Optimal Multitaper Spectrograms

Markus Ydreskog

Master's thesis
2023:E21

Lund University

Faculty of Engineering
Centre for Mathematical Sciences
Mathematical Statistics

Centrum Scientiarum Mathematicarum

Master's Theses in Mathematical Sciences 2023:E21
ISSN 1404-6342

LUTFMS-3472-2023

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

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

Optimal Multitaper Spectrograms

Master's thesis

Markus Ydreskog

Supervisor: Maria Sandsten
Assisting supervisor: Oskar Keding

Spring 2023

Abstract

Multitaper spectrograms have been proposed as a method of improving the spectrogram as a time-frequency representation (TFR). This thesis aimed to investigate both previously used and new methods for combining multitaper spectrograms of a Gaussian signal and a chirp. More specifically, the use of new measures to optimize the weights of a weighted sum of spectrograms was tested and evaluated. The optimization problems were first stated independently of the signal and then solved for a Gaussian and chirp signal. A theorem for the spectrogram of a Hermite function was proved. This allowed an exact solution to be obtained for the least squares optimization between the multitaper spectrogram and the Wigner distribution of a Gaussian signal. The theorem was also used to extend the method used for the Gaussian signal for a chirp based on the use of Hermite expansion. The optimized weights were then evaluated based on concentration and sensitivity to noise.

The results in the case of the Gaussian signal showed that some of the new methods gave improved concentration in comparison to the Wigner distribution and other commonly used methods. The solution derived for the least squares optimization was also seen to be much faster in computation compared to numerical methods. Furthermore, the optimized weights were less sensitive to noise than the Wigner distribution. The results for the chirp signal showed that the new method based on Hermite expansion gave improved results compared to using the weights calculated for the Gaussian signal.

The conclusion was that the new optimization methods were able to not only improve the localization of a time-frequency representation but also give desirable qualities such as non-negativity. This means that one could choose which method to use based on desired properties whether it be concentration or robustness to noise. The conclusion was also that the new method used for the chirp signal gave improved results and could possibly be extended to other signals not covered in the thesis. Further research could also be done pertaining to each of the new optimization methods used in the thesis.

Acknowledgment

This Master's thesis was written during the Spring of 2023 at LTH as part of the Engineering physics program. I want to thank my supervisor Maria Sandsten and assisting supervisor Oskar Keding for invaluable feedback and help. They also provided more meetings than what anyone could hope for which is greatly appreciated.

I would also like to thank my dog Nora for giving me well-needed play breaks during writing. Finally, I also want to thank anyone who feels they should be thanked.

1. Introduction

Time-frequency analysis concerns the study of a joint representation in time and frequency of a signal. The idea is that neither only the frequency content nor the temporal information sufficiently characterizes a signal. An analogy can be drawn to a piece of music where of course you need both the notes you play and at what time to play them to accurately describe it. Given the popularity of image-generating AI like Stable Diffusion a prime motivating example for the use of time-frequency analysis is the recent model called Riffusion [10]. This model used the image-based networks to generate spectrograms which is a type of time-frequency representation (TFR) that was then converted to audio allowing someone to create different songs using only prompts. Another more classical example of the use of time-frequency analysis is in radar where another TFR called the cross ambiguity function is used to determine the distance and velocity of a target [13]. A final example is the use of time-frequency analysis in medicine and specifically for analyzing electrical signals of the heart called Electrocardiograph (ECG) where yet another TFR called the Wigner distribution has been used [1]. The number of applications to time-frequency analysis is many and likewise the number of different representations.

So far a few different TFRs have already been encountered, each with its own advantages and disadvantages. The spectrogram although fast and easy to compute is not the most localized time-frequency representation and also depends on the choice of a so-called window. The Wigner distribution, however, does not depend on a choice of window and is more localized in time and frequency. On the other hand, the Wigner distribution is slower to compute and can introduce unwanted components in the time-frequency plane called cross-terms. The idea of multitaper spectrograms is to deal with some of the shortcomings of both of these representations providing a more localized representation than the spectrogram while reducing cross-terms, suppressing noise, and being computationally faster than the Wigner distribution.

The multitaper spectrogram uses a set of multiple windows for the spectrogram instead of one. It is therefore again dependent on a choice of windows but now also on how one should combine the spectrograms. One approach used in [18] for a Gaussian signal was optimizing the weights of a weighted sum of spectrograms against the Wigner distribution by minimizing the squared error. Another approach used in [16] was instead based on the calculus of variations and finding the function closest to the set of spectrograms under some measure. A common choice for the windows used in [18] are the Hermite functions which leave only the choice of weights.

In this thesis, the approach based on the use of optimization to calculate the weights of a weighted sum of spectrograms is investigated. More specifically the choice of different optimization criteria as opposed to only the squared error is investigated first for the commonly encountered Gaussian signal. A method of extending this to another common signal called a chirp is then proposed. The optimized weights in both of these cases are then evaluated based on commonly desired features such as noise reduction and concentration. The desired result is a method of combining multitaper spectrograms with equal or better concentration than the Wigner distribution while being less sensitive to noise.

The thesis is structured as follows: Section 2 provides a short theoretical background to time-frequency analysis and other theories needed. Section 3 introduces the different optimization criteria that are investigated. Section 4 derives expressions and numerically calculates the weights for the case of a Gaussian signal and evaluates them based on concentration and noise. Section 5 derives a method and weights of a signal consisting instead of a chirp and evaluates the weights using the same criteria. Finally, section 6 provides further discussion of the results and concludes the thesis.

Contents

2. Theoretical Background

2.1 Time Frequency Analysis

For simplicity, we restrict ourselves to 1-dimensional signals $x(t)$ where $t \in \mathbb{R}$ although all of the formulas generalize to higher dimensions. All integrals are taken over the whole domain of $\mathbb{R}^d$ unless explicitly stated. The short-time Fourier transform (STFT) for a signal $x(t)$ and a fixed function $g(t) \neq 0$ called a window is defined as

$$(V_g x)(t, f) = \int x(\tau) g^*(\tau - t) e^{-2\pi i \tau f} d\tau \tag{1}$$

with $x, g \in \mathbb{S}'(\mathbb{R})$ and $V_g x \in \mathbb{S}'(\mathbb{R}^2)$ where $\mathbb{S}'(\mathbb{R}^d)$ is the space of tempered distributions [12]. The STFT can be thought of as sliding the window g over the signal and taking the Fourier transform at each time. The spectrogram which is a type of time-frequency representation (TFR) is defined as the squared absolute value of the STFT i.e

$$(S_g x)(t, f) = |V_g x|^2 \tag{2}$$

The spectrogram has several useful properties such as being non-negative, real-valued, time and frequency shift invariant, and energy preserving if $\|g\|_2^2 = 1$. Time and frequency shift invariant mean that if the signal is frequency modulated and time-shifted, the TFR is simply shifted correspondingly. This can be written as

$$(S_g T_{t_0} M_{f_0} x)(t, f) = (S_g x)(t - t_0, f - f_0)$$

where M_{f_0} denotes a frequency modulation $M_{f_0} x(t) = e^{2\pi i f_0 t} x(t)$ and T_{t_0} a time shift $T_{t_0} x(t) = x(t - t_0)$. The effect of frequency modulation is shifting the signal in frequency by f_0 in the frequency domain and will thus sometimes be called a frequency shift instead. From the definitions above it is clear that the time-frequency representation of the STFT and thus also the spectrogram depends on the window $g(t)$. The time resolution in the time-frequency plane will depend on the essential support or length of the window $g(t)$ where a short window gives better time resolution [13]. Similarly the frequency resolution will depend on the bandwidth of the window in the frequency domain. However, due to the uncertainty principle, it is not possible to create a window that is both arbitrarily localized in time and frequency. Hence there is a trade-off when choosing the window whether good frequency or time resolution is desired.

Another time-frequency representation is the Wigner distribution defined for a signal as

$$(Wx)(t, f) = \int x(t + \frac{\tau}{2}) x^*(t - \frac{\tau}{2}) e^{-2\pi i f \tau} d\tau \tag{3}$$

where $x(t) \in \mathbb{S}'(\mathbb{R})$ and $Wx \in \mathbb{S}'(\mathbb{R}^2)$. The Wigner distribution like the spectrogram is real-valued, time and frequency shift invariant and energy preserving but has even more useful properties that can be found in [13]. Unlike the spectrogram, the Wigner distribution does not depend on choosing a window and provides better localization than the spectrogram.

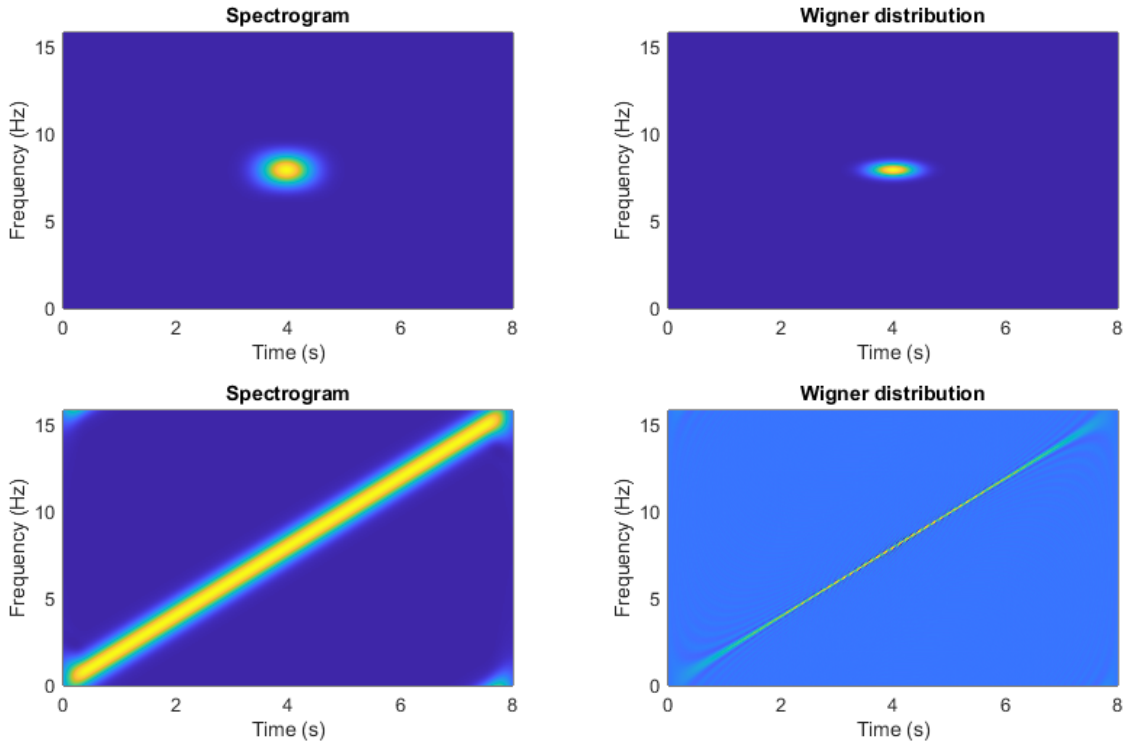

Figure 1: Spectrograms and Wigner distribution of a Gaussian signal (upper) and a chirp (lower)

In Figure 1 one can see the much better localization of the Wigner distribution compared to the spectrogram for both a Gaussian signal and a chirp.

Both the spectrogram and the Wigner distribution are members of the so-called quadratic class of TFRs. The quadratic class can be described by a sesquilinear form $G(x, g)$ meaning that it is linear in the first argument and conjugate linear in the second. The two ways to make G quadratic in x is either taking $|G(x, g)|^2$ such as for the spectrogram or $G(x, x)$ as in the Wigner distribution [12]. Another quadratic TFR of the form $G(x, x)$ is the ambiguity function defined as

$$(Ax)(\nu, \tau) = \int x(t + \frac{\tau}{2})x^*(t - \frac{\tau}{2})e^{-2\pi i t\nu}dt \tag{4}$$

which is usually written in ν and τ with the interpretation as a time-frequency correlation function or for its application in radar as range τ and Doppler-shift ν. The cross Wigner and cross ambiguity functions between two signals x and y is then defined as

$$(W_y x)(t, f) = \int x(t + \frac{\tau}{2})y^*(t - \frac{\tau}{2})e^{-2\pi i f\tau}d\tau$$
$$(A_y x)(\nu, \tau) = \int x(t + \frac{\tau}{2})y^*(t - \frac{\tau}{2})e^{-2\pi i t\nu}dt$$

The quadratic form $Q(x)$ ($|G(x, g)|^2$ or $G(x, x)$) is easily seen to follow the superposition rule

$$\Big\{Q(\alpha x_1 + \beta x_2) = |\alpha|^2 Q(x_1) + \alpha\beta^* G(x_1, x_2) + \alpha^*\beta G(x_2, x_1) + |\beta|^2 Q(x_2) \tag{5}$$

and it is evident that a superposition of two signals will give cross-terms $\alpha\beta^* G(x_1, x_2)$ and $\alpha^*\beta G(x_2, x_1)$ and auto terms $|\alpha|^2 Q(x_1)$ and $|\beta|^2 Q(x_2)$. Continuing the argument a superposition of N signals will give $\binom{N}{2}$ cross-terms between each signal component. However, although all quadratic representations will give cross-terms the effect will vary. For the spectrogram, the cross-terms will only appear in regions where the auto terms are overlapping meaning that if the two auto components are well separated no cross-terms will appear [13]. The

Wigner distribution on the other hand will always give rise to cross-terms appearing in between the auto-terms, while the cross-terms for the ambiguity function will generally be located away from the origin.

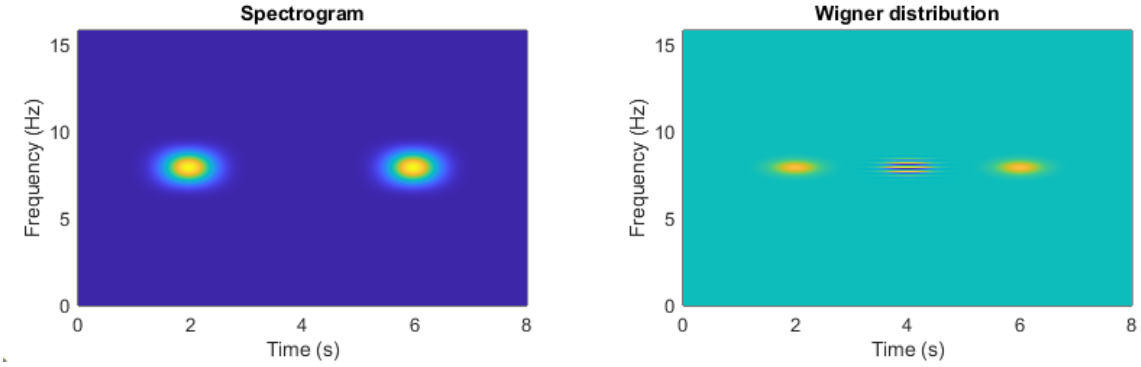

Figure 2: Spectrogram and Wigner distribution of a two-component signal consisting of two Gaussian functions located at different times.

In Figure 2, one can see that the spectrograms show no cross terms for the two well-separated components while the cross-terms of the Wigner distribution are clearly visible in between the two components. In [13], it is shown that there is a trade-off on how localized the auto-terms will be and how much interference is caused by the cross-terms.

Ideally, one would want a time-frequency representation as localized as the Wigner distribution while reducing the interference of the cross-terms. One way of doing this is a smoothing of the Wigner distribution, which brings us to Cohen's class of TFRs. A TFR $(Cx)(t,f)$ belonging to Cohen's class is given by the Wigner distribution smoothened with some kernel ψ, written as [13]

$$(Cx)(t,f) = Wx * \psi = \iint \psi(t-u, f-v)(Wx)(u,v)dudv \tag{6}$$

The Wigner distribution is trivially part of Cohen's class with the kernel $\psi = \delta$, where δ is the Dirac delta function, while the spectrogram is associated with the kernel $\psi(t,f) = g(-t-\frac{f}{2})g^*(-t+\frac{f}{2})$. As the ambiguity function is the Fourier transform of the Wigner distribution and the convolution operator is transformed into multiplication in the Fourier domain an equivalent definition of (6) is

$$(Cx)(t,f) = \iint \Psi(\tau,\nu)A(\tau,\nu)e^{2\pi i(t\nu - f\tau)}d\tau d\nu \tag{7}$$

where $\Psi(\tau,\nu)$ is the two-dimensional Fourier transform of $\psi(t,f)$. An advantage of the definition in (7) is that we know from earlier that the cross-terms of the ambiguity function were generally located far away from the origin so we can design the kernel $\Psi(\tau,\nu)$ based on this. Properties such as time-frequency shift invariance follow from properties of the convolution and of the Wigner distribution.

2.2 Stochastic Processes

So far only the case of a deterministic signal has been discussed. A stochastic process $x(t)$ is defined as weakly stationary if

$$\begin{cases} E[x(t)] = \mu \\ E[(x(t)-\mu)(x^*(s)-\mu)] = \gamma_x(t-s) = \gamma_x(\tau) \end{cases}$$

meaning that the mean μ is constant over time and the auto-covariance function γ_x is a function of only increments in time $\tau = t - s$ [8]. We will from now on assume $x(t)$ to be zero mean i.e. $\mu = 0$ without loss of generality. For a weakly stationary signal one can then define the Power Spectral Density (PSD) as

$$\Gamma_x(f) = \int \gamma_x(\tau)e^{-2\pi i f\tau}d\tau \tag{8}$$

which under certain conditions is guaranteed to exist (see [15]). The PSD can be thought of as a distribution of the variance or power at different frequencies and for a deterministic signal it reduces to the squared Fourier transform. Furthermore, any stochastic process $x(t)$ that is weakly stationary can be decomposed as

$$x(t) = \int e^{2\pi i f t} dX(f) \tag{9}$$

where the integral is taken as a Stieltjes integral [8]. From the decomposition, one finds the orthogonality relations

$$\begin{aligned} \int e^{2\pi i f_1 t} e^{2\pi i f_2 t} dt &= \delta(f_1 - f_2) \\ E[dX(f_1) dX^*(f_2)] &= \delta(f_1 - f_2)\Gamma_x(f_1) df_1 df_2 \end{aligned} \tag{10}$$

meaning that both the decomposing functions $e^{2\pi i f t}$ and spectral increments $dX(f)$ are orthogonal with respect to their corresponding inner products. The two main ideas of generalizing a spectral density for non-stationary signals lie in changing either of the orthogonality relations in (10). Starting by making the spectral increments correlated leads to the class of harmonizable signals with

$$\begin{aligned} E[dX(f_1) dX(f_2)] &= \Phi_x(f_1, f_2) df_1 df_2 \\ \iint |\Phi_x(f_1, f_2)| df_1 df_2 &< \infty \end{aligned} \tag{11}$$

where $\Phi_x(f_1, f_2)$ is called the Spectral Distribution Function. The auto-covariance function and the spectral distribution function form a Fourier-transform pair like the PSD and the auto-covariance function in the stationary case meaning

$$E[x(t)x^*(s)] = \iint \Phi_x(f_1, f_2) e^{2\pi i (f_1 t - f_2 s)} df_1 df_2 \tag{12}$$

When the stochastic process is weakly stationary $\Phi_x(f_1, f_2) = \delta(f_1 - f_2)\Gamma_x(f_1)$ and (12) coincides with the normal definition of the PSD. The other possibility is to change the decomposing function in (9) leading to

$$\begin{aligned} x(t) &= \int \psi(t, f) dX(f) \\ \int \psi(t, f_1)\psi^*(t, f_2) dt &= \delta(f_1 - f_2) \end{aligned} \tag{13}$$

which amounts to finding the eigenfunctions $\psi(t, f)$ of the autocovariance kernel. From (13) one finds the auto-covariance function

$$\begin{aligned} \gamma_x(t, s) &= \iint \psi_x(t, f_1)\psi_x^*(t, f_2) E[dX(f_1) dX(f_2)] = \iint \psi_x(t, f_1)\psi_x^*(t, f_2)\delta(f_1 - f_2)\Gamma_x(f_1) df_1 df_2 \\ &= \int \psi_x(t, f_1)\psi_x^*(t, f_1)\Gamma_x(f_1) df_1 \end{aligned}$$

leading to

$$\int \gamma_x(t, s)\psi_x(s, f) ds = \Gamma_x(f)\psi_x(t, f)$$

and from the relation

$$E[x(t)x^*(t)] = \int |\psi_x(t, f)|^2 \Gamma_x(f) df$$

one can find the time dependent power spectrum $|\psi(t, f)|^2\Gamma_x(f)$ [8]. Although this representation seems to extend the power spectral density for the stationary process well, the interpretation of f as a frequency is lost since the sinusoids $e^{2\pi i f t}$ used for the decomposition are replaced by the general function $\psi_x(t, f)$. As the purpose of time-frequency analysis is to find a joint representation in time and frequency the first extension of harmonizable signals is considered. From the class of harmonizable signals, one finds the Wigner-Ville spectrum (WVS)

$$\begin{aligned} (\mathbf{W}x)(t, f) &= \int \Phi_x(f - \frac{\nu}{2}, f + \frac{\nu}{2}) e^{-2\pi i \nu t} d\nu \\ &= \int \gamma_x(t + \frac{\tau}{2}, t - \frac{\tau}{2}) e^{-2\pi i f \tau} d\tau \end{aligned} \tag{14}$$

which reduces to the PSD for each time t. If the signal is deterministic, the WVS also becomes the regular Wigner distribution. As noted in [21], the WVS under mild conditions is the mean of the Wigner distribution of the realizations, i.e,

$$(\mathbf{W}x)(t,f) = E[(Wx)(t,f)]$$

and if one assumes some local stationarity the class of Cohen's TFRs defined in (6) can be considered an estimate for the WVS. Since the spectrogram belongs to Cohen's class it is thus a form of estimate of the WVS.

2.3 Multitapers

In the stationary case, a well-known method of reducing the variance without sacrificing the bias of the periodogram is Thomson's multitaper method [20]. This gives the inspiration to apply the same idea for the spectrogram as an estimate of the WVS. Thomson's multitaper method uses the prolate spheroidal wave functions as windows as they correspond to the eigenfunctions of an operator projecting onto a time and frequency band [6]. So for a time-limited signal, these are the functions with the highest energy concentration for a given frequency band and provide an orthonormal set of windows. In time-frequency analysis, since the signals of interest are non-stationary, one seeks an optimal concentration in both time and frequency. The eigenfunctions of an operator instead corresponding to a projection onto circular discs in the time-frequency plane are the Hermite functions. These functions can be defined in different ways and for convenience later we here define them following [9] as

$$\begin{aligned} h_n(t) &= \frac{2^{1/4}}{\sqrt{2^n \cdot n!}} \cdot \tilde{h}_n(\sqrt{2\pi}t) \\ \tilde{h}(t) &= e^{-\frac{t^2}{2}} H_n(t) \end{aligned} \tag{15}$$

where $H_n(t)$ are the Hermite polynomials defined as [2]

$$H_n(t) = (-1)^n e^{t^2} \frac{d^n e^{-t^2}}{dt^n} \tag{16}$$

The Hermite functions instead provide an orthonormal basis that is the most localized on circular regions of the time-frequency plane. The multitaper spectrogram of a given signal using the Hermite functions as windows can then be written as

$$(Sx)(t,f) = \sum_{n=0}^{N-1} w_n (S_{h_n} x)(t,f) \tag{17}$$

where w_n are some weights. The resulting TFR of the multitaper spectrogram will vary a lot depending on the choice of weights. A common choice is the use of the arithmetic mean $\frac{1}{N}$ as weights.

2.4 Entropy and Divergence Measures

The interpretation of a time-frequency representation as a density has led to concepts in statistics and information theory such as entropy and divergence to be used. We here briefly go over some basic concepts that will be used. The Rényi Entropy is a type of entropy that essentially measures the amount of information in the outcomes of a random variable. This entropy is also used to measure the concentration of a TFR and is defined as

$$H_\alpha(C) = \frac{1}{1-\alpha} \log_2 \iint C^\alpha dt df, \quad \alpha \geq 2 \tag{18}$$

where C is a TFR and α some parameter [19]. The Rényi entropy assumes larger values for less concentrated TFR. The effect of the parameter α has been studied in detail in [4] with the value $\alpha = 3$ often being used. To get a TFR as concentrated as possible, we want to have a low value for the Rényi entropy.

Another measure of concentration for a TFR C inspired by the kurtosis used in statistics is defined by [19]

$$\gamma(C) = \frac{\iint C^4 dt df}{(\iint C^2 dt df)^2} \tag{19}$$

The value of γ will be larger for more concentrated TFRs so in contrast to the Rényi entropy we want this measure to be as large as possible. The measure in (19) is however only useful for signals with one component and requires some adjustment for multi-component signals.

Using entropy one can define a measure of similarity between two probability distributions called divergence. A divergence could also be regarded as a type of non-symmetric distance. From the Rényi entropy, one can define the Rényi divergence for two TFRs C_1 and C_2 as [3]

$$D_\alpha(C_1, C_2) = \frac{1}{\alpha - 1} \log_2 \iint C_1^\alpha C_2^{1-\alpha} dtdf \tag{20}$$

The well-known Kullback-Leibler (KL) divergence also known as the relative entropy is obtained as the limit when α goes to one [3]. One can also define a symmetric divergence more akin to a distance from (20) by taking

$$\hat{D}_\alpha(C_1, C_2) = D_\alpha(C_1, C_2) + D_\alpha(C_2, C_1) \tag{21}$$

Using the Shannon entropy which is obtained as $\alpha \to 1$ of the Rényi entropy, another popular distance measure is given by the Jensen difference

$$J(C_1, C_2) = H(\frac{C_1 + C_2}{2}) - \frac{H(C_1) + H(C_2)}{2} \tag{22}$$

which due to the use of the Shannon entropy is only defined for positive TFRs. In [3], the Jensen difference is extended to all TFRs by changing the Shannon entropy H to the Rényi entropy H_α but with the added restriction $\alpha \in (0, 1)$ to ensure concavity of the distance measure. This distance can then be written as

$$J_\alpha(C_1, C_2) = H_\alpha(\frac{C_1 + C_2}{2}) - \frac{H_\alpha(C_1) + H_\alpha(C_2)}{2}, \quad \alpha \in (0, 1) \tag{23}$$

In [17], a distance measure based on the Jensen difference using the Rényi entropy is instead defined as

$$\hat{J}_\alpha(C_1, C_2) = H_\alpha(\sqrt{C_1 C_2}) - \frac{H_\alpha(C_1) + H_\alpha(C_2)}{2} \tag{24}$$

which removes the restriction on α but is only defined for positive TFRs. If the TFR is positive, this measure might be better due to allowing the use of α shown to be good for the Rényi entropy.

3. Choosing multitaper weights

Choosing the weights of the multitaper spectrogram in (17) is not an easy task and using equal weighting might seem natural. As mentioned earlier another way of choosing weights was by minimizing some optimization criteria. In this section, some previous work for the choice of weights is presented followed by the new optimization methods evaluated in the thesis.

3.1 Previous work

As shown in [19],[5], any TFR belonging to Cohen's class can be given as a multitaper spectrogram. One possible way to calculate the weights of the multitaper spectrogram defined in (17) is using optimization. For the simpler case of a unit energy Gaussian signal $x(t) = 2^{\frac{1}{4}} e^{-\pi t^2}$, the exact weights of an infinite window multitaper spectrogram resulting in the Wigner distribution are calculated in [18]. Since in practice one needs a fixed number of windows, preferably as few as possible, the weights are instead calculated by minimizing the squared error i.e

$$\min_w \quad \|(Wx)(t,f) - \sum_{n=0}^{N-1} w_n (S_{h_n} x)(t,f)\|_2^2 = \min_w \quad \iint \left((Wx)(t,f) - \sum_{n=0}^{N-1} w_n (S_{h_n} x)(t,f) \right)^2 dtdf \tag{25}$$

Another way of choosing weights is presented in [14] where weights for the multitaper spectrogram that weakly approximates an arbitrary Wigner distribution were derived. These weights are for a one-dimensional signal

given by

$$w_n = (-1)^n C_{N-1,n} \tag{26}$$

$$C_{k,j} = \sum_{m=j}^{k} 2^{-m} \binom{m}{j} \tag{27}$$

which were shown to give linear convergence in the distributional sense. These weights will be tested later as a comparison and are denoted as Cw.

In [16], the optimization problem was stated as finding among all positive TFRs the one closest to the set of spectrograms under some measurement. It was shown that under the KL divergence, the arithmetic mean was the optimal combination while the geometric mean was optimal in the cross-entropy sense. The paper was more related to optimally combining spectrograms of different window lengths but it gives an example of a different approach to optimizing for the Wigner distribution. The next section presents some new ways of optimizing the weights based on the work presented here.

3.2 Optimization problems

The Wigner distribution has very good localization, hence being able to represent it with a few multitapers is desirable from a computational aspect and also due to the possible cross-term and variance reduction mentioned in section 2. The optimization problem in (25) naturally leads to the questions of whether weights can be derived for signals other than a Gaussian and whether another optimization criterion such as the divergence measures in section 2.4 would be better. Instead of minimizing the distance in L_2-norm the following set of alternative optimization problems based on divergence are considered. Each optimization problem is given a name for easier referencing during evaluation. The least squares problem earlier in (25) is denoted LS. Using the distance measures based on the Jensen distance from (23) and (24) and the divergence based on Rényi entropy from (20), we define the optimization problems

$$\begin{aligned} \text{Jensen distance (J1):} \quad & \min_{w} \quad J_\alpha(W, \sum_{n=0}^{N-1} w_n h_n), \quad \alpha \in (0,1) \\ & \text{subject to} \quad \sum_{n=0}^{N-1} w_n = 1 \end{aligned} \tag{28}$$

$$\begin{aligned} \text{Jensen distance (J2):} \quad & \min_{w} \quad \hat{J}_\alpha(W, \sum_{n=0}^{N-1} w_n S_{h_n}) \\ & \text{subject to} \quad \sum_{n=0}^{N-1} w_n = 1 \end{aligned} \tag{29}$$

$$\begin{aligned} \text{Rényi divergence (RD):} \quad & \min_{w} \quad D_\alpha(W, \sum_{n=0}^{N-1} w_n S_{h_n}) \\ & \text{subject to} \quad \sum_{n=0}^{N-1} w_n = 1 \end{aligned} \tag{30}$$

The signals are assumed to be normalized so that their energy is 1 and constraints are added to the weights so that the multitaper spectrogram preserves the same energy for all but the least squares.

Another idea explored in this thesis is to not optimize against the Wigner distribution but instead finding the weights that maximize the concentration of the TFR. Two ways of measuring concentration was the Rényi entropy and kurtosis defined in (18) and (19). We want to in this case minimize the Rényi entropy but maximize

the kurtosis. The optimization problems can thus be stated as

$$\begin{aligned} \text{Rényi entropy (RE):} \quad & \min_{w} \quad H_\alpha(\sum_{n=0}^{N-1} w_n S_{h_n}) \\ & \text{subject to} \\ & \sum_{n=0}^{N-1} w_n = 1 \\ & \|w\|_2 \leq \lambda \end{aligned} \tag{31}$$

where an additional constraint on the weights is added for reasons explained later and

$$\begin{aligned} \text{Kurtosis (KT):} \quad & \min_{w} \quad -\gamma(\sum_{n=0}^{N-1} w_n S_{h_n}) \\ & \text{subject to} \quad \sum_{n=0}^{N-1} w_n = 1 \end{aligned} \tag{32}$$

where we have changed the maximization problem for kurtosis to a minimization problem by adding a minus sign. As earlier the optimization problems are given abbreviations (RE, KT).

4. Multitaper spectrogram of a Gaussian signal

The signal considered here is the frequency-modulated and time-shifted unit-energy Gaussian signal with frequency f_0 and time-center t_0

$$x(t) = T_{t_0} M_{f_0} \cdot 2^{\frac{1}{4}} e^{-\pi t^2}$$

The multitaper method is applied to this signal as it is a very commonly encountered signal, consisting of essentially a dampened sinusoid. Since all of the TFRs of Cohen's class are time-frequency shift-invariant, it is however sufficient to only consider the signal shifted to frequency and time 0, i.e,

$$x(t) = 2^{\frac{1}{4}} e^{-\pi t^2} \tag{33}$$

Using the definition, the Wigner distribution of the Gaussian signal is easily calculated to be

$$(Wx)(t,f) = 2e^{-2\pi(t^2+f^2)} \tag{34}$$

An expression of the Hermite windowed spectrogram of a Gaussian is derived in [7]. However, a more general formula for the Hermite windowed spectrogram of any Hermite function of which the Gaussian signal is a special case can be derived using a proof similar to [9] adapted to the STFT. In this case the signal, $x(t) = h_k(t)$ where $h_k(t)$ is the Hermite function defined in (15) and we have

Theorem 1.

$$(V_{h_j} h_k)(t,f) = \begin{cases} e^{\pi i t f} \sqrt{\frac{k!}{j!}} e^{-\frac{\pi}{2}|\omega|^2} (-\sqrt{\pi}\omega^*)^{j-k} L_k^{j-k}(\pi|\omega|^2), & j \geq k \\ e^{\pi i t f} \sqrt{\frac{j!}{k!}} e^{-\frac{\pi}{2}|\omega|^2} (\sqrt{\pi}\omega)^{k-j} L_j^{k-j}(\pi|\omega|^2), & j \leq k \end{cases} \tag{35}$$

Proof. The calculations for the STFT are done in the Fock space (see Appendix A)

$$
\begin{aligned}
(V_{h_j}h_k)(t,f) &= \langle h_k, M_f T_t h_j\rangle_{\mathrm{L}^2} = \langle \xi_k, \beta(\omega)\xi_j\rangle_{\mathcal{F}^2} = \{\omega = t - if\} \\
&= \frac{\sqrt{\pi^{j+k}}}{j!k!} e^{\pi i t f} \int e^{\pi z^*\omega - \frac{\pi}{2}|\omega|^2} ((z-\omega)^*)^j z^k e^{-\pi|z|^2} dz \\
&= e^{\pi i t f} \frac{\sqrt{\pi^{j+k}}}{j!k!} e^{-\frac{\pi}{2}|\omega|^2} \sum_{l=0}^{j} \binom{j}{l} (-1)^{j-l} (\omega^*)^{j-l} \int (z^*)^l z^k e^{-\pi|z|^2 + \pi z^*\omega} dz \\
&= e^{\pi i t f} \frac{\sqrt{\pi^{j+k}}}{j!k!} e^{-\frac{\pi}{2}|\omega|^2} \sum_{l=0}^{j} \binom{j}{l} (-1)^{j-l} (\omega^*)^{j-l} \pi^{-l} \frac{d^l}{d\omega^l} \int z^k e^{-\pi|z|^2 + \pi z^*\omega} dz \\
&= e^{\pi i t f} \frac{\sqrt{\pi^{j+k}}}{j!k!} e^{-\frac{\pi}{2}|\omega|^2} \sum_{l=0}^{j} \binom{j}{l} (-1)^{j-l} (\omega^*)^{j-l} \pi^{-l} \frac{d^l}{d\omega^l} \langle z^k, K_\omega\rangle_{\mathcal{F}^2} \\
&\quad \text{where } K_\omega \text{ is the reproducing kernel of the Fock space meaning } \langle z^k, K_\omega\rangle_{\mathcal{F}^2} = \omega^k \\
&= e^{\pi i t f} \frac{\sqrt{\pi^{j+k}}}{j!k!} e^{-\frac{\pi}{2}|\omega|^2} \sum_{l=0}^{j} \binom{j}{l} (-1)^{j-l} (\omega^*)^{j-l} \pi^{-l} \frac{d^l}{d\omega^l} \omega^k \\
&= e^{\pi i t f} \frac{\sqrt{\pi^{j+k}}}{j!k!} e^{-\frac{\pi}{2}|\omega|^2} \sum_{l=0}^{\min(j,k)} \binom{j}{l} (-1)^{j-l} (\omega^*)^{j-l} \pi^{-l} \frac{k!}{(k-l)!} \omega^{k-l}
\end{aligned}
$$

Consider first the case $j \geq k$ and substitute $l \to k - l$ which gives

$$
\begin{aligned}
(V_{h_j}h_k)(t,f) &= e^{\pi i t f} \sqrt{\frac{\pi^{j+k}}{j!k!}} e^{-\frac{\pi}{2}|\omega|^2} \sum_{l=0}^{k} \binom{j}{k-l} (-\omega^*)^{j-k+l} \pi^{l-k} \frac{k!}{l!} \omega^l \\
&= e^{\pi i t f} \sqrt{\frac{k!}{j!}} e^{-\frac{\pi}{2}|\omega|^2} \sum_{l=0}^{k} \binom{j-k+k}{k-l} (-\sqrt{\pi}\omega^*)^{j-k} \frac{(-\pi|\omega|^2)^l}{l!} \\
&= e^{\pi i t f} \sqrt{\frac{k!}{j!}} e^{-\frac{\pi}{2}|\omega|^2} (-\sqrt{\pi}\omega^*)^{j-k} L_k^{j-k}(\pi|\omega|^2)
\end{aligned}
$$

where L_k^{j-k} is the generalized Laguerre function of degree k. The generalized Laguerre function is defined as [9]

$$
L_k^j(t) = \sum_{l=0}^{k} \binom{k+j}{k-l} \frac{(-t)^l}{l!}
$$

For the case $j \leq k$, substitute $l \to j - l$

$$
\begin{aligned}
(V_{h_j}h_k)(t,f) &= e^{\pi i t f} \sqrt{\frac{\pi^{j+k}}{j!k!}} e^{-\frac{\pi}{2}|\omega|^2} \sum_{l=0}^{j} \frac{j!}{(j-l)!l!} (-\omega^*)^l \pi^{l-j} \frac{k!}{(k-(j-l)!)} \omega^{l-k} \\
&= e^{\pi i t f} \sqrt{\frac{j!}{k!}} e^{-\frac{\pi}{2}|\omega|^2} (\sqrt{\pi}\omega)^{k-j} \sum_{l=0}^{j} \binom{j+k-j}{j-l} \frac{(-\pi|\omega|^2)^l}{l!} \\
&= e^{\pi i t f} \sqrt{\frac{j!}{k!}} e^{-\frac{\pi}{2}|\omega|^2} (\sqrt{\pi}\omega)^{k-j} L_j^{k-j}(\pi|\omega|^2)
\end{aligned}
$$

□

Using Theorem 1 gives the spectrogram for a Hermite function

$$
(S_{h_j}h_k)(t,f) = |(V_{h_j}h_k)(t,f)|^2 = \begin{cases} \frac{k!}{j!} e^{-\pi|\omega|^2} (\pi|\omega|^2)^{j-k} (L_k^{j-k}(\pi|\omega|^2))^2, & j \geq k \\ \frac{j!}{k!} e^{-\pi|\omega|^2} (\pi|\omega|^2)^{k-j} (L_j^{k-j}(\pi|\omega|^2))^2, & j \leq k \end{cases} \tag{36}
$$

and we can find the special case of the Hermite windowed spectrogram of the Gaussian signal in (33) by setting $k = 0$ giving

$$(S_{h_j} h_0)(t, f) = \frac{1}{j!} e^{-\pi(t^2+f^2)} (L_0^j(\pi|\omega|^2))^2 (\pi(t^2 + f^2))^j = \frac{\pi^j}{j!} e^{-\pi(t^2+f^2)} (t^2 + f^2)^j \tag{37}$$

The case of $j = 0$ also results in the commonly used matched Gaussian (MG), where the window is matched to the Gaussian signal.

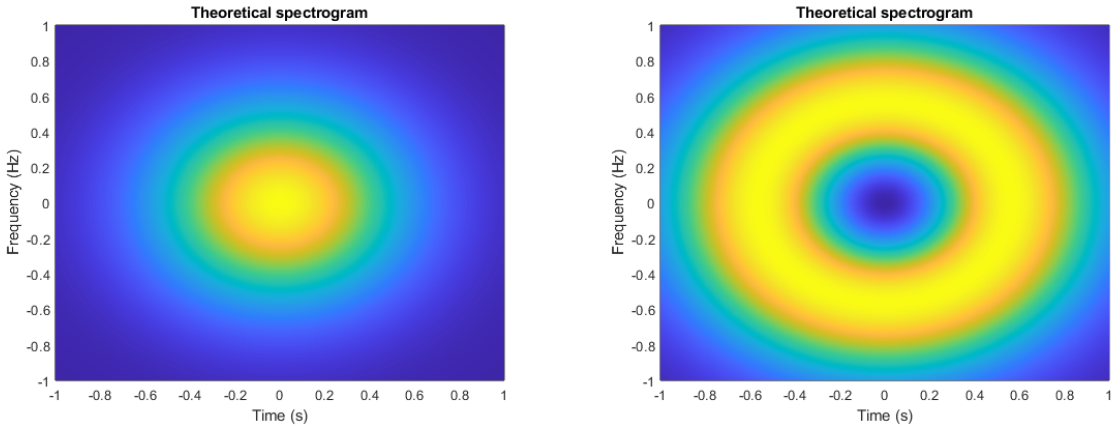

Figure 3: The resulting spectrogram of a Gaussian signal from the theoretical expression in (37)*. The left figure uses the first Hermite function $h_0(t)$ as window and right figure the second Hermite function $h_1(t)$.*

As Figure 3 shows, the resulting spectrogram using equation (37) matches what is expected. The spectrogram of $h_0(t)$, which is the same as the matched Gaussian results in a Gaussian shape as the left plot shows. The spectrogram using the higher order Hermite function instead results in a circular disc, as seen in the right plot.

4.1 Analytical solution to the least squares optimization

Given the expressions for the Wigner distribution in (34) and the multitaper spectrogram in (37), an explicit solution to the least-squares problem in (25) for a Gaussian signal can be computed analytically. We first note that we can obtain a solution to (25) by letting $\mathbf{w} = [w_0, \ldots, w_{N-1}]^T$ and $\mathbf{S} = [S_{h_0}x, \ldots, S_{h_{N-1}}x]$. We then have

$$\min_w \quad \iint (Wx - \mathbf{S}\mathbf{w})^2 \, dtdf$$

and taking the derivative with respect to $\mathbf{w}$ to be zero gives the minima $\mathbf{w}^*$ as

$$\begin{aligned} \iint 2\mathbf{S}^T(Wx - \mathbf{S}\mathbf{w}^*)dtdf = 0 \iff \\ \left(\iint \mathbf{S}^T\mathbf{S}dtdf\right)\mathbf{w}^* = \iint \mathbf{S}^T \cdot (Wx)dtdf \implies \\ \mathbf{w}^* = \left(\iint \mathbf{S}^T\mathbf{S}dtdf\right)^{-1} \iint \mathbf{S}^T \cdot (Wx)dtdf \end{aligned} \tag{38}$$

The integrals in (38) for the Gaussian $x(t)$ from (33) will then be of the form

$$\iint S_m x \cdot (Wx)dtdf = \iint 2\frac{\pi^m}{m!} e^{-3\pi(t^2+f^2)} (t^2 + f^2)^m dtdf \tag{39}$$

$$\iint S_m x \cdot (S_n x)dtdf = \iint \frac{\pi^{m+n}}{(m!n!)} e^{-2\pi(t^2+f^2)} (t^2 + f^2)^{m+n} dtdf \tag{40}$$

where S_m denotes the m:th element of the vector of spectrograms $\mathbf{S}$.

Disregarding the constants one can see from (39)-(40) that both integrals are of the form

$$I = \iint e^{-c(t^2+f^2)}(t^2+f^2)^k dtdf \tag{41}$$

This integral can be solved by using a binomial expansion as

$$I = \underbrace{\iint e^{-c(t^2+f^2)}(t^{2k}+f^{2k})dtdf}_{I_1} + \underbrace{\sum_{l=1}^{k-1}\binom{k}{l}\iint t^{2(k-l)}f^{2l}e^{-c(t^2+f^2)}dtdf}_{I_2}, \quad k>0$$

$$I = \iint e^{-c(t^2+f^2)}dtdf, \quad k=0 \tag{42}$$

where the sum is taken as 0 for $k=1$. Both integrals I_1 and I_2 can be split up to $\int f(x)dx \cdot \int f(y)dy$ where

$$\begin{aligned}\int f(y)dy &= \int y^{2k}e^{-cy^2}dy \\ &= \frac{1}{c^k 2^k}\cdot\frac{\sqrt{2\pi}}{\sqrt{2c}}\cdot(2k-1)!! = \frac{\sqrt{\pi}}{c^{k+\frac{1}{2}}2^k}(2k-1)!!\end{aligned} \tag{43}$$

Here $k!!$ denotes the double factorial (see Appendix B). This gives that I_2 can be computed as

$$\begin{aligned}I_2 &= \sum_{l=1}^{k-1}\binom{k}{l}\int t^{2(k-l)}e^{-ct^2}dt\cdot\int f^{2l}e^{-cf^2}df \\ &\underset{(43)}{=} \sum_{l=1}^{k-1}\binom{k}{l}\frac{\sqrt{\pi}}{c^{k-l+\frac{1}{2}}2^{k-l}}(2(k-l)-1)!!\cdot\frac{\sqrt{\pi}}{c^{l+\frac{1}{2}}2^l}(2l-1)!! \\ &= \sum_{l=l}^{k-1}\binom{k}{l}\frac{\pi}{c^{k+1}2^k}(2(k-l)-1)!!(2l-1)!!\end{aligned} \tag{44}$$

Similar calculations for I_1 give

$$\begin{aligned}I_1 &= \iint e^{-c(t^2+f^2)}t^{2k}dtdf + \iint e^{-c(t^2+f^2)}f^{2k}dtdf \\ &= 2\iint e^{-c(t^2+f^2)}t^{2k}dfdt = \frac{\sqrt{2\pi}}{\sqrt{2c}}\int e^{-ct^2}t^{2k}dt \\ &\underset{(43)}{=} \frac{2\sqrt{\pi}}{\sqrt{c}}\frac{\sqrt{\pi}}{c^{k+\frac{1}{2}}2^k}(2k-1)!! = \frac{2\pi}{c^{k+1}2^k}(2k-1)!!\end{aligned} \tag{45}$$

Combining (45) and (44) then finally gives

$$\begin{cases} I = \frac{\pi(2k-1)!!}{c^{k+1}2^{k-1}} + \sum_{l=1}^{k-1}\binom{k}{l}\frac{\pi}{c^{k+1}2^k}(2(k-l)-1)!!(2l-l)!!, & k>0 \\ I = \frac{\pi}{c}, \quad k=0 \end{cases} \tag{46}$$

Using the expression from (46) and we get for (39)-(40)

$$\begin{aligned}\iint S_m x\cdot(Wx)dtdf &= \begin{cases} 2\frac{\pi^m}{m!}\left(\frac{(2m-1)!!}{\pi^m 3^{m+1}2^{m-1}} + \sum_{l=1}^{m-1}\binom{m}{l}\frac{(2(m-l)-1)!!(2l-1)!!}{\pi^m 3^{m+1}2^m}\right), & m\in\mathbb{Z}_+ \\ \frac{2}{3}, \quad m=0 \end{cases} \\ &= \begin{cases} \frac{(2m-1)!!}{m!\cdot 3^{m+1}2^{m-2}} + \sum_{l=1}^{m-1}\binom{m}{l}\frac{(2(m-l)-1)!!(2l-1)!!}{m!\cdot 3^{m+1}2^{m-1}}, & m\in\mathbb{Z}_+ \\ \frac{2}{3}, \quad m=0 \end{cases}\end{aligned} \tag{47}$$

$$\iint S_m x \cdot (S_n x) dt df =$$

$$\begin{cases} \frac{\pi^{m+n}}{m!n!}\left(\frac{\pi(2(m+n)-1)!!}{\pi^{m+n}2^{m+n+1}2^{m+n-1}} + \sum_{l=1}^{m+n-1}\binom{m+n}{l}\frac{\pi}{\pi^{m+n+1}2^{m+n+1}2^{m+n}}(2(m+n-l)-1)!!(2l-1)!!\right), & m,n \in \mathbb{Z}_+ \\ \frac{1}{3}, & m,n = 0 \end{cases}$$

$$= \begin{cases} \frac{(2(m+n)-1)!!}{m!n!2^{2(m+n)}} + \sum_{l=1}^{m+n-1}\binom{m+n}{l}\frac{1}{m!n!2^{2(m+n)+1}}(2(m+n-l)-1)!!(2l-1)!!, & m,n \in \mathbb{Z}_+ \\ \frac{1}{3}, & m,n = 0 \end{cases} \tag{48}$$

which gives a solution to (38). Computing the least squares can thus be done without using numerical integration.

Weights $N = 8$	
"Numerical"	"Analytical"
1.9997	1.9997
-1.9948	-1.9948
1.9604	1.9607
-1.8233	-1.8241
1.4810	1.4827
-0.9343	-0.9364
0.3888	0.3902
-0.0777	-0.0780

Table 1: A comparison of the weights solving (25) *calculated using numerical integration and "fminunc" in MATLAB with the analytical expressions from* (47)*,* (48) *and* (38)

As an example, as seen in table 1, the 8 weights computed using numerical integration and "fminunc" on the left and the weights using the previously derived solutions on the right are almost identical confirming that the calculations are correct. Using the same error and computing the weights on the left took $t = 95$ s while the weights on the right took $t = 0.02999$ s. Using the analytical solution was thus more than a thousand times faster.

4.2 Analytical expression of Rényi entropy

Similarly, the Rényi entropy for the multitaper spectrogram of a Gaussian signal can also be computed analytically. This allows for faster computations for the RE optimization problem. We have that

$$H_\alpha(\sum_{n=0}^{N-1} w_n S_{h_n} x) = \frac{1}{1-\alpha}\log_2 \iint (\sum_{n=0}^{N-1} w_n S_{h_n} x)^\alpha = \frac{1}{1-\alpha}\log_2 \iint \sum_{|a|=\alpha}\binom{\alpha}{a}\mathbf{w}^a S_{h_a}^a$$

$$= \frac{1}{1-\alpha}\log_2 \sum_{|a|=\alpha}\binom{\alpha}{a}\mathbf{b}^a\mathbf{w}^a \cdot I \text{ where } b = \left(\frac{\pi^0}{0!}, \quad \frac{\pi^1}{1!}, \quad \ldots, \quad \frac{\pi^{N-1}}{(N-1)!}\right) \tag{49}$$

$$\tag{50}$$

using multi-index notation (see Appendix B) and I is taken from (46) with

$$\begin{cases} c = \alpha \cdot \pi \\ k = \sum_{n=0}^{N-1} a_n \cdot n \end{cases}$$

From (49), it is clear that the optimal weights of minimization problem RE will be very large since the logarithm is a monotonically increasing function and the only constraint is that the weights sum up to one. The expression involves a logarithm of a sum of products of weights so the minima would probably be obtained for a combination of very large positive and negative weights. Solving RE numerically for $N = 3$ gave the weights

$$w = (3.027 \cdot 10^6, -1.250 \cdot 10^6, -1.777 \cdot 10^6)$$

which are indeed very large. The Rényi entropy might therefore not be the best criterion since it pushes for large weights which could cause numerical issues and issues for a noisy signal. To solve the issue of large weights we included a constraint on the 2 norm of the weights i.e. $\|w\|_2 < \lambda$ as seen in (31). The use of $\lambda = 60$ seemed to be a good choice as the Rényi entropy was as low as for the unconstrained problem while still keeping weights somewhat low.

4.3 Weights

In this section, the weights using the different optimization criteria are presented. The weights were calculated in MATLAB using "fmincon" or "fminunc" depending on if there is an added constraint with the exception of LS which used the solution previously derived. The choice of $\alpha = 5$ both for the Rényi entropy H_α and the J2 weights was used as it gave a lower FWHM for $N = 3$ and $N = 8$ windows than the common $\alpha = 3$. The use of $\alpha = 0.8$ for J1 was based on the use in[3]. The RE weights used the expression in (49) instead of numerical integration. Because the Rényi divergence (RD) never managed to converge we discarded it as a method.

Weights $N = 3$					
LS	Cw	J1($\alpha = 0.8$)	J2($\alpha = 5$)	RE($\alpha = 5$)	KT
1.000	1.000	1.000	1.000	1.000	1.000
-0.769	-0.571	0.433	-0.956	-0.397	-1.726
0.308	0.143	-0.001	0.655	-0.583	1.067

Table 2: The optimized weights normalized on the first weight for comparison.

The weights calculated with $N = 3$ windows using the different optimization criteria are shown in 2. In table 2 we can see that the LS and Cw weights are the most similar. The weights are alternating positive and negative except for RE and J1. The weights also seem to be decreasing in absolute value except for KT and RE.

Weights $N = 8$					
LS	Cw	J1($\alpha = 0.8$)	J2($\alpha = 5$)	RE($\alpha = 5$)	KT
1.000	1.000	1.000	1.000	1.000	1.000
-0.9976	-0.9686	0.6178	-0.9261	-0.05	-5.2873
0.9805	0.8588	0.4188	0.6309	-0.1375	16.5753
-0.9122	-0.6392	0.3098	0.3004	-0.1729	-32.0795
0.7415	0.3647	0.2352	-0.2290	-0.1793	38.0389
-0.4683	-0.1451	0.1663	-0.3536	-0.1805	-25.6951
0.1951	0.0353	0.0896	-0.0744	-0.1808	7.9663
-0.0390	-0.0039	-0.0009	0.4261	-0.1810	-0.3516

Table 3: The optimized weights normalized on the first weight for comparison.

Table 3 instead shows the weights calculated using $N = 8$ windows. The behavior of the weights for $N = 8$ windows in table 3 closely resembles the $N = 3$ case. The LS and Cw weights are still the most similar and all of the weights except RE and KT are again decreasing in absolute value. The KT weights on the other hand have very large weights for the higher order windows with the 5th window having the largest weight. The RE weights also have all of the weights negative except for the first window.

4.4 Evaluation

We evaluate the weights using several criteria. The first is the concentration or how sharp the peak is which is evaluated using full width at half maximum (FWHM), Rényi entropy, and kurtosis of the theoretical multitaper spectrogram given by (37). Visual analysis is also done on a simulated signal without noise.

All of the simulations will be done on a signal consisting of a Gaussian in (33) with length $L = 512$ samples, sampling frequency 63.8750 Hz, time-shifted by 4 s, and frequency modulated by 15.9688 Hz.

The second evaluation criterion is the noise sensitivity of the different weights. The evaluation is done by counting the number of times the correct peak is identified for the simulated Gaussian with varying levels of SNR. We define SNR as

$$\text{SNR} = A^2/\sigma^2$$

where A is the amplitude of the signal (in this case $2^{1/4}$) and σ^2 the variance of the noise. The peak for the TFR is estimated by simply taking the maximum value. A visual comparison of the TFRs of one simulation is also performed.

All of the evaluations are done using first $N = 3$ windows and then $N = 8$ windows to compare performance for a small number of windows and a large amount. The arithmetic mean $\frac{1}{N}$ which is a common choice for weights as well as the matched Gaussian (MG) which uses a Gaussian function as a window are included for comparison.

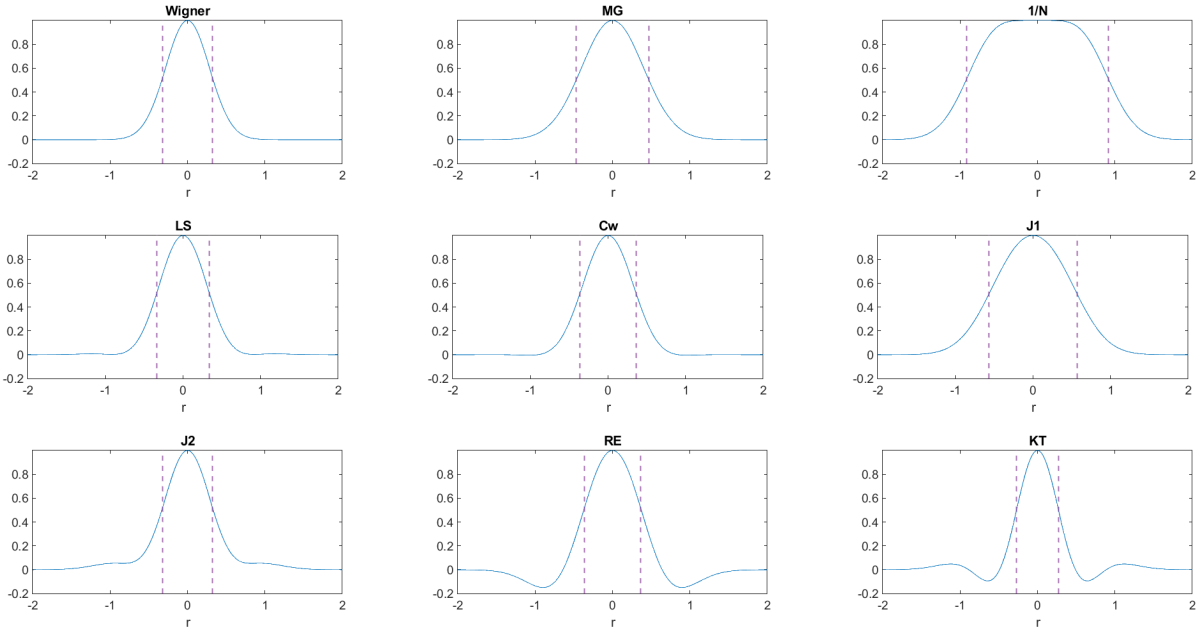

Figure 4: Radial sections of the Wigner distribution and the multitaper spectrograms with $N = 3$ windows using the optimized weights and normalized to height 1. The FWHM is marked with a dashed vertical line. Since the multitaper spectrogram of a Gaussian is radially symmetric the radial section gives a good comparison of the weights.

Weight	FWHM	$\gamma(C)$	$H_3(C)$
Wigner	0.6493	2.0000	-0.2075
MG	0.9379	1.0000	0.7925
$\frac{1}{N}$	1.8357	0.3446	2.0019
LS	0.6814	1.9039	-0.1912
Cw	0.7295	1.6782	-0.0729
J1($\alpha = 0.8$)	1.1463	0.7071	1.2159
J2($\alpha = 5$)	0.6493	1.8673	0.5112
RE($\alpha = 3$)	0.7295	1.3705	-7.2800
KT	0.5371	2.9177	-0.7431

Table 4: Comparison of concentration measures of the multitaper spectrograms with $N = 3$ windows using the different weights with the concentration of the Wigner distribution. The FWHM and Rényi entropy $H_3(C)$ should be small for concentrated TFRs C while the kurtosis $\gamma(C)$ should be high.

As seen in both Figure 4 and in table 4, the KT weights resulted in the sharpest peak. The KT weights had

a FWHM even lower than the Wigner distribution, so if the goal is estimating the time-frequency center of a component these weights seem the best. However, as seen in the figure the KT weights have some oscillations and negative values surrounding the peak, which could potentially be an issue. The LS weights had a FWHM only slightly larger than the Wigner distribution and the plot looked visually very similar to the Wigner distribution. Interestingly, the J2 weights resulted in a peak with FWHM matching the Wigner distribution, so in this regard, it performed even better than the LS weights. The restriction to positive values for the J2 weights is clearly seen in the figure by the heavy tail. The Cw weights also had a FWHM quite close to Wigner and significantly better than the MG weights. The J1 weights performed very poorly both for the FWHM as seen in the table and as observed in the figure by the wide peak. The RE weights performed better than MG with a FWHM equal to Cw. However, the RE weights also resulted in a large valley with negative values surrounding the peak. Using $\frac{1}{N}$ as weights is by far the worst, as evidenced by both the figure and the table.

For the kurtosis $\gamma(C)$, the KT weights as expected had the largest value since this was the optimization criteria used for the weights. The Wigner distribution had the second-largest value. Unlike the FWHM, the LS weights in this case had the second-best kurtosis measure among the optimized weights, but the J2 weights followed closely after. The performance of the weights in terms of kurtosis was similar to FWHM indicating that both concentration measures are somewhat similar. One difference though is that the RE weights had the same FWHM as Cw while the value of $\gamma(C)$ was much lower. The J1 weights again perform worse than MG but the arithmetic mean weights $\frac{1}{N}$ are by far the worst.

Finally, when the Rényi entropy $H_3(C)$ was used as a concentration measure, the RE weights that used it as an optimization criterion naturally performed the best. The KT weights had the second lowest Rényi entropy and the Wigner distribution came in third. In general, the concentration measures agree with each other. The Rényi entropy had a significantly lower value for the RE weights compared to the Wigner distribution and this could be due to the fact that the optimization problem might be unbounded if the only constraint is that the weights sum to 1. Using the Rényi entropy as an optimization criterion might therefore not be a good choice, as up to a certain point it seems to stop working as a concentration measure. Even though the RE weights have a very large negative Rényi entropy the weights do not perform better on the other concentration measures. The LS weights in this case performed better than the J2 weights which performed quite poorly for this measure. Since the RE weights had large negative values surrounding the peak as seen in Figure 4, while the J2 weights were required to be positive, it could be the reason why it performed worse for this measure.

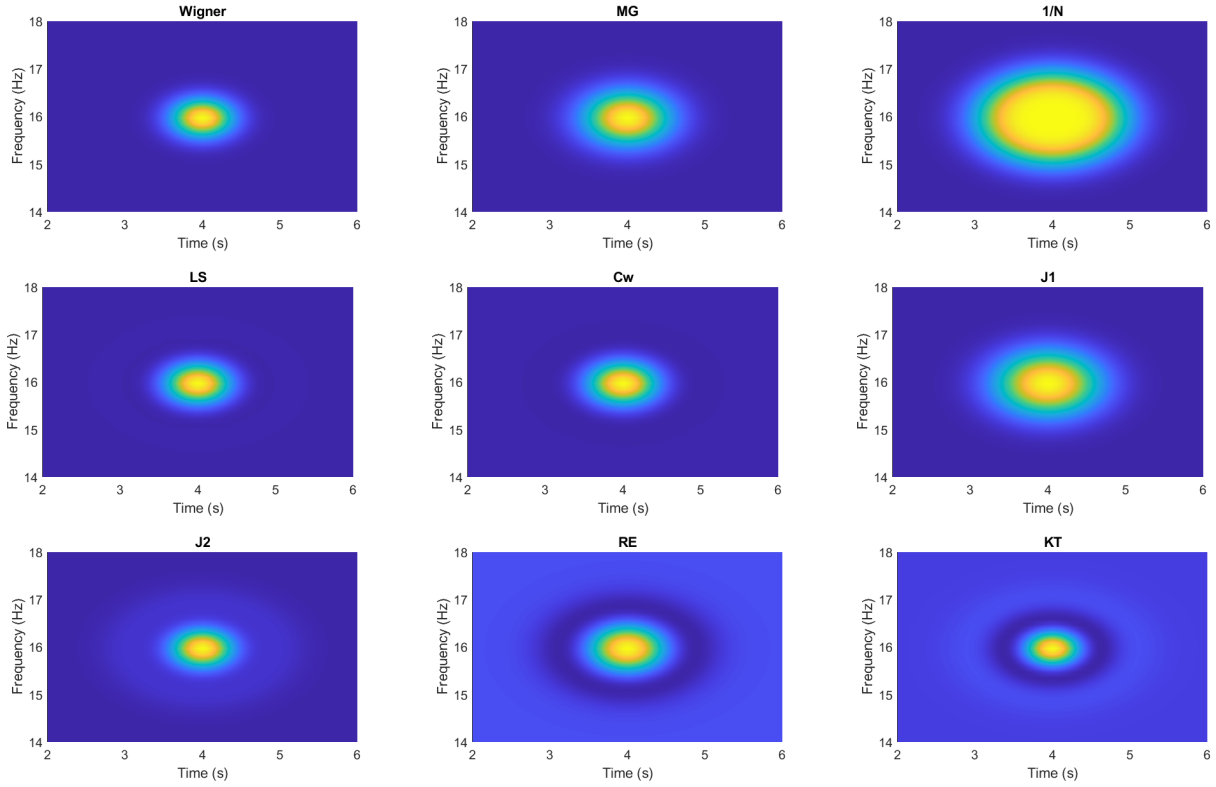

Figure 5: Wigner distribution and multitaper spectrograms of a simulated Gaussian using the optimized weights without noise. Each subplot has its own color scale since the weights are scaled differently and it is only the relative scale that matters.

In Figure 5, one can see a visual comparison of the multitaper spectrograms using the different weights. From Figure 5, it is evident that the LS, J2, and KT produced the narrowest peaks with Cw following closely. Although, KT has the sharpest peak one can also see the oscillatory behavior and negative values earlier noted in Figure 4. The heavy tail of J2 with is also clear by the very lightly shaded ball surrounding the peak. The TFR using the LS weights again most closely resembles the Wigner distribution. The negative values of the RE weights are also obvious by the dark blue surrounding the peak and the peak looks a bit more concentrated than MG as table 4 earlier showed. The very poor performance of the J1 weights seen in table 4 is also visible since the TFR looks very poorly localized. The $\frac{1}{N}$ weights are however much worse than the rest.

In regards to concentration, the clear winner for the weights was KT. The KT weights placed first for both FWHM and kurtosis and second for Rényi entropy. The KT weights also looked visually the most concentrated for the simulated signal. So if the purpose is to have a TFR with as localized a peak as possible, these weights are the best choice. Of course, there is a penalty to be paid somewhere due to the uncertainty principle which shows up in the oscillations that the TFR has. Interestingly, the J2 weights which had good results for FWHM, even matching the Wigner distribution performed worse than LS for the other measures. This is reasonable though since it was seen in Figure 4 that besides the FWHM the actual distribution is not that concentrated due to the heavy tails. The weights that could be regarded as the next best in terms of concentration would be the LS weights since they performed well for all three measures. The LS weights also looked visually close to the Wigner distribution both for the radial section and the simulated signal.

Peak detection, N=3, 1000 simulations				
Weight	100 dB	10 dB	4 dB	-1 dB
Wigner	1000	532	170	23
MG	1000	918	554	205
$\frac{1}{N}$	1000	2	2	0
LS	1000	914	552	200
Cw	1000	917	556	204
J1($\alpha = 0.8$)	1000	916	555	206
J2($\alpha = 5$)	1000	916	554	197
RE($\alpha = 3$)	1000	919	541	206
KT	1000	918	554	173

Table 5: Comparison of noise sensitivity for peak detection of the multitaper spectrograms using different weights and the Wigner distribution. The values are the number of times the correct peak was identified for 1000 simulations. The peaks are estimated as the maximum value of the TFR.

As for the robustness to noise as seen in table 5, the Wigner distribution performs very poorly for the noisy signal. When there is basically no noise, they all give the correct time-frequency center which is positive. Already at 10 dB the Wigner distribution starts to fail to detect the correct peak around 50% of the time while the multitapers except for $\frac{1}{N}$ have around a 90% accuracy. There is no real significant difference between the weights at this SNR though. At 4 dB all of the weights have around the same detection level with the exception again being $\frac{1}{N}$, which fails to detect the correct peak almost every time. The only difference starts to show at -1 dB, where the KT method has quite a bit worse detection in comparison to the other methods.

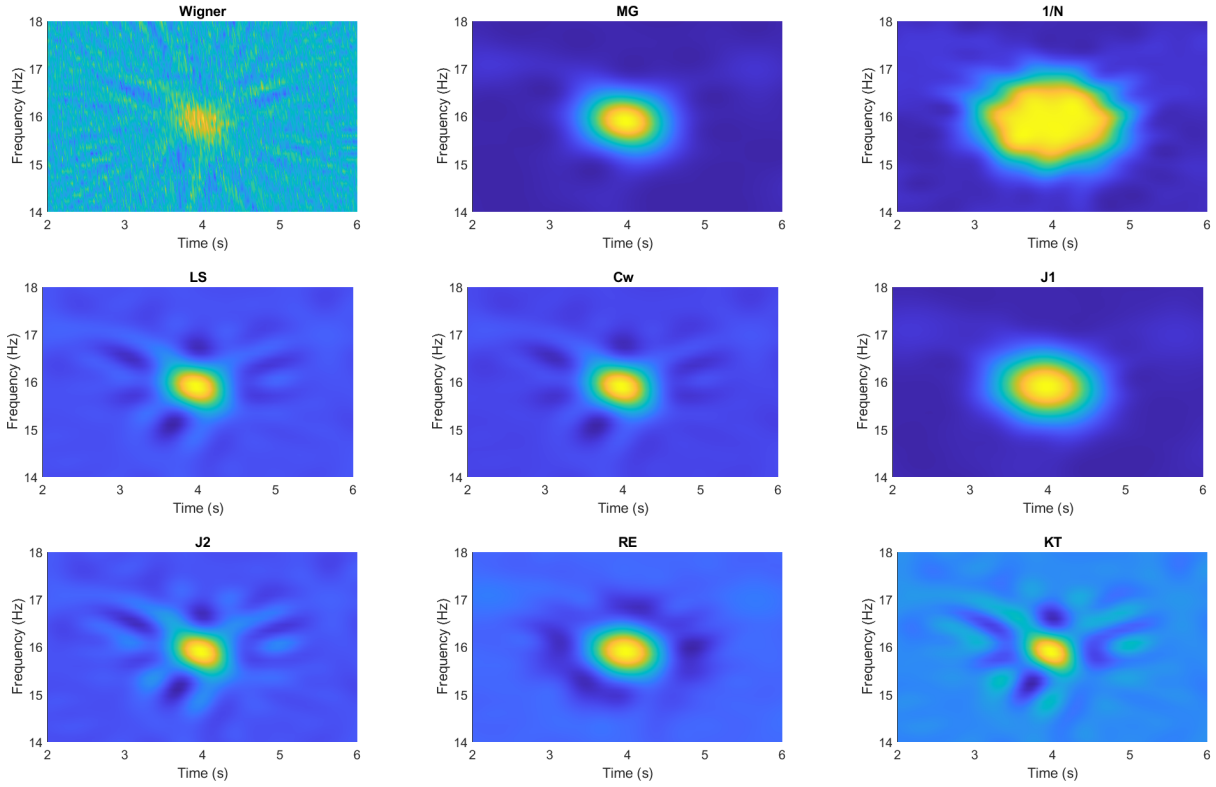

Figure 6: Wigner distribution and multitaper spectrograms of a simulated Gaussian using the optimized weights with SNR = 4 dB.

The worse performance to noise of the KT weights can be seen in Figure 6 with the large interference pattern surrounding the peak. The Wigner distribution however looks by far the worst with the peak quite difficult to

discern. The MG and J1 weights seem the least affected by the noise while J2 also seems to show some larger interference around the peak. The RE weights also look a bit less affected by the noise compared to LS, Cw, and J2. Since the RE, J1, and MG weights had the highest values in table 5 this also seems reasonable.

In general, there are no huge differences in noise reduction among the weights except for KT which performed much worse than the rest. However, all of the multitapers using the optimized weights perform noticeably better than the Wigner distribution for the noisy signal. The performance for the multitapers is not better than MG so for the case of $N = 3$ windows, there does not seem to be an advantage to using multitapers in terms of reducing noise.

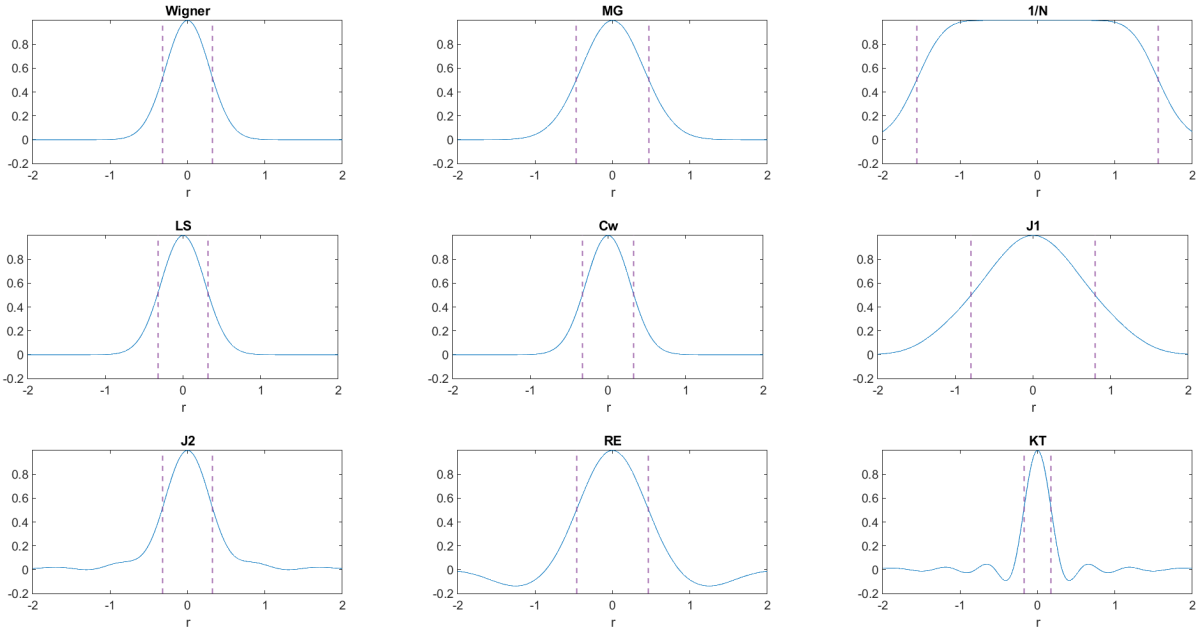

Figure 7: Radial sections normalized to height 1 of the Wigner distribution and the multitaper spectrograms with $N = 8$ windows using the optimized weights. The FWHM is marked with dashed vertical lines.

Weight	FWHM	$\gamma(C)$	$H_3(C)$
Wigner	0.6493	2.0000	-0.2075
MG	0.9379	1.0000	0.7925
$\frac{1}{N}$	3.1182	0.1286	3.2402
LS	0.6493	1.9999	-0.2075
Cw	0.6653	1.9922	-0.2061
J1($\alpha = 0.8$)	1.5952	0.3366	2.2942
J2($\alpha = 5$)	0.6493	1.8518	0.6053
RE($\alpha = 3$)	0.9218	0.8316	-7.8577
KT	0.3447	6.8401	-1.6794

Table 6: Comparison of concentration measures of the multitaper spectrograms with $N = 8$ windows using the different weights with the Wigner distribution.

In table 6, we can see that the LS weights now have the same FWHM as the Wigner distribution. In fact, all of the concentration measures for the LS weights are almost equal in value to the Wigner distribution, meaning that with 8 windows the LS weights seem to be a very good approximation. This is also seen in Figure 7, where the LS weights and Wigner distribution look visually indistinguishable. The Cw weights also have values close to the Wigner distribution but not quite as good as the LS weights. Interestingly, the J2 weights seem to have become worse when using 8 windows, since the kurtosis is lower and Rényi entropy higher while the FWHM remained the same. This might be because the tails are even heavier as it's difficult to ensure positivity using

8 windows giving a worse performance for the kurtosis and Rényi entropy. The J1 and the RE weights also seemed to have become worse in all aspects. Since KT had the best value for all concentration measures it seems like it again is the best choice when using many windows. The oscillations are however even more apparent in Figure 7. The LS weights again also seem to be the next best choice followed by Cw.

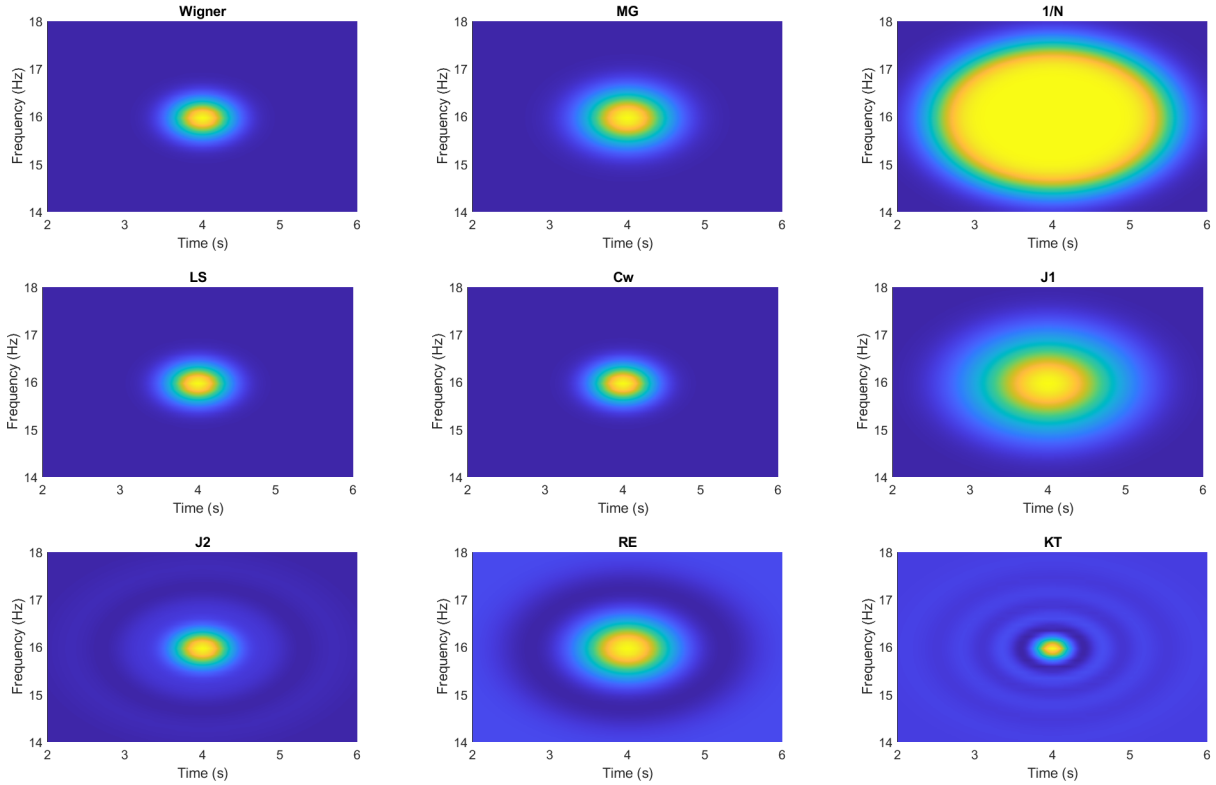

Figure 8: Wigner distribution and multitaper spectrograms of a simulated Gaussian using the optimized weights with no noise and $N = 8$ windows.

Both the very localized peak and the oscillations of the KT weights are visible for the simulated signal in Figure 8. Using $N = 8$ windows improved the concentration for the KT weights significantly and the peak is even visually much smaller compared to Wigner. The fact that the RE weights became worse is also very evident from the figure since the peak now looks wider than MG. The difference for the J2 weights for $N = 8$ compared to $N = 3$ is that there is now a small ring of lower values surrounding the peak. This could be slightly seen in Figure 7 where there seems to some small oscillations on the tail. The TFR of the LS and Cw weights look visually indistinguishable from the Wigner distribution. The $\frac{1}{N}$ weights also perform even worse compared to $N = 3$ as clearly evident by Figure 8. The J1 weights look worse than MG also for the simulated signal.

Peak detection, N=8, 1000 simulations				
Weight	100 dB	10 dB	4 dB	-1 dB
Wigner	1000	532	170	23
MG	1000	918	554	205
$\frac{1}{N}$	1000	1	1	0
LS	1000	917	496	158
Cw	1000	923	527	184
J1($\alpha = 0.8$)	1000	916	553	203
J2($\alpha = 5$)	1000	913	551	191
RE($\alpha = 3$)	1000	914	554	209
KT	1000	0	0	0

Table 7: Comparison of noise sensitivity for peak detection of the multitaper spectrograms using different weights and the Wigner distribution.

Looking at table 7 there is no big difference in performance for the low noise case in comparison to $N = 3$ windows except for the KT weights. The KT weights only managed to detect the peak when there was essentially no noise at all (SNR = 100 dB) but failed completely for all the other noise levels. In this case, it even performed worse than the Wigner distribution. So even though the KT weights gave excellent concentration when using 8 windows, the downside is that they are very sensitive to noise and are unusable for even just moderately noisy signals. As for the other weights, there was a slight performance degradation in the noisy case for all but RE. The largest difference was for the LS weights and the Cw weights which indicates that the robustness to noise decreases more when using multiple windows for these. The difference for the others was very small and could hardly be called significant.

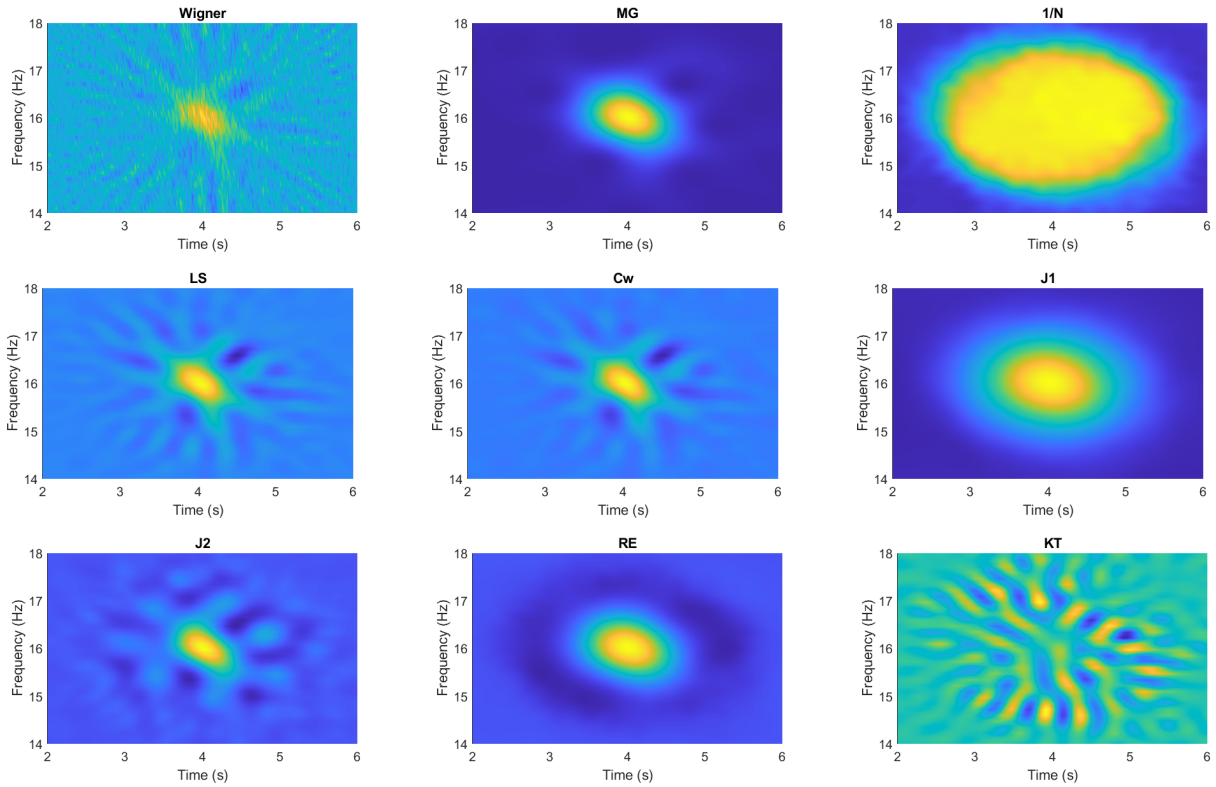

Figure 9: Wigner distribution and multitaper spectrograms of a Gaussian using the optimized weights with SNR = 4 dB and N = 8 windows.

The very poor performance for KT seen in table 7 is clearly visible in Figure 9, where the signal is not visible at all. The only visible part looks like some type of cross-term interference and noise. The weights in this case

perform visibly worse than the Wigner distribution since for the Wigner distribution one can slightly discern the signal from the noise. As for the other weights they generally seem slightly less robust to the noise when using more windows since some more interference is appearing around the time-frequency centers. The MG seems to now perform the best but the RE and J1 weights visibly seem to also better handle the noise. In contrast to $N = 3$ windows, the J2 weights in this case seem better than LS and Cw.

As general remarks to the different weights, the best choice in terms of concentration was clearly the KT weights. Although, in the case of many windows the weights were very sensitive to noise. However, even using only 3 windows it gave better concentration than the Wigner distribution while still being much less sensitive to noise. Another downside of the KT weights was the oscillations and negative values which could be hard to interpret. This is also an issue for all of the weights except for J2 but the negative values are much closer to zero for the LS and CW weights. If negative values are an issue, the J2 weights seem to be a good choice for a small number of windows since it had a FWHM as small as the Wigner distribution.

For the weights based on the Jenssen distance (J1 and J2) and the Rényi entropy using more windows than $N = 3$ only gave a worse performance in terms of concentration. In this case, it seems like the methods are only reasonable for a small number of windows. Since using a small number of windows is desirable for computational reasons this is not a big disadvantage. Using more windows for the LS and Cw weights could be a good choice though since in this case there was a significant difference. The LS weights gave a very good approximation to the Wigner distribution using 8 windows while still being much less sensitive to noise. The Cw weights were a bit worse but they have the advantage of being independent of the signal and rather approximates the Wigner distribution of any signal. Taking this into account, the Cw weights seem to be a quite good choice if using a higher amount of windows as they did not perform much worse than LS.

Using multiple windows did not really give an advantage in terms of noise reduction since MG performed equally well. The performance also in general became worse when using more windows. This is quite reasonable though since most weights were alternating positive and negative so using several windows does not average out the noise. The only exception was for the RE weights which did seem to be better for the noise using more windows. In this sense, the $\frac{1}{N}$ weights are the best for reducing the noise but since they perform so poorly in all other aspects it is not feasible to use them. The RE weights also had a slightly better performance than MG for both $N = 3$ and $N = 8$ windows which could mean that if one has a very noisy signal these could be a good choice. When using $N = 8$ windows both Figure 9 and table 7 showed better performance for the J2 weights in comparison to Cw and LS while the opposite held for $N = 3$ windows. In this sense, the J2 weights were better when using many windows however since the weights were simply worse than using only $N = 3$ windows, it does not really make sense to use the method for many windows.

The Jensen distance used for the J1 weights does not seem to work well at all since these weights performed poorly in practically all aspects. Although the J2 method gave good results, the resulting multitaper does not look as similar to the Wigner distribution as for the LS weights. Comparing the shape of the resulting multitaper in Figure 4 and Figure 7 the LS weights much better resemble the Wigner distribution in comparison to J2. The comparison is not entirely fair though since the J2 weights ensure that the TFR is always positive like the Wigner distribution is. Due to this, one could argue that the J2 weights give a better approximation than LS which allows negative values.

5. Multitaper spectrogram of a chirp

The optimization method for the multitaper spectrograms of a Gaussian is experimented with for a Gaussian enveloped chirp here defined as

$$y(t) = 2^{1/4} e^{-2\pi a^2 t^2} e^{-\pi t^2}, \quad a \in \mathbb{C} \tag{51}$$

Letting $a^2 = \frac{1}{2}(c - ib)$, it can be rewritten as

$$y(t) = 2^{1/4} e^{i\pi b t^2} e^{-\pi(1+c)t^2} \tag{52}$$

which is a linear chirp $e^{i\pi b t^2}$ with slope b multiplied with a Gaussian $e^{-\pi(1+c)t^2}$ with shape parameter $1 + c$. The Wigner distribution of the signal in (52) is easily calculated by using that the Wigner distribution is chirp

multiplication and scale invariant [13]

$$(We^{i\pi bt^2}y)(t,f) = (Wy)(t,f-bt)$$
$$(Wy(ct))(t,f) = \frac{1}{\sqrt{|c|}}(Wy(t))(ct,\frac{f}{c})$$

which for the signal $y(t)$ in (52) gives the Wigner distribution

$$(Wy)(t,f) = \frac{2}{(1+c)^{1/4}}\exp\Big(-2\pi\big((1+c)t^2+\frac{(f-bt)^2}{1+c}\big)\Big) \tag{53}$$

5.1 Hermite expansion

The multitaper spectrogram for the signal in (51) unfortunately does not seem to have an analytical solution as in the Gaussian case. The approach is, therefore, to try a series expansion in terms of Hermite functions since the multitaper spectrogram for the Hermite functions has a solution given by (36). Consider first a series expansion in terms of Hermite polynomials $H(t)$ of

$$\exp(-a^2t^2) = \sum_{l=0}^{\infty} c_l H_l(t) \tag{54}$$

where c_l are the corresponding weights. Since the Hermite polynomials form an orthogonal basis with respect to the weight function e^{-t^2} any square-integrable function using this weight function can be expressed as a series of Hermite polynomials.

Theorem 2.

$$\exp(-a^2t^2) = \frac{H_0(t)}{\sqrt{1+a^2}} + \sum_{l=1}^{\infty}\frac{(-1)^l a^{2l}}{2^{2l}l!(1+a^2)^{l+1/2}}H_{2l}(t) \tag{55}$$

Proof. We first make use of Mehler's formula [9]

$$\sum_{l=0}^{\infty}\frac{H_l(t_1)\cdot H_l(t_2)}{l!}(\frac{1}{2}w)^l = \frac{1}{\sqrt{1-w^2}}\exp\left(\frac{2t_1t_2w-(t_1^2+t_2^2)w^2}{1-w^2}\right), \quad w\in\mathbb{C}, \quad |w|<1 \tag{56}$$

Setting $t_2 = 0$ and using the property of the Hermite polynomials

$$H_l(0) = \begin{cases} 0, & \text{l odd} \\ (-2)^{l/2}(l-1)!!, & \text{l even} \end{cases}$$

simplifies (56) to

$$\sum_{l=0}^{\infty}\frac{H_{2l}(t_1)\cdot(-2)^l(2l-1)!!}{(2l)!}(\frac{1}{4}w^2)^l = \frac{1}{\sqrt{1-w^2}}\exp\left(\frac{-t_1^2w^2}{1-w^2}\right)$$

Substituting $\frac{w^2}{1-w^2} = a^2 \implies w^2 = \frac{a^2}{(1+a^2)}$ then gives

$$\sum_{l=0}^{\infty}\frac{H_{2l}(t_1)\cdot(-2)^l(2l-1)!!}{(2l)!}(\frac{1}{4}\frac{a^2}{1+a^2})^l = \sqrt{1+a^2}\exp\left(-a^2t_1^2\right) \underset{l>0}{\overset{(2l)!=(2l)!!(2l-1)!!}{\iff}}$$
$$\frac{H_0(t_1)}{\sqrt{1+a^2}} + \sum_{l=1}^{\infty}\frac{H_{2l}(t_1)\cdot(-1)^l 2^l(2l-1)!!}{(2l)!!(2l-1)!!2^{2l}(1+a^2)^{l+1/2}}a^{2l} = \exp(-a^2t_1^2) \underset{(2l)!!=2^l l!}{\iff}$$
$$\frac{H_0(t_1)}{\sqrt{1+a^2}} + \sum_{l=1}^{\infty}\frac{(-1)^l a^{2l}}{2^{2l}l!(1+a^2)^{l+1/2}}H_{2l}(t_1) = \exp(-a^2t_1^2) \tag{57}$$

□

Theorem 2 gives an expansion in terms of Hermite polynomials but for use in multitaper spectrograms, we want an expansion in terms of Hermite functions. Using (55), we can easily obtain an expansion in Hermite functions by multiplying both sides with $2^{1/4} \cdot e^{-\frac{t_1^2}{2}}$ and substituting $t_1 = \sqrt{2\pi}t$ which gives

$$
\begin{aligned}
&2^{1/4} \exp(-\pi t^2) \exp(-2\pi a^2 t^2) = \\
&\frac{2^{1/4} \exp(-\frac{(\sqrt{2\pi}t)^2}{2}) H_0(\sqrt{2\pi}t)}{\sqrt{1+a^2}} + \sum_{l=1}^{\infty} \frac{\sqrt{(2l)!}(-1)^l a^{2l}}{l!(1+a^2)^{l+1/2}\sqrt{2^{2l}}} \frac{2^{1/4}}{\sqrt{2^{2l}(2l)!}} \exp(-\frac{(\sqrt{2\pi}t)^2}{2}) H_{2l}(\sqrt{2\pi}t) \\
&\underset{(15)}{=} \frac{h_0(t)}{\sqrt{1+a^2}} + \sum_{l=1}^{\infty} \frac{\sqrt{(2l)!} \cdot (-1)^l a^{2l}}{2^l l!(1+a^2)^{l+1/2}} h_{2l}(t)
\end{aligned} \tag{58}
$$

In (58), one can see that we have obtained an expansion in Hermite functions of the chirp signal in (51) as desired.

The weights can be computed recursively as

$$
c_l = \sqrt{\frac{2l-1}{2l}} \frac{-a^2}{1+a^2} \cdot c_{l-1}, \quad c_0 = \frac{1}{\sqrt{1+a^2}}
$$

allowing for easy computation.

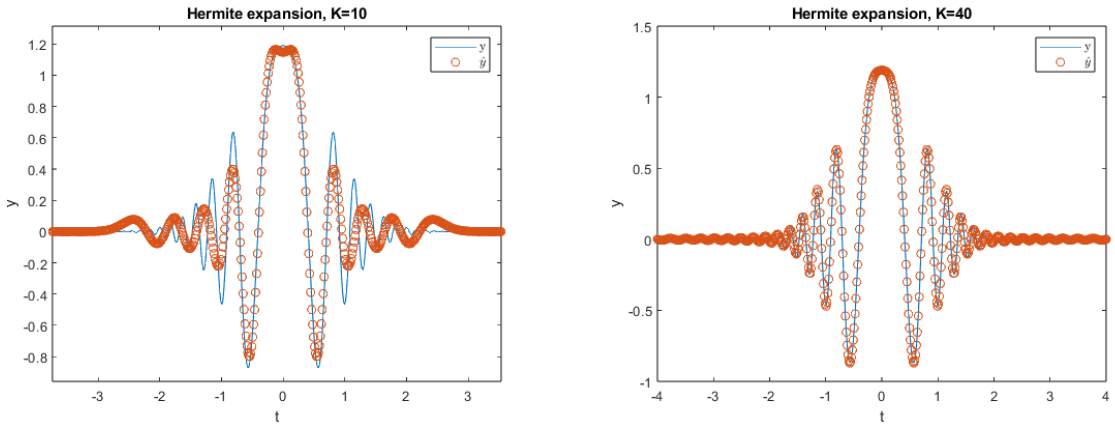

Figure 10: Hermite expansion $\hat{y}$ of a chirp y with different number of terms K included. The signal consists of the Gaussian chirp in (52) *with $b = 3$ and $c = -0.7$.*

As seen in Figure 10 the Hermite expansion using the derived weights approximates the chirp signal. For this signal only using 10 terms yields a poor approximation while using 40 terms yields a good approximation. The convergence rate of the Hermite expansion will depend on $|w|$ in (56). Since we have that $w^2 = \frac{a^2}{(1+a^2)}$, we get $|w|^2 = \frac{|a^2|}{|1+a^2|} \leq 1$ and the expansion will never diverge but will be slower depending on a.

Using Theorem 1, we can compute the STFT of a Hermite expansion as

$$
\begin{aligned}
(V_{h_j} \sum_k c_k h_k)(t,f) &= \sum_k c_k (V_{h_j} h_k)(t,f) = \\
&= e^{\pi i t f} e^{-\frac{\pi}{2}|\omega|^2} \left(\sum_{k \geq j} c_k \sqrt{\frac{j!}{k!}} (\sqrt{\pi}\omega)^{k-j} L_j^{k-j}(\pi|\omega|^2) + \sum_{k<j} c_k \sqrt{\frac{k!}{j!}} (-\sqrt{\pi}\omega^*)^{j-k} L_k^{j-k}(\pi|\omega|^2) \right)
\end{aligned}
$$

which gives the spectrogram

$$S_{h_j}\sum_k c_k h_k = |V_{h_j}\sum_k c_k h_k|^2$$
$$= e^{-\pi|\omega|^2}\left|\left(\sum_{k\geq j} c_k\sqrt{\frac{j!}{k!}}(\sqrt{\pi}\omega)^{k-j}L_j^{k-j}(\pi|\omega|^2)+\sum_{k<j} c_k\sqrt{\frac{k!}{j!}}(-\sqrt{\pi}\omega^*)^{j-k}L_k^{j-k}(\pi|\omega|^2)\right)\right|^2 \quad (59)$$

Using the weights obtained for the Gaussian chirp from (58) in (59) gives an approximation to the spectrogram. We can then form the multitaper spectrogram by using (59) in (17) and use it for the optimization problems in section 3.

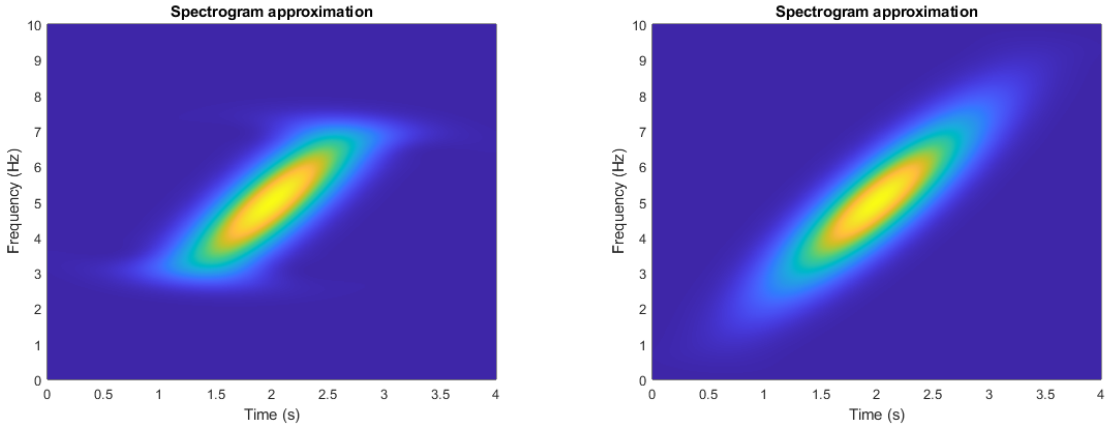

Figure 11: The theoretical spectrogram from (59) *for the Hermite expansion in Figure 10 time-shifted 2 s and frequency modulated 5 Hz. The spectrograms match the expected shape of a chirp. (Left) Approximation using a Hermite expansion of 10 terms. (Right) Approximation using Hermite expansion of 40 terms.*

Using the expression in (59) for the Hermite expansion in Figure 10 resulted in the spectrograms shown in Figure 11. The spectrograms clearly look like chirps validating that the theoretical expression is correct. As seen in the left plot, using only $K = 10$ terms in the expansion gave a worse approximation since the ends of the chirp look flatter. The right plot using $K = 40$ gave a better approximation with no visible deviation.

5.2 Weights

In this section, the weights optimized for the chirp signal in (51) are presented. Since there is a choice of parameter a in (51), or b and c in (52), the weights would be different for each parameter choice as each signal would require a new Hermite expansion. Since it would not be feasible to calculate new weights for each value of b and c, we instead check for which value the squared error is large when using old weights. These parameter values will then be used for the evaluation of the weights to see if the method gives any significant performance difference compared to just using the Gaussian weights. The weights optimized for the Gaussian signal will be denoted w_g while the weights optimized using the Hermite expansion of the actual chirp signal by w_h. This will also show the robustness of a mismatch in the signal model for the Gaussian weights since these weights are calculated with the assumption that the Hermite function window and the Gaussian signal are matched in the Gaussian shape parameter c. If the Gaussian weights are more robust, the error should not increase rapidly if the signal and the windows are mismatched meaning $1 + c \neq 0$. This is desirable since we often do not know the exact signal we are analyzing.

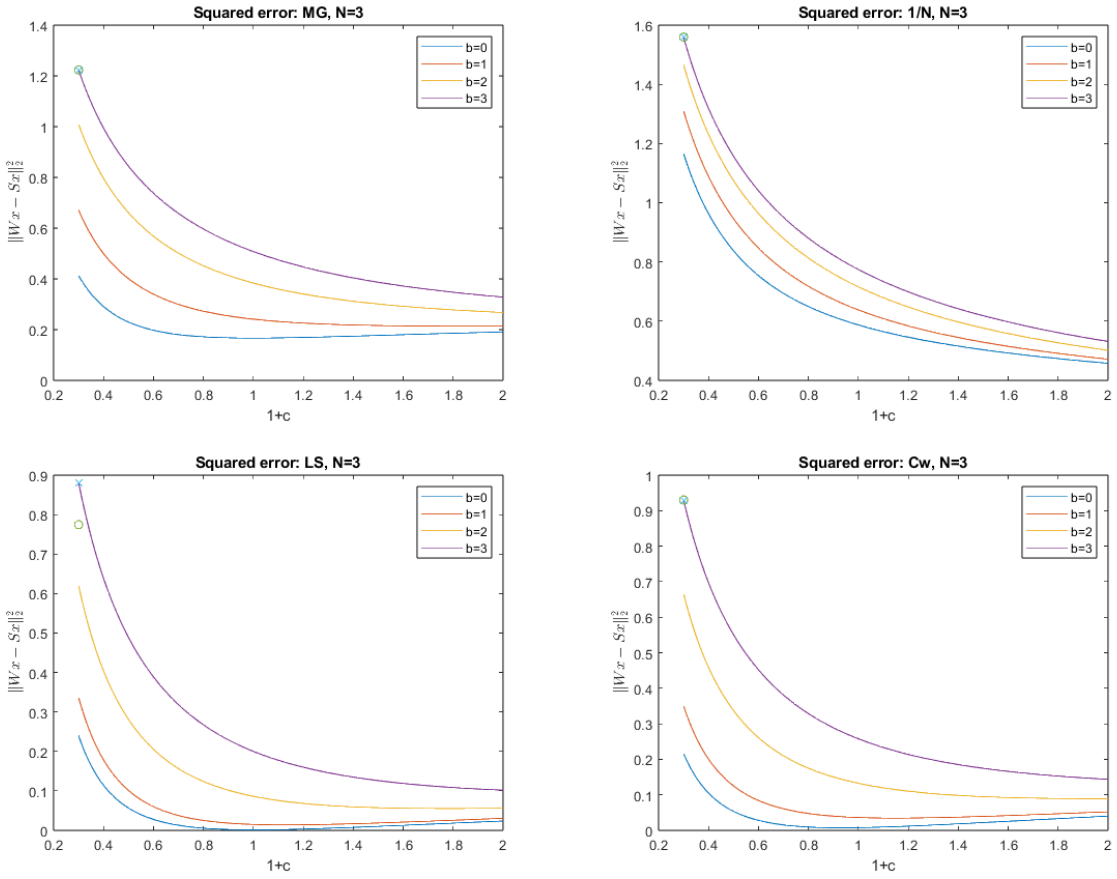

Figure 12: Squared error of Wigner and multitaper spectrograms with $N = 3$ windows and w_g weights plotted against mismatch in scale $1 + c$ for different linear chirp slopes b. The error using w_g for $1 + c = 0.3$ and $b = 3$ is marked with a cross and the error of using the w_h weights for that value with a circle.

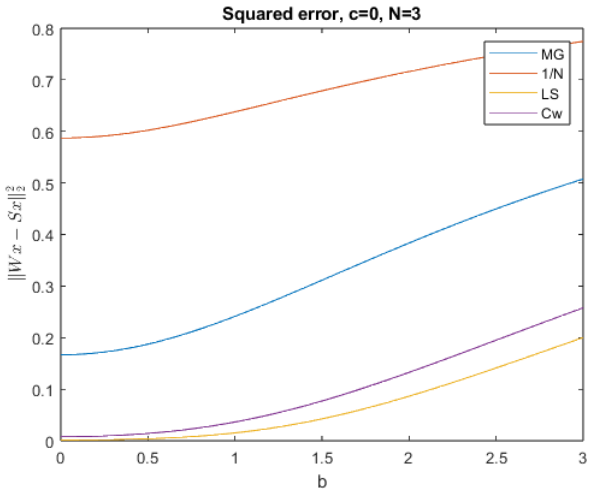

Figure 13: Squared error between Wigner and multitaper spectrograms using w_g plotted against linear chirp slope b for $1 + c = 1$

The J1 and J2 weights are not plotted since the squared error does not make sense as a comparison since they were optimized for the Jenssen distance. The Jensen distance was in fact undefined when using the Gaussian

weights. Since the RE and KT weights are not optimized against the Wigner distribution they are not included either. As seen in Figure 12 and 13 the error of the Gaussian weights generally grows with a mismatch in c and an increase of b. The error does not grow very fast with c though, especially as c is increased. This means that using a longer Hermite window than that of the signal does not give a much larger error. However, even for small c the error does not become very large until around $1 + c = 0.4$. This means that the error of using a mismatched window to the signal will not be large unless the window is much shorter than the signal. In general, the figure seems to indicate that the weights are quite robust to a mismatch in the scale of the window and the signal. Since the Cw weights approximate the Wigner distribution independent of the signal and LS does not perform worse, it seems like the chirp is simply more difficult to approximate with $N = 3$ windows and the error grows due to this.

As for the parameter b, one can see both in Figure 12 and Figure 13 that the error grows faster with the chirp slope b. There is not a big difference between the different w_g weights either. From Figure 12, one can see that the squared error of using the w_g is large for $b = 3$ and $c = -0.7$, so we will calculate w_h using these parameters. As one can see though, the improvement in the squared error of using w_h for the LS weights is neither huge nor insignificant. In practice, one would of course not know the values of b or c beforehand, but the method could still be used either by roughly approximating these or simply by using weights based on the general behavior of the signal. If, for example, one knows that the signal is typically close to a pure chirp, one could use weights that have been optimized to some small value of $1 + c$.

w_h, $c = -0.7$, $b = 3$, $N = 3$					
LS	Cw	J1($\alpha = 0.8$)	J2($\alpha = 5$)	RE($\alpha = 5$)	KT
1.000	1.000	1.000	1.000	1.000	1.000
-1.0548	-0.5714	0.0000	-0.8148	-0.7500	-0.9305
0.308	0.143	-0.0000	1.4275	-0.2288	0.5380

Table 8: The Hermite expansion weights denoted w_h normalized on the first weight for comparison. The weights were calculated numerically in MATLAB using a Hermite expansion of $K = 40$ terms.

Comparing the weights in table 8 with the weights in table 2, one can see that they are quite different. In this case, the LS and KT weights are the most similar. The LS weights now have a larger absolute value for the second weight than the first. The weights still seem alternating positive and negative except for RE. The J1 weights are now the same as MG since only the first weight is non-zero. The J1 weights will therefore be excluded from the results.

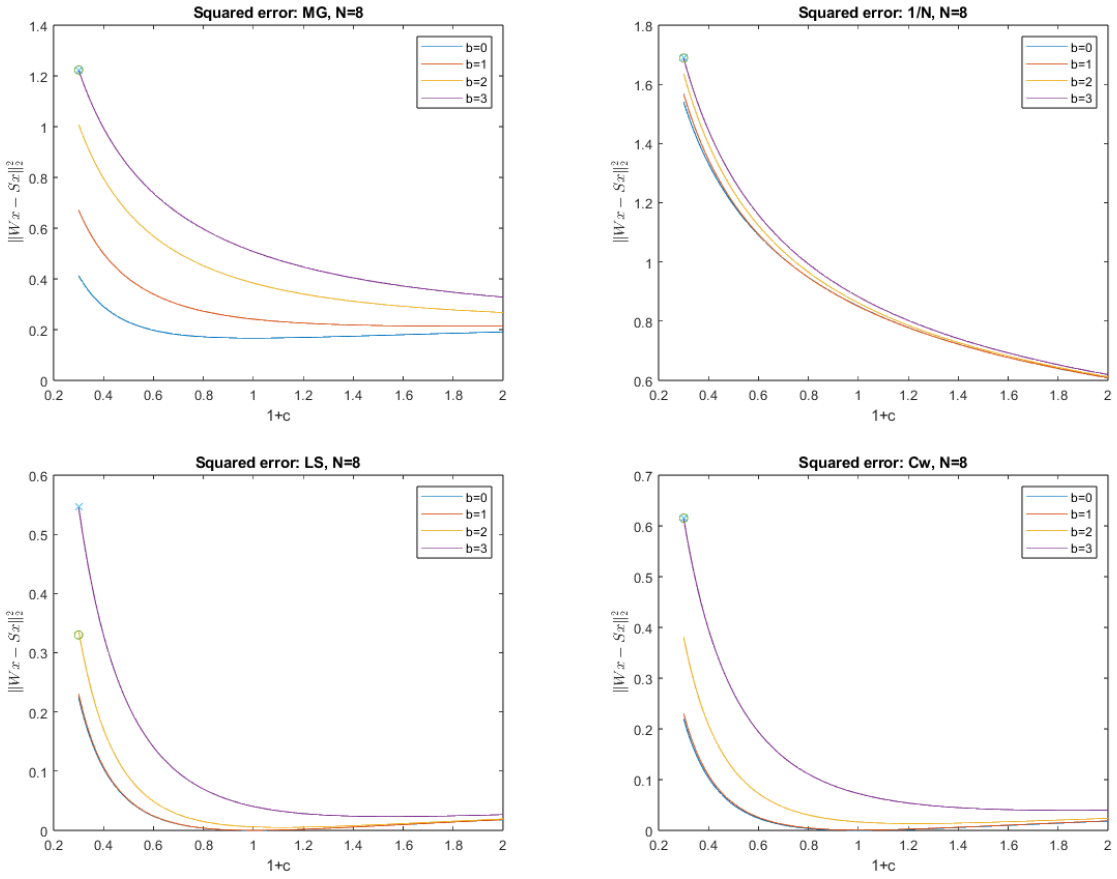

Figure 14: Squared error between Wigner and multitaper spectrograms with $N = 8$ windows using the w_g weights plotted against mismatch in scale $1 + c$ for different linear chirp slopes b. In the LS plot the largest error plotted using w_g is marked with a cross and the improvement using w_h with a circle.

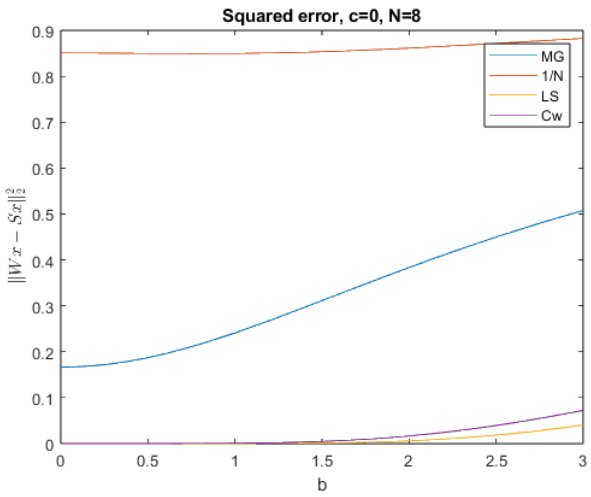

Figure 15: Squared error between Wigner and multitaper spectrograms with $N = 8$ windows using w_g plotted against linear chirp slope b for $1 + c = 1$.

The error in Figure 14 and 15 is again generally increasing with c and b, however, it seems to increase less with b for the case of $N = 8$ compared to $N = 3$. Another difference is that the improvement of using w_h seems

to be larger, which is reasonable since we are using more weights and can, in theory, better approximate the Wigner distribution.

w_h, $c = -0.7$, $b = 3$, $N = 8$					
LS	Cw	J1($\alpha = 0.8$)	J2($\alpha = 5$)	RE($\alpha = 5$)	KT
1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
-0.9948	-0.9686	0.0000	-0.5934	-0.2472	-0.9908
1.0018	0.8588	0.0000	0.8822	-0.1609	0.9945
-0.9921	-0.6392	0.0000	-0.3610	-0.2303	-0.9853
1.0060	0.3647	0.0000	0.8331	-0.0436	0.9891
-0.9813	-0.1451	0.0000	-0.1705	-0.2335	-0.9669
0.9769	0.0353	0.0000	0.8057	-0.1399	0.9474
-0.5723	-0.0039	0.0000	-0.0211	-0.2475	-0.5712

Table 9: The weights normalized on the first weight for comparison.

The weights calculated using numerical optimization with $N = 8$ windows are listed in table 9. Similar to the $N = 3$ case, the J1 weights resulted in the MG method. An interesting observation is the fact that the LS weights seem to be alternating between 1 and -1. In [18] the normalized theoretical weights when using an infinite amount of windows alternated between 1 and -1 which is similar to the case here. Comparing the KT weights to the Gaussian case and one can see that the weights, in this case, are very different. The KT weights in table 9 are similar to the LS weights while the weights in table 3 look completely different.

5.3 Evaluation

The evaluation is done for the weights calculated with the previously determined parameter choice $b = 3$, $c = -0.7$ and using $K = 40$ terms in the Hermite expansion. The concentration measures using the weights are then measured similarly to the Gaussian case with the exception of the FWHM. The TFRs in this case are not radially symmetric so the FWHM is computed along the line perpendicular to the linear chirp since ideally, one wants the TFR of a chirp to be as thin as possible. A visual analysis of a simulated signal is also done.

All of the simulations are done on the signal from (52) with $c = -0.7$, $b = 3$, a length of 256 samples, sampling frequency $f_s = 25.500$ and starting at time $t_0 = 5$ s and frequency $f_0 = 5$ Hz.

The evaluation of the robustness of noise is also done differently since only finding the time-frequency center does not make sense for the chirp. Instead, a line will be fitted to the maximum value of the TFR of the simulated signal for each time and then compared against the theoretical slope. The line is fitted using a standard linear regression. If the method is less sensitive to noise the estimated slope b should not change as the noise increases.

Finally, a visual analysis of the performance for multi-component signals will also be done by plotting the TFR of two parallel chirps at different distances apart. The evaluations are done for both $N = 3$ and $N = 8$ windows except for the multi-component signal where only $N = 3$ windows are evaluated. In this case, since there are two signals using many windows will give the same effect of introducing cross-terms so only using a few windows meaning $N = 3$ is reasonable.

w_g, $c = -0.7$, $b = 3$			
Weight	FWHM	$\gamma(C)$	$H_3(C)$
Wigner	0.0901	2.0000	-0.2075
MG	0.6575	0.3351	2.3655
$\frac{1}{N}$	1.5310	0.1496	3.3013
LS	0.3602	0.5966	1.1603
Cw	0.4188	0.5438	1.3514
J1($\alpha = 0.8$)	0.9592	0.2394	2.7587
J2($\alpha = 5$)	0.3288	0.6112	1.5107
RE($\alpha = 3$)	0.5539	0.2294	-5.1021
KT	0.2702	0.4502	-0.0715

Table 10: Comparison of concentration measures of the Wigner distribution and the multitaper spectrograms with $N = 3$ windows using the weights w_g

w_h, $c = -0.7$, $b = 3$			
Weight	FWHM	$\gamma(C)$	$H_3(C)$
Wigner	0.0901	2.0000	-0.2075
MG	0.6575	0.3351	2.3655
$\frac{1}{N}$	1.5310	0.1496	3.3013
LS	0.3288	0.6101	1.2678
Cw	0.4188	0.5438	1.3515
J1($\alpha = 0.8$)	0.6575	0.3351	2.3655
J2($\alpha = 5$)	0.3288	0.3648	2.7026
RE($\alpha = 3$)	0.4010	0.3279	-5.2650
KT	0.3288	0.6176	1.2894

Table 11: Comparison of concentration measures of the Wigner distribution and the multitaper spectrograms with $N = 3$ windows using the weights w_h.

Comparing the concentration measures of w_g in table 10 with the w_h weights in table 11, one can see that the values for the LS weights generally improved. The exception was for the Rényi entropy which actually increased a bit for w_h. However, the FWHM had a significant improvement compared to the old weights. The RE weights also improved for w_h in all measures with a much lower FWHM compared to the w_g RE. The KT weights on the other hand had a worse FWHM and Rényi entropy while the kurtosis naturally improved. For these weights, it seems w_g weights performed better. The J2 weights also had worse performance for all but the FWHM which remained the same. In this sense, the J2 weights were worse but at the same time the w_g weights did not satisfy the constraint of the TFR being positive and a comparison is not entirely fair.

Among the Hermite expansion weights, the LS weights performed the best after Wigner. The KT weights seem the next best followed by the Cw weights. Unlike the case of a pure Gaussian signal all of the weights in this case fall far behind the Wigner distribution in terms of concentration. All of the w_h weights still perform much better than MG and using $\frac{1}{N}$ is again a very poor choice.

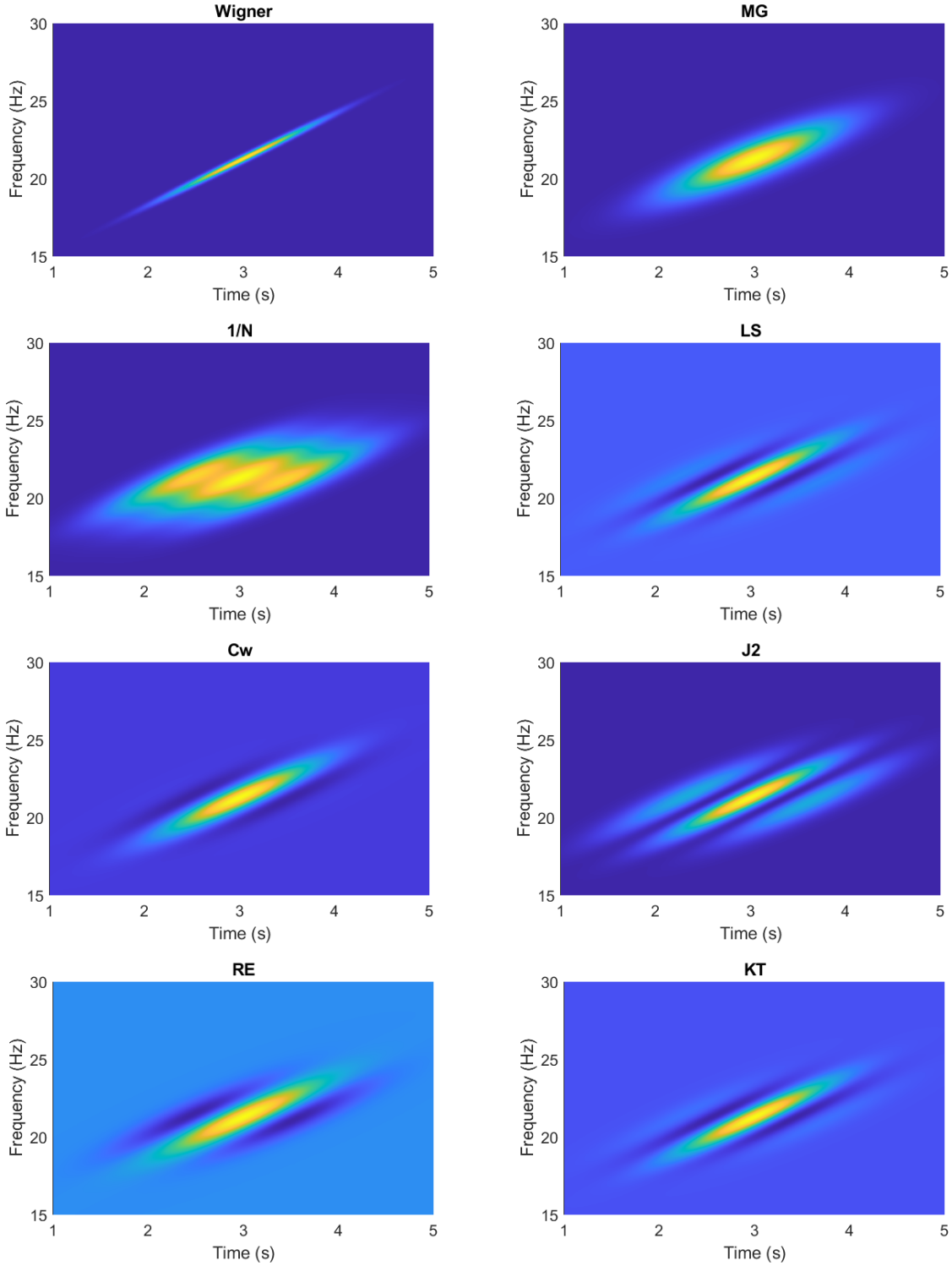

Figure 16: Wigner distribution and multitaper spectrograms of a simulated chirp with $b = 3$ *and* $1 + c = 0.3$ *using the optimized weights* w_h *and* $N = 3$ *windows without noise.*

In Figure 16, one can see that the LS weights seem to produce the most concentrated chirp with KT looking almost identical. The J2 weights also show a quite thin chirp but there are also some more apparent interference parallel to the chirp. This is probably again due to the fact that the TFR needs to be positive. One can also see that the RE weights produced a thinner peak than MG as the table showed, but it also has larger negative

values close to the chirp. The Cw weights seem to have produced the least interference or cross-terms while still looking quite localized. The very poor concentration of the $\frac{1}{N}$ weights seen in table 11 is very clear in the figure and it even looks like there are three parallel chirps. In the figure, it is also evident that the Wigner distribution is much more concentrated compared to the multitaper spectrograms and none of the weights give a result closely matching it. Since the squared error in Figure 12 was quite large even for the w_h weights it is reasonable that approximating the Wigner distribution of a chirp is difficult using only $N = 3$ windows.

w_g, 1000 simulations				
Weight	100 dB	10 dB	4 dB	-1 dB
Wigner	0.0000	0.0083	2.2500	8.1281
MG	0.4801	0.5395	0.7139	2.5503
$\frac{1}{N}$	7.3673	7.1079	6.5893	6.2217
LS	0.0750	0.0896	0.2634	2.6577
Cw	0.1106	0.1364	0.2900	2.5836
J2($\alpha = 5$)	0.0501	0.0675	0.5504	2.9063
RE($\alpha = 3$)	0.6103	0.4313	0.4376	2.7850
KT	0.0250	0.1024	1.0744	3.7574

Table 12: Comparison of the MSE of the estimated chirp slope using multitaper spectrograms with w_g weights and $N = 3$ windows with the Wigner distribution.

w_h, 1000 simulations				
Weight	100 dB	10 dB	4 dB	-1 dB
Wigner	0.0000	0.0083	2.2500	8.1281
MG	0.4801	0.5395	0.7139	2.5503
$\frac{1}{N}$	7.3673	7.1079	6.5893	6.2217
LS	0.0462	0.0625	0.5457	2.9141
Cw	0.1106	0.1364	0.2900	2.5836
J2($\alpha = 5$)	0.0457	1.4757	3.2927	5.3322
RE($\alpha = 3$)	0.1254	0.1486	0.2807	2.8253
KT	0.0529	0.0661	0.3949	2.7067

Table 13: Comparison of the MSE of the estimated slope using multitaper spectrograms with $N = 3$ windows and w_h weights with the Wigner distribution.

In table 12 and 13, we can see the MSE of the estimated chirp slope b compared to the true value for varying levels of SNR for a thousand simulations. Comparing table 12 and 13 and for low noise levels w_h outperforms w_g with the exception of the KT weights. The largest difference is for the RE weights which shows quite an improvement. When SNR decreases, it seems like w_g starts performing better with again the exception of KT. The KT weights for w_g perform very poorly for the noisy case, which should be expected as it was the behavior seen for the Gaussian signal. The KT weights for w_h on the other hand seem more robust to noise, performing even better than w_h LS at SNR = -1 dB.

Comparing the noise performance among w_h, the J2 weights performed the worst among the optimized weights. This might have been guessed from the simulated signal in Figure 16, where J2 displayed the most interference around the chirp. Similarly, the Cw weights which seemed to display the least amount of interference performs the best after MG. Since MG had the lowest error at the lowest SNR it seems like using multiple windows in this case is also not an advantage in terms of reducing noise. All of the weights again perform better than the Wigner distribution so there is clearly an advantage to using multitaper spectrograms rather than the Wigner distribution.

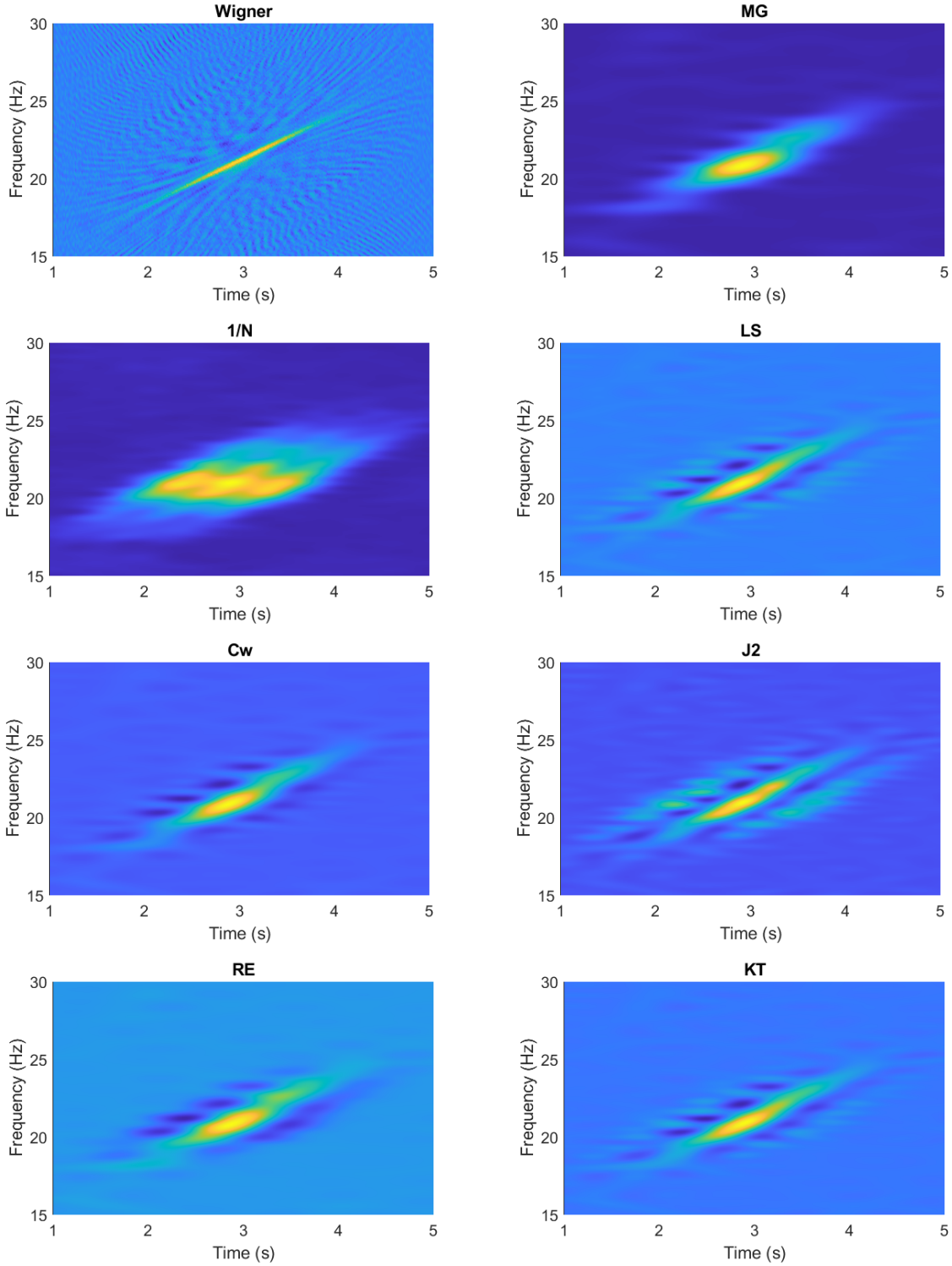

Figure 17: Wigner distribution and multitaper spectrograms of a chirp using the optimized weights with SNR = 4 dB and $N = 3$ windows.

The worse performance of J2 is very clear in Figure 17 where there is some strong interference close to the chirp. The same also holds for LS and KT which shows some more interference patterns compared to Cw. It is also clear that the Wigner distribution is more sensitive to noise since the chirp is now more difficult to discern and there are also inteference patterns that have formed in the image. That MG is the most robust to the noise is also clear since it looks very similar to the zero noise case.

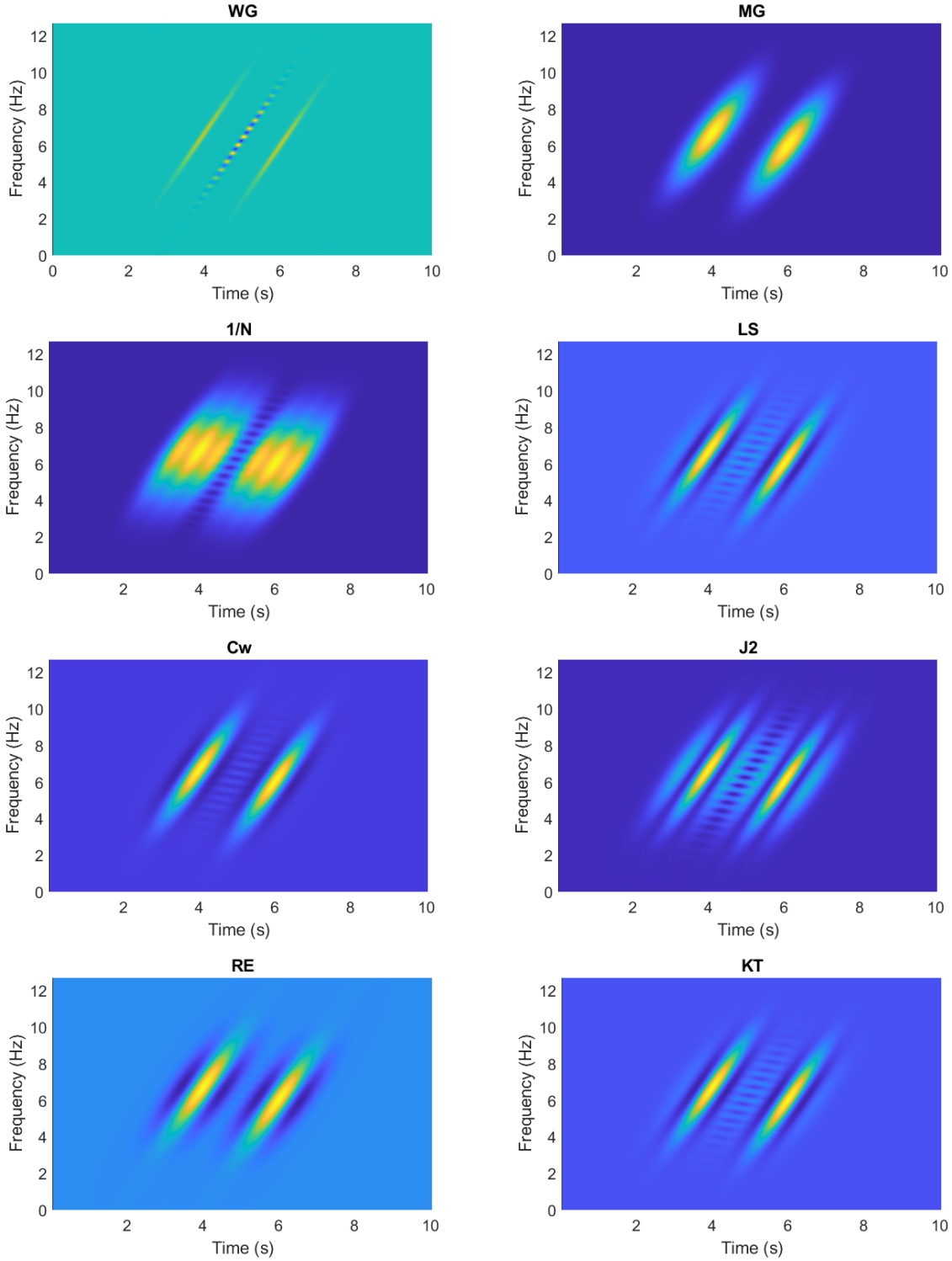

Figure 18: Wigner distribution and multitaper spectrograms using w_h with $N = 3$ windows for two simulated parallel chirps without noise.

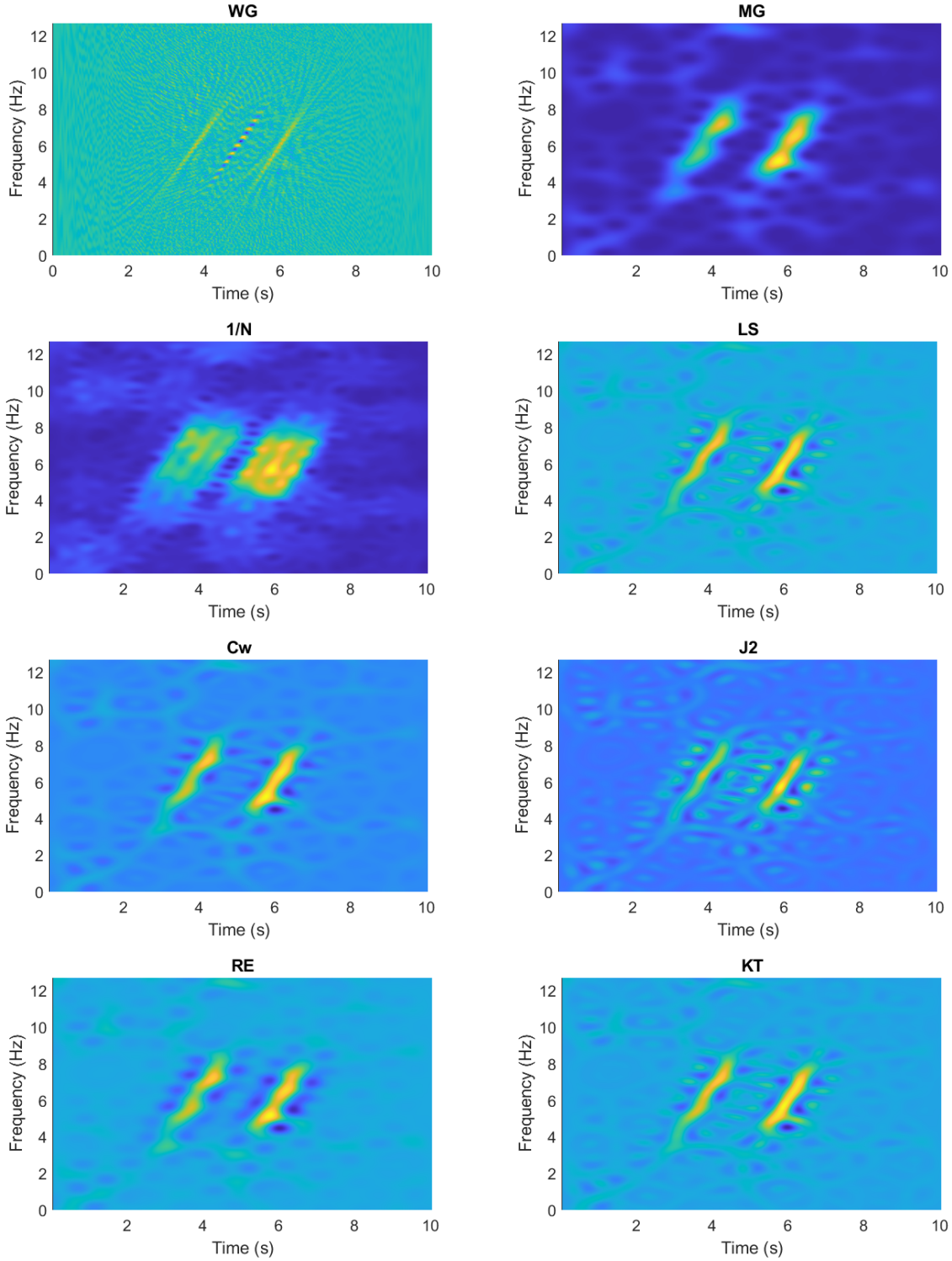

Figure 19: Wigner distribution and multitaper spectrograms using w_h with $N = 3$ windows for two simulated parallel chirps with SNR = 10 dB.

When the chirps are far away from each other, the weights as Figure 18 shows, clearly outperform the Wigner distribution. The cross-terms of the Wigner distribution are very visible between the chirps while they are barely visible for the multitaper spectrograms. The exception might be J2 that even at this separation starts to show some more visible cross-terms but it is still better than the Wigner distribution. For the noisy signal

in Figure 19 the same holds and the cross-terms are not visible for all but the J2 weights. The J2 weights still show some interference pattern in between the chirps and the chirps are also less visible compared to the other weights.

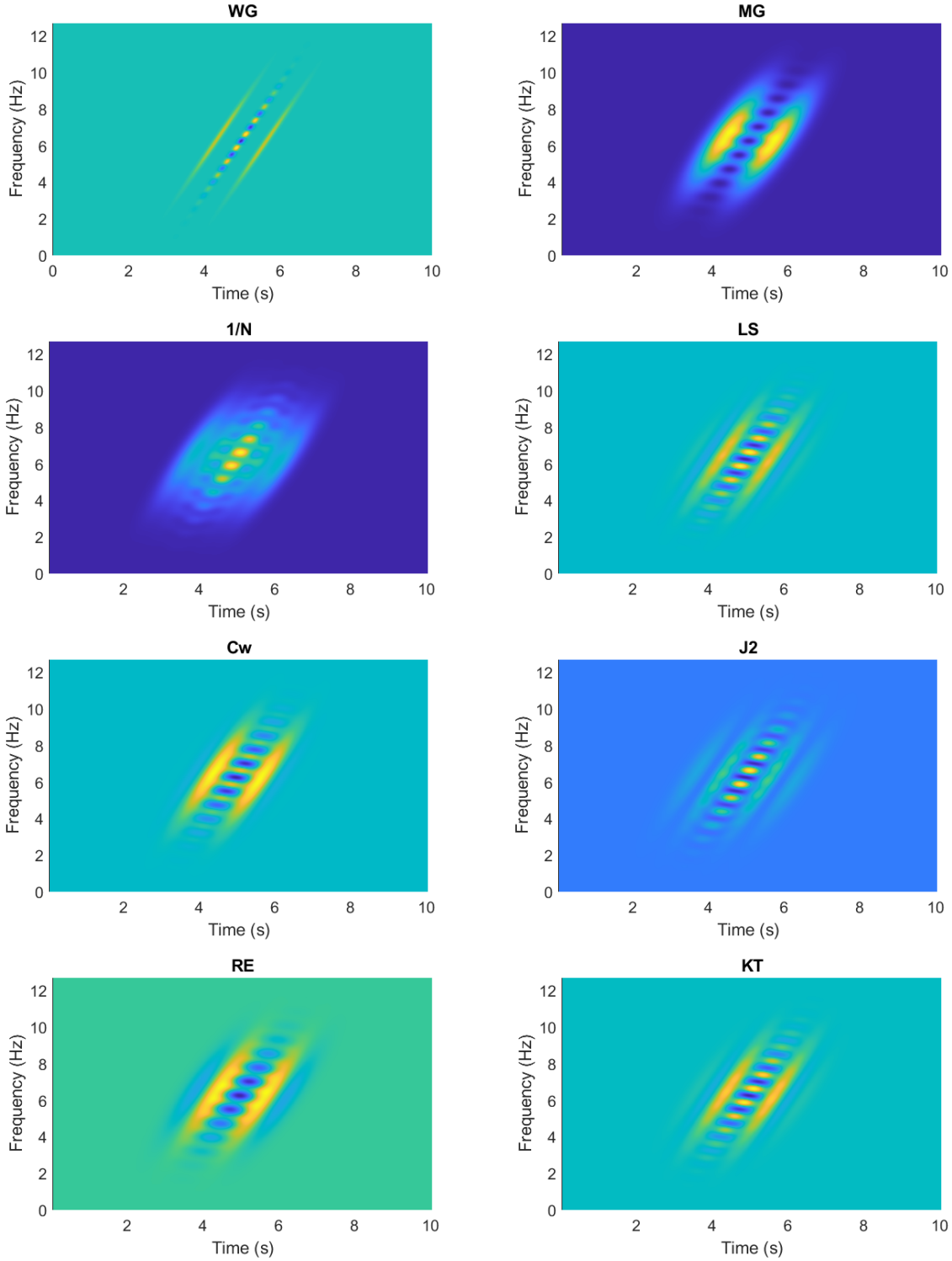

Figure 20: Wigner distribution and multitaper spectrograms using w_h with $N = 3$ windows for two simulated parallel chirps separated in time by $dt = 1.2$ s and frequency $df = 0.4$ Hz and with no noise.

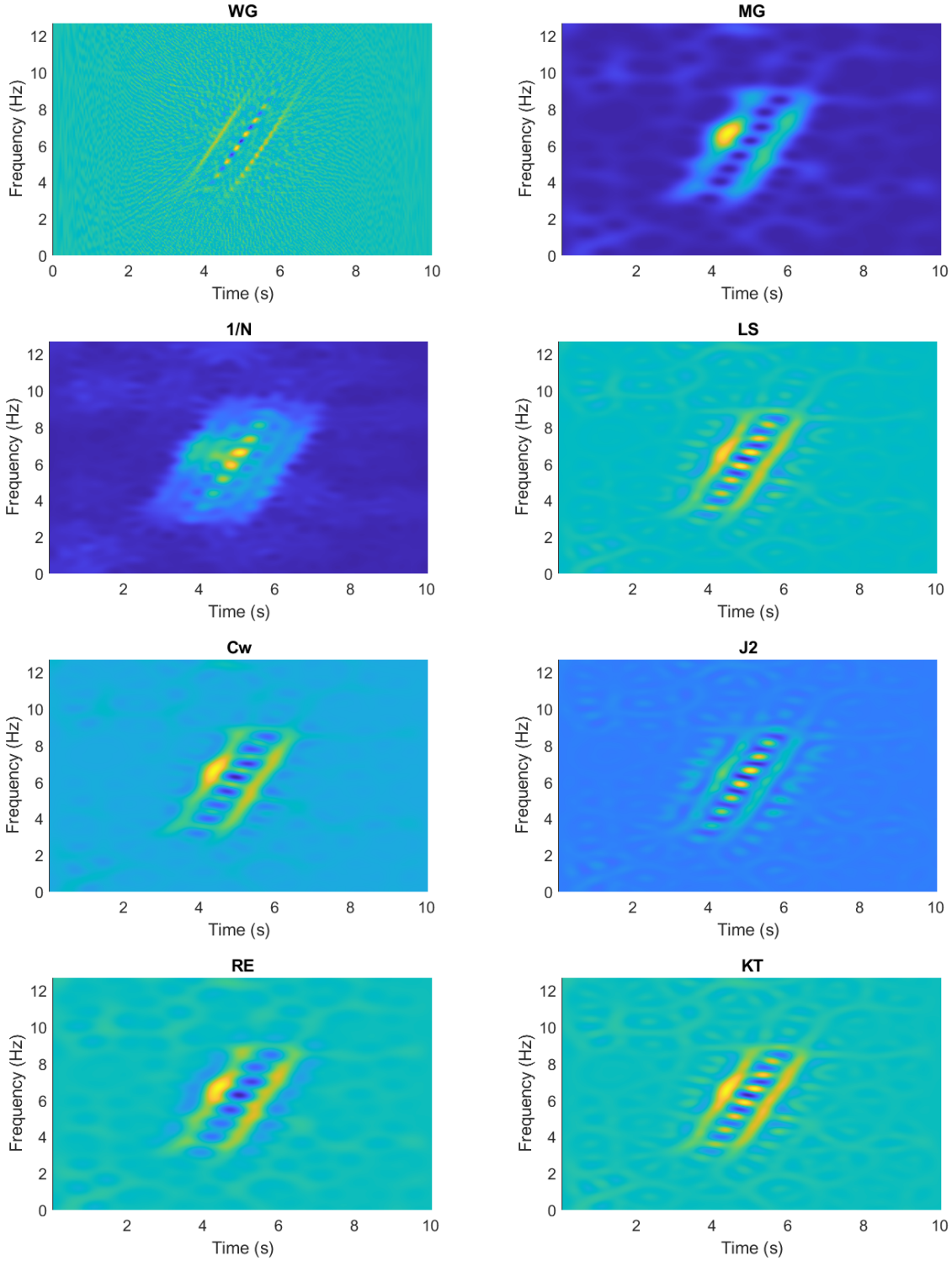

Figure 21: Wigner distribution and multitaper spectrograms using w_h with $N = 3$ windows for two simulated parallel chirps separated in time by $dt = 1.2$ s and frequency $df = 0.4$ Hz and with $SNR = 10$ dB.

In Figure 20, when the chirps are closer, the cross-terms of the multitaper spectrograms are much more apparent. In this case, there is no clear advantage to using multitapers compared to the Wigner distribution. The cross-terms of the J2 weights have made the two chirps unable to be resolved. The MG weights seem to have the best resolution since both chirps are visible and the interference due to cross-terms are minimal. The Cw and

RE weights seem to be the next best with both chirps clearly visible with the cross-terms not interfering too much. The LS and KT weights look a bit worse and the two chirps have started to become difficult to discern. The $\frac{1}{N}$ also fails to resolve the two chirps and visually looks the worst again.

When noise is added in Figure 21, the MG weights no longer seem the best at resolving the two chirps. In this case, the RE and Cw weights would be the best since both chirps are more visible while the cross-terms are still interfering minimally. The sensitivity to noise of the Wigner distribution is visible since the two chirps are very vague. The chirps for KT and LS are still quite visible but the cross-terms are also visibly distorting the result. The LS and KT weights in this case are arguably better than the Wigner distribution since the chirps are more easily distinguishable despite the cross-terms.

The performance of the multitaper spectrograms for a multi-component signal is better than the Wigner distribution both in the case of a noisy signal and when the two components are well separated. Among the different weights the RE and Cw weights were the best at resolving two close chirps both in the zero noise case and for the noisy signal. Since Cw also showed quite localized chirps when the signals were far apart these weights seem to be the best to use for multi-component signals.

w_g			
Weight	FWHM	$\gamma(C)$	$H_3(C)$
Wigner	0.0901	2.0000	-0.2075
MG	0.6575	0.3351	2.3655
$\frac{1}{N}$	2.7604	0.0804	4.1320
LS	0.2211	0.9311	0.4582
Cw	0.2702	0.8319	0.6481
J1($\alpha = 0.8$)	1.4410	0.1464	3.5244
J2($\alpha = 5$)	0.3874	0.4550	1.6691
RE($\alpha = 3$)	0.6304	0.2264	-5.7861
KT	0.9592	0.3315	-5.3225

Table 14: Comparison of concentration measures of the Wigner distribution and the multitaper spectrograms with $N = 8$ windows using the weights calculated from the Hermite expansion.

w_h			
Weight	FWHM	$\gamma(C)$	$H_3(C)$
Wigner	0.0901	2.0000	-0.2075
MG	0.6575	0.3351	2.3655
$\frac{1}{N}$	2.7604	0.0804	4.1320
LS	0.2073	1.0407	0.0046
Cw	0.2702	0.8319	0.6481
J2($\alpha = 5$)	0.2702	0.2912	3.4203
RE($\alpha = 3$)	0.3874	0.3593	-5.8252
KT	0.2073	1.0412	-0.1185

Table 15: Comparison of concentration measures of the Wigner distribution and the multitaper spectrograms with $N = 8$ windows using the weights calculated from the Hermite expansion.

Table 14 and 15 show the difference in concentration measure for w_g and w_h using $N = 8$ windows. The difference in concentration for w_g in table 14 and w_h in table 15 is larger compared to the $N = 3$ case. The w_h weights in this case outperform the Gaussian for all of the different weights. The most significant difference is for the KT weights where w_g now performs even worse than MG for all but the Rényi entropy. The KT weights from w_h have a much lower FWHM and quite a bit higher kurtosis. The Rényi entropy of the w_h KT is higher however, which again seems to indicate that the Rényi entropy is only useful up to a certain point. The RE weights of w_h are also significantly better in this case compared to w_g outperforming on all three measures. The J2 weights are also better in terms of FWHM while the kurtosis and Rényi entropy are worse. This is reasonable since these weights ensure the TFR is positive while this is not true for the Gaussian J2 weights. Similar to $N = 3$, comparing w_g and w_h for J2 is not entirely fair since the Jensen distance using w_g was undefined.

Comparing the w_h weights themselves, the KT weights are the best since they had the highest concentration values among the weights. The LS weights are very close behind with both FWHM and kurtosis almost equal in value. The J2 weights and Cw have the same FWHM but since Cw outperforms J2 on the other measures it seems like a better choice. RE had a higher kurtosis and lower Rényi entropy compared to the J2 weights but the FWHM was quite a bit higher. All of the optimized weights in general perform better than MG on most concentration values but worse than Wigner. The horrendous performance of using $\frac{1}{N}$ as weights is again evident by the poor values.

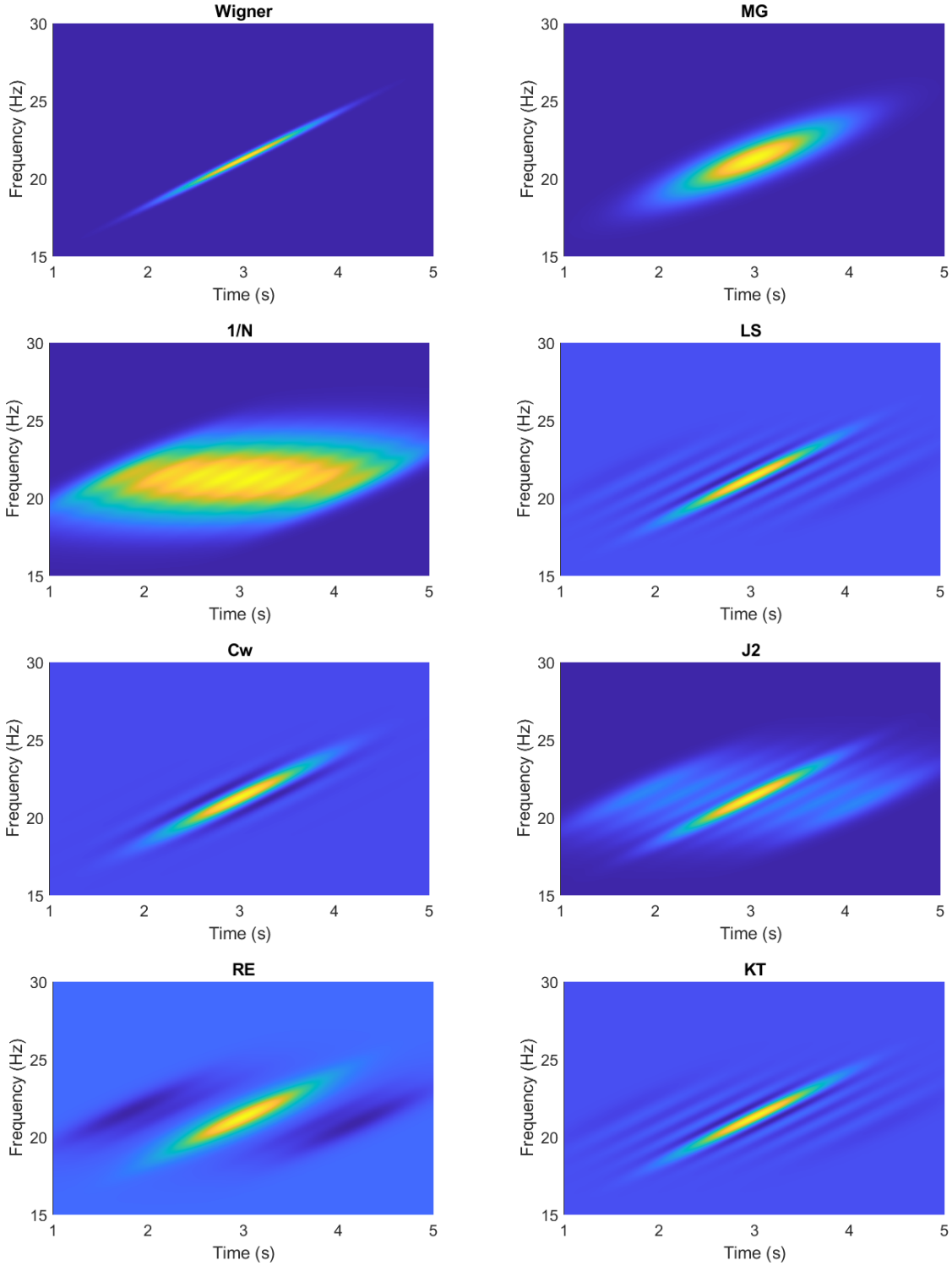

Figure 22: Wigner distribution and multitaper spectrograms of a chirp using w_h and $N = 8$ windows without noise.

In Figure 22, we can see the multitaper spectrogram of the simulated chirp signal without noise using the w_h weights and $N = 8$ windows. Compared to $N = 3$ windows, the chirps in Figure 22 show quite a visible improvement in the concentration of the chirp. All of the weights give a chirp that looks much thinner in comparison to MG. Compared to the Gaussian signal where more windows gave a negligible performance, in this case, using more windows seems beneficial.

When using many windows, the advantage of the Hermite expansion is very clear since the concentration measures improved significantly for most weights with $N = 8$ windows. One can also note that the concentration measures improved a lot by using $N = 8$ windows compared to $N = 3$. In the case of a chirp signal, there is clearly an advantage to using more windows. Using the Hermite expansion weights with multiple windows is in terms of concentration definitely a useful method for getting an improved TFR.

w_g, 1000 simulations				
Weight	100 dB	10 dB	4 dB	-1 dB
Wigner	0.0000	0.0083	2.2500	8.1281
MG	0.4801	0.5395	0.7139	2.5503
$\frac{1}{N}$	8.4953	8.7975	8.9253	8.9649
LS	0.0103	0.0165	0.4785	3.4772
Cw	0.0175	0.0252	0.3876	2.9616
J2($\alpha = 5$)	0.1168	0.2687	1.0864	3.2359
RE($\alpha = 3$)	0.4692	0.5419	0.6657	3.1135
KT	1.7524	2.7652	7.7097	11.5918

Table 16: Comparison of the MSE of the estimated slope using multitaper spectrograms with Gaussian weights with $N = 8$ windows and the Wigner distribution.

w_h, 1000 simulations				
Weight	100 dB	10 dB	4 dB	-1 dB
Wigner	0.0000	0.0083	2.2500	8.1281
MG	0.4801	0.5395	0.7139	2.5503
$\frac{1}{N}$	8.4953	8.7975	8.9253	8.9649
LS	0.0051	0.0120	0.6713	4.3948
Cw	0.0175	0.0252	0.3876	2.9616
J2($\alpha = 5$)	0.0172	0.1251	1.4082	4.6724
RE($\alpha = 3$)	0.0665	0.1040	0.3936	3.0231
KT	0.0051	0.0117	0.6438	4.2975

Table 17: Comparison of the MSE of the estimated slope using multitaper spectrograms with Hermite expansion weights with $N = 8$ windows and the Wigner distribution.

Table 16 and 17 show the noise performance using $N = 8$ windows. Comparing table 16 and 17, one can see that for low levels of noise the w_h weights generally perform better at estimating the chirp compared to w_g. The only exception is the KT weights which performed slightly worse for w_h at SNR = 100 dB. At SNR = 4 dB we can see that the w_g LS weights now have a smaller error than the w_h LS weights. The other weights except for J2 still have a smaller error for w_h. At SNR = -1 dB the LS weights for w_h still perform worse and the same holds for RE. The difference in error for the RE weights between w_g and w_h could be considered negligible though. The w_h KT weights perform much better than w_g which is reasonable since we observed the horrible performance for noisy signals earlier.

Comparing the w_h weights themselves in table 17, we can observe that the MG weights had the lowest error for the lowest SNR indicating that these are the most robust to noise. The Cw weights followed after with an error very close. Interestingly, the KT weights perform decently for the noisy signal which is the complete opposite of what occurred for the Gaussian signal. The J2 weights seem to be more sensitive to noise which was the case also for $N = 3$. The error of the LS weights is quite high so these do not seem as robust to noise when using many windows. The RE weights have an error close to Cw so these weights are also a bit more robust to the noise which was also observed when $N = 3$.

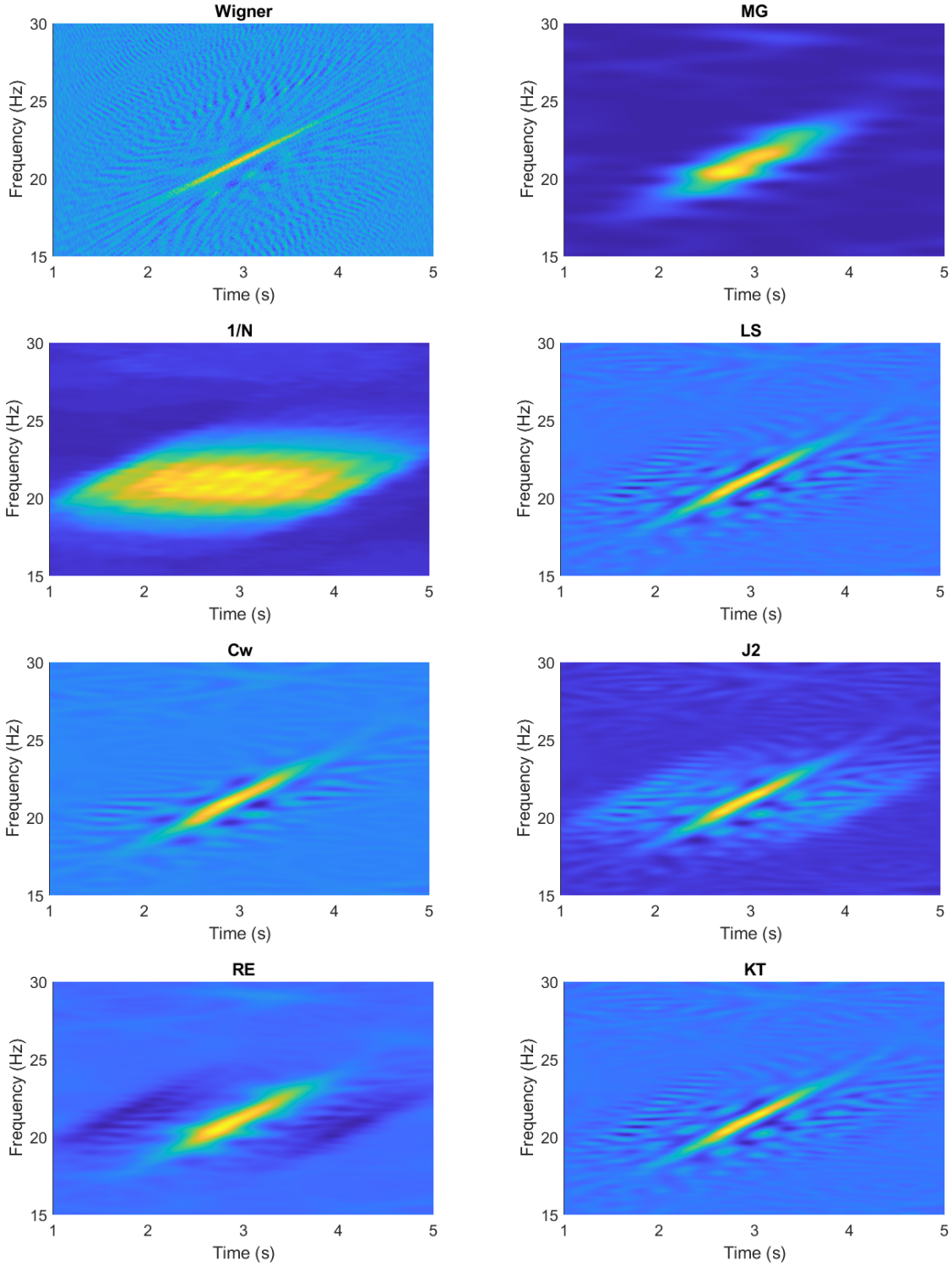

Figure 23: Wigner distribution and multitaper spectrograms of a chirp using w_h with SNR = 4 dB and $N = 8$ windows.

The worse performance of J2 to the noise is seen in Figure 23 where quite some interference is shown around the chirp. Both KT and LS also show some stronger interference patterns close to the chirp. The earlier observation of RE and Cw being robust to the noise is also visible. Using $N = 8$ windows rather than $N = 3$ does not seem to give a much worse performance for the noisy signal but the localization is visibly much better. The difference in performance is especially small for the Cw and RE weights, so using these weights with $N = 8$

windows seems to be advantageous, even in the case of a noisy signal. Although the MG weights showed good performance in table 17, it looks visually worse in the figure. The $\frac{1}{N}$ weights don't seem much affected by the noise but the concentration is extremely bad.

6. Discussion and conclusion

In the thesis, we improved upon previous work by deriving an exact solution to choosing weights by minimizing the squared error between the Wigner distribution and the multitaper spectrogram of a Gaussian signal. By using the exact solution we were able to compute the weights more than a thousand times faster than optimizing numerically. Due to the much shorter computational time, it might even be feasible to design an adaptive method that calculates the weights based on some feature of the measured signal. Computing the weights with a large number of windows is also more feasible due to the much shorter computational time. Furthermore, we presented new methods of optimizing the weights of the multitaper spectrograms by using different divergence measures such as the Jensen distance. We also presented and evaluated an approach based on maximizing the concentration of the signal using both the Rényi entropy and kurtosis as measures.

From the evaluation of the weights, we could conclude that the KT weights gave the most concentrated TFR outperforming even the Wigner distribution in all concentration measures. The only negative aspect of using kurtosis to calculate weights was that it produced some unwanted oscillations for the TFR and that it was quite sensitive to noise. The LS weights were seen to give a good approximation to the Wigner distribution and could in general be regarded as the most consistently performing weights. The Cw weights also consistently performed well both in terms of concentration and robustness to noise. In terms of concentration, they always performed worse than the LS weights but the advantage was that the weights did not require any optimization to compute. The J2 weights also performed well on most concentration measures and although they performed worse than LS it is important to recall the property that they always produced a positive TFR. Since negative values are quite difficult to interpret and the Wigner distribution for a Gaussian is non-negative this in some sense means that the J2 weights were better. The RE weights were not very good in terms of concentration in comparison to the other optimized weights but they did perform better than MG and also seemed to be more robust to noise. The only method that could be immediately discarded would be the J1 weights since these performed poorly in almost all aspects.

In the comparison between $N = 3$ and $N = 8$ windows in the Gaussian case, one can conclude that using more windows was only beneficial for the LS, KT, and Cw weights. The other weights only performed equally or worse in the evaluations by using more windows. Using $N = 8$ windows gave a very sharp concentration for the KT weights but they were practically unusable due to the sensitivity to noise. In fact, all of the weights performed worse for the noise evaluation when using $N = 8$ windows compared to $N = 3$. A conclusion is then that using multitapers with weights given by these optimization criteria is not an advantage for reducing noise. Multitapers are an advantage however in comparison to using the Wigner distribution since we could observe that the latter was much more sensitive to noise. There was a benefit to using more windows for the LS and Cw weights since these became significantly more concentrated and visually much closer to the Wigner distribution. Unfortunately, there is no single answer as to which weights one should use or how many windows. If one wants to use the multitapers for detection and knows that the signal is not very noisy then the KT weights are the best. If the signal is very noisy it might instead be better to use fewer weights with the LS or Cw weights.

As for the Hermite expansion method, there are several conclusions to be drawn. One is that the Wigner distribution of a chirp is very difficult to approximate. The w_h weights managed to decrease the squared error between the Wigner distribution and the multitaper spectrogram by a significant amount compared to using w_g but the error still remained large. The same was also true for the concentration measures where the Wigner distribution in this case outperformed all of the weights by a large margin. The Hermite expansion method improved the concentration compared to the Gaussian weights but all of the weights were still far behind the Wigner distribution. We can also conclude that using more windows was beneficial since both the squared error decreased and the concentration measures improved a lot with $N = 8$ windows compared to $N = 3$. Since we saw that the Wigner distribution was difficult to approximate it also makes sense that we then need to use more windows.

As for the weights, we can conclude that the KT weights and LS weights again yielded the most concentrated signal with both performing equally well. Again the J2 weights performed well for the FWHM but worse for

the other measures. It still had the advantage of giving a positive TFR. We can again conclude that using multiple windows was not an advantage for reducing the noise since the performance decreased when using $N = 8$ windows compared to $N = 3$. Since the concentration improved a lot by using more windows it seemed like it was worth it nonetheless to use more windows in the case of a chirp signal. We also saw that the w_h KT weights were much less sensitive to noise compared to w_g so using them even for $N = 8$ was feasible. We could also conclude that the RE seemed to perform better in this case by using multiple windows while also being robust to noise. If one has a very noisy signal while still wanting a more concentrated signal than MG they would be a good choice. The Cw weights were also a bit more robust to the noise while still having a high concentration so they would also be a good use case here. If the signal is not noisy then both LS and KT would be the best choice for a localized TFR. The J2 weights could again be useful if it is necessary to retain the positivity of the TFR but were also quite sensitive to noise.

It is worth mentioning that numerical difficulties could most likely have played a part in some of the results for the divergence and concentration measures with $N = 8$ windows. Since some of these required numerical integration optimizing with 8 weights was very slow and finding a minima seemed difficult. We also saw that we needed additional constraints for the Rényi entropy since otherwise the problem seemed unbounded. This was also observed in the values of the concentration measures where RE had a value for the Rényi entropy much lower than all of the rest while performing much worse on the other measures. The conclusion is thus that minimizing the Rényi entropy is not ideal or that the method would need to be modified further to yield good results.

The thesis leaves room for a lot of future work or research. Firstly, one could look at the analytical solution derived for the LS case. Since the analytical solution managed to compute weights very fast it might be possible to use it for some type of adaptive method that computes new weights based on certain features of the signal. One possibility would be modifying the method based on the level of noise of the signal since we observed that the LS method did not perform that well for noisy signals. The other optimization criterion would also benefit from further research since there is a lot to explore in how the parameter α for the Rényi entropy affects the weights as well as improving the numerical difficulties. In the thesis, only the two choices of windows $N = 3$ and $N = 8$ were investigated which is quite a big jump. We observed how especially sensitive to noise KT became using $N = 8$ windows. At the same time, the concentration was very good so exploring a bit more how the method might perform for fewer than 8 windows but more than 3 would be worthwhile. The idea of finding in some sense an optimal number of windows in relation to both concentration and noise is a topic of research on its own. Finally, there is the method of using Hermite expansion to optimize the multitaper weights. In this case, we only explored it on a chirp signal but it would be worthwhile to experiment with the method for other signals. We can thus conclude that there are many things worth exploring further.

A. Bargmann transform and Fock space

The Bargmann transform defined for $x(t) \in \mathcal{L}^2(\mathrm{R}^d)$ is [12]

$$(Bx)(z) = 2^{d/4} \int x(t) e^{2\pi t z - \pi t^2 - \frac{\pi}{2} z^2} dt, \quad z = t + if \in \mathrm{C}^d \tag{60}$$

Letting $\phi_0 = 2^{1/4} e^{-\pi t^2}$ denote standard Gaussian and $V_g x$ denoting the STFT of x with window g, the Bargmann transform can be expressed as

$$(V_{\phi_0} x)(t, -f) = e^{\pi i t f} e^{-\frac{\pi}{2}|z|^2} (Bx)(z), \quad z = t + if \tag{61}$$

In [12] B is shown to be a unitary operator from $\mathcal{L}^2(\mathrm{R}^d) \mapsto \mathcal{F}^2(\mathrm{C}^d)$ where $\mathcal{F}^2$ denotes the so called Fock space with the inner product defined by

$$\langle h_1, h_2 \rangle_{\mathcal{F}^2} = \int h_1(z) h_2^*(z) e^{-\pi |z|^2} dz, \quad z = t + if \tag{62}$$

Since the STFT can be written as an inner product it can be computed in the Fock space as

$$V_g x = \langle x, M_f T_t g^* \rangle_{\mathrm{L}^2} = \big\langle (Bx)(z), (B(M_f T_x) g^*)(z) \big\rangle_{\mathcal{F}^2} = \big\langle (Bx)(z), \beta_\omega (Bg^*)(z) \big\rangle_{\mathcal{F}^2} \tag{63}$$

We can find β_ω by using the relation between the Bargmann transform and the STFT in (61) giving

$$\beta_\omega (Bg)(z) = B(M_v T_u g)(z) = e^{-\pi i t f} e^{\frac{\pi}{2}|z|^2} e^{-2\pi i u v} V_{\phi_0}(T_u M_{-v} g)(x, -f) \tag{64}$$

$$= e^{-\pi i t f} e^{\frac{\pi}{2}|z|^2} e^{-2\pi i u v} (V_{\phi_0} g)(t - u)(-f + v) = \begin{cases} z = t + if \\ \omega = u + iv \end{cases} \tag{65}$$

$$= e^{-\pi i t f} e^{\frac{\pi}{2}|z|^2} e^{-2\pi i u v} e^{\pi i (f - v)(t - u)} e^{-\frac{\pi}{2}|z - \omega|^2} (Bg)(z - \omega) \tag{66}$$

$$= e^{\frac{\pi}{2}|z|^2} e^{-\pi u v} e^{-\pi v f} e^{-\pi i f u} e^{-\frac{\pi}{2}|z - \omega|^2} (Bg)(z - \omega) \tag{67}$$

$$= e^{-\frac{\pi}{2}|\omega|^2} e^{-\pi i u v} e^{\pi i \Im z \omega^*} e^{\pi \Re z \omega^*} (Bg)(z - \omega) \tag{68}$$

$$= e^{-\frac{\pi}{2}|\omega|^2 + \pi z \omega^*} e^{-\pi i u v} (Bg)(z - \omega) \tag{69}$$

Furthermore the Fock space is a reproducing kernel Hilbert space with the kernel given by

$$K_\omega(z) = e^{\pi \omega z} \tag{70}$$

The Fock space also has an orthonormal basis consisting of the functions

$$\xi_k(z) = \sqrt{\frac{\pi^{|k|}}{k!}} z^k \tag{71}$$

The Hermite functions defined in equation (15) can then be defined as the inverse mapping

$$h_k(t) = B^{-1} \xi_k(z) \tag{72}$$

which is used in Theorem 1.

B. Math Notation

Multi index

Multi-index notation is a convenient way to write multivariable formulas such as sums. An example for the multinomial expansion will be given here. The multinomial expansion $(\sum_{i=1}^{N} b_i)^n$ can be written using multi-index notation as

$$\sum_{|\mathbf{a}|=n} \binom{n}{\mathbf{a}} \mathbf{b}^{\mathbf{a}} \tag{73}$$

where $\mathbf{a} = (a_1, a_2, \ldots, a_N)$ is a vector with numbers $a_i \in \mathbb{N}$ and $\mathbf{b} = (b_1, b_2, \ldots, b_N)$. The binomial coefficient using multi-index notation is evaluated as $\binom{n}{\mathbf{a}} = \frac{n!}{a_1! a_2! \ldots a_N!}$ and the power as $\mathbf{b}^a = b^{a_1} \cdot b^{a_2} \cdots b^{a_N}$. The sum in (73) will thus sum over all combinations of $\mathbf{a}$ that sums to n. For example if $n = 2$ and $N = 2$ we get the possible combinations $\mathbf{a} = (2, 0)$, $(0, 2)$ and $(1, 1)$ giving

$$(b_1 + b_2)^2 = \mathbf{b}^{(2,0)} + \mathbf{b}^{(0,2)} + \frac{2!}{1!1!}\mathbf{b}^{(1,1)} = b_1^2 + b_2^2 + 2b_1 b_2 \tag{74}$$

as expected.

Double Factorial

The double factorial $n!!$ is defined as [11]

$$n!! = \begin{cases} n \cdot (n-2) \cdot (n-4) \cdots 4 \cdot 2, & n \text{ even} \\ n \cdot (n-2) \cdot (n-4) \cdots 3 \cdot 1, & n \text{ odd} \end{cases} \tag{75}$$

i.e the product of all odd or even integers up to n. Some useful formulas involving the double factorial are

$$\begin{aligned} (2m)!! &= 2^m m! \\ (2m-1)!! &= \frac{(2m)!}{2^m m!} \\ (2l)! &= (2l)!!(2l-1)!! \end{aligned}$$

References

[1] R.M.S.S. Abeysekera and B. Boashash. "Time-frequency domain features of ECG signals: their application in P wave detection using the cross Wigner-Ville distribution". In: *International Conference on Acoustics, Speech, and Signal Processing,* 1989, 1524–1527 vol.3. DOI: 10.1109/ICASSP.1989.266731.

[2] G.E. Andrews, R. Askey, and R. Roy. *Special Functions.* Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1999. DOI: 10.1017/CBO9781107325937.

[3] S. Aviyente. "Divergence measures for time-frequency distributions". In: *Seventh International Symposium on Signal Processing and Its Applications, 2003. Proceedings.* Vol. 1. 2003, 121–124 vol.1. DOI: 10.1109/ISSPA.2003.1224655.

[4] R.G. Baraniuk et al. "Measuring time-frequency information content using the Renyi entropies". In: *IEEE Transactions on Information Theory* 47.4 (2001), pp. 1391–1409. DOI: 10.1109/18.923723.

[5] S.G. Cunningham and J.W. Williams. "Kernel Decomposition of Time-Frequency Distributions". In: *IEEE Transactions on Signal Processing* 42.4 (1994).

[6] I. Daubechies. "Time-frequency localization operators: a geometric phase space approach". In: *IEEE Transactions on Information Theory* 34.4 (1988), pp. 605–612. DOI: 10.1109/18.9761.

[7] P. Flandrin. "A Note on Reassigned Gabor Spectrograms of Hermite Functions". In: *Journal of Fourier Analysis and Applications* (2013). DOI: 10.1007/s00041-012-9253-2.

[8] P. Flandrin. *Time-frequency/time Scale Analysis (Wavelet analysis and its applications ; v. 10).* n.d. ISBN: 9780122598708. URL: https://ludwig.lub.lu.se/login?url=https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,uid&db=cat07147a&AN=lub.6171362&site=eds-live&scope=site.

[9] G. B. Folland. "Harmonic analysis in phase space." In: Princeton University Press, 1989. Chap. 1, pp. 40–51. ISBN: 9780691085272.

[10] S. Forsgren and H. Martiros. *Riffusion - Stable diffusion for real-time music generation.* 2022. URL: https://riffusion.com/about.

[11] H. Gould and J. Quaintance. "Double Fun with Double Factorials". In: *Mathematics Magazine* 85.3 (2012), pp. 177–192. ISSN: 0025570X, 19300980. (Visited on 03/31/2023).

[12] K. Gröchenig. *Foundations of time-frequency analysis.* Applied and numerical harmonic analysis. Birkhäuser, 2001. ISBN: 0817640223. URL: https://ludwig.lub.lu.se/login?url=https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,uid&db=cat07147a&AN=lub.2509073&site=eds-live&scope=site.

[13] F. Hlawatsch and G.F. Boudreaux-Bartels. "Linear and quadratic time-frequency signal representations". In: *IEEE Signal Processing Magazine* 9.2 (1992), pp. 21–67. DOI: 10.1109/79.127284.

[14] J. Keller. "The spectrogram expansion of Wigner functions". In: *Applied and Computational Harmonic Analysis* 47.1 (2019), pp. 172–189. ISSN: 1063-5203. DOI: https://doi.org/10.1016/j.acha.2017.08.003. URL: https://www.sciencedirect.com/science/article/pii/S106352031730091X.

[15] G. Lindgren, G. Rootzén, and M. Sandsten. *Stationary stochastic processes for scientists and engineers.* CRC Press, Taylor Francis Group, 2014. ISBN: 9781466586185. URL: https://ludwig.lub.lu.se/login?url=https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,uid&db=cat02271a&AN=atoz.ebs10447704e&site=eds-live&scope=site.

[16] P. Loughlin, J. Pitton, and B. Hannaford. "Approximating time-frequency density functions via optimal combinations of spectrograms". In: *IEEE Signal Processing Letters* 1.12 (1994), pp. 199–202. DOI: 10.1109/97.338752.

[17] O. Michel, R.G. Baraniuk, and P. Flandrin. "Time-frequency based distance and divergence measures". In: *Proceedings of IEEE-SP International Symposium on Time- Frequency and Time-Scale Analysis.* 1994, pp. 64–67. DOI: 10.1109/TFSA.1994.467363.

[18] M. Sandsten. "Matched Gaussian multitaper spectrogram". In: *European Signal Processing Conference.* 2013.

[19] M. Sandsten. "Time-Frequency Analysis of Time-Varying Signals and Non-Stationary Processes". Compendium Lund University. 2022.

[20] D.J. Thomson. "Spectrum estimation and harmonic analysis". In: *Proceedings of the IEEE* 70.9 (1982), pp. 1055–1096. DOI: 10.1109/PROC.1982.12433.

[21] J. Xiao and P. Flandrin. "Multitaper Time-Frequency Reassignment for Nonstationary Spectrum Estimation and Chirp Enhancement". In: *IEEE Transactions on Signal Processing* 55.6 (2007), pp. 2851–2860. DOI: 10.1109/TSP.2007.893961.