, vol. 64, no. 6, pp. 1383–1390, March 2016.

[38] B.N. Li, M.C. Dong, and M.I. Vai, "Modeling cardiovascular physiological signals using adaptive Hermite and wavelet basis functions," *IET Signal Processing*, vol. 4, no. 5, pp. 588–597, 2010.

[39] T. H. Linh, S. Osowski, and M. Stodolski, "On-line heart beat recognition using Hermite polynomials and neuro-fuzzy network," *IEEE Trans. on Instrumentation and Measurement*, vol. 52, no. 4, pp. 1224–1231, August 2003.

[40] A.I. Rasiah, R. Togneri, and Y. Attikiouzel, "Modeling 1-D signals using Hermite basis functions," in *IEEE Proc.-Vis. Image Signal Process.* IEEE, 1997, vol. 144, pp. 345–354.

Scientific publications

Author contributions

Co-authors are abbreviated as follows:
Maria Sandsten (MS), Josefin Starkhammar (JS).

Paper A: Optimal time–frequency distributions using a novel signal adaptive method for automatic component detection

I had the idea for the paper and developed the presented algorithm, however with invaluable guidance from MS. I did the simulations and had the main responsibility for the writing.

Paper B: Objective detection and time-frequency localization of components within transient signals

MS, JS and I came up with the idea for the paper. I developed the presented detection algorithm, with valuable support from MS. I did the testing on the simulated data, while JS did the testing on the dolphin echolocation data. JS also collected and worked with the ultra sound pulse-echo data. Everyone collaborated with the writing, MS had the main responsibility for the introduction, including finding related research, JS had the main responsibility for the examples section, and I for the rest.

Paper C: The scaled reassigned spectrogram adapted for detection and localisation of transient signals

MS and I developed the idea for the paper. I did the calculations presented and simulations, with support from MS. The data for the dolphin echolocation example was recorded by JS, who also contributed to the text in the example section. MS provided information on previous research and I had main responsibility for the writing.

Paper A

Paper A

Optimal time–frequency distributions using a novel signal adaptive method for automatic component detection

Isabella Reinhold, 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: time-frequency, multi-component signal, detection, performance measure, Heart Rate Variability

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 suppres-

sion. 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, 12, 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

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

where the time-lag kernel $G(t, \tau)$ is specific for each different quadratic TFD. The convolution of the kernel in (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

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

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

$$G_2(f) = \delta(f), \quad (4)$$

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

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

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 β 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, λ, for the auto-terms.
3. Determine between which frequencies all peaks above λ 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 λ, which are separated with at least δ. 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.

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

The selection of the parameter λ 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 λ always will relate to the amplitudes of the signal content. For simulations

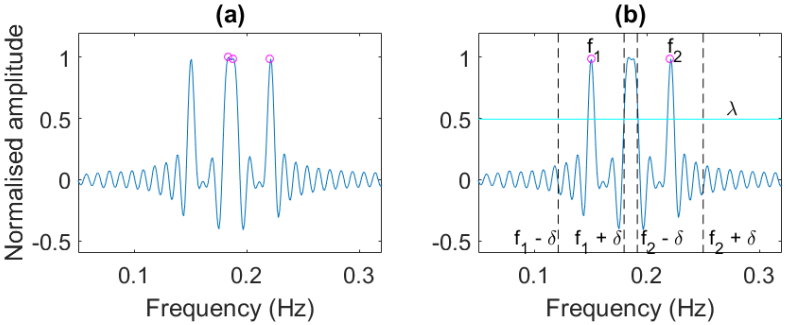

Figure 1: Time slice of the WVD of a two-component linear FM signal, with IFs 0.15 and 0.22; (a) Potential auto-terms detected with Sucic's ADAT, marked with circles; (b) Detected auto-terms with Reinhold's ADAT, the auto-terms are marked by circles and the labels f_1 and f_2 respectively. The horizontal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

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 λs.

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 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.

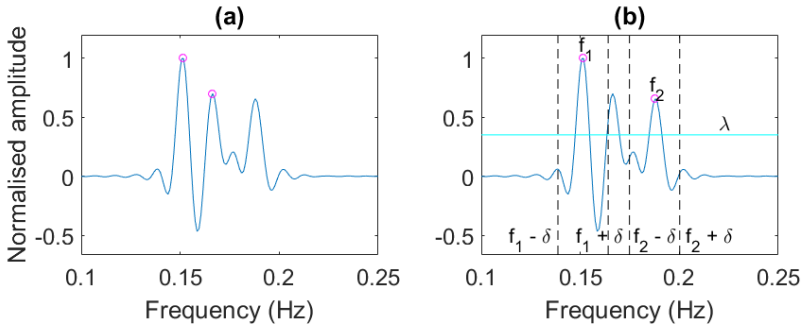

Figure 2: Time slice of a LID-TFD of a two-component non-linear FM signal, with IFs 0.15 and 0.19; (a) Auto-terms detected by Sucic's ADAT, marked by circles; (b) Auto-terms detected by Reinhold's ADAT, marked by circles and the labels f_1 and f_2 respectively. The horizontal line shows the threshold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

In this example only four peaks are above the threshold λ, 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 Δf_a. Reinhold's ADAT will then detect three peaks above λ which are separated by at least δ, 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 δ, 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 λ, 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

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

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. In this example the NIR performance is calculated in the frequency band 0 - 0.25. Figures 4(a) and (b) shows the detected

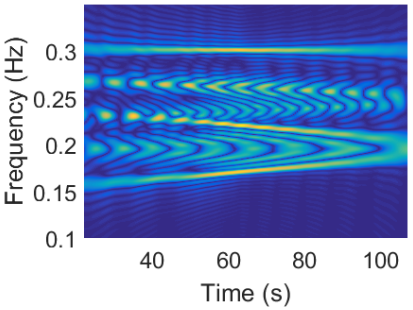

Figure 3: MBD of the signal (9), which is a three component signal with two close components and one further away.

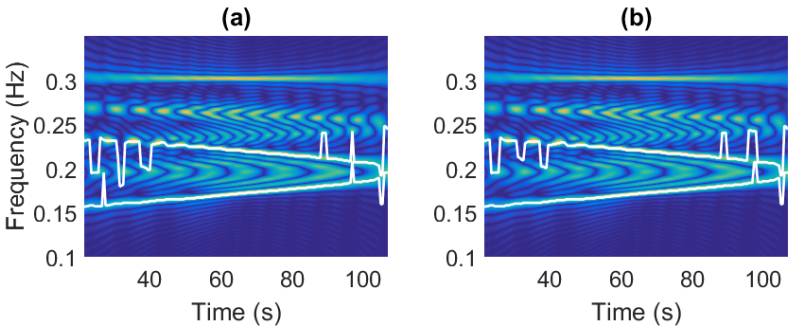

Figure 4: Detection of auto-terms in the example three-component signal (9), when only considering the frequency band 0 - 0.25 in the ADAT algorithms; (a) Sucic's ADAT; (b) Reinhold's ADAT.

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 meth-

ods 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 (10). β is the scaling parameter for the MBD kernel (6).

Setup	a_0	f_0	f_I	k	β
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, β, of the MBD kernel in (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 (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.

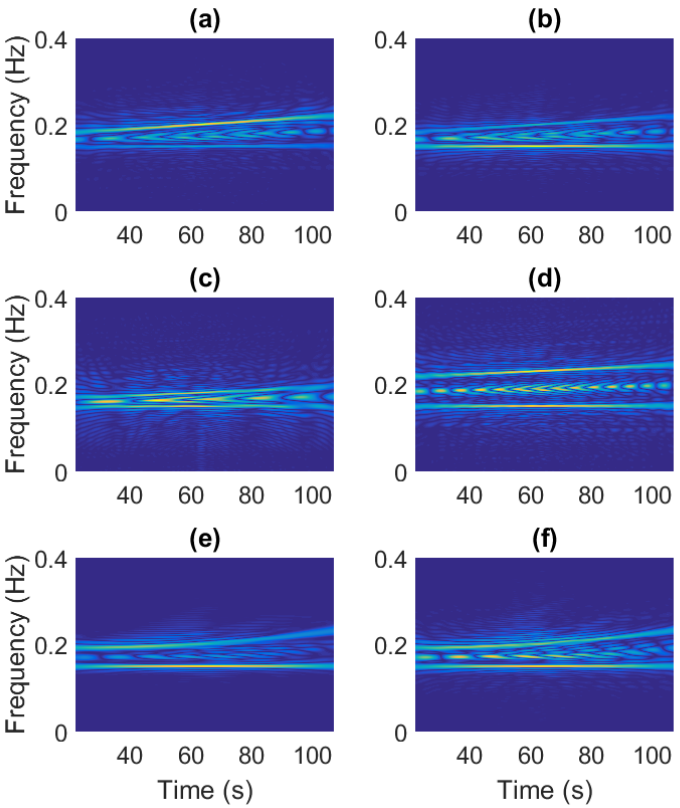

Figure 5: Some MBDs of the signal in (10) with parameters according to the setups in Table 1 evaluated when calculating the rate of detection. The figures show the part of the distribution which is assessed; (a) Setup 1 with $a_0 = 0.6$; (b) Setup 1 with $a_0 = 1.4$; (c) Setup 2 with $f_0 = 0.16$; (d) 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 TFDs. Figures 5(c) and 5(d) show the MBD when the signal components are closest together ($f_0 = 0.16$) and furthest apart

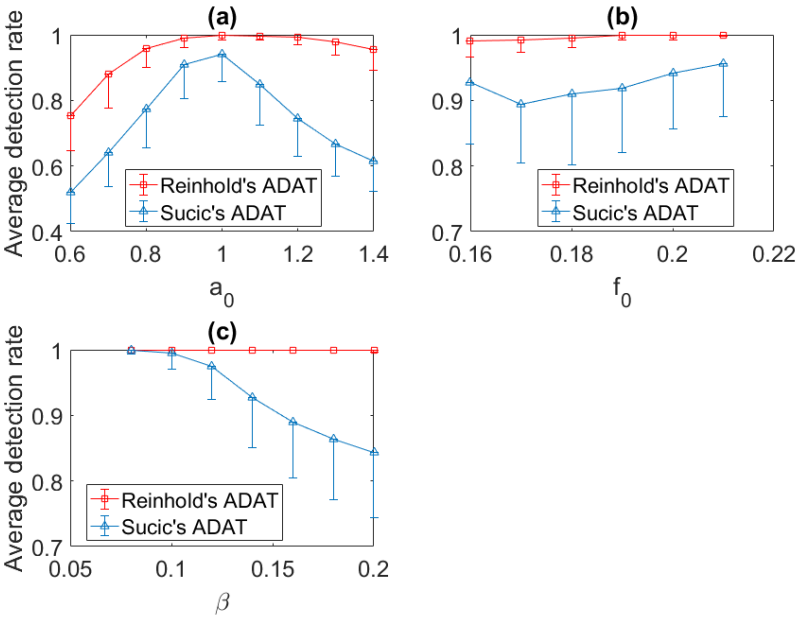

Figure 6: The average rate of detection for the ADAT algorithms according to the three setups in Table 1. The lower bound is a one sided confidence interval with 5% significance. The results are obtained from 500 simulations with different realisations of stationary Gaussian white noise with SNR 5 dB;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

($f_0 = 0.21$). 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 β 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 a MBD with the largest scaling parameter

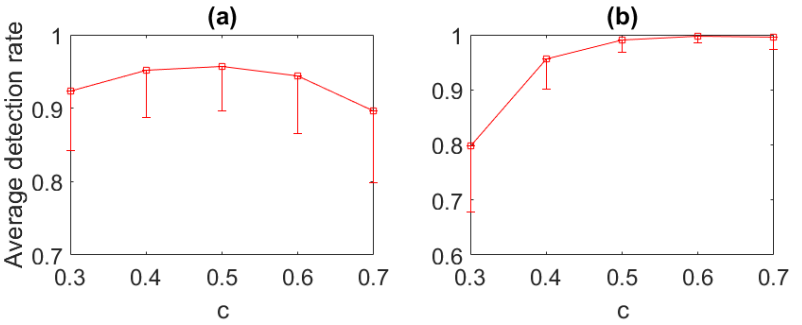

Figure 7: The average rate of detection for Reinhold's ADAT according to the setups in Table 1. The lower bound is a one sided confidence interval with 5% significance. The results are obtained from 500 simulations with different realisations of stationary Gaussian white noise with SNR 5 dB; (a) Setup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

($\beta = 0.20$), 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 β increases, however for Reinhold's the average rate of detection is 1 for all evaluated β. 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 λ. 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 (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 λ 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 (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.

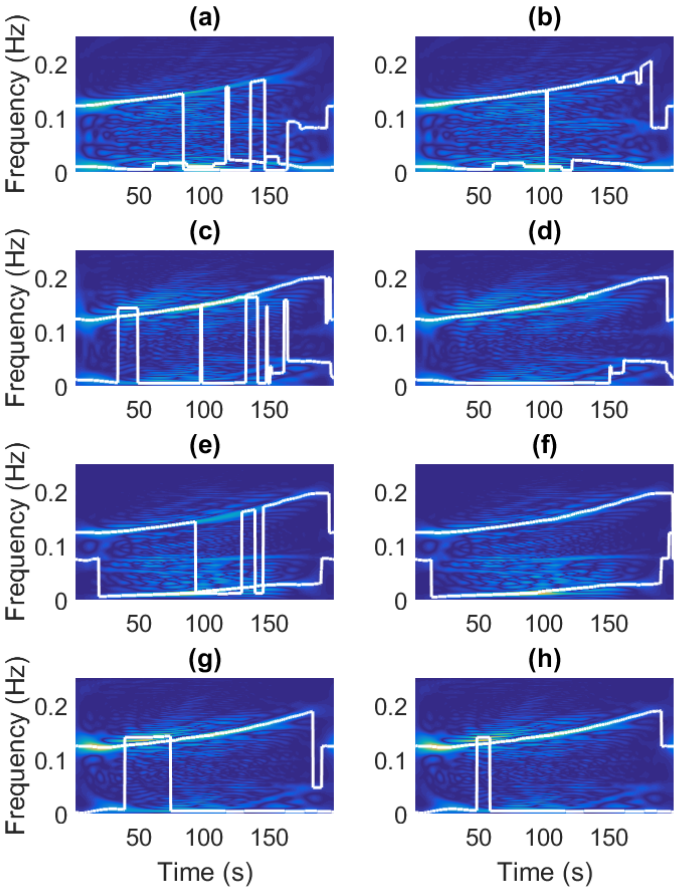

Figure 8: Detection of auto-terms in measured HRV using Sucic's and Reinhold's ADAT; (a) subject 1, Sucic's ADAT; (b) subject 1, Reinhold's ADAT; (c) subject 2, Sucic's ADAT; (d) subject 2, Reinhold's ADAT; (e) subject 3, Sucic's ADAT; (f) subject 3, Reinhold's ADAT; (g) subject 4, Sucic's ADAT; (h)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

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, δ, 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 (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 β 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 β. This procedure is repeated

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

for an interval of β with an appropriate length of the increments. The optimal kernel parameter β 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 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].

The interval performance measure is calculated for different time intervals for the four HRV signals using the measure in (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 β. 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

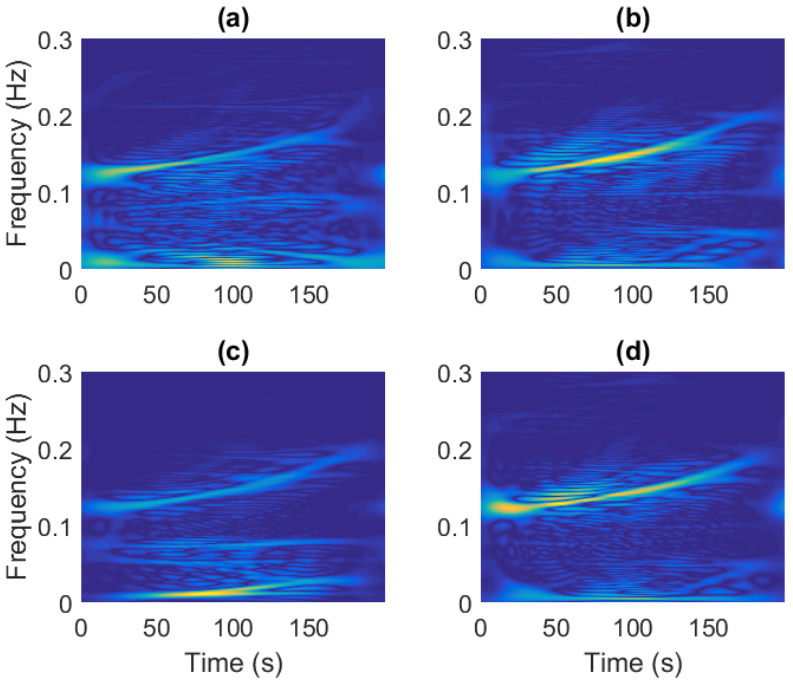

Figure 9: Optimal MBD for four examples of HRV signals; (a) subject 1; (b) subject 2; (c) subject 3; (d) subject 4.

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, λ and δ need to fulfil the following:

- Outer peaks which have a distance larger than δ to their closest auto-term are smaller than λ.
- Peaks with less distance than δ to their closest auto-term have smaller amplitudes than that auto-term.
- If the cross-term is the largest peak, the parameter δ is smaller than the actual distance between the cross-term and auto-terms.

It is therefore important that λ and δ depend on the examined TFD time slice. It is reasonable to choose $\delta = \widehat{\Delta f_a}/2 - \varepsilon$, where ε 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 λs. 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 perform-

ance 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.

[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 B

Paper B

Objective detection and time-frequency localization of components within transient signals

Isabella Reinhold[1], Maria Sandsten[1], 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 utilized time-frequency representation 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 analyzed. 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: reassignment, non-stationary signals, time-frequency analysis, pulse-echo

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, 2, 3, 4, 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 pathways inside the animals' forehead. These pathways affect the emitted signal in different directions, however to what extent the signal can be controlled by the animal and what functions it serves, are not yet fully understood [9, 10].

Well known TF techniques can successfully be applied for relatively simple transient signals of narrowband excitation, however for broadband excitations where the multiple components appear very close in the TF domain, the signals are increasingly difficult to analyze [11]. 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 com-

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

A technique to improve the localization of single TF components and enhance the readability of the TF representation of multi-component signals is introduced by Kodera *et al.* [12] and later reintroduced by Auger and Flandrin [13]. 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 Daubechies *et al.* [14], related to the empirical mode decomposition [15], reassigns all energy in frequency at a certain time point. However these methods only work well for longer chirps and constant frequency signals, and are based on the assumption of a linear frequency modulation, essentially of infinite length. Methods exist that convert the possible non-linear instantaneous frequency into a linear one and in Wang *et al.* [16] a nonlinear squeezing transform, especially designed for weak signal detection, is proposed.

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 Guo [17] and Kheirati Roonizi [18]. 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 [19]. The Gabor and Stockwell transforms are widely used and adapted in many fields of research, e.g. estimating the direction of arrival [20], automatically adapting the TF resolution of transients [21], detecting epileptic seizures [22] or double-talk in acoustic echo cancellation [23].

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 Hansson-Sandsten and Brynolfsson [24]. This reassignment technique, the scaled reassigned spectrogram, (ReSTS) finds the TF centers of individual signal components by utilizing that many transients have a Gaussian like shape in time. The ReSTS is a high TF res-

olution method and is therefore well suited for detection and localization of transient signal components, also when they are closely located in the TF domain. There exist many methods for localizing and counting signal components, e.g. [25, 26, 27, 28], however these methods are developed for longer lasting signals.

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 ReSTS 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 reassigned spectrogram for transient signals

In order to detect and localize individual transient pulses in a multicomponent 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) and (b), the spectrogram has adequate resolution and separates the components, but for the last signal, Figure 1 (c), the overlap in time is too large and as a result the spectrogram does not fully resolve the two components.

However the components of the last signal can be resolved using the reassigned spectrogram for transient signals (ReSTS) [24], which is an adaptation of the reassigned spectrogram [12, 13]. The ReSTS can in theory give perfect TF localization to transient signals and is obtained by first calculating the spectrogram of a signal $x(t)$ using a desired time window $h(t)$

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

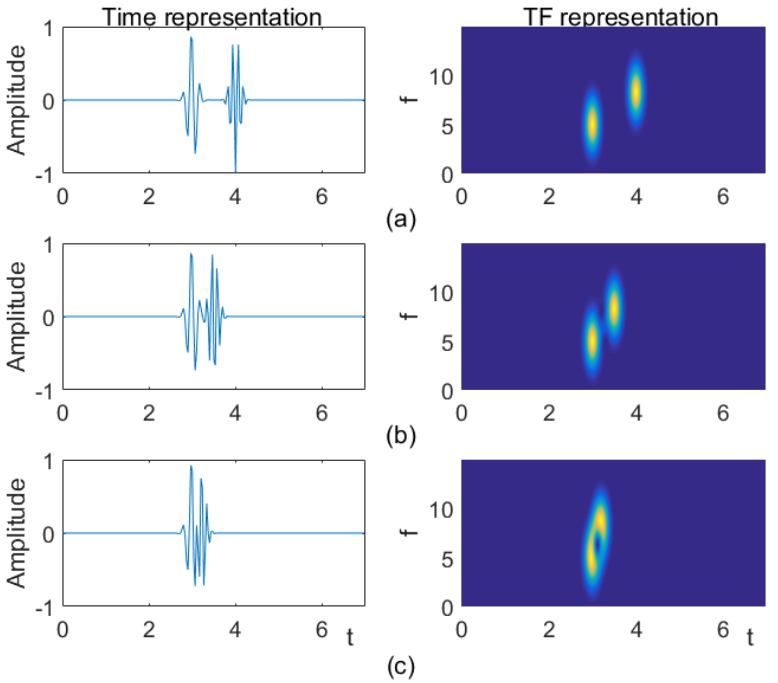

Figure 1: Illustration of a transient signal with two Gaussian pulses with different TF centers; (a) pulses clearly separated in time and frequency; (b) pulses overlapping in time but clearly separated in frequency; (c) pulses overlapping so much in time that the frequency separation is not clear.

where $*$ denotes complex conjugate, $\omega = 2\pi f$ and the integral runs from $-\infty$ to ∞. 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, $\int\int f(t,\omega)\delta(t-t_0,\omega-\omega_0)dtd\omega = f(t_0,\omega_0)$. The ReSTS is then defined as

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

and thus maps signal energy from a point (t_0,ω_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_ω makes the ReSTS adaptable to transient signals, and separates it from the normal reassigned spectrogram that have $c_t = c_\omega = 1$ [13, 24, 29].

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

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

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

Hansson-Sandsten and Brynolfsson [24] 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 time-frequency center. This result can easily be expanded to Gaussian

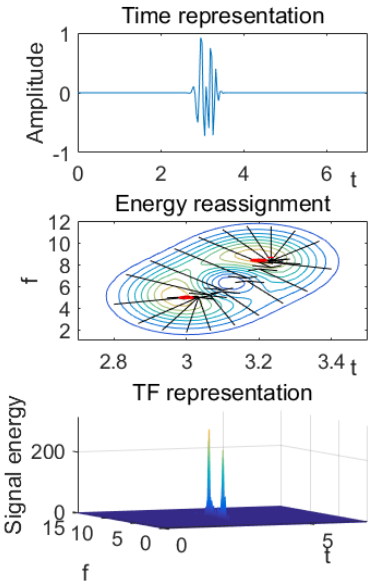

Figure 2: Illustration of the reassignment of signal energy and the resulting TF representation, adapted for transient signals, for the signal in Figure 1 (c). The signal has two transient pulses which overlap heavily in time, however when the signal energy is reassigned to the center of mass, the ReSTS shows two clear peaks at the correct TF centers.

signals with other TF centers since the spectrogram obeys time-frequency shift-invariance and due to the linearity of the Fourier transform, the reassignment coordinates are linear [29].

For multi-component signals (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 (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. To our knowledge there exist no researched methods for finding multiple peaks in a matrix.

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 more energy, and thus higher amplitude, if they are the result of signal components. 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 δ_t and δ_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 ρ assumes high SNR, where signal and noise peak amplitudes are clearly separated, a small ρ 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}, δ_t, δ_f, ρ
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*
δ_t : *Smallest time separation*
δ_f : *Smallest frequency separation*
ρ : *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 n=1:N **do**
    find coordinates of maximum in ReS, $(T(n), F(n)$
    save maximum amplitude in vector, $A(n)$
    define the rectangle area, B_{TF}, with center $(T(n), F(n))$ 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. Δ_n is the difference in amplitude of all known noise peaks, normalized with the constant ρ. Δ_s is the difference in amplitude of all the possible signal peaks. Δ 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 Δ 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 Δ_s/Δ is closer to 1 compared to Δ/Δ_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 reassigned spectrogram for transient signals

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 (5) that is a linear combination of Gaussian pulses (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 sampling frequency 100 MHz and the scaling of the Gaussian pulses (4), $\sigma = 0.5$ μs, which gives an approximate signal length of 1.2 μs, full width at half maximum, or approximately 5 periods. The evaluated time distances range from to 0.5 μs to 2.0 μs, 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 $(t_1, t_2) = (4.00, 5.50)$ μs 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 Choi-Williams distribution (CWD), [30], Figure 3 (c).

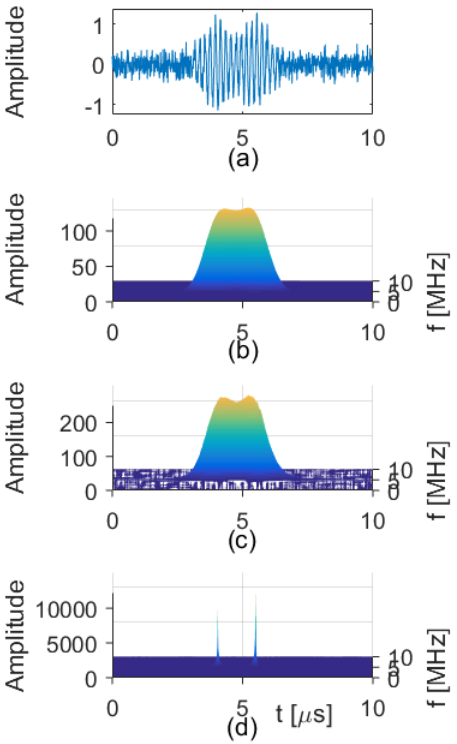

Figure 3: Illustrations for a realization of the simulated signal (5) with $(t_1, t_2) = (4.00, 5.50)$ μs, white Gaussian noise and SNR = 5 dB; (a) the time signal; (b) the spectrogram; (c) the Choi-Williams distribution; (d) the ReSTS.

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, $\sigma = 0.50$ μs, and the CWD, are evaluated by simulating 200 realizations of each test signal with different time distances between components. Each simulation has different noise disturbance realizations and phase shifts for both signal components. The TF distributions are calculated for each realization and the maximum in an

Table 1: Mean and standard deviation of the estimated time centers from 200 simulations of the signal (5) with white Gaussian noise, SNR = 5 dB and random phase for each pulse.

		Spectrogram				CWD			
		Mean [μs]		SD [μs]		Mean [μs]		SD [μs]	
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) μs	**4.01**	**5.99**	**0.03**	**0.03**	**4.08**	**5.93**	**0.05**	**0.06**
2	(4.00, 5.50) μs	4.12	5.37	0.19	0.21	4.13	5.37	0.11	0.10
3	(4.00, 5.00) μs	4.37	4.63	0.24	0.24	4.23	4.78	0.22	0.21
4	(4.00, 4.50) μs	4.23	4.26	0.07	0.07	4.14	4.35	0.16	0.15

		ReSTS			
		Mean [μs]		SD [μs]	
Test	True (t_1, t_2)	$\hat{t_1}$	$\hat{t_2}$	$\hat{t_1}$	$\hat{t_2}$
1	(4.00, 6.00) μs	4.00	6.00	0.02	0.02
2	(4.00, 5.50) μs	4.01	5.50	0.04	0.04
3	(4.00, 5.00) μs	**4.02**	**4.97**	**0.09**	**0.10**
4	(4.00, 4.50) μs	4.14	4.35	0.12	0.12

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 - 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.

It can be seen that all three methods separate the signal components when time distance between the components is 2.00 μs, even though the CWD gives some deviations in the time centers, see Table 1. When the time distance is 1.50 μs 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 for this type of signal with low SNR and random phase shifts, the spectrogram does not succeed to resolve the Gaussian pulses that are $2\sigma = 1.0$ μs apart, which is the theoretical minimum time separation needed between two Gaussian pulses for the peaks to be resolved [31]. The estimated time centers for the ReSTS are

Table 2: Mean and coefficient of variation of the peak amplitudes from 200 simulations of the signal (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	**89**	**89**	**0.04**	**0.04**	**192**	**192**	**0.11**	**0.11**
2	95	95	0.18	0.18	201	201	0.19	0.19
3	131	131	0.45	0.45	227	227	0.40	0.40
4	169	169	0.61	0.61	268	269	0.61	0.61

	ReSTS			
	Mean		CV	
Test	A_1	A_2	A_1	A_2
1	3056	3177	0.33	0.34
2	1079	1077	0.36	0.32
3	**384**	**366**	**0.48**	**0.44**
4	1250	1310	1.07	1.08

however still very good when $(t_1, t_2) = (4.00, 5.00)$ μs, it is not until the time distance is 0.5 μs 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 δ_t and δ_f

The proposed automatic component detection algorithm requires the user to define the parameters δ_t and δ_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 δ_t and δ_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 δ_t and δ_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 δ_t is the uncertainty in time, i.e. the length of the pulse, and δ_f is the uncertainty in frequency. Two Gaussian pulses of equal amplitude and time length, are separable if

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

where σ is the scaling parameter of the Gaussian [31]. This means that these distances correspond to the best resolution the spectrogram can possibly achieve.

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

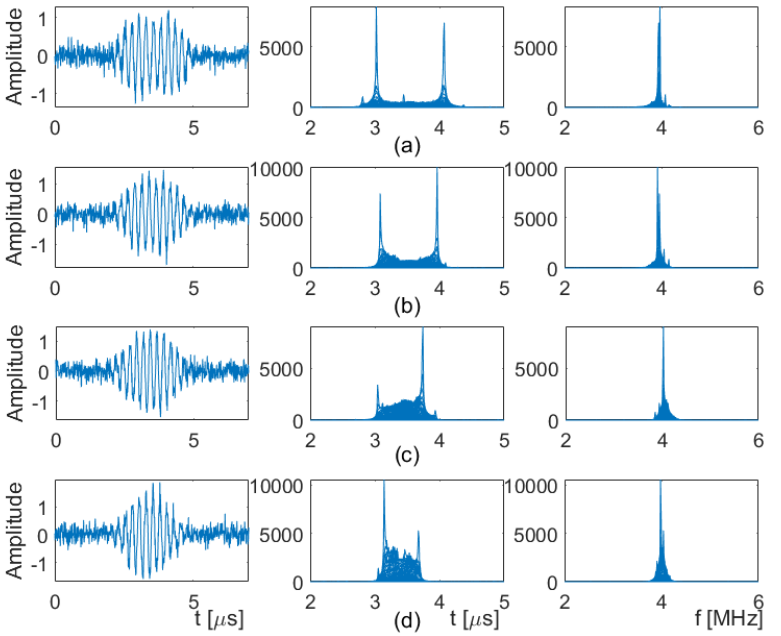

Figure 4: Realizations of the test signals (5) with two Gaussian pulses and white Gaussian noise, SNR = 5 dB, where $f_1 = f_2 = 4.00$ MHz and random phases of the components. The time signal is shown to the left and the ReSTS is viewed so that the amplitude and time axes (middle) or amplitude and frequency axes (right) can be seen; (a) signal with $(t_1, t_2) = (3.00, 4.10)$ μs; (b) signal with $(t_1, t_2) = (3.00, 4.00)$ μs; (c) signal with $(t_1, t_2) = (3.00, 3.90)$ μs; (d) signal with $(t_1, t_2) = (3.00, 3.80)$ μs.

The minimum required time separation is evaluated by keeping the frequency of the two signal components constant and decreasing the time distance between the TF centers of the components. Figure 4 shows realizations of the signals with decreasing time separation of the TF centers, it shows the time signal (left) and the ReSTS from two viewpoints so that only the time (middle) or frequency (right) axis is seen. 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: Mean and standard deviation of the estimated time centers and the mean and coefficient variation of the peak amplitudes from 200 simulations of the signal (5) with $f_1 = f_2 = 4.00$ MHz, white Gaussian noise, SNR = 5 dB and random phase for each pulse.

	Mean [μs]		SD [μs]		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) μs	3.03	4.08	0.08	0.07	6764	6667	0.41	0.38
(3.00, 4.00) μs	**3.04**	**3.97**	**0.09**	**0.09**	**5758**	**5695**	**0.47**	**0.40**
(3.00, 3.90) μs	3.06	3.84	0.11	0.11	5314	5261	0.54	0.52
(3.00, 3.80) μs	3.06	3.73	0.12	0.14	5169	5350	0.59	0.67

Table 3 shows the results of 200 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 $t_2 = 3.90$ μs the mean of $\hat{t}_2$ is closer to 3.80 μs than the true value. Also when the time distance is less than 1.0 μs 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 4 (c) and (d). It can therefore be safer to set δ_t corresponding to 1.0 μs 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 μs corresponds to 2σ, 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, $\sigma = 0.50$ μs and the sampling frequency is 100 MHz. For the calculated ReSTS matrices, the distance between two frequency values is 0.012 MHz.

Figure 5 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 two viewpoints

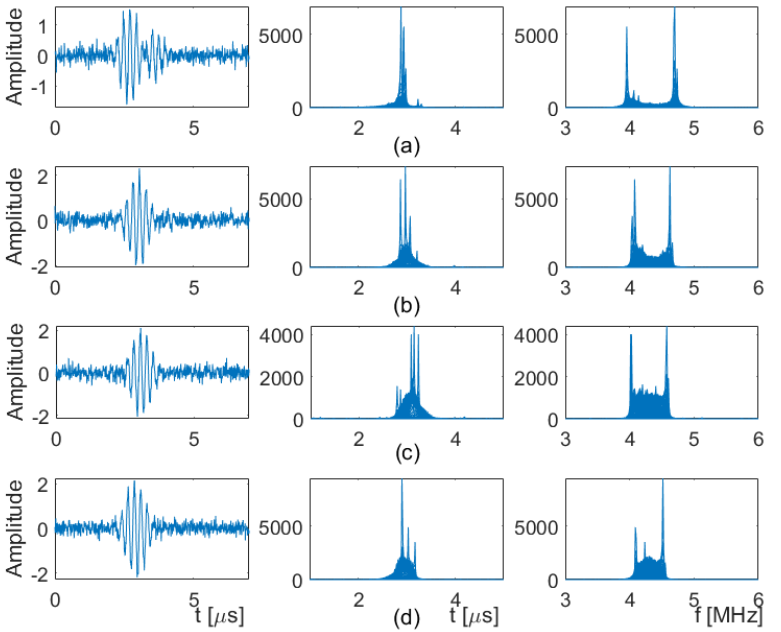

Figure 5: Realizations of the test signals (5) with two Gaussian pulses and white Gaussian noise, SNR = 5 dB, where $t_1 = t_2 = 3.00$ μs and random phase of the components. The time signal is shown to the left and the ReSTS is viewed so that the amplitude and time axes (middle) or amplitude and frequency axes (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.

so that only the time (middle) or frequency (right) axis is seen. It can be noted that the energy seem more scattered for smaller frequency distances.

Table 4 shows the estimated frequency centers and peak amplitudes from 200 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 $t_1 = t_2 = 3.00$ μs. 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

Table 4: Mean and standard deviation of the estimated frequency centers and the mean and coefficient variation of the peak amplitudes from 200 simulations of the signal (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	6895	6704	0.39	0.38
(4.00, 4.65) MHz	**4.02**	**4.64**	**0.06**	**0.06**	**5831**	**6126**	**0.41**	**0.43**
(4.00, 4.60) MHz	**4.03**	**4.58**	**0.06**	**0.07**	**5227**	**5248**	**0.50**	**0.47**
(4.00, 4.55) MHz	4.04	4.51	0.07	0.07	5294	5525	0.56	0.51

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 δ_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 δ_t and δ_f according to (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.

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 de-

tection algorithm can now be evaluated. This is done using test signals (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 500 simulations are done for each signal and SNR, with different noise realizations and random phase shifts for the signal components. The sampling frequency is 100 MHz and the scaling of the Gaussian pulses $\sigma = 0.5$ μs. 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 500 simulations are averaged to get the detection rate. Thus 1 means that detection was correct for all 500 simulations and 0 that detection was incorrect for all simulations.

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 cen-

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

k	t_k [μs]	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

Table 6: Detection rate of the automatic component detection algorithm for multi-component, transient signals (5), disturbed by white noise and with parameters according to Table 5. The detection rates are obtained from 500 simulations with different noise realizations and where the signal components have random phase shifts. If all detections for all 500 simulations are correct, the rate is 1.

		Detection rate for	
SNR [dB]	ρ	2 components	5 components
15	2/3	1	1
5	2/3	1	0.95
5	1/3	1	0.98

ter if the time and frequency difference is less than $\delta_t = 2\sigma = 1$ μs 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 ρ. When the SNR is higher, the algorithm correctly identifies all components even for the five components signals.

Since there, to our knowledge, exist no other researched methods, the performance of our algorithm can not be compared to other known methods. 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 [32]. 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 gives

a detection rate around 0.90, however the detection rate decreases when SNR increases, which certainly is undesirable. The Gaussian threshold, has detection rates around 0.25 - 0.30, which also decreases when SNR increases. The level out approach has a detection rate around 0.60 for the signals with two components and approximately 0.50 for the signals 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 [24]. 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 setup 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 and was measured at 100 MHz sampling rate with a Textronix TDS 2002C oscilloscope. The reflective object was a plexi glass 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 6.

Measurements were taken in each transition between two steps of the plexi glass phantom. This resulted in a total of four measured time signals con-

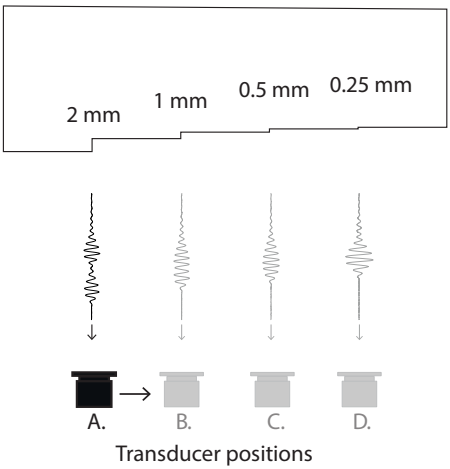

Figure 6: Setup for the pulse-echo measurement of two surfaces, separated by 2, 1, 0.5 or 0.25 mm (positions A-D).

taining 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 7 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 plexi glass phantom, which are shown in the figure and should be compared with the step sizes in Figure 6.

The simple piezoceramic transducer with an approximate pulse length of 1 μs, full width at half maximum, can resolve surfaces separated by at least 1.5 mm. Thus only for case A, exemplified in Figure 7 (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 7 (a) and (b), the algorithm correctly identifies the TF centers of the two pulses. For case C shown in Figure 7 (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 7 (d), only one component is detected, which is expected.

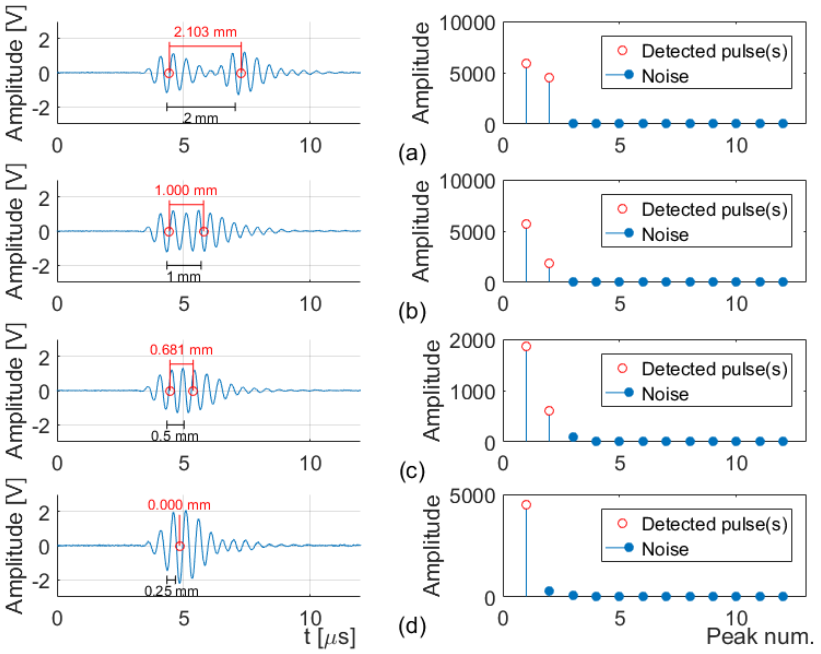

Figure 7: Time signals and peak detections from ReSTS, of pulse-echo measurements of two surfaces that are separated by 2, 1, 0.5 or 0.25 mm (a) - (d). The red circles indicate the detected time centers, obtained by the automatic detection method from the ReSTS. The time separation between the pulse centers correspond to an estimation of the step sizes of the plexi glass phantom, Figure 6. The estimates are shown in red (above), the true step sizes are shown in black (under).

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

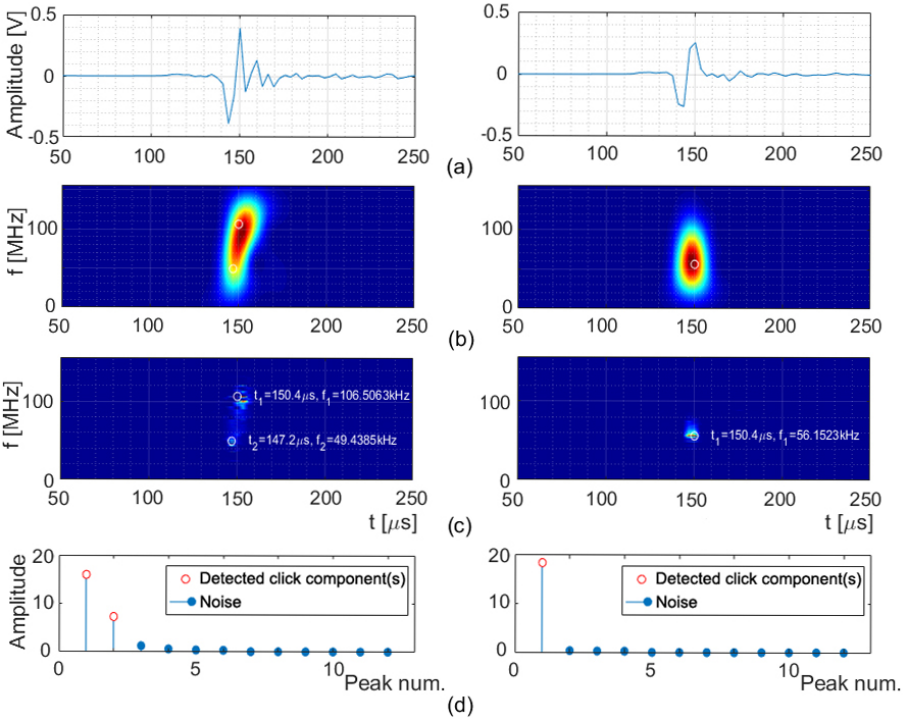

Figure 8: Two dolphin echolocation signals, to the left a bimodal, to the right a unimodal; (a) time signal; (b) spectrogram; (c) ReSTS representation and detected signal TF centers; (d) pulse detections.

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 Starkhammar *et al.* [10].

Figure 8 left hand side, shows recordings from a bimodal transient signal, expected to contain more than one component. Figure 8 right hand side, shows recordings from a unimodal transient signal, expected to contain only one component. The algorithm outcome in Figure 8 (left) shows that

it detects two signal components separated in both time and frequency, while in Figure 8 (right) only one component is detected.

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

From a biosonar perspective, it is interesting to compare the signal component time and frequency centers in Figure 8. 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, 33, 34].

7 Conclusions

In this paper an automatic component detection method for short, multicomponent 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.

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] 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.

[12] 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.

[13] 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.

[14] 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.

[15] N. E. Huang and Z. Wu, "A review on Hilbert-Huang transform: Method and its applications to geophysical studies," *Reviews of Geophysics*, vol. 46, no. 2, pp. n/a–n/a, 2008, RG2006.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

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

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

[31] 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.

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

[33] 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.

[34] 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.

Paper C

Paper C

The scaled reassigned spectrogram adapted for detection and localisation of transient signals

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

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

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

Abstract

The reassigned spectrogram can be used to improve the readability of a time-frequency representation of a non-stationary and multi-component signal. However for transient signals the reassignment needs to be adapted in order to achieve good localisation of the signal components. One approach is to scale the reassignment. This paper shows that by adapting the shape of the time window used with the spectrogram and by scaling the reassignment, perfect localisation can be achieved for a transient signal component. It is also shown that without matching the shape of the window, perfect localisation is not achieved. This is used to both identify the time-frequency centres of components in a multi-component signal, and to detect the shapes of the signal components. The scaled reassigned spectrogram with the matching shape window is shown to be able to resolve close components and works well for multi-components signals with noise. An echolocation signal from a beluga whale (*Delphinapterus leucas*) provides an example of how the method performs on a measured signal.

Keywords: Hermite functions, non-stationary signals, time-frequency analysis, reassignment, signal resolution

1 Introduction

For non-stationary signals the reassigned spectrogram (Re-Spect) can improve the readability of the time-frequency representation [1, 2]. The concentration of a component is increased by reassigning mass to the centre of gravity, squeezing the signal terms to be more localised, while cross-terms are reduced by a smoothing of the specific distribution. Recently, the theoretical expressions for the reassigned Gabor spectrograms of Hermite functions have been derived [3, 4]. Although the Re-Spect gives perfect localisation of linear chirps, this is not achieved for transient signals which are common in for example marine biosonar research. Transient signals can effectively be modelled by a linear combination of Hermite basis functions [5, 6, 7, 8]. Perfect localisation of a Gaussian function (first Hermite function) can be achieved by the adaptable reassignment methods, the Levenberg-Marquardt reassignment [9] and the scaled reassigned spectrogram (ScRe-Spect) [10].

This paper builds on the ScRe-Spect to show that perfect localisation in time and frequency can be achieved with reassignment for higher order Hermite functions. Perfect localisation is possible if the shape of the time window used with the spectrogram is matched with the shape of the signal component. It also shows that when the matching shape window is not used, perfect localisation is not possible with the ScRe-Spect. This can be used to detect the shape of and localise the time-frequency centres of individual transient components in a non-stationary signal.

To illustrate the use of the ScRe-Spect with matching shape window this paper includes an example of an echolocation signal from a beluga whale (*Delphinapterus leucas*). In this field, there is a need for signal processing methods that allows for analysis of the time dependence of each frequency component within each echolocation signal [11, 12, 13].

In this paper, section 2 calculates the reassignment coordinates for first and second Hermite signal components with first and second Hermite time windows. The results are also extended to multi-component signals. The performance of the ScRe-Spect is evaluated in section 3, by simulating transient multi-component signals with noise. Section 4 shows

the performance of the method on a measured echolocation signal from a dolphin. Section 5 concludes the paper.

2 The scaled reassigned spectrogram

The spectrogram of the signal $x(t)$ using the window $h(t)$ is found from the short-time Fourier transform (STFT)

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

The Re-Spect, with reassignment to $\hat{t}_x$ and $\hat{\omega}_x$, is defined as

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

where $\delta(t,\omega)$ is the two-dimensional Dirac impulse defined as

$$\int\int f(t,\omega)\delta(t-t_0,\omega-\omega_0)dtd\omega = f(t_0,\omega_0). \tag{3}$$

As shown in [10], the scaling factors c_t and c_ω can be introduced and the reassignment coordinates can be computed as

$$\begin{aligned}\hat{t}_x(t,\omega) &= t + c_t\Re\left(\frac{F_x^{th}(t,\omega)}{F_x^h(t,\omega)}\right),\\ \hat{\omega}_x(t,\omega) &= \omega - c_\omega\Im\left(\frac{F_x^{dh/dt}(t,\omega)}{F_x^h(t,\omega)}\right),\end{aligned} \tag{4}$$

where $\Re$ and $\Im$ are the real and imaginary parts respectively and F_x^h, F_x^{th} and $F_x^{dh/dt}$ are STFTs with different time windows. If $c_t = c_\omega = 1$ the Re-Spect is obtained [4, 14]. Since the reassignment coordinates are calculated from STFTs using the same signal values needed for the spectrogram and only differs in the choice of time window, the computational complexity of the ScRe-Spect is not drastically increased compared the spectrogram.

2.1 The scaled reassigned spectrogram of multi-component transient signals

A multi-component transient signal can be modelled as a sum of Hermite functions

$$x(t) = \sum_{k=1}^{n} a_k x_k(t - t_k) e^{i\omega_k t}, \tag{5}$$

where $x_k(t)$ are Hermite basis functions, t_k and ω_k are the time and frequency centres and a_k the amplitudes. Due to the linearity of the Fourier transform, the reassignment vector is also linear [14]. For the calculations it is assumed that the reassignment can be calculated for each component individually. This is true for a signal with somewhat separated components, in time or frequency. The spectrogram also obeys time-frequency shift-invariance, meaning that further analysis can be restricted to signals of the form $x_k(t) = g_k(t)$ instead of $x_k(t) = g_k(t - t_k)e^{-i\omega_k t}$.

This paper analyses the unit energy Gaussian function

$$x_1(t) = \sigma^{-1/2}\pi^{-1/4} e^{-\frac{t^2}{2\sigma^2}}, \tag{6}$$

and the unit energy second Hermite function

$$x_2(t) = 2^{1/2}\sigma^{-3/2}\pi^{-1/4} t e^{-\frac{t^2}{2\sigma^2}}, \tag{7}$$

as the amplitude of the signal has no effect on the reassignment coordinates, compare (4). The ScRe-Spect with perfect localisation of $x_1(t)$ using a Gaussian window is shown in [10]. The next section will show that the ScRe-Spect of a signal $x(t) = x_1(t) + x_2(t)$ with a Gaussian window will only give perfect localisation to $x_1(t)$ while the energy of $x_2(t)$ remains scattered, and when a second Hermite window is used, perfect localisation will be obtained for $x_2(t)$ but not for $x_1(t)$.

2.2 Scaled reassignment with a Gaussian window

The reassignment coordinates for $x_1(t)$ using a window $h_1(t) = x_1(t)$, where the window and signal have the same time length, are

$$\begin{aligned} \hat{t}_{x_1}^{h_1}(t,\omega) &= t - c_t\frac{t}{2}, \\ \hat{\omega}_{x_1}^{h_1}(t,\omega) &= \omega - c_\omega\frac{\omega}{2}, \end{aligned} \tag{8}$$

and perfect localisation is achieved when $c_t = c_\omega = 2$ [10]. To find the reassignment coordinates for $x_2(t)$ with $h_1(t)$, we need to calculate the necessary STFTs, also assuming that the time length of the signal and window are the same

$$\begin{aligned} F_{x_2}^{h_1}(t,\omega) &= \frac{\sqrt{2}}{\sigma^2\sqrt{\pi}}\int s e^{-\left(s^2+(s-t)^2\right)/(2\sigma^2)}e^{-i\omega s}ds \\ &= \frac{\sqrt{2}}{\sigma^2\sqrt{\pi}}e^{-t^2/(2\sigma^2)}\int s e^{-s^2/\sigma^2+(t/\sigma^2-i\omega)s}ds \\ &= \frac{t-i\sigma^2\omega}{\sqrt{2}\sigma}e^{-\left(t^2/\sigma^2+\sigma^2\omega^2+i2t\omega\right)/4}, \end{aligned} \tag{9}$$

$$\begin{aligned} F_{x_2}^{th_1}(t,\omega) &= \frac{\sqrt{2}}{\sigma^2\sqrt{\pi}}\int s(s-t) e^{-\left(s^2+(s-t)^2\right)/(2\sigma^2)}e^{-i\omega s}ds \\ &= \frac{2\sigma^2 - t^2 - \sigma^4\omega^2}{2\sqrt{2}\sigma}e^{-\left(t^2/\sigma^2+\sigma^2\omega^2+i2t\omega\right)/4}. \end{aligned} \tag{10}$$

The derivative of $h_1(t)$ is $dh_1(t)/dt = -th_1(t)/\sigma^2$, thus

$$F_{x_2}^{dh_1/dt}(t,\omega) = -\frac{1}{\sigma^2}F_{x_2}^{th_1}(t,\omega). \tag{11}$$

This gives the reassignment coordinates

$$\begin{aligned} \hat{t}_{x_2}^{h_1}(t,\omega) &= t - c_t\left(\frac{t}{2} - \frac{\sigma^2 t}{t^2+\sigma^4\omega^2}\right), \\ \hat{\omega}_{x_2}^{h_1}(t,\omega) &= \omega - c_\omega\left(\frac{\omega}{2} - \frac{\sigma^2\omega}{t^2+\sigma^4\omega}\right). \end{aligned} \tag{12}$$

It can be seen that there exist no c_t or c_ω so that $\left(\hat{t}_{x_2}^{h_1}(t,\omega), \hat{\omega}_{x_2}^{h_1}(t,\omega)\right) = (0,0)$, i.e. the centre of the component, $\forall\ t, \omega$. Instead of perfect localisation of the component $x_2(t)$, the reassigned energy will be located on ellipses, which is also found for the reassigned Gabor spectrogram in [4].

2.3 Scaled reassignment with a second Hermite window

Using the second Hermite window, $h_2(t) = x_2(t)$, with the spectrogram, makes it possible to get perfect localisation to $x_2(t)$ with scaled reassignment. For the calculations of the reassignment coordinates it is assumed that the time length of the window is the same as for the signal,

$$\begin{aligned} F_{x_2}^{h_2}(t,\omega) &= \frac{2}{\sigma^3\sqrt{\pi}} \int s(s-t) e^{-\left(s^2+(s-t)^2\right)/(2\sigma^2)} e^{-i\omega s} ds \\ &= \frac{2\sigma^2 - t^2 - \sigma^4\omega^2}{2\sigma^2} e^{-\left(t^2/\sigma^2+\sigma^2\omega^2+i2t\omega\right)/4}, \end{aligned} \tag{13}$$

$$\begin{aligned} F_{x_2}^{th_2}(t,\omega) &= \frac{2}{\sigma^3\sqrt{\pi}} \int s(s-t)^2 e^{-\left(s^2+(s-t)^2\right)/(2\sigma^2)} e^{-i\omega s} ds \\ &= \frac{w(t,\omega) - 2\sigma^2 t - i6\sigma^4\omega}{4\sigma^2} e^{-\left(t^2/\sigma^2+\sigma^2\omega^2+i2t\omega\right)/4}, \end{aligned} \tag{14}$$

where $w(t,\omega) = t^3 + i\sigma^2 t^2\omega + \sigma^4 t\omega^2 + i\sigma^6\omega^3$. Since $dh_2/dt = h_2(t)/t - th_2(t)/\sigma^2$, we get

$$\begin{aligned} F_{x_2}^{dh_2/dt}(t,\omega) &= \frac{2}{\sigma^3\sqrt{\pi}} \int s e^{-\left(s^2+(s-t)^2\right)/(2\sigma^2)} e^{-i\omega s} ds \\ &- \frac{1}{\sigma^2} F_{x_2}^{th_2}(t,\omega) \\ &= \frac{6\sigma^2 t + i2\sigma^4\omega - w(t,\omega)}{4\sigma^4} e^{-\left(t^2/\sigma^2+\sigma^2\omega^2+i2t\omega\right)/4}. \end{aligned} \tag{15}$$

This gives the following reassignment coordinates

$$\begin{aligned} \hat{t}_{x_2}^{h_2}(t,\omega) &= t - c_t\frac{t}{2}, \\ \hat{\omega}_{x_2}^{h_2}(t,\omega) &= \omega - c_\omega\frac{\omega}{2}. \end{aligned} \tag{16}$$

It can be seen that by choosing $c_t = c_\omega = 2$ indeed $\left(\hat{t}_{x_2}^{h_1}(t,\omega), \hat{\omega}_{x_2}^{h_1}(t,\omega)\right) = (0,0)\,,\ \forall\, t, \omega$ and perfect localisation in time and frequency is achieved.

Perfect time and frequency localisation is not possible when using the second Hermite window with the Gaussian component $x_1(t)$. This is shown by calculating the reassignment coordinates,

$$F_{x_1}^{h_2}(t,\omega) = -\frac{t + i\sigma^2\omega}{\sqrt{2}\sigma} e^{-\left(t^2/\sigma^2 + \sigma^2\omega^2 + i2t\omega\right)/4}, \tag{17}$$

$$F_{x_1}^{th_2}(t,\omega) = \frac{2\sigma^2 + \left(t + i\sigma^2\omega\right)^2}{2\sqrt{2}\sigma} e^{-\left(t^2/\sigma^2 + \sigma^2\omega^2 + i2t\omega\right)/4}. \tag{18}$$

We remind us that $dh_2/dt = h_2(t)/t - th_2(t)/\sigma^2$, which gives

$$F_{x_1}^{dh_2/dt}(t,\omega) = \frac{2\sigma^2 - \left(t + i\sigma^2\omega\right)^2}{2\sqrt{2}\sigma^3} e^{-\left(t^2/\sigma^2 + \sigma^2\omega^2 + i2t\omega\right)/4}. \tag{19}$$

This gives the reassignment coordinates

$$\begin{aligned} \hat{t}_{x_1}^{h_2}(t,\omega) &= t - c_t\left(\frac{t}{2} + \frac{\sigma^2 t}{t^2 + \sigma^4\omega^2}\right), \\ \hat{\omega}_{x_1}^{h_2}(t,\omega) &= \omega - c_\omega\left(\frac{\omega}{2} + \frac{\sigma^2\omega}{t^2 + \sigma^4\omega^2}\right). \end{aligned} \tag{20}$$

It can be seen that perfect localisation is not possible, even if small values of both $\hat{t}_{x_1}^{h_2}(t,\omega)$ and $\hat{\omega}_{x_1}^{h_2}(t,\omega)$ are possible if $c_t = c_\omega = 1$ and t and ω are assumed not to be large, the signal energy will be located on ellipses.

2.4 Detection and localisation of a multi-component signal

The reassignment coordinates and the resulting ScRe-Spect of a multi-component signal $x(t) = a_1 x_1(t - t_1)e^{i2\cdot f_1\pi t} + a_2 x_2(t - t_2)e^{i2\cdot f_2\pi t}$, where $x_1(t)$ is a Gaussian function and $x_2(t)$ a second Hermite function, can easily be calculated by a linear operation of the reassignment coordinates [14].

An illustration of the reassignment for such a multi-component signal, with $\left(t_1, f_1\right) = (4,4)$, $\left(t_2, f_2\right) = (8,8)$ and $c_t = c_\omega = 2$, is shown in Fig.

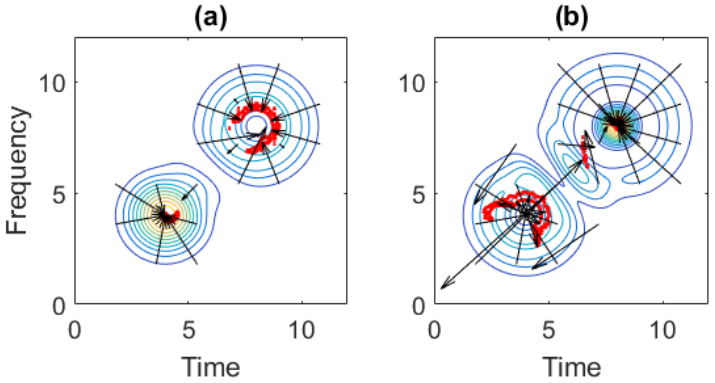

Figure 1: Illustration of the spectrogram and the scaled reassignment of a signal with one Gaussian, $(4, 4)$, and one second Hermite, $(8, 8)$, signal component. The spectrogram is shown by the gradient lines. The arrows represent the reassignment of the signal energy and the red areas show the most energy dense areas in the ScRe-Spect; (a) scaled reassignment with a Gaussian window; (b) scaled reassignment with a second Hermite window.

1. The gradient lines show the spectrogram, in Fig. 1(a) with a Gaussian window and in Fig. 1(b) with a second Hermite window. The arrows show how the signal energy is reassigned and the red areas mark the most energy dense parts in the ScRe-Spect. Fig. 1(a) shows that the energy from the Gaussian component is reassigned to a small area in the centre of the component, while the energy from the second Hermite component is reassigned to a circle around the centre of the component. In Fig. 1(b) the energy from the Gaussian component is reassigned to a circle and the energy from the second Hermite component is reassigned to a small area in the centre of the component. There is also some interaction between the components resulting in some small interference after reassignment.

In Fig. 1(a) it can also be seen that some energy is moved away from the centre of the second Hermite component. This is due to that the reassignment coordinates (12) grow large when $t, \omega \to 0$ and $c_t = c_\omega = 2$. This can also be seen in Fig. 1(b) for the Gaussian component. Similarly the reassignment coordinates (20) grow when $t, \omega \to 0$ and $c_t = c_\omega = 2$.

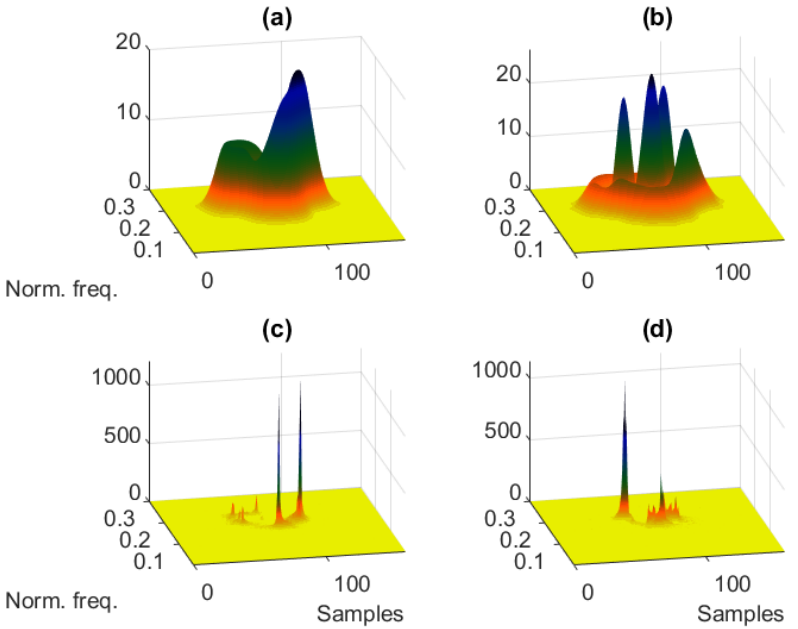

Figure 2: Detection and time-frequency localisation of individual components in a multi-component transient signal; (a) the spectrogram with Gaussian window; (b) the spectrogram with second Hermite window; (c) the ScRe-Spect with Gaussian window; (d) the ScRe-Spect with second Hermite window.

3 Simulations

A multi-component signal with Gaussian and second Hermite components can be resolved with the ScRe-Spect using Gaussian and second Hermite time windows. This is illustrated by the simulated signal

$$\begin{aligned} x(n) = {} & x_1(n-80)e^{i2\pi 0.14n} + x_2(n-100)e^{i2\pi 0.18n} \\ & + x_3(n-60)e^{i2\pi 0.19n} + e(n), \end{aligned} \tag{21}$$

where $x_1(n)$ and $x_2(n)$ are Gaussian functions and $x_3(n)$ a second Hermite function, all with lengths around 60 samples, and $e(n)$ is white Gaussian noise, SNR = 15 dB, where SNR is the average total signal energy to the variance of the noise. Fig. 2 shows the spectrogram and ScRe-Spect. The

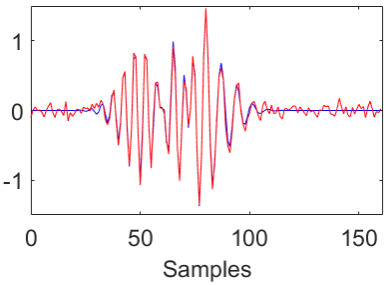

Figure 3: Two component signal (22), $f_1 = 0.14$, with and without noise, SNR 15 dB.

spectrogram and the ScRe-Spect, the illustrations are made in 3D to clearly show the difference in amplitude of the peaks, i.e. the energy density. In Fig. 2(a), showing the spectrogram with a Gaussian window, only one clear peak is visible and it is located at $(n, f) = (97, 0.167)$, however in Fig. 2(c), the ScRe-Spect with a Gaussian window, two peaks are clearly seen at $(80, 0.140)$ and $(101, 0.180)$, which are very close to the true centres of the two Gaussian components. Fig. 2(b), the spectrogram with a second Hermite window, shows many peaks and is hard to interpret, however Fig. 2(d) shows only one large peak at $(60, 0.190)$, also present in the corresponding spectrogram, which is the true centre of the second Hermite component.

3.1 Close components

Detection and localisation of components becomes harder if they are close in time and frequency. For a simulated signal with two components

$$x(n) = x_1(n - 80)e^{i2\pi f_1 n} + x_2(n - 60)e^{i2\pi 0.2n} + e(n), \qquad (22)$$

where $x_1(n)$ is a Gaussian function, $x_2(n)$ a second Hermite function and $e(n)$ is white Gaussian noise, SNR 15 dB, we can vary the normalised frequency f_1 to change the frequency distance between the components. The components overlap in time, see Fig. 3 where $f_1 = 0.14$.

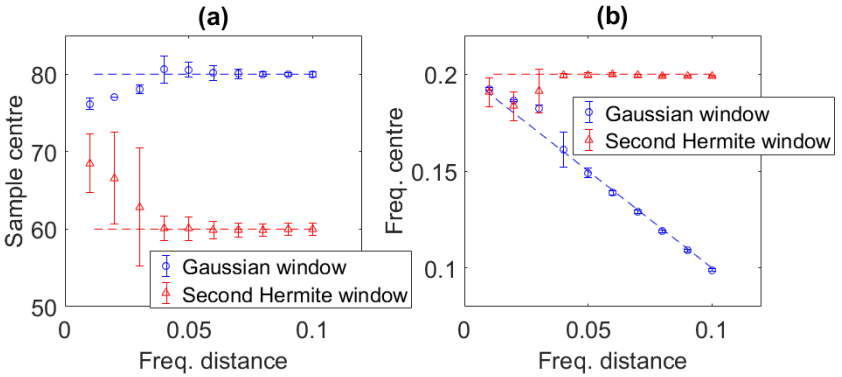

Figure 4: The frequency distance between the two components in signal (22) is varied to get the average sample-frequency centres of the largest peaks in the ScRe-Spect with Gaussian and second Hermite window respectively, as well as a 95% confidence interval, obtained from 200 noise simulations. The true sample-frequency centres of both signal components are also marked; (a) the sample centres of the largest peaks; (b) the frequency centres of the largest peaks.

We let $0.1 \leq f_1 \leq 0.19$ and simulate 200 signals with different noise for each f_1, to find the average sample-frequency (time-frequency) centres from the ScRe-Spect for both signal components. The sample-frequency centres are obtained by finding the largest peak in the ScRe-Spect, with a Gaussian window for the Gaussian component and a second Hermite window for the second Hermite component. The average sample centres and the 95% confidence intervals are shown in Fig. 4(a). Fig. 4(b) shows the average frequency centres and the 95% confidence intervals. It can be seen that the largest peak in the ScRe-Spect with a matched shape window accurately represents the centre of the signal component until the normalised frequency distance is only 0.03.

3.2 Noise sensitivity

Detection and localisation of components also becomes harder if the signal has low SNR. For a sampled signal (22) with $f_1 = 0.14$, the variance of the white Gaussian noise $e(n)$ can be varied to evaluate the noise sensitivity of

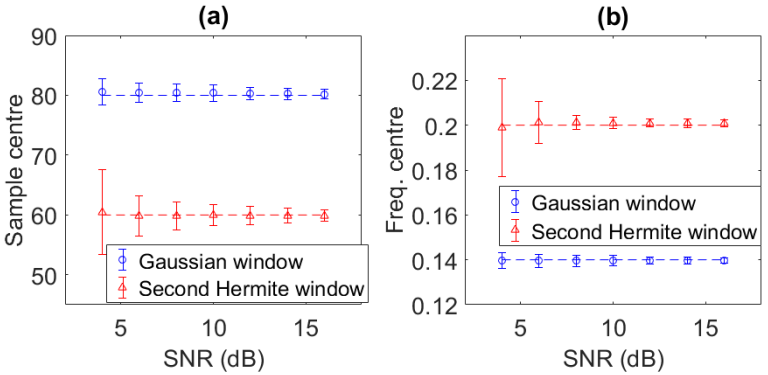

Figure 5: The SNR of the signal (22), $f_1 = 0.14$, is varied to get the average sample-frequency centres of the largest peaks in the ScRe-Spect with Gaussian and second Hermite window respectively, as well as a 95% confidence interval, obtained from 200 noise simulations. The true sample-frequency centres of both signal components are also marked; (a) the sample centre of the largest peaks; (b) the frequency centres of the largest peaks.

the ScRe-Spect. The SNR is varied between 2 and 16 dB and the number of simulations for each SNR is 200. The average sample centres and 95% confidence intervals from the ScRe-Spect with matching shape window for the two components are shown in Fig. 5(a). Fig. 5(b) shows the average frequency centres and the 95% confidence intervals. It can be seen that the method becomes unreliable for the second Hermite component around SNR 6 dB, while the localisation of the Gaussian component remains good even for low SNR.

4 Transient echolocation signal example

This section provides an example of a transient echolocation signal from a beluga whale (*Delphinapterus leucas*). The signal is sampled with 1 MHz and recorded by one of 47 simultaneously sampling hydrophones as described in [15]. The signal was chosen because it is recorded at the centre of the echolocation beam, based on the peak amplitude level the signal is sample by the hydrophone closest to the centre beam axis of the an-

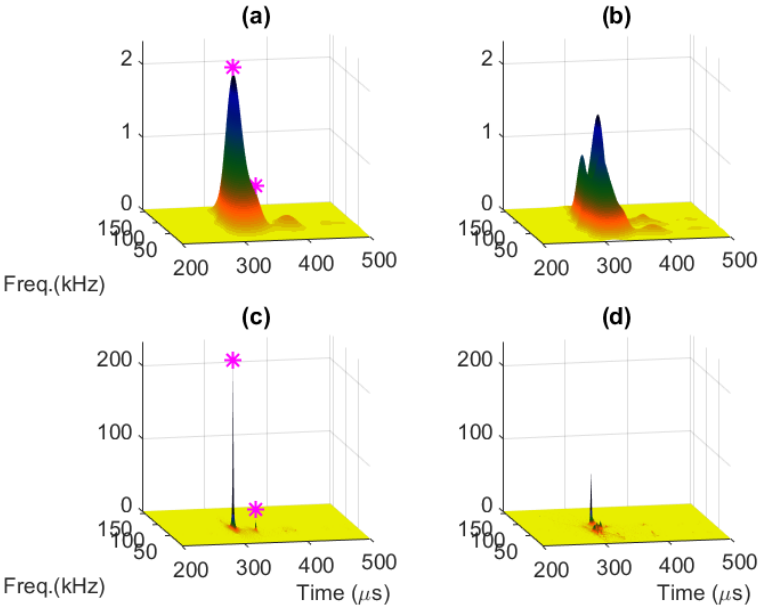

Figure 6: Detection and time-frequency localisation of individual components in a dolphin echolocation signal; (a) the spectrogram with Gaussian window; (b) the spectrogram with second Hermite window; (c) the ScRe-Spect with Gaussian window; (d) the ScRe-Spect with second Hermite window.

imal. Fig. 6(a) shows the spectrogram with a Gaussian window, Fig. 6(c) the ScRe-Spect with a Gaussian window, Fig. 6(b) the spectrogram with a second Hermite window and Fig. 6(d) the ScRe-Spect with a second Hermite window. Fig. 6(c) shows two peaks, one clear at (307 μs, 65 kHz) and one with smaller amplitude at (338 μs, 53 kHz), these peaks corresponds well to the spectrogram in Fig. 6(a). In Fig. 6(d) the signal energy is more scattered even if peaks appear, the peaks also do not correspond well to the spectrogram in Fig. 6(b). This suggests that the signal components in the dolphin echolocation signal more closely resembles Gaussian functions than second Hermite functions.

With additional measurements, this method could determine if the larger and smaller Gaussian like components originate from two locations within the sound generation mechanism of the dolphin or are the result of internal

reflection of the sound wave. Assuming an approximate sound velocity of 1490 m/s in the tissue, the time separation of the components corresponds to a total difference in sound wave travel distance of approximately 4.6 cm.

5 Conclusions

It is shown that perfect time-frequency localisation of a Gaussian and second Hermite transient signal component can be achieved by the ScRe-Spect using a matching shape time window. If a time window which do not match the shape of the signal component is used, the component energy is instead scattered in ellipses around the time-frequency centre of the component. It is shown that this can be used to find the time-frequency centres and the shapes of the individual transient signal components within a multi-component signal.

The results from simulated multi-component signals with noise show that the ScRe-Spect can resolve and correctly identify the time-frequency centres and component shapes even if the signal components are close in time and frequency. The ScRe-Spect is also robust to noise disturbances. The performance is evaluated on a measured dolphin echolocation signal, which gives good time-frequency localisation of what seems to be two Gaussian-like signal components.

References

[1] 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.

[2] 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.

[3] F. Auger, E. Chassande-Mottin, and P. Flandrin, "On phase-magnitude relationships in the short-time Fourier transform," *IEEE Signal Processing Letters*, vol. 19, no. 5, pp. 267–270, May 2012.

[4] P. Flandrin, "A note on reassigned Gabor spectrograms of Hermite functions," *J Fourier Analysis and Applications*, vol. 19, no. 2, pp. 285–295, 2013.

[5] A.I. Rasiah, R. Togneri, and Y. Attikiouzel, "Modeling 1-D signals using Hermite basis functions," in *IEEE Proc.-Vis. Image Signal Process.* IEEE, 1997, vol. 144, pp. 345–354.

[6] T. H. Linh, S. Osowski, and M. Stodolski, "On-line heart beat recognition using Hermite polynomials and neuro-fuzzy network," *IEEE Trans. on Instrumentation and Measurement*, vol. 52, no. 4, pp. 1224–1231, August 2003.

[7] B.N. Li, M.C. Dong, and M.I. Vai, "Modeling cardiovascular physiological signals using adaptive Hermite and wavelet basis functions," *IET Signal Processing*, vol. 4, no. 5, pp. 588–597, 2010.

[8] R. Ma, Z. Huang L. Shi, and Y. Zhou, "EMP signal reconstruction using associated-Hermite orthogonal functions," *IEEE Trans. on Electromagnetic Compatibility*, vol. 64, no. 6, pp. 1383–1390, March 2016.

[9] 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.

[10] 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.

[11] 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.

[12] J. Starkhammar and M. Hansson-Sandsten, "Evaluation of seven time-frequency representation algorithms applied to broadband echolocation signals," *Advances in Acoustics and Vibration*, vol. 2015, pp. 1–13, 2015.

[13] J. Starkhammar, I. Reinhold, P. Moore, D. Houser, and M. Sandsten, "Intra-click time-frequency patterns across the echolocation beam of a beluga whale," *The Journal of the Acoustical Society of America*, vol. 140, no. 4, pp. 3239–3239, 2016.

[14] 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.

[15] J. Starkhammar, M. Amundin, J. Nilsson, T. Jansson, S. Kuczaj, M. Almqvist, and H.W. Persson, "47-channel burst-mode recording hydrophone system enabling measurements of the dynamic echolocation behavior of free-swimming dolphins," *J Acoust Soc Am.*, vol. 126, no. 3, pp. 959–962, 2009.