Estimation of RF from HRV

Daniel Ågärd

Master's thesis
2023:E22

Lund University

Faculty of Engineering
Centre for Mathematical Sciences
Mathematical Statistics

Centrum Scientiarum Mathematicarum

Abstract

In this master's thesis the ability to estimate the respiratory frequency from heart rate variability measurement is analyzed. The goal was to implement a solution that is easily transferable to real time. Starting from the initial processing of data, continuing with two different spectrogram implementations, a single spectrogram and a multitaper spectrogram, combined with three different methods of spectral estimates for each time step in the two spectrograms, the respiratory frequency is estimated. A relatively limited real data set, in combination with the necessity to evaluate the different permutations of methods in a controlled environment, created the need to start on simulated data. The different permutations of methods were evaluated on the simulated data with sane defaults in order to find the best performing methods. The chosen methods were then applied to real data containing 97 different individuals. In order to maximize the different methods' capabilities the real data was divided into two data sets, one for training, and one for validation, containing 31 and 66 individuals each. The best performing methods found in the simulations were then evaluated with different parameter choices, and the weights for a multitaper spectrogram method were optimized.

The conclusion is that the respiratory frequency is possible to estimate with a low margin of error from the traditional high frequency band, $0.12 - 0.4$Hz, of the heart rate variability. The ever present time-delay of time-frequency estimates when using a spectrogram is the main contributor to the errors when estimating the actual frequency of a signal, when no noise is present. This is also the case when estimating the true respiratory frequency from the heart rate variability. If the minimization of time-delay in the frequency estimate is needed, a standard spectrogram, combined with a high heart rate will maximize the possibility of accurately estimating the respiratory frequency from the heart rate variability. For most applications outside a controlled environment however, the signal-to-noise ratio is a problem. With the small drawback of a few seconds more of extra time-delay, any multitaper spectrogram solution will perform equal or better than a single spectrogram method for time-frequency estimation. If a "real time" estimation is of no concern, a simple offset in time in post-processing of the estimated respiratory frequency will yield a result with the best of two worlds.

Acknowledgements

This Master of Science in Engineering thesis was conducted at LTH during the winter 2022/2023, as a part of the Engineering mathematics program at LTH. I would like to thank the people that made this master's thesis possible. Professor Maria Sandsten at Mathematical Statistics LTH, for being my supervisor, and providing invaluable feedback/insight throughout the work. Associated Professor Peter Jönsson at Kristianstad University, for providing the real world data that was used in order to evaluate the different methods discussed in this project.

Popular Science Summary

Estimating the breathing frequency from Heart Rate

More and more devices are today collecting data about individuals' heart rate. Smartphone applications are today using data collected from different heart rate monitoring devices, e.g. smart-watches and chest straps, and using a combination of resting heart rate and the heart rate variability in order to help people make smarter training and lifestyle choices. In this study the exploration of heart rate continues, namely what more is possible to determine with the use of only heart beats. The breathing speed of an individual is closely connected to the speed in which the heart beats. The goal was to explore the implications of this connection, and evaluate under which assumptions this connection can be used in order to predict the breathing speed of an individual.

The correlation between the heart rate variability (HRV) and the respiratory frequency (RF) is established, and the HRV is possible to estimate with good precision from the RF of an individual. The opposite is not yet fully explored, the problem with the HRV as a predictor for the RF is that the nervous system of an individual regulates a plethora of different body functions. The high frequency HRV band (HF-HRV) is the main component related to the connection between the RF and the HRV. Using this connection the precision of the estimated RF from HRV was mainly determined by time-delay, the disturbance present, and the speed of which the heart rate and the RF was changing.

With the goal of a low-latency solution on data from a controlled environment, and sufficiently close estimation, different methods of RF estimation were explored. The time-delay needed in order to estimate the RF "properly" was determined to be about $15 - 25$ seconds on data collected with the usage of an ECG. In order to make the solution more robust to disturbances, a couple more methods were evaluated, with mixed performance, and a robust solution against disturbances was found with equal or even lower time delay needed. When a chest strap was used instead of an ECG signal in order to estimate the RF, the time-delay needed exceeded 45 seconds.

Further development of the precise tracking of RF from HRV in combination with the increased precision of measuring heart rate related metrics in consumer devices, will in the future enable individuals to generate more relevant metrics about their health with less apparatus needed. A low time delay between RF and HRV will enable individuals during a data-driven workout to not only keep track of the heart rate, but also how much they are breathing.

The goal of a low-latency solution for frequency estimation once again brought to light the problems associated with knowing both time and frequency at the same time.

Contents

Introduction

More and more devices that measure heart rate are emerging on the consumer market, and it is today easier than ever before to keep track of your heart rate, and heart rate variability (HRV). During exercise people are tracking an increasing number of data fields, just in order to get an edge versus competitors, or just for the love of tracking data.

The HRV is the variation in time interval between heart beats. The HRV is assumed to display the heart's ability to adapt to changing conditions. Spectral analysis of the HRV gives insight into the cardiac health of an individual, and the state of their autonomic nervous system (ANS) [1]. HRV is commonly measured as spectral power in the frequency domain, and is traditionally divided into two frequency bands, low frequency (LF), $0.04 - 0.15$Hz, and high frequency (HF), $0.15 - 0.4$Hz. Activities related to the autonomic nervous system (ANS) are commonly deemed to have correlation to the LF-, and HF-HRV. It is generally accepted that HF-HRV correlates with activities in the parasympathetic cardiac modulation, determined by the parasympathetic nervous system (PNS). The respiratory sinus arrhythmia (RSA) is driven mainly by the PNS, and is the change in heart rate at the respiratory frequency (RF), represented as a frequency increase during inspiration, and the opposite during expiration. The main component of the HF-HRV band is the RSA, and thus the change in heart rate inside the HF-HRV band is closely related to the RF.

A performance comparison between 5 different kernel-based time-frequency distributions (TFD) was performed in the "more traditional" HF-HRV band, $0.12 - 0.4$ Hz [2]. However a narrower definition of the HF-HRV, $0.15 - 0.4$ Hz, gives a higher coherence between the RF and the power spectrum of the HRV [3]. Since a regular spectrogram was one of the 5 TFDs compared in [2] it generates the question. Is it possible to use the full HF-HRV band, and still get a robust estimate of the RSA. Even though a wider band may include extra unwanted information not related to the RSA. What it will enable however, is the estimation of a wider range of respiration rates. Instead of exploring different types of TFDs, the focus was on the spectrogram, a member in the quadratic class of TFDs, and a derivative of the spectrogram, namely the multitaper

spectrogram. This choice was made to keep the computational complexity low, in order to enable an easier real-time implementation.

The real data had the participants perform a chirp breathing task, i.e. a time dependent frequency increase/decrease in respiration rate, meanwhile their heart rate was also recorded with an electrocardiogram. In order to isolate the RSA component in the HF-HRV band with as low latency as possible, a multitude of different time-frequency methods were evaluated.

The purpose of this master's thesis is to show that a robust estimate of the RF is possible inside the full HF-HRV band, and to explore what type of time-frequency methods, and parameter-choices are preferable for a low time-delay robust implementation of the RF estimation. This would make "real time" RF another useful metric for individuals, without the need for additional measuring devices.

The thesis is structured in the following way:

- **Chapter 2: Theory**
  A brief description of the essential tools/methods used in this thesis.
- **Chapter 3: Method**
  Description of the methods and parameters used, combined with the evaluation of them on simulated data.
- **Chapter 4: Results**
  The results of the different candidate methods and parameter choices compared against each other.
- **Chapter 5: Discussion**
  Discussions around the different methods results, their pros and cons.
- **Chapter 6: Conclusion**
  The conclusion drawn from the work done, evaluation of the project goals, takeaways, and suggestions for future related work.

Theory

2.1 Spectrum

The generally accepted way to perform a spectral estimation of a discrete signal, where the latent model is unknown is by using a non-parametric method [4], i.e. the spectrum. In order to do a frequency domain sampling of an aperiodic finite energy sequence $x(n)$, containing equidistantly sampled points of data, the discrete Fourier transform (DFT) is used. For a sequence $x(n)$ of length N that is padded with $L - N$ zeroes, the DFT is defined as

$$X(l) = \sum_{n=0}^{N-1} x(n) e^{-j \cdot 2\pi n l / L}, \quad l = 0, 1, 2, ..., L-1 \tag{2.1}$$

where the absolute value $|X(l)|$ for each frequency bin is the amplitude of that bin's content, and where the normalized frequency is $\nu = \frac{l}{L}$. Due to the finiteness of the DFT bins combined with the finiteness of the sequence $x(n)$, artificial frequency components that are not actually present in the original signal are visible. The term for this phenomena is spectral leakage, and it is a non desirable part of the DFT. In order to reduce this leakage, windowing may be used in the time domain, with the goal to reduce the discontinuities amplitude at the beginning and the end of a signal [5]. An example of spectral leakage can be seen in Figure 2.1. Not applying a window to the signal $x(n)$ in Figure 2.1 will yield the exact same result as applying a rectangular window, except from a multiplication factor, to $x(n)$. A rectangular window is defined as

$$h_R(n) = \begin{cases} 1/\sqrt{M}, & 0 \leq n \leq M-1 \\ 0, & \text{otherwise} \end{cases} \tag{2.2}$$

where the length M of h_R is the same as the original sequence N. The mainlobe is the peak located at $f_n = 0.1$ in Figure 2.1(**b**). In order to reduce the spectral leakage the most commonly used window is the Hanning window defined as

$$h_H(n) = \begin{cases} \frac{1}{2}(1 - \cos \frac{2\pi}{M-1} n), & 0 \leq n \leq M-1 \\ 0, & \text{otherwise} \end{cases} \tag{2.3}$$

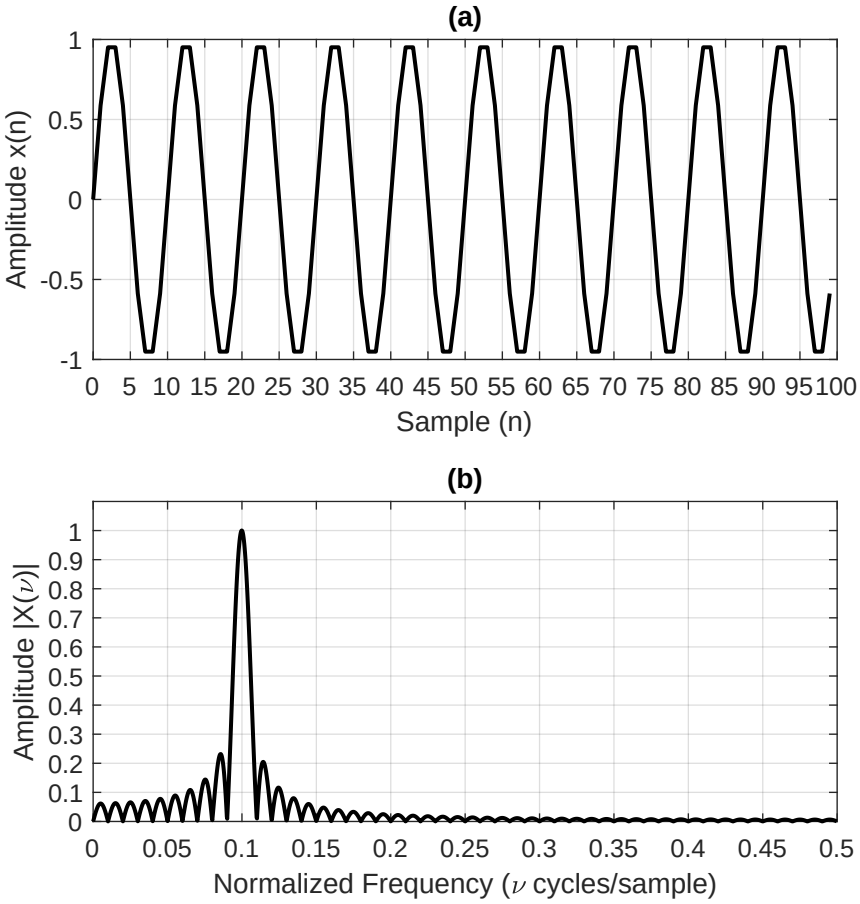

Figure 2.1: **(a)** A time-series of a sinusoidal signal with an amplitude of 1.0, normalized frequency $\nu = 0.1$ cycles/sample, and a length of $N = 100$ samples. **(b)** The corresponding absolute value of the DFT of the single frequency time-series.

An example of an implemented Hanning window, Figure 2.2(**a**), versus a rectangular window that can be seen in Figure 2.2(**b**), where a three-component sinusoidal signal is analyzed. Note that the exact location of the spectral peaks are lost when applying a window, because of the increased mainlobe size compared to a rectangular window. For example a Hanning window has a factor two wider mainlobe compared to a rectangular window

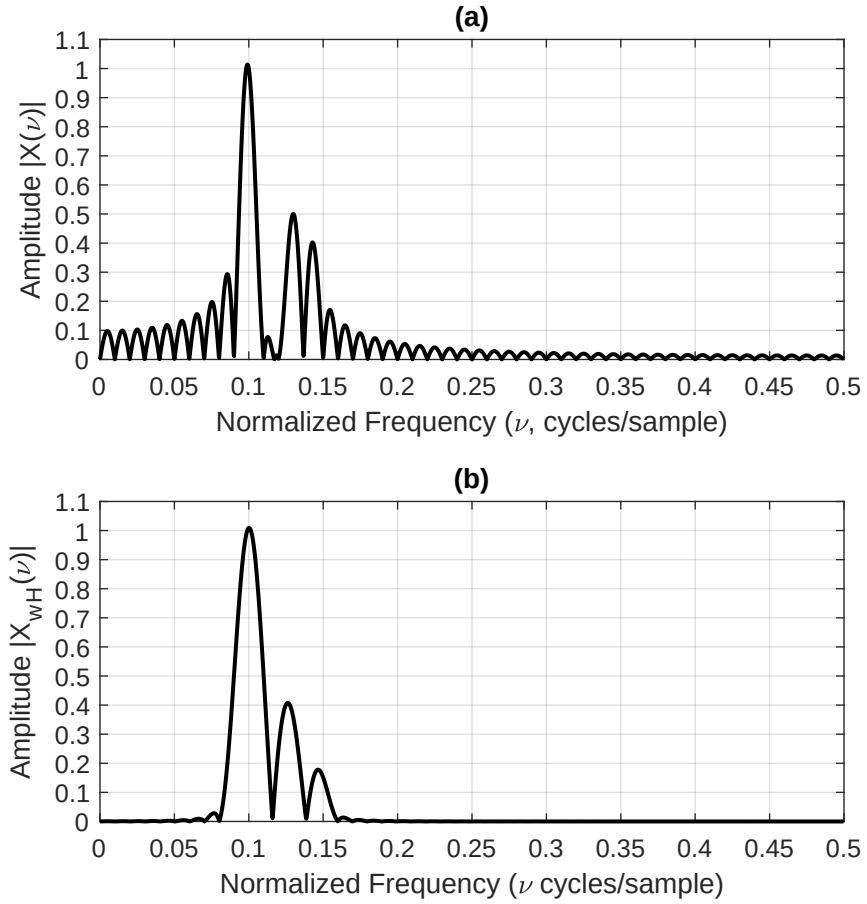

Figure 2.2: A time-series is generated by combining three sinusoidal waves with amplitude α and a normalized frequency ν. $(\alpha_1, \nu_1) = (1, 0.1)$, $(\alpha_2, \nu_2) = (0.5, 0.13)$, and $(\alpha_3, \nu_3) = (0.3, 0.14)$. In **(a)** a rectangular window is applied before the DFT, and in **(b)** a Hanning window is applied before the DFT.

2.2 Spectrogram

If the RF and HRV signal were stationary, a spectral decomposition of the full signal length would provide a good estimation of the spectral components in the signals. However, nothing about heart rate or breathing frequency are inherently stationary. In order to be able to detect a difference between two signals where the frequency content of the signals are time-varying, e.g. a linear chirp, a signal with a linearly time dependent frequency component, and an impulse signal, with the exact same spectral estimate as the chirp. A tool-set that does time-frequency analysis becomes a necessity. The window function $h(n)$, used previously in the spectrum are now sliding across the

signal $x(n)$ in time, moving a fixed amount of samples for each instant. In order to make this process a reasonable one, the window function should not increase the energy of the signal, and therefore

$$\sqrt{\sum_{n=0}^{M-1} h^2(n)} = 1 \tag{2.4}$$

The discrete-time discrete-frequency (DTDF) spectrogram is then defined as

$$S_x(n,l) = |\sum_{n_1=0}^{N-1} x_{n_1} h^*(n_1 - n + M/2) e^{-j2\pi n_1 \frac{l}{L}}|^2 \tag{2.5}$$

2.3 About the Gaussian window function

In order to enable the usage of more time-frequency methods with the goal of achieving an even better spectral estimate for each time instant of a given signal, the Hanning window is abandoned for the more analytical and "simple" truncated version of the Gaussian window. The time-frequency difference when comparing the Hanning window to the Gaussian window is small, if the appropriate Gaussian parameter is chosen. The properties of the Gaussian window makes it a way better candidate for further modifications of the spectrogram. However in order to make the discontinuities at the edges of the window sufficiently small, a quite narrow mainlobe is used in conjunction with an at least theoretical truncation.

2.3.1 Frequency resolution of the Gaussian window

A discrete Gaussian window of length M is defined as

$$h_G(n) = e^{-\frac{1}{2}(\frac{10}{M})^2 n^2}, \qquad -(\frac{M}{2} - 1) \leq n < \frac{M}{2} \tag{2.6}$$

and thus is the standard deviation of the Gaussian window $\sigma_t = \frac{M}{10}$ samples. The corresponding standard deviation in frequency $\sigma_f = \frac{L}{2\pi\sigma} = \frac{10L}{2\pi M}$, and are calculated by doing a DTDF transform of the window function $h_H(n)$. According to [6] the theoretical limit of resolution, i.e. when a bimodal distribution is perceived as a unimodal distribution, between two closely separated Gaussian functions or distributions is 2σ. This establishes the maximum frequency resolution to $2\sigma_f = 2 \cdot \frac{10L}{2\pi M}$. For example a window size of 100 combined with a 2^9 frequency bins gives a frequency resolution of 16.2975 specified in the number of frequency bins. Corresponding to a normalized frequency resolution of 0.0318 cycles/sample. This result may be helpful when later choosing the number of needed frequency bins for each spectral estimate in the spectrogram.

2.4 Multitaper spectrogram

When trying to estimate the spectrogram of a real signal, noise is often a recurring theme. In order to combat this, a multitaper spectrogram may be used. By the usage of Hermite functions, and more specifically their property of orthogonality one may reduce the noise in a signal by essentially using different realizations of the same data. A multitaper spectrogram is defined as

$$S_x(n,l) = \sum_{k=0}^{K-1} \lambda_k S_{x,k}(n,l) \tag{2.7}$$

where K is the number of multitapers used and $S_{x,k}$ is the DTDF spectrogram defined in Equation 2.5, with corresponding weights λ_k to the orthogonal window functions. The window functions are the energy normalized Hermite functions of window length M in accordance to

$$h_k(n) = \frac{H_k e^{-\frac{1}{2}(\frac{10}{M})^2 n^2}}{\sqrt{\frac{M}{10} \cdot \sqrt{\pi} \cdot 2^k k!}} \quad , \qquad k = 0, 1, 2, ..., K-1 \tag{2.8}$$

where H_k is the physicist's Hermite polynomials defined as

$$H_k(x) = (-1)^k e^{x^2} \frac{d^k}{dx^k} e^{-x^2} \tag{2.9}$$

and the discretized x points used are equally spaced numbers in $(-10, 10]$ such that the cardinality of the set is the same as the window length. If a precise energy normalization is required λ_k should be chosen such that $\sum \lambda_k = 1$. The first 4, i.e $k \in (0, 3)$, energy normalized Hermite functions, Equation 2.8, of window length $M = 100$ is shown in Figure 2.3. A comparison between a spectrogram using a single Gaussian window, Figure 2.4(**a**), and a four window multitaper is presented in Figure 2.4(**b**). Notice the amplitude reduction of the noise when using multiple realizations of the same signal. Adjusting the multitaper weights, i.e. changing the constant value for each λ_k in relation with each other changes the spectral density estimation. When changing the multitaper weights the amplitude estimation in the spectral domain may no longer be correct. This is especially true if no constraints on total energy of the weights are applied. However, by using different weights, a more task appropriate time-frequency resolution for the multitaper application may be found.

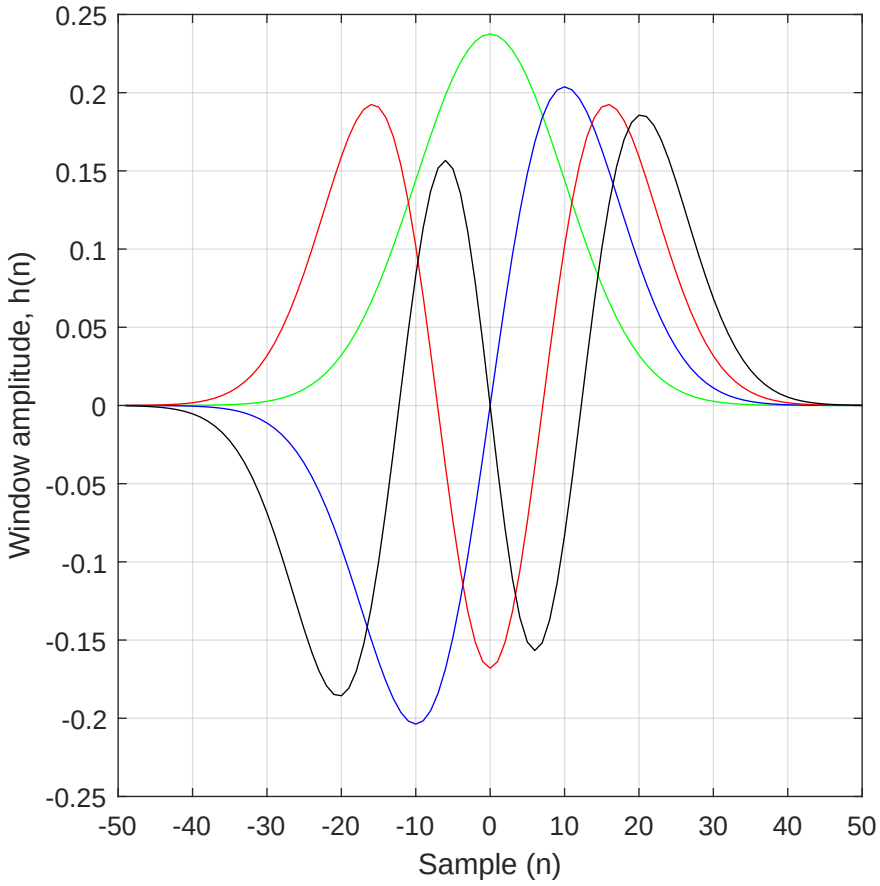

Figure 2.3: The first 4 energy normalized Hermite functions. Green is h_0, blue is h_1, red is h_2, and black is h_3.

2.5 Time lag in frequency estimation

The fact that time and frequency are closely connected makes the exact estimation of frequency at a given time an impossibility. When using a window in order to get time resolution the drawback is that a time-delay of half the window size in time is introduced to the spectral estimate. This is due to the fact that in order to estimate the spectrum of a given signal at a given time, samples on both sides of the given time need to be included in order to get an estimate. Combine this with the fact that the frequency resolution is inversely proportional to the time used in order to estimate the spectrum, cementing the need for some time-lag in the frequency estimation of a signal. An artificial exaggerated example of this problem is presented in Figure 2.5.

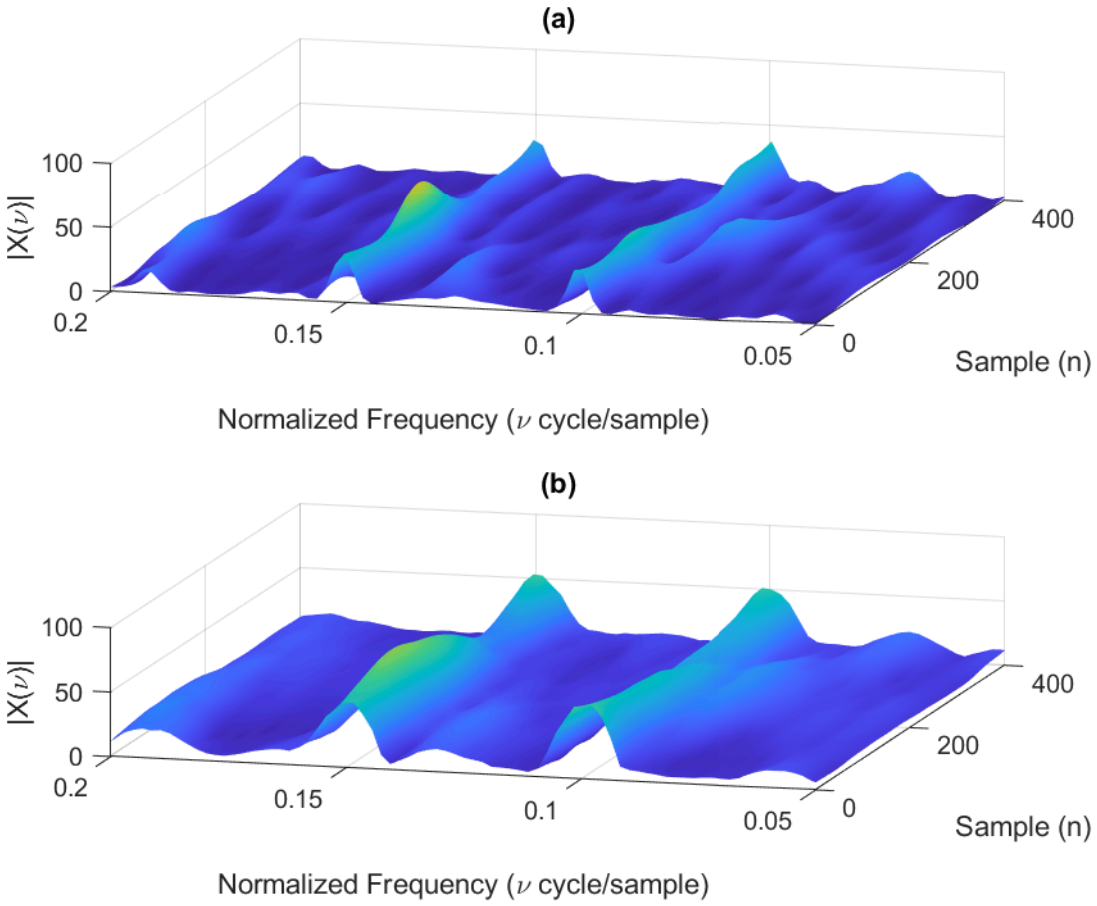

Figure 2.4: A time-series is generated by combining two sinusoidal waves with amplitude α and a normalized frequency ν, $(\alpha_1, \nu_1) = (1.0, 0.1)$, and $(\alpha_2, \nu_2) = (1.0, 0.15)$, combined with Gaussian distributed noise, $G(0, 2)$. In **(a)** a single Gaussian window with length 50, and in **(b)** a multitaper spectrogram with 4 orthogonal Hermite windows.

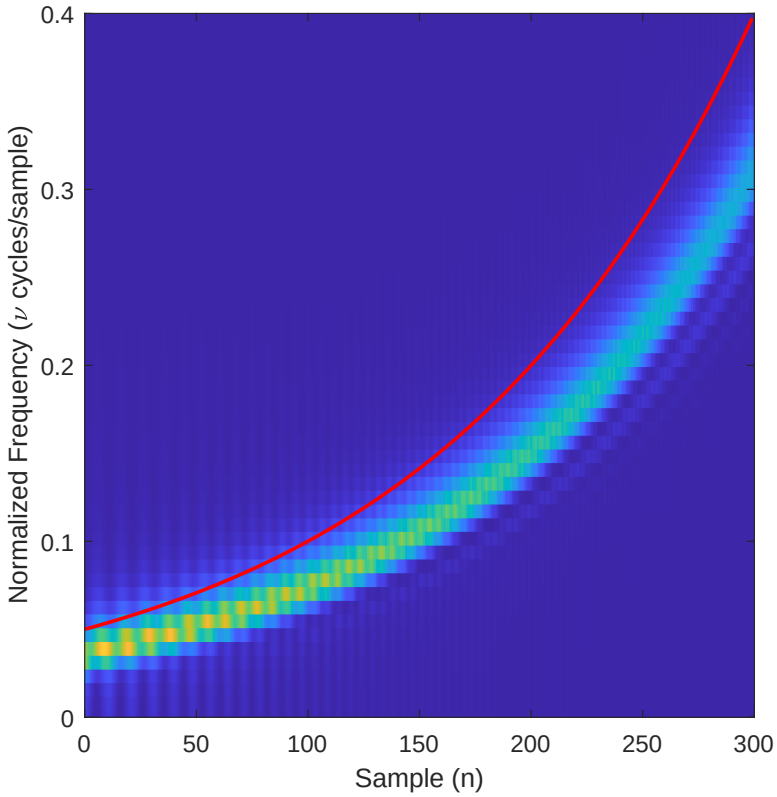

Figure 2.5: An exponential chirp starting at a normalized frequency of $f_n = 0.05$ at sample $n = 0$, rising to $f_n = 0.4$ at sample $n = 300$. The window length is 300 samples in order to exaggerate the effect of the frequency estimation time-delay that occurs when using windowing. The red line is the actual signal frequency.

Method

In order to validate different methods' ability to correctly estimate the spectral component of the two connected signals, a simulated HRV/RF model is a good starting point. In order to motivate and explain the different steps of the applied methods, a set of default parameters will be used on simulated data, together with visualizations in the form of figures. In the final implementation of the spectrogram peak estimation, the discrete parameter values will be evaluated with appropriate step size, and the continuous values, i.e. the multitaper weights will be optimized for using the Nelder-Mead simplex algorithm, with the costfunciton defined as the , [7], with bound constraints. All the pre-processing described below on the artificial data will be performed, in combination with the different spectral estimation methods on real data after the initial evaluation is done. The sampling frequency is from now on $F_s = 4$ Hz, instead of normalized frequency, in accordance with the real data, just to make comparisons easier between the real and simulated data. Conversion from normalized frequency to "real" frequency is done by $f = f_n \cdot F_s$. However in order to implement an artificial model, and validate it against non-stationary signals, frequency sweeps are needed. Methods of generating linear and exponential in time frequency sweeps are presented in Appendix A.1.

3.1 Construction of simulated data

With an understanding on how to generate chirps of appropriate time-frequency behavior, the focus is now on how to generate simulated HRV data combined with the corresponding respiratory data. The model used was an integral pulse frequency modulation (IPFM) model presented in [8]. A high level description of the IPFM is that using the fact that RSA exists, one may use a sinusoidal chirp respiratory signal in combination with noise to generate a jitter in a stationary heart beat frequency. The resulting heart rate signal will behave in a similar way to the HRV collected from an ECG. In Figure 3.1(**a**) is a real HRV signal shown, and an artificially created HRV signal is visualized in Figure 3.1(**b**) .

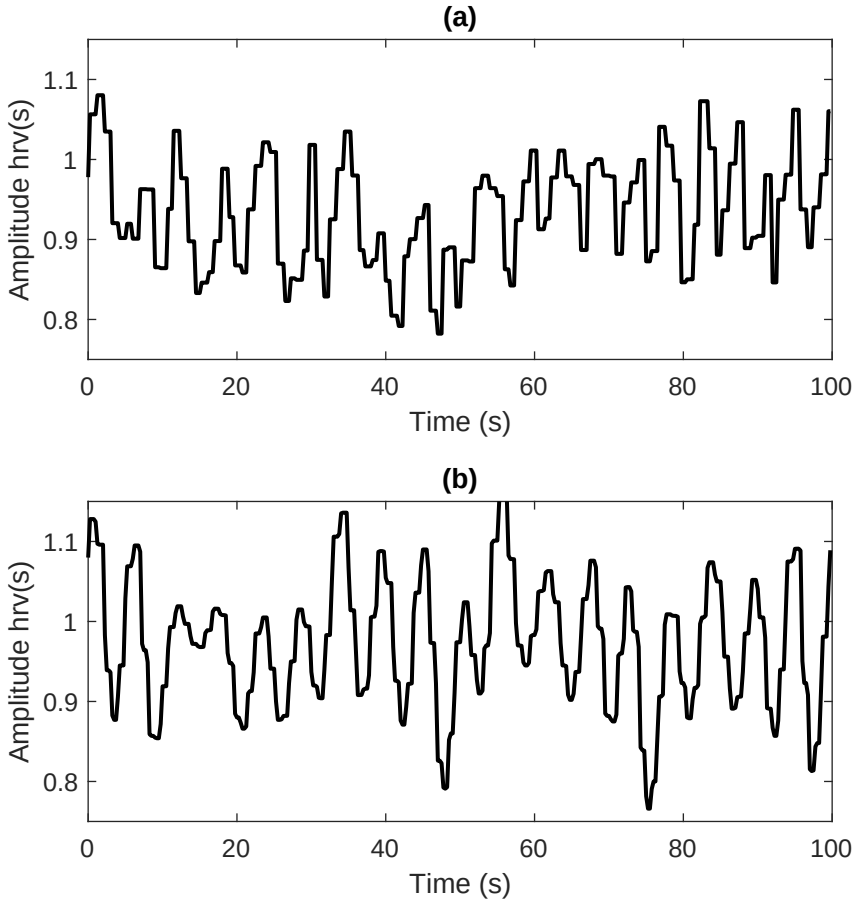

Figure 3.1: Comparison between **(a)** an recorded HRV, and **(b)** an artificially generated HRV. The RF is stationary at 0.15 Hz in both time-series, and the heart rate is at about 60 bpm (1 Hz).

3.1.1 Validating spectral estimate methods on simulated data

In order to perform a spectral estimate on a time-series it needs to be of zero mean. In this time-frequency estimation it is known that the RSA component of a HRV is located in the HF-HRV band, and thus everything in the HRV-LF band needs to be removed by a filter, both in the respiratory data, and the HRV data. The filter used was an order 120, finite-impulse-response (FIR), high-pass filter, with a stop-band at 0.08 Hz and a pass-band at 0.15 Hz. Two corresponding snippets of simulated time-series, respiratory and HRV are presented in Figure 3.2. The non processed signals, RF and HRV are shown in Subfigure 3.2(**a-b**) and the same signals after processing in Subfigure 3.2(**c-d**).

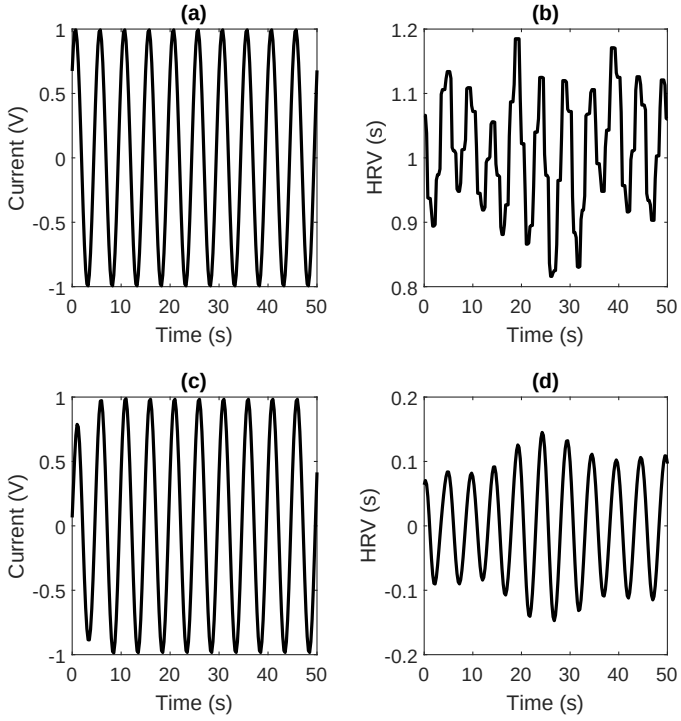

Figure 3.2: Snippets of artificial RF and HRV data, **(a)** unprocessed respiratory time-series, **(b)** unprocessed HRV time-series. The mean value of respective time-series was removed, and high-pass filtered with a stop-band at 0.08 Hz. Processed; **(c)** respiratory time-series, and in **(d)** HRV time-series.

The spectral estimation was after the initial processing performed with two different methods. A spectrogram using a single window in accordance with Equation 2.5, and a spectral estimation using a multitaper with four orthogonal windows, in accordance with Equation 2.7. Four windows were seen as a good middle ground between a single window, and "many" windows, since the combination of many windows and small window size will introduce low frequency artifacts which are impossible to remove. The spectral estimation in the spectrogram moved one sample in time for each new estimate, and the window size was set to a fixed 100 samples. The length of the real data is about 300s, with a sample rate of $F_s = 4$. Half the window size is needed in order to remove transient effects from the spectral estimation on each side of the data, leading to a data loss of 25 seconds. Add on top of that the time-variability of the HRV signal in the real data, and a compelling argument for the necessity of data extension is presented. The added benefit of generating spectrograms without losing information at the edges of the data is also compelling argument.

3.2 Naive data extension

Two naive ways of performing data extension is either by expanding the function in an odd, or even manner. Each of the two methods have pros and cons. In Figure 3.3 is the asymmetrical and the symmetrical extensions performed on a sinusoidal signal with normalized frequency $f_n = 0.01$. The estimated frequency peak in Figure 3.3(**b**) is at $f_n = 0.0098$. The estimated frequency peak in the symmetrical extension is at $f_n = 0.0039$. In Figure 3.4 are the same two extensions performed on a sinusoidal signal with a $\pi/2$ phase offset. The peak with the most amplitude in the frequency spectrum for the asymmetrical extension is located at $f_n = 0.0159$, and the peak for the symmetrical extension is located at $f_n = 0.0103$. The conclusion here is that an asymmetrical extension will on average generate a frequency increase at the edges of the signal, while a symmetrical extension will generate a frequency decrease at the edges. Since the extension of the real data aims to pre-saturate the first windows of the spectrogram, the natural choice became a symmetrical data extension. The extension length was chosen to be of a length corresponding to half of the spectrogram window size. This data extension is mainly used as a tool in order to get more real time data spectral estimates, with the limited data set available for this estimation of RF from HRV.

Next step in the process is to estimate the spectral density of both the artificial time-series, by using the multitaper spectrogram, Equation. 2.7, and the naive symmetrical extension presented above. The first 400 samples, and its initial extension of the artificial signals, and their corresponding spectrogram, for window-length $M = 100$ samples, is shown in Figure 3.5 for the respiratory data, and in Figure 3.6 for the corresponding HRV data.

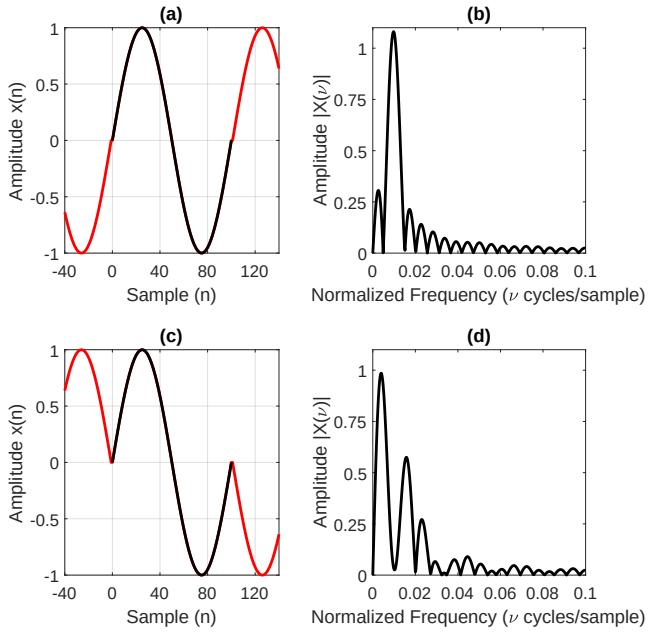

Figure 3.3: Naive data extension of a sinusoidal signal with normalized frequency $f_n = 0.01$. **(a)** Asymmetrical extension of the signal together with the frequency spectrum in **(b)**. In **(c)** is the symmetrical extension of the signal, and on the right **(d)** the corresponding frequency spectrum.

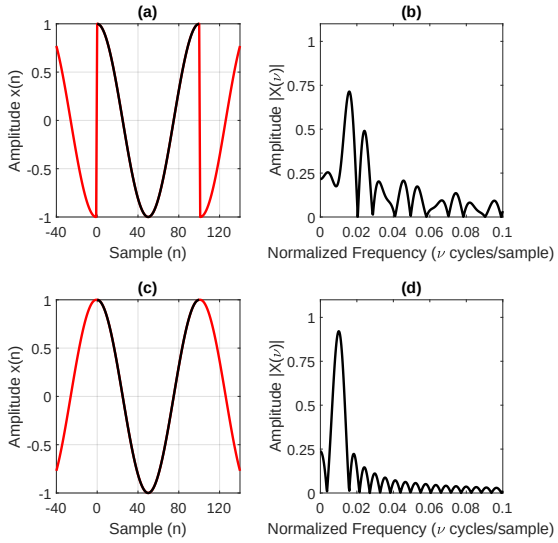

Figure 3.4: Naive data extension of a sinusoidal signal with normalized frequency $f_n = 0.01$, and a phase offset $\pi/2$. **(a)** Asymmetrical extension of the signal together with the frequency spectrum in **(b)**. In **(c)** is the symmetrical extension of the signal, and on the right **(d)** the corresponding frequency spectrum.

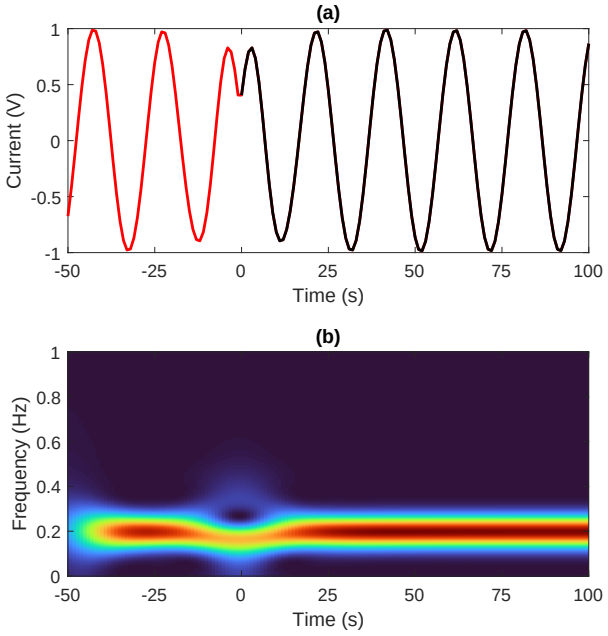

Figure 3.5: The spectral estimation **(b)** of the artificial respiratory time-series **(a)**. Notice the symmetric extension's effect on the spectral estimate at time 0s.

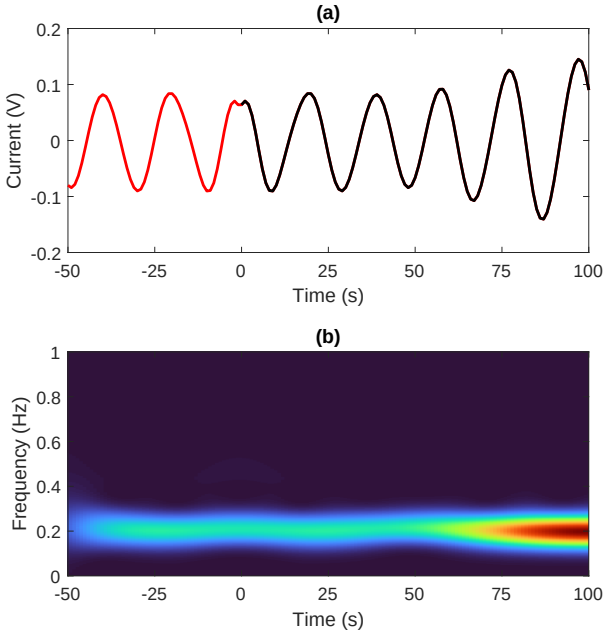

Figure 3.6: The spectral estimation **(b)** of the artificial HRV time-series **(a)**. This extension is almost the "best" case for a symmetric extension.

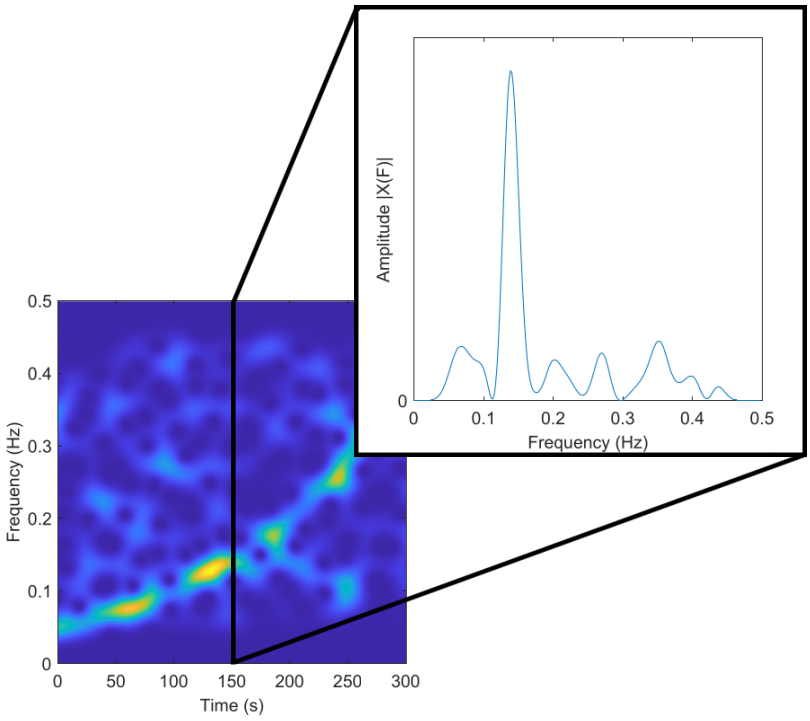

Figure 3.7: An example showing the estimated spectral amplitudes at the sample located at 150s.

3.3 Peak detection in the spectral domain

In order to evaluate the estimated RF from the processed HRV signal, the highest amplitude peak in each single spectral estimate needs to be determined for each time-step in the single window and multitaper spectrograms used, $K = 4$. In Figure 3.7, a single spectral estimate for sample located at time 150s is shown. Three peak detection algorithms were evaluated, each with an increasing order of intricacy. All algorithms are described for a single window at a single time-step n, i.e. the same scenario that Figure 3.7 shows. First `NP` - Naive Peak Detection, the algorithm selects the highest peak with respect to $S_x(n, l)$ without considering any info about other local peaks.

The second algorithm $\texttt{SP}_A$ - Squared Peak Detection, uses information of a chosen number of A peaks location and height, performs a sorting of the A peaks w.r.t height, and chooses the highest peak as the central peak location. The other $A-1$ sorted peaks contributes to a movement of the central peak location in a way that is proportional to that peak's height, divided by the squared distance in bins from the central peak location. The third algorithm $\texttt{DP}_A$ - Democratic Peak Detection also uses information about A peaks location and height. The A peaks here contributes to the peak location selection process in the sense of a weighted mean.

`NP` Naive Peak Detection

Initialize localpeaks lp = []
for each $l \in (0, F_s/2)$ **do**
  if $S_x(n, l-1) < S_x(n, l)$ & $S_x(n, l+1) < S_x(n_i, l)$ **then**
    lp.add($[S_x(n, l), l]$)
  end if
end for
Sort lp //w.r.t $S_x(n, l)$
(peak, loc) = lp(0) //Create a tuple containing the amplitude and location of //the highest peak. Where loc corresponds to bin for peak freq.

`SP`$_A$ Squared Peak Detection (with A peaks used)

Initialize localpeaks lp = []
for each $l \in (0, F_s/2)$ **do**
  if $S_x(n, l-1) < S_x(n, l)$ & $S_x(n, l+1) < S_x(n_i, l)$ **then**
    lp.add($[S_x(n, l), l]$)
  end if
end for
Sort lp //w.r.t $S_x(n, l)$
peaklocation = lp.(0).loc + $\lfloor \sum_{i=1}^{A} \frac{1}{(\text{lp(0).loc} - \text{lp(i).loc})^2} \text{lp(i).peak} \rfloor$

`DP`$_A$ Democratic Peak Detection (with A peaks used)

Initialize localpeaks lp = []
for each $l \in (0, F_s/2)$ **do**
  if $S_x(n, l-1) < S_x(n, l)$ & $S_x(n, l+1) < S_x(n_i, l)$ **then**
    lp.add($[S_x(n, l), l]$)
  end if
end for
Y $\leftarrow \sum_{i=1}^{A} \text{lp(i).peak}$ //Total amplitude denoted Y
peaklocation = $\lfloor \frac{1}{Y} \sum_{i=1}^{A} \text{lp(i).peak} \cdot \text{lp(i).loc} \rfloor$

The goal of the last two methods is to increase the robustness of the spectral estimates for each time step in the algorithm in a naive way, by the assumption that there exist other spectral information that is valuable, and not only the location of the highest peak. The goal is reduced bias in the the peak detection.

3.4 Evaluation of the three peak detection methods on simulated data

The given naming conventions from section 3.3 are used together with different parameter values in the evaluation. Apart from the name of the algorithm described above, the superscript s/m indicates if it was a single spectra used, or a multitaper spectra. Subscript indicates how many peaks that were used in the algorithm. If a superscript is not present it is a single spectra method, and if a subscript is not present, it is a single peak method.

The previously mentioned parameters that are changeable for the spectrogram, namely windows size M, and the weights for Hermite window inside the multitaper, are later being optimized using real data. The parameters are during this evaluation locked at "sane defaults". Window size $M = 100$ and if a multitaper is used the weights of the 4 orthogonal windows are $1, 0.75, 0.5, 0.25$. The window function is multiplied with the signal in the time-domain, and a relatively fast changing window with respect to the window length may introduce unwanted low-frequency components into the signal. Therefore only one multitaper window $h_0(n)$ allowed to influence the spectral estimation with full effect, and when the slopes of the multitaper window functions, Figure 2.3, becomes steeper, the effect of that windows influence on the spectral estimate is reduced accordingly. Even with "sane defaults" an evaluation of the different peak detection methods for in the spectral domain will help bring a light on problems associated with estimation of frequency in a temporal sense.

Using the IPFM method, a stationary heart rate frequency of 120 BPM and a linear chirp as the base RF, ranging from 0.2Hz to 0.8Hz, in the span of 300 seconds. A frequency jitter was added to artificial HRV signal with a mean squared error (MSE) of 1610 ms. This was repeated for 1000 artificial HRV and respiratory pairs, using the previously described processing on both. First a comparison between the estimated RF versus the true frequency used in order to generate the HRV for 7 different permutations of methods, `NP`, naive peak detection with a multitaper using 4 Hermite windows ($\texttt{NP}^m$), detection of 5 peaks in combination with the usage of the `SP` detection algorithm ($\texttt{SP}_5$), detection of 5 peaks in combination with a 4 window multitaper ($\texttt{SP}_5^m$), as well as detection up to 10 peaks ($\texttt{SP}_{10}^m$), detection of 5 peaks in combination with the `DP` detection algorithm ($\texttt{DP}_5^m$), as well as detection of 10 peaks ($\texttt{DP}_{10}^m$). The evaluation of `DP` on the estimate from a single window spectrogram is not considered due to the abundance of noise present in a single window spectral estimation, and the `DP`

algorithm's inability to suppress it.

The different performance metrics used, mean-squared-error (MSE), Bias^2, and variance (VAR), are calculated from the error at each sample between the estimated RF versus the estimated RF from HRV, for all simulated data, between the true RF, the spectrogram estimated RF and the spectrogram HRV estimation of the RF. Results of errors for all time steps in all simulations presented in Table 3.1, the unit is Hz.

	NP	NP^m	SP$_5$	SP$_5^m$	SP$_{10}^m$	DP$_5^m$	DP$_{10}^m$
MSE	0.0230	0.0240	0.0250	0.0250	0.0250	0.0086	0.0082
Bias^2	0.0027	0.0040	0.0036	0.0036	0.0036	0.0032	0.0029
Var	0.0200	0.0210	0.0210	0.0210	0.0210	0.0055	0.0053

Table 3.1: The performance of the different methods estimating the true RF from simulated HRV data spectrogram. All units are in Hz.

	NP	NP^m	SP$_5$	SP$_5^m$	SP$_{10}^m$	DP$_5^m$	DP$_{10}^m$
MSE	0.0230	0.0240	0.0250	0.0250	0.0250	0.0082	0.0078
Bias^2	0.0026	0.0039	0.0036	0.0036	0.0036	0.0029	0.0026
Var	0.0200	0.0200	0.0210	0.0210	0.0210	0.0053	0.0052

Table 3.2: The performance of the different methods estimating the spectrogram estimated RF from the simulated HRV data spectrogram. All units are in Hz.

It is possible in Table 3.1 to decompose how large part of the MSE is attributed from the time-lag effect of the spectrogram. The chirp that generates the artificial RF and HRV is linear which results in an easy calculation of the time-lag frequency component of the MSE. The frequency lag is 0.0250 Hz. The MSE contribution is therefore $6.25e^{-4}$, the rest of the MSE is most likely from the phase difference between the two signals in the IPFM method, combined with the jitter in the HRV signal, and the loss of the precise peak location due to windowing. When comparing the estimation of HF-HRV with the RF, no frequency lag is present, and the contribution to the MSE is only from the errors described in the last sentence times two. The small difference between the two tables are to be expected, the slow rate in frequency increase of the generation chirp inside the IPFM method reduces the effect of time-lag, in the spectral estimation, and that is why the true RF is close to the estimated RF from the artificial HRV signal.

The goal here is to find a couple of methods that are suitable for evaluation on real data, with the task to estimate the spectrogram estimated RF from the HRV data spectrogram. The measurements in Table 3.2 is most closely related to this task. The performance of the single window methods; NP and SP$_5$ concludes that NP is far superior compared to its competitor. Regarding the multitaper methods; NP^m, SP$_5^m$, SP$_{10}^m$, DP$_5^m$, and DP$_{10}^m$. The performance of both DP$_5^m$ and DP$_{10}^m$, are unrivaled, however the only

difference between the two of them is the number of peaks used in the peak detection. A parameter that is easily changeable if needed when evaluating the method on real data. When inspecting the performance of the other multitaper implementations, it becomes clear that the performance of $\texttt{NP}^m$ is marginally better than the other methods, excluding the democratic ones. Since the goal is low complexity, and the squared peak detection methods did not increase the performance, the KISS principle is respected.

The conclusion here is that the usable methods are; $\texttt{NP}$, $\texttt{NP}^m$, and $\texttt{DP}^m_{10}$. The simplicity of the $\texttt{NP}$ makes it a valid candidate for real data. The two other methods, $\texttt{NP}^m$ and $\texttt{DP}^m_{10}$, implements a multitaper, which should perform better than a basic spectrogram on real data, due to the noise present in real data recordings. Questions about optimum window size, and the possible benefit of optimizing window weights, are left to be answered by real data.

Results

All RF estimates below are done on the real HRV data. The subjects were asked to breathe in sync with a metronome that increased its frequency from $0.2 - 0.35$ Hz during the time span of 5 minutes. The speed of the frequency increase is about half of the speed used in the evaluation of the different methods on simulated data in Section 3.4.

4.1 Data

The sample rate is 4 Hz, the same as in the simulation example. Both the respiratory frequency and time in between the heart beats were recorded. The complete data set of 97 recordings is here divided into two disjoint sets, one for the purpose of "training", with a size of 31 recordings, i.e. optimizing of multitaper weights, and one set for validation, containing 66 recordings. Furthermore, the validation data set has a subset containing "extra bad" data, size 28, containing partly corrupted recordings of RF, or both RF and HRV. This partition of real data is done since the "extra bad" data introduced a non realistic problem into the evaluation process, the issue that no true answer of what the RF actually was is available. After initial processing, the spectrogram/multitaper used for frequency estimation had a frequency resolution of 512 samples in between zero and the Nyqvist frequency. A visual representation of the 31 training data recordings are shown in Figure 4.1.(a), and the corresponding figure for the validation data is presented in Figure 4.1.(b). Note that in order to generate a valid representation of the data, the time duration of the figure is that of the shortest recording.

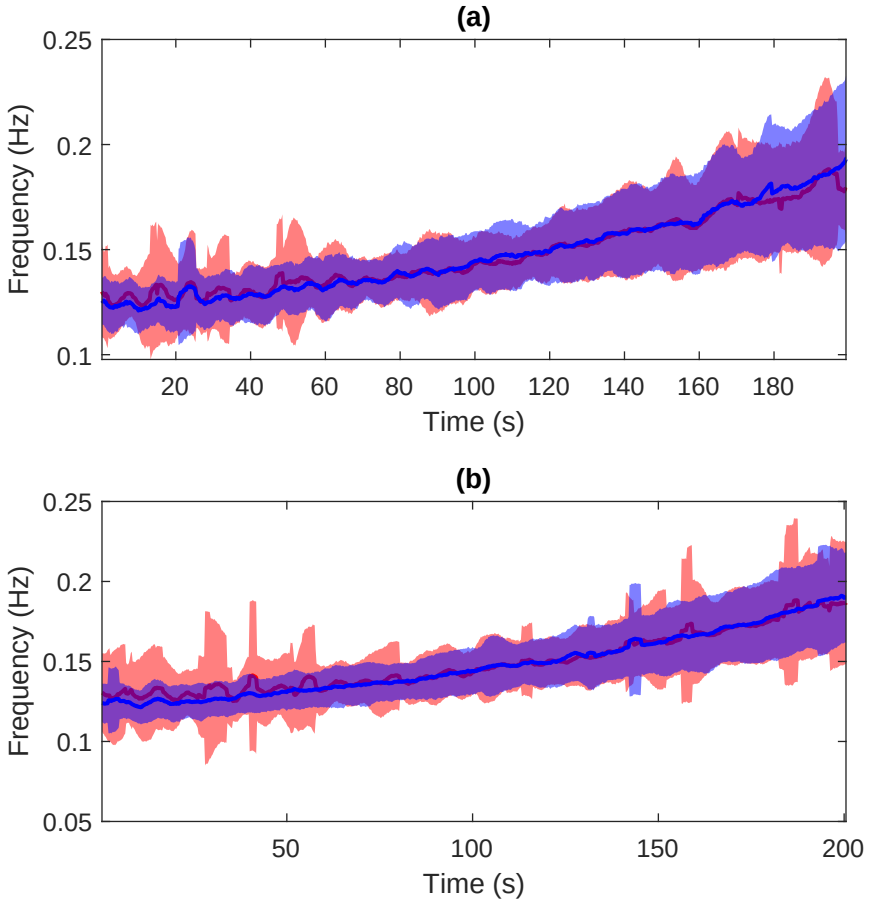

Figure 4.1: Visual representation of the training data in **(a)**, and a visual representation of the validation data in **(b)**. In both figures a spectrogram with a window size of $M = 100$ was used. The blue line is the estimated respiratory frequency and the blue shading represents the standard deviation of the former. The red line is the estimated HRV-HF frequency, and the red shade is the standard deviation of the former.

4.2 Evaluation

The different performance metrics used for evaluation here are the same that was presented for simulated data in Section 3.4. The result of the naive peak detection in the spectral region is presented in Table 4.1. One example on how the different window sizes match up against each other are shown in Figure 4.2. The evaluation of naive peak detection in the spectral domain for a multitaper implementation with "sane defaults", i.e. the weights for the first 4 energy normalized Hermite functions are; $(\lambda_0, \lambda_1, \lambda_2, \lambda_3) = (1, 0.75, 0.5, 0.25)$.

NP	val.-bad data			val. data		
Window size	**MSE**	**Bias**^{2}	**Var**	**MSE**	**Bias**^{2}	**Var**
60	$3.3e^{-3}$	$1.2e^{-6}$	$3.3e^{-3}$	$3.5e^{-3}$	$7.1e^{-7}$	$3.5e^{-3}$
80	$1.5e^{-3}$	$5.0e^{-7}$	$1.5e^{-3}$	$1.4e^{-3}$	$4.6e^{-7}$	$1.4e^{-3}$
100	$7.3e^{-4}$	$2.0e^{-6}$	$7.3e^{-4}$	$8.0e^{-4}$	$1.8e^{-6}$	$8.0e^{-4}$
120	$5.6e^{-4}$	$2.5e^{-6}$	$5.5e^{-4}$	$6.5e^{-4}$	$2.2e^{-6}$	$6.5e^{-4}$
140	$5.2e^{-4}$	$2.5e^{-6}$	$5.2e^{-4}$	$6.1e^{-4}$	$2.4e^{-6}$	$6.1e^{-4}$
160	$5.1e^{-4}$	$3.0e^{-6}$	$5.1e^{-4}$	$5.7e^{-4}$	$2.6e^{-6}$	$5.7e^{-4}$

Table 4.1: The performance of single peak spectral estimation, using a single window spectrogram, NP, on validation data without the bad data, and on all validation data.

$\texttt{NP}^m$	val.-bad data			val. data		
Window size	**MSE**	**Bias**^{2}	**Var**	**MSE**	**Bias**^{2}	**Var**
60	-	-	-	-	-	-
80	-	-	-	-	-	-
100	$3.0e^{-3}$	$1.9e^{-6}$	$3.0e^{-3}$	$3.4e^{-3}$	$4.4e^{-7}$	$3.4e^{-3}$
120	$1.2e^{-3}$	$1.8e^{-9}$	$1.2e^{-3}$	$1.1e^{-3}$	$4.7e^{-8}$	$1.1e^{-3}$
140	$4.5e^{-4}$	$1.7e^{-6}$	$4.5e^{-4}$	$4.6e^{-4}$	$1.6e^{-6}$	$4.6e^{-4}$
160	$3.6e^{-4}$	$3.7e^{-6}$	$3.5e^{-4}$	$3.8e^{-4}$	$3.4e^{-6}$	$3.8e^{-4}$

Table 4.2: The performance of single peak spectral estimation using a 4 window multitaper spectrogram, i.e. $\texttt{NP}^m$, on validation data without bad data, and on all validation data.

The results of the estimation error are shown in Table 4.2. The window sizes $M = 60$ and $M = 80$, combined with the Hermite windows inside a multitaper generates low frequency components in the spectral domain as mentioned before in Section 3.4, and any sort of peak detection is not valid due to the high energy of the artificial components. In Figure 4.3 the RF estimated from the HRV multitaper is presented, notice the failure in detection for the two shortest window sizes. In both Table 4.1 and in Table 4.1 there exists one outlier. In Table 4.1 it occurs when the window size is 80 for the validation data with the bad data exclusion, and in Table 4.2 it occurs at the window size of 100 also in the validation excluded the bad data. The outlier in the expected bias-variance tradeoff demonstrates the sensitivity of using short window lengths, in this case with $F_s = 4$ Hz. Only $4 - 5$ full periods of the RSA is observed inside each window, assuming a RF of 0.2 Hz.

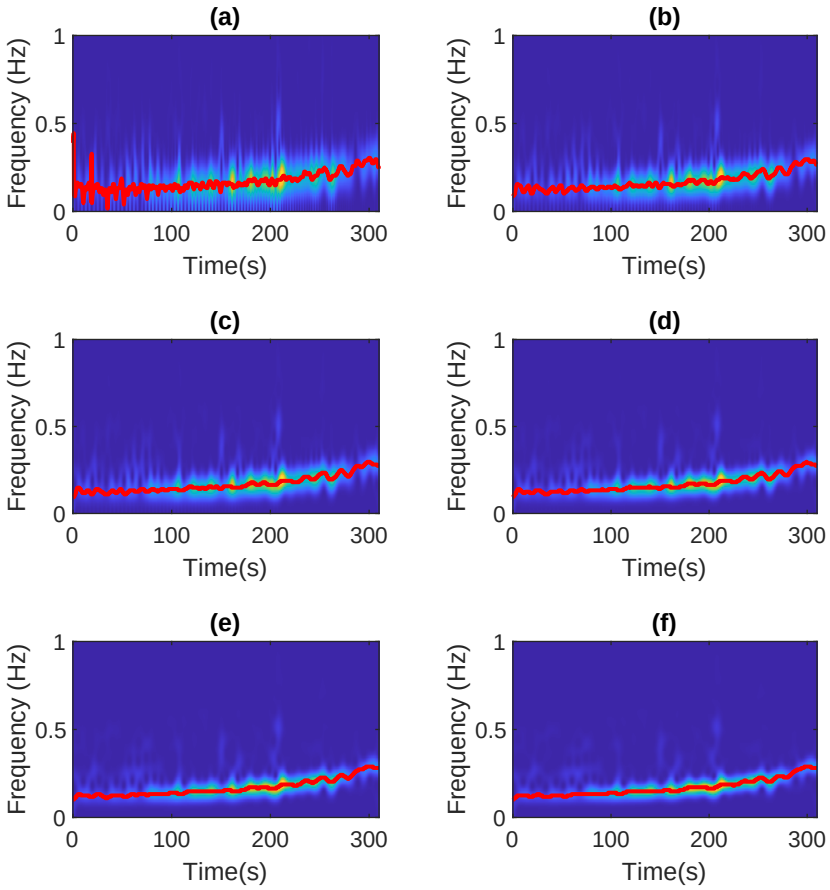

Figure 4.2: An example of the estimated RF using a single window single peak method, visualized on the spectrogram from the recorded HRV signal, using the different window sizes: (a) $M = 60$; (b) $M = 80$; (c) $M = 100$; (d) $M = 120$; (e) $M = 140$; (f) $M = 160$

The result of a multitaper consisting of 4 windows using the method $\texttt{DP}_{10}^{m}$ is outlined in Table 4.3. The results for this type of weighted frequency domain peak detection were not as successful as initial tests on artificial data suggested, and the equal weighting of the frequency peaks in the spectral domain, with only regard to the amplitude of the peaks, is not robust against the noise present in real data.

$\texttt{DP}_{10}^{m}$	**val.-bad data**			**val. data**		
Window size	**MSE**	**Bias**^{2}	**Var**	**MSE**	**Bias**^{2}	**Var**
60	-	-	-	-	-	-
80	-	-	-	-	-	-
100	$1.6e^{-2}$	$1.0e^{-2}$	$5.9e^{-3}$	$1.6e^{-2}$	$9.6e^{-3}$	$6.0e^{-3}$
120	$1.4e^{-2}$	$9.9e^{-3}$	$3.9e^{-3}$	$1.3e^{-2}$	$9.3e^{-3}$	$3.5e^{-3}$
140	$1.2e^{-2}$	$9.2e^{-3}$	$2.7e^{-3}$	$1.1e^{-2}$	$8.7e^{-3}$	$1.1e^{-2}$
160	$1.1e^{-2}$	$8.3e^{-3}$	$2.2e^{-3}$	$1.0e^{-2}$	$7.9e^{-3}$	$2.1e^{-3}$

Table 4.3: The performance of $\texttt{DP}_{10}^{m}$'s RF estimation, on validation data without bad data, and on all validation data.

$\texttt{DP}_{10}^{m}$ clearly underperforms against the two other methods of spectral estimation. NP is the best performing estimation method for window sizes below 140 samples, and a shorter window size is preferable in order to make the frequency estimations more real time. However it is well known that the multitaper is much more resilient to noise. What if an optimization of the multitaper weights can bridge the gap between a single spectral estimate, and a spectral estimate from a multitaper. Let's define the method $\texttt{ONP}^{m}$ as $\texttt{NP}^{m}$ but with an optimization step of the multitaper weights performed on the training data set.

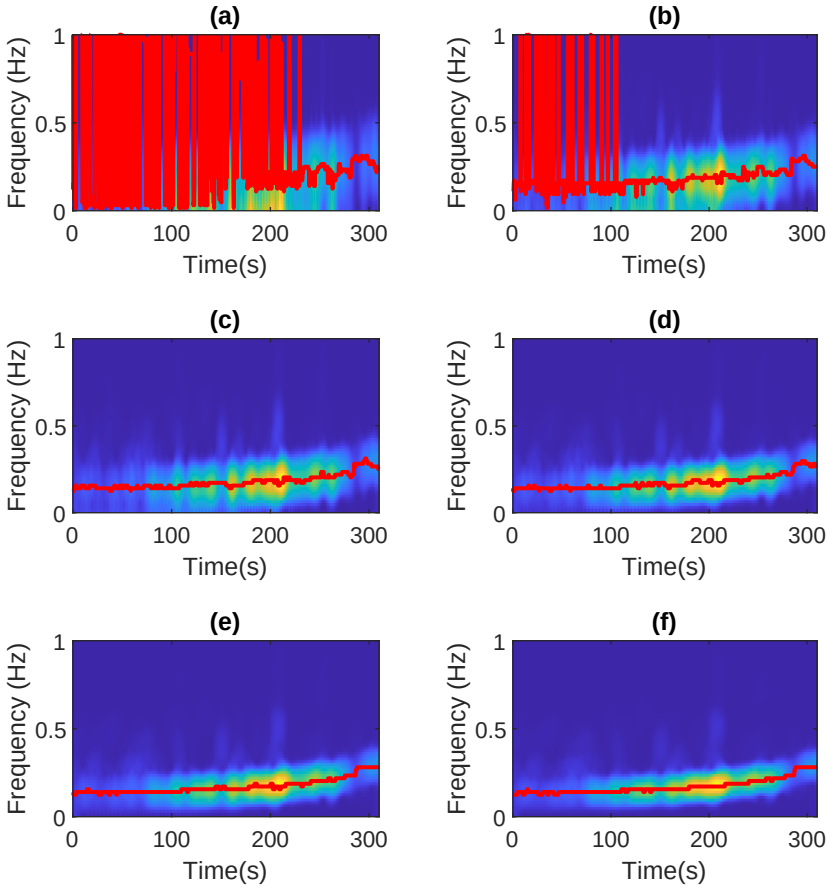

Figure 4.3: An example of the estimated RF visualized on the spectrogram of the recorded HRV signal, using the multitaper single peak method for the recorded HRV. With the different window sizes: (a) $M = 60$; (b) $M = 80$; (c) $M = 100$; (d) $M = 120$; (e) $M = 140$; (f) $M = 160$

4.3 Optimization of multitaper weights

The optimization was done using a bound constraints modification of the Nelder-Mead algorithm, limiting the all the weights λ_k, $k = 1, 2, 3$, to $(-1, 1)$, keeping $\lambda_0 = 1$, and defining the cost as the sum of the mean squared error for all time-series included in the training data, thus giving equal weight to all time-series in the optimization, regardless of length. This procedure was then performed on all the different choices in window length previously selected. The improvement in performance on the validation data that is achieved after optimization of the multitaper weights is significant, and cannot be discarded. The improved RF estimation from the optimization is most significant when the window size is shorter, and becomes negligible with an increase of the window length, as shown in Table 4.4 Comparing the different RF estimation methods against

$\texttt{ONP}^m$	**val.-bad data**			vs. $\texttt{NP}^m$
Window size	**MSE**	**Bias**^{2}	**Var**	**p-value**
60	$1.7e^{-3}$	$7.6e^{-7}$	$1.7e^{-3}$	-
80	$7.3e^{-4}$	$1.1e^{-6}$	$7.3e^{-4}$	-
100	$5.5e^{-4}$	$1.7e^{-6}$	$5.5e^{-4}$	0.0038
120	$4.9e^{-4}$	$1.6e^{-6}$	$4.8e^{-4}$	0.0318
140	$4.3e^{-4}$	$2.3e^{-6}$	$4.3e^{-4}$	0.7196
160	$3.6e^{-4}$	$3.2e^{-6}$	$3.6e^{-4}$	0.6849

Table 4.4: The performance of $\texttt{ONP}^m$ RF estimation, on validation data without bad data, combined with the p-value indicating the probability of obtaining an at least as extreme performance(MSE) difference with the assumption of a null-hypothesis between the two methods.

each other, Table 4.5, the performance gain of the optimized multitaper weights is significant. The same example as before is shown in Figure 4.4, where the lost energy normalization when using optimized weights can be observed. However for this use-case the energy is not important, the only interesting feature is the spectral peaks, and energy normalization would be easily implementable if needed. Finally the movement of the multitaper weights λ_k is visualized in Figure 4.5

	MSE, all val. data			**p-value**	**p-value**
Win. size	NP	NP^m	ONP^m	NP vs. ONP^m	NP^m vs. ONP^m
60	$3.5e^{-3}$	-	$1.7e^{-3}$	$9.0469e^{-8}$	-
80	$1.4e^{-3}$	-	8.0^{-4}	$2.2113e^{-4}$	-
100	$7.9e^{-4}$	$3.4e^{-3}$	$6.4e^{-4}$	6.0176^{-5}	$1.6331e^{-6}$
120	$6.5e^{-4}$	$1.1e^{-3}$	$5.6e^{-4}$	$1.9e^{-3}$	0.0135
140	$6.1e^{-4}$	$4.6e^{-4}$	$4.6e^{-4}$	$1.2e^{-3}$	0.8067
160	$5.7e^{-4}$	$3.8e^{-4}$	$3.8e^{-4}$	$8.3088e^{-4}$	0.8452

Table 4.5: The performance of ONP^m RF estimation, on all validation data, combined with the p-value indicating the probability of obtaining an at least as extreme performance (MSE) difference with the assumption of a null-hypothesis between the two methods.

4.4 Proof of concept with data from Polar H10

In the spirit of exploration a consumer grade heart rate monitor, the Polar H10, was used in order to estimate the RF as in "real-time" as possible. In order to emulate a use case, where "real-time" data is displayed for a user during exercise. Breathing was done in synchronization to a metronome with a frequency of 0.2 Hz, while riding a bicycle outside. The goal here was to let the heart rate be non-stationary, and keeping the RF as stationary as possible. The respiratory frequency estimation on top of the "default" weight multitaper spectrogram, window length 120, for the processed HRV signal is presented in Figure 4.6(**a**). The sample rate was $F_s = 4$ Hz, and an IIR filter with a group delay of 1 sample at 0.2 Hz was used. Inside each window of the multitaper filtering and mean removal took place before the DFT was computed, in order to handle the highly non stationary HRV signal, originating from the heart rate 4.6(**b**).

Data collection was also performed at higher RF during workload on a bike. However the amount of discontinuities in the RF estimation made the need for longer window lengths a necessity, and with that the assumption of zero mean inside each window is violated.

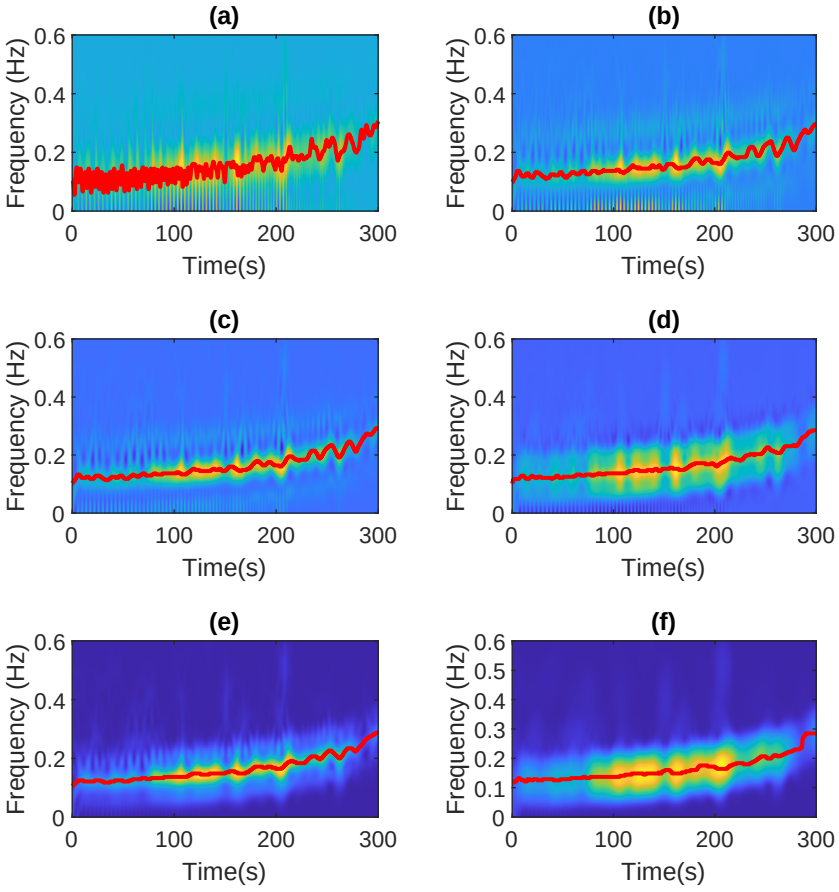

Figure 4.4: An example of the estimated RF visualized on the spectrogram of the HRV signal, using the multi window single peak optimized method. With window sizes: (a) $M = 60$; (b) $M = 80$; (c) $M = 100$; (d) $M = 120$; (e) $M = 140$; (f) $M = 160$

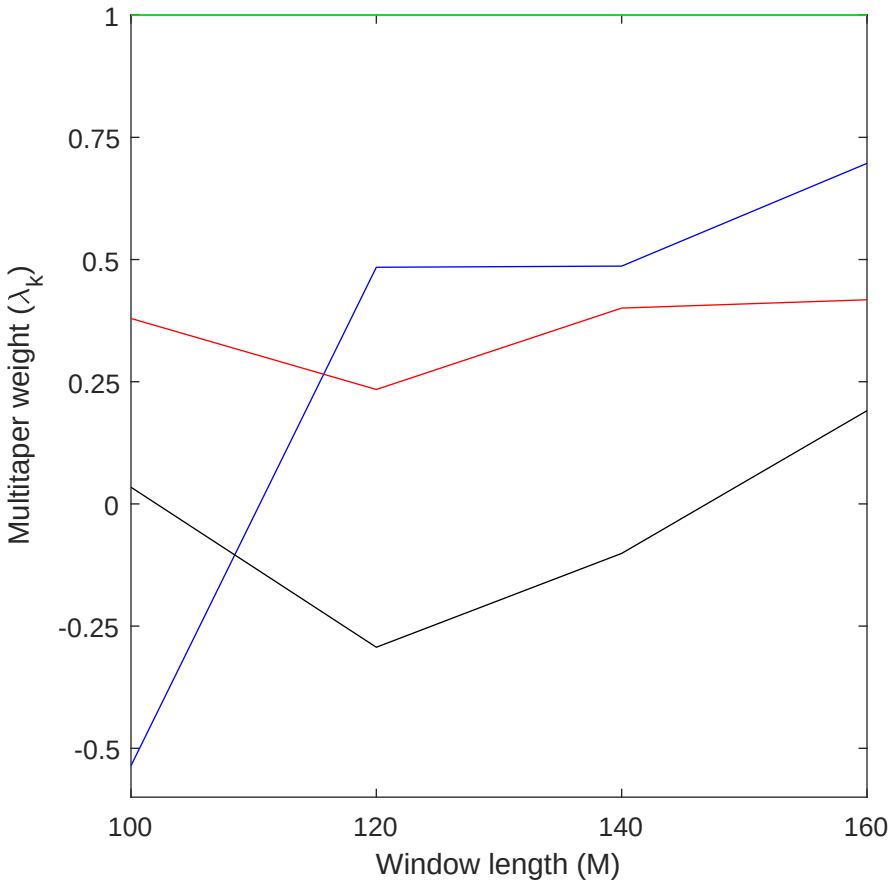

Figure 4.5: The four multitaper weights' movement when the window size increases. Green is fixed at $\lambda_0 = 1.0$, blue is λ_1, red is λ_2, and black is λ_3.

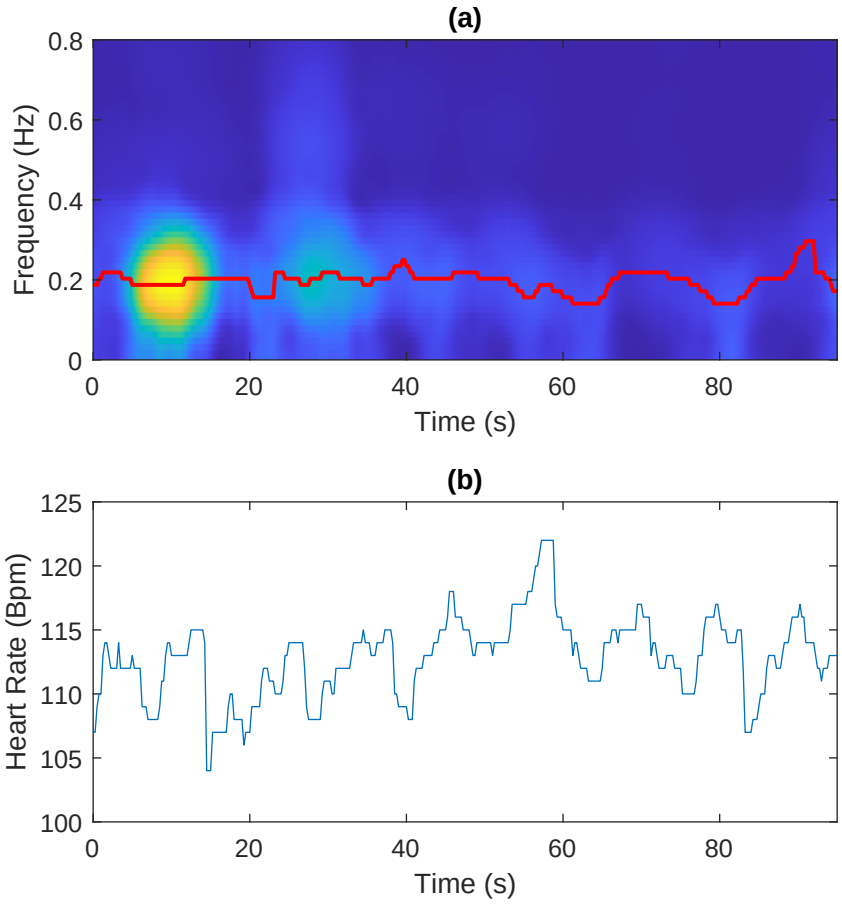

Figure 4.6: In **(a)** the estimated RF, on top of "default" weight multitaper spectrogram of processed HRV signal, the respiratory frequency was 0.2 Hz. In **(b)** the heart rate in beats per minute.

Discussion

The RF can clearly be estimated inside the HF-HRV band, if the heart rate is short time stationary, i.e. if the heart rate is stationary inside each window of the spectrogram, and if that is the case then the main problem is noise, and a multitaper solution is preferred. The inherent problem a multitaper approach exhibits is the addition of artificial low-frequency components to a signal, when the window length causes the multitaper windows' slopes to be non-stationary, with respect to the frequency that is being estimated. Causing artificial components to be added to the signal due to the multiplication of the signal and the window. For this application the window limit is about 100 samples at $F_s = 4$ Hz, with a "sane defaults" choice of multitaper weights. The effect of this problem is diminished if the frequency content searched for in the signal is close to the Nyquist frequency, but applicable here since the RF estimation is only done inside the HF-HRV band. Another way to negate this effect is to make the multitaper weights move, based on the window length, and what type of spectral properties that are needed. The "sane defaults" that were used for the multitaper weights in the simulation, are close to optimal weights when the window length makes the multitaper windows "stationary" in relation to the signal that are being resolved for.

Outside of the laboratory environment the real time aspect of the RF might be more of interest. All methods here are real-time implementable and the main concern is the time-lag in the frequency estimation. The estimation of the frequency will always be half the window size in samples behind the true frequency. Spectrogram reassignment [9] is a way to work around this, using information about the instantaneous frequency to better estimate the "true" frequency. By making the multitaper weights movable, the needed time-delay should decrease even further.

All real data used here had a heart rate of about $60 - 80$ beats per minute (bpm), and was short time stationary in frequency. A way to further decrease the impact of noise on the signal, and thus enabling the use of shorter windows is by increasing the heart rate, and in that way enabling more "sampling" of the RSA inside each window.

The optimization of the multitaper weights was done on the training data for the use of the validation data, relying on the similarity between the two, and the found optimal weights will most likely not be the same for a stationary signal when reducing the window length. Depending on where in the frequency spectrum the signal resides, different window weights are needed in order to enhance that part of the spectrum.

The observant reader realizes that all this time the heart rate was assumed to be short time stationary in frequency during the full duration of the signal. If the heart rate evolves over time, new problems arise, since there is a need for a zero mean in order to perform a spectral decomposition. One way around this is to only assume that the heart rate is stationary during a single window inside the spectrogram/multitaper, and perform the mean removal and filtering inside the solution for the spectral estimation. This process was conducted in Section 4.4, where the minimum heart rate was 104 bpm, and the maximum heart rate was 122 bpm, Figure 4.6(**b**). The obvious downside here is that the short window length needed is competing with the need for longer window length in order to achieve a finer frequency resolution.

The frequency resolution of the Gaussian window of length $M = 100$ is about 0.12 Hz, that is, if there exists two signals of equal strength within 0.12 Hz, $F_s = 4$ Hz, from each other, they are not distinguishable, i.e. it is not possible to perceive if it is a bimodal or unimodal Gaussian distribution. However this ended up not applying here, since the RSA component is the main feature of the HF-HRV band, and thus the RF estimation is not limited by this fundamental limit of resolution. The estimation will however get more smeared out in the spectral domain, making peak detection harder, and more prone to errors.

Lastly about the usage of a high-pass FIR filter. The filter used for all the computations was an order 120 FIR filter, generating a time lag of 60 samples. However the filter is easily replaceable by an Infinite-Impulse Response (IIR) filter. By placing the stop-band at 0.08 Hz, the group delay of a Chebyshev Type 2 filter(inverse Chebyshev filter), at relevant frequencies (> 0.12 Hz) of respiration, is below 5 samples by margin, and therefore not a big contributor to the inherent time-lag problem in the spectral estimation.

Conclusion

6.1 Possibility of RF estimation from HF-HRV

Estimation of the RF inside the traditionally defined HF-HRV band is possible, and depending on the signal quality in combination with the accepted time-delay of the frequency estimation, and the HRV signal, different methods are applicable. The most robust solution is a multitaper spectrogram with optimized weights. Exactly how to optimize the weights for estimation of an unknown RF from HF-HRV is not determined. However, on controlled data optimized weights perform significantly better than other methods of similar window length.

A low time delay solution, about 20s is feasible as long as the heart rate is short time stationary, and the breathing an individual performs is close to stationary in frequency for the length of a window. As mentioned above the RSA is the main component of the HRV-HF band, and therefore the possibility of confusion with other signals inside the frequency band is low.

6.2 Future work

6.2.1 Time-lag in time-frequency estimation

The effect of the time lag in the time-frequency estimation on the real time data used in this work was small. Both due to the relatively slow increase in RF, and the almost linear RF chirp. If this methodology of time-frequency estimation is to be applied on more general HRV data, a spectrogram reassignment using the instantaneous phase of the signal needs to be used, in order to produce more "real-time" RF results.

6.2.2 Multitaper weights

The optimizing of multitaper weights provided positive results for the available data set, even though training was done on "good" data. However, in order to get similar

results for more general HRV signals, the multitaper weights need to move in order to keep the looked-for signal component of the HF-HRV band as prominent as possible. A comparison of the performance between a reassigned multitaper spectrogram with moving weights, and the other TFDs presented in [2], would be an interesting read. Also the scheme for optimization of the multitaper weights could be improved since the current data set was limited. For example by the usage of k-fold cross validation.

6.2.3 RF estimation above the HF-HRV band

The RSA is the prominent spectral component in the HF-HRV band. This band is traditionally defined between $0.12 - 0.4$Hz, however during heavy exercise individuals RF may exceed 0.4Hz. The correlation between the RF and the HRV above the HF-HRV band is not yet established. The heart rate for the individuals in this work was between 60 and 80 beats per minute (bpm). The low amount of beats leads to a maximum resolvable RF frequency of between $0.5 - 0.67$ Hz, due to aliasing. By using data with a higher heart rate, one may determine if a RF component, a component that is not directly connected to the RSA, exists above the HF-HRV band or not.

Appendix

A.1 Generation of chirps

A chirp is a signal with a time dependent frequency component, the frequency can either increase or decrease with respect to time.

Linear chirp

A linear chirp changes its frequency between two values in a linear fashion. However the generation of such a chirp is not as straightforward as one may think, and thus one methodology of generating such a chirp is presented here.
Starting with a sinusoidal signal

$$x(t) = A\sin(\omega t + \phi), \quad t \in \mathbb{R} \tag{A.1}$$

The frequency of the signal needs to be time-varying, i.e. the angular frequency needs to be non constant in time. One way to achieve this is to make the phase non-linear in time. For example by defining

$$\theta(t) = 2\pi\alpha t^2 + 2\pi f_0 t + \phi \tag{A.2}$$

since $\omega = \frac{d}{dt}\theta(t)$ we now have a way of making the instantaneous frequency ω time-varying. Combining A.2 and A.1 together with the fact that $\frac{d}{dt}\theta(t) = \omega(t) = 2\pi(kt+f0)$ we get after integration that

$$\theta(t) = 2\pi(\frac{kt}{2} + f_0)t + \phi \tag{A.3}$$

using this in combination with A.1 yields

$$x(t) = A\sin(2\pi(\frac{kt}{2}t + f_0)t + \phi) \tag{A.4}$$

where $k = \frac{f_1 - f_0}{T}$, with T is the duration of the chirp.

Exponential chirp

The core concept used in the generation of the linear chirp is that the phase of the sinusoidal function $\theta(t)$ needs to be continuous in order to generate a non discontinuous chirp. For an exponential (in time) chirp, the only problem is to derive the correct constants in order to get the chirp to behave in an appropriate manner. For example by letting

$$\omega(t) = f_0 e^{\frac{\log(\frac{f_1}{f_0})}{T}t} \tag{A.5}$$

be the instantaneous frequency, the chirp will be continuous with some reservations, i.e. $\log(\frac{f_1}{f_0}) = 0$), The function of the chirp signal becomes

$$x(t) = A\sin(2\pi\frac{\omega(t) - f_0}{\log(\frac{f_1}{f_0})}T)\ , \qquad 0 \leq t \leq T \tag{A.6}$$

Bibliography

[1] E. Kristal-Boneh, M. Raifel, P. Froom, and J. Ribak, "Heart rate variability in health and disease", Scand. J Work Environ., Health, vol. 21, pp. 86-95, 1995.

[2] N. Ali Khan, P. Jönsson and M. Sandsten, "Performance Comparison of Time-Frequency Distributions for Estimation of Instantaneous Frequency of Heart Rate Variability Signals", Applied Sciences 7, NO. 3: 221, Feb. 2017.

[3] M. Hansson-Sandsten, and P. Jönsson, "Multiple Window Correlation Analysis of HRV Power and Respiratory Frequency", IEEE Transactions on Biomedical Engineering, VOL. 54, NO. 10, Oct. 2007.

[4] J. G. Proakis, D. K. Manolakis. *Digital Signal Processing* (4th Edition). Pearson, Harlow, 2014.

[5] G. Lindgren, H. Rootzén, and M. Sandsten. *Stationary stochastic processes for scientists and engineers*. CRC Press, 2014.

[6] H. Holzmann and S. Vollmer, "A likelihood ratio test for biomodality in two-component mixtures with application to regional income distribution in the EU", AStaA Advances in Statistical Analysis: *A Journal of the German Statistical Society*, VOL.92, NO. 1, pp. 57-69, 2008.

[7] J. A. Nelder and R. Mead, "A Simplex Method for Function Minimization", The Computer Journal, VOL. 7, NO. 4, Jan. 1965.

[8] R. Bailón, G. Laouini, C. Grao, M. Orini, P. Laguna, and O. Meste, "The Integral Pulse Frequency Modulation Model With Time- Varying Threshold: Application to Heart Rate Variability Analysis During Exercise Stress Testing", IEEE Transactions on Biomedical Engineering, VOL. 58, NO. 13, March. 2011.

[9] M. Sandsten. "Time-Frequency Analysis of Time-Varying Signals and Non-Stationary Processes - An introduction", 2020

Master's Theses in Mathematical Sciences 2023:E22
ISSN 1404-6342

LUTFMS-3473-2023

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

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