On the statistics and practical application of the reassignment method for Gabor spectrograms

Erik M. Månsson

Master's Thesis

2019

Lund University
Faculty of Engineering
Centre for Mathematical Sciences
Mathematical Statistics

Abstract

The reassignment method is a technique for improving the concentration of signals in spectrograms and other time-frequency representations (TFR). It achieves this by displacing the points in a TFR according to the *reassignment vector* for every point. By doing so, the reassignment method gives perfect concentration of infinite constant frequency sinusoids, impulses and linear chirps.

A downside to the reassignment method is that it is fairly sensitive to noise. While this is well known, the subject of how noise affects the reassignment method is largely unexplored. Some important groundwork has been laid by Chassande-Mottin et al. In their report from 1996, they derived the density function of the reassignment vector, given that the signal is subjected to additive white Gaussian noise (AWGN). For Gabor (Gaussian windowed) spectrograms, a closed form expression of the density function is given.

This thesis largely builds on top of said result, and aims to extend the general knowledge about the statistics of reassigned spectrograms. The focus lies on Gabor spectrograms, and a rather practical approach is taken. First, some statistical properties of the reassignment vector are explored. From this, a Gaussian approximation is suggested which makes the density function for the reassignment vector feasible to work with.

Then, we look at how the reassigned spectrogram behaves as a whole when subjected to AWGN. The signals examined are those previously mentioned, all perfectly localized by the reassignment method. It shows that in the context of reassigning the spectrogram, these signals are equivalent. The resulting reassigned spectrogram turns out to be of infinite variance since the distribution is heavy tailed. However, its shape can still be related to the width of a Gaussian. By doing so, a simple formula is proposed which states the ratio of concentration given by the reassigned spectrogram compared to the original spectrogram.

Finally, based on the previous findings, an idea for a new method of resampling reassigned noisy spectrograms is proposed. This method attempts to mitigate the issue that the reassigned spectrogram "deteriorates" when resampled in a naive manner.

Acknowledgements

First and foremost, I would like to express a special thanks to my supervisor Prof. Maria Sandsten for her incredible support throughout this thesis project. She has gone above and beyond, giving me many fruitful discussions on my thoughts, invaluable feedback, and guiding me though the process. I would also like to especially thank my supervisor at Acconeer, Dr. Bo Lincoln, for always tirelessly discussing my work with me and continuously giving me feedback.

A big thanks to Acconeer and Dr. Peter Almers for the opportunity to work on my thesis, and also giving me the freedom to take it whatever direction I wanted. Finally, I would like to thank the rest of the algorithm team at Acconeer – Dr. Rikard Nelander, Dr. Daniel Jung, and David Montgomery – for their support and just being awesome to work with.

Lund, May 2019
Erik Månsson

Contents

Notation and abbreviations

Abbreviations

FT	Fourier Transform
STFT	Short-Time Fourier Transform
TFR	Time-Frequency Representation
RV	Reassignment Vector
NRV	Normalized Reassignment Vector
MWR	Matched Window Reassignment
AWGN	Additive White Gaussian Noise
SNR	Signal to Noise Ratio
PDF	Probability Density Function
RLF	Reassignment Localization Factor
KDE	Kernel Density Estimation
RMSE	Root Mean Square Error

S	Spectrogram
NS	Normalized Spectrogram
RS	Reassigned Spectrogram
NRS	Normalized Reassigned Spectrogram
SRS	Smoothed Reassigned Spectrogram

Statistical measures

E	Expectation value
V	Variance
D	Standard deviation
Skew	Skewness
Mode	Mode
ModeSkew	Pearson mode skewness

Notation

z^* is the complex conjugate of z. ν is used as the symbol for frequency. Typically, f denotes a probability density function (PDF). The Gaussian distribution with mean μ and variance σ^2 is written as $\mathcal{N}(\mu, \sigma^2)$. The brackets (), [], and {} are used interchangeably for readability. Additionally, we define the following symbols:

x is distributed as ...	$x \sim \mathcal{N}(\mu, \sigma^2)$
x is approximately distributed as ...	$x \dot{\sim} \mathcal{N}(\mu, \sigma^2)$
x is proportional to y	$x \propto y$
x is approximately proportional to y	$x \propto\!\!\!\!\approx y$
x is defined as y	$x \triangleq y$

Integrals without limits imply integration over all real numbers $\mathbf{R}$, i.e., from $-\infty$ to ∞:

$$\int = \int_{\mathbf{R}} = \int_{-\infty}^{\infty}$$

In a similar fashion, we write integration over the whole complex plane $\mathbf{C}$ as:

$$\int_{\mathbf{C}} f(z)\,\mathrm{d}z = \iint f(z)\,\mathrm{d}(\mathrm{Re}\{z\})\,\mathrm{d}(\mathrm{Im}\{z\})$$

Let δ denote the Dirac delta function with the following properties:

$$\int \delta(x) f(x)\,\mathrm{d}x = f(0), \quad \iint \delta(x, y) f(x, y)\,\mathrm{d}x\,\mathrm{d}y = f(0, 0), \quad \int_{\mathbf{C}} \delta(z) f(z)\,\mathrm{d}z = f(0)$$

The complex Gaussian distribution is written as

$$x \sim \mathcal{CN}(0, 1) \Rightarrow \mathrm{Re}\{x\} \sim \mathrm{Im}\{z\} \sim \mathcal{N}(0, 1/2)$$

where $\mathrm{Re}\{x\}$ and $\mathrm{Im}\{z\}$ are independent.

The mode is defined as the value with maximum probability. For a distribution with a probability density function $f(z)$,

$$\mathrm{Mode}[z] = \operatorname*{argmax}_z f(z)$$

Chapter 1
Introduction

1.1 Background

Non-stationary stochastic processes are all around us, from the changing temperature of our globe to the response from a complex radar system. The need for analyzing these kind of signals arises practically everywhere, and is ever so challenging. For the last of decades, perhaps more so than ever. With increasing computing power available at our fingertips, the methods we use grow more and more complex and sophisticated.

One such method is the reassignment method, originally presented by Kodera et al. in 1976 [1], and then re-introduced by Auger and Flandrin in 1995 [2]. It was not until then that the general interest for the method sparked. Since then, lots of work has been done around this method.

A modern notable example is the method of synchrosqueezing, presented in 2011 by Daubechies et al. [3]. Synchrosqueezing is a special case of the reassignment method, applied in the context of wavelet transforms. We also have the matched window reassignment by Sandsten et al., presented in 2018 [4]. They showed that all transient signals with a known envelope can be perfectly localized using a modification to the reassignment method.

A very important article for this thesis is *On the statistics of spectrogram reassignment vectors* (translated) by Chassande-Mottin et al. published in 1996 [5]. In it, they derive key results on the density function of reassignment vectors for signals subjected to noise. Overall, it allows us to better understand the shortcomings of the method.

Robustness is key for successful application of many time-frequency methods, and the reassignment method gets no exception. Since it can be rather fragile, finding improvements for the reassignment method could potentially be useful for many applications.

1.2 Purpose

The purpose of this thesis is mainly to explore and learn more about the reassignment method. The focus is on its real world application, in which signals are discrete and always noisy to some extent. Previous research has shown that reassigned spectrogram sometimes behave poorly these situations. By learning about its behavior when subjected to noise, the aim is to be able to come up with suggestions for improving the method when applied in practice. Since this thesis is done in collaboration with Acconeer, a radar sensor company, we would also like to investigate how the method can be used for their sensors.

1.3 Outline

Chapter 2 walks through the necessary concepts and theory to understand the reassignment method.

Chapter 3 explores the statistical properties of the reassignment vector for Gabor (Gaussian windowed) spectrograms.

Chapter 4 takes a step back from just looking at the reassignment vector, and instead analyzes the reassigned spectrogram as a whole.

Chapter 5 discusses the issue of resampling noisy reassigned spectrograms, and proposes a new method to mitigate the issue.

Chapter 6 discusses some final thoughts on the subject and summarizes the work done in this thesis. Finally, some suggestions for further work in the field is presented.

Chapter 2
Spectrograms and the reassignment method

In this chapter, we will walk through the necessary concepts to understand the reassignment method. It is assumed that the reader is familiar to the concept of basic time-frequency analysis. A great introduction to the subject is the book *Time-frequency analysis of time-varying signals and non-stationary processes* [6] by Sandsten, found freely online. The primary focus here will be on the reassignment for Gabor (Gaussian windowed) spectrograms.

2.1 The short-time Fourier transform and spectrogram

We start by defining the Fourier transform (FT) of a signal x as

$$X(\nu) = \int x(s) \exp(-i2\pi\nu s)\, \mathrm{d}s \tag{2.1}$$

and introducing the Fourier transform operator $\mathcal{F}$:

$$X = \mathcal{F}x \tag{2.2}$$

The short-time Fourier transform (STFT) of a signal x using a window function h is defined as

$$F_x^h(t,\nu) = \int x(s)\, h^*(s-t) \exp(-i2\pi\nu s)\, \mathrm{d}s \tag{2.3}$$

with the corresponding spectrogram defined as

$$S_x^h(t,\nu) = \left|F_x^h(t,\nu)\right|^2 \tag{2.4}$$

As in [7, 2, 5], we require the STFT window h to be of unit energy:

$$\int |h(t)|^2\, \mathrm{d}t = 1 \tag{2.5}$$

The Gabor spectrogram is simply the spectrogram using a Gaussian window function. We define this function with a scaling parameter (standard deviation) λ as

$$h(t) = \frac{1}{\pi^{1/4}\sqrt{\lambda}} \exp\left(-\frac{t^2}{2\lambda^2}\right) \tag{2.6}$$

where the normalizing constant satisfies that the window is of unit energy:

$$\int |h(t)|^2 \,\mathrm{d}t = \frac{1}{\sqrt{\pi}\lambda} \int \exp\left(-\frac{t^2}{\lambda^2}\right) \mathrm{d}t \overset{(A.1)}{=} \frac{1}{\sqrt{\pi}\lambda} \cdot \sqrt{\pi}\lambda = 1 \tag{2.7}$$

The Fourier transform of the window is

$$H(\nu) = (\mathcal{F}h)(\nu) = \frac{1}{\pi^{1/4}\sqrt{\lambda}} \int e^{-t^2/2\lambda^2} e^{-i2\pi\nu s} \,\mathrm{d}s \overset{(A.7)}{=} \sqrt{2\lambda}\pi^{1/4} e^{-2(\pi\lambda\nu)^2} \tag{2.8}$$

As in [7, 5] we define the duration Δt_x and bandwidth $\Delta\nu_x$ for a signal x as

$$\Delta t_x^2 = \int t^2 \,|x(t)|^2 \,\mathrm{d}t \quad \text{and} \quad \Delta\nu_x^2 = \int \nu^2 \,|X(\nu)|^2 \,\mathrm{d}\nu \tag{2.9}$$

For the Gaussian window h, these become

$$\Delta t_h^2 = \int t^2 |h(t)|^2 \,\mathrm{d}t = \frac{1}{\sqrt{\pi}\lambda} \int t^2 e^{-t^2/\lambda^2} \,\mathrm{d}t \overset{(A.8)}{=} \frac{1}{\sqrt{\pi}\lambda} \cdot \frac{\sqrt{\pi}\lambda^3}{2} = \frac{\lambda^2}{2} \tag{2.10}$$

$$\Delta\nu_h^2 = \int \nu^2 |H(\nu)|^2 d\nu = 2\lambda\sqrt{\pi} \int \nu^2 e^{-(2\pi\lambda\nu)^2} d\nu \overset{(A.8)}{=} 2\lambda\sqrt{\pi} \cdot \frac{\sqrt{\pi}}{2(2\pi\lambda)^3} = \frac{1}{2^3\pi^2\lambda^2} \tag{2.11}$$

i.e.,

$$\Delta t_h = \frac{\lambda}{\sqrt{2}} \quad \text{and} \quad \Delta\nu_h = \frac{1}{\sqrt{2}\,2\pi\lambda} \tag{2.12}$$

Let us show an example. Let $x(t)$ be a signal consisting of two linear chirps:

$$x(t) = \exp\left[i2\pi \cdot \left(0.15 - \frac{0.0005}{2} \cdot t\right) \cdot t\right] + \exp\left[i2\pi \cdot \left(0.30 + \frac{0.0010}{2} \cdot t\right) \cdot t\right] \tag{2.13}$$

The signal and its Gabor spectrogram ($\lambda = 5$) can be seen in figure 2.1.

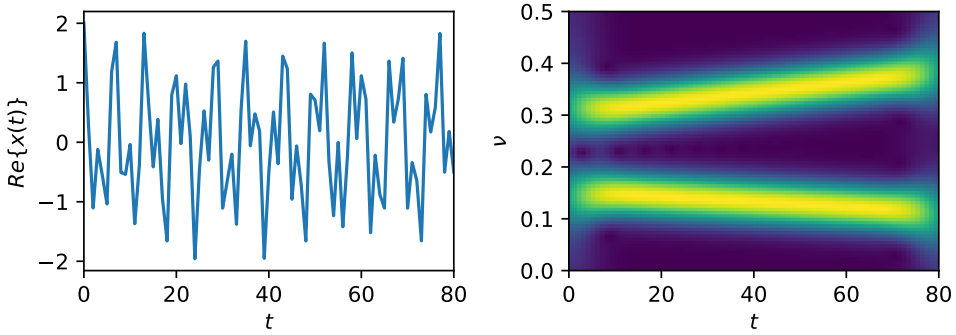

Figure 2.1: An illustration of the signal x (2.13), consisting of two linear chirps. The left plot shows the signal, and the right plot shows its Gabor spectrogram ($\lambda = 5$).

2.2 The reassignment method for spectrograms

Ideally, ignoring all theoretical limitations, we would want the time-frequency representation (TFR) of a signal to be perfectly localized. A fully concentrated estimate of the power and instantaneous frequency. Take for example the signal from before (2.13), shown in figure 2.1. What we mean by the ideal TFR is shown in figure 2.2.

Unfortunately, this is not attainable due to the Heisenberg-like uncertainty principle [7]. Broadening the window decreases its bandwidth, which improves the localization in the frequency domain. But this also has the effect of increasing its duration, worsening the localization in the time domain. Simply put, you can not achieve perfect localization both in time and frequency.

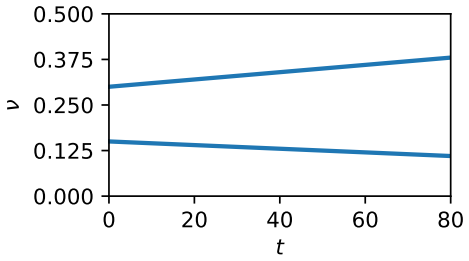

Figure 2.2: The ideal TFR of the signal (2.13) shown in figure 2.1.

The *reassignment method* aims to improve the concentration of a TFR by reallocating its energy distribution in the time-frequency plane. It was originally introduced by Kodera et al. in 1976 [1] and then re-introduced by Auger and Flandrin in 1995 [2]. It works by estimating the offset in center of gravity for each point in the TFR, then moving them accordingly. For many signals, the method achieves perfect localization [2, 8, 4], as we shall soon see for ourselves.

For the original definitions of the reassignment method, see [1]. In this thesis, we will limit ourselves to the reassignment of spectrograms, and therefore use the formulation found in [2]. In fact, we will limit further analysis to Gabor spectrograms specifically, but the following formulation of the reassignment holds for all spectrograms.

Let $\mathcal{D}$ and $\mathcal{T}$, respectively, be the operators of differentiation and multiplication by the running variable:

$$\mathcal{D}h(t) = \frac{\mathrm{d}}{\mathrm{d}t}h(t), \ \mathcal{T}h(t) = t \cdot h(t) \tag{2.14}$$

Figure 2.3 shows the result of these operators being applied to the Gaussian window.

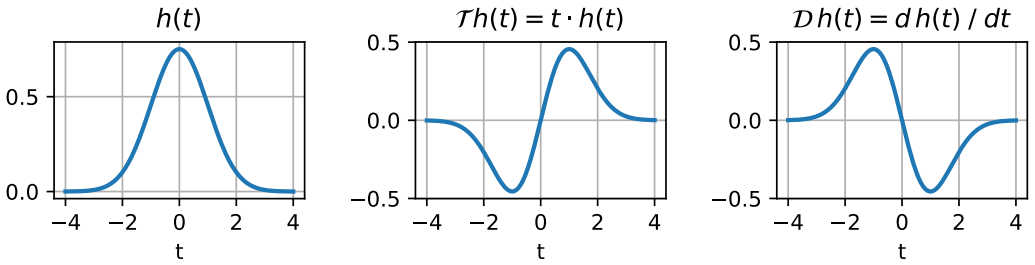

Figure 2.3: The Gaussian window h with unit scale ($\lambda = 1$) with the operators $\mathcal{T}$ and $\mathcal{D}$ applied to it.

The reassigned time and frequency $(\hat{t}_x^h, \hat{\nu}_x^h)$ for a point (t, ν) in the spectrogram S_x^h can be calculated from

$$\hat{t}_x^h(t,\nu) = t + c_t \tilde{t}_x^h(t,\nu) \tag{2.15}$$

$$\hat{\nu}_x^h(t,\nu) = \nu + c_\nu \tilde{\nu}_x^h(t,\nu) \tag{2.16}$$

where

$$\tilde{t}_x^h(t,\nu) = \mathrm{Re}\left\{\frac{F_x^{\mathcal{T}h}(t,\nu)}{F_x^h(t,\nu)}\right\} \tag{2.17}$$

$$\tilde{\nu}_x^h(t,\nu) = -\frac{1}{2\pi}\,\mathrm{Im}\left\{\frac{F_x^{\mathcal{D}h}(t,\nu)}{F_x^h(t,\nu)}\right\} \tag{2.18}$$

The reassigned spectrogram can then be expressed as

$$RS_x^h(t,\nu) = \iint S_x^h(t',\nu') \cdot \delta[t - \hat{t}_x(t',\nu'), \nu - \hat{\nu}_x(t',\nu')]\,\mathrm{d}t'\,\mathrm{d}\nu' \tag{2.19}$$

In the original definitions [1, 2], the parameters c_t and c_ν were not included. Hence, if $c = c_t = c_\nu = 1$ we obtain the ordinary reassignment. These parameters were introduced by Sandsten and Brynolfsson in 2015 [8] which has proven useful in their *matched window reassignment* [4]. This will be further discussed in section 2.4.

Figure 2.4 and 2.5 show two different representations of the reassigned spectrogram of the signal/spectrogram seen in figure 2.1. Apart from the prominent edge effects, the reassigned spectrogram does indeed locate the components in the signal extremely well. The representations will be discussed more thoroughly in chapter 5.

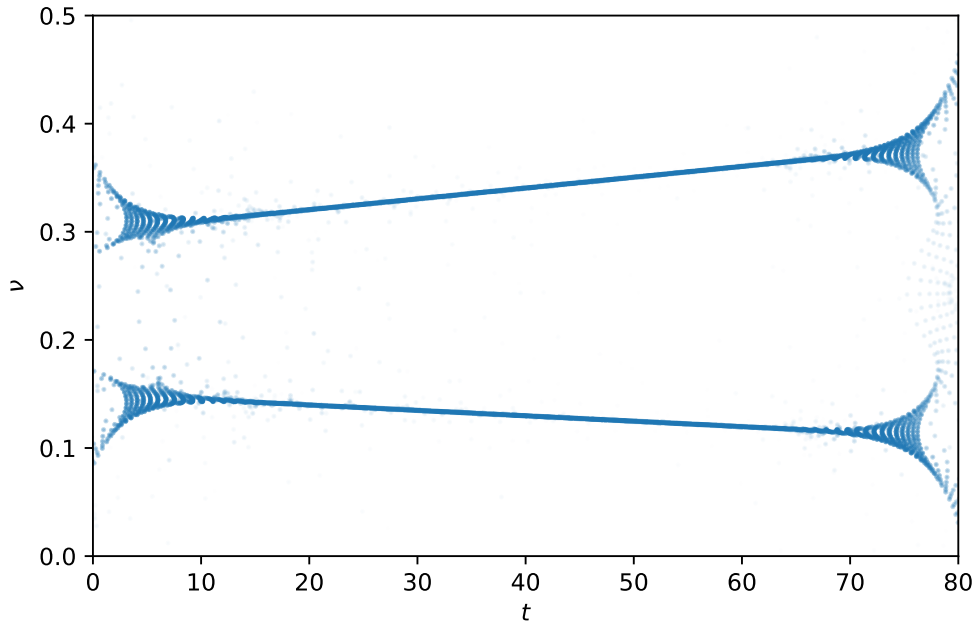

Figure 2.4: The reassigned points from the signal/spectrogram in figure 2.1. We refer to this as the *raw* representation of the reassigned spectrogram.

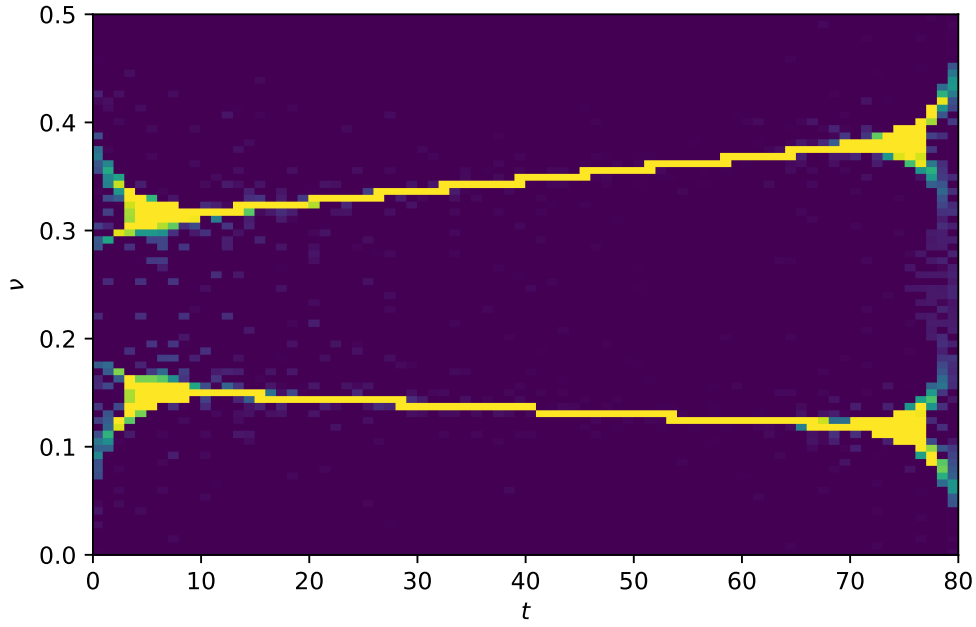

Figure 2.5: What we refer to as the *binned* reassigned spectrogram from the signal/spectrogram in figure 2.1.

For reasons we will soon see, as done in [5], we define the dimensionless *normalized reassignment vector* (NRV) as

$$\begin{aligned} r(t,\nu) &= \frac{\tilde{t}_x^h(t,\nu)}{\Delta t_h} + i\,\frac{\tilde{\nu}_x^h(t,\nu)}{\Delta \nu_h} \\ &= \frac{1}{\Delta t_h}\operatorname{Re}\left\{\frac{F_x^{\mathcal{T}h}(t,\nu)}{F_x^h(t,\nu)}\right\} - \frac{i}{2\pi\Delta\nu_h}\operatorname{Im}\left\{\frac{F_x^{\mathcal{D}h}(t,\nu)}{F_x^h(t,\nu)}\right\} \end{aligned} \tag{2.20}$$

Going backwards from the NRV,

$$\tilde{t}_x^h(t,\nu) = \Delta t_h \cdot \operatorname{Re}\{r(t,\nu)\} \quad \text{and} \quad \tilde{\nu}_x^h(t,\nu) = \Delta\nu_h \cdot \operatorname{Im}\{r(t,\nu)\} \tag{2.21}$$

which is quite practical for implementation.

For the Gaussian window,

$$\mathcal{D}h(t) = \frac{\mathrm{d}}{\mathrm{d}t}h(t) \overset{(A.5)}{=} -\frac{1}{\lambda^2}\cdot t \cdot h(t) = -\frac{1}{\lambda^2}\cdot \mathcal{T}h(t) \tag{2.22}$$

which, since the Fourier transform is linear, also means that

$$F_x^{\mathcal{D}h}(t,\nu) = -\frac{1}{\lambda^2}F_x^{\mathcal{T}h}(t,\nu) \tag{2.23}$$

Thus, for the Gabor spectrogram,

$$\begin{aligned} r(t,\nu) &\overset{(2.20)}{=} \frac{1}{\Delta t_h}\operatorname{Re}\left\{\frac{F_x^{\mathcal{T}h}(t,\nu)}{F_x^h(t,\nu)}\right\} - i\frac{1}{2\pi\Delta\nu_h}\operatorname{Im}\left\{\frac{F_x^{\mathcal{D}h}(t,\nu)}{F_x^h(t,\nu)}\right\} \\ &\overset{(2.23)}{=} \frac{1}{\Delta t_h}\operatorname{Re}\left\{\frac{F_x^{\mathcal{T}h}(t,\nu)}{F_x^h(t,\nu)}\right\} + i\frac{1}{2\pi\lambda^2\Delta\nu_h}\operatorname{Im}\left\{\frac{F_x^{\mathcal{T}h}(t,\nu)}{F_x^h(t,\nu)}\right\} \\ &\overset{(2.12)}{=} \frac{\sqrt{2}}{\lambda}\operatorname{Re}\left\{\frac{F_x^{\mathcal{T}h}(t,\nu)}{F_x^h(t,\nu)}\right\} + i\frac{\sqrt{2}}{\lambda}\operatorname{Im}\left\{\frac{F_x^{\mathcal{T}h}(t,\nu)}{F_x^h(t,\nu)}\right\} \\ &= \frac{\sqrt{2}}{\lambda}\cdot\frac{F_x^{\mathcal{T}h}(t,\nu)}{F_x^h(t,\nu)} \end{aligned} \tag{2.24}$$

This eventually leads to another result derived by Chassande-Mottin et al. [5]. It is presented in chapter 3, and more or less lays the foundation for this whole thesis. This definition also helps us greatly in deriving the theoretical reassignment for some signals.

2.3 The spectrogram and reassignment of some common signals

There are three signals that we would now like to discuss – the impulse, constant frequency sinusoid, and linear chirp. Those signals frequently appear in the context of reassignment. This is with good reason, as they are all perfectly localized by the

reassignment method. In fact, it can be argued that in this context, the signals are in some sense equivalent. They are all infinitely spread in the spectrogram, and they all form an infinite line in it. Chapter 4 deals with this in much further detail. But for now, let us just convince ourselves that the reassignment actually perfectly localizes these signals.

Calculation of the Gabor spectrogram and reassignment of these signals is not hard, although slightly tedious. For the full calculations, see appendix B. The results match what is presented in [5], except for r_0 of the linear chirp which we believe to be off by a factor $\sqrt{4\pi}$. Later on, in chapter 3, r_0 will defined as the noise-free NRV. But thus far we have not added any noise to our signal model, so in the current context $r_0 = r$.

We have already seen by example that the reassignment perfectly localizes the linear chirp. Again, showing that this is true is not hard, but tedious. Instead, we will cover the impulse and sinusoid, starting with the former. We define it as

$$s(t) = \delta(t - t_0) \tag{2.25}$$

Its spectrogram and NRV is, respectively,

$$S(t, \nu) = \frac{1}{\sqrt{\pi}\lambda} \exp\left(-\frac{(t - t_0)^2}{\lambda^2}\right) \quad \text{and} \quad r(t, \nu) = -\frac{\sqrt{2}}{\lambda}(t - t_0) \tag{2.26}$$

which means that (2.21)

$$\tilde{t} = -(t - t_0) \Leftrightarrow \hat{t} = t_0 \quad \text{and} \quad \tilde{\nu} = 0 \Leftrightarrow \hat{\nu} = \nu \tag{2.27}$$

which in turn means that the localization is perfect. Also, it allows us to reformulate the spectrogram as

$$S(t, \nu) \propto \exp\left[-\frac{1}{2}\left(\frac{\tilde{t}}{\Delta t}\right)^2\right] \tag{2.28}$$

Let us compare this to the sinusoid:

$$s(t) = \exp(i2\pi\nu_0 t) \tag{2.29}$$

$$S(t, \nu) = 2\sqrt{\pi}\lambda \exp\left[-4\pi^2\lambda^2(\nu - \nu_0)^2\right] \quad \text{and} \quad r(t, \nu) = -i2\sqrt{2}\pi\lambda(\nu - \nu_0) \tag{2.30}$$

$$\tilde{t} = 0 \Leftrightarrow \hat{t} = t \quad \text{and} \quad \tilde{\nu} = -(\nu - \nu_0) \Leftrightarrow \hat{\nu} = \nu_0 \tag{2.31}$$

Here, we can reformulate the spectrogram as

$$S(t, \nu) \propto \exp\left[-\frac{1}{2}\left(\frac{\tilde{\nu}}{\Delta \nu}\right)^2\right] \tag{2.32}$$

which is suspiciously similar to the reformulation of the spectrogram of the impulse. In fact, in chapter 4, we will show that this kind of reformulation also works for the linear chirp. This is why we argue that in some way, the three signals can be seen as equivalent in the context of reassignment.

2.4 The matched window reassignment

We previously introduced the scaling parameters of the reassignment – c_t and c_ν. Now, their purpose will be shown. Consider a Gaussian windowed constant frequency signal

$$s(t) = \exp\left(-\frac{(t-t_0)^2}{2\sigma^2}\right) \cdot \exp(i2\pi\nu_0 t) \tag{2.33}$$

where (t_0, ν_0) is its center time and frequency, and σ is its shape parameter (standard deviation). As noted in [8, 7], this signal is time-frequency and shift-invariant. Hence, for further analysis we are allowed to let $(t_0, \nu_0) = (0, 0)$, and the results will still apply for any (t_0, ν_0). To clarify, we will analyze the simplified signal

$$s(t) = \exp\left(-\frac{t^2}{2\sigma^2}\right) \tag{2.34}$$

Calculations (see appendix B.4) give that for this signal,

$$r(t,\nu) = -\frac{\sqrt{2}\lambda}{\lambda^2+\sigma^2}(t + i2\pi\sigma^2\nu) \Rightarrow \tilde{t} = -\frac{\lambda^2}{\lambda^2+\sigma^2} \cdot t \Rightarrow \hat{t} = t \cdot \left(1 - c_t \cdot \frac{\lambda^2}{\lambda^2+\sigma^2}\right) \tag{2.35}$$

If we let

$$c_t = \frac{\lambda^2+\sigma^2}{\lambda^2} \tag{2.36}$$

then perfect localization in time is achieved. But perhaps more interesting is matching the signal window and STFT window, i.e., letting $(\sigma = \lambda)$. Then,

$$r(t,\nu) = -\frac{1}{\sqrt{2}\lambda}(t + i2\pi\lambda^2\nu) \Rightarrow \tilde{t} = -\frac{1}{2} \cdot t \Rightarrow \hat{t} = t \cdot \left(1 - c_t \cdot \frac{1}{2}\right) \tag{2.37}$$

If we now let $c_t = 2$, perfect reassignment is always achieved! The same result, given from the same steps, is obtained for the frequency. Therefore, we can let

$$c_t = c_\nu = c = 2 \tag{2.38}$$

This was originally discovered by Sandsten and Brynolfsson in 2015 [8]. In 2018, Sandsten et al. found that this works for any arbitrary window matched to the signal [4]. They refer to this method the *matched window reassignment*, or MWR for short.

Let us end this chapter some more illustrative examples. Let $s(t)$ be a Gaussian with $\sigma = 10$ and constant center frequency $\nu_0 = 0.17$:

$$s(t) = \exp\left(-\frac{t^2}{2 \cdot 10^2}\right) \cdot \exp(i2\pi \cdot 0.17 \cdot t) \tag{2.39}$$

First, we let our signal $x(t) = s(t)$. The left plot of figure 2.6 shows the real part of this signal and its envelope. The right plot shows the corresponding matched Gabor spectrogram, where the center frequency ν_0 is marked. The reassigned spectrogram simply becomes a dot, perfectly concentrated in a single point. Therefore, we chose not to show that plot since the point would be barely visible.

What we instead like to do is to show the slice of the spectrogram, as shown by the white line in figure 2.6. The slice is taken at the frequency that matches the center frequency ν_0 of the signal. It is shown in the leftmost plot in figure 2.7, with its points marked for clarity. The two right plots shows the reassigned counterpart of the spectrogram slice. Evidently, perfect localization is indeed obtained.

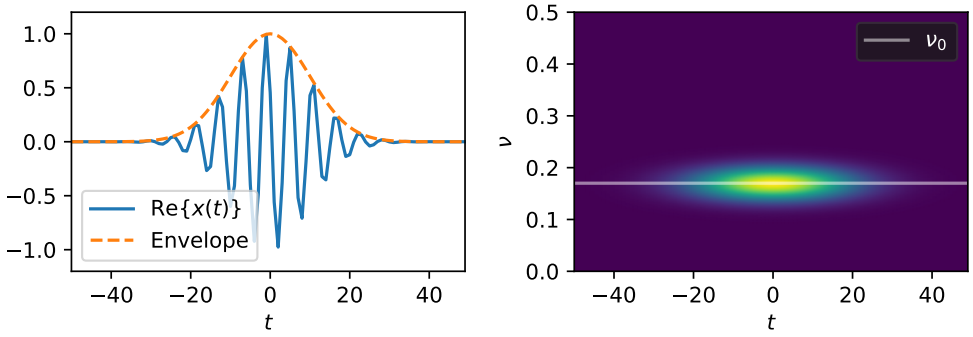

Figure 2.6: The signal (2.39) and its corresponding matched Gabor spectrogram. The center frequency of the Gaussian is marked in the spectrogram.

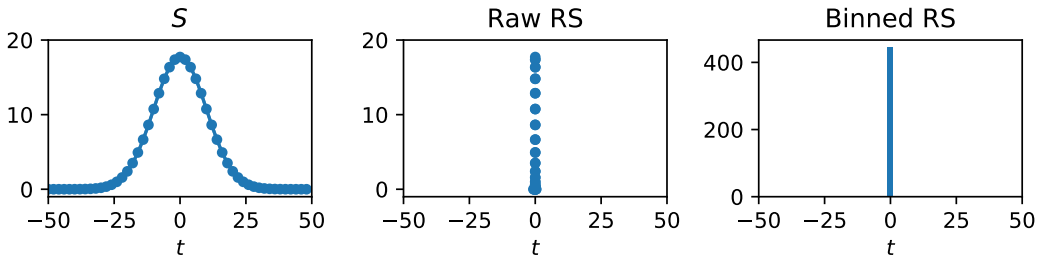

Figure 2.7: The leftmost plot shows the slice of the spectrogram at the frequency matching that of the signal, as seen in figure 2.6. The two other plots show the reassigned counterpart, both raw and binned.

As a final demonstration, let us show what can happen if noise is added to the signal. Now, let $x(t) = s(t) + n(t)$ where $n(t)$ is additive white Gaussian noise (AWGN) as defined in chapter 3, section 3.1. For reference, the standard deviation of the noise $\sigma_n = 0.3$.

Figure 2.8 and 2.9 shows the signal and spectrogram in the same way as before. It can be seen that the noise is strong enough that it doubtlessly affects the spectrogram and its reassignment. This is what the following chapters will be all about. Take special note of how the amplitude of the binned reassigned spectrogram is much lower than before. This is because the signal energy now happens to spread over multiple bins. Why this happens and what we can do about it is the topic of chapter 5. But first, in the next chapter, we will start with how noise affects the normalized reassignment vector.

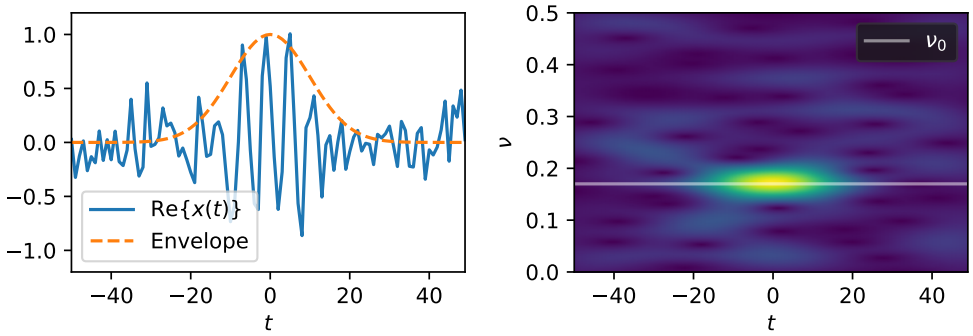

Figure 2.8: Same as figure 2.6 but with noise added to the signal.

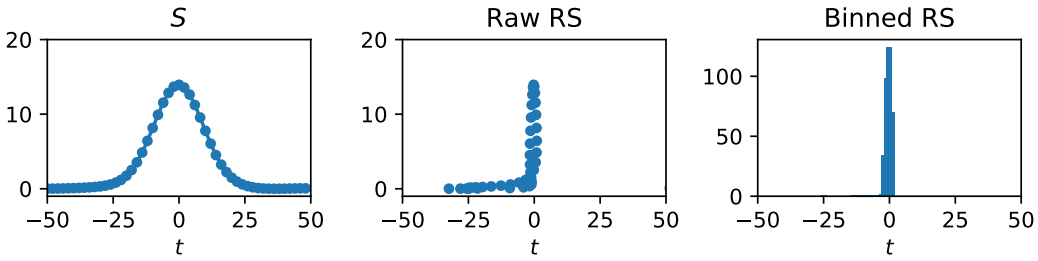

Figure 2.9: Same as figure 2.7 but with noise added to the signal.

Chapter 3

Statistics of the normalized reassignment vector for Gabor spectrograms

In 1996, Chassande-Mottin et al. released a paper [5] deriving the probability density function (PDF) of the normalized reassignment vector for spectrograms when subjected to Gaussian white noise. The paper is written in French, but luckily they have blessed the non-French speakers with a summary in English [9]. In this chapter, the statistical properties of the reassignment vector for Gabor (Gaussian windowed) spectrograms is explored. Also, an approximation of the PDF is presented which could be useful for some applications.

3.1 Definitions and assumptions

Throughout the rest of this thesis, we will work with a signal model where a deterministic signal is subjected to additive white Gaussian noise. We define this simple model as

$$x(t) = s(t) + n(t) \tag{3.1}$$

where $s(t)$ is an analytic deterministic complex signal, and $n(t)$ is analytic (and circular) white Gaussian noise with variance σ_n^2, as defined in [5] and further discussed in [10].

The noise being analytic means that $\mathrm{Re}\{n(t)\} \sim \mathrm{Im}\{n(t)\} \sim \mathcal{N}(0, \sigma_n^2/2)$, and that $\mathrm{Re}\{n(t)\}$ and $\mathrm{Im}\{n(t)\}$ form a Hilbert transform pair. This in turn means that the noise has strictly positive frequency content and that $\mathrm{Re}\{n(t)\}$ and $\mathrm{Im}\{n(t)\}$ are correlated. In MATLAB this can be generated with (`T` being the signal length)

```
n = hilbert(randn(T, 1) * sqrt(noise_variance / 2))
```

This is in contrast to uncorrelated circular white Gaussian noise, where $\mathrm{Re}\{n(t)\}$ and $\mathrm{Im}\{n(t)\}$ are independent. In this case, the power is spread in the whole spectra, and as such the power density is halved. This is an important distinction since the signal to noise ratio later defined will be doubled if uncorrelated noise is used in place of the analytic. This noise can be generated with

```
n = (randn(T, 1) + 1j * randn(T, 1)) * sqrt(noise_variance / 2)
```

The Gabor spectrogram is obtained by transforming the signal x with a STFT using the Gaussian window h, as presented in chapter 2. The *local signal to noise ratio* (SNR) ρ is defined such that it relates to the power ratio between the deterministic signal s and

noise n at some time t and frequency ν in the spectrogram:

$$\rho = \rho(t, \nu) = \frac{S_s^h(t, \nu)}{2\sigma_n^2} = \frac{\left|F_s^h(t, \nu)\right|^2}{2\sigma_n^2} \tag{3.2}$$

This exact definition comes from [5]. It can be discussed if the 2 should be left out of the expression or not, but in this thesis we decided use their definition. Note that $\rho \geq 0$. In the case when $\rho = 0$, $x(t) = n(t)$, and when $\rho = \infty$, $x(t) = s(t)$.

In this chapter, we are going to work exclusively with the normalized reassignment vectors (NRV:s) r from the signal x, which now have some probability distribution since x has a random component. The definition of r can be found in (2.24). To be able to work with this model, we also need to define r_0 as the *noise-free* NRV. That is, r_0 is the theoretical value of the NRV for the deterministic signal component $s(t)$. This was calculated for some common signals in the previous chapter under section 2.3.

3.2 Probability density function

3.2.1 Introduction

With the definitions and assumptions above, the probability density function (PDF) f of the NRV r is [5]:

$$\begin{aligned} f(r) &= \frac{1}{\pi\left(1 + |r|^2\right)^2}\left[1 + \rho\left(1 + |r_0|^2 - \frac{|r - r_0|^2}{1 + |r|^2}\right)\right]\exp\left(-\rho\frac{|r - r_0|^2}{1 + |r|^2}\right) \\ &= \frac{1}{\pi\left(1 + |r|^2\right)^2}\left(1 + \rho\frac{\left|1 + r r_0^*\right|^2}{1 + |r|^2}\right)\exp\left(-\rho\frac{|r - r_0|^2}{1 + |r|^2}\right) \end{aligned} \tag{3.3}$$

Please note that since f is a PDF, $f : \mathbf{C} \to \mathbf{R}$. In the next chapter, f will be written as f^r for clarity.

In the case when $\rho = 0$, i.e., when $x(t) = n(t)$, the PDF reduces to:

$$f(r) = \frac{1}{\pi\left(1 + |r|^2\right)^2} \tag{3.4}$$

When $\rho = \infty$, i.e., when $x(t) = s(t)$, per definition:

$$f(r) = \delta(r - r_0) \tag{3.5}$$

When $r_0 = 0$:

$$f(r) = \frac{1}{\pi}\left(\frac{1}{\left(1 + |r|^2\right)^2} + \frac{\rho}{\left(1 + |r|^2\right)^3}\right)\exp\left(-\rho\frac{|r|^2}{1 + |r|^2}\right) \tag{3.6}$$

Note that for both the case when $\rho = 0$ and when $r_0 = 0$, $f(r) = f(|r|)$, meaning that the distribution only depends on the length of r and is therefore circularly symmetric.

The PDF $f(r)$ depends on r_0 and ρ implicitly, which is reasonable since f is a PDF. However, sometimes we need to be explicit about its dependence on r_0. Therefore $g(r, r_0)$ is defined which is the same PDF but with explicit dependence on r_0. At other times it will prove useful to let $r = x+iy$, in which case we will write the joint distribution as $f(r) = f(x, y) = f_{x,y}(x, y)$. The marginal distribution of $\mathrm{Re}\{r\} = x$ and $\mathrm{Im}\{r\} = y$ are written as $f_x(x)$ and $f_y(y)$ respectively. To be clear,

$$f_x(x) = \int f_{x,y}(x, y)\,\mathrm{d}y \quad \text{and} \quad f_y(y) = \int f_{x,y}(x, y)\,\mathrm{d}x \tag{3.7}$$

3.2.2 Properties

Rotational invariance

The rotation of r_0 in the complex plane rotates the PDF $f(r)$ correspondingly. Except for the rotation, the distribution stays the same. This property is seen from that

$$\begin{aligned} g\left(re^{i\phi}, r_0e^{i\phi}\right) &= \frac{1}{\pi\left(1+\left|re^{i\phi}\right|^2\right)^2}\left(1+\rho\frac{\left|1+re^{i\phi}r_0^*e^{-i\phi}\right|^2}{1+\left|re^{i\phi}\right|^2}\right)\exp\left(-\rho\frac{\left|re^{i\phi}-r_0e^{i\phi}\right|^2}{1+\left|re^{i\phi}\right|^2}\right) \\ &= \frac{1}{\pi\left(1+|r|^2\right)^2}\left(1+\rho\frac{\left|1+rr_0^*\right|^2}{1+|r|^2}\right)\exp\left(-\rho\frac{|r-r_0|^2}{1+|r|^2}\right) = g\left(r, r_0\right) \end{aligned} \tag{3.8}$$

which implies that

$$g\left(r, r_0\right) = g\left(re^{-i\arg(r_0)}, r_0e^{-i\arg(r_0)}\right) = g\left(re^{-i\arg(r_0)}, |r_0|\right) \tag{3.9}$$

This is a very useful property, as it allows us to study the PDF with $r_0 \in \mathbf{R}$, and all properties will apply no matter how r_0 is rotated. Therefore, when $r_0 \in \mathbf{R}$, we can look at the real and imaginary arguments to the PDF as the axes along and about r_0 in the general case when $r_0 \in \mathbf{C}$.

Symmetry about the noise-free NRV

Let $r_0 = x_0 \in \mathbf{R}$. Then,

$$f(r^*) = \frac{1}{\pi\left(1+|r^*|^2\right)^2}\left(1+\rho\frac{\left|1+r^*r_0^*\right|^2}{1+|r^*|^2}\right)\exp\left(-\rho\frac{|r^*-r_0|^2}{1+|r^*|^2}\right) = f(r) \tag{3.10}$$

since

$$|r^*| = |r| \text{ and } r_0^* = x_0^* = x_0 = r_0 \tag{3.11}$$

This means that when $r_0 \in \mathbf{R}$ the PDF is symmetric around the real axis. Combining this with the previous result (3.9), we can see that the PDF is symmetric about $r_0 \in \mathbf{C}$:

$$\begin{gathered} g\left(re^{i\phi}, r_0e^{-i\arg(r_0)}\right) \overset{(3.10)}{=} g\left(re^{-i\phi}, r_0e^{-i\arg(r_0)}\right) \overset{(3.9)}{\Rightarrow} \\ g\left(re^{i\arg(r_0)}e^{i\phi}, r_0\right) = g\left(re^{i\arg(r_0)}e^{-i\phi}, r_0\right) \end{gathered} \tag{3.12}$$

Dependence of the real and imaginary variable

Let $r = x + iy$. The marginal distributions for the simplest possible case when $\rho = 0$ are [5]:

$$f_x(x) = \frac{1}{2\left(1+x^2\right)^{3/2}} \quad \text{and} \quad f_y(y) = \frac{1}{2\left(1+y^2\right)^{3/2}} \tag{3.13}$$

$$f_x(x)\cdot f_y(y) = \frac{1}{4\left[(1+x^2)(1+y^2)\right]^{3/2}} \neq \frac{1}{\pi\left(1+x^2+y^2\right)^2} = f_{x,y}(x,y) \tag{3.14}$$

which shows that in general, $\mathrm{Re}\{r\}$ and $\mathrm{Im}\{r\}$ are dependent. Due to the previous results, this means that the distributions about and along r_0 are generally not independent. This has also been verified numerically for $(\rho, r_0) \neq 0$.

3.2.3 Visualized example

Before diving deeper into the properties of the PDF, it is a good idea to get a feeling for how it looks first. Again, let $r_0 = x_0 \in \mathbf{R}$. The left part in figure 3.1 shows what the PDF looks like for a reasonable value of x_0 and SNR ρ. The expectation value (mean) is close to $x_0 = 1$, but the mode is not (in the real variable x). Therefore it is easy to see that the distribution is noticeably skewed in x. The figure also shows the marginal distributions for the real and imaginary variables. We shall see later on that the marginal distributions are useful to work with, especially the marginal distribution in the real variable.

The right part of the figure shows the PDF for an extreme case where the SNR ρ is very low and r_0 is very large. This case clearly shows that the distributions about and along r_0 are dependent. Other than that, as we shall also see later on, there is little point in further studying this case.

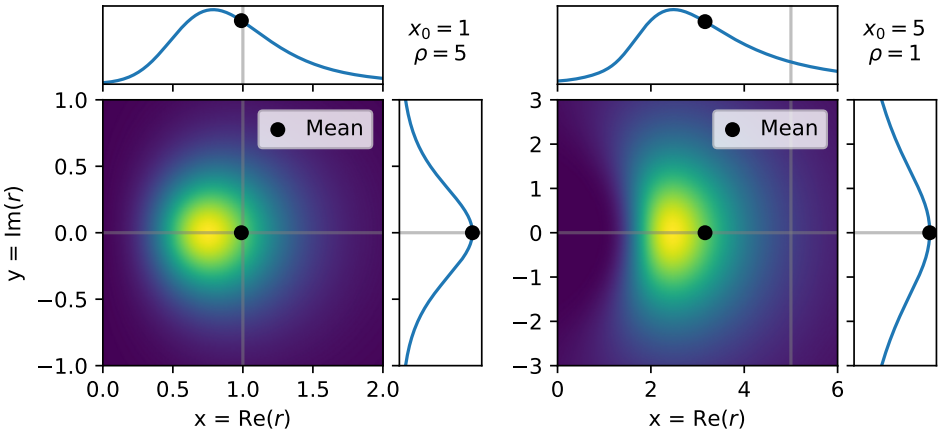

Figure 3.1: The PDF of the NRV with its marginal distributions visualized for two different sets of parameters r_0 and ρ.

3.2.4 A note

In the report by Chassande-Mottin et al. [5], they pointed out that $f(r)$ is approximately Gaussian for small r_0. That is, the $\exp(\cdot)$ in $f(r)$ is the dominating factor. Later in this chapter and also in the following chapters, we will make good use of this property. Note that this only holds for reasonable values of local SNR ρ, as we shall soon see. Also, loosely speaking, the distribution approaches another shape for large r_0, which can be seen later in for example figure 3.6, 3.7, and 3.9. It is hard to give a good definition of what a small or large r_0 is, but it will show that sometimes, the small r_0 approximation is quite useful.

3.3 Statistical properties

3.3.1 Expectation value

As Chassande-Mottin et al. [5] also has pointed out, it is rather easy to see that $\rho = 0 \Rightarrow \mathrm{E}[r] = 0$ and also that, per definition, as $\rho \to \infty$, $\mathrm{E}[r] \to r_0$. It follows that when $r_0 = 0 \Rightarrow \mathrm{E}[r] = 0 \; \forall \; \rho$. Furthermore, due to the PDF being symmetric about r_0, it must also follow that for $r_0 \neq 0$,

$$\arg \mathrm{E}[r] = \arg r_0 \tag{3.15}$$

However, finding a closed expression for the expectation value analytically is seemingly impossible, or at least *very* hard. By numerical integration, Brynolfsson [11] has demonstrated how the expectation value of r varies as a function of SNR ρ, given a single value of r_0. Inspired by this, we shall now see how $\mathrm{E}[r]$ behaves for arbitrary values of r_0. But

first, let μ_n be the *normalized* expectation value of r for $r_0 \neq 0$:

$$\mu_n = \mathrm{E}[r] \,/\, r_0 = \mathrm{E}[r/r_0] \tag{3.16}$$

Given (3.15), it must follow that $\mu_n \in \mathbf{R}$. A batch of different $r_0 \in \mathbf{C}$ was randomly selected, $r_0 \in \mathbf{C}$, $0 < |r_0| < 10$. For these arbitrary r_0, the normalized expectation value μ_n was calculated over a wide range of SNRs, $10^{-3} \leq \rho \leq 10^3$. The results show that μ_n is indeed strictly real (within a reasonable margin of integration error). More interestingly, it showed that μ_n does not depend on r_0, and therefore can be seen as a function of only SNR ρ. This is a very nice property, as it allows us to draw some conclusions about the distribution that holds for all r_0. The relationship $\mu_n(\rho)$ can be seen in figure 3.2.

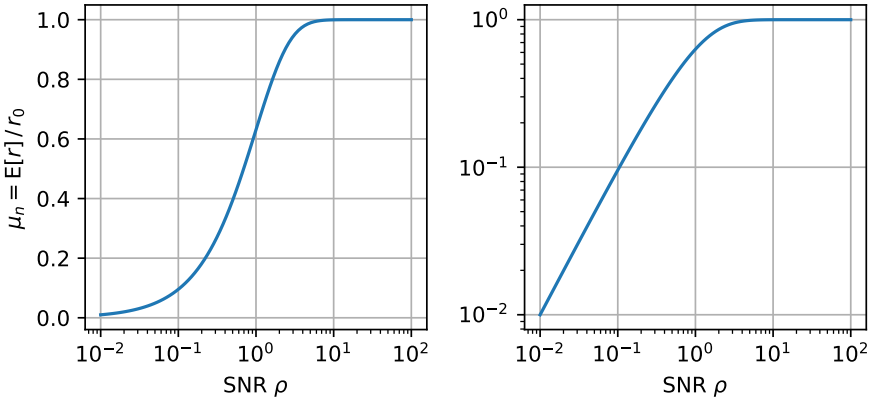

Figure 3.2: Normalized expectation value $\mu_n(\rho)$ of the NRV over a wide range of SNRs ρ. The difference between the plots is the scaling of the y-axis.

One can argue that the reassignment is only "useful" when $\mathrm{E}[r] \approx r_0$, i.e., when $\mu_n \approx 1$, otherwise it will be biased. Following this line of thought when studying $\mu_n(\rho)$, it can be seen that some minimum SNR ρ could be defined from which the reassignment is "useful". Up to at least $\rho = 1$, μ_n is heavily biased. Thus, studying the case when $\rho < 1$ is not relevant for the purpose of this thesis. Table 3.1 shows at what SNR ρ the normalized expectation value μ_n reaches a certain value. For $\rho \approx 5$, there is a $\sim 1\%$ bias, and for $\rho \approx 10$ the bias is less than 0.1%. Given those results, a rule of thumb is suggested which states that $\rho \gtrsim 10 = 10\,\mathrm{dB}$ gives a practically unbiased reassignment vector for most applications. This simple rule is easy to keep in mind when doing further analysis.

3.3.2 Variance

Chassande-Mottin et al. [5] states that as

$$\rho \to 0^+ \Rightarrow \mathrm{V}[r] \to \infty \tag{3.17}$$

μ_n	0.9	0.99	0.999	0.999999
ρ	2.38	5.09	8.36	22.4
	3.77 dB	7.07 dB	9.22 dB	13.51 dB

Table 3.1: The inverse relationship of $\mu_n(\rho)$ and ρ for some values of μ_n. This was solved numerically using a root-finding algorithm.

and again that per definition as

$$\rho \to \infty \Rightarrow \mathrm{V}[r] \to 0 \tag{3.18}$$

In a similar fashion as for the expectation value, the variance of the NRV is explored for a range of reasonable SNR ρ and r_0. This was again done by numerical evaluation. Since the complex variance is the sum of the variance of the real and imaginary parts, it also makes sense to also look at the two separately. The relationship between variance and ρ, can be seen in figure 3.3.

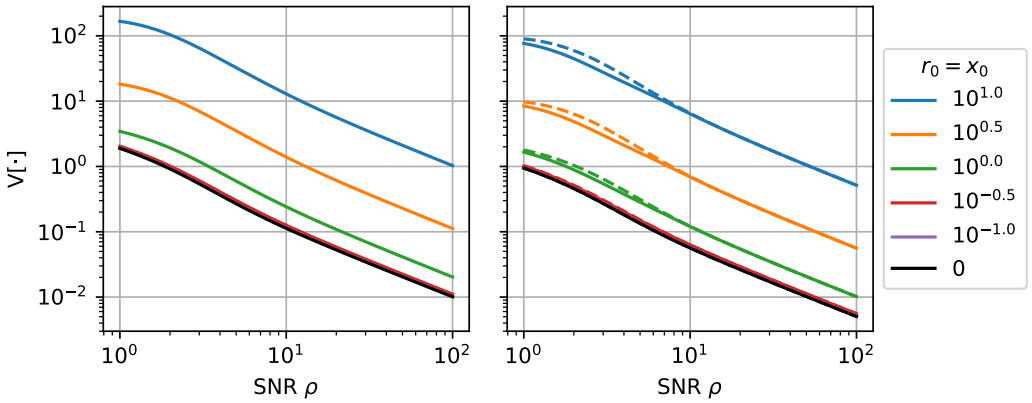

Figure 3.3: Variance of the NRV r as a function of SNR ρ. The plot to the left shows the combined complex variance, and the plot to the right shows the real (solid) and imaginary (dashed) parts separated.

Studying the curves in the figure, for reasonably large SNR ρ ($\gtrsim 10$),

$$\mathrm{V}[r] \approx 2\,\mathrm{V}[\mathrm{Re}\{r\}] \approx 2\,\mathrm{V}[\mathrm{Im}\{r\}] \approx \kappa(|r_0|) \cdot \rho^{-1} \tag{3.19}$$

where $\kappa(|r_0|)$ is some unknown function, seemingly growing monotonically with $|r_0|$. This function explains the relationship between $\mathrm{V}[r]$ and $|r_0|$. To explore what $\kappa(|r_0|)$ could be, a similar evaluation was made with a fixed large SNR ρ, instead letting r_0 vary over a reasonable range. The results for $\rho = 10 = 10\,\mathrm{dB}$ and $\rho = 100 = 20\,\mathrm{dB}$ can be seen in figure 3.4.

Studying the figure, we can see that the relationship g for *large* ρ is approximately

$$\kappa(|r_0|) \approx 1 + |r_0|^2 \tag{3.20}$$

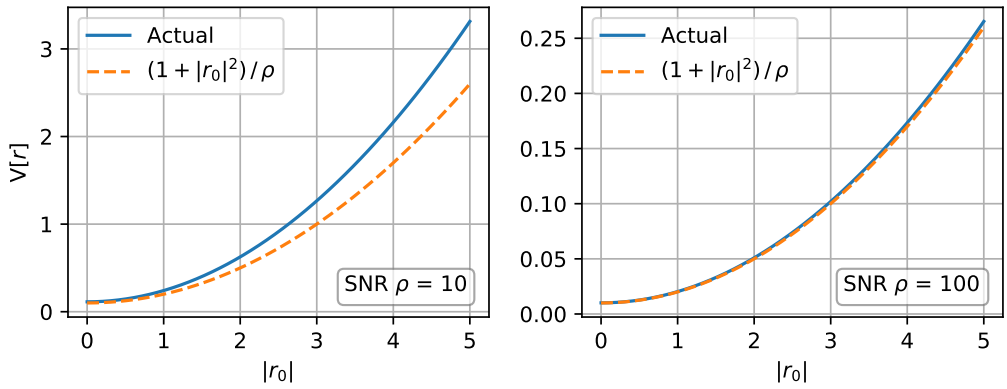

Figure 3.4: Variance of the NRV as a function of $|r_0|$ for a fixed SNR ρ. The actual variance (solid blue line) is compared to an approximation (orange dashed line).

Putting (3.19) and (3.20) together,

$$\mathrm{V}[r] \approx 2\,\mathrm{V}[\mathrm{Re}\{r\}] \approx 2\,\mathrm{V}[\mathrm{Im}\{r\}] \approx \frac{1 + |r_0|^2}{\rho} \triangleq \hat{\mathrm{V}}[r] \tag{3.21}$$

While this approximation is quite loose, it will show itself useful when discussing how the PDF can be approximated in section 3.4.

3.3.3 Mode and Skewness

As previously seen in figure 3.1, the distribution $f(r)$ can be noticeably skewed, with the mode clearly offset from the mean. In this section, we will explore how the mode and skewness depends on r_0 and SNR ρ. As done previously, let $r_0 = x_0 \in \mathbf{R}$, utilizing the symmetric properties of the PDF. We will start by analyzing the mode of the distribution in the same way as we did for the expectation value. When doing so, the *normalized* expectation value was defined as $\mu_n = \mathrm{E}_n[r] = \mathrm{E}[r]/r_0$. In the same way, we now define the *normalized* mode as $\mathrm{Mode}_n[r] = \mathrm{Mode}[r]/r_0$. However, now we have to be careful as $\mathrm{Mode}[\mathrm{Re}\{r\}] \neq \mathrm{Re}\{\mathrm{Mode}[r]\}$ due to the dependence of the real and imaginary variable. Therefore, it makes sense both to look at the mode for the joint distribution, and for the marginal distribution.

Figure 3.5 shows how the mode of the joint and marginal distribution depends on $r_0 = x_0$ and SNR ρ. The lines have different values of x_0, sweeping a wide range of reasonable x_0. The same range of x_0 is swept both for the joint and marginal distribution. It is clear that the normalized mode does not have the same nice property that the normalized mean has in that it does not depend on r_0. Furthermore, the mode of the marginal distribution differs slightly from that of the joint distribution. This is expected as the

real and imaginary variables are generally dependent. However, they do not differ significantly, which means that the properties of marginal distribution largely applies to the joint distribution as well.

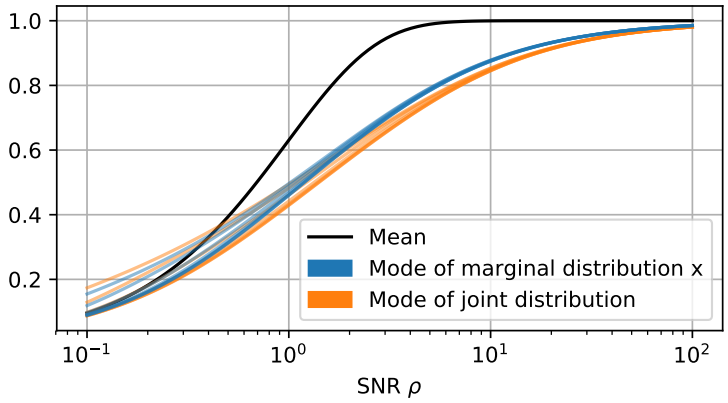

Figure 3.5: Normalized mode of the joint PDF and the marginal distribution of x compared to normalized expectation value μ_n (mean). The different lines show a wide range of values of x_0.

The main takeaway from studying the figure is that the mode converges to r_0 much slower than the mean does. Since the normalized mode is *roughly* the same for different r_0, the distance between mode and mean grows almost linearly with r_0. This means that when r_0 is "large enough", the distribution will be considerably skewed for even moderately large SNR ρ. We shall soon investigate what "large enough" means, but first we need to consider the consequences of a skewed distribution.

If the distribution of r is skewed, it is per definition asymmetric. In this case, it is skewed towards zero, which means that sampling a value less than the mean is more likely than a value greater than the mean. This is why we do not really care that the mean is correct if the distribution is skewed, because then the samples will not spread nicely around the mean anyway. Instead, in the extremely skewed case, most values will be somewhat smaller than the main, while a few samples will be much greater than the main.

For a bivariate/complex distribution like this one, defining skewness is not as straightforward as for the expectation value or variance. However, since the distribution (when $r_0 = x_0 \in \mathbf{R}$) is symmetric around the real axis, it is per definition not skewed in the imaginary variable y. This makes things much more simple, as we only have to study the skewness for the real variable x. As such, we can define skewness as the third standardized moment of the marginal distribution of x:

$$\text{Skew}[x] = (\text{D}[x])^{-3} \cdot \iint (x - \mu_x)^3 f(x + iy)\,\mathrm{d}x\,\mathrm{d}y \tag{3.22}$$

Another way to define skewness is the Pearson mode skewness coefficient [12]:

$$\text{ModeSkew}[x] = \frac{\text{E}[x] - \text{Mode}[x]}{\text{D}[x]} \tag{3.23}$$

Figure 3.6 and 3.7 shows the central moment skewness and the Pearson mode skewness respectively. The two measures differ somewhat in scale, but both tell the same overall story. Interestingly, and perhaps not too unexpectedly, x_0 has a major impact on the skewness. For small x_0, the distribution is barely skewed for any reasonable SNR ρ. In contrast, for large x_0 the distribution is highly skewed up to fairly large SNR ρ.

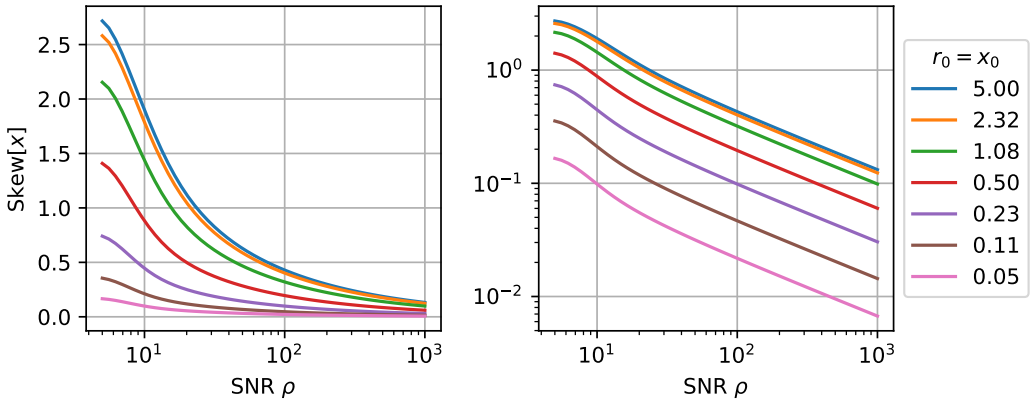

Figure 3.6: Skewness of the marginal distribution in the real variable $f_x(r)$.

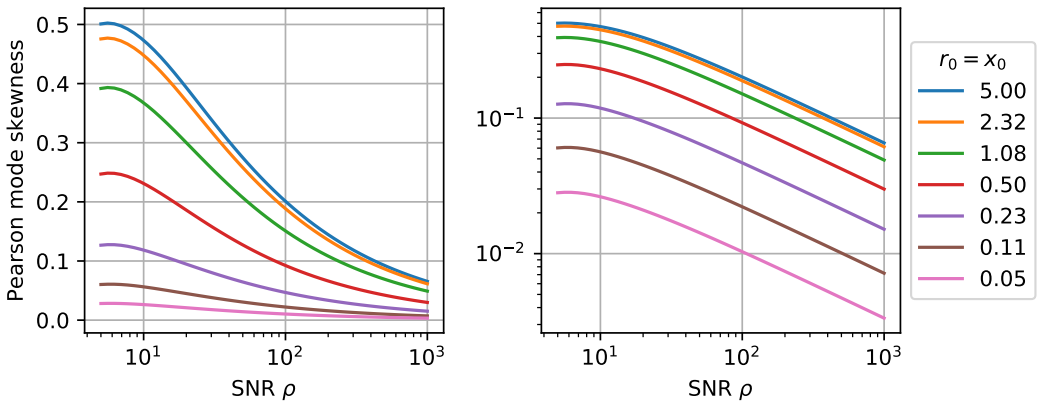

Figure 3.7: Pearson mode skewness of the marginal distribution in the real variable $f_x(r)$.

3.4 Approximative distribution

It was previously noted that for small r_0 and large ρ, the $\exp(\cdot)$ factor in the distribution (3.3) will dominate, and as such the distribution will be approximately Gaussian. We have learned that for $\rho \gtrsim 10$, the expectation value of r is approximately r_0. Furthermore, we have also seen that the variance can be approximately expressed as (3.21):

$$\mathrm{V}[r] \approx \hat{\mathrm{V}}[r] = \frac{1 + |r_0|^2}{\rho} \tag{3.24}$$

Putting the pieces together,

$$r \dot{\sim} \mathcal{CN}\left(r_0, \frac{1 + |r_0|^2}{\rho}\right) \tag{3.25}$$

From this a Gaussian approximation $\hat{f}$ of the PDF f is proposed, normalized such that $\int_{\mathbb{C}} \hat{f} = 1$:

$$f(r) \approx \hat{f}(r) \triangleq \frac{\rho}{\pi\left(1 + |r_0|^2\right)} \exp\left(-\rho \frac{|r - r_0|^2}{1 + |r_0|^2}\right) \tag{3.26}$$

Note the implied independence of $\mathrm{Re}\{r\}$ and $\mathrm{Im}\{r\}$.

So how good is this approximation? Naturally, this is a very hard question to answer, and we will not dive very deep into this issue. Holding on to the rule of thumb of $\rho \gtrsim 10$, we at least know that the expectation value is essentially unbiased. Let us also have a look at how good the variance approximation is. Figure 3.8 shows the approximative standard deviation relative to the actual value. For $\rho \gtrsim 10$, the error is less than $\sim$15% for any $|r_0|$, and for slightly larger ρ the upper bound of the error quickly shrinks below $\sim$5%. Based on this, we suggest that as $\rho \to \infty$, $\hat{\mathrm{D}}[r] \to \mathrm{D}[r]$.

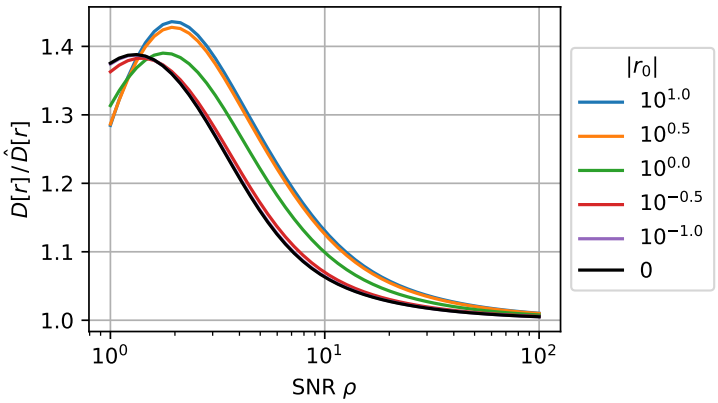

Figure 3.8: Comparison of the approximative variance relative to the actual variance. Note that the plot shows standard deviation.

Another thing one could do is to look at the *statistical distance* between the two distributions. More specifically the f-divergence, which measures the distance between two probability density functions. Liese and Vajda has a good paper from 2006 discussing f-divergences [13]. One common measure is the Kullback-Leibler divergence, proposed in their paper from 1951 [14]. Another common measure is the Hellinger distance, proposed by Hellinger back in 1909 [15].

Kullback-Leibler divergence:

$$D_{KL} \triangleq \int f_1(x) \log\left(\frac{f_1(x)}{f_2(x)}\right) \mathrm{d}x + \int f_2(x) \log\left(\frac{f_2(x)}{f_1(x)}\right) \mathrm{d}x \tag{3.27}$$

Hellinger distance:

$$D_H^2 \triangleq 1 - \int \sqrt{f_1(x) f_2(x)}\, \mathrm{d}x \tag{3.28}$$

These measures can easily be extended to allow for PDF:s in the complex domain ($f : \mathbf{C} \to \mathbf{R}$) – just integrate over $\mathbf{C}$ instead.

Figure 3.9 shows the two measures comparing f and $\hat{f}$. While it is hard to relate to the measures in an absolute sense, it is clear that the divergence goes towards zero as ρ grows large. It can also be seen that $|r_0|$ has a big effect on the divergence measure. For large $|r_0|$, the divergence seems to go against some upper limit, although still converging to zero as ρ grows large.

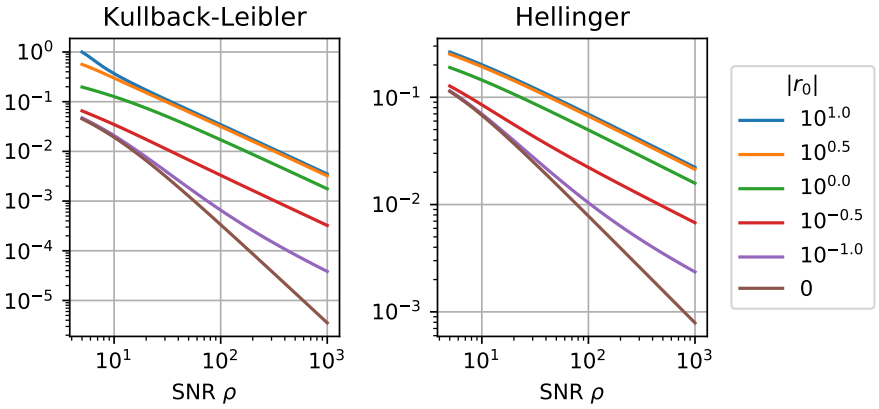

Figure 3.9: Kullback-Leibler divergence and Hellinger distance between f and $\hat{f}$.

Chapter 4

Statistics of the reassigned Gabor spectrogram

In the previous chapter, we explored the statistics of the NRV for Gabor spectrograms. We looked at the distribution of the NRV at an arbitrary point in an arbitrary Gabor spectrogram for an arbitrary signal subjected to AWGN. As such, the conclusions drawn were general but did not really paint the picture of how the reassigned spectrogram behaves as a whole given some actual signal. In this chapter, we shall use our knowledge about some basic common signals presented in chapter 2 to draw further conclusions about the reassigned spectrogram.

4.1 Approach

The question we are going to try to answer in this chapter is rather simple: what is the distribution of the reassigned Gabor spectrogram for some known signals subjected to AWGN? The signal model is assumed to be the same as in the previous chapter (see section 3.1). The signals to be evaluated are the ones examined in chapter 2 – the impulse, sinusoid, linear chirp, and the matched Gaussian.

First, we will focus on the impulse, sinusoid, and linear chirp. It will show that these three signals behave in the same way in the normalized spectrogram, which is presented and discussed in the next section. Because of that, it is possible to generalize and draw conclusions about the three simultaneously. These conclusions will include an approximation of the reassigned spectrogram distribution, which enables us to compare it to the original spectrogram.

After that, the matched Gaussian case is studied. It will show that this signal behaves similarly to the previous signals, but in two dimensions. However, due to the symmetric properties of the normalized spectrogram and reassignment vector, the resulting distribution is circularly symmetric. This allows us to once again examine the results in only one variable. Finally, the theory is tested against some simulations, both for the sinusoid and matched Gaussian case.

4.2 The normalized spectrogram

As seen in the previous chapter, it makes a lot of sense to look at the normalized version the reassignment vector. However, to make practical use of the known properties of the NRV, it has to be denormalized again. Optionally, it is possible to look at a normalized version of the spectrogram instead. In that context, the scaling of the NRV matches that of the spectrogram. We will simply name the normalized version of the spectrogram as the *normalized spectrogram*, or NS for short.

The NRV is normalized with the Gaussian window duration Δt and bandwidth $\Delta\nu$ (2.12). Correspondingly, let the normalized time and frequency be defined as $t_n = t/\Delta t$ and $\nu_n = \nu/\Delta\nu$ respectively. The normalized spectrogram is defined as the spectrogram S expressed as a function of the normalized variables. For example, taking the Gabor spectrogram of the impulse signal (B.1) centered in $t_0 = 0$:

$$\begin{aligned} S(t,\nu) &\overset{(B.4)}{=} \frac{1}{\sqrt{\pi}\lambda}\exp\left[-\left(\frac{t}{\lambda}\right)^2\right] \overset{(2.12)}{=} \frac{1}{\sqrt{\pi}\lambda}\exp\left[-\frac{1}{2}\left(\frac{t}{\Delta t}\right)^2\right] \Rightarrow \\ NS(t_n,\nu_n) &= \frac{1}{\sqrt{\pi}\lambda}\exp\left(-\frac{t_n^2}{2}\right) \end{aligned} \tag{4.1}$$

Note that here, $|t_n|$ is the normalized distance from the signal in the normalized spectrogram. We define this general normalized distance from the signal as d_n. Since the noise-free NRV r_0 is perpendicular to the signal in the normalized spectrogram, $|r_0|^2 = d_n^2$. This is also shown by equation (B.8). The main takeaway is that

$$NS \propto \exp\left(-\frac{d_n^2}{2}\right) = \exp\left(-\frac{|r_0|^2}{2}\right) \tag{4.2}$$

For a sinusoid (B.10) centered in $\nu_0 = 0$, where $d_n = |\nu_n|$, the normalized spectrogram becomes (as shown in appendix B.2)

$$NS_s^h(t_n,\nu_n) = \frac{1}{\sqrt{2\pi}\Delta\nu}\exp\left(-\frac{\nu_n^2}{2}\right) = \frac{1}{\sqrt{2\pi}\Delta\nu}\exp\left(-\frac{d_n^2}{2}\right) = \frac{1}{\sqrt{2\pi}\Delta\nu}\exp\left(-\frac{|r_0|^2}{2}\right) \tag{4.3}$$

So for both the impulse and the sinusoid, the normalized spectrogram can be written as (4.2). In fact, this is also the case for a linear chirp, which will now be demonstrated. The linear chirp previously discussed (B.19) is defined such that the instantaneous frequency $\nu = \beta t$. In the normalized spectrogram, given that $t = \Delta t \cdot t_n$ and $\nu = \Delta\nu \cdot \nu_n$, this corresponds to

$$\Delta\nu \cdot \nu_n = \beta \cdot (\Delta t \cdot t_n) \Rightarrow \nu_n = \frac{\Delta t}{\Delta\nu} \cdot \beta \cdot t_n \overset{(2.12)}{=} 2\pi\lambda^2\beta t_n \tag{4.4}$$

The squared distance from the line to a point $(\check{t}_n, \check{\nu}_n)$ in the NS is [16]

$$d_n^2 = \frac{(\check{\nu}_n - 2\pi\lambda^2\beta\check{t}_n)^2}{1 + (2\pi\lambda^2\beta)^2} \tag{4.5}$$

and the normalized spectrogram of the chirp is (B.25):

$$NS_s^h(t_n, \nu_n) = \frac{2\sqrt{\pi}\lambda}{\sqrt{1+(2\pi\lambda^2\beta)^2}} \exp\left(-\frac{1}{2} \cdot \frac{(\nu_n - 2\pi\lambda^2\beta t_n)^2}{1+(2\pi\lambda^2\beta)^2}\right) \tag{4.6}$$

Substitution with d_n^2 (4.5) in (4.6) again gives (4.2). With this it is also easy to see that the NRV is perpendicular to the signal in the normalized spectrogram, see appendix B.3. The left plot of figure 4.1 illustrates what this looks like. As for the NRV, the normalized spectrogram can also be written as a function of a complex number; $NS(z) = NS(\mathrm{Re}\{z\}, \mathrm{Im}\{z\})$.

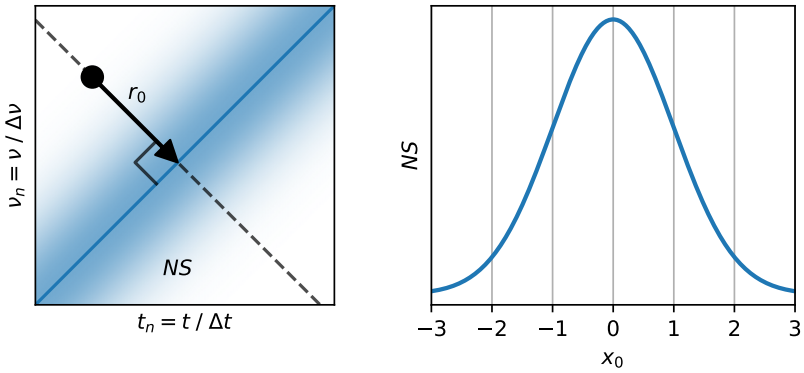

Figure 4.1: An illustration of that r_0 always is perpendicular to the signal in the normalized spectrogram. The left plot shows the normalized spectrogram, and the right part shows a slice of the normalized spectrogram perpendicular to the signal (along r_0).

Remember the rotational invariance property of the NRV (see section 3.2.2). With that, we are allowed to look at a rotated version of the normalized spectrogram such that the signal becomes a vertical line, like the impulse. By doing so, r_0 becomes horizontal – perpendicular to the vertical line that the signal causes. In other words, we rotated the normalized spectrogram such that $r_0 \in \mathbf{R}$, and as such we can let $r_0 = x_0$. The reason to why this is done is the same as when the properties of the NRV was studied – it allows us to use the real and imaginary axes as the axes along and about r_0 respectively.

Now, take a slice along the axis of r_0, which in the rotated normalized spectrogram corresponds to the real axis. No matter where this slice is taken, when expressed as a function of x_0, it will always be the same. The result can be seen in figure 4.1, where the slice taken along the dashed line in the left plot is viewed in the right plot. To be clear, this holds for all the three signals we are now working with – the impulse, sinusoid, and linear chirp.

4.3 Reassigning the normalized spectrogram

We are now working with a slice from the normalized spectrogram taken perpendicular to the signal, as discussed in the previous section. Given some thought, this slice can be seen as the marginal distribution in x of the rotated normalized spectrogram. Again, for the normalized Gabor spectrogram, this slice will always look the same in terms of the NRV x_0, see equation (4.2).

If the slice is seen as a marginal distribution, it makes sense to define a corresponding PDF of how the normalized spectrogram NS is distributed in x:

$$f_x^{NS}(x) \triangleq \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \tag{4.7}$$

The corresponding PDF for the normalized reassigned spectrogram (NRS) is defined as f_x^{NRS}. In the noise-free case the reassignment has perfect localization, and as such, per definition $f_x^{NRS}(x) = \delta(x)$. But of course, as thoroughly discussed in the previous chapter, this is not the case when the signal is subjected to AWGN.

The reassigned distribution f_x^{NRS} is literally a redistributed version of the original distribution f_x^{NS}. Every point x_0 in the original distribution gets redistributed according to the distribution of the NRV (3.3), which in this chapter is written as f^r instead of just f. Figure 4.2 illustrates how points in the original spectrogram distribution (dashed) are redistributed to form a new reassigned distribution. Taking care of the direction of the NRV, some point x in the reassigned distribution can be expressed as:

$$f_x^{NRS}(x) = \int f_x^{NS}(x_0) \cdot f_x^r(x_0 - x)\, \mathrm{d}x_0 \tag{4.8}$$

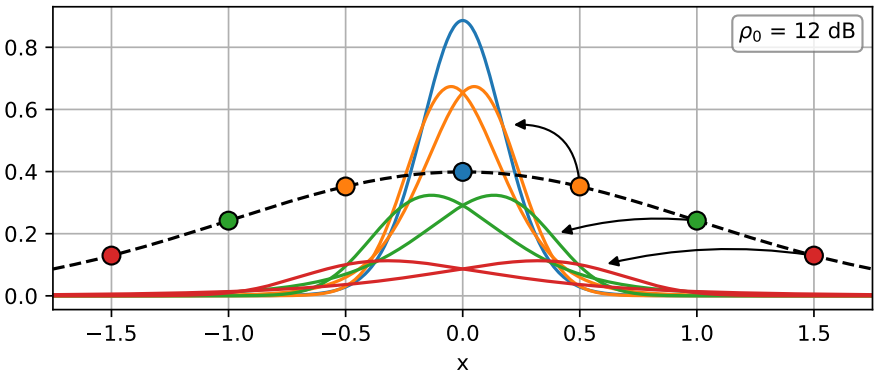

Figure 4.2: An illustration of how the normalized Gabor spectrogram distribution (dashed black line) is redistributed.

The final piece of the puzzle is in the local SNR ρ. Since it is proportional to the spectrogram, it too can be written as a function of x:

$$\rho(x) = \rho_0 \cdot \exp\left(-\frac{x^2}{2}\right), \ \rho_0 = \frac{NS(0)}{2\sigma_n^2} \tag{4.9}$$

where ρ_0 is defined as the peak SNR. This is incorporated into figure 4.2, showing the case when $\rho_0 = 12\,\text{dB}$.

To put the pieces together, let $r = x + iy$ and $r_0 = x_0$. Again, x is a point in the reassigned spectrogram NRS, and x_0 is a point in the original spectrogram NS. Note that f_x^{NS} (4.7) and ρ (4.9) become functions of x_0. Plugging $\rho(x_0)$ (4.9) into f^r (3.3), the marginal distribution becomes:

$$\begin{aligned} f_x^r(x) &\overset{(3.3)}{=} \int \frac{1}{\pi(1+x^2+y^2)^2}\left[1+\rho\frac{(1+xx_0)^2+(yx_0)^2}{1+x^2+y^2}\right]\exp\left(-\rho\frac{(x-x_0)^2+y^2}{1+x^2+y^2}\right)\mathrm{d}y \\ &\overset{(4.9)}{=} \int \frac{1}{\pi(1+x^2+y^2)^2}\left[1+\rho_0\cdot\exp\left(-\frac{x_0^2}{2}\right)\cdot\frac{(1+xx_0)^2+(yx_0)^2}{1+x^2+y^2}\right] \\ &\qquad \exp\left[-\rho_0\cdot\exp\left(-\frac{x_0^2}{2}\right)\cdot\frac{(x-x_0)^2+y^2}{1+x^2+y^2}\right]\mathrm{d}y \end{aligned} \tag{4.10}$$

This is then, together with f_x^{NS} (4.7), plugged into the expression for f_x^{NRS} (4.8), which makes for a quite cumbersome double integral in y and x_0.

Instead of f^r we can also use the approximative NRV distribution $\hat{f}^r$ (3.26) presented in the previous chapter. In that case the resulting reassigned distribution is defined as $\hat{f}_x^{NRS}$. Figure 4.3 shows what the redistribution looks like using the approximate NRV distribution.

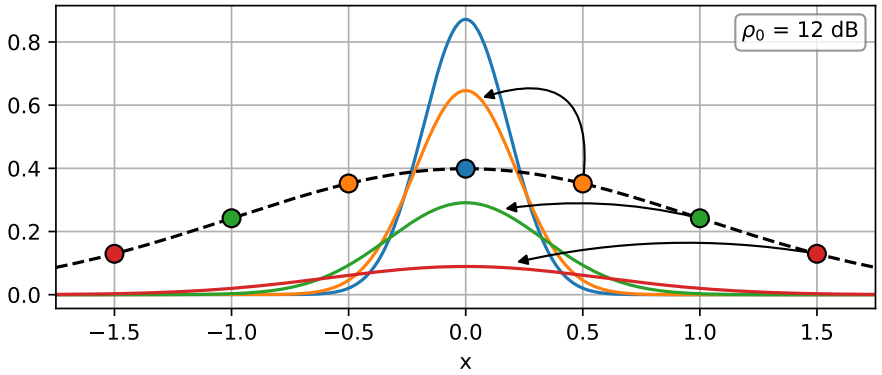

Figure 4.3: An illustration of how the normalized Gabor spectrogram distribution is redistributed with the approximative NRV distribution.

With the approximative NRV distribution $\hat{f}^r$, the expression for the marginal distribution $\hat{f}_x^r$ becomes somewhat nicer:

$$\begin{aligned} \hat{f}_x^r(x) &\overset{(3.26)}{=} \sqrt{\frac{\rho}{\pi\,(1+x_0^2)}}\exp\left[-\rho\frac{(x-x_0)^2}{1+x_0^2}\right] \\ &\overset{(4.9)}{=} \sqrt{\frac{\rho_0}{\pi\,(1+x_0^2)}}\exp\left[-\frac{x_0^2}{4}\right]\exp\left[-\rho_0\cdot\exp\left(-\frac{x_0^2}{2}\right)\cdot\frac{(x-x_0)^2}{1+x_0^2}\right] \end{aligned} \tag{4.11}$$

Still, when plugged into plugged into the expression for f_x^{NRS} (4.8), forming $\hat{f}_x^{NRS}$, that integral also becomes rather cumbersome. However, the resulting integral is only in one variable instead of two, which makes for much faster numerical evaluation.

Figure 4.4 shows the true and approximative marginal distribution f_x^{NRS} and $\hat{f}_x^{NRS}$ respectively evaluated for two different peak SNRs ρ_0. It turns out that for $\rho_0 \gtrsim 10\,\mathrm{dB}$, the approximative distribution matches the true distribution almost perfectly. Even for as low as $\rho_0 \approx 0\,\mathrm{dB}$, the distributions still match reasonably well. Note that neither f_x^{NRS} nor $\hat{f}_x^{NRS}$ are Gaussian distributions.

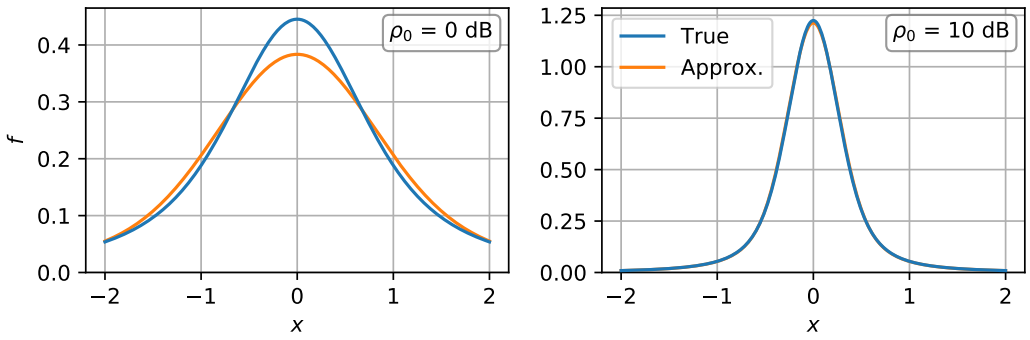

Figure 4.4: The marginal distribution of the NRS f_x^{NRS} obtained from the true and approximative NRV distributions for two different peak SNRs ρ_0.

4.4 Analysis and approximation of the reassigned spectrogram

The fact that the approximative distribution $\hat{f}_x^{NRS}$ is practically identical to the true distribution f_x^{NRS} comes with a nice implication. Basically, it tells us that the distribution can be seen as a weighted sum of zero-mean Gaussians. Studying these Gaussians, see (4.11), it can be seen that their standard deviations are inversely proportional to $\sqrt{\rho_0}$. Consequently, the NRS distribution will also scale inversely with $\sqrt{\rho_0}$.

It would make sense if it followed that the standard deviation $\mathrm{D}_x^{NRS}[x] \propto 1/\sqrt{\rho_0}$, but this is not the case. When numerically evaluating the variance, the calculations do not converge. From this, we draw the conclusion that the distribution does not have finite variance. This is explainable with the fact that the variance of the NRV distribution grows infinite as the SNR ρ shrinks to zero.

While it is interesting that the resulting distribution has infinite variance, it is not a very useful result. Perhaps a better measure of spread can be found by relating the shape of the distribution to some other distribution? Comparisons to other common distributions has shown that the NRS distribution closely resembles a t-distribution with two degrees of freedom. A comparison between the true and matching t-distribution can be seen in figure 4.5.

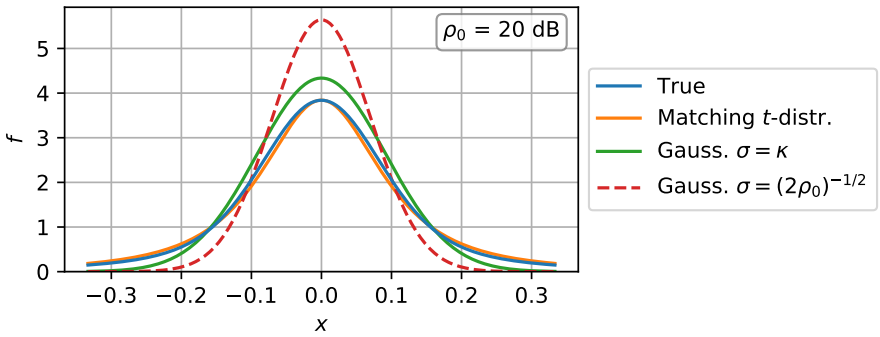

Figure 4.5: A comparison of distributions in relation to the true marginal distribution of the NRS.

The t-distribution with two degrees of freedom has infinite variance, which matches that of the true distribution. Let us define this t-distribution approximation as $\dot{f}_x^{NRS}$. Its PDF can be written as

$$\dot{f}_x^{NRS}(x) \triangleq \frac{1}{2\sqrt{2}\,\kappa_x}\left[1 + \frac{1}{2}\left(\frac{x}{\kappa_x}\right)^2\right]^{-3/2} \tag{4.12}$$

where κ_x is some scaling factor. Given that the NRS distribution $\hat{f}_x^{NRS}$ scales inversely with $\sqrt{\rho_0}$, it must follow that $\kappa_x \propto 1/\sqrt{\rho_0}$. Fitting $\dot{f}_x^{NRS}$ to the true distribution gave that $\kappa_x \approx 0.9/\sqrt{\rho_0}$.

For a t-distribution, as the degrees of freedom grows infinite, it approaches the Gaussian distribution. By letting the number of degrees of freedom grow infinite in our approximative t-distribution $\dot{f}_x^{NRS}$, a new Gaussian distribution $\ddot{f}_x^{NRS}$ is formed. Its standard deviation $\ddot{\mathrm{D}}_x^{NRS}[x] = \kappa_x$. Figure 4.5 shows this distribution (green) compared to the true (blue) and matching t-distribution (orange). This approximation might be a bit on the optimistic side, so rounding the scale to $\ddot{\mathrm{D}}_x^{NRS}[x] \approx 1/\sqrt{\rho_0}$ makes for an acceptable and simple rule of thumb.

By having this Gaussian approximation, we are now able to compare it against the original normalized spectrogram, which has unit variance. The *reassignment localization factor* γ is suggested to be defined as the scaling ratio between the two. Here,

$$\gamma_x \triangleq \frac{\mathrm{D}_x^{NS}}{\ddot{\mathrm{D}}_x^{NRS}} = \frac{1}{\kappa_x} \approx \frac{\sqrt{\rho_0}}{0.9} \approx \sqrt{\rho_0} \tag{4.13}$$

In the next chapter, one potential use of this ratio will be shown.

As a final thought, let us revisit figure 4.2 and 4.3. It can be seen that the heavy tails of the NRS distribution mainly comes from the redistribution of far out points in the original spectrogram. In practice, these points could likely be ignored to some extent. Taking this to the extreme, ignoring all points but the very center, we end up with a single NRV distribution (with $r_0 = 0$). This distribution is approximately Gaussian with $\mathrm{D}_x \approx 1/\sqrt{2\rho_0}$. Think of this as the very best case scenario, although unattainable. Figure 4.5 shows this distribution (dashed red) compared to the previously discussed distributions.

Let us recap the steps taken. In the previous chapter, the NRV distribution f^r was approximated with a Gaussian distribution $\hat{f}^r$. It showed that this approximation holds very well in the context of the marginal NRS distribution f_x^{NRS} for the signals examined. From this, conclusions were drawn about the NRS distribution, and it was approximated as a t-distribution with two degrees of freedom, $\dot{f}_x^{NRS}$. This approximation was not very useful in the sense of that it was hard to compare to the original spectrogram which is Gaussian distributed. Therefore, a Gaussian approximation was formed from the t-distribution approximation, $\ddot{f}_x^{NRS}$.

The scale of the Gaussian approximation $\ddot{f}_x^{NRS}$ came from the scaling κ_x of the t-distribution $\dot{f}_x^{NRS}$. That was a bit optimistic, so it was suggested rounding up the scale to $\ddot{\mathrm{D}}_x^{NRS} \approx 1/\sqrt{\rho_0}$ for a simple rule of thumb. Having both the scale of the reassigned and original distribution, the reassignment localization factor γ was suggested to be defined as the ratio between the two. The scale of the RS was also put in context by relating it to the NRV distribution in the center point of the original spectrogram, which has $\mathrm{D}_x \approx 1/\sqrt{2\rho_0}$.

4.5 The matched Gaussian case

In section 2.4 the *matched window reassignment* was introduced. It was seen that scaling the reassignment vector by a factor of $c = 2$ gives perfect localization of any transient signal with an envelope matching the STFT window. We define this scaled NRV as

$$q = c \cdot r = 2r \Leftrightarrow r = q/2 \tag{4.14}$$

Correcting for the factor $c = 2$, the distribution of q becomes

$$f^q(q) = \frac{1}{4} f^r(q/2) \tag{4.15}$$

Naturally, the normalized distance from the signal is

$$|d_n|^2 = t_n^2 + \nu_n^2 \tag{4.16}$$

Since the scaled reassignment gives perfect localization, $\left|q\right| = |d_n|$. As for the previously discussed signals, the normalized spectrogram (B.35) can be expressed as a function of $|d_n|^2$, and therefore also of $\left|q\right|^2$:

$$NS = \sqrt{\pi}\lambda \exp\left(-\frac{|d_n|^2}{4}\right) = \sqrt{\pi}\lambda \exp\left(-\frac{\left|q\right|^2}{4}\right) \tag{4.17}$$

The corresponding PDF $f_q^{NS} \colon \mathbf{C} \to \mathbf{R}$ is now defined such that $\int_C f_q^{NS} = 1$:

$$f_q^{NS}(z) \triangleq \frac{1}{4\pi} \exp\left(-\frac{|z|^2}{4}\right) \tag{4.18}$$

The local SNR ρ function is also defined correspondingly:

$$\rho(z) \triangleq \rho_0 \cdot \exp\left(-\frac{|z|^2}{4}\right), \; \rho_0 = \frac{NS(0)}{2\sigma_n^2} \overset{\text{(B.35)}}{=} \frac{\sqrt{\pi}\lambda}{2\sigma_n^2} \tag{4.19}$$

The expression (4.8) is expanded to allow for the complex PDF:

$$\begin{aligned} f_q^{NRS}(x, y) &= \iint f_q^{NS}(x_0, y_0) \cdot f^q(x_0 - x, y_0 - y)\, \mathrm{d}x_0\, \mathrm{d}y_0\,, \text{ or} \\ f_q^{NRS}(q) &= \int_{\mathbf{C}} f_q^{NS}(q_0) \cdot f^q(q_0 - q)\, \mathrm{d}q_0 \end{aligned} \tag{4.20}$$

Before putting the pieces together, take note of the symmetric nature of the reassigned normalized spectrogram. Since (4.18) is circularly symmetric and (4.15) is rotationally invariant, it must follow that f_q^{NRS} is circularly symmetric:

$$f_q^{NRS}(q) = f_q^{NRS}\left(qe^{i\phi}\right) = f_q^{NRS}\left(\left|q\right|\right) \tag{4.21}$$

Thus, in our analysis, we can let $q \in \mathbf{R}$. However, we can not assume the reassigned distribution f_q^{NRS} to simply be a scaled version of f_x^{NRS} (4.8).

As before, we are able to use the approximative NRV distribution $\hat{f}^r$ in place of f^r. In that case, the NRS distribution is written as $\hat{f}_q^{NRS}$. Figure 4.6 shows both f_q^{NRS} and $\hat{f}_q^{NRS}$ for two different peak SNRs ρ_0. Note that now the cases when $\rho_0 = 10\,\mathrm{dB}$ and $\rho_0 = 20\,\mathrm{dB}$ are shown instead. The results show that for the matched Gaussian case, in this context, the Gaussian approximation $\hat{f}^r$ is slightly worse than for the previously examined signals. However, it still holds reasonably well for $\rho_0 \gtrsim 10\,\mathrm{dB}$. Also, it is worth noting that it is on the pessimistic side.

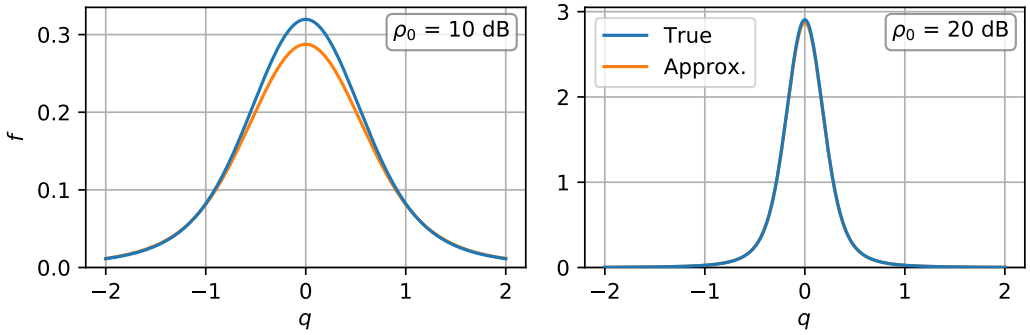

Figure 4.6: The true and approximative NRS distribution f_q^{NRS} and $\hat{f}_q^{NRS}$ respectively, showing two cases for the peak SNRs ρ_0.

Numerical evaluation of f_q^{NRS} showed that it is of infinite variance, just as for the previously examined signals (evaluating f_x^{NRS}). By comparison to other common distributions, it was again found that the distribution resembles a t-distribution with two degrees of freedom. In this case it is the bivariate/complex and circularly symmetric version. Its PDF, named $\dot{f}_q^{NRS}$, can be written as

$$\dot{f}_q^{NRS}(z) = \frac{1}{2\pi\kappa_q^2}\left[1 + \frac{1}{2}\left(\frac{|z|}{\kappa_q}\right)^2\right]^{-2} \tag{4.22}$$

Since the Gaussian approximation $\hat{f}^r$ also works well for the matched Gaussian case, it again follows that $\kappa_q \propto 1/\sqrt{\rho_0}$. Fitting $\dot{f}_q^{NRS}(z)$ to the true distribution gave that $\kappa_q \approx 2.3/\sqrt{\rho_0}$. Figure 4.7 shows the true f_q^{NRS} distribution compared to the t-distribution approximation $\dot{f}_q^{NRS}$. As before, by letting the number of degrees of freedom grow infinite for the t-distribution approximation $\dot{f}_q^{NRS}$, a new Gaussian approximation $\ddot{f}_q^{NRS}$ is formed. This approximation is also shown in figure 4.7. Its PDF can be written as

$$\ddot{f}_q^{NRS}(z) = \frac{1}{2\pi\kappa_q^2}\exp\left(\frac{|z|^2}{2\kappa_q^2}\right) \tag{4.23}$$

where κ_q is the scaling factor for the matched Gaussian case. This fills two purposes, the first of which is the same as before – to be able to compare the width to the original distribution. Its second purpose is to obtain a distribution which factorizes into the marginal distributions, which the t-distribution does not.

For the matched Gaussian case, the normalized spectrogram does not have unit variance as for the previous signals. Instead, $\mathrm{D}_q^{NS} = 2$. Compared to the Gaussian approximation $\ddot{f}_q^{NRS}$ with $\ddot{\mathrm{D}}_q^{NRS} = \sqrt{2}\kappa_q$, the reassignment localization factor γ_q becomes

$$\gamma_q = \frac{\mathrm{D}_q^{NS}}{\ddot{\mathrm{D}}_q^{NRS}} = \frac{2}{\sqrt{2}\kappa_q} \approx \frac{\sqrt{2}}{2.3}\cdot\sqrt{\rho_0} \approx 0.6\cdot\sqrt{\rho_0} \tag{4.24}$$

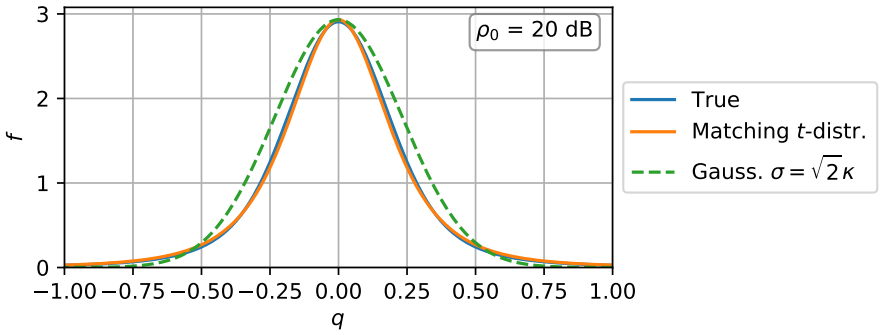

Figure 4.7: A comparison of distributions in relation to the true distribution of the NRS in the matched Gaussian case.

4.6 Comparison to simulations

Concluding this chapter, the derived distributions are tested against simulations. Before looking at the results from the simulations, let us recap on the approach of deriving the NRS distributions. To do so, we redistributed the NS according to the NRV distribution in each point. Noise was modelled by assuming a peak SNR ρ_0 which shapes the NRV distributions. However, no actual noise floor was added to the spectrogram prior to (or post) redistribution. Yet, in the simulations, we expect the simulated spectrograms to have a noise floor.

Throughout the chapter, we have only worked with normalized spectrograms. Going back the the denormalized variables is straightforward, multiplying the normalized time t_n and frequency v_n by the window duration Δt and bandwidth $\Delta \nu$ respectively. Taking for example the sinusoid, this means scaling the normalized marginal distribution by the bandwidth $\Delta \nu$.

Two signals are examined – the sinusoid and the matched Gaussian, starting with the former. In this case, the reassignment is only done in frequency since we are looking at the marginal distribution. Figure 4.8 shows the results from the simulation, which is the mean of 10 000 realizations. In the log-scaled plot, the noise floor is clearly seen, but its overall effect on the shape of the distribution is negligible. In practice, the t-distribution approximation $\dot{f}_x^{NRS}$ seems to hold very well.

For the matched Gaussian, the simulations are done in both variables. To compare to the circularly symmetric theoretical t-distribution approximation $\dot{f}_q^{NRS}$ (4.22), t_n and v_n are reduced to one variable $|q| = |d_n| = \sqrt{t_n^2 + v_n^2}$. Figure 4.9 shows the results from the simulation, which is the mean of 2000 realizations. In contrast to the previous result, it is presented in the normalized variables, making it easier to translate to the distribution

of one variable $|d_n|$. Again, the simulations seem to align well with the theory, and we draw the conclusion that the distribution derived in this chapter are conceptually legitimate.

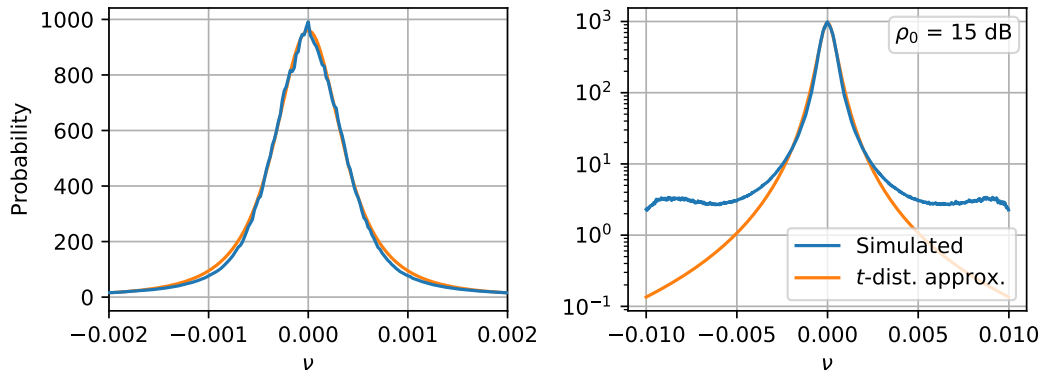

Figure 4.8: A comparison of the simulated and theoretical marginal distributions of the reassigned spectrogram for the sinusoid. $\rho_0 = 15\,\text{dB}$ and $\lambda = 50$. The shown simulation is the mean of 10 000 realizations. The difference between the plots is the scaling of the axes.

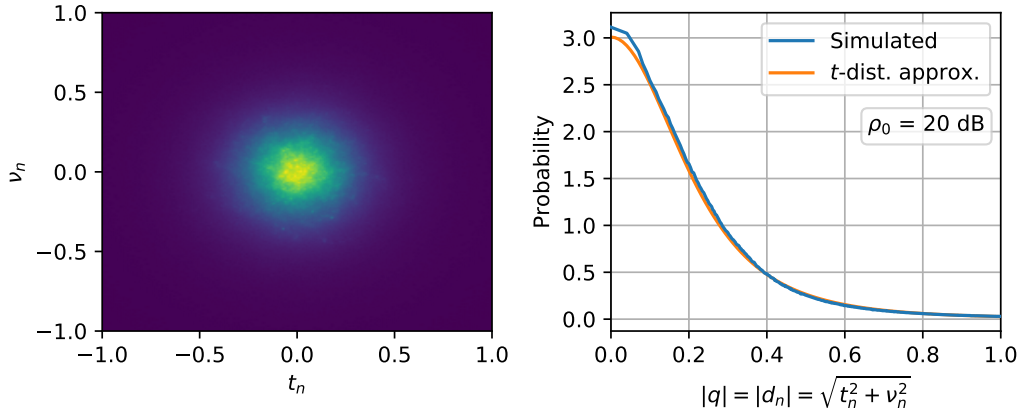

Figure 4.9: The mean of 2000 realizations of the normalized reassigned spectrogram for the matched Gaussian signal at $\rho_0 = 20\,\text{dB}$, $\lambda = 20$. A comparison to the theoretical t-distribution approximation $\dot{f}_q^{NRS}$ is shown in the right plot.

Chapter 5

On the resampling of noisy reassigned Gabor spectrograms

In this chapter, we touch upon the subject of resampling discrete reassigned spectrograms when the signal is subjected to noise. First, we will go through the problem of naively resampling (binning), and how this is sometimes avoided by not resampling at all. Then, a method is presented that utilizes the now known properties of reassigned spectrograms to resample them in a better way. Finally, an evaluation of the method is presented, discussing the performance of the method.

5.1 The issue of resampling

In current literature, the reassigned spectrogram (RS) of a discrete signal is often represented in one of two ways. In the first way, hereafter referred to as the *raw* representation, the RS is not resampled. Instead, the individual reassigned points from the spectrogram are shown. Sometimes, they are plotted with an intensity proportional to their energy. This method is used by for example Fitz et al. in [17]. An example is shown in chapter 2, figure 2.4. In the second way, hereafter referred to as the *binned* representation, the RS is naively resampled by assigning each point to the closest time and frequency point in the original spectrogram. Figure 2.5 shows an example of this method.

The representations are best understood by example. Figure 5.1 shows the reassigned spectrogram for a matched Gaussian pulse. To make it easier to illustrate and understand, we only look at the slice of the spectrogram which frequency matches that of the signal. That is, we only look at the spectrogram and its reassignment in the time domain. Plot A shows the real part of the signal, with (orange) and without (blue) noise. Note that the peak SNR $\rho_0 = 25\,\text{dB}$, which is more than enough for all our previous approximations to hold well. The Gaussian's width $\lambda = 10$ and the reassignment localization factor (4.24) $\gamma_q \approx 0.6 \cdot \sqrt{\rho_0} \approx 19$. Thus, the corresponding width of the reassigned spectrogram is approximately $\lambda / \gamma_q \approx 0.5$.

Plot B shows the spectrogram, and plot C and D, respectively, show the raw and binned representation of the reassigned spectrogram. Note that the two right plots (C and D) are zoomed in by a factor of 8. If it is not already, it will soon be clear that both representations have their drawbacks.

The raw representation (C) is fairly good for visualization, as the spread of the points does not affect the perceived energy in the signal. However, it does not show the total energy of the signal, and is not a very suitable representation for further processing. Contrastingly, the binned representation (D) does the exact opposite. For the noise-free signal, the total energy is correctly obtained in the peak, but when it is subjected to noise, the binned RS spreads and deteriorates. For reference, the dashed line shows the total energy in the spectrogram.

Deterioration of the binned RS could perhaps be compensated with wider bins. But then, the localization of the peak would be subject to a larger quantization error. Also, no matter the bin width, if the reassigned points align between two bins, the energy is often split between the two. We have observed that if this happens, the robustness of the binned RS is drastically worsened.

To put it shortly, both of these representations have their often unacceptable drawbacks. The question naturally arises – how do we represent/resample the reassigned spectrogram in a suitable and balanced way? Is it possible to find a middle ground where the total energy can be obtained in the peak while simultaneously being robust against noise?

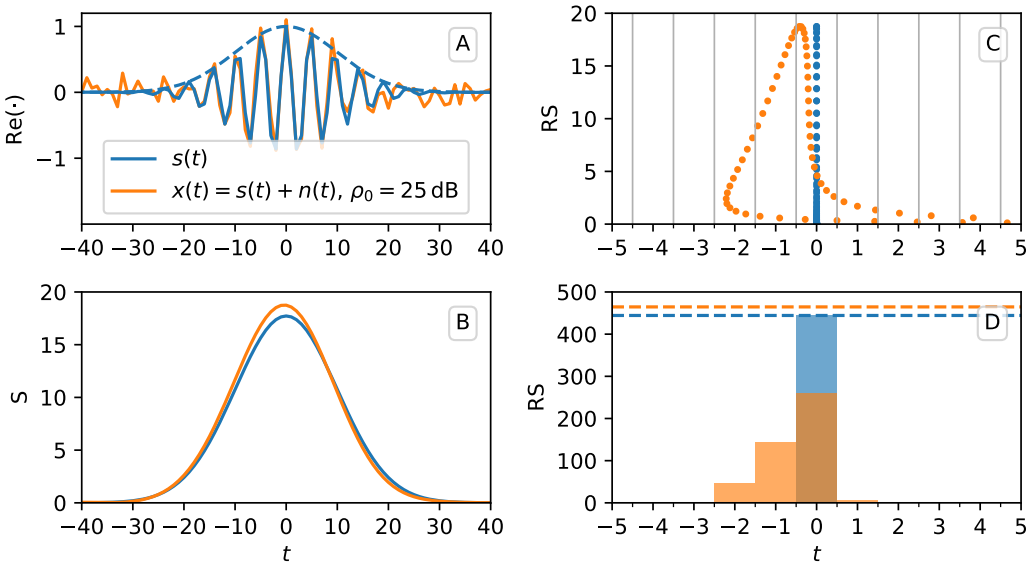

Figure 5.1: An example of the reassigned spectrogram (RS) of a matched Gaussian pulse, shown both with and without noise. Plot A shows the signals. B shows the spectrograms. C and D shows the raw and binned RSs respectively. The figure is thoroughly explained and discussed in section 5.1.

5.2 Proposal of the smoothed reassigned spectrogram

Now, we present an idea for a method attempting to mitigate the issue presented in the previous section. Let us call this method the *smoothed reassigned spectrogram*, or SRS for short. Its main goal is to be able to obtain a stable measure of the signal energy/power in its peak. The basic idea is to place a kernel on each reassigned point with some scale relating to that of the RS or RV[1] distribution. Essentially this means smoothing the reassigned points before sampling to obtain a well behaved representation of the RS. In contrast to the binned method, this representation can be sampled as finely as desired without deteriorating. The concept is similar to *kernel density estimation* (KDE) or *kernel smoothing*.

The SRS makes the most sense in the discrete case, where it actually will be applied. To express it, let the discrete spectrogram be written as $S_i(t_i, \nu_i)$, $i \in \{1..N\}$, where N is the number of points in the discrete spectrogram. The value S_i will sometimes be referred to as the mass. Let its corresponding reassigned time and frequency be written as $(\hat{t}_i, \hat{\nu}_i)$. Then,

$$SRS(t', \nu') \triangleq \sum_i S_i \cdot K_i(t' - \hat{t}_i, \nu' - \hat{\nu}_i) \tag{5.1}$$

where K is a kernel, possibly a Gaussian:

$$K^G(t, \nu) = \exp\left[-\frac{1}{2}\left(\frac{t^2}{\sigma_{K,t}^2} + \frac{\nu^2}{\sigma_{K,\nu}^2}\right)\right] \tag{5.2}$$

or a parabola:

$$K^P(t, \nu) = \max\left[0,\ 1 - \frac{1}{2}\left(\frac{t^2}{\sigma_{K,t}^2} + \frac{\nu^2}{\sigma_{K,\nu}^2}\right)\right] \tag{5.3}$$

where $\sigma_{K,t}^2$ and $\sigma_{K,\nu}^2$ are scaling parameters. The scale of a kernel K_i may depend on S_i or $S(\hat{t}_i, \hat{\nu}_i)$, but more on that later. In continuous variables, the SRS can be expressed as

$$SRS(t, \nu) \triangleq \iint RS(t', \nu') \cdot K(t - t', \nu - \nu')\, \mathrm{d}t'\, \mathrm{d}\nu' \tag{5.4}$$

If K is fixed, this is simply a convolution with the RS, which is convenient for theoretic calculations.

Before showing an example, let us think of a desirable property of the SRS. Consider a Gaussian centered in (t_0, ν_0). Its noise-free RS is perfectly localized in the center point with the total signal energy E. That is, $RS(t, \nu) = E \cdot \delta(t - t_0, \nu - \nu_0)$. The SRS should retain that $SRS(t_0, \nu_0) = E$, which corresponds to the mass in $RS(t_0, \nu_0)$. Given this requirement, $K(0, 0) = 1$.

[1] Reassignment Vector

It is not uncommon that in other contexts, kernels are constrained to be of unit energy. Combining this constraint with $K(0,0) = 1$ would imply a fixed kernel width. This goes against the idea of relating the kernel width to that of the RS or RV distribution. Hence, the kernels can not be constrained to be of unit energy. It follows that the SRS does not have the energy conservation property that the spectrogram and reassigned spectrogram both have. But remember – the goal is to reliably find the energy/power in the peak.

Before further discussing kernels, let us first show a practical example. Figure 5.2 shows two examples of placing kernels on the reassigned points of the spectrogram. The figure shows the same signal as figure 5.1 but with some points removed to reduce clutter in the plot. As before, to simplify, we only look in the time domain. In both examples, Gaussian kernels are used, but with different scales. The top plots use kernels as wide as the RS distribution (theoretically), and the bottom plots use a 3 times wider kernel. Those kernels were picked arbitrarily – the point here is only to illustrate how the kernel scale affects the SRS.

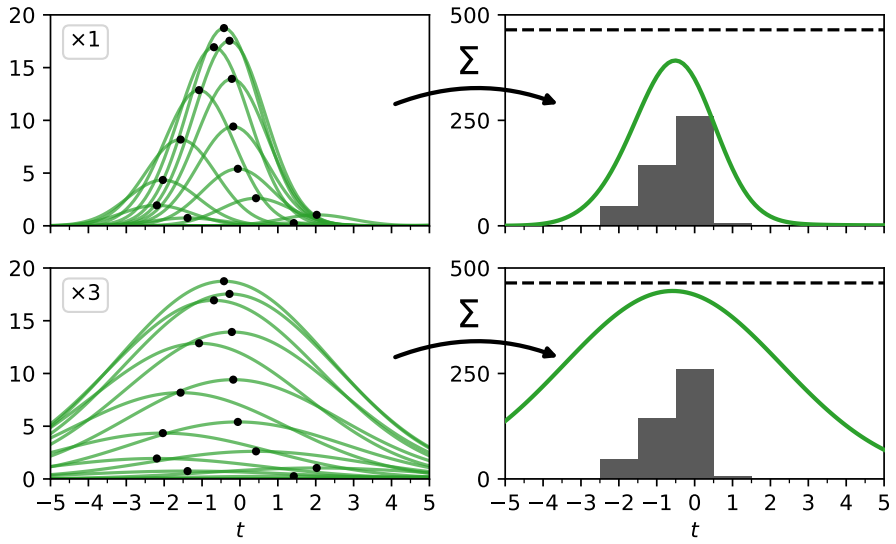

Figure 5.2: Two examples of placing Gaussian kernels on each reassigned point. The two examples have different kernel scales, which affects the resulting SRS, shown in green in the right plot. For comparison, in the right plots, the bars show the binned RS, and the dashed lines shows the total energy.

The main point of figure 5.2 is to show that if the smoothing kernels are wide enough, the peak of the SRS will provide a more stable estimate of the signal power. By "more stable", we mean less biased and with lower variance over multiple realizations. But of course, if a too wide kernel is used, there is no point in reassigning at all.

Now that we have an idea of how the SRS behaves, let us discuss kernels. There is two more desired properties of the SRS that we would like to suggest. It should retain that a noise-free Gaussian with a total signal energy E is unambiguously located in its center (t_0, ν_0). I.e., $SRS(t, \nu) < E$ for $(t, \nu) \neq (t_0, \nu_0)$. For the kernel, this means that $K(t, \nu) < 1$ for $(t, \nu) \neq (0, 0)$. This rules out using, for example, a rectangular window. Finally, the SRS should never be negative, so $K \geq 0$.

5.3 Kernel selection

Both the kernels mentioned – the Gaussian (5.2) and the parabola (5.3) – fulfill the properties stated in the previous section. Also, the scaling factors $\sigma_{K,t}$ and $\sigma_{K,\nu}$ can be chosen such that the kernels become circularly symmetric in the normalized variables. We will limit the scope to these two kernels, but of course, other choices are possible.

The subject on kernel selection for kernel smoothing and KDE is still today widely discussed, and stands without universal consensus. Going down that rabbit hole would also be outside the scope of this thesis. Instead, a rather brief, practical, and nonrigorous approach is taken.

As done in the previous chapter, we continue to approach this problem by only looking at the marginal distribution. For the matched Gaussian case, extension to both the time and frequency domain is trivial since the NRS distribution is circularly symmetric. For the other signals examined, this is a subject of further research. But for now we suggest treating this case as also circularly symmetric in the normalized variables.

In the previous chapter, section 4.4, we approximated the RS marginal distribution as a Gaussian. Using this Gaussian approximation together with a fixed Gaussian kernel results in some nice properties. For a fixed kernel, the SRS becomes a convolution of the RS and kernel, which means that it too will be Gaussian. Here, we define the RS with unit energy and variance as

$$RS(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \tag{5.5}$$

and the kernel K scaled with a factor a relative to the RS as

$$K(x) = \exp\left(-\frac{x^2}{2a^2}\right) \tag{5.6}$$

The SRS is the convolution of the RS and kernel K:

$$SRS(x) \overset{[18]}{=} \frac{1}{\sqrt{1 + 1/a^2}} \cdot \exp\left(-\frac{x^2}{2\,(1 + a^2)}\right) = E_0^{SRS} \cdot \exp\left(-\frac{x^2}{2\,\sigma_{SRS}^2}\right) \tag{5.7}$$

As the kernel grows wider, i.e., as $a \to \infty$, then $E_0^{SRS} \to 1$ and asymptotically, $\sigma_{SRS} = \sqrt{1+a^2} \to a$. In other words, the energy is recovered, and the SRS width becomes that of the kernel. The relationship between the kernel width a and energy recovered E_0^{SRS} can be seen in figure 5.3. If we want to recover 95 % of the energy, we pick $a \approx 3.0$. This gives an SRS scaling of $\sqrt{1+a^2} \approx 3$. To clarify, if we use Gaussian kernels ~ 3 times wider than the estimated RS width, we expect to recover 95 % of the signal energy in the SRS peak. The bottom plots of figure 5.2 shows an example of exactly this.

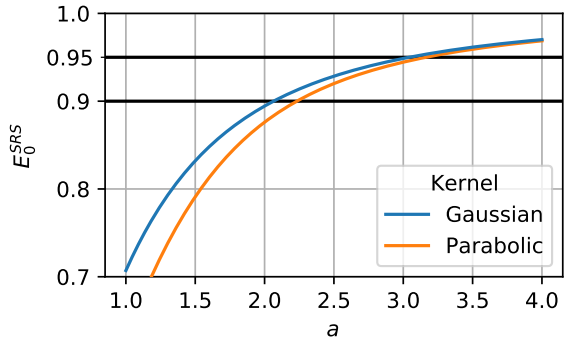

Figure 5.3: The relationship between kernel scale a and expected energy recovered E_0^{SRS}.

This is fine, but how is the RS distribution width obtained? A simple solution would be to estimate the global peak SNR ρ_0 to calculate the reassignment localization factor γ, which in turn is used to estimate the width. This works fine if there is only one component in the signal, but what if there are multiple components with different energy? Clearly, in that case, the kernel width will need to be adapted locally.

To adapt the kernel width locally, we would need to somehow use the local SNR instead of the global peak SNR. The obvious straight forward solution would be to simply use the pointwise SNR, estimated from the points' individual energy. But what will happen then is that points with low SNR will get spread out more, creating heavy tails around peaks and increasing the noise floor of the SRS. This is likely not a desired behavior.

What do we propose to do instead? One idea is to not use the energy of the point itself, but the local SNR at where the point is reassigned to. That is, $\rho(\hat{t}, \hat{\nu})$ instead of $\rho(t, \nu)$. All points (in the marginal) for a signal component are ideally reassigned to the same central point. Thus, all those points will ideally be assigned the same SNR – the peak SNR for that component. This technique will later be demonstrated on some real world data.

A desirable kernel property not mentioned before is finite support. This makes the SRS practically feasible to compute. Otherwise, if the kernels have infinite support, every point in the RS contributes to every point in the SRS. This results in some rather bad

time complexity, and should be avoided. Of course, the Gaussian kernel can simply be cut to make it finite support, but there are other kernels which have finite support by nature. For example, the parabola previously presented 5.3.

The parabola is often used in KDE since it is optimal in the sense that it minimizes spread [19]. We will not dive into what this means for our application, but it could be an argument for its feasibility. In the context of the marginal distribution, we define the kernel as

$$K(x) = \begin{cases} 1 - x^2/2a^2 & \text{for } x^2 < 2a^2 \\ 0 & \text{otherwise} \end{cases} \tag{5.8}$$

Note that it matches the first two terms of the Taylor series of the Gaussian kernel. In other words, it resembles the Gaussian kernel close to its center. One drawback of any other kernel than the Gaussian is that the resulting SRS distribution will not be Gaussian. However, we can still calculate the expected energy recovery:

$$E_0^{SRS} = SRS(0) = \int_{-\sqrt{2}a}^{\sqrt{2}a} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \cdot \left[1 - \frac{x^2}{2a^2}\right] \mathrm{d}x \tag{5.9}$$

Again, the relationship between scale a and recovered energy E_0^{SRS} is plotted in figure 5.3. It can be seen that in this regard, the parabolic kernel behaves similarly to the Gaussian. Recovering 95 % of the energy again gives a scale $a \approx 3$. Finally, studying the plot, we suggest that reasonable choices of scale a fall somewhere between 2 and 4.

5.4 Suggestions for practical enhancements

To further enhance the reassigned spectrogram, there are some more things we can do. In this section, we suggest a set of criteria which a point must meet to be included. These criteria have hard thresholds – either a point is included or ignored. The reason for this is that it drastically speeds up the average computation time for the SRS.

In chapter 3, section 3.3.1, we learned that points with an SNR $\rho \lesssim 5\,\mathrm{dB}$ have a significantly biased expectation value. If we know that the reassignment will be incorrect, there is no point of reassigning at all. Therefore, we suggest that points below a certain SNR ρ, somewhere around $\rho \lesssim 5\,\mathrm{dB}$, are ignored.

We can also look at the SNR where the point is reassigned $(\hat{t}, \hat{\nu})$. If that SNR is too low, it is likely that the point was not properly reassigned. Therefore we suggest ignoring points which are reassigned to a location with a SNR below some threshold. This threshold should be at least as high as the previous. As this SNR is suggested to be used for the kernel width, this rule also prevents kernels from becoming too wide.

In chapter 3 and 4, we learned that points that are reassigned long distances have higher variance and contribute to the heavy tails of the RS distribution. In section 4.4, it was mentioned that these far out points can be ignored to some extent. But how do we

decide what is a "far out" point? Consider the marginal distribution of the spectrogram, which is a Gaussian. Let σ be its standard deviation. If we want to recover $\sim 95\,\%$ of its energy, this corresponds to reassigning all points from within $\pm 2\sigma$. Therefore, we suggest ignoring points which are reassigned a distance longer than $\sim 2\sigma$.

To summarize, ignore points which do not meet the following criteria:

- The point's SNR $\rho(t, \nu) \gtrsim 5\,\text{dB}$.
- The SNR at the location which the point is reassigned to $\rho(\hat{t}, \hat{\nu}) \gtrsim 5\,\text{dB}$.
- The point is reassigned a distance smaller than $\sim 2\sigma$, where σ is the standard deviation of the spectrogram marginal distribution.

There is one more idea that we would like to present. It briefly touches the topic on what happens for multi-component signals. Chassande-Mottin et al. have done some research on the topic, presented in [20]. One takeaway is that the reassignment breaks down between signal components which are too close. Let us take a look at what happens. Figure 5.4 shows an example where two Gaussians are placed fairly close to each other in time. The signal can be seen in plot A. Plot B shows the RS and SRS with the enhancements previously discussed. The SRS uses parabolic kernels 3 times wider than the theoretical RS distribution width. It can be seen that there is some interference between the signal components.

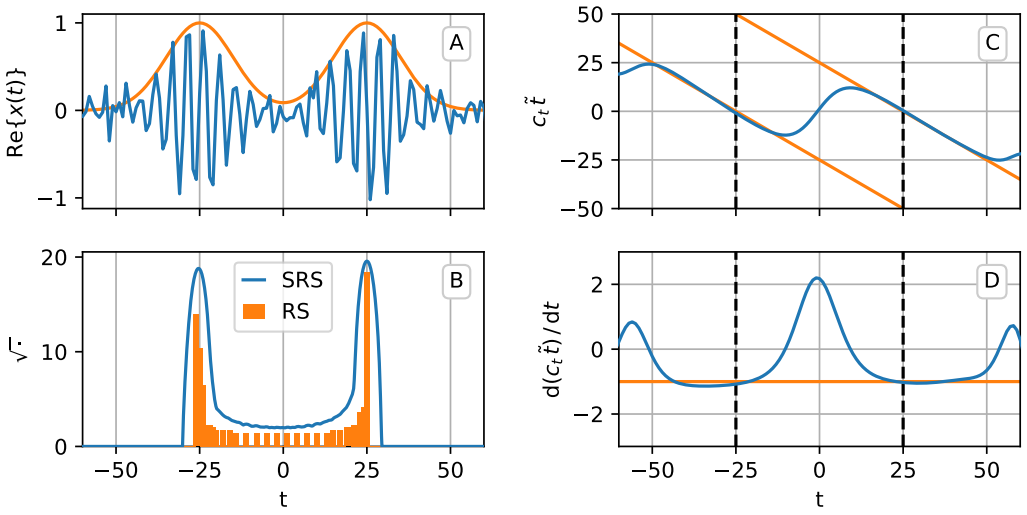

Figure 5.4: A typical scenario for a signal with multiple close components. Plot A shows the signal which is two Gaussians close in time. Plot B shows the RS and SRS. Plot C shows the actual (blue) and ideal (orange) scaled reassignment in time $c_t\tilde{t}$. Plot D shows the derivative of $c_t\tilde{t}$.

To understand what is going on, we look at plot C. The reassignment in time ($c_t\tilde{t}$) is shown by the blue line. Remember, $c_t = 2$ for the matched window reassignment. The ideal reassignment in time for each component is shown in orange. What happens

between the centers of the components is a transition from reassigning points to one center to the other. During the transition, the points are incorrectly reassigned. While it might not be possible to fix this due to the nature of reassigning spectrograms, these points could perhaps be ignored instead.

So how can we classify which points to be ignored? We know that ideally, $c_t\tilde{t} = t_0 - t$, which is shown in plot C. This means that ideally, $\mathrm{d}\left(c_t\tilde{t}\right)/\mathrm{d}t = (c_t\tilde{t})' = -1$, which is what plot D shows. It can be seen that when the reassignment is well behaved, $(c_t\tilde{t})' \approx -1$. Therefore, we suggest discriminating points which have $(c_t\tilde{t})' \not\approx -1$. Exactly how to set the limits has not been thoroughly studied. For the results later shown, the limits used are $-1.5 < (c_t\tilde{t})' < -0.5$.

The result of taking the example signal discussed (figure 5.4) and applying this rule can be seen in figure 5.5. It can be seen that most of the interference between the two components is removed. Having this smoothed reassigned spectrogram, it would be trivial to locate the components.

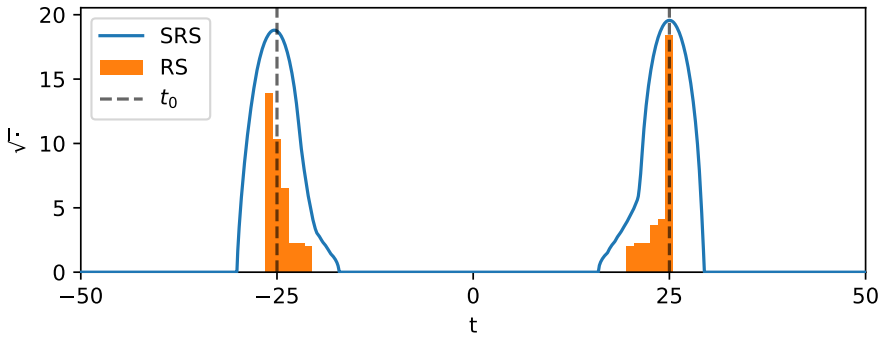

Figure 5.5: The RS and SRS from figure 5.4 after applying the rule related to $(c_t\tilde{t})'$.

5.5 Performance evaluation

Remember, the main goal of the SRS is to be able to obtain a stable measure of the signal energy/power in its peak. Now that we have built up some theory around the SRS, let us put it to the test. We will use the same signal as before – a single matched Gaussian with $\lambda = 10$. Again, for simplicity, we only look at the spectrogram and its reassignment in the time domain. All enhancements proposed in the previous section are applied, and a parabolic kernel with a relative scale of 3 is used. That is, the kernel scale

$$\sigma_{K,t} = 3 \cdot \frac{\lambda}{\gamma_q} = 3 \cdot \frac{10}{0.6} \cdot \left[\rho(\hat{t})\right]^{-1/2} \tag{5.10}$$

With this kernel, the expected energy recovery is $\sim 94\,\%$. Also, since we ignore points reassigned further than 2σ, that in itself limits the energy recovery by $\sim 4\,\%$.

In the following simulation study, the spectrogram, binned RS and SRS was realized $N = 5000$ times. From each realization, the peak value from each of the three TFRs was saved. To avoid corner cases, the center time and frequency (t_0, ν_0) was randomized every realization. For the binned RS and SRS, the estimated energy $\hat{E}$ was obtained directly from the peak. To estimate the total energy $\hat{E}$ from the spectrogram peak,

$$\hat{E} = \max_t[S(t)] \cdot \lambda\sqrt{2\pi} \tag{5.11}$$

The spectrogram is the baseline which the binned RS and SRS are compared against. The point here is that we wish to retain this information when reassigning.

A wide range of peak SNRs ρ_0 was swept, from 12 dB to 44 dB. For every peak SNR, the mean, 25th and 75th percentiles, and root mean square error (RMSE) was calculated. The RMSE is defined as

$$\text{RMSE} = \left[\frac{1}{N}\sum_{n=1}^{N}(E - \hat{E}_n)^2\right]^{1/2} \tag{5.12}$$

where E is the theoretical energy in the noise-free signal. Figure 5.6 shows the RMSE, and figure 5.7 shows the mean and 25th/75th percentiles.

We start by looking at what happens for the binned RS. Clearly, it suffers from significant deterioration, worsening as noise is added. Even for a peak SNR ρ_0 as high as 25 dB, its mean peak value is heavily biased. At even higher SNRs, when its no longer much more biased than the SRS, it still has a much higher RMSE.

Perhaps more interesting is looking at the SRS. In contrast to the binned RS, it has a fairly constant bias. For the lower SNRs, the RMSE of the SRS is only slightly worse than the spectrogram baseline. At higher SNRs, the RMSE goes to some lower limit caused by the bias. Overall, its behavior is very similar to that of the spectrogram. This is great since it means that using the SRS, we can reassign the spectrogram while retaining information of signal energy/power.

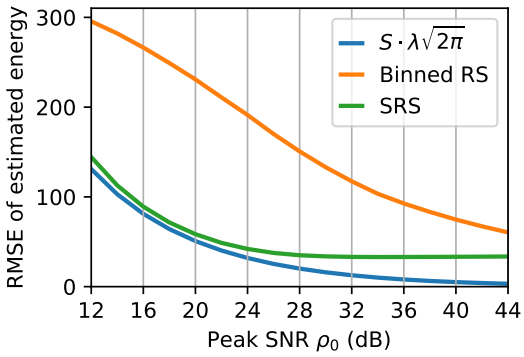

Figure 5.6: RMSE of the energy estimated from the spectrogram, binned RS, and SRS.

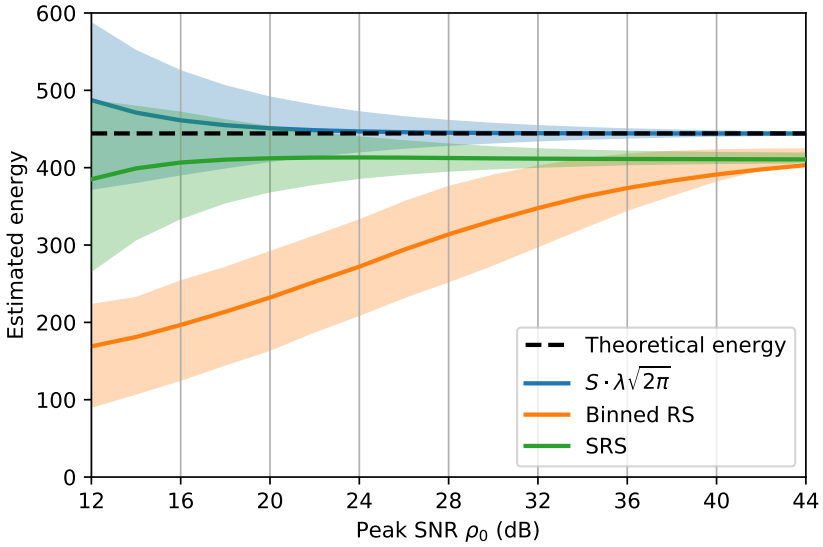

Figure 5.7: The energy estimated from the peak the spectrogram, binned RS, and SRS. Solid lines show the mean and the fill shows the span between the 25th and 75th percentile. For reference, the theoretical noise-free signal energy is shown by the dashed black line.

5.6 Application on real world radar data

One of the original purposes of this thesis was to find new techniques for processing data from the Acconeer A1 radar sensor. The A1 is a tiny pulsed coherent 60 GHz radar with very low power consumption, ideal for small battery powered devices. It is a time-of-flight based system, meaning that it measures the time between transmitting and recieving a reflected pulse, which corresponds to some distance in space. Measurements are taken by correlating the recieved pulse with a time delayed version of the transmitted pulse. By repeating measurements over a range of delays, like a sliding correlator, the reflected pulse is reconstructed. This is referred to as *sweeping*, and the data from sweeping a range of delays is simply called a *sweep*. The resolution in delay is ~ 3 ps which corresponds to ~ 0.5 mm in distance.

In a sweep, a reflection from an object will appear as the radar pulse correlated with a matched filter. Since the pulse envelope approximately has the shape of a Gaussian, the output from the matched filter also will be Gaussian. The top plot in figure 5.8 shows a sweep where two objects has been placed in front of the sensor. It also shows the matched Gaussian STFT window, obtained by matching it by hand ($\lambda = 27$ mm). Other methods of matching the STFT window exists, such as minimizing the Rényi entropy [8].

The two objects placed in front of the sensor were of different material and shape, thus creating slightly different pulses in the sweep. Also, the two objects were close enough that the pulses interfere with each other. On top of that, there are some nonlinear effects and slight correlation between the points in the sweep. Overall, a not so favorable situation for the RS and SRS, even though a maximum SNR of 41 dB is reached. In the figure, the standard binned RS (orange) can be seen breaking down, especially for the peak further away. However, the SRS recovers the peaks rather well. For both the binned RS and SRS, all enhancements presented in the previous section are used.

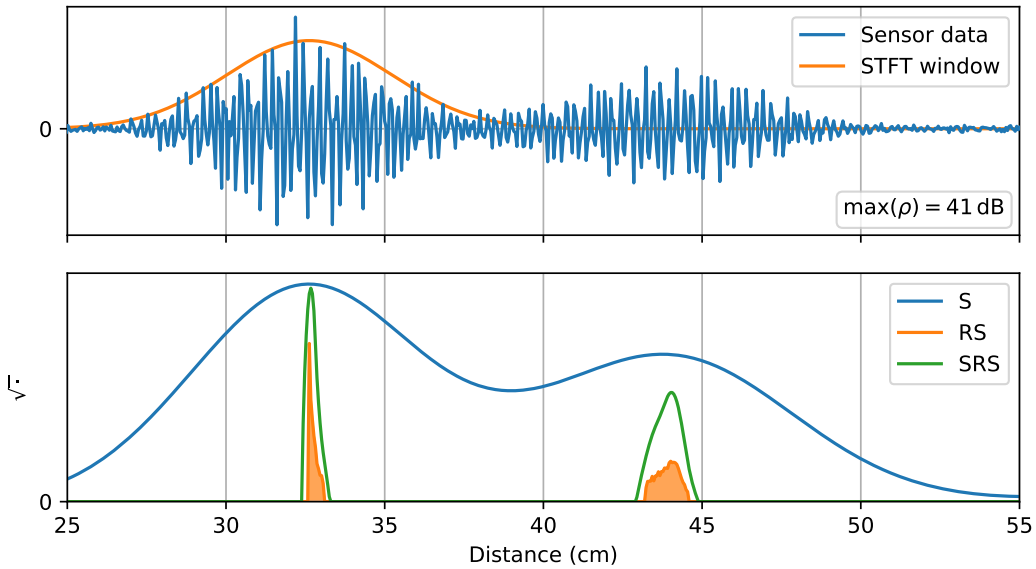

Figure 5.8: An example of processing data from the Acconeer A1 sensor. The top plot shows the data in which two pulses can be seen. It also shows the manually matched STFT window. The bottom plot shows the slice of the spectrogram (blue), binned RS (orange), and SRS (green), matching the frequency of the signal. Note that the spectrogram has been scaled up from demonstration purposes.

Chapter 6
Conclusions and future work

6.1 Conclusions

In this thesis, we explored the reassignment method for Gabor spectrograms. The main interest was analyzing what happens when deterministic signals are subjected to additive white Gaussian noise (AWGN). A practical approach was taken, and many of the conclusions drawn were a result of numerical evaluation. While perhaps nonrigorous, this allowed us to progress faster, greatly extending the scope of this thesis.

In chapter 3, we started by looking the statistics of the normalized reassignment vector (NRV). A key observation was that the NRV has some nice symmetric properties, simplifying future analysis. We saw that over a certain local SNR, it always holds that the NRV is practically unbiased. The variance of its distribution was studied, and an approximation for it was made. Then we looked at its mode and skewness, trying to better understand its behavior. Finally, from the knowledge gained, a Gaussian approximation of the distribution was made. It assumes that the NRV is unbiased, meaning that the SNR is reasonably high. The main purpose of the approximation was to simplify future analysis.

Chapter 4 took a step back from just looking at the distribution of the NRV to looking at that of the whole reassigned spectrogram. A normalized spectrogram was presented in which the impulse, sinusoid, and linear chirp are, in the context of reassignment, practically equivalent. The distribution of their reassigned spectrogram was derived, both using the true and approximative NRV distribution from chapter 3. It showed that the approximative NRV distribution holds well in this context, allowing us do draw further conclusions.

One conclusion drawn about these signals' reassigned distribution was that it is not of finite variance. By first making a t-distribution approximation, we could then from that form a Gaussian approximation. While not super accurate, it still allowed us to relate the width of the reassigned distribution to the original spectrogram distribution. This relation was defined as the *reassignment localization factor*. It basically tells us that the localization provided by the reassignment is proportional to the square root of the SNR.

Chapter 5 brought up the subject on how reassigned spectrogram are represented/resampled in practice. A big shortcoming with naively resampling the reassigned spectrogram is that it deteriorates when subjected to noise. To mitigate this, a new method

was presented – the *smoothed reassigned spectrogram*. By utilizing the knowledge from previous chapters, applying kernels to the reassigned points, a more well behaved distribution was obtained.

6.2 Future work

There are many areas in this thesis that are open for future work. Overall, a more rigorous approach could be taken to many of the presented results. For example, one could attempt analytically showing that the normalized expectation value of the NRV r is independent of the noise-free NRV r_0. Or, finding a closed expression for said value, variance, or skewness. This also goes for the calculations on the distribution of the reassigned spectrogram. Perhaps one could generalize to any arbitrary window, maybe using the matched window reassignment. Of course, we also think it would be interesting to dive deeper into the proposed *smoothed reassigned spectrogram*. What are the theoretical statistical properties? How do we optimally select kernels? The questions are many, but there is only so much that can be done in one thesis.

Appendix A

Formulas for Gaussian functions

For $a, b \in \mathbf{C}$, $\mathrm{Re}(a) > 0$:

$$\int \exp\left(-ax^2\right) \mathrm{d}x = \sqrt{\frac{\pi}{a}} \tag{A.1}$$

$$\int \exp\left(-ax^2\right) \exp(-2bx) \, \mathrm{d}x = \sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2}{a}\right) \tag{A.2}$$

$$\int x \exp\left(-ax^2\right) \exp(-2bx) \, \mathrm{d}x = -\frac{b}{a}\sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2}{a}\right) \tag{A.3}$$

$$\int_{\mathbf{C}} \exp\left(-a|z|^2\right) \mathrm{d}z = \frac{\pi}{a} \tag{A.4}$$

For all a:

$$\frac{\mathrm{d}}{\mathrm{d}x} \exp\left(-ax^2\right) = -2ax \cdot \exp\left(-ax^2\right) \tag{A.5}$$

Hence, for $a \neq 0$:

$$x \cdot \exp\left(-ax^2\right) = -\frac{1}{2a} \cdot \frac{\mathrm{d}}{\mathrm{d}x} \exp\left(-ax^2\right) \tag{A.6}$$

For $a > 0$:

$$\left(\mathcal{F} \exp\left(-at^2\right)\right)(\nu) = \int \exp\left(-at^2\right) \exp(-i2\pi\nu t) \, \mathrm{d}t \overset{(A.2)}{=} \sqrt{\frac{\pi}{a}} \exp\left(-\frac{\pi^2\nu^2}{a}\right) \tag{A.7}$$

For $a > 0$:

$$\int x^2 \exp\left(-ax^2\right) \mathrm{d}x = \frac{\sqrt{\pi}}{2a^{3/2}} \tag{A.8}$$

Appendix B

Calculations for the spectrogram and reassignment of some common signals

B.1 Impulse

$$s(t) = \delta(t - t_0) \tag{B.1}$$

$$\begin{aligned} F_s^h(t,\nu) &= \frac{1}{\pi^{1/4}\sqrt{\lambda}} \int \delta(\tau - t_0) \exp\left(-\frac{(\tau - t)^2}{2\lambda^2}\right) \exp(-i2\pi\nu\tau)\,\mathrm{d}\tau \\ &= \frac{1}{\pi^{1/4}\sqrt{\lambda}} \exp\left(-\frac{(t - t_0)^2}{2\lambda^2}\right) \exp(-i2\pi\nu t_0) \end{aligned} \tag{B.2}$$

$$F_s^{\mathcal{T}h}(t,\nu) = (t_0 - t) \cdot F_s^h(t,\nu) \tag{B.3}$$

$$S_s^h(t,\nu) = \left|F_s^h(t,\nu)\right|^2 = \frac{1}{\sqrt{\pi}\lambda} \exp\left(-\frac{(t - t_0)^2}{\lambda^2}\right) \tag{B.4}$$

$$r_0 = \frac{\sqrt{2}}{\lambda} \frac{F_s^{\mathcal{T}h}(t,\nu)}{F_s^h(t,\nu)} = -\frac{\sqrt{2}}{\lambda}(t - t_0) \tag{B.5}$$

$$|r_0|^2 = \frac{2}{\lambda^2}(t - t_0)^2 \tag{B.6}$$

With $t_0 = 0$:

$$NS_s^h(t_n,\nu_n) = \frac{1}{\sqrt{2\pi}\Delta t} \exp\left(-\frac{t_n^2}{2}\right) \tag{B.7}$$

$$t_n^2 = \frac{t^2}{\Delta t^2} \overset{(2.12)}{=} \frac{2}{\lambda^2} t^2 = |r_0|^2 \tag{B.8}$$

$$NS_s^h(r_0) = \frac{1}{\sqrt{2\pi}\Delta t} \exp\left(-\frac{|r_0|^2}{2}\right) = NS_s^h(0) \cdot \exp\left(-\frac{|r_0|^2}{2}\right) \tag{B.9}$$

B.2 Sinusoid

$$s(t) = \exp(i2\pi\nu_0 t) \tag{B.10}$$

$$\begin{aligned}
F_s^h(t,\nu) &= \frac{1}{\pi^{1/4}\sqrt{\lambda}} \int \exp[i2\pi(\nu_0-\nu)\tau] \exp\left[-\frac{(\tau-t)^2}{2\lambda^2}\right] \mathrm{d}\tau \\
&= \frac{1}{\pi^{1/4}\sqrt{\lambda}} \int \exp[i2\pi(\nu_0-\nu)(\tau-t)] \exp\left[-\frac{\tau^2}{2\lambda^2}\right] \mathrm{d}\tau \\
&= \frac{\exp[i2\pi(\nu-\nu_0)t]}{\pi^{1/4}\sqrt{\lambda}} \int \exp[-i2\pi(\nu-\nu_0)\tau] \exp\left[-\frac{\tau^2}{2\lambda^2}\right] \mathrm{d}\tau \\
&\overset{(A.7)}{=} \frac{\exp[i2\pi(\nu-\nu_0)t]}{\pi^{1/4}\sqrt{\lambda}} \sqrt{2\pi\lambda^2} \exp\left[-2\pi^2\lambda^2(\nu-\nu_0)^2\right] \\
&= \sqrt{2\lambda}\pi^{1/4} \exp\left[-2\pi^2\lambda^2(\nu-\nu_0)^2\right] \exp[i2\pi(\nu-\nu_0)t]
\end{aligned} \tag{B.11}$$

$$\begin{aligned}
F_s^{\mathcal{T}h}(t,\nu) &= \frac{\exp[i2\pi(\nu-\nu_0)t]}{\pi^{1/4}\sqrt{\lambda}} \int \exp[-i2\pi(\nu-\nu_0)\tau]\tau \exp\left(-\frac{\tau^2}{2\lambda^2}\right) \mathrm{d}\tau \\
&\overset{(A.6)}{=} -\lambda^2 \frac{\exp[i2\pi(\nu-\nu_0)t]}{\pi^{1/4}\sqrt{\lambda}} \\
&\quad \cdot \int \exp[-i2\pi(\nu-\nu_0)\tau] \frac{\mathrm{d}}{\mathrm{d}\tau}\left[\exp\left(-\frac{\tau^2}{2\lambda^2}\right)\right] \mathrm{d}\tau \\
&= -\lambda^2[i2\pi(\nu-\nu_0)] \cdot F_s^h(t,\nu)
\end{aligned} \tag{B.12}$$

$$S_s^h(t,\nu) = \left|F_s^h(t,\nu)\right|^2 = 2\sqrt{\pi}\lambda \exp\left[-4\pi^2\lambda^2(\nu-\nu_0)^2\right] \tag{B.13}$$

$$r_0 = \frac{\sqrt{2}}{\lambda} \frac{F_s^{\mathcal{T}h}(t,\nu)}{F_s^h(t,\nu)} = -i2\sqrt{2}\pi\lambda(\nu-\nu_0) \tag{B.14}$$

$$|r_0|^2 = 8\pi^2\lambda^2(\nu-\nu_0)^2 \tag{B.15}$$

With $\nu_0 = 0$:

$$NS_s^h(t_n,\nu_n) = \frac{1}{\sqrt{2\pi}\Delta\nu} \exp\left(-\frac{\nu_n^2}{2}\right) \tag{B.16}$$

$$\nu_n^2 = \frac{\nu^2}{\Delta\nu^2} \overset{(2.12)}{=} 8\pi^2\lambda^2\nu^2 = |r_0|^2 \tag{B.17}$$

$$NS_s^h(r_0) = \frac{1}{\sqrt{2\pi}\Delta\nu} \exp\left(-\frac{|r_0|^2}{2}\right) = NS_s^h(0) \cdot \exp\left(-\frac{|r_0|^2}{2}\right) \tag{B.18}$$

B.3 Linear chirp

$$s(t) = \exp\left(i2\pi\frac{\beta t}{2}t\right) = \exp\left(i\pi\beta t^2\right) \tag{B.19}$$

$$\begin{aligned}
F_s^h(t,\nu) &= \frac{1}{\pi^{1/4}\sqrt{\lambda}}\int \exp\left(i\pi\beta\tau^2\right)\exp\left(-\frac{(\tau-t)^2}{2\lambda^2}\right)\exp(-i2\pi\nu\tau)\,\mathrm{d}\tau \\
&= \frac{\exp\left(-t^2/2\lambda^2\right)}{\pi^{1/4}\sqrt{\lambda}}\int \exp\left[-\left(\frac{1}{2\lambda^2}-i\pi\beta\right)\tau^2\right]\exp\left[-\left(-\frac{t}{\lambda^2}+i2\pi\nu\right)\tau\right]\mathrm{d}\tau \\
&= \frac{\exp\left(-t^2/2\lambda^2\right)}{\pi^{1/4}\sqrt{\lambda}}\sqrt{\frac{2\pi\lambda^2(1+i2\pi\lambda^2\beta)}{1+(2\pi\lambda^2\beta)^2}}\exp\left[\frac{2\lambda^2(1+i2\pi\lambda^2\beta)}{1+(2\pi\lambda^2\beta)^2}\left(\frac{t-i2\pi\lambda^2\nu}{2\lambda^2}\right)^2\right] \\
&= \frac{\sqrt{2\pi}\lambda}{\pi^{1/4}\sqrt{\lambda}}\sqrt{\frac{1+i2\pi\lambda^2\beta}{1+(2\pi\lambda^2\beta)^2}}\exp\left[\frac{(1+i2\pi\lambda^2\beta)}{1+(2\pi\lambda^2\beta)^2}\frac{(t-i2\pi\lambda^2\nu)^2}{2\lambda^2}-\frac{t^2}{2\lambda^2}\right] \\
&= \sqrt{2\lambda}\pi^{1/4}\sqrt{\frac{1+i2\pi\lambda^2\beta}{1+(2\pi\lambda^2\beta)^2}} \\
&\quad\cdot\exp\left[\frac{(1+i2\pi\lambda^2\beta)(t^2-i4\pi\lambda^2\nu t-(2\pi\lambda^2)^2\nu^2)-t^2-(2\pi\lambda^2)^2(\beta t)^2}{2\lambda^2(1+(2\pi\lambda^2\beta)^2)}\right] \\
&= \sqrt{2\lambda}\pi^{1/4}\sqrt{\frac{1+i2\pi\lambda^2\beta}{1+(2\pi\lambda^2\beta)^2}} \\
&\quad\cdot\exp\left[-\frac{2\pi}{2}\frac{2\pi\lambda^2(\nu-\beta t)^2+i[2\nu t-\beta t^2+(2\pi\lambda^2)^2\beta\nu^2]}{1+(2\pi\lambda^2\beta)^2}\right]
\end{aligned} \tag{B.20}$$

$$
\begin{aligned}
F_s^{\mathcal{T}h}(t,\nu) &= \frac{\exp\left(-t^2/2\lambda^2\right)}{\pi^{1/4}\sqrt{\lambda}} \\
&\qquad \cdot \int \exp\left[-\left(\frac{1}{2\lambda^2} - i\pi\beta\right)\tau^2\right] \exp\left[-\left(-\frac{t}{\lambda^2} + i2\pi\nu\right)\tau\right](\tau - t)\,\mathrm{d}\tau \\
&= -\frac{(-t + i2\pi\lambda^2\nu)(1 + i2\pi\lambda^2\beta)}{1 + (2\pi\lambda^2\beta)^2} F_s^h(t,\nu) - t \cdot F_s^h(t,\nu) \\
&= -\left[\frac{(-t + i2\pi\lambda^2\nu)(1 + i2\pi\lambda^2\beta)}{1 + (2\pi\lambda^2\beta)^2} + t\right] \cdot F_s^h(t,\nu) \\
&= -\left[\frac{-t + i2\pi\lambda^2\nu - i2\pi\lambda^2\beta t - (2\pi\lambda^2)^2\beta\nu + t + (2\pi\lambda^2\beta)^2 t}{1 + (2\pi\lambda^2\beta)^2}\right] \cdot F_s^h(t,\nu) \\
&= -\left[\frac{2\pi\lambda^2(-2\pi\lambda^2\beta + i)(\nu - \beta t)}{1 + (2\pi\lambda^2\beta)^2}\right] \cdot F_s^h(t,\nu)
\end{aligned}
\tag{B.21}
$$

$$
S_s^h(t,\nu) = \frac{2\sqrt{\pi}\lambda}{\sqrt{1 + (2\pi\lambda^2\beta)^2}} \exp\left(-\frac{(2\pi\lambda)^2}{1 + (2\pi\lambda^2\beta)^2}(\nu - \beta t)^2\right) \tag{B.22}
$$

$$
r_0 = -\frac{\sqrt{2}\,2\pi\lambda}{1 + (2\pi\lambda^2\beta)^2}\,(\nu - \beta t)\,(-2\pi\lambda^2\beta + i) \tag{B.23}
$$

$$
|r_0|^2 = \frac{2\,(2\pi\lambda)^2}{1 + (2\pi\lambda^2\beta)^2}(\nu - \beta t)^2 \tag{B.24}
$$

$$
NS_s^h(t_n,\nu_n) = \frac{2\sqrt{\pi}\lambda}{\sqrt{1 + (2\pi\lambda^2\beta)^2}} \exp\left(-\frac{1}{2} \cdot \frac{(\nu_n - 2\pi\lambda^2\beta t_n)^2}{1 + (2\pi\lambda^2\beta)^2}\right) \tag{B.25}
$$

$$
NS_s^h(r_0) = \frac{2\sqrt{\pi}\lambda}{\sqrt{1 + (2\pi\lambda^2\beta)^2}} \exp\left(-\frac{|r_0|^2}{2}\right) = NS_s^h(0) \cdot \exp\left(-\frac{|r_0|^2}{2}\right) \tag{B.26}
$$

B.4 Gaussian

$$s(t) = \exp\left(-\frac{t^2}{2\sigma^2}\right) \tag{B.27}$$

$$\begin{aligned}
F_s^h(t,\nu) &= \frac{1}{\pi^{1/4}\sqrt{\lambda}} \int \exp\left(-\frac{\tau^2}{2\sigma^2}\right) \exp\left(-\frac{(\tau-t)^2}{2\lambda^2}\right) \exp(-i2\pi\nu\tau)\,\mathrm{d}\tau \\
&= \frac{\exp\left(-t^2/2\lambda^2\right)}{\pi^{1/4}\sqrt{\lambda}} \int \exp\left(-\frac{\tau^2}{2\sigma^2}\right) \exp\left(-\frac{\tau^2}{2\lambda^2}\right) \exp\left(\frac{t\tau}{\lambda^2}\right) \exp(-i2\pi\nu\tau)\,\mathrm{d}\tau \\
&= \frac{\exp\left(-t^2/2\lambda^2\right)}{\pi^{1/4}\sqrt{\lambda}} \int \exp\left[-\frac{\sigma^2+\lambda^2}{2\sigma^2\lambda^2}\tau^2\right] \exp\left[\left(\frac{t}{\lambda^2} - i2\pi\nu\right)\tau\right]\mathrm{d}\tau \\
&\overset{(A.2)}{=} \frac{\exp\left(-t^2/2\lambda^2\right)}{\pi^{1/4}\sqrt{\lambda}} \frac{\sqrt{2\pi}\sigma\lambda}{\sqrt{\sigma^2+\lambda^2}} \exp\left(\frac{1}{4}\left[\frac{t}{\lambda^2} - i2\pi\nu\right]^2 \cdot \frac{2\sigma^2\lambda^2}{\sigma^2+\lambda^2}\right) \\
&= \frac{\sqrt{2\lambda}\sigma\pi^{1/4}}{\sqrt{\sigma^2+\lambda^2}} \exp\left(-\frac{t^2}{2\lambda^2}\right) \exp\left(\frac{1}{2}\left[\frac{t^2}{\lambda^4} - i\frac{4\pi\nu t}{\lambda^2} - (2\pi\nu)^2\right] \cdot \frac{\sigma^2\lambda^2}{\sigma^2+\lambda^2}\right) \\
&= \frac{\sqrt{2\lambda}\sigma\pi^{1/4}}{\sqrt{\sigma^2+\lambda^2}} \exp\left(-\frac{t^2}{2}\left[\frac{1}{\lambda^2} - \frac{\sigma^2}{\lambda^2(\sigma^2+\lambda^2)}\right]\right) \exp\left(-i\frac{2\pi\sigma^2\nu t}{\sigma^2+\lambda^2}\right) \\
&\quad \cdot \exp\left(-\frac{1}{2}\frac{\sigma^2\lambda^2}{\sigma^2+\lambda^2}(2\pi\nu)^2\right) \\
&= \frac{\sqrt{2\lambda}\sigma\pi^{1/4}}{\sqrt{\sigma^2+\lambda^2}} \exp\left(-\frac{1}{2}\left[\frac{t^2+(\sigma\lambda 2\pi\nu)^2}{\sigma^2+\lambda^2}\right]\right) \exp\left(-i\frac{\sigma^2 2\pi\nu t}{\sigma^2+\lambda^2}\right)
\end{aligned} \tag{B.28}$$

$$\begin{aligned}
F_s^{\mathcal{T}h}(t,\nu) &= \frac{\exp\left(-t^2/2\lambda^2\right)}{\pi^{1/4}\sqrt{\lambda}} \int \exp\left[-\frac{\sigma^2+\lambda^2}{2\sigma^2\lambda^2}\tau^2\right] \exp\left[\left(\frac{t}{\lambda^2} - i2\pi\nu\right)\tau\right](\tau-t)\,\mathrm{d}\tau \\
&\overset{(A.3)}{=} \left(\left[\frac{t}{\lambda^2} - i2\pi\nu\right]\frac{\sigma^2\lambda^2}{\sigma^2+\lambda^2}\right) \cdot F_s^h(t,\nu) - t \cdot F_s^h(t,\nu) \\
&= \left(\frac{1}{\sigma^2+\lambda^2}[\sigma^2 t - \sigma^2 t - \lambda^2 t - i\sigma^2\lambda^2 2\pi\nu]\right) \cdot F_s^h(t,\nu) \\
&= -\left(\frac{\lambda^2}{\sigma^2+\lambda^2}[t + i\sigma^2 2\pi\nu]\right) \cdot F_s^h(t,\nu)
\end{aligned} \tag{B.29}$$

$$S_s^h(t,\nu) = \frac{2\sqrt{\pi}\lambda\sigma^2}{\lambda^2+\sigma^2}\exp\left(-\frac{t^2+(2\pi\lambda\sigma\nu)^2}{\lambda^2+\sigma^2}\right) \tag{B.30}$$

$$r_0 = -\frac{\sqrt{2}\lambda}{\lambda^2+\sigma^2}(t + i2\pi\sigma^2\nu) \tag{B.31}$$

In the case when $\sigma = \lambda$, i.e., the matched window case:

$$S_s^h(t,\nu) = \lambda\sqrt{\pi}\exp\left(-\frac{t^2+(2\pi\lambda^2\nu)^2}{2\lambda^2}\right) \tag{B.32}$$

$$r_0 = -\frac{1}{\sqrt{2}\lambda}(t + i2\pi\lambda^2\nu) \tag{B.33}$$

$$|r_0|^2 = \frac{1}{2\lambda^2}\left(t^2+(2\pi\lambda^2\nu)^2\right) \tag{B.34}$$

$$NS_s^h(t_n,\nu_n) = \lambda\sqrt{\pi}\exp\left(-\frac{t_n^2+\nu_n^2}{4}\right) \tag{B.35}$$

Appendix C

Popular science summary

This thesis explores and improves the reassignment method, used for analyzing signals that change over time. With our newfound knowledge, a special smoothing can be applied, enhancing its performance.

Signals varying over time, so called non-stationary signals, appear all around us. From the changing temperature of our globe to the response from a complex radar system. The need for analyzing them arises practically everywhere, and is ever so challenging. A fundamental tool in this field of research, called time-frequency analysis, is the spectrogram. It takes the signal and decomposes it into its frequency content at any given point in time.

A problem with the spectrogram is that it is not very precise. Let's say we want to analyze some bird chirps. Perhaps we want to know both when the chirp happened, and what pitch it had. This corresponds to localizing the signal in time and in frequency respectively. Unfortunately, we can not obtain sharp measures of both simultaneously. Either the time or frequency will, to some extent, appear to be smeared out.

This is what the so called reassignment method seek to compensate. For some signals, we can obtain a perfect concentration of a signal both in time and frequency. In other words, it allows us to clearly see both when and with what pitch a bird has chirped, which is great! However, the reassignment method comes with a drawback – it is quite sensitive to noise and disturbances.

If there is some way to work around this, we would get the best of two worlds. And this is exactly what our thesis project is all about. We present a method that mitigates the effects that noise has on the reassignment. To get there, we had to build up a thorough understanding of the method.

In some sense, the (Gabor) spectrogram can be seen as normal distributed. However, we saw that this is actually not the case for the reassigned spectrogram. Actually, it turns out to be of infinite variance (a measure of spread), which might seem counterintuitive. This is not as bad as it seems, and we suggest that the reassigned spectrogram can be roughly seen as normal distributed as well.

Knowing how the reassigned spectrogram behaves, we had an idea on how to modify it to make it more robust against noise. By smoothing it corresponding to its theoretical distribution, its behavior becomes much more predictable. Our method is basically a way to sacrifice some concentration for the gain of robustness. The concept is similar to kernel density estimation or kernel smoothing. We call it the smoothed reassigned spectrogram, or SRS for short.

References

[1] Kunihiko Kodera, Claude De Villedary, and Roger Gendrin. "A new method for the numerical analysis of non-stationary signals." In: *Physics of the Earth and Planetary Interiors* 12.2 (1976), pp. 142–150.

[2] Francois Auger and Patrick Flandrin. "Improving the readability of time-frequency and time-scale representations by the reassignment method." In: *IEEE Transactions on Signal Processing* 43.5 (May 1995), pp. 1068–1089.

[3] Ingrid Daubechies, Jianfeng Lu, and Hau-Tieng Wu. "Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool." In: *Applied and computational harmonic analysis* 30.2 (2011), pp. 243–261.

[4] Maria Sandsten, Johan Brynolfsson, and Isabella Reinhold. "The Matched Window Reassignment." In: *2018 26th European Signal Processing Conference (EUSIPCO)*. Sept. 2018, pp. 2340–2344.

[5] Eric Chassande-Mottin, Patrick Flandrin, and Francois Auger. "Statistique Des Vecteurs De Reallocation Du Spectrogramme." In: (Jan. 1996).

[6] Maria Sandsten. "Time-frequency analysis of time-varying signals and non-stationary processes." In: *Lund University* (2018). URL: `http://www.maths.lu.se/fileadmin/maths/personal_staff/mariasandsten/TFkompver2.pdf`.

[7] Leon Cohen. *Time-frequency analysis*. Vol. 778. Prentice hall, 1995.

[8] Maria Hansson-Sandsten and Johan Brynolfsson. "The Scaled Reassigned Spectrogram with Perfect Localization for Estimation of Gaussian Functions." In: *IEEE Signal Processing Letters* 22.1 (2015), pp. 100–104.

[9] Eric Chassande-Mottin, Patrick Flandrin, and Francois Auger. "On the Statistics of Spectrogram Reassignment Vectors." In: *Multidimensional Systems and Signal Processing* 9.4 (Oct. 1998), pp. 355–362.

[10] Patrick Flandrin. *Explorations in Time-frequency Analysis*. Cambridge University Press, 2018.

[11] Johan Brynolfsson and Maria Sandsten. "Parameter estimation of Oscillating Gaussian functions using the scaled reassigned spectrogram." In: *Signal Processing* 150 (2018), pp. 20–32.

[12] Eric W. Weisstein. *Pearson Mode Skewness. From MathWorld—A Wolfram Web Resource*. URL: `http://mathworld.wolfram.com/PearsonModeSkewness.html` (visited on 03/20/2019).

[13] Friedrich Liese and Igor Vajda. "On Divergences and Informations in Statistics and Information Theory." In: *IEEE Transactions on Information Theory* 52.10 (Oct. 2006), pp. 4394–4412.

[14] Solomon Kullback and Richard A Leibler. "On information and sufficiency." In: *The annals of mathematical statistics* 22.1 (1951), pp. 79–86.

[15] Ernst Hellinger. "Neue begründung der theorie quadratischer formen von unendlichvielen veränderlichen." In: *Journal für die reine und angewandte Mathematik* 136 (1909), pp. 210–271.
[16] Eric W. Weisstein. *Point-Line Distance – 2-Dimensional. From MathWorld—A Wolfram Web Resource*. URL: `http://mathworld.wolfram.com/Point-LineDistance2-Dimensional.html` (visited on 04/27/2019).
[17] Kelly R. Fitz and Sean A. Fulop. "A Unified Theory of Time-Frequency Reassignment." In: *CoRR* abs/0903.3080 (2009).
[18] Paul Bromiley. "Products and convolutions of Gaussian probability density functions." In: *Tina-Vision Memo* 3.4 (2003), p. 1.
[19] Vassiliy A Epanechnikov. "Non-parametric estimation of a multivariate probability density." In: *Theory of Probability & Its Applications* 14.1 (1969), pp. 153–158.
[20] Éric Chassande-Mottin, François Auger, and Patrick Flandrin. "Reassignment." In: *Time-Frequency Analysis: Concepts and Methods* (2008), pp. 249–277.