

*F*-TESTS FOR FREQUENCY MODULATED TIME SERIES  
USING MULTITAPER METHODS

by

BENJAMIN OTT

A thesis submitted to the Graduate Program in  
Mathematics & Statistics  
in conformity with the requirements  
for the Degree of Masters of Science

Queen's University  
Kingston, Ontario, Canada  
September 2023

Copyright © Benjamin Ott, 2023

# Abstract

We propose three new semiparametric multitaper tests for the detection of modulated line components where the modulation is assumed to be created by a polynomial of degree  $P$ . The first test,  $F_4$ , is a modification of the  $F$ -test found in [3] and uses a set of weights to obtain a better in-band concentration for a wider range of tapers. The second test,  $F'_4$ , builds on the aforementioned test by fixing a problem that arises when choosing to evaluate the  $F$ -test with an even number of tapers. We derive the distributions for these two tests along with an examination of the final test based on an aggregation of  $F'_4$  tests, where the aggregation is over different multitaper orders. We derive an upper bound for the distribution of the Aggregate  $F$ -test statistic, and via simulation, we approximate the  $F$ -test's confidence level needed to obtain a pre-specified Type 1 error of the Aggregated test. These new  $F$ -tests are compared to the  $F$ -test in [3], and [30] via simulation comparing the probability of detecting a frequency-modulated signal in different types of noise and modulation widths. We also examine the functionality of these new test statistics through a study of the SoHO GOLF instrument data.

**Keywords:** *frequency modulation, F-test, aggregated test, time series, multitaper*

# Acknowledgements

To ...

**Glen Takahara:** Thanks for bringing me back down to earth and making me put pen to paper to prove the long list of ideas we had. I know I didn't always jump at the chance to prove them, but looking back, I am glad I did it. Also, thanks for taking the time to meet with me whenever I asked; I really appreciate it.

**Wesley Burr:** Thanks for dragging me out of Physics and sticking me into a new field that I have grown to love and be passionate about. I will now admit to you that coding is more of a hobby than a job for me at this point; thanks for proving me wrong.

**My parents:** Without you, none of this would ever have been possible.

And, most of all

**To Hannah**

Firstly, thanks for being my rock; No matter how hard my day was, talking to you at the end of the day always made me remember why I am doing all this work.

Lastly, thanks for sitting there patiently while I tried to explain what I was working on; I know the math never got through, but it always helped me get through the problem!

# Table of Contents

<b>Abstract</b>	<b>i</b>
<b>Acknowledgements</b>	<b>ii</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Background</b>	<b>4</b>
2.1 Time Series . . . . .	4
2.2 Frequency Domain . . . . .	6
2.2.1 Spectral Representation Theorem . . . . .	6
2.2.2 The Power Spectrum . . . . .	7
2.2.3 Direct Spectral Estimators and Multitaper . . . . .	8
2.2.4 First Derivatives of DPSSs and Sine Tapers . . . . .	14
2.3 Frequency Modulation and Test Statistics . . . . .	15
<b>3 Review Of <math>\tilde{F}_3</math> and Proposed <math>F_4</math> and <math>F'_4</math> Test Statistics</b>	<b>17</b>
3.1 Derivation of $\tilde{F}_3$ . . . . .	17
3.2 Issues with $\tilde{F}_3$ in practice . . . . .	24
3.3 Weighted $F_4$ Test Statistic . . . . .	26
3.4 Distribution of $F_4$ Under the Null Hypothesis . . . . .	30
3.4.1 Eigencoefficient Distribution . . . . .	31
3.4.2 Real and Imaginary Parts Of Eigencoefficients . . . . .	32

3.4.3	Standard Inverse and its Derivative . . . . .	36
3.4.4	Instantaneous Frequency . . . . .	38
3.4.4.1	Distribution of Denominator of $\psi(t; f)$ . . . . .	38
3.4.4.2	Distribution of Numerator of $\psi(t; f)$ . . . . .	40
3.4.4.3	Distribution of $\psi(t; f)$ . . . . .	41
3.5	$F'_4$ Test Statistic . . . . .	43
3.5.1	Isolation of $K^{\text{th}}$ Taper . . . . .	45
3.5.1.1	Eigencoefficients . . . . .	45
3.5.1.2	Standard Inverse . . . . .	47
3.5.1.3	Instantaneous Frequency . . . . .	49
3.5.1.4	Eigencoefficients of the Instantaneous Frequency . . . . .	51
3.5.1.5	$\hat{c}_{P/0}$ and $H$ . . . . .	52
3.5.1.6	Norm of $\Psi(f)$ . . . . .	53
3.5.1.7	Decomposing the $F_4$ Test Statistic . . . . .	53
3.5.2	Distribution of $F'_4(f, K)$ . . . . .	55
<b>4</b>	<b>Aggregate Test</b>	<b>58</b>
4.1	Distribution Discussion . . . . .	59
<b>5</b>	<b>Simulation and Comparison</b>	<b>63</b>
5.1	Test Parameters . . . . .	64
5.2	Power of Tests Under White Background Noise . . . . .	65
5.3	Power of Tests With Coloured Noise . . . . .	70
5.4	Aggregate Test Estimate . . . . .	77
5.5	Conclusion . . . . .	79
<b>6</b>	<b>SoHO Time Series Analysis</b>	<b>82</b>
6.1	Pre-Processing . . . . .	83

6.2	Analysis . . . . .	85
<b>7</b>	<b>Conclusion and Future Work</b>	<b>90</b>
7.1	Conclusion . . . . .	90
7.2	Future Work . . . . .	91
	<b>Bibliography</b>	<b>94</b>

# List of Figures

2.1	Comparison of the Periodogram (Blue) vs the Direct Spectral Estimator (Red). These estimates were applied to an AR(4) Process with the true SDF shown in black [19]. . . . .	10
3.1	Real part of the standard inverse projection filter, $Z$ shown in red, applied to a linear polynomial shown in black . . . . .	19
3.2	Real part of the standard inverse projection filter, $Z$ shown in red, applied to a quadratic polynomial shown in black. . . . .	20
3.3	$\tilde{F}_3$ applied to a frequency modulated series (carrier frequency $f = 0.1$ ) with input parameter $K = 79$ . Dotted red lines represent $f \pm \frac{(K+1)}{2N}$ , and the horizontal red line is the quantile of an $F_{1,79-1}$ corresponding to a significance level of $\frac{1}{2}$ . This examples uses $N = 6000$ and $\phi(\tau) = \frac{0.002}{\frac{N}{2}} (t - \frac{N}{2})$ , and $Z_t$ is Gaussian white noise with a low signal to noise ratio . . . . .	25
3.4	$\tilde{F}_3$ applied to FM series (carrier frequency $f = 0.1$ ) with input parameter $K = 79$ . Dotted red lines represent $f \pm \frac{(K+1)}{2N}$ , and the horizontal red line is the quantile of an $F_{1,79-1}$ corresponding to a significance level of $\frac{1}{2}$ . Where $N = 2000$ and $\phi(\tau) = \frac{0.001}{\frac{N}{2}} (t - \frac{N}{2})$ , and $Z_t$ is Gaussian white noise with a signal to noise ratio of 1 . . . . .	26

3.5  $\tilde{F}_3$  applied to FM series (carrier frequency  $f = 0.1$ ). The left-hand plot is  $K = 8$  and the right-hand plot is  $K = 9$ . The horizontal red line is the quantile of an  $F_{1,K-1}$  corresponding to a significance level of  $\frac{1}{2}$ . In this example,  $N = 6000$  and  $\phi(\tau) = \frac{0.001}{\frac{N}{2}} (t - \frac{N}{2})$ , and  $Z_t$  is Gaussian white noise with a signal to noise ratio of 1. . . . . 27

3.6 Selected spectral windows for  $K = 80$  for both Sinusoidal (in red) and DPSS tapers (in black). The left-hand plot is the  $k = 0$  and the right-hand plot is the  $k = 9$  spectral window. The red dotted line is the width of the window given by  $W = \frac{(K-1)}{2N}$ . . . . . 28

3.7 Spectral windows for  $K = 80$  for both Sinusoidal (in red) and DPSS tapers (in black). The left-hand plot is the  $k = 69$  and the right-hand plot is the  $k = 79$  spectral window. The red dotted line is the width of the window given by  $W = \frac{(K-1)}{2N}$ . . . . . 29

3.8 Weighting scheme generated for  $K = 80$ , with a penalty value of 0.10. 31

3.9 Theoretical density of the  $F_{1,K-P}$  distribution in red overlayed on the simulated density of  $F_4$  under the null hypothesis, when the process was Gaussian white noise. The number of simulations was 10,000. . . 42

3.10  $F_4$  applied to FM series (carrier frequency  $f = 0.1$ ) with input parameter  $K = 79$ . Dotted red lines represent  $f \pm \frac{(K+1)}{2N}$ , and the horizontal red line is the quantile of an  $F_{1,79-1}$  corresponding to a significance level of  $\frac{1}{N}$  with  $N = 6000$ . Also,  $\phi(\tau) = \frac{0.002}{\frac{N}{2}} (t - \frac{N}{2})$ , and  $\{Z_t\}$  is Gaussian white noise with a low signal-to-noise ratio. . . . . 43

3.11  $\tilde{F}_3$  in red applied to FM series under background SNR = 0.8 Gaussian white noise with input parameter  $K = 20$  across all frequencies.  $F_4$  in black applied to the same FM series with the same input parameter. The horizontal red line is the quantile of an  $F_{1,19}$  corresponding to a significance level of  $\frac{1}{N}$  where  $N = 6000$ . . . . . 44

3.12	Probability of detection over 2000 tests. The $\tilde{F}_3$ in red is applied to the FM series under background SNR = 0.5 Gaussian white noise with input parameter $K \in \{5, \dots, 80\}$ at carrier frequency. The $F_4$ in green is applied to the same FM series with the same input parameters. The 2000 tests were all conducted at a confidence level of $\alpha = \frac{1}{N}$ . . . . .	45
3.13	$\tilde{F}_3$ in red applied to FM series under background SNR = 0.5 Gaussian white noise with input parameters $K \in \{5, \dots, 80\}$ at the carrier frequency. $F_4$ in green and $F'_4$ in blue applied to the same FM series with the same input parameters. The 2000 tests were all conducted with $\alpha = \frac{1}{N}$ . . . . .	57
4.1	Estimated correlation between 70 pairs of $(T_a)_i$ ranging from $K = 10$ to $K = 80$ . Twenty thousand simulations were conducted with unique Gaussian white noise as the time series. . . . .	60
4.2	Estimated complementary CDF under the null hypothesis for the Aggregate test statistic with $\mathcal{K} = \{5, \dots, 80\}$ and $\beta = \frac{1}{N}$ , $N = 2000$ . 10,000 simulations were conducted with unique Gaussian white noise as the time series for each simulation. . . . .	61
5.1	Spectral multitaper estimate for $K = 9$ under $m_P = 0.01$ linear modulation and $N = 2000$ . . . . .	66
5.2	Probability of detection conducted through 10,000 simulations with modulation width $m_P = 0.0005$ and $N = 2000$ . The pink curve, Harm-F, is the classic Thomson [30] Harmonic $F$ -Test Statistic, which has little modulation detection capacity. . . . .	67
5.3	Probability of detection conducted through 10,000 simulations with modulation width $m_P = 0.001$ and $N = 2000$ . . . . .	68

5.4	Probability of detection conducted through 10000 simulations with modulation width $m_P = 0.002$ and $N = 2000$ . This is the first case where the $F_4$ and $F'_4$ tests perform at acceptable levels, while the Harmonic $F$ -test is essentially unusable. Note the extremely strong performance of the Aggregate test removes the subjectivity of the $F_4$ and $F'_4$ tests which vary with $K$ . . . . .	69
5.5	Probability of detection conducted through 10000 simulations with modulation width $m_P = 0.003$ and $N = 2000$ . Essentially the same performance and conclusion as suggested by Figure 5.4. . . . .	70
5.6	Probability of detection conducted through 10,000 simulations with modulation width $m_P = 0.01$ and $N = 2000$ . . . . .	71
5.7	Probability of detection conducted through 10,000 simulations with modulation width $m_P = 0.01$ , $N = 2000$ and using a penalty in the weighting of 0.05. . . . .	71
5.8	Probability of detection, $m_P = 0.001$ and $N = 2000$ . Background noise was generated with an AR = (0.4, -0.2, 0.6), MA = (0.2, -0.5) with innovations from a zero mean Gumbel distribution. . . . .	73
5.9	Probability of detection, $m_P = 0.002$ and $N = 2000$ . Background noise was generated with an AR = (0.4, -0.2, 0.6), MA = (0.2, -0.5) with innovations from a zero mean Gumbel distribution. . . . .	74
5.10	Spectral multitaper estimate for $K = 20$ and $N = 2000$ under $m_P = 0.002$ linear modulation with AR = (0.4, -0.2, 0.6), MA = (0.2, -0.5) with innovations from a zero mean Gumbel distribution. . . . .	75
5.11	Realization of test statistic $F'_4$ for $K = 20$ and $N = 2000$ . The test is applied to the same data as the spectrum provided above in Figure 5.10, with the same parameters. . . . .	75

5.12	Probability of detection, $m_P = 0.003$ and $N = 2000$ . Background noise was generated with an AR = (0.4, -0.2, 0.6), MA = (0.2, -0.5) with innovations from a zero mean Gumbel distribution. . . . .	76
5.13	Spectral multitaper estimate for $K = 20$ under $m_P = 0.003$ linear modulation with AR = (0.5, 0.3, -0.1), MA = (0.6) with innovations from a zero mean Gumbel distribution. . . . .	76
5.14	Probability of detection, $m_P = 0.003$ and $N = 2000$ . Background noise was generated with an AR = (0.5, 0.3, -0.1), MA = (0.6) with innovations from a zero mean Gumbel distribution. . . . .	77
5.15	Probability of detection with modulation width $m_P = 0.005$ and $N = 2000$ . Background noise was generated with an AR = (0.5, 0.3, -0.1), MA = (0.6) with innovations from a zero mean Gumbel distribution. .	78
5.16	Probability of detection with modulation width $m_P = 0.01$ and $N = 2000$ . Background noise was generated with an AR = (0.5, 0.3, -0.1), MA = (0.6) with innovations from a zero mean Gumbel distribution. .	78
5.17	Probability of detection with modulation width $m_P = 0.01$ and $N = 2000$ . Background noise from AR = (0.5, 0.3, -0.1), MA = (0.6) with innovations from a zero mean Gumbel distribution. . . . .	79
5.18	Probability of detection width $m_P = 0.01$ and $N = 2000$ . Background noise was generated with an AR = (0.5, 0.3, -0.1), MA = (0.6) with innovations from a zero mean Gumbel distribution. . . . .	80
6.1	The left figure shows what the SoHo satellite would look like after being unfolded [17]. The right figure shows the location of the $L_1$ Lagrange point with respect to the Earth and Sun [18]. . . . .	83
6.2	Multitaper spectral estimate of the original data for $K = 7$ , $NW = 4$ , and $\text{pad} = 2^{\lfloor \log_2(10N) \rfloor}$ . . . . .	83

6.3	Multitaper spectral estimate of the processed data for $K = 7$ , $NW = 4$ , and $\text{pad} = 2^{\lfloor \log_2(10N) \rfloor}$ . . . . .	84
6.4	$\tilde{F}_3$ result conducted with $\alpha = \frac{1}{N}$ , $K = 2NW - 2$ , and $\text{pad} = 2^{\lfloor \log_2(10N) \rfloor}$ . . . . .	85
6.5	Overlay of $F'_4$ in green compared to the results with $\tilde{F}_3$ shown in purple. Both tests were conducted at a significance level of $\alpha = \frac{1}{N}$ . . . . .	86
6.6	$\tilde{F}_3$ result conducted with $\alpha = \frac{1}{N}$ , $K = 2NW - 2$ , $\text{pad} = 2^{\lfloor \log_2(10N) \rfloor}$ for $K = 5$ to $K = 80$ . . . . .	87
6.7	$F'_4$ result conducted with $\alpha = \frac{1}{N}$ , $\text{pad} = 2^{\lfloor \log_2(10N) \rfloor}$ for $K = 5$ to $K = 80$ . . . . .	87
6.8	Aggregate test result conducted with $\text{pad} = 2^{\lfloor \log_2(10N) \rfloor}$ for $K = 5$ to $K = 80$ . . . . .	88
7.1	Instantaneous frequency at carrier frequency overlaid with linear model estimated slope for a linear modulation bandwidth of 0.001 . . . . .	92
7.2	Instantaneous frequency at carrier frequency overlaid with linear model estimated slope for a linear modulation bandwidth of 0.005 . . . . .	92

## List of Tables

5.1	Simulation input parameters for white background noise. Each simu- lation is conducted with 10,000 iterations. . . . .	65
5.2	Simulation input parameters for coloured background noise. Each simu- lation is conducted with 10,000 iterations. . . . .	72

# 1. Introduction

A time series: a set of observations in which every measurement has an exact place in an itemized list. Organized in such a way that when one puts the entire series together, we have the power to discover and prove fundamental concepts about the universe. Unfortunately, the cobblestone path laid out over decades of time series research is not as straightforward and smooth as some may hope. Although traditional time series techniques are available in books like [6] and [5], the idea of translating the time series into the frequency domain is an intriguing one. By conducting spectral estimation of time series in the frequency domain, we gain the unique ability to detect periodic components that are fundamental to the time series in question.

One of the most well-known spectral estimators was coined by Sir Arthur Schuster called the Periodogram [26]. Unfortunately, this spectral estimator has been shown to be badly biased [8]. In Dr. Dave Thompson's 1982 paper, [30], not only did he propose a new spectral estimator, but he also proposed an  $F$ -test to conduct hypothesis testing on specific frequencies found with the multitaper spectral estimator. This Harmonic  $F$ -test allows statisticians to collaborate with experts in the field of study to discover what inherent properties created these periodicities in the recorded time series.

With all that said, not every time series are created equal. Some may have different signal-to-noise (SNR) ratios which can cause smaller peaks in the spectral estimate, but the Harmonic  $F$ -test is quite robust to these types of scenarios. One

area where the Harmonic  $F$ -test tends to suffer is when the time series has been frequency modulated. Frequency modulated time series occur in many fields, such as Solar Physics through [11], [32], oceanography [13], and geophysics [12]. But what is frequency modulation? In the frequency domain, any time series can be represented by a combination of sinusoids combined with background noise where each sinusoid has a specific period,  $T = \frac{1}{f}$  where  $f$  is the carrier frequency for that specific sinusoid. In non-frequency modulated time series,  $f$  does not change for any time  $t$ . When the time series is frequency modulated,  $f$  varies by time  $t$ ; this causes the signal to change period as time increases. If this change in period is large enough, the Harmonic  $F$ -test will not detect the carrier frequency causing an increase in Type 2 errors.

In [33], Thompson derived an alternative test statistic based on work in [20] and [7] to allow for the detection of frequency modulation in time series by combining the idea of multitaper spectral estimation and isolating the frequency modulation through a standard inverse projection filter. This filter captures the frequency domain information contained in the time series at a specific time increment  $t$ . If one models the time derivative of this filtered series by a polynomial of a certain degree, a test statistic can then be created. The test statistic will be large for frequencies where the model fits well, i.e. there is polynomial frequency modulation present, and the test statistic will be small for frequencies where there is no frequency modulation, as the model will not fit well.

Later, Blanchette. et al modified this test statistic to increase its performance under both white and coloured background noise. They compared their new test statistic,  $\tilde{F}_3$ , under simulation, in addition to applying  $\tilde{F}_3$  to SoHO GOLF (Solar & Heliospheric Observatory mission Global Oscillations at Low Frequencies) data [3] [10].

The things that I have contributed to this area of research included in this thesis are:

- Created the  $F_4$   $F$ -test to allow for a wider functioning range of tapers
- Created the  $F'_4$   $F$ -test which fixes low probability of detection when using an even selection of tapers in  $\tilde{F}_3$  as well as  $F_4$
- Created the Aggregate test, which allows for a range of tapers to be chosen, along with an upper bound for the distribution.
- A comparison between  $\tilde{F}_3$ ,  $F_4$ ,  $F'_4$ , and the Aggregate test is presented
- A study of the SoHO GOLF satellite data is also conducted with a comparison between the aforementioned test statistics.

This thesis is organized as follows. In Chapter 2, we review the necessary background needed to derive a frequency modulated test statistic, including an overview of the multitaper method as well as a discussion of different possible types of tapers used in multitaper spectral estimation. Section 3.1, we review the test statistic from [3], which is denoted by  $\tilde{F}_3$ . In Section 3.3 through Section 3.5, we propose and discuss our modification of the  $\tilde{F}_3$  test statistic, denoted by  $F_4$  and  $F'_4$ , and show that their asymptotic distributions. Chapter 4 proposes our aggregate test statistic, denoted by  $T_a$ , which aggregates the  $F'_4$  test statistic over multitaper orders. In Chapter 5, we compare, via simulation, the performance of the  $T_a$ ,  $F'_4$ ,  $F_4$  and  $\tilde{F}_3$  test statistics through their ability to detect a frequency modulated signal at a fixed significance level of  $1/N$ , where  $N$  is the length of the time series. In Chapter 6, we apply the new test statistics to the SoHO GOLF data and compare results to the ones found in [3]. This is followed in Chapter 7 with a conclusion and a discussion of future work to improve our created test statistics.

## 2. Background

Although most of this thesis is concerned with objects in the frequency domain, the process of study is, at its essence, a problem in the time domain. Perhaps an obvious first step is to then examine the definition of a time series and talk about some of its properties. A time series can be created by taking a number of measurements, preferably at equidistant intervals, over a period of time. As these measurements are indexed (predominantly by time), they follow a strict order. If a person were to shuffle the series, it would not have the same meaning anymore; in the same way, saying 4 pm comes before 2 pm is inaccurate. As the reader may need to become more familiar with the field of time series, we will start with a discussion of this type of object.

### 2.1 Time Series

A time series is a stochastic process indexed by time, i.e., a family of random variables  $\{X_t, t \in \mathcal{T}\}$ , where  $\mathcal{T}$  is an index set in which the indices correspond to ordered times. We use lowercase,  $\{x_t, t \in T\}$ , to denote a realization of  $\{X_t, t \in T\}$  where  $T \subseteq \mathcal{T}$ . If  $T$  is discrete, then the time increment between any two adjacent times is usually taken to be constant and denoted by  $\Delta t$ , which can be any value in  $\mathbb{R}^+$  not including zero. Without loss of generality, the value  $t$  represents a time index rather than an actual time value. It will be assumed that  $\Delta t = 1$  for the remainder of this

thesis, which simplifies notation. Thus, a time series of length  $N$  will be of the form  $\{X_t, t = 0, \dots, N - 1\}$  or  $\{x_t, t = 0, \dots, N - 1\}$ , and denoted simply by  $\{X_t\}$  or  $\{x_t\}$  if the length of the series  $N$  is understood.

## Weak Stationarity and Covariance

The *covariance function* of a time series  $\{X_t\}$  is denoted by  $\gamma_X(r, s)$ , where  $\gamma_X(r, s)$  is defined as the covariance between  $X_r$  and  $X_s$ :

$$\gamma_X(r, s) := E[(X_r - E[X_r])(X_s - E[X_s])] = \text{Cov}(X_r, X_s) \text{ for } r, s \in \mathbb{Z} \quad (2.1)$$

A time series is defined to be weakly stationary if [6]:

$$E|X_t|^2 \text{ is finite } \forall t \in \mathbb{Z} \quad (2.2)$$

$$E[X_t] = \mu \quad \forall t \in \mathbb{Z} \quad (2.3)$$

$$\gamma_X(r, s) = \gamma_X(r + t, s + t) \quad \forall r, s, t \in \mathbb{Z}. \quad (2.4)$$

Under the assumption of weak stationarity, the *covariance function* (CVF) is given by

$$\gamma_X(\tau) := \gamma_X(t + \tau, t) = \text{Cov}(X_{t+\tau}, X_t) = \text{Cov}(X_\tau, X_0) \quad \forall \tau, t \in \mathbb{Z}. \quad (2.5)$$

Under stationarity, the function  $\gamma_X(\tau)$  is called the *autocovariance function* (ACVF). Here  $\tau$  represents the **lag**, an integer representing the difference between time points. When assuming that the process  $\{X_t\}$  is weakly stationary, one only has to look at the covariance between  $X_0$  and  $X_\tau$ , Eqn. (2.4) states that the covariance between  $X_{r+t}$  and  $X_{s+t}$  is the same for any choice of  $t$ .

## 2.2 Frequency Domain

The time series,  $\{X_t\}$ , has a frequency domain representation, which allows us to isolate periodic structure and examine which specific frequencies contribute meaningfully to the overall time series. This alternate representation becomes particularly important when building a statistical test to determine intrinsic information about an object. As an example, take the Earth revolving around the Sun. It is commonly known that the Earth completes a revolution around the sun approximately every 365.25 days. This can be converted into a frequency of  $f_e = \frac{1}{365.25\text{d}\cdot 24\text{hr}\cdot 60\text{min}\cdot 60\text{s}} \approx 3.17 \times 10^{-8}\text{Hz}$ . By conducting a statistical test on a hypothetical time series representing, say, the position of the earth relative to the sun that covered a time span of several years, we would expect to obtain a significant result from our test at the frequency  $f_e$ . Without prior knowledge about the number of days to complete a full revolution, this test result would provide strong formal statistical evidence of an important scientific reality that we all use. This is a toy example, but the ability to conduct these tests allows the scientific community to gain knowledge of the fundamental frequencies contained in their time series that may have been missed without analysis conducted in the frequency domain. This is known as harmonic analysis and has a rich history dating back to Stokes and Rayleigh in the late 19th century [28], [24].

### 2.2.1 Spectral Representation Theorem

One version of the spectral representation theorem states the following. Let  $\{X_t\}$  be a discrete parameter, zero mean, weakly stationary process that is real-valued. Then there exists a (complex-valued) orthogonal increments process,  $\{Z(f)\}$  defined on  $[-\frac{1}{2}, \frac{1}{2}]$ , such that, for all  $t \in \mathbb{Z}$ ,  $X_t$  can be expressed in the form:

$$X_t = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{i2\pi ft} dZ(f), \quad (2.6)$$

where  $\{Z(f)\}$  has three properties for  $f, f' \in [-\frac{1}{2}, \frac{1}{2}]$  (see [19]):

$$E[dZ(f)] = 0 \quad \forall f; \quad (2.7)$$

$$dS^{(I)}(f) := E[|dZ(f)|^2] = S(f)df; \quad (2.8)$$

$$\text{Cov}(dZ(f), dZ(f')) = E[dZ(f)dZ^*(f')] = 0 \text{ for any two frequencies } f \neq f', \quad (2.9)$$

where  $S^{(I)}(f)$  is a bounded and nondecreasing function called the **Integrated Spectrum** and  $S(f)$  is called the **Power Spectrum**, or simply the **spectrum**.

With the spectral representation theorem, we can define the spectrum in terms of the ACVF.

## 2.2.2 The Power Spectrum

Using the spectral representation theorem (2.6) and multiplying by its complex conjugate, we can obtain a relation between the power spectrum  $S(f)$  and  $\gamma_X(\tau)$ :

$$X_{t+\tau}X_t^* = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{i2\pi f(t+\tau)} dZ(f) \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-i2\pi f't} dZ^*(f').$$

Taking expectations on each side,

$$E[X_{t+\tau}X_t^*] = \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{i2\pi t(f-f')} e^{i2\pi \tau f} E[dZ(f)dZ^*(f')].$$

Using Eqn. (2.9), the integral will be zero for all  $f \neq f'$ , so:

$$E[X_{t+\tau}X_t^*] = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{i2\pi \tau f} E[dZ(f)dZ^*(f)] \quad (2.10)$$

$$= \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{i2\pi \tau f} E[|dZ(f)|^2]. \quad (2.11)$$

Substituting Eqn. (2.8) and Eqn. (2.5) into Eqn. (2.11), and noting that  $S(f)$  is the **power spectrum**, the second central moment of the process  $\{dZ(f)\}$ , we obtain a relation between the spectrum and the ACVF:

$$\gamma_X(\tau) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{i2\pi \tau f} S(f)df. \quad (2.12)$$

The power spectrum can then be written as the Fourier transform of the ACVF

provided that  $\{\gamma_X(\tau)\}$  is square summable<sup>1</sup>,

$$S(f) = \sum_{\tau=-\infty}^{\infty} \gamma_X(\tau) e^{-i2\pi\tau f} \quad \forall f \in \left[-\frac{1}{2}, \frac{1}{2}\right]. \quad (2.13)$$

Therefore, we now have a way of relating the CVF in the time domain to  $S(f)$  in the frequency domain: they are Fourier transform pairs. Realistically, in practice, one doesn't know the true CVF. Thus, to use this representation as a way to estimate the spectrum, we would want to use an estimator for the CVF that is as accurate as possible. This turns out to be problematic.

## 2.2.3 Direct Spectral Estimators and Multitaper

### Periodogram

The periodogram is the classic estimator of the second central moment of  $dZ$  that should be mentioned in the progression to the multitaper estimator. The periodogram is constructed as the Fourier transform of the sample CVF. Using Eqn. (2.1), the sample CVF at lag  $\tau$  is

$$\hat{\gamma}_X(\tau) := \frac{1}{N} \sum_{t=0}^{N-|\tau|-1} (X_{t+|\tau|} - \bar{X})(X_t - \bar{X}), \quad (2.14)$$

where  $\bar{X}$  is the sample mean of  $\{X_t\}$ . It should also be noted that  $\hat{\gamma}_X(\tau)$  can also be defined with the multiplicative factor of  $\frac{1}{N-|\tau|}$ . This alternative definition is usually referred to as the 'unbiased' estimator (under the condition of knowing  $\mu$ ). Eqn. (2.14) is the 'biased' estimator of  $\gamma_X(\tau)$  and is more commonly used in practice as it has a lower bias when the true mean is unknown [19]. It also has the advantage that it gives an estimator of the covariance matrix of  $\{X_t\}$  that is guaranteed to be nonnegative definite. Under the assumption of zero mean and using Eqn. (2.13) and Eqn. (2.14),

---

<sup>1</sup>The spectrum  $S(f)$  can be defined for all  $f \in \mathbb{R}$ . However, from Eqn. (2.13), we note that  $S(f)$  is periodic with period 1 and  $S(-f) = S(f)$  (since  $\gamma_X(-\tau) = \gamma_X(\tau)$ ).

we can obtain the periodogram estimate of the spectrum:

$$\hat{S}^{(P)}(f) := \sum_{\tau=-(N-1)}^{N-1} \hat{\gamma}_X(\tau) e^{-i2\pi\tau f}. \quad (2.15)$$

Note that the index is truncated such that  $\hat{\gamma}_X(\tau) = 0$  for  $|\tau| \geq N - 1$  (compare to Eqn. (2.13)). An alternative equivalent definition of the periodogram under the assumption of having a zero mean time series is [19]:

$$\hat{S}^{(P)}(f) = \frac{1}{N} \left| \sum_{n=0}^{N-1} X_n e^{-i2\pi n f} \right|^2. \quad (2.16)$$

Regardless of the representation, the periodogram can be shown to be an asymptotically unbiased estimator of the spectrum. However, as we will discuss in the following, the asymptotics are extremely slow, so much so that in practice, the periodogram is a badly biased estimator.

## Tapering

Although the periodogram is still used in practice to this day, it has issues with bias. This bias is demonstrated in Fig. 2.1 (blue curve). By multiplying a weight  $h_t$  to each  $X_t$  in the original time series (and hence, to the transforms), the bias can be reduced, making the spectral estimate closer to the true spectrum of the series. These weights are referred to as *tapers*.

$$\hat{S}^{(D)}(f) = \left| \sum_{t=0}^{N-1} h_t X_t e^{-i2\pi t f} \right|^2, \quad (2.17)$$

where  $\{h_t\}$  is the taper. Any spectrum estimator written in this form is called a *Direct Spectral Estimator* (DSE). If one lets  $h_t = \sqrt{\frac{1}{N}} \forall t$ , we see that the periodogram is a particular case of a DSE. The connection between tapering and bias is that the mean of a direct spectral estimator is the convolution between the true spectrum and the discrete Fourier transform of the taper. Specifically, it can be shown that the mean

of  $\hat{S}^{(D)}(f)$  can be written in the form (Eqn. 186(e) in [19]):

$$E[\hat{S}^{(D)}(f)] = \int_{-1/2}^{1/2} \mathcal{H}(f - f')S(f')df' \quad (2.18)$$

where  $S(f)$  is the true spectrum and  $\mathcal{H}(f)$  is the discrete Fourier transform of the taper, also called the *spectral window* of the taper:

$$\mathcal{H}(f) = \sum_{n=0}^{N-1} h_n e^{-i2\pi n f}$$

Although any arbitrary set of weights could be selected for  $h_t$ , perhaps the most famous set of tapers is the Discrete Prolate Spheroidal Sequences (DPSSs) [27], also called Slepian sequences after their discoverer, David Slepian, and denoted by  $\{v_t^{(k)}\}$  for the  $k$ th sequence. The main advantage of the Slepian sequences is that they are

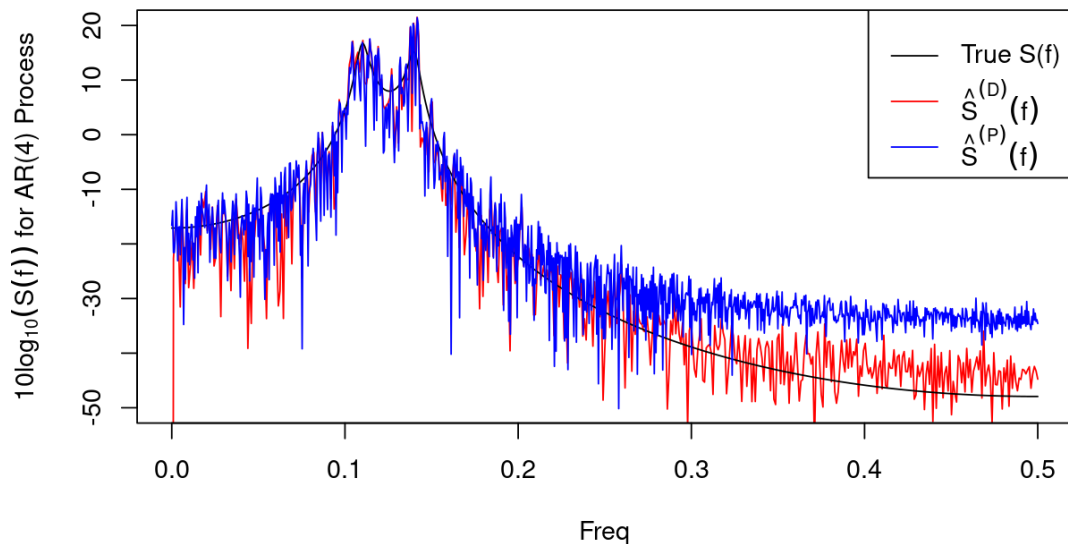


Figure 2.1: Comparison of the Periodogram (Blue) vs the Direct Spectral Estimator (Red). These estimates were applied to an AR(4) Process with the true SDF shown in black [19].

the (successive) solutions to the maximization of:

$$\lambda_k = \frac{\int_{-W}^W |V_k(f)|^2 df}{\int_{-\frac{1}{2}}^{\frac{1}{2}} |V_k(f)|^2 df} \quad (2.19)$$

which is the concentration equation that gives the proportion of energy for the  $k$ th sequence contained in the frequency band  $(-W, W)$  [27], for  $0 < W < \frac{1}{2}$  and  $1 \leq k \leq N$ .

In this,  $V_k(f)$  is the spectral window of the  $k$ th Slepian sequence:

$$V_k(f) := \sum_{n=0}^{N-1} v_n^{(k)} e^{-i2\pi n f} \quad (2.20)$$

The  $k$ th sequence is obtained by maximizing Eqn. (2.19) subject to the constraint that the  $k$ th sequence be orthogonal to the first  $k-1$  sequences. Note that the solution  $\{v_t^{(k)}\}$  to Eqns.(2.19) and (2.20) depends on  $N$  and  $W$ . We sometimes use the notation  $\{v_t^{(k)}(N, W)\}$  to denote the  $k$ th Slepian sequence to emphasize its dependence on the time series length  $N$  and the bandwidth  $W$ . By maximizing Eqn. (2.19), we ensure the maximum energy concentration is contained in a band of  $(-W, W)$  [27]. Referring to Eqns.(2.18) and (2.19), we see that bias is reduced if  $W$  is small and  $\lambda_k$  is close to 1. It can be shown that the concentrations  $\lambda_k$  are all very close to 1 for  $k < \lfloor 2NW - 2 \rfloor$  and drop sharply down to 0 for  $k > \lfloor 2NW \rfloor$ . In the multitaper approach, to be discussed shortly, we would like to use as many orthogonal tapers as possible while keeping the bias low. If  $K$  denotes the number of tapers used, a typical choice for Slepian tapers is  $K = \lfloor 2NW - 1 \rfloor$  or  $K = \lfloor 2NW - 2 \rfloor$ . The number of eligible tapers  $K$  decreases with  $W$ , so there is a tradeoff between the number of tapers used (which can be used to reduce variance) and the bandwidth  $W$ , which can be used to reduce bias.

With the condition of normalization (Eqn. 9 from [27]):

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} |V_k(f)|^2 = 1, \quad (2.21)$$

Eqn. (2.19) reduces to just

$$\lambda_k = \int_{-W}^W |V_k(f)|^2 df. \quad (2.22)$$

The eigenvalue-eigenvector problem that produces these  $\lambda_k$  is:

$$\sum_{m=0}^{N-1} \frac{\sin(2\pi W(n-m))}{\pi(n-m)} v_m^{(k)} = \lambda_k(N, W) v_n^{(k)}(N, W) \quad \text{for } n = 0, \dots, N-1. \quad (2.23)$$

Slepian derives the solutions to Eqn. (2.23) as the real solutions to the maximization problem [27]. The maximization causes a reduction in the number of unwanted frequency components in the integral in Eqn.(2.18) outside of the band  $(f - W, f + W)$ . The effect of the unwanted frequencies is usually referred to as *spectral leakage* [34].

A disadvantage to the Slepian tapers is the speed at which they can be calculated [19]. When one only needs to compute them once, it may go unnoticed. However, the additional time can pose a problem if they must be computed many times. This will become more important in the following chapters.

Another well-known set of tapers is the sinusoidal (sine) tapers. These tapers set out to solve a different optimization problem. The sinusoidal tapers are an approximation to the minimum (local) bias tapers (MBT) [25]. For  $k \geq 1$ , the local bias is defined as

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} |V_k(f)|^2 f^2 df, \quad (2.24)$$

where  $V_k(f)$  is the spectral window of the  $k$ th taper. The MBT aims to minimize Eqn. (2.24) in the subspace of vectors orthogonal to to the first  $k - 1$  MBT  $\{v_t^{(1)}, \dots, v_t^{(k-1)}\}$ , subject to the taper having a unit norm. The MBT do not have a convenient closed form but are closely approximated by the sinusoidal tapers, which are defined as:

$$v_n^{(k)} := \sqrt{\frac{2}{N+1}} \sin\left(\frac{\pi(k+1)(n+1)}{N+1}\right) \quad (2.25)$$

These tapers have a convenient closed-form, approach the MBT at a rate of  $\frac{1}{N}$  as  $N \rightarrow \infty$ , are extremely fast to calculate and have other advantages for our specific use case that will come up later. The sinusoidal tapers are defined for  $k \in \{0, \dots, N-1\}$  and  $n \in \{0, \dots, N-1\}$ , but in practice we use  $K < N$  of the tapers. The choice of  $K$

is discussed in the next chapter.

Note that the DPSSs and the sine tapers are not trying to optimize the same problem. The DPSS tapers provide maximum concentration inside the band  $(-W, W)$ , as compared to the sine tapers, which try to minimize the local bias in the spectral window. Thus we cannot directly state what set of tapers is better; it depends entirely on the problem in context as to what tapers will obtain the optimum result. That being said, the DPSSs and sine tapers are created to be  $K$  dimensional orthonormal bases in  $\mathbb{R}^N$  [25, 27]. It is advantageous to have orthonormality when deriving test statistics.

## Multitaper

We define the inner part of Eqn. (2.17) with the general taper  $\{h_t\}$  replaced by the  $k$ th DPSS or sine taper  $\{v_t^{(k)}\}$  as the  $k$ th *eigencoefficient*  $y_k(f)$ :

$$y_k(f) := \sum_{n=0}^{N-1} v_n^{(k)} X_n e^{-i2\pi n f}, \quad (2.26)$$

the multitaper spectral estimate is then created by averaging over the  $K$  modulus squared eigencoefficients:

$$\hat{S}^{(MTM)}(f) := \frac{1}{K} \sum_{k=0}^{K-1} \hat{S}_k(f). \quad (2.27)$$

where

$$\hat{S}_k(f) := |y_k(f)|^2 = \left| \sum_{n=0}^{N-1} v_n^{(k)} X_n e^{-i2\pi n f} \right|^2 \quad (2.28)$$

is called the  $k$ th *eigenspectrum*. Note that the  $k$ th eigenspectrum is just the direct spectrum estimator obtained by using the  $k$ th DPSS or sine taper. By averaging over multiple eigencoefficients, the resulting spectral estimate has a smaller variance by a factor of  $\frac{1}{K}$  than any individual eigenspectra, Eqn. (2.28), as well as obtaining all the bias reduction benefits that come with using the DPSSs or sine tapers rather than  $\frac{1}{\sqrt{N}}$  [30].

## 2.2.4 First Derivatives of DPSSs and Sine Tapers

In the following chapters, defining the derivatives with respect to time of both the DPSS and the Sinusoidal tapers will be necessary. First, we define the continuous-time version of the DPSSs as we cannot differentiate Eqn. (2.23) directly.

$$v^{(k)}(t) := \frac{1}{\lambda_k} \sum_{n=0}^{N-1} \rho(t-n)v_n^{(k)}(N, W) \quad (2.29)$$

where

$$\rho(t) = \begin{cases} \frac{\sin(2\pi Wt)}{\pi t} & t \neq 0 \\ 2W & t = 0. \end{cases}$$

For  $t \in \mathbb{C}$ , the function  $\rho(t)$  is entire. Thus, since  $v^{(k)}(t)$  is a linear combination of entire functions, the derivative of the continuous-time DPSSs must exist  $\forall t \in \mathbb{C}$ . Now, taking the derivative of Eqn. (2.29) (as in [1]) we find:

$$\dot{v}^{(k)}(t) := \frac{1}{\lambda_k} \sum_{n=0}^{N-1} \rho'(t-n)v_n^{(k)}(N, W) \quad (2.30)$$

$$\rho'(t) = \begin{cases} \frac{\cos(2\pi Wt)2W}{t} - \frac{\sin(2\pi Wt)}{\pi t^2} & t \neq 0 \\ 0 & t = 0 \end{cases}$$

Notice that here the main difference between  $v^{(k)}(t)$  and  $v_n^{(k)}$  is that the latter is only defined for  $n \in \{0, 1, 2 \dots, N-1\}$  whereas the former is defined for all complex numbers. Moving onto the derivative of the sinusoidal tapers, we have a much more straightforward derivative due to their closed form:

$$\dot{v}^{(k)}(t) = \sqrt{\frac{2}{N+1}} \cos\left(\frac{(k+1)\pi(t+1)}{N+1}\right) \frac{(k+1)\pi}{N+1}. \quad (2.31)$$

Here,  $\dot{v}^{(k)}(t)$  is also defined for all  $t \in \mathbb{C}$  as for the DPSS time derivative.

## 2.3 Frequency Modulation and Test Statistics

Now that we have defined some necessary background information, we can start the discussion of frequency-modulated (FM) signals. In a time series with a frequency modulated signal, the frequency of the signal shifts around over time in some interval, say  $(f - m_P, f + m_P)$ , where  $m_P$  is the modulation bandwidth. We assume that our time series  $\{X_t\}$  with  $M$  frequency modulated signals has the form

$$X_t = \sum_{m=0}^{M-1} \mu_m \cos \left( 2\pi f_m t + 2\pi \int_0^t \phi_m(\tau) d\tau \right) + Z_t, \quad (2.32)$$

where  $\{Z_t\}$  is a zero-mean stationary noise process,  $\mu_m$  and  $f_m$  are the amplitude and carrier base frequency of the  $m$ th modulated signal, respectively, and  $\phi_m(\tau) := \sum_{p=0}^P a_p \tau^p$  is a polynomial of degree at most  $P$  whose range in the time span of the data is assumed to lie in a given bandwidth around 0. In this thesis we consider only polynomial modulation of the form expressed in Eqn.2.32. We desire to detect the  $f_m$ , the carrier base frequencies, of the frequency-modulated signals. Assuming the  $M$  carrier frequencies  $f_m$  are spaced apart by at least the resolution bandwidth [19], for analysis we may assume there is only a single modulated signal:

$$X_t := \mu \cos \left( 2\pi f_0 t + 2\pi \int_0^t \phi(\tau) d\tau \right) + Z_t. \quad (2.33)$$

Detecting (unmodulated) periodic signals in noise has a long history in many fields. The non-parametric multitaper spectral estimator and the Harmonic  $F$ -test created in [30] are well known for their ability to detect periodic signals in stationary noise. However, under conditions of the carrier frequency being slowly modulated by a polynomial of degree  $P$ , the Harmonic  $F$ -test is often unable to detect the carrier frequency of the frequency modulated (FM) signal.

We propose semi-parametric tests to detect an FM line component modulated by a polynomial up to a given degree  $P$ . Two tests are a modification of the test in [3]. The modified tests are denoted by  $F_4$  and  $F'_4$  in Chapter 3, and the test in [3] is denoted

by  $\tilde{F}_3$ . We also propose a further test statistic, which is an aggregation of these test statistics over a range of multitaper orders, and is denoted by  $T_a$ . The aggregate test  $T_a$  is shown to be more effective under broader conditions than the one constructed in [3], which is based on the test given in [33]. The test statistics in [3] and those proposed in this thesis are constructed within the multitaper framework. However, whereas the test statistic in [3] uses the Discrete Prolate Spheroidal Sequences (DPSSs) [27] as tapers due to their optimal energy concentration properties, we adopt the Sinusoidal tapers [25] as these provide optimal local bias properties as well as a (roughly four-fold) increase in computational efficiency due to their closed form. Along with the ability to down weight specific regions of the bandwidth, this is useful for providing a mechanism for controlling Type I error in our new aggregate test.

### 3. Review Of $\tilde{F}_3$ and Proposed $F_4$ and $F'_4$ Test Statistics

Reviewing the previously developed estimator  $\tilde{F}_3$  [1,3] is necessary before proceeding to the proposed new test statistics  $F_4$  and  $F'_4$  to give the reader a better understanding of the modifications made and the pitfalls of the test statistic  $\tilde{F}_3$  when put to use in practice. The idea behind  $\tilde{F}_3$  is to capture modulation in a signal using the instantaneous frequency (defined in Section 3.1) and to model it by a polynomial of degree  $P$ . In doing so, the test can separate the carrier frequency  $f_0$  from the modulation, which can make the signal undetectable with Thomson's Harmonic  $F$ -test [30].

#### 3.1 Derivation of $\tilde{F}_3$

The test statistic  $\tilde{F}_3$  is created by first computing a vector of eigencoefficients,  $Y_f$ , at a specific frequency  $f$ . Let  $\{X_t; t = 0, \dots, N - 1\}$  be a time series of length  $N$  of the form in Eqn.(2.33). Define

$$Y_f := \begin{pmatrix} y_0(f) \\ y_1(f) \\ \vdots \\ y_{K-1}(f) \end{pmatrix}, \quad X_f := \begin{pmatrix} X_0 e^{-i2\pi f 0} \\ X_1 e^{-i2\pi f 1} \\ \vdots \\ X_{N-1} e^{-i2\pi f (N-1)} \end{pmatrix},$$

$$V := \begin{pmatrix} v_0^{(0)} & v_0^{(1)} & \cdots & v_0^{(K-1)} \\ \vdots & \ddots & & \vdots \\ v_{N-1}^{(0)} & \cdots & \cdots & v_{N-1}^{(K-1)} \end{pmatrix}, \quad \dot{V} := \begin{pmatrix} \dot{v}_0^{(0)} & \dot{v}_0^{(1)} & \cdots & \dot{v}_0^{(K-1)} \\ \vdots & \ddots & & \vdots \\ \dot{v}_{N-1}^{(0)} & \cdots & \cdots & \dot{v}_{N-1}^{(K-1)} \end{pmatrix}$$

The  $k$ th element of the  $K$ -vector  $Y_f$  is given by the  $k^{\text{th}}$  eigencoefficient defined in Eqn. (2.26), where  $h_n$  is replaced by  $v_n^{(k)}$ . The  $N \times K$  matrices  $V$  and  $\dot{V}$  contain the  $K$  tapers and their time derivatives (see Subsection 2.2.4) as their columns. For the  $\tilde{F}_3$  test statistic from [1, 3] the tapers are taken to be the DPSS tapers discussed in Section 2.2.3. For our  $F_4$  and  $F'_4$  test statistics starting in Subsection 3.3, we take the tapers to be the sine tapers, also discussed in Section 2.2.3. Note that  $Y_f$  can also be written without the summation as

$$Y_f = V^T X_f \tag{3.1}$$

Using the eigencoefficients from the multitaper framework, we have transformed the time series into the frequency domain. Next, using the eigencoefficients we construct a length  $N$ , complex-valued, standard inverse projection filter vector  $Z = (Z(0), \dots, Z(N-1))^T$  of the time series at a frequency  $f$  given by

$$Z := VY_f = U + iQ \tag{3.2}$$

and its derivative  $\dot{Z} = (\dot{Z}(0), \dots, \dot{Z}(N-1))^T$

$$\dot{Z} := \dot{V}Y_f = \dot{U} + i\dot{Q}, \tag{3.3}$$

where  $V$  is the  $N \times K$  matrix of DPSS tapers,  $\dot{V}$  is the time derivative of the DPSS given in Eqn. (2.30), and  $K$  is the number of tapers used (we have suppressed the dependence on  $f$  in  $Z$ ,  $U$ ,  $Q$ ,  $\dot{Z}$ ,  $\dot{U}$ , and  $\dot{Q}$  in the preceding notation).

Because our end goal is to gain information on the polynomial frequency modulation as it changes over time, we want to capture the frequency domain information contained in the eigencoefficients and transfer it back to the time domain; this allows us to look at the modulation in a more effective manner. To illustrate this point

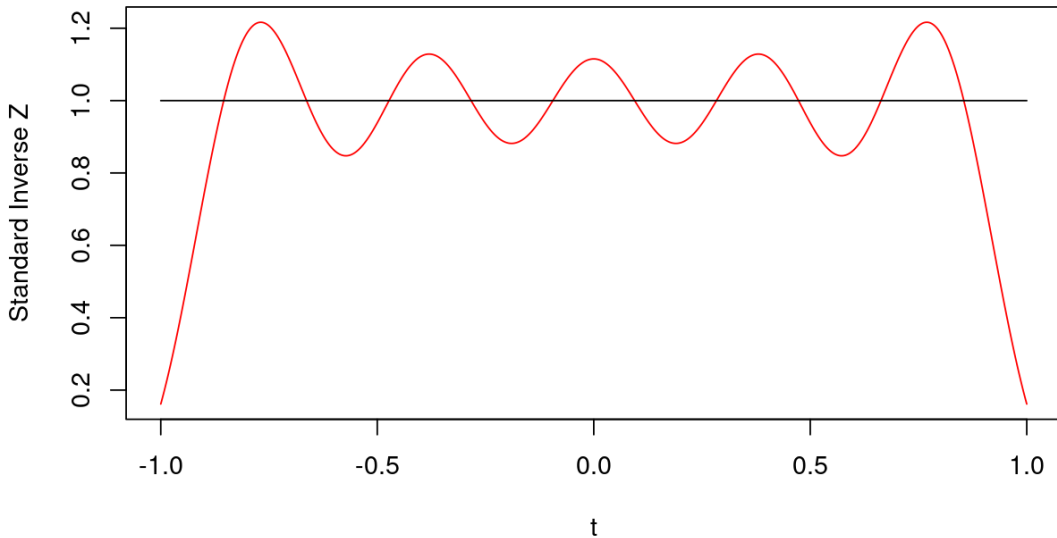


Figure 3.1: Real part of the standard inverse projection filter,  $Z$  shown in red, applied to a linear polynomial shown in black

observe both Fig. 3.1 and Fig. 3.2: the  $Z$  allows for polynomials to be passed through the filter. The ripples in the red lines in these figures are Gibbs ripples [22], [31], but note the tracking of the true polynomial modulation, shown in black. This is advantageous for detecting frequency modulation as we can isolate the degree  $P$  polynomial modulation in the series.

From the standard inverse  $Z$  we can define an instantaneous frequency vector of length  $N$ , which we denote by  $\psi(f)$ , of the time series  $\{X_t\}$ , where at time  $t$ , the component  $\psi(t; f)$  is an estimate of the instantaneous frequency in a prespecified band of width  $2W = \frac{K-1}{N}$  around a frequency  $f$  at time  $t$ . This is given by writing the standard inverse in polar form and taking the derivative of the instantaneous phase,  $\theta(t)$ , with respect to time. First, we write the standard inverse at time  $t$ ,  $Z(t; f)$ , in polar form:

$$Z(t; f) = M(t)e^{i\theta(t)}, \quad (3.4)$$

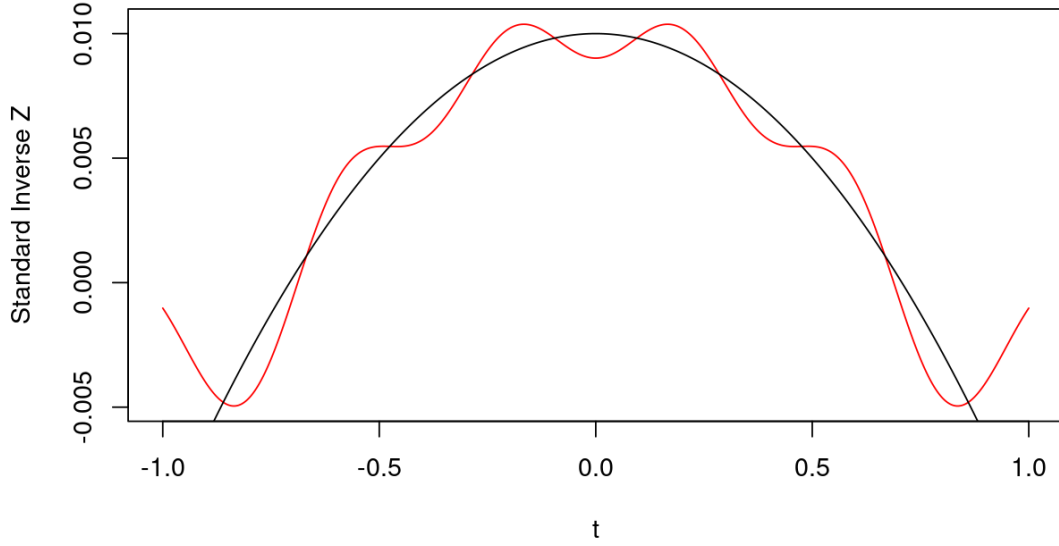


Figure 3.2: Real part of the standard inverse projection filter,  $Z$  shown in red, applied to a quadratic polynomial shown in black.

where

$$M(t) = \sqrt{U(t)^2 + Q(t)^2} \quad \text{and} \quad \theta(t) = \arctan\left(\frac{Q(t)}{U(t)}\right). \quad (3.5)$$

The instantaneous frequency at time  $t$  is then defined as the time derivative of  $\theta(t)$ :

$$\begin{aligned} \psi(t; f) &:= \frac{1}{2\pi} \frac{d}{dt} \theta(t) \\ &= \frac{1}{2\pi} \left[ \frac{1}{1 + \frac{Q^2(t)}{U^2(t)}} \cdot \frac{\dot{Q}(t)U(t) - \dot{U}(t)Q(t)}{U^2(t)} \right] \\ &= \frac{1}{2\pi} \left[ \frac{\dot{Q}(t)U(t) - \dot{U}(t)Q(t)}{U^2(t) + Q^2(t)} \right] \end{aligned}$$

Thus, making the dependence on  $f$  explicit,

$$\psi(t; f) = \frac{U(t; f)\dot{Q}(t; f) - \dot{U}(t; f)Q(t; f)}{2\pi M(t; f)^2}, \quad (3.6)$$

where  $U(t; f)$  is the  $t^{\text{th}}$  component of  $U$ , and similarly for  $\dot{U}(t; f)$ ,  $Q(t; f)$ , and  $\dot{Q}(t; f)$ .

The distribution of  $\psi(t; f)$  has been shown asymptotically (as  $N \rightarrow \infty$ ,  $W \rightarrow 0$ ) to be

$$\psi(t; f) \sim \frac{\text{Laplace}(0, 1)}{\chi_2^2} \quad (3.7)$$

up to a proportionality factor, when  $\{X_t\}$  is purely white noise [1]. Additionally,  $\psi(t; f)$  can be scaled so it has the same distribution at each time  $t$ , which allows for it to be very well-behaved even though it is a non-linear function of a non-stationary process. Computing the eigencoefficients of this new instantaneous frequency series which is representative of the polynomial trends in our original time series,

$$\Psi_k(\nu; f) := \sum_{t=0}^{N-1} v_t^{(k)} \psi(t; f) e^{-i2\pi\nu t}, \quad (3.8)$$

where  $v_t^{(k)}$  is the  $k^{\text{th}}$  data taper, we need only observe the case of  $\nu = 0$  due to the concentration of the polynomial trends near 0 in the instantaneous frequency series.

Thus, we define

$$\Psi(f) := V^T \psi(f), \quad (3.9)$$

where  $\Psi(f) = (\Psi_0(0; f), \dots, \Psi_{K-1}(0; f))^T$  and  $\psi(f) = (\psi(0; f), \dots, \psi(N-1; f))^T$ .

Next, we want to estimate this instantaneous frequency series by a signal that comprises a degree  $P$  deterministic polynomial  $q(n)$  and a zero mean, weakly stationary noise  $\{Z_n\}$ . Then assuming our signal has the form

$$\psi(n; f) = q(n) + Z_n = \sum_{p=0}^P a_p \left( \frac{2n}{N-1} - 1 \right)^p + Z_n. \quad (3.10)$$

As  $Z_n$  is likely not Gaussian in practice, thus we will take the Fourier transform and look at the regression problem in the frequency domain to try and estimate  $q(n)$ .

Computing the eigencoefficients of  $\psi$ ,

$$\Psi_k(\nu; f) = \sum_{n=0}^{N-1} v_n^{(k)} \sum_{p=0}^P a_p \left( \frac{2n}{N-1} - 1 \right)^p e^{-i2\pi\nu n} + z_k(\nu) \quad (3.11)$$

where the  $z_k(\nu)$  are the eigencoefficients of  $\{Z_n\}$ . Looking only at  $\nu = 0$  due to the concentration of the instantaneous frequency lying at zero, and letting  $x_k$  denote

$\Psi_k(0; f)$  for simpler notation, we have

$$x_k = \sum_{p=0}^P a_p \sum_{n=0}^{N-1} v_n^{(k)} \left( \frac{2n}{N-1} - 1 \right)^p + z_k(0) \quad (3.12)$$

$$= \sum_{p=0}^P a_p \sum_{n=0}^{N-1} v_n^{(k)} D_{n,p} + z_k(0) \quad (3.13)$$

$$= \sum_{p=0}^P a_p A_{k,p} + z_k(0), \quad (3.14)$$

where  $D_{n,p} := \left( \frac{2n}{N-1} - 1 \right)^p$  and  $A_{k,p} := \sum_{n=0}^{N-1} v_n^{(k)} D_{n,p}$ . Letting  $x, a, z$  be the vectors containing  $x_k, a_p, z_k(0)$  respectively,  $A$  the  $K \times (P+1)$  matrix with elements  $A_{k,p}$ , and  $R$  the  $N \times (P+1)$  matrix with elements  $D_{n,p}$ , we obtain:

$$x = Aa + z; \text{ where } A = V^T R. \quad (3.15)$$

The first two columns of  $A$  are orthogonal due to the interesting property that the elements in each of these columns alternate between 0 and a nonzero number, but the first column starts with a nonzero number while the second column starts with a 0. This is because the DPSS tapers, (as well as the sine tapers), are alternately even and odd functions about the midpoint of their time span. However,  $A$  doesn't have orthonormal columns and so we conduct a Gram-Schmidt-like procedure. Denote  $H$  as the orthonormalized matrix and  $G$  as the associated polynomial matrix. It should be noted here that  $A_{\cdot 0}$  is the first column of the  $A$  matrix associated with the  $p = 0$  polynomials; thus, the  $p = 1$  column of  $A$  is actually the second column of the  $A$  matrix.

For  $p = 0$

$$H_{k,0} = \frac{A_{k,0}}{\sqrt{\langle A_{\cdot 0}, A_{\cdot 0} \rangle}} \quad (3.16)$$

and

$$G_{n,0} = \frac{R_{n,0}}{\sqrt{\langle A_{\cdot 0}, A_{\cdot 0} \rangle}}. \quad (3.17)$$

For  $1 \leq p \leq P$

$$H_{k,p} = \frac{A_{k,p} - \sum_{l=0}^{p-1} \langle A_{\cdot,p}, A_{\cdot,l} \rangle A_{k,l}}{\sqrt{\langle A_{\cdot,p}, A_{\cdot,p} \rangle}} \quad (3.18)$$

and

$$G_{n,p} = \frac{D_{n,p} - \sum_{l=0}^{p-1} \langle A_{\cdot,p}, A_{\cdot,l} \rangle R_{n,l}}{\sqrt{\langle A_{\cdot,p}, A_{\cdot,p} \rangle}}. \quad (3.19)$$

It should be noted here that  $H$  and  $G$  are both rank  $P + 1$  matrices, but only  $H$  has orthonormal columns; therefore,  $H^T H = I_{P+1}$  but  $G^T G \neq I_{P+1}$ . Our official regression problem can now be stated as:

$$x = Hc + z. \quad (3.20)$$

Using ordinary least squares,

$$\hat{c} = (H^T H)^{-1} H^T x = I_{P+1} H^T x. \quad (3.21)$$

The fitted ordinary least squares representation of the polynomial part of the instantaneous frequency signal in the frequency domain is given by

$$\hat{x} = H\hat{c} = HH^T x \quad (3.22)$$

with residuals

$$r := x - \hat{x} = (I_K - HH^T)x. \quad (3.23)$$

Going back to our original notation from Eqn.3.9, making explicit the dependence on the frequency  $f$  and the order of the polynomial  $P$ , we write the regression model in Eqn.3.20 as

$$\Psi(f) = H_P c_P + z \quad (3.24)$$

and Eqn.3.21 becomes

$$\hat{c}_P := (H_P^T H_P)^{-1} H_P^T \Psi(f) = H_P^T \Psi(f). \quad (3.25)$$

The residuals are given by

$$r_P = (I_K - H_P H_P^T) \Psi(f). \quad (3.26)$$

Below we introduce the notation  $\hat{c}_{P/0}$ , which is  $\hat{c}_P$  without the first component. This was done to improve the performance in the center of the frequency band in [1]. To give an example, if  $P = 1$ , then  $\hat{c}_{P/0}$  would be constructed by using the submatrix  $H_{P/0}$  made up of only the second column of  $H$ . Thus,

$$\hat{c}_{P/0} := (H_{P/0}^T H_{P/0})^{-1} H_{P/0}^T \Psi(f) = H_{P/0}^T \Psi(f) \quad (3.27)$$

$$r_{P/0} := (I_K - H_{P/0} H_{P/0}^T) \Psi(f) \quad (3.28)$$

where  $\hat{c}_{P/0}$  has dimension  $P \times 1$  with elements  $\hat{c}_{p/0}$  and  $r_{P/0}$  has dimension  $K \times 1$ . The test statistic  $\tilde{F}_3$  is then defined as:

$$\tilde{F}_3(P; f, K) := \frac{(\hat{c}_{p/0})^2 (K - P)}{\|r_{P/0}\|^2} \quad (3.29)$$

or equivalently

$$\tilde{F}_3(P; f, K) = \frac{(\hat{c}_{p/0})^2 (K - P)}{\|\Psi(f)\|^2 - \|\hat{c}_{p/0}\|^2} \quad (3.30)$$

Note the subtle notational difference of the lower-case  $p$  and the upper-case  $P$  in these two definitions. As  $N \rightarrow \infty, W \rightarrow 0$ ,  $\tilde{F}_3$  converges in distribution to an  $F$  distribution with 1 and  $K - P$  degrees of freedom [3].

## 3.2 Issues with $\tilde{F}_3$ in practice

Now that  $\tilde{F}_3$  has been derived, we can explain why one would want to improve this test. Due to the use of tapers in the  $F$ -test, specifically DPSSs in the derivation above, one must choose the number of tapers  $K$  used to conduct the test. This would not become an issue if, for any number of tapers you chose, the test was coherent, in the sense that the tests would come to the same conclusion regardless of the  $K$  chosen. Unfortunately, this is not the case; in fact, the amount of underlying modulation in

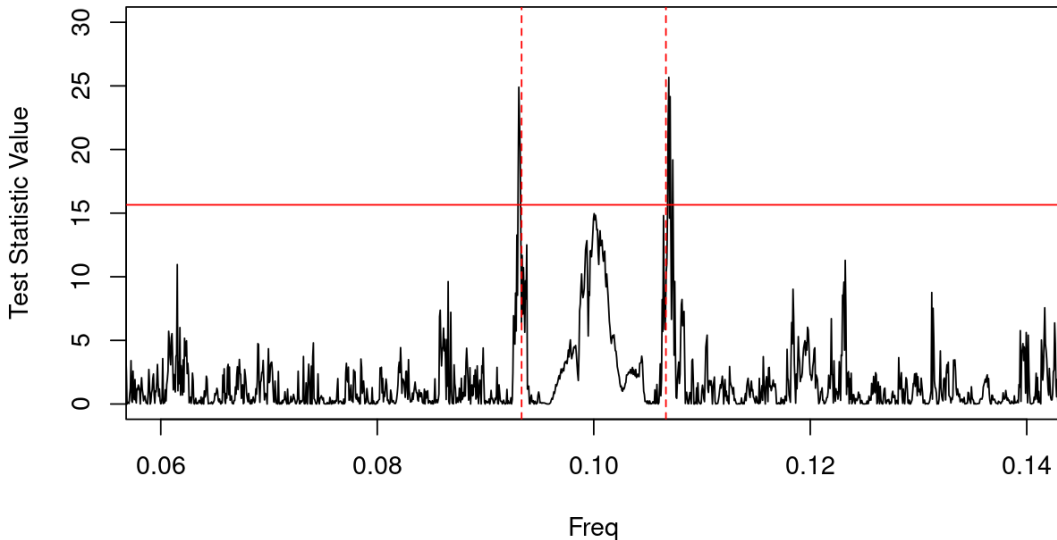


Figure 3.3:  $\tilde{F}_3$  applied to a frequency modulated series (carrier frequency  $f = 0.1$ ) with input parameter  $K = 79$ . Dotted red lines represent  $f \pm \frac{(K+1)}{2N}$ , and the horizontal red line is the quantile of an  $F_{1,79-1}$  corresponding to a significance level of  $\frac{1}{2}$ . This examples uses  $N = 6000$  and  $\phi(\tau) = \frac{0.002}{\frac{N}{2}} \left(t - \frac{N}{2}\right)$ , and  $Z_t$  is Gaussian white noise with a low signal to noise ratio

the time series affects the optimal choice of  $K$ .

One example of where the choice of  $K$  is very influential is in the case of a very large  $K$ . Looking at the two figures Fig. 3.3 and Fig. 3.4 we can see that the choice of  $K$  causes the carrier frequency to be detected in multiple spots simultaneously. In Fig. 3.3, the carrier frequency of 0.1 is not detected at any point, but at the sides of the spectral window  $f \pm \frac{(K+1)}{2N}$  we obtain significant detections. In Fig. 3.4 we do detect the carrier frequency at 0.1, but we again detect a significant frequency on the right side of the spectral window. This behaviour is obviously not ideal; in the case of Fig. 3.3 only modeling the splitting into the side spikes would obtain the truth.

A second issue we desired to fix with  $\tilde{F}_3$  has to do with even versus odd choices of  $K$ . If one were to take a time series and conduct the  $\tilde{F}_3$  test at an even  $K$  and the

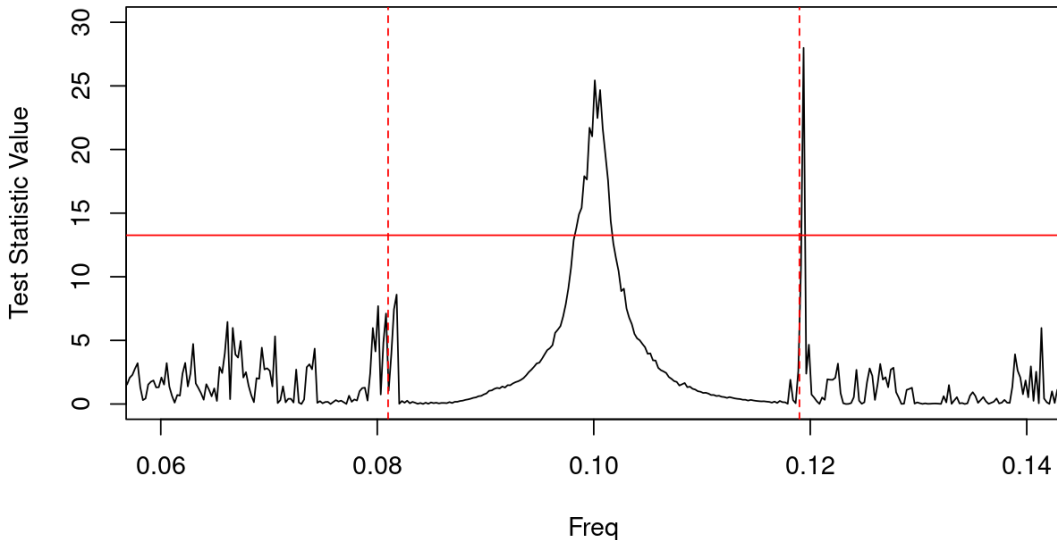


Figure 3.4:  $\tilde{F}_3$  applied to FM series (carrier frequency  $f = 0.1$ ) with input parameter  $K = 79$ . Dotted red lines represent  $f \pm \frac{(K+1)}{2N}$ , and the horizontal red line is the quantile of an  $F_{1,79-1}$  corresponding to a significance level of  $\frac{1}{2}$ . Where  $N = 2000$  and  $\phi(\tau) = \frac{0.001}{\frac{N}{2}} (t - \frac{N}{2})$ , and  $Z_t$  is Gaussian white noise with a signal to noise ratio of 1

(incrementally) next odd  $K$ , the probability of detecting the carrier frequency with the even choice of  $K$  is very much reduced. An example of this behaviour can be seen in Fig. 3.5. The left-hand figure of this plot shows that the carrier frequency of 0.1 is not detected at a  $\alpha = \frac{1}{N}$  significance level. On the other hand, the  $K = 9$  plot on the right in the figure does successfully detect the carrier frequency.

### 3.3 Weighted $F_4$ Test Statistic

A new test, which we have denoted  $F_4$ , was created to help mitigate the issue of large  $K$  values causing false detections near the sides of the window. This was achieved by first examining the spectral window. By replacing the time series in the expression for the eigencoefficients (2.26) (technically, replacing the time series with a series equal

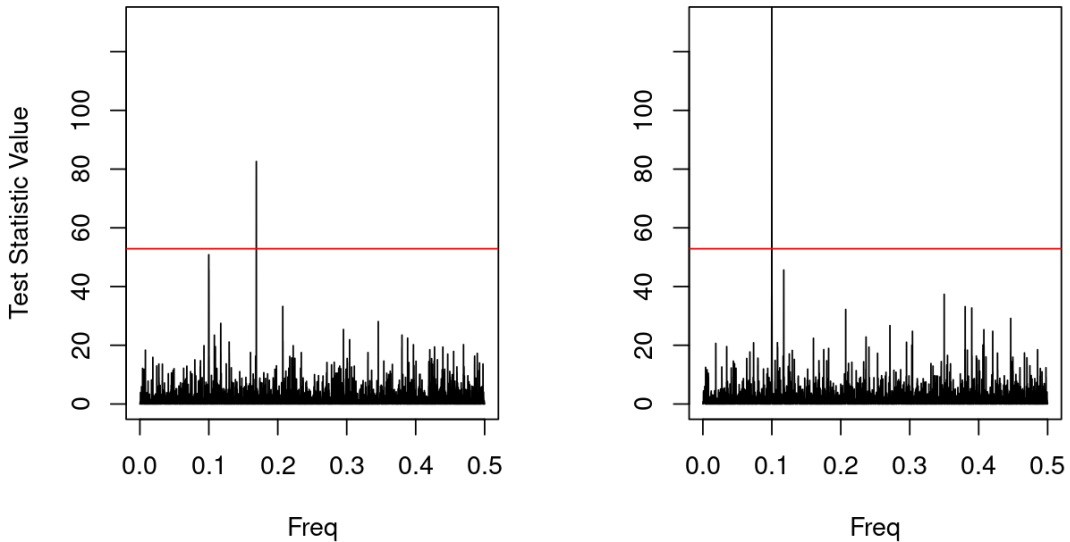


Figure 3.5:  $\tilde{F}_3$  applied to FM series (carrier frequency  $f = 0.1$ ). The left-hand plot is  $K = 8$  and the right-hand plot is  $K = 9$ . The horizontal red line is the quantile of an  $F_{1,K-1}$  corresponding to a significance level of  $\frac{1}{2}$ . In this example,  $N = 6000$  and  $\phi(\tau) = \frac{0.001}{\frac{N}{2}} (t - \frac{N}{2})$ , and  $Z_t$  is Gaussian white noise with a signal to noise ratio of 1.

to 1 everywhere) and computing magnitude spectrum, we can obtain the spectral window for a specific taper given by:

$$\mathcal{H}(f) := \left| \sum_{n=0}^{N-1} v_n^{(k)} e^{-i2\pi fn} \right|^2 \quad (3.31)$$

The spectral window gives us insight into the behaviour of only the tapers within the window  $(f - W, f + W)$  without the influence of the time series. For a given high  $K$ , such as  $K = 80$ , our goal was to reduce the concentration closer to the  $W$  band while keeping a high concentration near the centre frequency that is being tested. Looking first at Fig. 3.6 we can see that the lower order tapers making up the windows of both the DPSS and Sinusoidal tapers are concentrated near the center of the band ( $f = 0$  in this case). Comparing this to Fig. 3.7, we see that the DPSSs

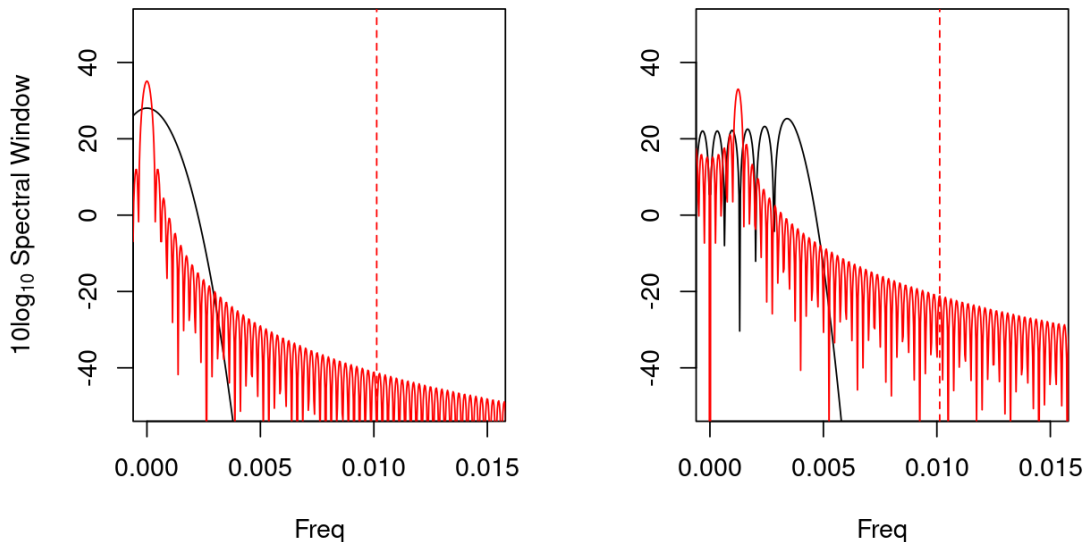


Figure 3.6: Selected spectral windows for  $K = 80$  for both Sinusoidal (in red) and DPSS tapers (in black). The left-hand plot is the  $k = 0$  and the right-hand plot is the  $k = 9$  spectral window. The red dotted line is the width of the window given by  $W = \frac{(K-1)}{2N}$ .

maintains high concentration along the entirety of the bandwidth for the higher-order tapers. In contrast, the Sinusoidal tapers only focus the concentration near the band-edge. Therefore, by down-weighting these higher-order tapers, we could reduce the false detections at the edges of the band. We discovered that due to the consistently high concentration of the DPSSs across the entire window for higher-order tapers, the down-weighting did not work as intended. Rather, it reduced the detection probability at the center frequency as well. In contrast to this, for the Sinusoidal tapers down-weighting did work as intended due to the concentration at higher orders being focused in a specific part of the window near the edge of the band, therefore allowing us to tune the weighting in such a way to have very little concentration at the band edges for the estimator.

To create  $F_4$ , consider  $K$  tapers, enumerated from 0 to  $K - 1$ . The matrix of

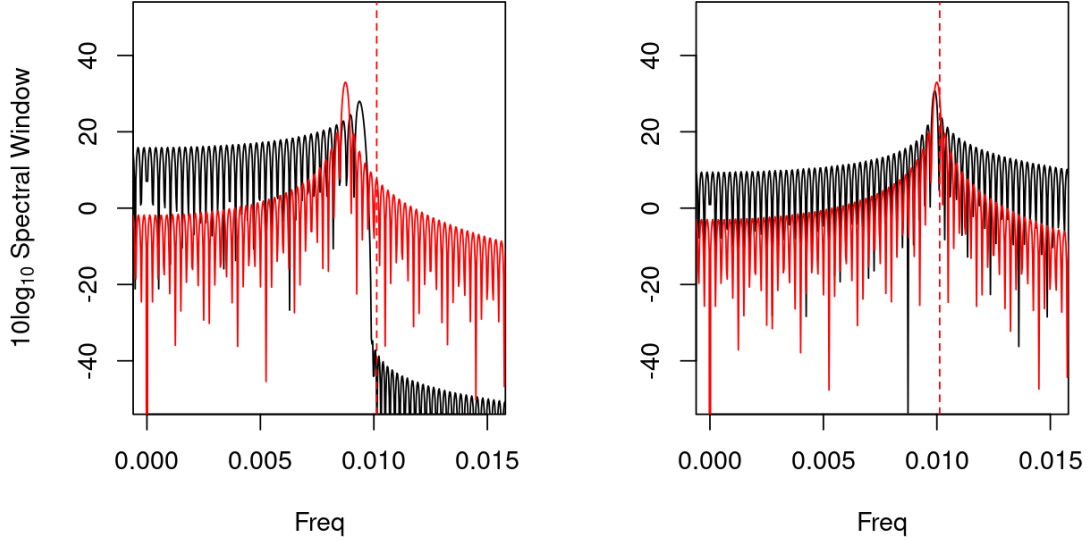


Figure 3.7: Spectral windows for  $K = 80$  for both Sinusoidal (in red) and DPSS tapers (in black). The left-hand plot is the  $k = 69$  and the right-hand plot is the  $k = 79$  spectral window. The red dotted line is the width of the window given by  $W = \frac{(K-1)}{2N}$ .

Sine tapers ( $V$ ) and Sine taper derivatives ( $\dot{V}$ ) are created by evaluating Eqn. (2.25) and Eqn. (2.31), respectively, on a discrete mesh over  $t \in \{0, \dots, N-1\}$  and  $k \in \{0, \dots, K-1\}$ , and define a matrix of weights  $\mathbf{W}$ :

$$V := \begin{pmatrix} v_0^{(0)} & v_0^{(1)} & \cdots & v_0^{(K-1)} \\ \vdots & \ddots & & \vdots \\ v_{N-1}^{(0)} & \cdots & \cdots & v_{N-1}^{(K-1)} \end{pmatrix}, \dot{V} := \begin{pmatrix} \dot{v}_0^{(0)} & \dot{v}_0^{(1)} & \cdots & \dot{v}_0^{(K-1)} \\ \vdots & \ddots & & \vdots \\ \dot{v}_{N-1}^{(0)} & \cdots & \cdots & \dot{v}_{N-1}^{(K-1)} \end{pmatrix}, \mathbf{W} := \begin{pmatrix} w_0 & 0 & \cdots & 0 \\ 0 & w_1 & \cdots & 0 \\ 0 & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & w_{K-1} \end{pmatrix}.$$

Notice that these taper matrices are identical in dimension to the ones defined for  $\tilde{F}_3$ , other than the type of tapers that are used in their creation, and that we have used the notation of boldface ( $\mathbf{W}$ ) to distinguish the weight matrix from the multitaper bandwidth  $W$ . Now, let

$$\tilde{Y}_f := \mathbf{W}Y_f \tag{3.32}$$

where  $Y_f$  is defined in Eqn. (3.1).  $F_4$  can now be computed by following the same method to create  $\tilde{F}_3$ , replacing Eqn. (3.1) with Eqn. (3.32) and using the newly defined  $V$  and  $\dot{V}$  instead of the DPSS taper matrices. The resulting  $F$ -test has the form

$$F_4(P; f, K) := \frac{(\hat{c}_{P/0}(f))^2 (K - P)}{\|\Psi(f)\|^2 - \|\hat{c}_{P/0}\|^2}. \quad (3.33)$$

Note that this is identical in form to Eqn. (3.30), with the difference being in both  $\Psi(f)$  and  $\hat{c}_{P/0}$  terms through the modified tapers. The weights used for the remainder of the thesis are such that the elements on the diagonal of  $\mathbf{W}$  are given by

$$\left\{ \frac{1}{1 + \frac{M-1}{K-1} \cdot i} \mid i = 0, \dots, K - 1 \right\}, \quad (3.34)$$

where  $M = \max(1, \text{penalty} \cdot K)$ . So in one extreme with  $\text{penalty} \leq \frac{1}{K}$ ,  $M = 1$  and all weights are  $1/1 = 1$ ; for  $\text{penalty} > \frac{1}{K}$ , the weights decrease as  $\frac{1}{1+ax}$ , where  $a > 0$ , from 1 for  $i = 0$  to a quantity less than 1. As an example, Fig. 3.8 illustrates the weights we used in the simulations in Chapter 5. One can see that the weight is 1 for the first taper, and the higher-order tapers are down-weighted significantly. We did not downweight the higher order tapers to zero as, under simulation, that resulted in a lower detection probability.

### 3.4 Distribution of $F_4$ Under the Null Hypothesis

The first concern with this new  $F$ -test was to verify the distribution of the test statistic  $F_4$  under the null hypothesis, as the weighting could, for example, affect the degrees of freedom of the test statistic. Therefore we will now derive the null distribution of the  $F_4$  test statistic under the null hypothesis that there is no frequency modulation present at frequency  $f$ . For this analysis we make the following assumptions.

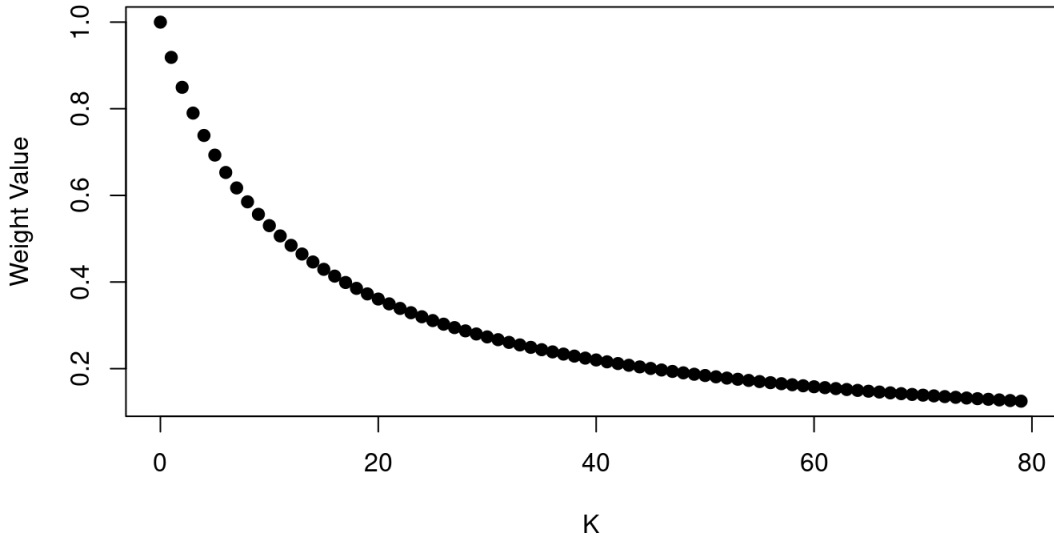


Figure 3.8: Weighting scheme generated for  $K = 80$ , with a penalty value of 0.10.

## Assumptions

1.  $\{X_t\}$  is a zero-mean white noise process;
2. It is reasonable to treat the eigencoeficients as asymptotically Gaussian;
3. All weights are non-zero.

### 3.4.1 Eigencoeficient Distribution

We start with the modified eigencoeficients including the weighting from above, where  $w_k$  is the  $k^{\text{th}}$  element on the diagonal of the weight matrix  $\mathbf{W}$ . In the following, let  $f = f_0$  be a fixed frequency and, in an abuse of notation, let the  $k$ th weighted eigencoeficient at  $f_0$  be denoted by  $y_k(f_0)$  (not to be confused with Eqn. (2.26)):

$$y_k(f_0) = w_k \sum_{n=0}^{N-1} v_n^{(k)} X_n e^{-i2\pi f_0 n}. \quad (3.35)$$

Also, since  $\{X_t\}$  is a zero-mean process,

$$\mathbb{E}[y_k(f_0)] = 0. \quad (3.36)$$

Since the tapers are orthogonal and  $\{X_t\}$  is a white noise process,  $\text{Cov}[y_j(f_0), y_k(f_0)^*] = 0$  for  $i \neq j$ , where superscript  $*$  represents the complex conjugate. If  $j = k$  then:

$$\begin{aligned} \text{Cov}[y_k(f_0), y_k(f_0)^*] &= w_k w_k \sum_{n=0}^{N-1} v_n^{(k)} v_n^{(k)} \mathbb{E}[X_n^2] \\ &= w_k^2 \sum_{n=0}^{N-1} v_n^{(k)} v_n^{(k)} \text{Var}[X_n] \\ &= w_k^2 S(f_0) \sum_{n=0}^{N-1} \left(v_n^{(k)}\right)^2 \\ &= w_k^2 S(f_0), \end{aligned}$$

where  $S(f_0)$  is the spectrum at  $f_0$ , equal to the variance  $\sigma^2$  of the white noise process. The final equality follows from the orthonormality of the tapers. Thus, the covariance matrix of the eigencoefficients is  $\Sigma_{f_0} = S(f_0) \mathbf{W}^2$ .

Asymptotically (fixed  $K, N \rightarrow \infty, W \rightarrow 0$ ) the eigencoefficients  $y_k(f)$  are complex Gaussian functions of frequency, so the specific  $y_k(f_0)$  at frequency  $f_0$  is a complex Gaussian as well:

$$y_k(f_0) \xrightarrow{d} \begin{cases} \mathcal{N}\left(0, w_k^2 S(f_0)\right); & f_0 = 0, \frac{1}{2} \\ \mathcal{N}_{\mathbb{C}}\left(0, w_k^2 S(f_0)\right); & \text{else} \end{cases}$$

Due to the lack of correlation and being approximately complex Gaussian, we conclude that  $y_i(f_0), y_j(f_0)$  are approximately independent for  $i \neq j$ .

### 3.4.2 Real and Imaginary Parts Of Eigencoefficients

Without loss of generality, assume  $N$  to be odd with  $t = \frac{-(N-1)}{2}, \dots, \frac{(N-1)}{2}$ . For now we let  $u(t)$  and  $v(t)$  be arbitrary tapers with  $U_t = u(t)X_t$  and  $V_t = v(t)X_t$  and let  $w_u$  and  $w_v$  be the weights corresponding to the respective tapers. Abusing our notation

somewhat, we will use  $u$  and  $v$  as subscripts in place of the index variable  $k$ , so rather than being an index number, the subscript will refer directly to the specific choice of taper. Then:

$$\begin{aligned}
y_u(f_0) &= w_u \sum_{t=-(N-1)/2}^{(N-1)/2} U_t e^{-i2\pi f_0 t} \\
&= w_u \sum_{t=-(N-1)/2}^{(N-1)/2} (U_t^E + U_t^O) e^{-i2\pi f_0 t} \\
&= y_u^E(f_0) + y_u^O(f_0)
\end{aligned}$$

where  $U_t^E := \frac{U_t + U_{-t}}{2} = \frac{u(t)X_t + u(-t)X_{-t}}{2}$  and  $U_t^O := \frac{U_t - U_{-t}}{2} = \frac{u(t)X_t - u(-t)X_{-t}}{2}$  are the even and odd parts of  $X_t$ . Let us first show that  $y_u^E(f_0) \in \mathbb{R}$  by expanding out our definition of  $U_t^E$ :

$$\begin{aligned}
y_u^E(f_0) &= w_u \sum_{t=-(N-1)/2}^{(N-1)/2} U_t^E e^{-i2\pi f_0 t} \\
&= \frac{w_u}{2} \left( \sum_{t=-(N-1)/2}^{(N-1)/2} U_t e^{-i2\pi f_0 t} + \sum_{t=-(N-1)/2}^{(N-1)/2} U_{-t} e^{-i2\pi f_0 t} \right) \\
&= \frac{w_u}{2} \left( \sum_{t=-(N-1)/2}^{(N-1)/2} U_t e^{-i2\pi f_0 t} + \sum_{t'=--(N-1)/2}^{(N-1)/2} U_{t'} e^{i2\pi f_0 t'} \right); \text{ with } t' = -t \quad (3.37)
\end{aligned}$$

$$= w_u \sum_{s=-(N-1)/2}^{(N-1)/2} U_s \cos(2\pi f_0 s), \quad (3.38)$$

using Euler's formula  $e^{i\theta} = \cos \theta + i \sin \theta$ . This shows that  $y_u^E(f_0) \in \mathbb{R}$ . This same method can be used to show that  $y_u^O(f_0) \in i\mathbb{R}$ , where the cosine terms cancel, and the sine terms survive in Euler's formula. Also notice that since  $\{X_t\}$  is zero-mean, so are  $y_u^E(f_0)$  and  $y_u^O(f_0)$ .

Looking at  $\text{Cov}[y_u^E(f_0), y_v^O(f_0)]$ ,

$$\text{E}[y_u^E(f_0) y_v^O(f_0)^*] = w_u w_v \sum_{s=-(N-1)/2}^{(N-1)/2} \sum_{t=-(N-1)/2}^{(N-1)/2} \text{E}[U_s^E V_t^O] e^{-i2\pi f_0 (s-t)}. \quad (3.39)$$

Expanding out  $E[U_s^E V_t^O]$ :

$$\begin{aligned}
E[U_s^E V_t^O] &= E\left[\left(\frac{u(s)X_s + u(-s)X_{-s}}{2}\right)\left(\frac{v(t)X_t - v(-t)X_{-t}}{2}\right)\right] \\
&= \frac{1}{4}E\left[(u(s)X_s + u(-s)X_{-s})(v(t)X_t - v(-t)X_{-t})\right] \\
&= \frac{1}{4}\left[u(s)v(t)\gamma_X(t-s) - u(s)v(-t)\gamma_X(-t-s)\right. \\
&\quad \left.+ u(-s)v(t)\gamma_X(t+s) - u(-s)v(-t)\gamma_X(-t+s)\right],
\end{aligned}$$

where  $\gamma_X(\cdot)$  is the autocovariance function of  $\{X_t\}$ . Since  $\{X_t\}$  is white noise,  $\gamma_X(\tau) = 0$  for all  $\tau \neq 0$  and  $\gamma_X(0) = \sigma^2$ . So all of the ACVF terms above are equal to 0 unless  $t = s$  or  $t = -s$ . Therefore,

$$\begin{aligned}
E[y_u^E(f_0)y_v^O(f_0)^*] &= w_u w_v \sum_{s=-(N-1)/2}^{(N-1)/2} \sum_{t=-(N-1)/2}^{(N-1)/2} E[U_s^E V_t^O] e^{-i2\pi f_0(s-t)} \\
&= w_u w_v \frac{1}{4} \sum_{s=-(N-1)/2}^{(N-1)/2} [u(s)v(-s) - u(-s)v(s)] \gamma_X(0) \\
&\quad + w_u w_v \frac{1}{4} \sum_{s=-(N-1)/2}^{(N-1)/2} [u(-s)v(-s) - u(s)v(s)] \gamma_X(0) \\
&= 0,
\end{aligned} \tag{3.40}$$

since the sum over  $s$  of  $u(s)v(-s)$  and over  $u(-s)v(s)$  are the same, and similarly the sum over  $u(-s)v(-s)$  and over  $u(s)v(s)$  are the same. Thus, with  $u = v$ , we have that the real and imaginary parts of  $y_u(f_0)$  are uncorrelated, and so approximately independent, from the real and imaginary parts being approximately Gaussian from section 3.4. We now want to show that the real and imaginary parts have approximately the same variance, thus making them approximately independent and identically distributed.

For this, we will assume that all tapers are unit vectors, which includes the Slepian DPSSs and the sine tapers as possible candidates. Starting with finding the variance

of  $y_u^E(f_0)$ ,

$$\begin{aligned}
\text{Var}[y_u^E(f_0)] &= \text{E}[y_u^E(f_0)y_u^E(f_0)^*] \\
&= w_u^2 \sum_{t=-(N-1)/2}^{(N-1)/2} \sum_{s=-(N-1)/2}^{(N-1)/2} \text{E}[U_t^E U_s^E] e^{-i2\pi f_0 t} e^{+i2\pi f_0 s} \\
&= w_u^2 \sum_{t=-(N-1)/2}^{(N-1)/2} \text{Var}[U_t^E], \tag{3.41}
\end{aligned}$$

as for  $t \neq s$ ,  $\text{E}[U_t^E U_s^E] = 0$  by  $\{X_t\}$  being a white noise process. Expanding further,

$$\begin{aligned}
\text{Var}[U_t^E] &= \frac{1}{4} \text{Var}[u(t)X_t + u(-t)X_{-t}] \\
&= \frac{1}{4} \left( u(t)^2 \text{E}[X_t^2] + u(-t)^2 \text{E}[X_{-t}^2] + 2u(t)u(-t) \text{E}[X_t X_{-t}] \right) \\
&= \begin{cases} u(0)^2 \sigma^2 & \text{for } t = 0 \\ \frac{(u(t)^2 + u(-t)^2) \sigma^2}{4} & \text{else.} \end{cases} \tag{3.42}
\end{aligned}$$

Substituting Eqn. (3.42) into Eqn. (3.41),

$$\begin{aligned}
\text{Var}[y_u^E(f_0)] &= \left( \frac{w_u^2 \sigma^2}{4} \sum_{t=-(N-1)/2}^{(N-1)/2} (u(t)^2 + u(-t)^2) \right) + \frac{1}{2} w_u^2 \sigma^2 u(0)^2 \\
&= \frac{1}{2} w_u^2 \sigma^2 (1 + u(0)^2). \tag{3.43}
\end{aligned}$$

Next, finding the variance for  $y_u^O(f_0)$ ,

$$\text{Var}[y_u^O(f_0)] = w_u^2 \sum_{t=-(N-1)/2}^{(N-1)/2} \text{Var}[U_t^O], \tag{3.44}$$

where,

$$\text{Var}[U_t^O] = \begin{cases} 0 & \text{for } t = 0 \\ \frac{(u(t)^2 + u(-t)^2) \sigma^2}{4} & \text{else.} \end{cases} \tag{3.45}$$

Substituting Eqn. (3.45) into Eqn. (3.44),

$$\begin{aligned}
\text{Var}[y_u^O(f_0)] &= \left( \frac{w_u^2 \sigma^2}{4} \sum_{t=-(N-1)/2}^{(N-1)/2} (u(t)^2 + u(-t)^2) \right) - \frac{1}{2} w_u^2 \sigma^2 u(0)^2 \\
&= \frac{1}{2} w_u^2 \sigma^2 (1 - u(0)^2). \tag{3.46}
\end{aligned}$$

This is an interesting result, showing that the variance of the real and imaginary parts are very similar, with differences on the order of 0.5%<sup>1</sup>, which means that these variances are approximately equal, with exact equality for tapers which are odd (that is, are zero at the midpoint of their time domain). For the remainder of this thesis, we will take

$$\text{Var}[y_u^E(f_0)] = \text{Var}[y_u^O(f_0)] := \frac{1}{2}\text{Var}[y_u(f_0)] = \frac{1}{2}w_u^2\mathcal{S}(f_0) \quad (3.47)$$

when using the DPSS or Sinusoidal tapers. This approximation remains valid even for  $\{X_t\}$  not equal to white noise as long as the spectrum of  $\{X_t\}$  is locally white (i.e., approximately flat) around  $f_0$ .

### 3.4.3 Standard Inverse and its Derivative

Using (3.32), the standard inverse and its derivative are  $N \times 1$  vectors defined as:

$$\mathbf{Z} := V\tilde{Y}_f = U + iQ \quad (3.48)$$

$$\dot{\mathbf{Z}} := \dot{V}\tilde{Y}_f = \dot{U} + i\dot{Q} \quad (3.49)$$

where the  $t^{\text{th}}$  element of these vectors for a given  $f$  is:

$$Z(t; f) = U(t) + iQ(t) = \sum_{k=0}^{K-1} v_t^{(k)} y_k(f) \quad (3.50)$$

$$\dot{Z}(t; f) = \dot{U}(t) + i\dot{Q}(t) = \sum_{k=0}^{K-1} \dot{v}_t^{(k)} y_k(f), \quad (3.51)$$

and  $y_k(f)$  is the  $k$ th weighted eigencoefficient in Eqn. (3.35). We assume that  $\{X_t\}$  is mean 0, with a locally white spectrum, and is weakly stationary.

---

<sup>1</sup>For DPSSs with  $N = 501$ ,  $K = 80$  and  $W = (K + 1)/(2N)$ , the  $u(0)^2$  quantity is on the order of 0.0025 for even tapers, and exactly 0 for odd tapers.

## Distributions of the Standard Inverse and its Derivative

From subsection 3.4.1

$$\tilde{Y}_f \approx \begin{cases} \mathcal{N}(0, \Sigma_f); & f = 0, \frac{1}{2} \\ \mathcal{N}_{\mathbb{C}}(0, \Sigma_f); & \text{else,} \end{cases}$$

where  $\Sigma_f = S(f)\mathbf{W}^2$ . Then for  $Z = V\tilde{Y}_f$ ,

$$\mathbb{E}[Z] = \mathbb{E}[V\tilde{Y}_f] = V\mathbb{E}[\tilde{Y}_f] = 0$$

and the covariance matrix of  $Z$  is

$$\text{Cov}[Z] = \mathbb{E}[ZZ^H] = \mathbb{E}[V\tilde{Y}_f(V\tilde{Y}_f)^H] = \mathbb{E}[V(\tilde{Y}_f\tilde{Y}_f^H)V^T] = V\mathbb{E}[\tilde{Y}_f\tilde{Y}_f^H]V^T = V\Sigma_fV^T,$$

where  $H$  is the complex conjugate transpose operator. Since  $Z$  is a linear transformation of  $\tilde{Y}_f$ ,  $Z$  will also be approximately multivariate complex Gaussian with

$$Z \approx \begin{cases} \mathcal{N}(0, V\Sigma_fV^T); & f = 0, \frac{1}{2} \\ \mathcal{N}_{\mathbb{C}}(0, V\Sigma_fV^T); & \text{else.} \end{cases}$$

Similarly, the distribution of  $\dot{Z} = \dot{V}\tilde{Y}_f$  will also be approximately multivariate complex Gaussian with

$$\dot{Z} \approx \begin{cases} \mathcal{N}(0, \dot{V}\Sigma_f\dot{V}^T); & f = 0, \frac{1}{2} \\ \mathcal{N}_{\mathbb{C}}(0, \dot{V}\Sigma_f\dot{V}^T); & \text{else.} \end{cases}$$

### Real and Imaginary Parts of $Z$ and $\dot{Z}$

Let  $Y_R = \Re(\tilde{Y}_f)$  and  $Y_I = \Im(\tilde{Y}_f)$  be the real and imaginary parts of  $\tilde{Y}_f$ . Then

$$U := VY_R \quad \dot{U} := \dot{V}Y_R \quad Q := VY_I \quad \dot{Q} := \dot{V}Y_I$$

It should be noted here that these are  $N \times 1$  vectors of time at a given frequency  $f$ . Looking at the covariances between the real and imaginary parts of  $Z$  and  $\dot{Z}$ , we have

first that the cross-covariance matrix between  $U$  and  $Q$  is

$$\text{Cov}[U, Q] = E[UQ^T] = E[VY_R Y_I^T V^T] = VE[Y_R Y_I^T]V^T = 0$$

as the real and imaginary parts of the weighted eigencoefficients are uncorrelated (from subsection 3.4.2). Again referring back to subsection 3.4.2, the covariance matrices of  $U$  and  $Q$  are

$$E[UU^T] = E[QQ^T] = \frac{1}{2}V\Sigma_f V^T$$

and by the same reasoning,

$$E[\dot{U}Q^T] = E[U\dot{Q}^T] = E[\dot{U}\dot{Q}^T] = 0$$

$$E[\dot{U}\dot{U}^T] = E[\dot{Q}\dot{Q}^T] = \frac{1}{2}\dot{V}\Sigma_f \dot{V}^T$$

$$\text{Cov}[Z, \dot{Z}] = E[V\tilde{Y}_f \tilde{Y}_f^H \dot{V}^T] = V\Sigma_f \dot{V}^T \quad (3.52)$$

### 3.4.4 Instantaneous Frequency

Recall from Eqn. (3.6), the instantaneous frequency is given by

$$\psi(t; f) = \frac{U(t)\dot{Q}(t) - \dot{U}(t)Q(t)}{2\pi M^2(t)}.$$

In the following we will consider the distributions of the numerator and denominator separately, and then the distribution of  $\psi(t; f)$ .

#### 3.4.4.1 Distribution of Denominator of $\psi(t; f)$

Here we are assuming that the noise process is Gaussian and weakly stationary, so the Gaussian approximations in subsection 3.4.3 hold true exactly. If the noise is not Gaussian, then the distributions below are considered to be asymptotic.

First, let

$$\begin{aligned}
\tilde{s}_t^2 &:= \text{Var}[U(t)] = \text{Var} \left[ \sum_{k=0}^{K-1} v_t^{(k)} y_k^R(f) \right]; \text{ from Eqn. (3.50)} \\
&= \sum_{k=0}^{K-1} \left( v_t^{(k)} \right)^2 \text{Var}[y_k^R(f)] \\
&= \sum_{k=0}^{K-1} \left( v_t^{(k)} \right)^2 \frac{1}{2} \text{Var}[y_k(f)] \\
&= \sum_{k=0}^{K-1} \left( v_t^{(k)} \right)^2 \frac{1}{2} w_k^2 S(f),
\end{aligned}$$

where  $y_k^R(f)$  is the real part of  $y_k(f)$  (we denoted this by  $y_k^E(f)$  in subsection 3.4.2).

Thus, letting  $\tilde{\sigma}_t^2 := \sum_{k=0}^{K-1} (v_t^{(k)} w_k)^2$ , we have

$$\tilde{s}_t^2 = \frac{\tilde{\sigma}_t^2 S(f)}{2}, \quad (3.53)$$

and

$$\dot{s}_t^2 := \text{Var}[\dot{Q}(t)] = \text{Var}[\dot{U}(t)] = \sum_{k=0}^{K-1} \left( \dot{v}_t^{(k)} \right)^2 \frac{1}{2} w_k^2 S(f). \quad (3.54)$$

Letting  $\dot{\sigma}_t^2 := \sum_{k=0}^{K-1} \left( \dot{v}_t^{(k)} w_k \right)^2$ , this gives

$$\dot{s}_t^2 = \frac{\dot{\sigma}_t^2 S(f)}{2}. \quad (3.55)$$

Note that  $Q(t)$  can be shown to also have variance given by 3.53. As both  $U(t)$  and  $Q(t)$  are linear combinations of Gaussian random variables with zero mean and identical variance at any time,  $t$ , they are distributed as Gaussian random variables as well. Furthermore, they are independent. The squared sum of  $U(t)$  and  $Q(t)$ , will result in

$$M^2(t) = U^2(t) + Q^2(t) \propto \chi_2^2 \quad (3.56)$$

Scaling, we obtain that at time  $t$ :

$$\frac{M^2(t)}{\tilde{s}_t^2} \sim \chi_2^2 \quad \forall t. \quad (3.57)$$

### 3.4.4.2 Distribution of Numerator of $\psi(t; f)$

Recall that  $U(t) \stackrel{d}{=} Q(t)$ , and  $\dot{U}(t) \stackrel{d}{=} \dot{Q}(t)$ . As well we proved that  $U(t)$  and  $\dot{Q}(t)$  and  $\dot{U}(t)$  and  $Q(t)$  are uncorrelated, respectively. Each term in the numerator is then a product of two uncorrelated normal random variables,  $U(t)\dot{Q}(t)$  and  $\dot{U}(t)Q(t)$ . Let  $A$  denote either product. Then  $A$  has the known PDF of a modified Bessel function of the second kind with zero mean and since  $\rho_{U\dot{Q}}(t) = \rho_{\dot{U}Q}(t) = 0 = \rho_0$  for the pairs  $U(t), \dot{Q}(t)$  and  $\dot{U}(t), Q(t)$  we obtain:

$$f(a) = \frac{1}{\pi\sigma_1\sigma_2\sqrt{1-\rho_0^2}} e^{\left(\frac{\rho_0 a}{\sigma_1\sigma_2(1-\rho_0^2)}\right)} K_0\left(\frac{|a|}{\sigma_1\sigma_2(1-\rho_0^2)}\right) \quad (3.58)$$

$$= \frac{1}{\pi\sigma_1\sigma_2} K_0\left(\frac{|a|}{\sigma_1\sigma_2}\right), \quad a \in \mathbb{R} \quad (3.59)$$

where  $K_0$  is the 0<sup>th</sup> order modified Bessel function of the second kind and  $\sigma_1 = \tilde{\sigma}_t$  and  $\sigma_2 = \dot{\tilde{\sigma}}_t$  [16]. If we standardize:

$$Z_1(t) = \frac{U(t)}{\tilde{\sigma}_t}, \quad Z_2(t) = \frac{\dot{U}(t)}{\dot{\tilde{\sigma}}_t}, \quad Z_3(t) = \frac{Q(t)}{\tilde{\sigma}_t}, \quad Z_4(t) = \frac{\dot{Q}(t)}{\dot{\tilde{\sigma}}_t} \quad (3.60)$$

where  $Z_1, \dots, Z_4$  become standard Gaussian variables, then for  $A = Z_1Z_4$  or  $A = Z_2Z_3$  Eqn. (3.59) becomes

$$f(a) = \frac{1}{\pi} K_0(|a|), \quad a \in \mathbb{R} \quad (3.61)$$

which is a variance gamma distribution with parameters  $\text{VG}(0,1,0,1/2,1)$ . We can conclude that

$$\frac{Z_1Z_4 - Z_2Z_3}{\sqrt{1-\rho_t^2}} \sim \mathbf{Laplace}(0, 1) \quad (3.62)$$

where  $\rho_{U\dot{U}}(t) = \rho_{Q\dot{Q}}(t) = \rho_t$  is the correlation at time  $t$  between  $U(t)$  and  $\dot{U}(t)$  or  $Q(t)$  and  $\dot{Q}(t)$  using Proposition 2.2.5 of [16]. Computing  $\rho_t$  we have

$$\rho_t := \frac{\mathbb{E}[U(t)\dot{U}(t)]}{\tilde{s}_t \dot{\tilde{s}}_t} = \frac{\mathbb{E}[Q(t)\dot{Q}(t)]}{\tilde{s}_t \dot{\tilde{s}}_t} \quad (3.63)$$

$$= \frac{\sum_{k=0}^{K-1} v_t^{(k)} \dot{v}_t^{(k)} \mathbb{E}[y_k^R(f)y_k^R(f)]}{\tilde{s}_t \dot{\tilde{s}}_t}; \text{ from Eqn.(3.50)} \quad (3.64)$$

$$= \frac{\sum_{k=0}^{K-1} v_t^{(k)} \dot{v}_t^{(k)} \frac{1}{2} w_k^2 S(f)}{\tilde{s}_t \dot{\tilde{s}}_t} \quad (3.65)$$

$$= \frac{\sum_{k=0}^{K-1} v_t^{(k)} \dot{v}_t^{(k)} \frac{1}{2} w_k^2 S(f)}{\frac{\sqrt{\tilde{\sigma}_t^2 S(f)} \sqrt{\dot{\tilde{\sigma}}_t^2 S(f)}}{\sqrt{4}}} \quad (3.66)$$

$$= \frac{\sum_{k=0}^{K-1} v_t^{(k)} \dot{v}_t^{(k)} \frac{1}{2} w_k^2 S(f)}{\frac{\tilde{\sigma}_t \dot{\tilde{\sigma}}_t S(f)}{2}} \quad (3.67)$$

$$= \frac{\sum_{k=0}^{K-1} v_t^{(k)} \dot{v}_t^{(k)} w_k^2}{\tilde{\sigma}_t \dot{\tilde{\sigma}}_t} \quad (3.68)$$

Therefore,

$$U(t)\dot{Q}(t) - \dot{U}(t)Q(t) \propto \mathbf{Laplace}(0, 1) \quad (3.69)$$

with proportionality factor

$$\frac{1}{\sqrt{\frac{\tilde{\sigma}_t^2 S(f)}{2}} \sqrt{\frac{\dot{\tilde{\sigma}}_t^2 S(f)}{2}} \sqrt{1 - \rho_t^2}} = \frac{1}{\frac{\tilde{\sigma}_t \dot{\tilde{\sigma}}_t S(f)}{2} \sqrt{1 - \left( \frac{\sum_{k=0}^{K-1} v_t^{(k)} \dot{v}_t^{(k)} w_k^2}{\tilde{\sigma}_t \dot{\tilde{\sigma}}_t} \right)^2}} \quad (3.70)$$

$$= \frac{1}{\frac{S(f)}{2} \sqrt{\tilde{\sigma}_t \dot{\tilde{\sigma}}_t - \left( \sum_{k=0}^{K-1} v_t^{(k)} \dot{v}_t^{(k)} w_k^2 \right)^2}}. \quad (3.71)$$

### 3.4.4.3 Distribution of $\psi(t; f)$

Using the two sections above, we have determined that

$$\psi(t; f) \propto \frac{\mathbf{Laplace}(0, 1)}{\chi_2^2} \quad (3.72)$$

with proportionality factor  $\alpha$  found by combining equations (3.6), (3.53) and (3.71),

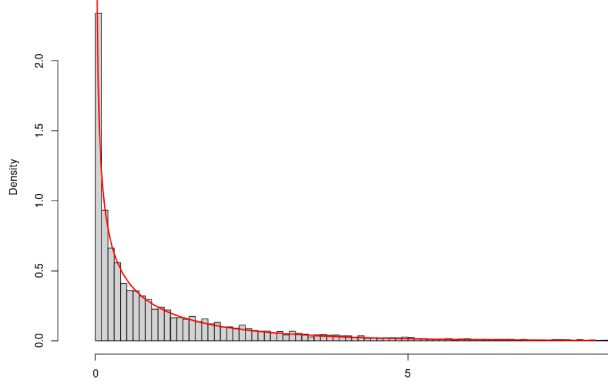


Figure 3.9: Theoretical density of the  $F_{1,K-P}$  distribution in red overlaid on the simulated density of  $F_4$  under the null hypothesis, when the process was Gaussian white noise. The number of simulations was 10,000.

which gives us:

$$\alpha = \frac{\frac{2\pi\tilde{\sigma}_t^2 S(f)}{2}}{\frac{S(f)}{2} \sqrt{\tilde{\sigma}_t \dot{\sigma}_t - \left( \sum_{k=0}^{K-1} v_t^{(k)} \dot{v}_t^{(k)} w_k^2 \right)^2}} \quad (3.73)$$

$$= \frac{2\pi\tilde{\sigma}_t^2}{\sqrt{\tilde{\sigma}_t \dot{\sigma}_t - \left( \sum_{k=0}^{K-1} v_t^{(k)} \dot{v}_t^{(k)} w_k^2 \right)^2}} \quad (3.74)$$

Note here that we have obtained the identical distribution for  $\psi(t; f)$  found in [3] except for the scaling factor. Therefore, as the weighting has been absorbed into the scaling factor  $\alpha$ ;  $\hat{c}_{P/0}$ ,  $\|r_{P/0}\|^2$ , and  $\Psi(f)$  will have the same distributions found in [1] under the assumptions stated in Section 3.4, provided that  $\psi(t; f)$  can be scaled so it has the same distribution at each time  $t$ , allowing  $\psi(t; f)$  to remain well-behaved. Then asymptotically,  $F_4$  still converges to an  $F$  distribution with degrees of freedom 1 and  $K - P$ . This is further illustrated by the simulation shown in Fig. 3.9, where the theoretical density of the  $F_{1,K-P}$  distribution agrees well with a histogram density estimate of the distribution of  $F_4(P; f, K)$ , under the assumption of the time series being Gaussian white noise.

As an example of the effectiveness, we applied  $F_4$  to the identical series shown in Fig. 3.3. We can see in Fig. 3.10 that the center frequency is now successfully

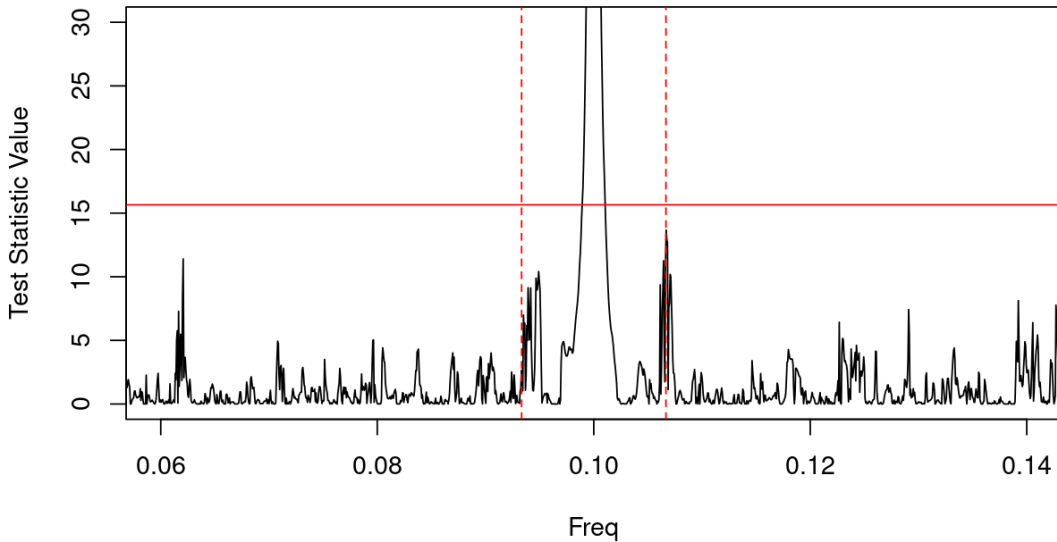


Figure 3.10:  $F_4$  applied to FM series (carrier frequency  $f = 0.1$ ) with input parameter  $K = 79$ . Dotted red lines represent  $f \pm \frac{(K+1)}{2N}$ , and the horizontal red line is the quantile of an  $F_{1,79-1}$  corresponding to a significance level of  $\frac{1}{N}$  with  $N = 6000$ . Also,  $\phi(\tau) = \frac{0.002}{2} (t - \frac{N}{2})$ , and  $\{Z_t\}$  is Gaussian white noise with a low signal-to-noise ratio.

detected, and the frequencies found at the edges of the band are reduced below the threshold for rejection of the null hypothesis.

Furthermore, using a lower number of tapers,  $K = 20$ , as shown in Fig. 3.11 across all the frequencies tested,  $F_4$  only has one false detection away from the carrier, whereas  $\tilde{F}_3$  falsely detects multiple other frequencies. Therefore,  $F_4$  seems to not only be effective at reducing side spikes, but may also help mitigate false detections away from the carrier.

### 3.5 $F'_4$ Test Statistic

Now that we have discussed a way to mitigate the spikes found at the band's edges, we can discuss the even and odd behaviour of the  $\tilde{F}_3$  and  $F_4$  test statistics. Unfortunately,

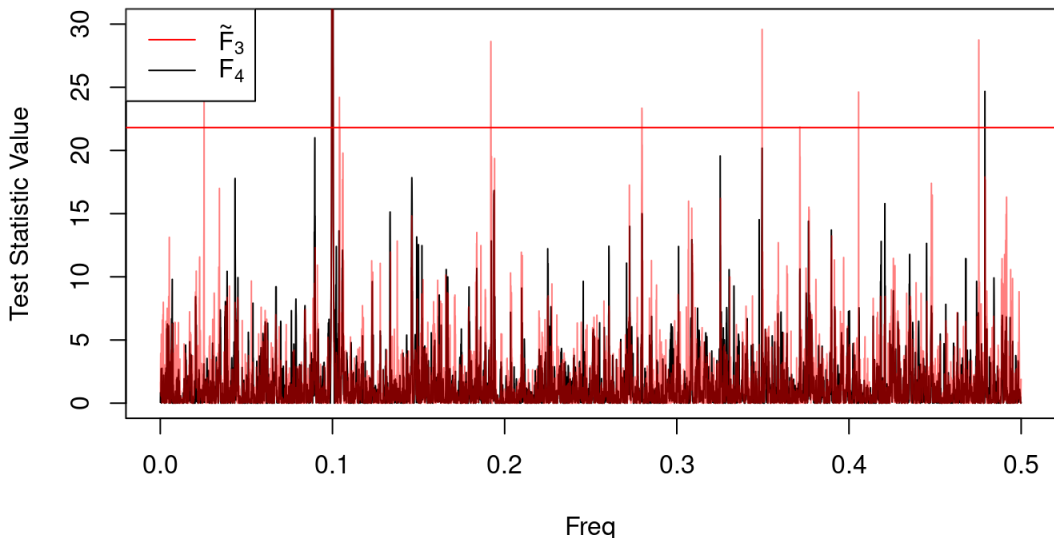


Figure 3.11:  $\tilde{F}_3$  in red applied to FM series under background SNR = 0.8 Gaussian white noise with input parameter  $K = 20$  across all frequencies.  $F_4$  in black applied to the same FM series with the same input parameter. The horizontal red line is the quantile of an  $F_{1,19}$  corresponding to a significance level of  $\frac{1}{N}$  where  $N = 6000$ .

looking at Fig. 3.12 we see that the even/odd behaviour still happens with the  $F_4$  test across the set of  $K$ 's as the probability of detection between any two consecutive  $K$ 's can vary immensely. This was very interesting as it confirmed the behaviour was not a property of only the DPSSs, but was more of an intrinsic property of both the  $F$ -tests. The Sine tapers have a very interesting property that when using a fixed  $N$ , for any two choices of  $K$  and  $K'$ , where  $K' \geq K$ , the first  $K$  columns of both taper matrices  $V$  will be identical. Therefore, the only difference between any consecutive even and odd choices of  $K$  for an identical time series,  $\{X_t\}$ , is the last column of the taper matrix. With this in mind, we aim to isolate the contribution due to the last  $k^{\text{th}}$  taper for any choice of  $K$ .

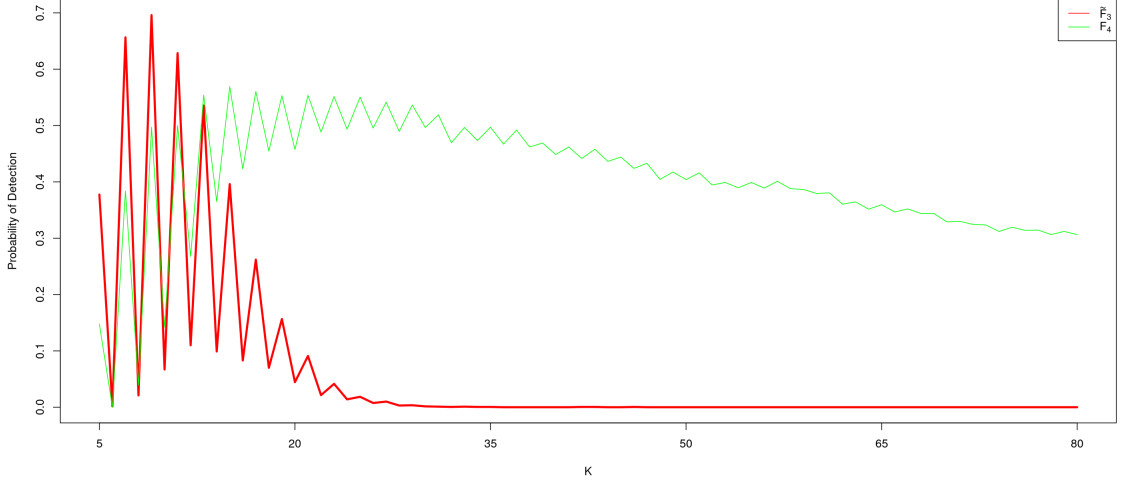


Figure 3.12: Probability of detection over 2000 tests. The  $\tilde{F}_3$  in red is applied to the FM series under background SNR = 0.5 Gaussian white noise with input parameter  $K \in \{5, \dots, 80\}$  at carrier frequency. The  $F_4$  in green is applied to the same FM series with the same input parameters. The 2000 tests were all conducted at a confidence level of  $\alpha = \frac{1}{N}$ .

### 3.5.1 Isolation of $K^{\text{th}}$ Taper

For the sake of clarity, I will define all vectors and matrices again.

#### Assumptions

In the following, we assume  $\phi(\tau) = \frac{m_P}{2}(t - \frac{N}{2})$ . That is, that we have a signal with linear modulation ( $P = 1$ ), with modulation width  $m_P$ .

#### 3.5.1.1 Eigencoefficients

As a reminder,

$$\tilde{Y}_f := \begin{pmatrix} y_0(f) \\ y_1(f) \\ \vdots \\ y_{(K-1)}(f) \end{pmatrix}, X_f := \begin{pmatrix} X_0 e^{-i2\pi f 0} \\ X_1 e^{-i2\pi f 1} \\ \vdots \\ X_{N-1} e^{-i2\pi f (N-1)} \end{pmatrix}, V := \begin{pmatrix} v_0^{(0)} & v_0^{(1)} & \dots & v_0^{(K-1)} \\ \vdots & \ddots & & \vdots \\ v_{N-1}^{(0)} & \dots & \dots & v_{N-1}^{(K-1)} \end{pmatrix}$$

$$\dot{V} := \begin{pmatrix} \dot{v}_0^{(0)} & \dot{v}_0^{(1)} & \cdots & \dot{v}_0^{(K-1)} \\ \vdots & \ddots & & \vdots \\ \dot{v}_{N-1}^{(0)} & \cdots & \cdots & \dot{v}_{N-1}^{(K-1)} \end{pmatrix}, \text{ and } \mathbf{W} := \begin{pmatrix} w_0 & 0 & \cdots & 0 \\ 0 & w_1 & \cdots & 0 \\ 0 & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & w_{K-1} \end{pmatrix}.$$

In this,  $V$  is the matrix of arbitrary tapers used to form the eigencoefficients and  $\dot{V}$  is created by evaluating the time derivative of the tapers. The  $K \times 1$  vector of eigencoefficients is then given by:

$$\tilde{Y}_f = \mathbf{W}V^T X_f. \quad (3.75)$$

Writing the elements of  $\tilde{Y}_f$ ,

$$\begin{aligned} y_k(f) &= w_k \sum_{n=0}^{N-1} v_n^{(k)} X_n e^{-i2\pi f n} \\ &= w_k \sum_{n=0}^{N-1} v_n^{(k)} X_n [\cos(2\pi f n) - i \sin(2\pi f n)] \\ &= w_k \sum_{n=0}^{N-1} v_n^{(k)} X_n \cos(2\pi f n) + i(-1)w_k \sum_{n=0}^{N-1} v_n^{(k)} X_n \sin(2\pi f n) \end{aligned} \quad (3.76)$$

Next, we split the real and imaginary parts of  $y_k(f)$ :

$$\Re(y_k(f)) = w_k \sum_{n=0}^{N-1} v_n^{(k)} X_n \cos(2\pi f n) \quad (3.77)$$

$$\Im(y_k(f)) = w_k(-1) \sum_{n=0}^{N-1} v_n^{(k)} X_n \sin(2\pi f n). \quad (3.78)$$

and write the set across the  $K$  tapers as:

$$\Re(\tilde{Y})_f = \begin{pmatrix} \Re(y_0(f)) \\ \Re(y_1(f)) \\ \vdots \\ \Re(y_{(K-1)}(f)) \end{pmatrix}, \Im(\tilde{Y})_f = \begin{pmatrix} \Im(y_0(f)) \\ \Im(y_1(f)) \\ \vdots \\ \Im(y_{(K-1)}(f)) \end{pmatrix}. \quad (3.79)$$

We now define the reduced versions by removing the last column. These reduced versions will be indicated with a prime.

$$V' := \begin{pmatrix} v_0^{(0)} & v_0^{(1)} & \cdots & v_0^{(K-2)} \\ \vdots & \ddots & & \vdots \\ v_{N-1}^{(0)} & \cdots & \cdots & v_{N-1}^{(K-2)} \end{pmatrix}, \dot{V}' := \begin{pmatrix} \dot{v}_0^{(0)} & \dot{v}_0^{(1)} & \cdots & \dot{v}_0^{(K-2)} \\ \vdots & \ddots & & \vdots \\ \dot{v}_{N-1}^{(0)} & \cdots & \cdots & \dot{v}_{N-1}^{(K-2)} \end{pmatrix}$$

The resulting set of  $K - 1$  eigencoefficients have the form:

$$\tilde{Y}'_f := \begin{pmatrix} y_0(f) \\ y_1(f) \\ \vdots \\ y_{(K-2)}(f) \end{pmatrix}.$$

The goal of this reduction is to allow for the splitting of contribution from the last taper from the first  $K - 1$  tapers.

### 3.5.1.2 Standard Inverse

Recall that the standard inverse vector (size  $N \times 1$ ) and standard inverse derivative vector (size  $N \times 1$ ) are given by:

$$Z = V\tilde{Y}'_f = U + iQ \tag{3.80}$$

$$\dot{Z} = \dot{V}\tilde{Y}'_f = \dot{U} + i\dot{Q}. \tag{3.81}$$

$U$ ,  $Q$ ,  $\dot{U}$ , and  $\dot{Q}$  can also be computed directly with

$$U = V\Re(\tilde{Y}'_f) \tag{3.82}$$

$$Q = V\Im(\tilde{Y}'_f) \tag{3.83}$$

$$\dot{U} = \dot{V}\Re(\tilde{Y}'_f) \tag{3.84}$$

$$\dot{Q} = \dot{V}\Im(\tilde{Y}'_f) \tag{3.85}$$

The reduced versions of  $U$ ,  $Q$ ,  $\dot{U}$ , and  $\dot{Q}$  are then given by

$$U' = V' \mathfrak{R}(\tilde{Y}'_f) \quad (3.86)$$

$$Q' = V' \mathfrak{I}(\tilde{Y}'_f) \quad (3.87)$$

$$\dot{U}' = \dot{V}' \mathfrak{R}(\tilde{Y}'_f) \quad (3.88)$$

$$\dot{Q}' = \dot{V}' \mathfrak{I}(\tilde{Y}'_f) \quad (3.89)$$

where each element of each of the vectors is at a time  $t$  and can be expanded as

$$U(t) = v_t^{(0)} w_0 \sum_{n=0}^{N-1} v_n^{(0)} X_n \cos(2\pi f n) + \cdots + v_t^{(K-1)} w_{K-1} \sum_{n=0}^{N-1} v_n^{(K-1)} X_n \cos(2\pi f n) \quad (3.90)$$

$$U'(t) = v_t^{(0)} w_0 \sum_{n=0}^{N-1} v_n^{(0)} X_n \cos(2\pi f n) + \cdots + v_t^{(K-2)} w_{K-2} \sum_{n=0}^{N-1} v_n^{(K-2)} X_n \cos(2\pi f n) \quad (3.91)$$

Therefore we can write  $U(t)$  in terms of  $U'(t)$ , with Eqn. (3.77), as

$$\begin{aligned} U(t) &= U'(t) + v_t^{(K-1)} w_{K-1} \sum_{n=0}^{N-1} v_n^{(K-1)} X_n \cos(2\pi f n) \\ &= U'(t) + v_t^{(K-1)} \mathfrak{R}(y_{(K-1)}) \end{aligned} \quad (3.92)$$

Repeating this derivation with  $\dot{U}(t)$ ,  $Q(t)$ , and  $\dot{Q}(t)$  we obtain

$$\begin{aligned} \dot{U}(t) &= \dot{U}'(t) + \dot{v}_t^{(K-1)} w_{K-1} \sum_{n=0}^{N-1} v_n^{(K-1)} X_n \cos(2\pi f n) \\ &= \dot{U}'(t) + \dot{v}_t^{(K-1)} \mathfrak{R}(y_{(K-1)}) \end{aligned} \quad (3.93)$$

$$\begin{aligned} Q(t) &= Q'(t) + v_t^{(K-1)} w_{K-1} (-1) \sum_{n=0}^{N-1} v_n^{(K-1)} X_n \sin(2\pi f n) \\ &= Q'(t) + v_t^{(K-1)} \mathfrak{I}(y_{(K-1)}) \end{aligned} \quad (3.94)$$

$$\begin{aligned} \dot{Q}(t) &= \dot{Q}'(t) + \dot{v}_t^{(K-1)} w_{K-1} (-1) \sum_{n=0}^{N-1} v_n^{(K-1)} X_n \sin(2\pi f n) \\ &= \dot{Q}'(t) + \dot{v}_t^{(K-1)} \mathfrak{I}(y_{(K-1)}). \end{aligned} \quad (3.95)$$

### 3.5.1.3 Instantaneous Frequency

Now computing the instantaneous frequency at time  $t$  and frequency  $f$  for the version with the final term:

$$\psi(t; f) := \frac{\psi_{num}(t)}{\psi_{den}(t)} = \frac{U(t)\dot{Q}(t) - \dot{U}(t)Q(t)}{2\pi(U(t)^2 + Q(t)^2)} \quad (3.96)$$

and for the reduced version, without the final term:

$$\psi'(t; f) := \frac{\psi'_{num}(t)}{\psi'_{den}(t)} = \frac{U'(t)\dot{Q}'(t) - \dot{U}'(t)Q'(t)}{2\pi(U'(t)^2 + Q'(t)^2)} \quad (3.97)$$

Writing  $\psi_{num}$  in terms of equations (3.92) through (3.95):

$$\begin{aligned} \psi_{num}(t) = & \left( U'(t) + v_t^{(K-1)} \Re(y_{(K-1)}) \right) \left( \dot{Q}'(t) + \dot{v}_t^{(K-1)} \Im(y_{(K-1)}) \right) \\ & - \left( \dot{U}'(t) + \dot{v}_t^{(K-1)} \Re(y_{(K-1)}) \right) \left( Q'(t) + v_t^{(K-1)} \Im(y_{(K-1)}) \right) \end{aligned}$$

Expanded, this gives:

$$\begin{aligned} \psi_{num}(t) = & (U'(t)\dot{Q}'(t) - \dot{U}'(t)Q'(t)) \\ & + U'(t)\dot{v}_t^{(K-1)} \Im(y_{(K-1)}) + \dot{Q}'(t)v_t^{(K-1)} \Re(y_{(K-1)}) \\ & - \dot{U}'(t)v_t^{(K-1)} \Im(y_{(K-1)}) - Q'(t)\dot{v}_t^{(K-1)} \Re(y_{(K-1)}) \quad (3.98) \\ & + v_t^{(K-1)} \Re(y_{(K-1)})\dot{v}_t^{(K-1)} \Im(y_{(K-1)}) \\ & - \dot{v}_t^{(K-1)} \Re(y_{(K-1)})v_t^{(K-1)} \Im(y_{(K-1)}) \end{aligned}$$

Notice here that the first term shown in brackets in Eqn. (3.98) is  $\psi'_{num}$  in Eqn. (3.97); thus, all other terms in the expansion are contributions from the last ( $k = K - 1$ ) taper. Also, notice both elements on the last two lines of Eqn. (3.97) cancel each other. Letting

$$A(t) := \dot{Q}'(t)\Re(y_{(K-1)}) - \dot{U}'(t)\Im(y_{(K-1)}) \quad (3.99)$$

$$B(t) := U'(t)\Im(y_{(K-1)}) - Q'(t)\Re(y_{(K-1)}), \quad (3.100)$$

we can rewrite  $\psi_{num}$  by combining the above as

$$\psi_{num}(t) = \psi'_{num}(t) + \left( v_t^{(K-1)} A(t) + \dot{v}_t^{(K-1)} B(t) \right). \quad (3.101)$$

Then, writing  $\psi_{den}$  in terms of equations (3.92) through (3.95).

$$\psi_{den}(t) = 2\pi \left( U^2(t) + Q^2(t) \right) \quad (3.102)$$

$$= 2\pi \left[ \left( U'(t) + v_t^{(K-1)} \Re(y_{(K-1)}) \right)^2 + \left( Q'(t) + v_t^{(K-1)} \Im(y_{(K-1)}) \right)^2 \right]. \quad (3.103)$$

Expanding out Eqn. (3.103), if we let

$$C(t) := 2\pi \left[ 2U'(t)v_t^{(K-1)} \Re(y_{(K-1)}) + 2Q'(t)v_t^{(K-1)} \Im(y_{(K-1)}) \right. \\ \left. + (v_t^{(K-1)} \Re(y_{(K-1)}))^2 + (v_t^{(K-1)} \Im(y_{(K-1)}))^2 \right] \quad (3.104)$$

$$= 2\pi v_t^{(K-1)} \left[ 2U'(t) \Re(y_{(K-1)}) + 2Q'(t) \Im(y_{(K-1)}) \right. \\ \left. + v_t^{(K-1)} \Re(y_{(K-1)})^2 + v_t^{(K-1)} \Im(y_{(K-1)})^2 \right] \quad (3.105)$$

$$= 2\pi v_t^{(K-1)} G(t) \quad (3.106)$$

where  $G(t)$  is the expanded quantity in the square brackets, then  $\psi_{den}(t)$  can be written as a sum of  $\psi'_{den}(t)$ , and  $C(t)$  to obtain

$$\psi_{den}(t) = \psi'_{den}(t) + C(t) = \psi'_{den}(t) + 2\pi v_t^{(K-1)} G(t). \quad (3.107)$$

Now we can write  $\psi(t; f)$  in terms of the reduced version by combining equations (3.96), (3.101), and (3.107):

$$\psi(t; f) = \frac{\psi'_{num}(t) + \left( v_t^{(K-1)} A(t) + \dot{v}_t^{(K-1)} B(t) \right)}{\psi'_{den}(t) + 2\pi v_t^{(K-1)} G(t)}. \quad (3.108)$$

Rewriting in an alternate form, we have

$$\psi(t; f) = \psi'(t; f) + \psi_{\Delta}(t; f). \quad (3.109)$$

where

$$\psi_{\Delta}(t; f) := \frac{\left( v_t^{(K-1)} A(t) + \dot{v}_t^{(K-1)} B(t) \right) \psi'_{den}(t) - 2\pi v_t^{(K-1)} G(t) \psi'_{num}(t)}{\psi'_{den}(t)^2 + 2\pi v_t^{(K-1)} G(t) \psi'_{den}(t)}. \quad (3.110)$$

In this expression,  $\psi_{\Delta}(t; f)$  is the total contribution from only the  $K - 1^{\text{st}}$  taper to the instantaneous frequency at time  $t$  and frequency  $f$ .

### 3.5.1.4 Eigencoefficients of the Instantaneous Frequency

Recall the full eigencoefficient vector of the instantaneous frequency is given by

$$\mathbf{\Psi}(f) := V^T \boldsymbol{\psi}(f) = \begin{pmatrix} \Psi_0(f) \\ \Psi_1(f) \\ \vdots \\ \Psi_{(K-1)}(f) \end{pmatrix} \quad (3.111)$$

where each element can be written using Eqn. (3.109) as:

$$\Psi_k(f) = \sum_{n=0}^{N-1} v_n^{(k)} \psi(n; f) \quad (3.112)$$

$$= \sum_{n=0}^{N-1} v_n^{(k)} \left( \psi'(n; f) + \psi_{\Delta}(n; f) \right) \quad (3.113)$$

$$= \sum_{n=0}^{N-1} v_n^{(k)} \psi'(n; f) + \sum_{n=0}^{N-1} v_n^{(k)} \psi_{\Delta}(n; f) \quad (3.114)$$

$$= \Psi'_k(f) + \Psi_k^d(f) \quad (3.115)$$

with

$$\Psi'_k(f) := \sum_{n=0}^{N-1} v_n^{(k)} \psi'(n; f) \quad (3.116)$$

$$\Psi_k^d(f) := \sum_{n=0}^{N-1} v_n^{(k)} \psi_{\Delta}(n; f). \quad (3.117)$$

Lastly, we define some vectors that will be useful in the next section:

$$\tilde{\boldsymbol{\Psi}}(f) := \begin{pmatrix} \Psi'_0(f) \\ \vdots \\ \Psi'_{(K-2)}(f) \\ \Psi'_{(K-1)}(f) \end{pmatrix}, \quad \boldsymbol{\Psi}'(f) := \begin{pmatrix} \Psi'_0(f) \\ \vdots \\ \Psi'_{(K-2)}(f) \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Psi}_{\Delta}(f) := \begin{pmatrix} \Psi_0^d(f) \\ \vdots \\ \Psi_{(K-1)}^d(f) \end{pmatrix}. \quad (3.118)$$

Notice here that these vectors were constructed in a way such that the expression

$$\mathbf{\Psi}(f) = \tilde{\boldsymbol{\Psi}}(f) + \boldsymbol{\Psi}_{\Delta}(f) \quad (3.119)$$

is valid, by equations (3.115) and (3.118).

### 3.5.1.5 $\hat{c}_{P/0}$ and $H$

Start by defining  $H := H_{P/0}$  as a normalized vector of length  $K$  since  $P = 1$ , where the (1<sup>st</sup>, 3<sup>rd</sup>, 5<sup>th</sup>, ...) elements are all zero. Define  $H' := H'_{P/0}$  in the same way, but having only  $K - 1$  elements. That is:

$$H := \begin{pmatrix} H_0 \\ \vdots \\ H_{(K-1)} \end{pmatrix}, \quad H_{/(K-1)} := \begin{pmatrix} H_0 \\ \vdots \\ H_{(K-2)} \end{pmatrix}, \quad \text{and } H' := \begin{pmatrix} H'_0 \\ \vdots \\ H'_{(K-2)} \end{pmatrix}. \quad (3.120)$$

Recall Eqn. (3.25) and its definition of  $\hat{c}$ . Given the above, we can now compute  $\hat{c} := \hat{c}_{P/0}$  and relate it to the reduced set of tapers with Eqn. (3.118).

$$\hat{c}(f) := H^T \Psi(f) \quad (3.121)$$

$$= H^T \left( \tilde{\Psi}(f) + \Psi_{\Delta}(f) \right) \quad (3.122)$$

$$= H^T \tilde{\Psi}(f) + H^T \Psi_{\Delta}(f) \quad (3.123)$$

$$= H_{/(K-1)}^T \Psi'(f) + H_{(K-1)} \Psi'_{(K-1)} + H^T \Psi_{\Delta}(f). \quad (3.124)$$

The reduced  $\hat{c}$ , denoted  $\hat{c}'$ , can be computed easily with

$$\hat{c}'(f) := (H')^T \Psi'(f), \quad (3.125)$$

by taking the difference of equations (3.124) and (3.125):

$$\hat{c}_{\Delta}(f) := \hat{c}(f) - \hat{c}'(f) \quad (3.126)$$

$$= H_{/(K-1)}^T \Psi'(f) + H_{(K-1)} \Psi'_{(K-1)} + H^T \Psi_{\Delta}(f) - (H')^T \Psi'(f) \quad (3.127)$$

$$= \left( H_{/(K-1)}^T - (H')^T \right) \Psi'(f) + H^T \Psi_{\Delta}(f) + H_{(K-1)} \Psi'_{(K-1)} \quad (3.128)$$

This allows us to write  $\hat{c}(f)$  as the reduced  $\hat{c}'$  plus the contribution only due to the last taper:

$$\hat{c}(f) = \hat{c}'(f) + \hat{c}_{\Delta}(f). \quad (3.129)$$

### 3.5.1.6 Norm of $\Psi(f)$

Using (3.118) and (3.119),  $\|\Psi(f)\|^2$  can also be broken down:

$$\|\Psi(f)\|^2 = \left\| \left( \tilde{\Psi}(f) + \Psi_{\Delta}(f) \right) \right\|^2 \quad (3.130)$$

$$= \left( \tilde{\Psi}(f) + \Psi_{\Delta}(f) \right)^T \left( \tilde{\Psi}(f) + \Psi_{\Delta}(f) \right) \quad (3.131)$$

$$= \tilde{\Psi}^T(f) \tilde{\Psi}(f) + \Psi_{\Delta}^T(f) \Psi_{\Delta}(f) + \Psi_{\Delta}^T(f) \tilde{\Psi}(f) + \tilde{\Psi}^T(f) \Psi_{\Delta}(f) \quad (3.132)$$

$$= \left[ (\Psi'(f))^T \Psi'(f) + \Psi'_{(K-1)} \Psi'_{(K-1)} \right] + \|\Psi_{\Delta}(f)\|^2 \quad (3.133)$$

$$+ \sum_{k=0}^{K-1} \Psi'_k(f) \Psi_k^d(f) + \sum_{k=0}^{K-1} \Psi_k^d(f) \Psi'_k(f)$$

$$= \|\Psi'(f)\|^2 + \|\Psi_{\Delta}(f)\|^2 + 2 \sum_{k=0}^{K-1} \Psi'_k(f) \Psi_k^d(f) + \Psi'_{(K-1)} \Psi'_{(K-1)} \quad (3.134)$$

This now gives us all of the components necessary to develop an expression for the contributions of the final term, based on the  $K - 1$  taper.

### 3.5.1.7 Decomposing the $F_4$ Test Statistic

Now that we have all the pieces, let's start with  $F_4(f)$  for  $P = 1$  and put it all together.

$$F_4(1; f, K) := \frac{(\hat{c}(f))^2 (K - 1)}{\|\Psi(f)\|^2 - \|\hat{c}(f)\|^2} \quad (3.135)$$

$$= \frac{\hat{c}^2(f) (K - 1)}{\|\Psi(f)\|^2 - \hat{c}^2(f)} \quad (3.136)$$

$$= \frac{(K - 1)}{\frac{\|\Psi(f)\|^2}{\hat{c}^2(f)} - 1} \quad (3.137)$$

Then expanding the ratio  $\frac{\|\Psi(f)\|^2}{\hat{c}^2(f)}$  using equations (3.129), and (3.134):

$$\frac{\|\Psi(f)\|^2}{\hat{c}^2(f)} = \frac{\|\Psi'(f)\|^2 + \|\Psi_{\Delta}(f)\|^2 + 2 \sum_{k=0}^{K-1} \Psi'_k(f) \Psi_k^d(f) + \Psi'_{(K-1)} \Psi'_{(K-1)}}{(\hat{c}'(f) + \hat{c}_{\Delta}(f))^2} \quad (3.138)$$

$$= \frac{\|\Psi'(f)\|^2 + \|\Psi_{\Delta}(f)\|^2 + 2 \sum_{k=0}^{K-1} \Psi'_k(f) \Psi_k^d(f) + \Psi'_{(K-1)} \Psi'_{(K-1)}}{(\hat{c}'(f))^2 + 2\hat{c}'(f)\hat{c}_{\Delta}(f) + (\hat{c}_{\Delta}(f))^2}. \quad (3.139)$$

The difference between the reduced test ratio and the full test ratio is given by

$$r_\Delta := r - r' := \frac{\|\Psi(f)\|^2}{\hat{c}^2(f)} - \frac{\|\Psi'(f)\|^2}{(\hat{c}'(f))^2}. \quad (3.140)$$

Substituting in 3.139,

$$r_\Delta = \frac{\left(\|\Psi_\Delta(f)\|^2 + 2 \sum_{k=0}^{K-1} \Psi'_k(f) \Psi_k^d(f) + \Psi'_{(K-1)} \Psi'_{(K-1)}\right) - r' (2\hat{c}'(f)\hat{c}_\Delta(f) + (\hat{c}_\Delta(f))^2)}{(\hat{c}'(f))^2 + 2\hat{c}'(f)\hat{c}_\Delta(f) + (\hat{c}_\Delta(f))^2}, \quad (3.141)$$

Now substituting Eqn. (3.140) into 3.137, we have

$$\begin{aligned} F_4(1; f, K) &= \frac{(K-1)}{r-1} \\ &= \frac{(K-1)}{(r' + r_\Delta) - 1} \end{aligned} \quad (3.142)$$

$$= \frac{(K-1)}{(r' - 1) + r_\Delta} \quad (3.143)$$

The difference between Eqn. (3.142) and the reduced test can then be finally defined as:

$$\begin{aligned} F_\Delta(1; f, K) &:= F_4(1; f, K) - F'_4(1; f, K) \\ &= \frac{(K-1)}{(r' + r_\Delta) - 1} - \frac{(K-1)}{r' - 1} \\ &= \frac{r_\Delta(1-K)}{(r' + r_\Delta - 1)(r' - 1)} \end{aligned} \quad (3.144)$$

Therefore,

$$F_4(1; f, K) = F'_4(1; f, K) + \frac{r_\Delta(1-K)}{(r' + r_\Delta - 1)(r' - 1)} \quad (3.145)$$

There are two ways of representing  $F_4(1; f, K)$ . The more interpretable is Eqn. (3.143), as this makes it clear that  $r_\Delta$  in the denominator can reduce  $F_4(1; f, K)$  if it is positive.

### 3.5.2 Distribution of $F'_4(f, K)$

First, recall the arbitrary case of  $Y = a + bX$  where  $b > 0$ , and  $X$  has known PDF of  $f_X(x)$  and CDF  $F_X(x)$ . The CDF of  $Y$  can be evaluated by:

$$\begin{aligned} F_Y(x) &= P(Y \leq x) \\ &= P(a + bX \leq x) \\ &= P\left(X \leq \frac{x-a}{b}\right) \end{aligned} \tag{3.146}$$

so,

$$F_Y(x) = F_X\left(\frac{x-a}{b}\right) \tag{3.147}$$

The PDF can be easily found:

$$f_Y(x) = F'_Y(x) = \frac{d}{dx} \left[ F_X\left(\frac{x-a}{b}\right) \right],$$

and thus

$$f_Y(x) = \frac{1}{b} f_X\left(\frac{x-a}{b}\right). \tag{3.148}$$

Now, to find the distribution for the  $F'_4$  test. First, let  $\tilde{K}$  be an arbitrary odd number of tapers and conduct the  $F_4(1; f, \tilde{K})$  test with weight matrix  $\tilde{w} = \text{diag}(\tilde{w}_0, \dots, \tilde{w}_{\tilde{K}-1})$ .

This will have the form:

$$F_4(1; f, \tilde{K}) = \frac{\tilde{K} - 1}{(\tilde{r}' + \tilde{r}_\Delta) - 1}, \tag{3.149}$$

and if we define

$$X := \frac{\tilde{K} - 1}{(\tilde{r}' + \tilde{r}_\Delta) - 1} \sim F(x, 1, \tilde{K} - 1). \tag{3.150}$$

Now let  $K = \tilde{K} + 1$ . Conducting the weighted  $F$ -test with parameter  $K$ , the weight matrix is defined as  $w = \text{diag}(\tilde{w}_0, \dots, \tilde{w}_{\tilde{K}-1}, w_{K-1})$ , and

$$F_4(1; f, K) = \frac{K - 1}{(r' + r_\Delta) - 1}. \tag{3.151}$$

The respective  $F'_4$  test for Eqn. (3.151) would be given by

$$Y := F'_4(1; f, K) = \frac{K-1}{r'-1} \quad (3.152)$$

Due to the use of Sine Tapers, the first  $K-1$  tapers are identical to the tapers used in Eqn. (3.149). In addition, by construction, the first  $K-1$  diagonal elements of the weight matrices are equivalent as well. Due to both of these facts,

$$r' = \tilde{r}' + \tilde{r}_\Delta. \quad (3.153)$$

We can use this to rewrite (3.152) in terms of Eqn. (3.150) as

$$\begin{aligned} F'_4(1; f, K) &= \frac{K-1}{r'-1} \\ &= \frac{K-1}{\tilde{r}' + \tilde{r}_\Delta - 1} \\ &= \frac{K-1}{\tilde{K}-1} \left( \frac{\tilde{K}-1}{\tilde{r}' + \tilde{r}_\Delta - 1} \right) \\ Y &= \frac{K-1}{\tilde{K}-1} X = bX, \end{aligned} \quad (3.154)$$

for  $b = \frac{K-1}{\tilde{K}-1}$ . Then using Eqn. (3.148), and the known distribution of  $X$  in Eqn. (3.150)

we can conclude that the prime variant of the test has PDF

$$f_Y(x; d_1, d_2) = \frac{1}{b} \frac{\sqrt{\frac{(d_1 \frac{x}{b})^{d_1} d_2^{d_2}}{(d_1 \frac{x}{b} + d_2)^{d_1 + d_2}}}}{\frac{x}{b} B\left(\frac{d_1}{2}, \frac{d_2}{2}\right)}$$

where  $d_1 = 1, d_2 = \tilde{K} - P$ ; as we know that  $X$  is  $F$ -distributed with PDF

$$f_X(x; d_1, d_2) = \frac{\sqrt{\frac{(d_1 x)^{d_1} d_2^{d_2}}{(d_1 x + d_2)^{d_1 + d_2}}}}{xB\left(\frac{d_1}{2}, \frac{d_2}{2}\right)}.$$

In practice, the first  $K-1$  weights are not identical, but they become equal as  $K \rightarrow \infty$ .

Therefore we assume that the weights are close enough for consecutive  $K$ 's that this distribution still holds.

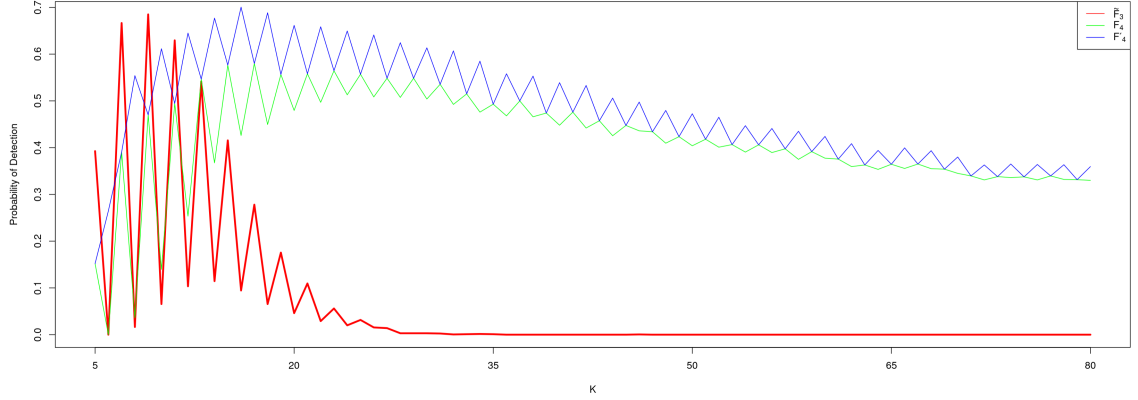


Figure 3.13:  $\tilde{F}_3$  in red applied to FM series under background SNR = 0.5 Gaussian white noise with input parameters  $K \in \{5, \dots, 80\}$  at the carrier frequency.  $F_4$  in green and  $F'_4$  in blue applied to the same FM series with the same input parameters. The 2000 tests were all conducted with  $\alpha = \frac{1}{N}$ .

We can now define the new prime test as

$$F'_4(P; f, K) = \begin{cases} \frac{K-1}{r'-1}; & K \text{ is even} \\ \frac{K-1}{(r'+r_\Delta)-1}; & \text{else} \end{cases} \quad (3.155)$$

which has a PDF under the null hypothesis of

$$f_{F'_4(P;f,K)}(x) = \begin{cases} f_y(x, 1, K-1-P); & K \text{ is even} \\ f_x(x, 1, K-P); & \text{else.} \end{cases} \quad (3.156)$$

Updating Fig. 3.12 to form Fig. 3.13, we see that the  $F'_4$  test doesn't affect the outcome of odd  $K$  selections. Still, under the case of testing even  $K$ s, the probability of detection is improved and surpasses the detection probability for odd  $K$  in most cases. Although it is not shown here, one can use the DPSS tapers with  $F'_4$  and the result of increased probability of detection for even  $K$  is consistent with the sinusoidal set of tapers, but without the added benefit of the weights to aid with the detection probability.

## 4. Aggregate Test

Although the choice of  $K$  for the  $F'_4$  test is not as crucial as for the  $F_4$  or the  $\tilde{F}_3$  tests, based on Figure 3.13, this is still a point where there are “analyst degrees-of-freedom”, and there are clearly better choices of  $K$  than others. It would thus be better if we could create a test in which we pool information across a set of  $K$  values. In doing so, the test should have a more consistent probability of detection for a wider range of modulation widths, as well as allow for the creation of a standard range of  $K$  that any analyst could use, regardless of the use case. We define our proposed solution to this problem, an *Aggregate test statistic*, denoted by  $T_a$ , as follows.

Let  $\mathcal{K} := \{K_1, \dots, K_{|\mathcal{K}|}\}$ , where each  $K_j \in \mathcal{K}$  is a specified number of tapers that would be used in the  $F'_4$ -test corresponding to the  $j^{\text{th}}$  choice of tapers and bandwidth. The test statistic  $T_a$  is then given by

$$T_a := \sum_{i=1}^{|\mathcal{K}|} \mathbb{1}_{(F'_4(P;f,K_i) \geq F_i)}, \quad (4.1)$$

where the CDF used to find the quantile,  $F_i$ , corresponding to a given level of significance, say  $\beta$ , is given by regularized incomplete beta functions in the form:

$$\begin{cases} F(x) = \mathcal{I}_{\frac{x}{x+(K_i-P)}} \left( \frac{1}{2}, \frac{K_i-P}{2} \right) = 1 - \beta; & \text{if } K_i \text{ is Odd} \\ F(x) = \mathcal{I}_{\frac{\frac{x}{b}}{\frac{x}{b}+(K_i-1-P)}} \left( \frac{1}{2}, \frac{K_i-1-P}{2} \right) = 1 - \beta; & \text{if } K_i \text{ is Even} \end{cases} \quad (4.2)$$

The parameter  $b$  is given by  $b = \frac{K-1}{K-2}$  from Section 3.5.2. The Aggregate test rejects the null hypothesis of there being no modulated signal at frequency  $f$  if  $T_a \geq R$ , where

$R$  is a specified threshold number of  $F'_4$  tests that must reject the null hypothesis.

## 4.1 Distribution Discussion

A challenge in applying the Aggregate test for a given threshold  $R$  is that, for a given desired level of significance  $\alpha$  for the Aggregate test (a typical choice is  $\alpha = \frac{1}{N}$ ), one must determine the level of significance  $\beta$  for each of the individual tests to end up with the desired overall level of significance for the Aggregate test. Notice that the Aggregate test statistic, Eqn. (4.1), can be considered a sum of Bernoulli random variables. Thus, if for any pair  $K_i, K_j$ , the two  $F'_4(P; f, K_i)$  and  $F'_4(P; f, K_j)$  are independent, then the Aggregate test would be distributed as a Binomial( $|\mathcal{K}|, \beta$ ), and we could then easily determine how to set  $\beta$  to give an overall size  $\alpha$  Aggregate test, for any given  $\alpha$ . Unfortunately, this is not the case. Letting  $(T_a)_i$  denote the  $i^{\text{th}}$  indicator in the sum in  $T_a$ , since each of the  $(T_a)_i$  tests are conducted on the same starting data, only differing in the number of tapers used in the test, there is reason to suppose that the  $(T_a)_i$  are correlated. This is further supported by Fig. 4.1 in which the correlations from 10,000 simulations of 70 unique choices of  $K$  were examined under Gaussian white noise. There is a positive correlation for pairs of  $K$  close to each other, e.g.,  $K_i = 49$  and  $K_j = 51$ , and a negative correlation for points further away, e.g.,  $K_i = 20$  and  $K_j = 80$ . With the correlation  $\text{Cor} [(T_a)_i, (T_a)_j] \neq 0$  for any  $i, j$  pair, the distribution of  $T_a$  is unknown, as it is the sum of non-independent Bernoulli( $\beta$ ) random variables. Using simulation, we can estimate the complementary CDF for any given value of  $|\mathcal{K}|$  and  $\beta$ . The simulated CDF for  $|\mathcal{K}| = 76$  and  $\beta = \frac{1}{2000}$  is shown in Fig. 4.2. Also shown in Figure 4.2 is the complementary CDF of the corresponding Binomial(76, 1/2000) distribution, which would be the distribution of  $T_a$  if the  $(T_a)_i$  were indeed independent. Using the independence assumption would lead to using a Type I error probability  $\beta$  for the individual tests that was far too

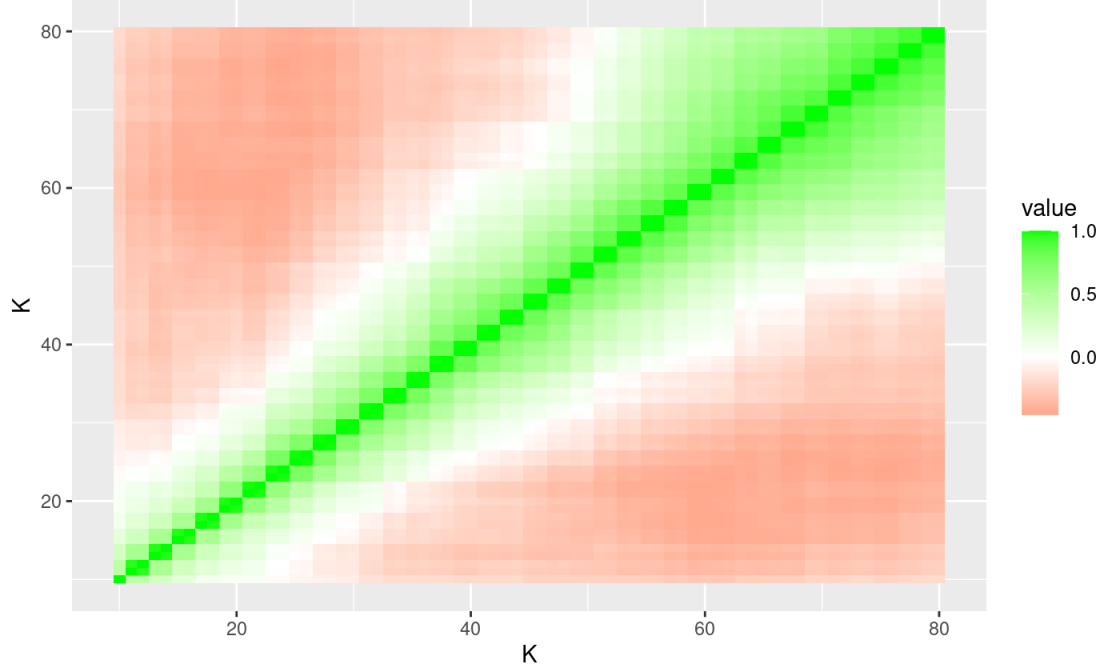


Figure 4.1: Estimated correlation between 70 pairs of  $(T_a)_i$  ranging from  $K = 10$  to  $K = 80$ . Twenty thousand simulations were conducted with unique Gaussian white noise as the time series.

large, giving an overinflated Type I error probability  $\alpha$  for the Aggregated test.

We can use a union upper bound on the Type I error for  $R = 1$  and any  $\beta$ . The idea is to set the Type I error for the entire Aggregate test to  $\alpha$ , then try and estimate the Type I error  $\beta$  needed for each  $(T_a)_i$ , thus setting respective rejection regions for the input tests. We will take each  $(T_a)_i$  to have the same Type I error and call it  $\beta$ . The union upper bound for the case of  $R = 1$  is

$$P[T_a \geq 1] = P[F'_4(K_1) > F_1] \cup P[F'_4(K_2) > F_2] \cup \dots \cup P[F'_4(K_{|\mathcal{K}|}) > F_{|\mathcal{K}|}] \quad (4.3)$$

$$\leq \sum_{i=1}^{|\mathcal{K}|} P[F'_4(K_i) > F_i] \quad (4.4)$$

$$= \sum_{i=1}^{|\mathcal{K}|} \beta = \beta \cdot |\mathcal{K}|. \quad (4.5)$$

Considering disjoint union, the upper bound will always be greater or equal to the true probability of the union. Next, we can estimate the point at which the complementary

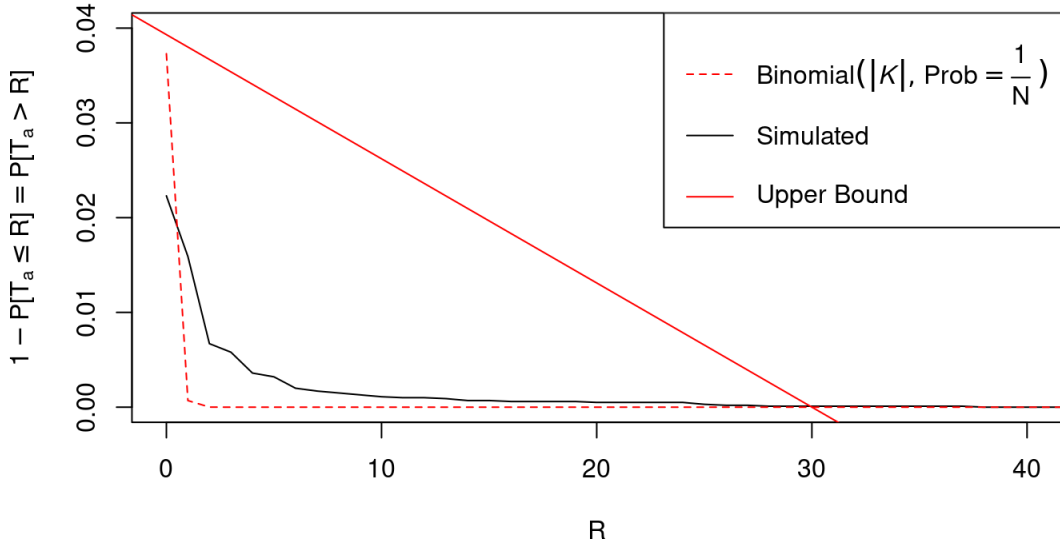


Figure 4.2: Estimated complementary CDF under the null hypothesis for the Aggregate test statistic with  $\mathcal{K} = \{5, \dots, 80\}$  and  $\beta = \frac{1}{N}$ ,  $N = 2000$ . 10,000 simulations were conducted with unique Gaussian white noise as the time series for each simulation.

CDF is negligibly different from zero. Based on Fig. 4.2, we could choose, for example,  $R = 30$ . If we now take these two points ( $R = 1, y = \beta|\mathcal{K}|$ ) and ( $R = 30, y = 0$ ) we can draw a line, shown on Figure 4.2, and solve for  $\beta$  in terms of  $\alpha$ .

First, using the point ( $R = 30, y = 0$ ) in the form  $Y = mR + b$

$$\begin{aligned} 0 &= m(30) + b \\ b &= -30m \end{aligned} \tag{4.6}$$

and using the second point ( $R = 1, y = \beta|\mathcal{K}|$ ) and Eqn. (4.6),

$$\begin{aligned} \beta|\mathcal{K}| &= m \cdot 1 - 30m \\ &= m(-29) \\ m &= -\frac{\beta|\mathcal{K}|}{29} \end{aligned} \tag{4.7}$$

the equation of the line is given by

$$\begin{aligned} Y &= -\frac{\beta|\mathcal{K}|}{29}R + 30\frac{\beta|\mathcal{K}|}{29} \\ &= \beta|\mathcal{K}|\left(\frac{30-R}{29}\right) \end{aligned} \tag{4.8}$$

Since we want to set the Type I error always to be  $\alpha$ , we can solve for the estimated  $\beta$  for any choice of  $R$  to obtain a choice of beta for a given choice of tapers,  $\alpha$ , and  $R$

$$\beta = \frac{29\alpha}{|\mathcal{K}|\left(30-R\right)}, \tag{4.9}$$

which is valid for  $R < 30$ . In practice, we expect that  $R$  should be small ( $R = 1$  or  $R = 2$ ) to maintain good detection probability at a carrier frequency. This is explored further in the next chapter.

Note that while the simulation in Figure 4.2 was fairly general in that it covered a reasonable range of  $K$ , suitable for common real-world analyses, this was not exhaustive, and the bound should be considered quite crude. Using an upper bound such as the simple one described above gives a value of  $\beta$  that is smaller than necessary, leading to an overall size of the Aggregate test that is smaller than  $\alpha$ , and so gives a conservative test. A tighter upper bound would lead to an Aggregate test with a Type I error closer to (but still smaller than) the nominal prespecified  $\alpha$ , but this investigation is outside of the scope of the current thesis.

## 5. Simulation and Comparison

Now that we have discussed all proposed modifications, we should look at some examples of the newly created tests under simulation. To directly compare the Harmonic- $F$ ,  $\tilde{F}_3$ ,  $F_4$ ,  $F'_4$ , and the Aggregate test, we set the Type I error of the tests all to  $\frac{1}{N}$  as recommended by Thomson [30, 33]. This is trivial for the four  $F$ -tests but not as straightforward for the Aggregate test due to not knowing a good approximation to the distribution of the test statistic  $T_a$ . For example, we found that under a particular simulation, the Type I error of the Aggregate test while using the upper bound approximation, Eqn. (4.9), results in a Type I error of approximately  $1.5 \times 10^{-4}$  when we would be expecting  $\frac{1}{N} = \frac{1}{2000} = 5 \times 10^{-4}$ . Thus, in the following simulations, the Aggregate test will have a lower Type I error than the four aforementioned  $F$ -tests. As Eqn. (4.9) is an upper bound to the complementary CDF of the distribution of the Aggregate test statistic the performance of the Aggregate test using this bound is conservative and so the simulations below will remain a good representation of the performance of the Aggregate test in practice. In the simulations in this chapter we used a threshold of  $R = 2$  and  $\mathcal{K} = \{5, \dots, 80\}$  for the Aggregate test.

During the study of  $\tilde{F}_3$  one can, in theory, pick any degree polynomial  $P$  to test. This open choice of degree creates a wide field of study which was beyond the scope of this thesis. Therefore, all the improvements above were discovered when only observing time series with linear modulation present. We believe this choice does not take away from our test statistic, as any higher degree polynomial can be

approximated by a line if one takes a small enough subset of the time series, i.e., if the time series is blocked then the modulation occurring in a finite time, while polynomial, may be approximated well by a linear modulation.

## 5.1 Test Parameters

To aid in understanding the simulations below, we will first explain the effect of changing parameters in the expression

$$X_t = \mu \cos \left( 2\pi f t + 2\pi \int_0^t \phi(\tau) d\tau \right) + Z_t. \quad (5.1)$$

The first parameter,  $\mu$ , is used to set the signal-to-noise ratio (SNR). In our simulations, if we say  $\text{SNR} = 0.5$ , this was calculated with the cosine amplitude divided by the noise variance,  $Z_t$ :

$$\text{SNR} := \frac{\mu}{\text{Var}[Z_t]}. \quad (5.2)$$

$\phi(\tau)$  dictates the modulation type in the simulation. If one wants to create linear modulation centred at the carrier frequency with modulation width  $m_P$ , we set  $\phi$  with

$$\phi(\tau) = \frac{m_P}{\frac{N}{2}} \left( t - \frac{N}{2} \right). \quad (5.3)$$

When experimenting with non-centred linear modulation, i.e. the  $x$ -intercept of 5.3 not being located at the carrier frequency. It was found that changing the linear offset effectively changes the detected center carrier frequency; this is similar to the effect of aliasing. Thus, the only linear modulation of interest is where the modulation is centred at the carrier frequency. In Eqn. (5.3), we must choose the modulation width and the length of the time series. Below, we will compare the performance of different modulations and time series lengths under Gaussian white and coloured background noise.

Simulation Parameters			
Bandwidth $m_p$	Carrier $f$	Amplitude $\mu$	Polynomial Degree
0.0005	0.15	0.5	Linear
0.001	0.15	0.5	Linear
0.002	0.15	0.5	Linear
0.003	0.15	0.5	Linear
0.005	0.15	0.5	Linear

Table 5.1: Simulation input parameters for white background noise. Each simulation is conducted with 10,000 iterations.

## 5.2 Power of Tests Under White Background Noise

Firstly, we want to examine the detection probability of these new tests in contrast to  $\tilde{F}_3$ . We chose to use an SNR of 0.5 because most time series with frequency modulation will realistically have an SNR of less than 1. Thus, the functionality is much more critical below 1 than above 1. The details of the simulations using Gaussian white noise are provided in Table 5.1.

To illustrate the type of time series on which we are conducting these modulated  $F$ -tests, we took one iteration of the time series with modulation used in the simulation. We observed its spectral estimate using the multitaper spectral estimator with  $K = 9$ , presented in Fig. 5.1. We can see a spike found around the carrier of  $f = 0.15$ , but the time series is quite noisy, as it is a time series with white background noise and small SNR.

To explore the modulation widths chosen in this simulation study shown in Table. 5.1, we first looked at  $m_p = 0.0005$  as shown in Fig. 5.2. Here can see that the detection probability of the Harmonic  $F$ -test (Harm-F) is far higher than any other modulated tests examined, including our new proposed tests. We have found that with a low SNR and a minimal modulation width, there is virtually no degradation of detection due to the frequency modulation when using the Harmonic  $F$ -test under

linear modulation. Therefore, we can consider  $m_P = 0.0005$  as a lower bound for where frequency modulation tests do not need to be used as the Harmonic  $F$ -test is sufficient. Further, as neither of the alternate  $F$ -tests can detect modulation at this low level, the Aggregate test similarly does not perform well.

Now observe Fig. 5.3, created with a linear modulation width of  $m_P = 0.001$ . The detection probability has now increased for the three modulated  $F$ -tests in contrast to the Harmonic  $F$ -test, which we can see has a low probability of detection for a low choice of  $K$ . However, when more tapers are used, eg.,30, the Harmonic  $F$ -test still detects the carrier frequency with a high probability. We believe that the decrease in the likelihood of detection is now due to the modulation being wide enough to pass through the Harmonic  $F$ -test without detection. When the number of tapers is increased, the bandwidth of the test,  $W$ , is also increased allowing the test to correctly identify the carrier frequency. For  $\tilde{F}_3$  and  $F_4$ , the even and odd

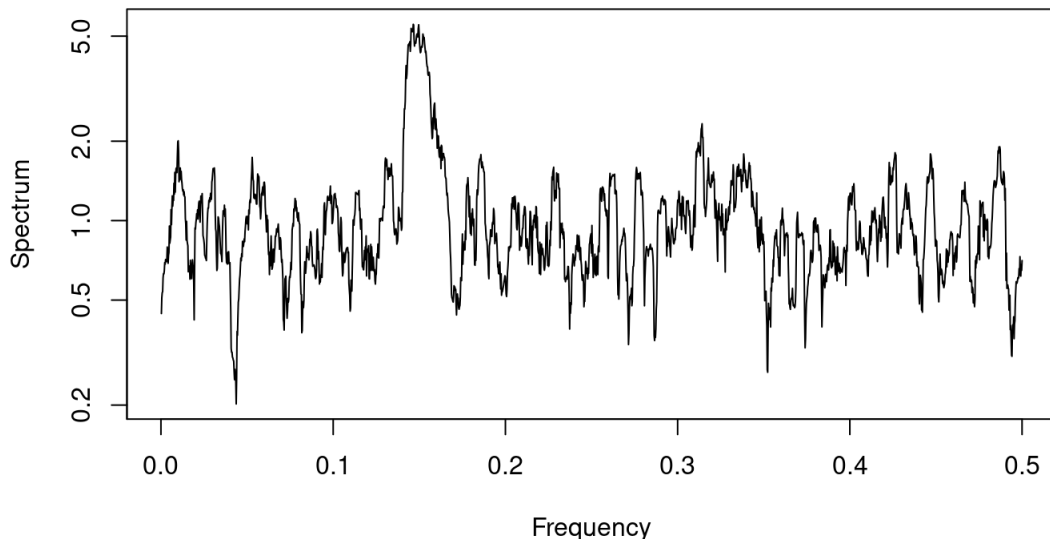


Figure 5.1: Spectral multitaper estimate for  $K = 9$  under  $m_P = 0.01$  linear modulation and  $N = 2000$ .

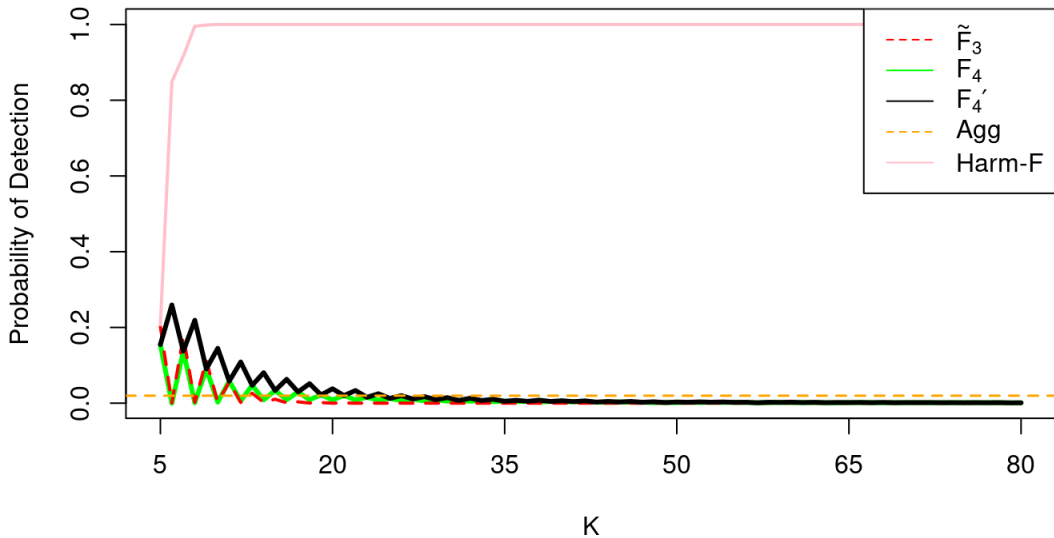


Figure 5.2: Probability of detection conducted through 10,000 simulations with modulation width  $m_P = 0.0005$  and  $N = 2000$ . The pink curve, Harm-F, is the classic Thomson [30] Harmonic  $F$ -Test Statistic, which has little modulation detection capacity.

behaviour is also noticeable in Fig. 5.3. It is especially prominent for  $\tilde{F}_3$  in particular. Furthermore, we observed that a small range of tapers will work optimally for the  $\tilde{F}_3$  test, while in contrast, the new  $F'_4$  has a broader range of tapers that will give a more consistent probability of detection. Notice that the removal of the last taper during the derivation of  $F'_4$  has caused the even-odd behaviour to flip; now, the odd tapers have the lowest probability of detection compared to the even tapers. We attempted removing the last taper from odd choices of  $K$  as well, but this was found to lower the probability of detection. The Aggregate test was found to perform better here when compared to a modulation width of  $m_P = 0.0005$ . This was most likely due to the higher probability of detection for  $F'_4$ . Provided that a higher number of tapers is used with the Harmonic  $F$ -test, it remains the best option for the cases with small modulation widths.

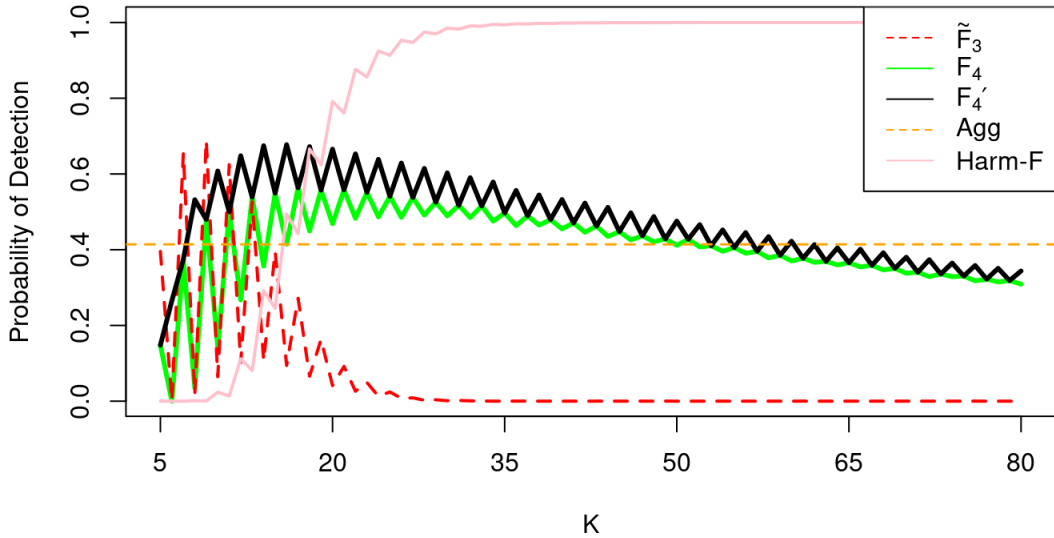


Figure 5.3: Probability of detection conducted through 10,000 simulations with modulation width  $m_P = 0.001$  and  $N = 2000$ .

Fig. 5.4 and Fig. 5.5 demonstrate the advantages of using the tests created in this thesis. Although  $\tilde{F}_3$  does perform well for  $K \in \{13, 15, 17\}$  and  $m_P = 0.002$ ; when  $m_P = 0.003$ , the optimal  $K$  has now shifted to the right when compared to the previous figures. The Aggregate test has a very high detection probability of 0.997 and 0.999 for Fig. 5.4 and Fig. 5.5, respectively. Additionally, there is an increase in possible  $K$  that provides  $\tilde{F}_3$  with high detection probability in  $m_P = 0.003$ , but  $F'_4$  has the largest range of  $K$ 's that give a high probability of detection.

Observing, Fig. 5.6, it demonstrates the effect of a large modulation bandwidth to  $m_P = 0.01$ . The shift in optimal choices of tapers is much more noticeable. For any test that uses a single choice of taper such as  $\tilde{F}_3$ ,  $F_4$ , or  $F'_4$ , any  $K$  below 20 provides an extremely poor detection rate. In practice, the  $m_P$  is unknown so when comparing  $m_P = 0.1$  and  $m_P = 0.002$  for  $\tilde{F}_3$ , the optimal choice of  $K$  never overlap. If an analyst chose  $K = 20$ , they might detect a small modulation width, but no detections for a

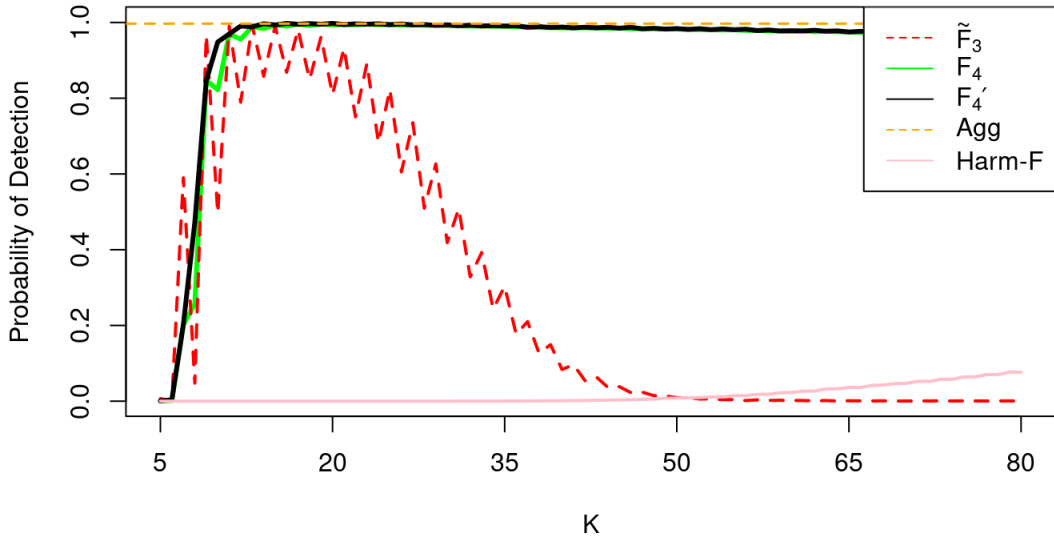


Figure 5.4: Probability of detection conducted through 10000 simulations with modulation width  $m_p = 0.002$  and  $N = 2000$ . This is the first case where the  $F_4$  and  $F'_4$  tests perform at acceptable levels, while the Harmonic  $F$ -test is essentially unusable. Note the extremely strong performance of the Aggregate test removes the subjectivity of the  $F_4$  and  $F'_4$  tests which vary with  $K$ .

large modulation width will be found. In contrast, if a person chose  $K = 40$ , they will have a high detection probability for a large modulation but no detection if the modulation is small. This uncertainty in the choice of tapers is why the Aggregate test is generally more applicable.

Comparing Fig. 5.6 to Fig. 5.7, the only difference is the penalty term applied to the eigencoefficients in  $F'_4$ . The first uses 0.1, similar to the other plots in this section, and the second plot uses 0.05. It is believed that the decrease in detection of  $F'_4$  was due to an over-penalization at this wider modulation bandwidth. In the second figure, the penalty term was reduced, and one can see that the probability of detection was increased for both the Aggregate test and the  $F'_4$ . This shows that the penalty term is a tuning parameter that needs to be examined in more detail in the

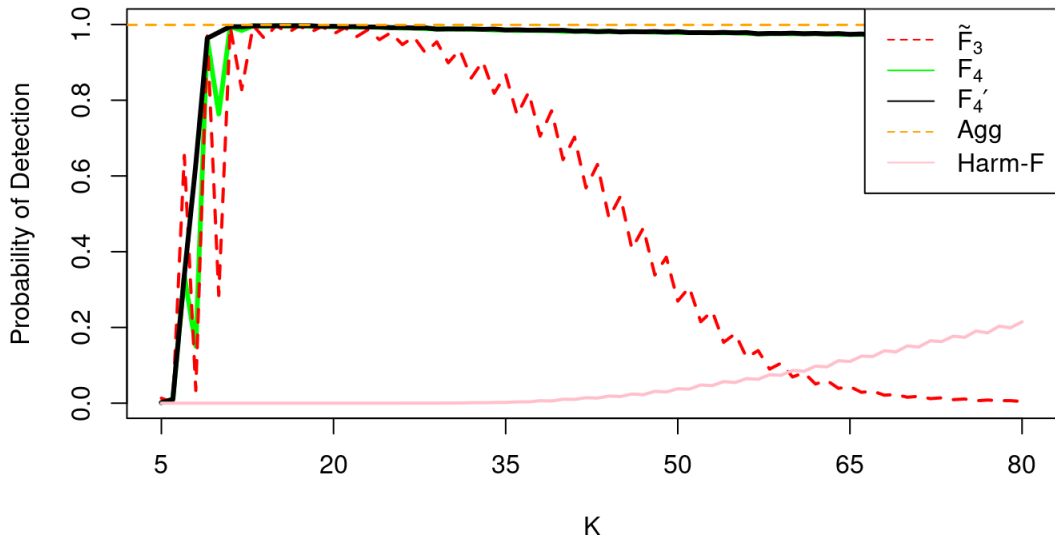


Figure 5.5: Probability of detection conducted through 10000 simulations with modulation width  $m_P = 0.003$  and  $N = 2000$ . Essentially the same performance and conclusion as suggested by Figure 5.4.

future.

### 5.3 Power of Tests With Coloured Noise

To examine the aforementioned test statistics under more realistic conditions, we will now study simulations from which the noise was not Gaussian. We chose to use samples from a Gumbel distribution with zero mean. The input parameters were taken from AR and MA processes. The details of the simulations are provided in Table 5.2.

Consider first Fig. 5.8. Here, the coloured noise decreases the probability of detection compared to the white noise cases with similar parameters. There is consistency in that for  $m_P = 0.001$ , the Harmonic  $F$ -test continues to perform the highest in this case of low SNR with a small length of time series of  $N = 2000$ , although the largest

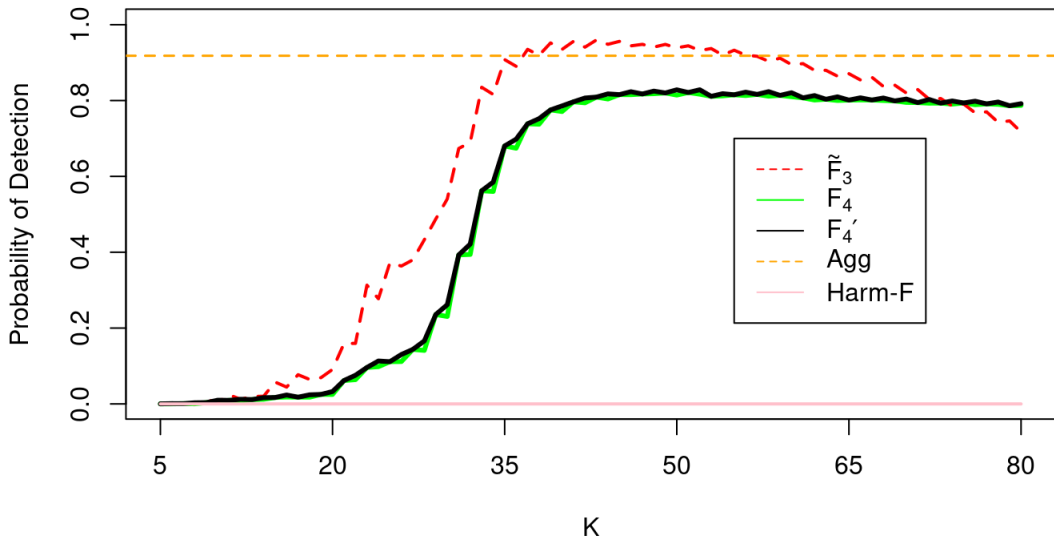


Figure 5.6: Probability of detection conducted through 10,000 simulations with modulation width  $m_p = 0.01$  and  $N = 2000$ .

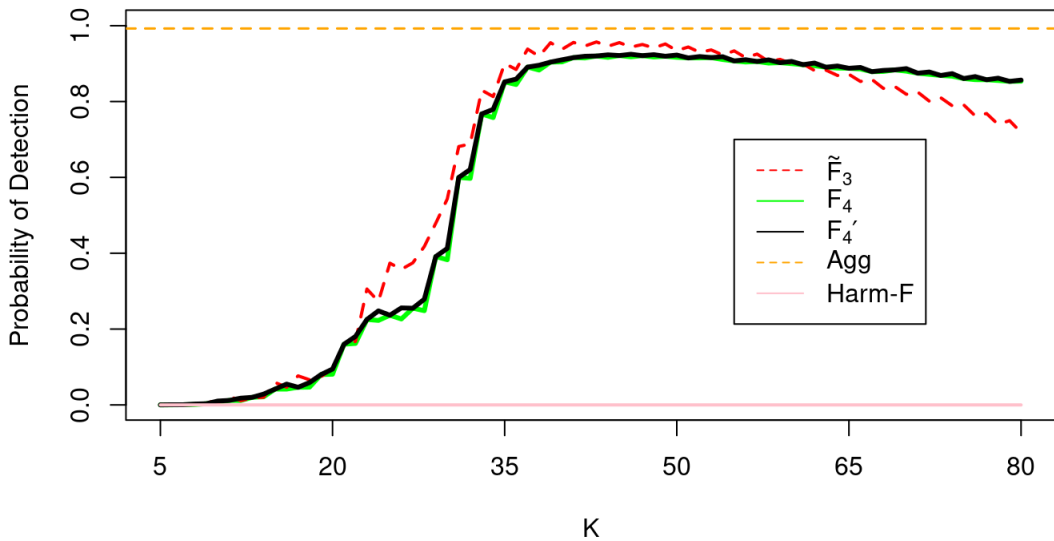


Figure 5.7: Probability of detection conducted through 10,000 simulations with modulation width  $m_p = 0.01$ ,  $N = 2000$  and using a penalty in the weighting of 0.05.

probability of detection being approximately 0.8.

Fig. 5.9, displays an improvement in the performance of  $F'_4$ , when compared to  $\tilde{F}_3$ . It is interesting that with the coloured noise, the even-odd issue is not as prominent when compared to the white noise cases shown earlier. This suggests that the even-odd problem is an inherent challenge with the modulation and how it interacts with the noise in the last taper. Once again, the Aggregate test is not performing as well as the  $F'_4$  test due to the upper bound estimate for the distribution. To illustrate the difference in the noise, we also generated the multitaper spectral estimate for the case of linear modulation with AR = (0.4, -0.2, 0.6), MA = (0.2, -0.5) with innovations from a zero mean Gumbel distribution, presented in Fig. 5.10. This spectrum is rather interesting as the main peak, found at the modulation carrier of 0.15, but there are other parts of the spectrum with higher power. Observing one instance of this simulation for  $F'_4$  shown in Fig. 5.11, we see that even though the spectrum in Fig. 5.10 has peaks at 0.3 and near 0, the  $F$ -test only detects at the carrier frequency of 0.15 and one other point at 0.4067. This could be due to the steep slope in the

Simulation Parameters				
Bandwidth $m_P$	Carrier $f$	Amplitude $\mu$	Polynomial Degree	ARMA
0.001	0.15	0.5	Linear	AR = (0.4, -0.2, 0.6), MA = (0.2, -0.5)
0.002	0.15	0.5	Linear	AR = (0.4, -0.2, 0.6), MA = (0.2, -0.5)
0.003	0.15	0.5	Linear	AR = (0.4, -0.2, 0.6), MA = (0.2, -0.5)
0.003	0.15	0.5	Linear	AR = (0.5, 0.3, -0.1), MA = (0.6)
0.005	0.15	0.5	Linear	AR = (0.5, 0.3, -0.1), MA = (0.6)
0.01	0.15	0.5	Linear	AR = (0.5, 0.3, -0.1), MA = (0.6)

Table 5.2: Simulation input parameters for coloured background noise. Each simulation is conducted with 10,000 iterations.

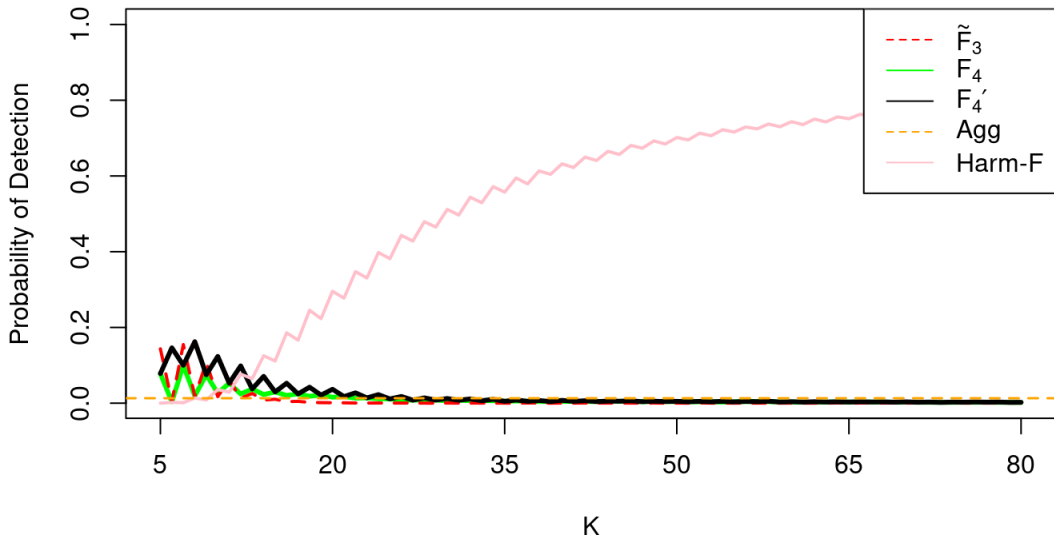


Figure 5.8: Probability of detection,  $m_P = 0.001$  and  $N = 2000$ . Background noise was generated with an AR = (0.4, -0.2, 0.6), MA = (0.2, -0.5) with innovations from a zero mean Gumbel distribution.

spectrum around 0.4, as the instantaneous frequency could be mistaking this slope for modulation.

To expand on the types of background noise used in this study, we will now examine two cases of  $m_P = 0.003$ . The first simulation, shown in Fig. 5.12, has the same ARMA coefficients and the same shape of spectrum as illustrated in Figure 5.10. The other, shown in Fig. 5.14, has a multitaper estimate shown in Fig. 5.13. We see that this alternative choice of ARMA coefficients gives us a spectral estimate that is coloured with a slope from  $f = 0$  to  $f = 0.5$ . Even with these two types of noise, the respective probability of detection plots are very similar but not identical. The fact that there is a similarity between the two different simulations does give us some confidence in the robustness of the test statistics created in this thesis. Again, the Aggregate test does perform better under the higher cases of modulation, but not as well as the  $F'_4$  test.

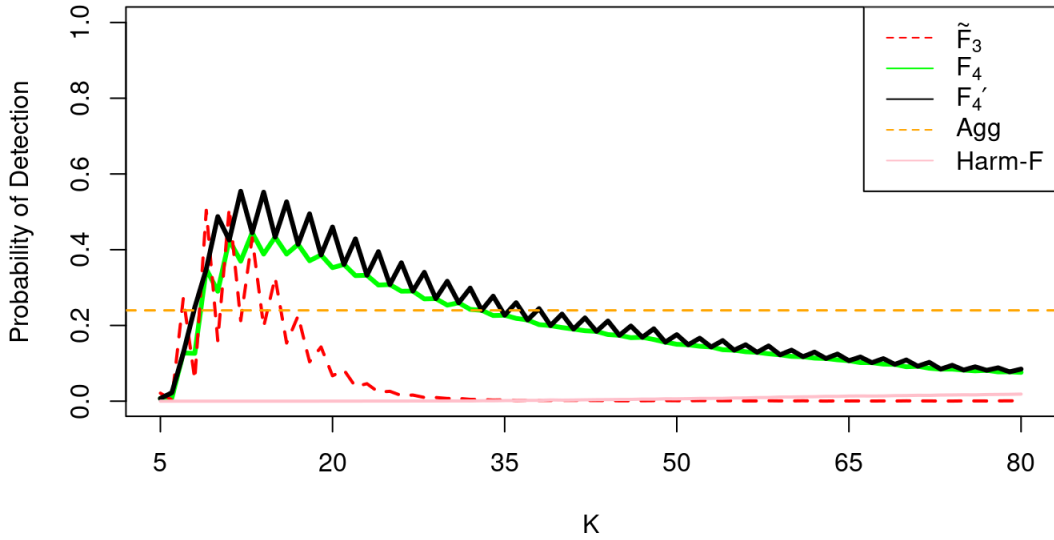


Figure 5.9: Probability of detection ,  $m_P = 0.002$  and  $N = 2000$ . Background noise was generated with an AR = (0.4, -0.2, 0.6), MA = (0.2, -0.5) with innovations from a zero mean Gumbel distribution.

Lastly, we looked at the higher levels of modulation, presented in Fig. 5.15 and Fig. 5.16. These two probability-of-detection plots are created with modulation widths of  $m_P = 0.005$  and  $m_P = 0.01$ , respectively. The same pattern emerges as with the white noise case, as the point in which the probability of detection is highest moves toward higher  $K$ 's as the modulation width is increased. For  $m_P = 0.005$ , the Aggregate test performs better than the  $F'_4$  test. This indicates that the detections that were found in the  $F'_4$  test would be above the significance cut off-of  $\alpha = \frac{1}{N}$ , as they would also have to be detected at the significance level,  $\beta$ , for the Aggregate test. As for Fig. 5.16, we see the  $\tilde{F}_3$  has the highest probability of detection by a large margin. It stands to reason that this would be either due to the penalty term being too aggressive for the sets of tapers between  $K = 35$  to  $50$ , or the DPSSs in band concentration is aiding in the detection for wider modulation widths.

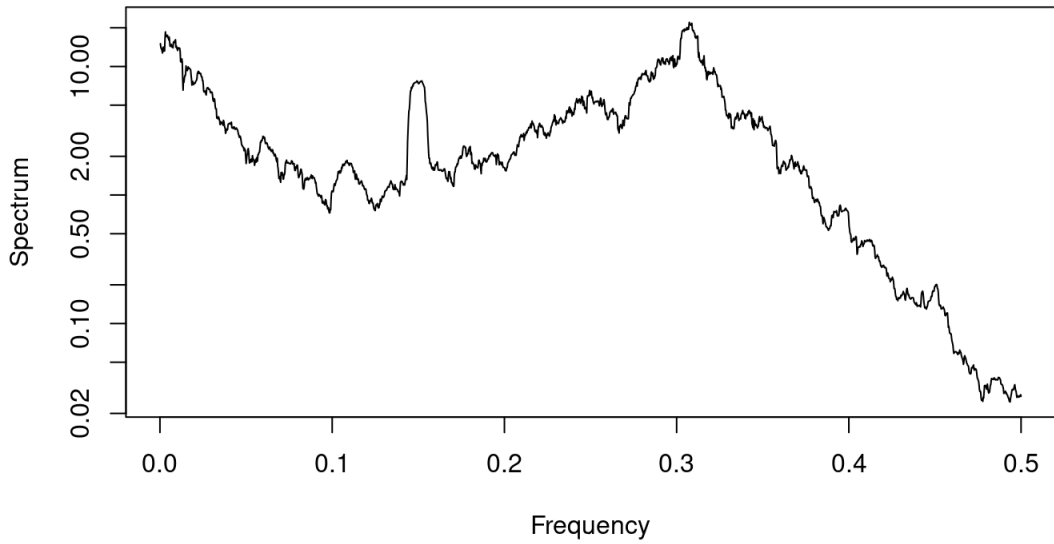


Figure 5.10: Spectral multitaper estimate for  $K = 20$  and  $N = 2000$  under  $m_P = 0.002$  linear modulation with AR =  $(0.4, -0.2, 0.6)$ , MA =  $(0.2, -0.5)$  with innovations from a zero mean Gumbel distribution.

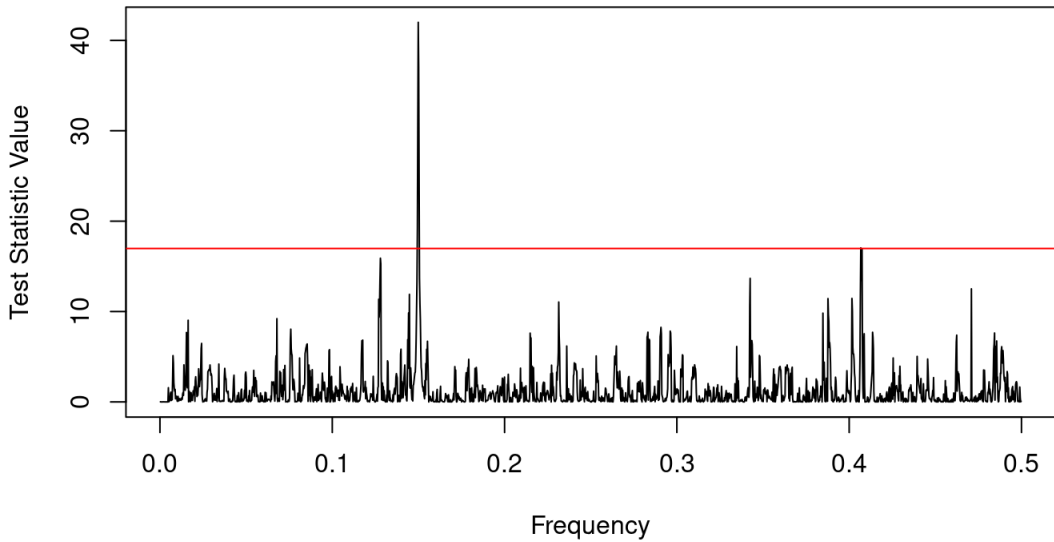


Figure 5.11: Realization of test statistic  $F'_4$  for  $K = 20$  and  $N = 2000$ . The test is applied to the same data as the spectrum provided above in Figure 5.10, with the same parameters.

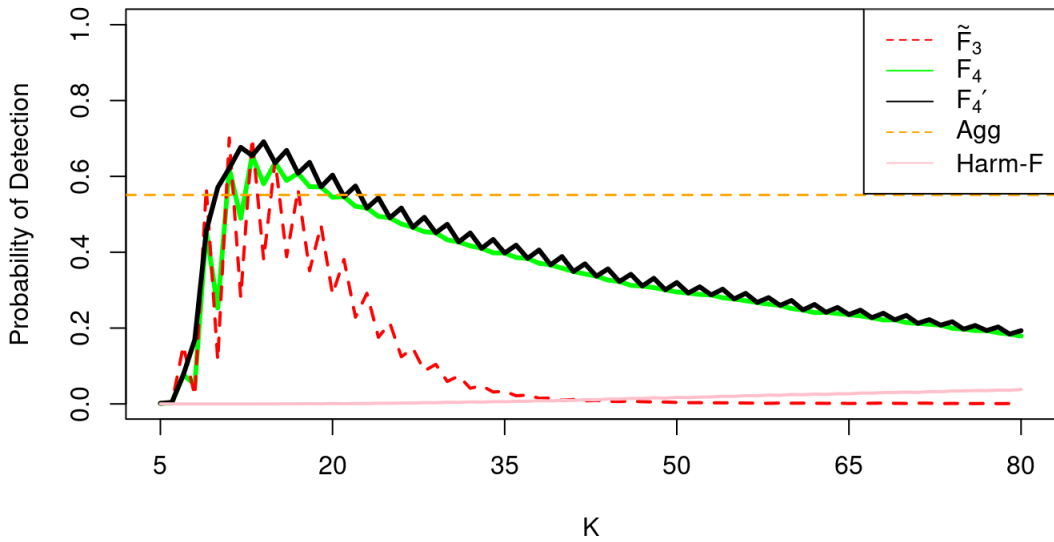


Figure 5.12: Probability of detection,  $m_P = 0.003$  and  $N = 2000$ . Background noise was generated with an AR = (0.4, -0.2, 0.6), MA = (0.2, -0.5) with innovations from a zero mean Gumbel distribution.

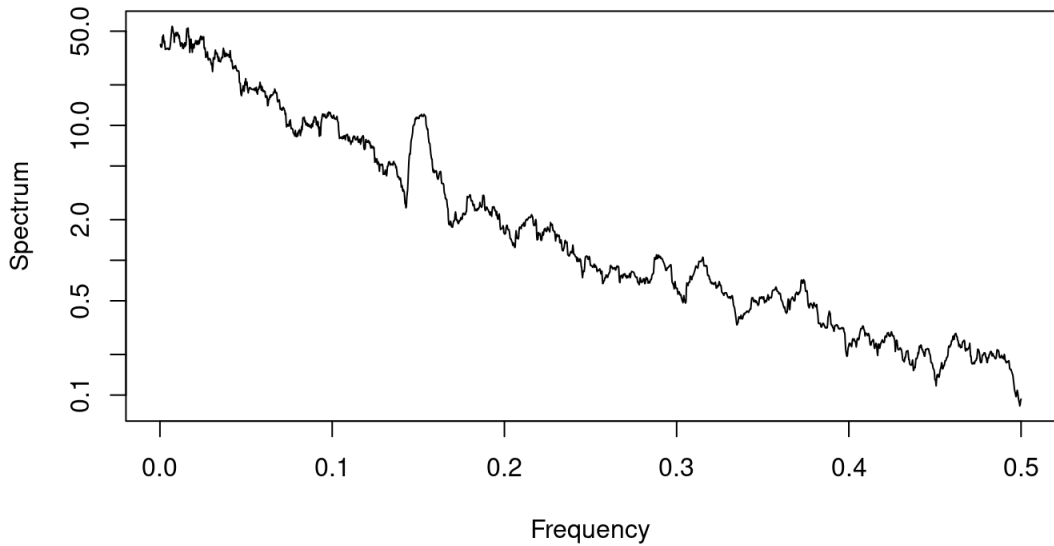


Figure 5.13: Spectral multitaper estimate for  $K = 20$  under  $m_P = 0.003$  linear modulation with AR = (0.5, 0.3, -0.1), MA = (0.6) with innovations from a zero mean Gumbel distribution.

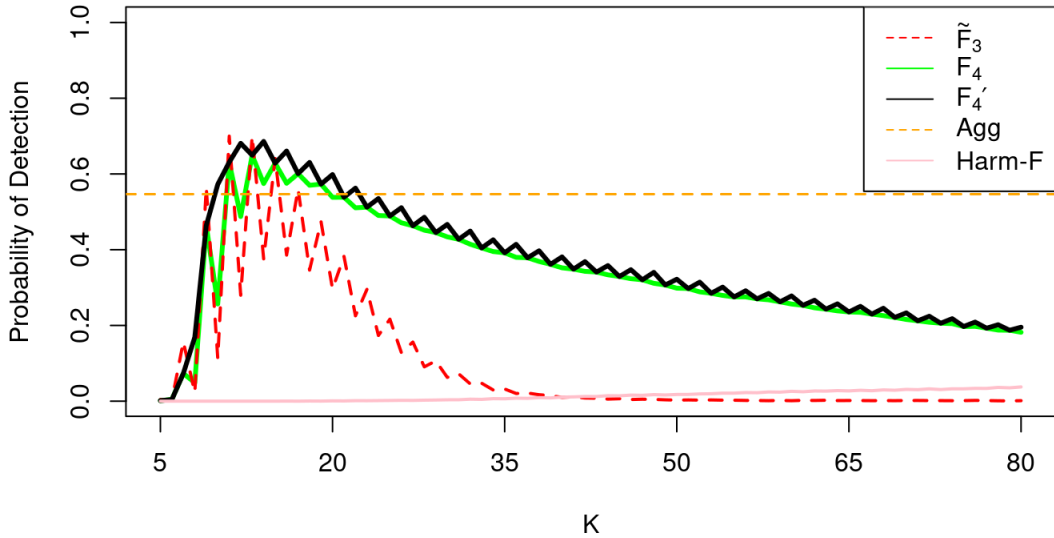


Figure 5.14: Probability of detection,  $m_P = 0.003$  and  $N = 2000$ . Background noise was generated with an AR = (0.5, 0.3, -0.1), MA = (0.6) with innovations from a zero mean Gumbel distribution.

## 5.4 Aggregate Test Estimate

To examine the proposed Aggregate test more fully, we also tried estimating the input parameter  $\beta$  in order to obtain a Type I error closer to  $\frac{1}{N}$ . To do this, we chose a range of input parameters for  $\beta$  and conducted 20,000 simulations under the null hypothesis. We discovered for  $N = 2000$ , if we used  $\beta = 2 \times 10^{-5}$ , the resulting Type I error would be  $4 \times 10^{-4}$ . This is slightly lower than the desired Type I error of  $\frac{1}{N} = \frac{1}{2000} = 5 \times 10^{-4}$ , but higher than the upper bound Type I error of  $1 \times 10^{-4}$ .

We will now conduct two simulations to examine how this change will effect the probability of detection. First, in Fig. 5.17, we see in blue that the Aggregate test conducted with the estimated  $\beta$  has higher power than that using the upper-bound estimate. This was expected as the upper-bound estimate of  $\beta$  is much lower at  $7.631 \times 10^{-6}$ . The  $F'_4$  is superior for an optimal choice of  $K$  when compared to the

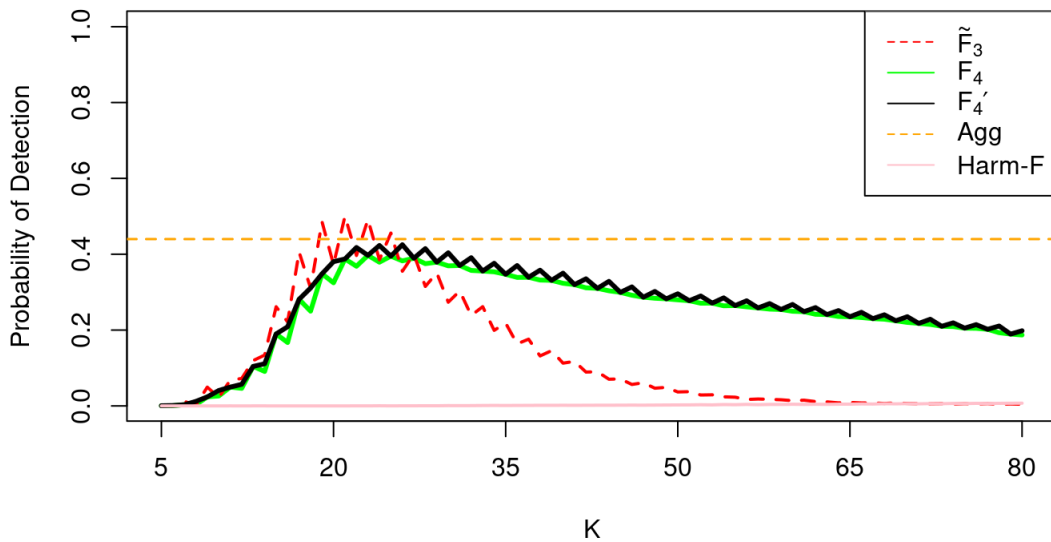


Figure 5.15: Probability of detection with modulation width  $m_P = 0.005$  and  $N = 2000$ . Background noise was generated with an AR = (0.5, 0.3, -0.1), MA = (0.6) with innovations from a zero mean Gumbel distribution.

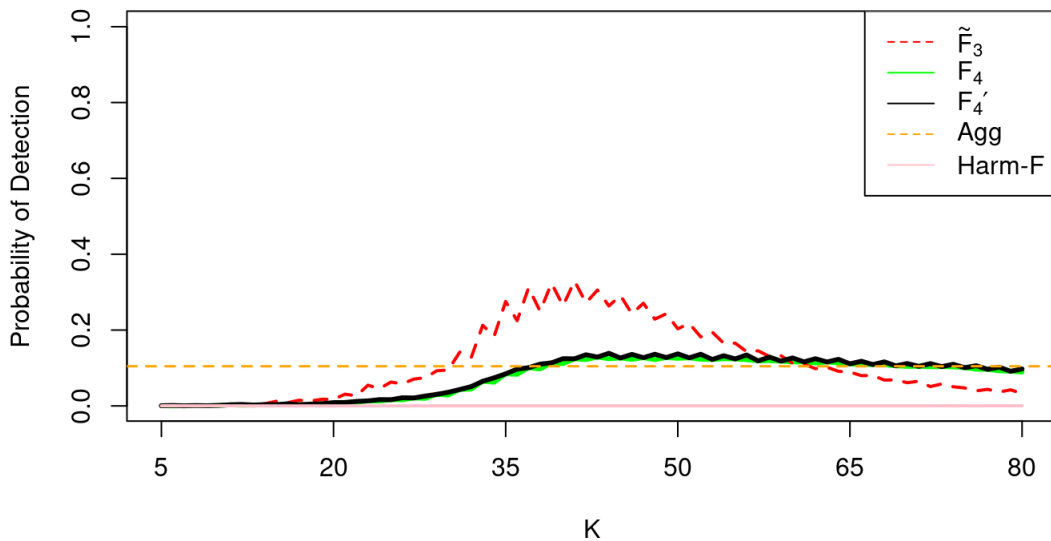


Figure 5.16: Probability of detection with modulation width  $m_P = 0.01$  and  $N = 2000$ . Background noise was generated with an AR = (0.5, 0.3, -0.1), MA = (0.6) with innovations from a zero mean Gumbel distribution.

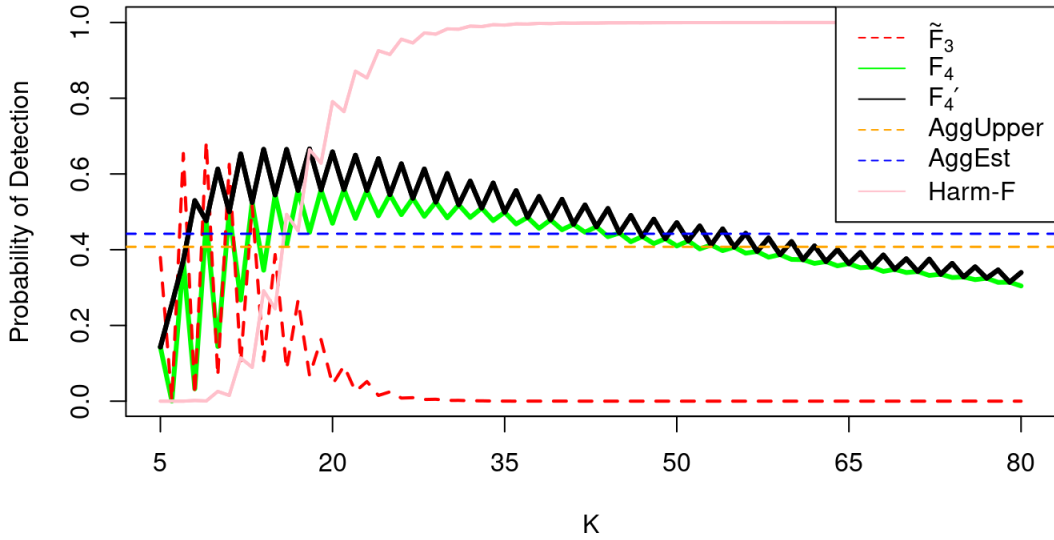


Figure 5.17: Probability of detection with modulation width  $m_P = 0.01$  and  $N = 2000$ . Background noise from AR = (0.5, 0.3, -0.1), MA = (0.6) with innovations from a zero mean Gumbel distribution.

Aggregate test. This is also apparent in Fig. 5.18. We obtain a 4 percent increase in the probability of detection, but it is not enough to surpass the  $F'_4$  test. This is believed to be attributed to the low probability of detection of all tests. As  $F'_4$  does not surpass 0.8, the Aggregate test also will be reduced accordingly. The advantage of using the Aggregate test is that it does not require the analyst to specify a particular  $K$ .

## 5.5 Conclusion

In conclusion, the types of modulation can be divided into three sections. In the case of small modulation,  $m_P \leq 0.001$ , the Harmonic  $F$ -test yields the highest probability of detection, provided a large enough number of tapers is chosen, for both white and coloured noise. In the case of moderate modulation with white noise,  $F'_4$  obtained

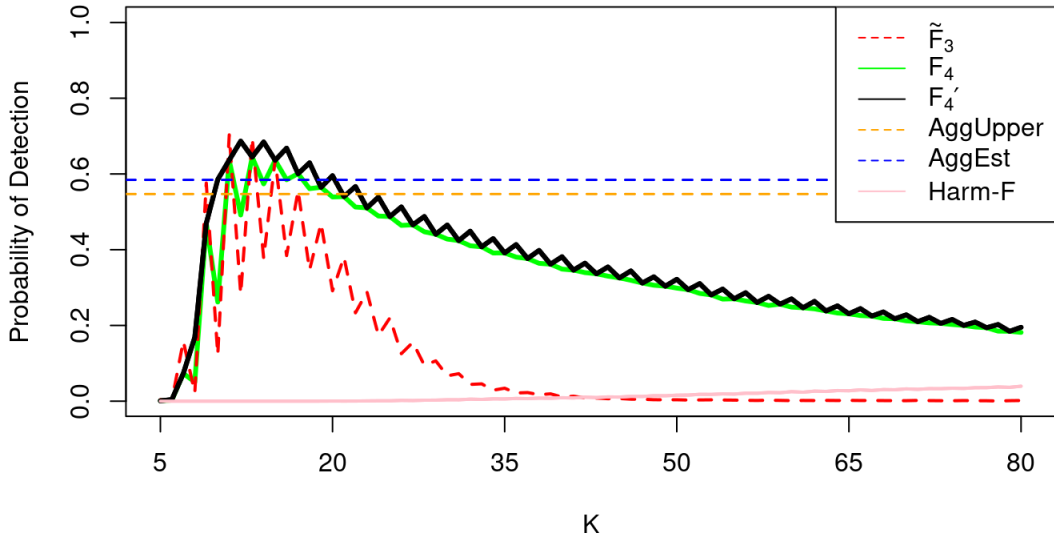


Figure 5.18: Probability of detection width  $m_P = 0.01$  and  $N = 2000$ . Background noise was generated with an AR = (0.5, 0.3, -0.1), MA = (0.6) with innovations from a zero mean Gumbel distribution.

the highest probability of detection and works quite well for a wide range of taper choices. Under coloured noise, again  $F'_4$  has the highest probability of detection, but its performance suffers due to the different types of noise: the probability of detection is only around 0.6 to 0.8 for the examples shown. This lower probability of detection causes a reduction in the detection rate of the Aggregate test as well. Under large amounts of modulation with white background noise, the Aggregate test (with the correct tuning parameters) has the highest probability of detection, closely followed by the  $\tilde{F}_3$  test. With the larger amounts of modulation for our examples and choice of  $N$ , with  $m_P \geq 0.005$ , the penalty term needs to be more carefully chosen to attain the optimal detection probability for both the Aggregate test and the  $F'_4$  test. Under large amounts of modulation with coloured noise, all the tests have a low detection for all choices of tapers. We also demonstrated that if a superior estimate for the upper bound on the Aggregate test existed, we would be able to increase the probability of

detection in simulation, and hopefully, in practice with real data.

## 6. SoHO Time Series Analysis

In [3], Blanchette *et al.* applied the  $\tilde{F}_3$  modulated  $F$ -test to solar data recorded from the SoHO GOLF (Solar and Heliospheric Observatory, Global Oscillations at Low Frequencies) instrument [4]. The SoHO satellite was used to study our Sun's inner structure by measuring the spectrum generated by the global oscillations in the range of frequencies between  $10^{-7}$  and  $10^{-2}$  Hz. There are three separate instruments for measuring the properties of the sun onboard; we will be only using data from the GOLF helioseismology instrument. This instrument is mainly used to find low-frequency modes, such as  $g$ -modes which give insight into the Sun's core. The satellite sits at the  $L_1$  Lagrange point, which allows the satellite to have an uninterrupted view of the sun. This satellite started recording on April, 11<sup>th</sup> 1997 and is still in use to this day. Fig. 6.1 discusses what the SoHO satellite looks like and where the  $L_1$  Lagrange point is located with respect to Earth.

Although we do not have the special experience or domain expertise needed to draw any meaningful conclusions about the Sun from the test results, it may be interesting to test  $F'_4$  and the aggregate test on the SoHO data to see how the results compare to the  $\tilde{F}_3$ .

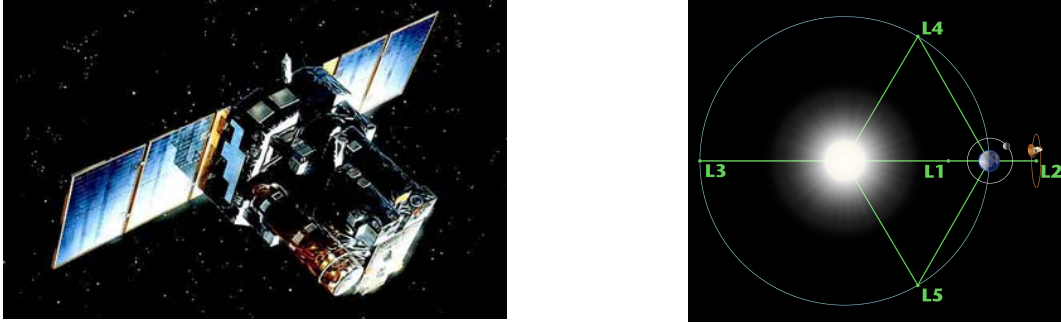


Figure 6.1: The left figure shows what the SoHo satellite would look like after being unfolded [17]. The right figure shows the location of the  $L_1$  Lagrange point with respect to the Earth and Sun [18].

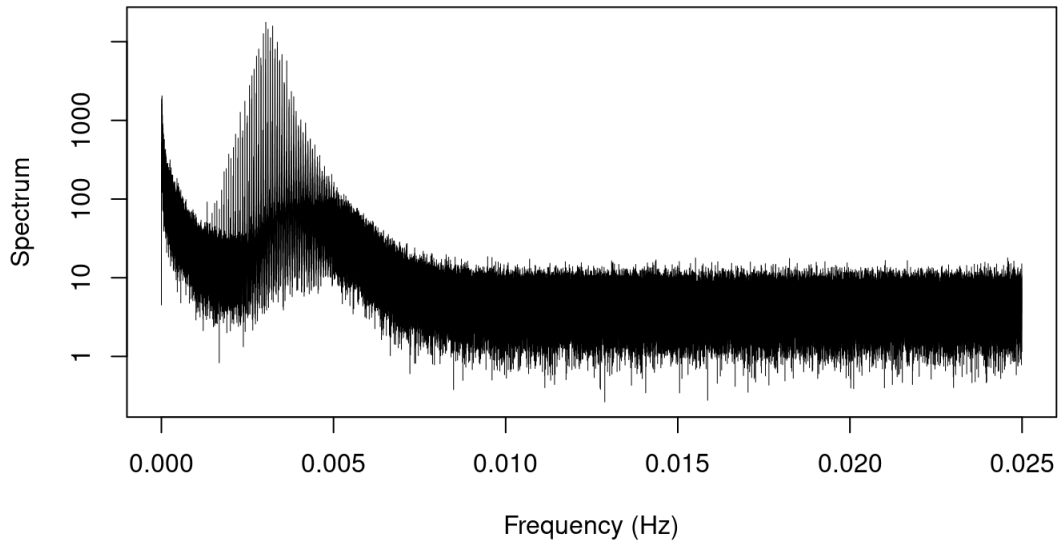


Figure 6.2: Multitaper spectral estimate of the original data for  $K = 7$ ,  $NW = 4$ , and  $\text{pad} = 2^{\lceil \log_2(10N) \rceil}$ .

## 6.1 Pre-Processing

We were fortunate to obtain the pre-isolated time series from [2]. The official solar data record can still be found at [4]. As they have updated the time series to include an extra four years, the indexing used in isolation may have differed, so we conducted

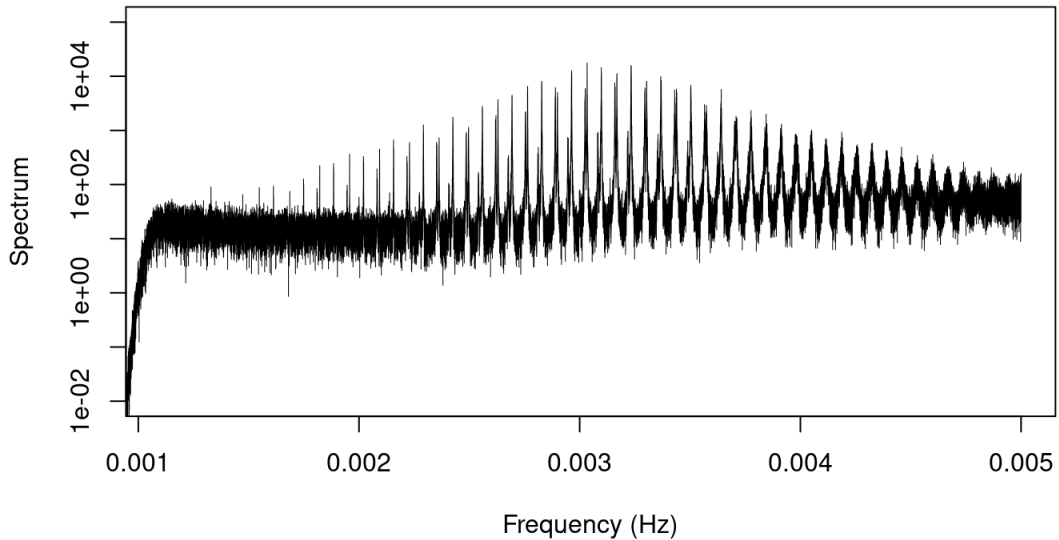


Figure 6.3: Multitaper spectral estimate of the processed data for  $K = 7$ ,  $NW = 4$ , and  $\text{pad} = 2^{\lceil \log_2(10N) \rceil}$ .

our analysis using only the `Solar_20s.RData` file found at the incorporated GitHub reference to ensure consistency [2].

First, using the `multitaper` [23] package in R [21], the multitaper spectral estimate was computed, presented in Fig. 6.2. As most large values found in the spectrum are located before  $f = 0.005$ , it was decided to first mean-centre the data, then use a 5:1 decimation filter and undersample the 20-second data down to one sample every 100 seconds; this simultaneously causes the Nyquist frequency to change to 0.005. The code used to decimate the time series was written originally by David J. Thompson in the F77 language, and provided to Blanchette *et al.*. The resulting finalized time series was then created by applying a high pass filter of 1 minus the decimation filter to make it comparable to the one found in [1]. The multitaper spectral estimate of the finalized time series is presented in Fig. 6.3. We will be not testing any frequencies under 1 mHz (0.001 Hz), as we can see that the high pass filter has made frequencies

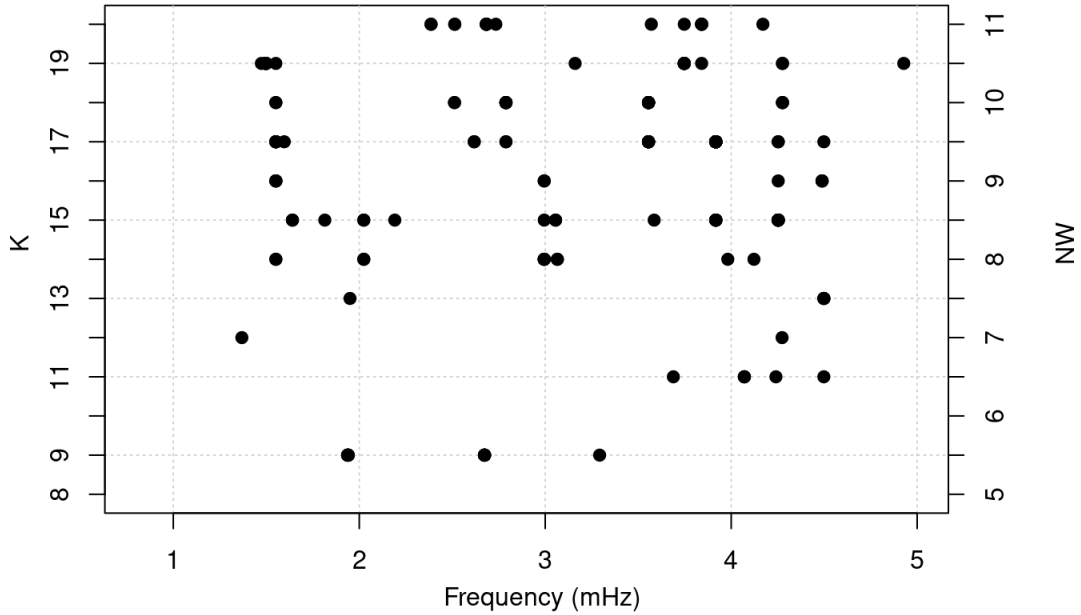


Figure 6.4:  $\tilde{F}_3$  result conducted with  $\alpha = \frac{1}{N}$ ,  $K = 2NW - 2$ , and  $\text{pad} = 2^{\lceil \log_2(10N) \rceil}$ .

under this threshold to have essentially zero power.

## 6.2 Analysis

The first step in the analysis was to verify that we could obtain the exact results previously found in [1]. This was crucial as the pre-processing method could be slightly different, potentially altering the results obtained in this thesis. It also verifies that the R code created for the  $\tilde{F}_3$  developed in this thesis is computationally similar to that used previously.

First, examining the plot used for verification shown in Fig. 6.4. The  $NW$  values corresponding to a specific  $K$  are shown on the right-hand side to allow for comparison to the previous work. As  $K$  is increased, we see some interesting patterns. Firstly, for  $K = 8$ , and 10, we now know there are no detections here due entirely to the even-odd issue that  $\tilde{F}_3$  experiences for small numbers of tapers. In addition, there seems to be

some consistency in significant frequencies across  $K = 14$  to  $K = 19$ . The prominent ones are the frequencies around 1.6, 2.8, 3.6, 4.3, and 4.5mHz.

By examining an overlay of the new  $F'_4$  test results on top of  $\tilde{F}_3$ , we gain some insight into the differences between the two tests. The interesting observation that can be gleaned from Fig. 6.5 is that  $F'_4$  detects some significant frequencies for  $K = 6, 8$ , and 10. The new test with the even-odd correction finds some significant frequencies where the aforementioned test did not. Some frequencies around  $f = 1.5, 2.8, 3.6$ , and 4.3 are still significant, with lots of agreement primarily found around 1.5. Some new frequencies have been found that were not detected by  $\tilde{F}_3$ , such as 4.7.

It is surprising how much these two  $F$ -tests disagree when it comes to the detected frequencies. Although there were some detections in common, the majority were not. A possible explanation could be that the amount of noise in the time series is causing

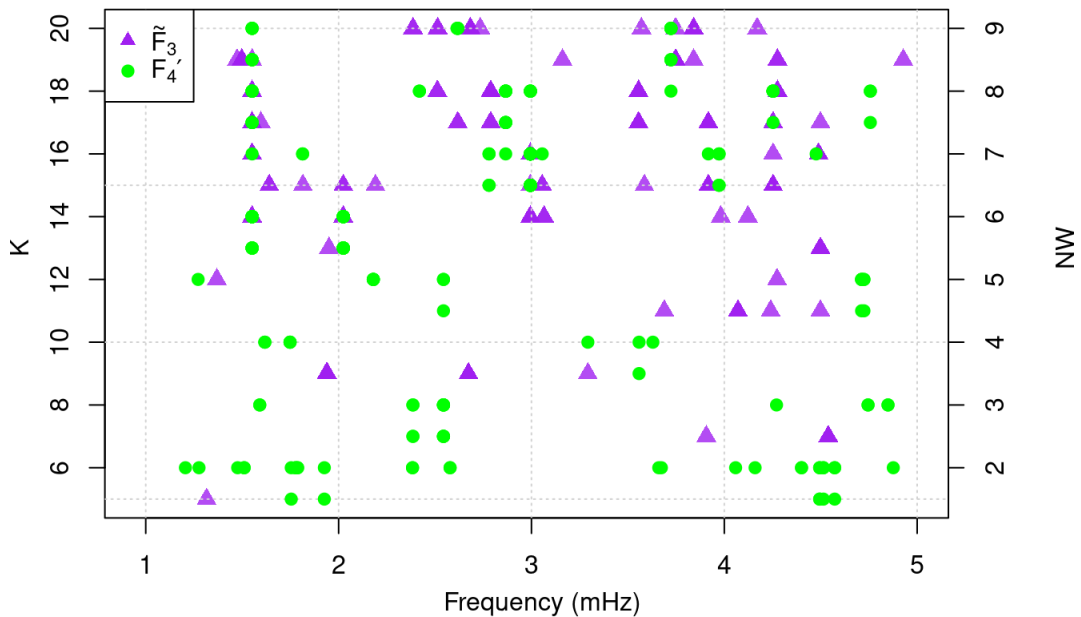


Figure 6.5: Overlay of  $F'_4$  in green compared to the results with  $\tilde{F}_3$  shown in purple. Both tests were conducted at a significance level of  $\alpha = \frac{1}{N}$ .

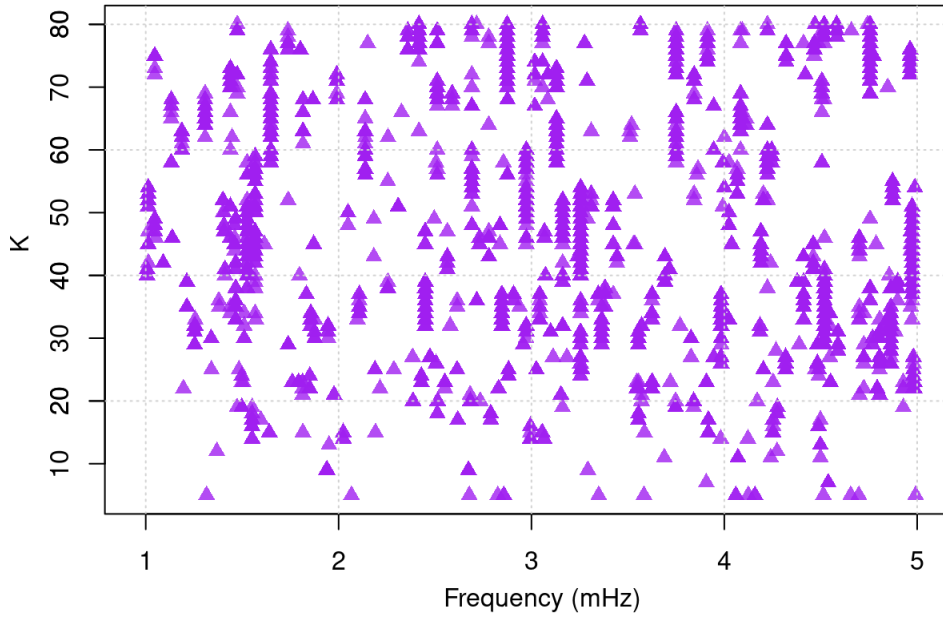


Figure 6.6:  $\tilde{F}_3$  result conducted with  $\alpha = \frac{1}{N}$ ,  $K = 2NW - 2$ ,  $\text{pad} = 2^{\lceil \log_2(10N) \rceil}$  for  $K = 5$  to  $K = 80$

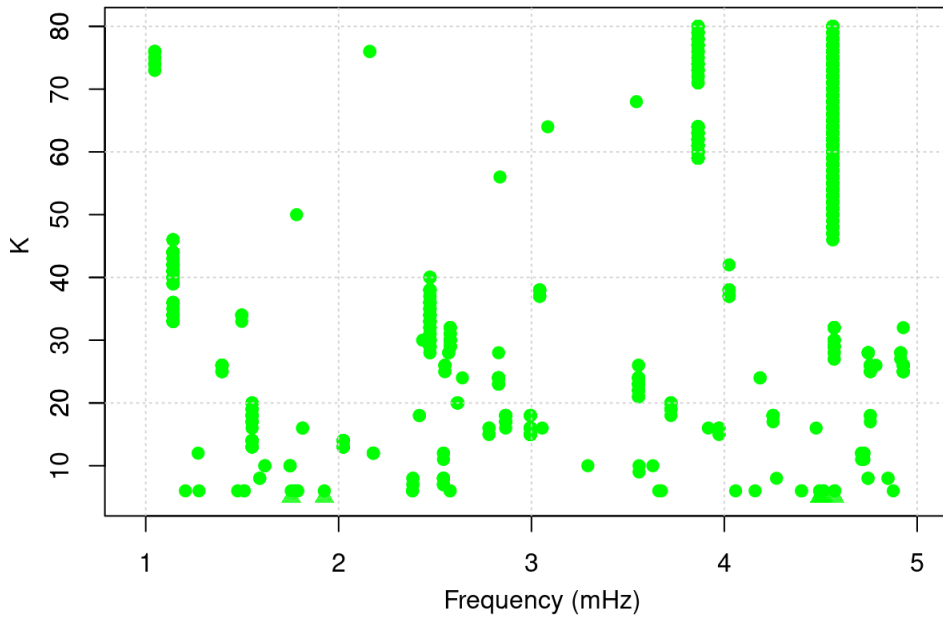


Figure 6.7:  $F'_4$  result conducted with  $\alpha = \frac{1}{N}$ ,  $\text{pad} = 2^{\lceil \log_2(10N) \rceil}$  for  $K = 5$  to  $K = 80$

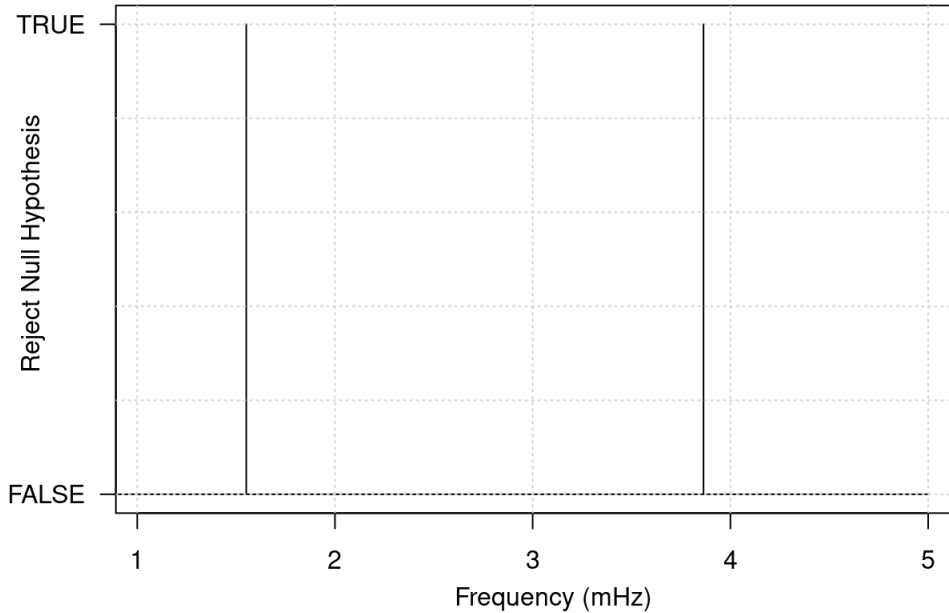


Figure 6.8: Aggregate test result conducted with  $\text{pad} = 2^{\lfloor \log_2(10N) \rfloor}$  for  $K = 5$  to  $K = 80$

the two tests to detect different signals. Another theory is that signals in the time series have slower modulation width, thus, one would have to use a higher number of tapers to see the primary frequencies that make up the series. This becomes increasingly interesting if you expand the number of tapers you look at.

In both Fig. 6.6 and Fig. 6.7, we expanded the number of tapers used for each test from  $K \in \{8, \dots, 20\}$  to  $K \in \{5, \dots, 80\}$ . First, observing the  $\tilde{F}_3$ , there is not much to be gleaned from this plot; unfortunately, the test doesn't seem to find any coherent results other than a significant amount of detections around the 1.55 mHz frequency. In comparison, the  $F'_4$  test plot does show some significant frequencies that are consistent for the majority of different  $K$  tested, such as  $f = 4.6$  mHz.

The final test to conduct is the aggregate test, presented in Fig. 6.8. The aggregate test using  $K \in \{5, \dots, 80\}$ , with the upper bound approximation for the distribution given by Eqn. (4.9), finds only two significant frequencies at  $f = 1.55$ , and 3.86. This

result was expected due to the lower Type I error used in the aggregate test. It has the advantage that there is no debate as to what number of tapers to use, as you will get only one set of significant frequencies. That being said, due to the large Type I error, one can not help but feel like significant frequencies that could have been detected (and explored, scientifically), have been removed due to the upper bound of the distribution.

With all this being said, it would take an expert in helioseismology to dissect these test statistics and decide which one produced the most reliable detection of true significant modulated frequencies, and the physical science explanation of the results. We note in passing that knowledge of these frequencies informs complex physics models of the solar core rotation, which is still unsettled science. We do find it interesting that there are some commonalities between the test statistics, as they all detected that 1.55 mHz is a significant frequency: this is around 10.75 minutes, which may have a physical interpretation in terms of the solar core dynamics.

## 7. Conclusion and Future Work

### 7.1 Conclusion

In conclusion, we have created three new test statistics. We first reviewed the derivation for the statistic created in [3] and discussed the multitaper framework, including two different sets of commonly used tapers. By changing from the DPSS tapers to the Sinusoidal tapers, we were able to create a weighting scheme that, under simulation, allows for an increase in the probability of detection for a greater range of  $K$ . We derived the distribution for the weighted  $F_4$  and showed that the different set of tapers in combination with arbitrary weights does not change the distribution under the null compared to the previous  $\tilde{F}_3$  test.

An isolation of the last taper's contribution to the  $F_4$  test was also derived. By removing this last term for even  $K$ , in both  $\tilde{F}_3$  and  $F_4$ , we discovered an increase in detection probability. A new test  $F'_4$  was created, which implemented this removal process for only even  $k$  and combined with the weights used in  $F_4$ . Under testing, we found that depending on the type of modulation, the choice of  $K$  is still important to the functionality of the test; thus, we wanted to make a combination of tests in which a range of tapers was used. The aggregate test was created, and an upper bound was found for the distribution; we discovered under simulation it suffers due to the upper bound, and we showed that if this bound could be improved, the aggregate does work better.

We also applied the newly created test statistics to the SoHO GOLF data set and compared them to the previously created statistics. Lastly, we created an R package that will aid in the future use of these test statistics. This package makes use of [23], [29], [35], [9], and [21], with built-in packages `parallel`, `splines` and `stats4`.

## 7.2 Future Work

The simulations conducted in Chapter 5 demonstrate the potential power and limitations of the aggregate test. Although the upper bound created does allow for the test to be conducted, it is quite clear that a better upper bound needs to be found. One potential way to improve the upper bound is to use the method found in [14]. Another alternative is to use the Conway-Maxwell-Binomial Distribution, in [15], to find an approximation to the true distribution of the aggregate test.

An alternative approach would be to improve the  $F'_4$  test statistic first. As the aggregate test is based on  $F'_4$ , it will benefit as well. This could be done by improving the set of weights used on the tapers. In our study, we only looked at two types of weights; the first being the one included in this thesis and the alternative being a step function that would down-weight any tapers above a certain value to zero. It was not included here due to its poor performance. It is possible that looking at a less heavy down-weighting scheme or more of a linear down-weight would assist with an increased probability of detection.

It would be of interest to create an estimate for the modulation bandwidth. This could help with choosing the correct number of tapers in  $F'_4$  or using a different range of tapers based on the bandwidth in the aggregate test. One possible way to do this is through the instantaneous frequency. When plotting the instantaneous frequency at the carrier frequency  $f$ , we have seen some evidence that the range corresponds to the modulation bandwidth. The main issue with this method is that you need to

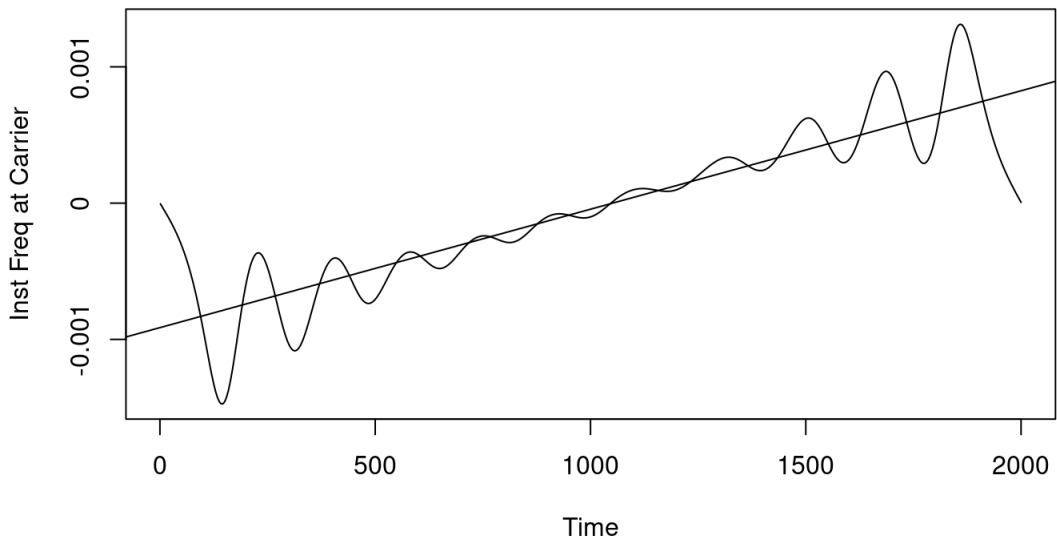


Figure 7.1: Instantaneous frequency at carrier frequency overlaid with linear model estimated slope for a linear modulation bandwidth of 0.001

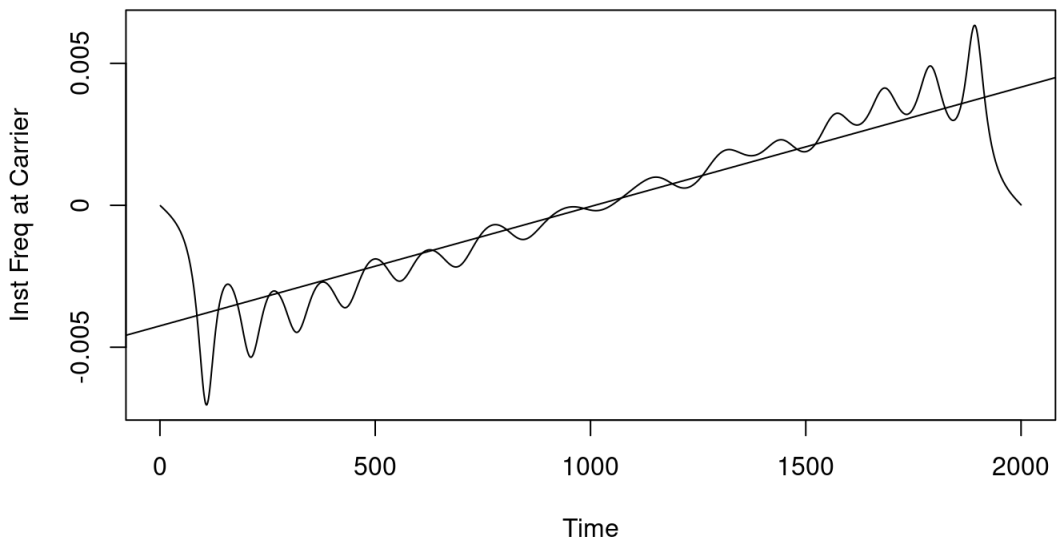


Figure 7.2: Instantaneous frequency at carrier frequency overlaid with linear model estimated slope for a linear modulation bandwidth of 0.005

know the carrier frequency before you can find the modulation bandwidth. If one knows the carrier frequency, we can see from the two examples (shown in Fig. 7.1 and Fig. 7.2) that in both cases, the modulation bandwidth is around half the range of the estimated linear model slope. In addition, from Fig. 5.18, discussed in Chapter 5, by getting a rough estimate of the modulation bandwidth it would be possible to adjust the weights accordingly to improve the performance of the test statistics covered in this thesis.

Although we did discover that when  $K$  is even, the last taper is having the most negative effect on the test statistic, we were never able to pinpoint the specific contribution from the last term causing the degradation in the test statistic. This is a very high priority for future work, as being able to directly correct for this factor would have immediate gains in performance across all  $K$ .

An additional area of future work is in the type of modulation that can be explored. In this thesis, we only looked at the case of linear modulation; other types of modulation, such as quadratic polynomial modulation, could behave differently. It would be interesting to see how the new test statistics perform against the  $\tilde{F}_3$ .

Finally, the idea of consistency of the tests could be explored. One could study how the probability of detection changes as  $N$  and  $K$  approach  $\infty$ . This could be examined through a fixed selection of modulation degree  $P$  and fixed modulation width  $m_P$ . One could also examine how the modulation width of polynomial structure affects the consistency of the test, i.e. if the probability of detection approaches one as  $N$  and  $K$  approach  $\infty$ .

# Bibliography

- [1] Kian Blanchette. Multitaper statistical tests for the detection of frequency-modulated signals. Master's thesis, Queen's University, 2020.
- [2] Kian Blanchette. A collection of code from multitaper statistical test for detection of frequency-modulated signals (on github), 2021. <https://github.com/kianBlanchette/Thesis/>.
- [3] Kian Blanchette, Wesley S Burr, and Glen Takahara. An F-test for polynomial frequency modulation. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pages 5010–5014, 2021.
- [4] Patrick Boumier. GOLF SoHO: How to Access GOLF Data. <https://www.ias.u-psud.fr/golf/templates/access.html>.
- [5] David R Brillinger. *Time Series: Data Analysis and Theory*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 2001.
- [6] Peter J Brockwell and Richard A Davis. *Time Series: Theory and Methods*. Springer, 1991.
- [7] Ferhat Çakrak and Patrick J Loughlin. Multiple window time-varying spectral analysis. *IEEE Transactions on Signal Processing*, 49(2):448–453, 2001.
- [8] Chris Chatfield and Haipeng Xing. *The Analysis of Time Series: An Introduction with R*. Chapman & Hall/CRC Texts in Statistical Science. CRC Press, 2019.
- [9] Hadley Wickham et al. *ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*, 2023. R package version 3.4.2.

- [10] A.H. Gabriel, G. Grec, J. Charra, J-M Robillot, T. Roca Cortés, S. Turck-Chièze, R. Bocchia, P. Boumier, M. Cantin, E. Cespédes, et al. Global oscillations at low frequency from the SOHO mission (GOLF). In *The SOHO Mission*, pages 61–99. Springer, 1995.
- [11] RA García, A Jiménez, S Mathur, J Ballot, A Eff-Darwich, SJ Jiménez-Reyes, PL Pallé, J Provost, and S Turck-Chièze. Update on g-mode research. *Astronomische Nachrichten: Astronomical Notes*, 329(5):476–484, 2008.
- [12] Per Gloersen and Norden Huang. Comparison of interannual intrinsic modes in hemispheric sea ice covers and other geophysical parameters. *IEEE Transactions on Geoscience and Remote Sensing*, 41(5):1062–1074, 2003.
- [13] Arthur P Guillaumin, Adam M Sykulski, Sofia C Olhede, Jeffrey J Early, and Jonathan M Lilly. Analysis of non-stationary modulated time series with applications to oceanographic surface flow measurements. *Journal of Time Series Analysis*, 38(5):668–710, 2017.
- [14] David Hunter. An upper bound for the probability of a union. *Journal of Applied Probability*, 13(3):597–603, 1976.
- [15] Joseph B Kadane. Sums of Possibly Associated Bernoulli Variables: The Conway–Maxwell–Binomial Distribution. *Bayesian Analysis*, 11(2), 2016.
- [16] Samuel Kotz, Tomasz Kozubowski, and Krzysztof Podgórski. *The Laplace distribution and generalizations*. Number 183. Springer Science & Business Media, 2001.
- [17] Gunter D. Krebs. Soho - gunter’s space page. [https://space.skyrocket.de/doc\\_sdat/soho.htm](https://space.skyrocket.de/doc_sdat/soho.htm), 2023. Accessed: July 15, 2023.
- [18] NASA. What are Lagrange Points? – NASA Solar System Exploration. <https://solarsystem.nasa.gov/faq/88/what-are-lagrange-points/>, 2023. Accessed: July 15, 2023.
- [19] Donald B Percival and Andrew T Walden. *Spectral Analysis for Univariate Time*

- Series*. Cambridge University Press, mar 2020.
- [20] James W Pitton. Adapting multitaper spectrograms to local frequency modulation. In *Proceedings of the Tenth IEEE Workshop on Statistical Signal and Array Processing (Cat. No. 00TH8496)*, pages 108–112. IEEE, 2000.
  - [21] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2023.
  - [22] Humera Rafique. Simulation of harmonic analysis, synthesis and Gibbs effect of periodic signals. In *2019 16th International Multi-Conference on Systems, Signals & Devices (SSD)*, pages 282–287. IEEE, 2019.
  - [23] Karim Rahim and Wesley S Burr. *multitaper: Spectral Analysis Tools using the Multitaper Method*, 2023. R package version 1.0-17.
  - [24] Lord Rayleigh. XXI. On the spectrum of an irregular disturbance. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 5(26):238–243, 1903.
  - [25] Kurt S. Riedel and Alexander Sidorenko. Minimum bias multiple taper spectral estimation. *IEEE Transactions on Signal Processing*, 43(1):188–195, Jan 1995.
  - [26] Arthur Schuster. On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenonema. *Terrestrial Magnetism*, 3:13-41, 1898.
  - [27] D. Slepian. Prolate Spheroidal Wave Functions, Fourier Analysis, and Uncertainty—V: The Discrete Case. *Bell System Technical Journal*, 57(5):1371–1430, 1978.
  - [28] Alexander Rawson Stokes. A numerical Fourier-analysis method for the correction of widths and shapes of lines on X-ray powder photographs. *Proceedings of the Physical Society*, 61(4):382, 1948.
  - [29] Cleve Moler Thomas Yee. *VGAM: Vector Generalized Linear and Additive Models*, 2023. R package version 1.1-8.

- [30] David J Thomson. Spectrum Estimation and Harmonic Analysis. *Proceedings of the IEEE*, 70(9):1055–1096, 1982.
- [31] David J Thomson. Inverse-constrained projection filters. In *In Wavelets: Applications in Signal and Image Processing IX*, volume 4478. SPIE, 2001.
- [32] David J Thomson. Spectra of MDI magnetograms. In D. Danesy, editor, *Proceedings of the SOHO 14/GONG 2004 Workshop (ESA SP-559). Helio-and Asteroseismology: Towards a Golden Future. 12-16 July, 2004. New Haven, Connecticut, USA*, volume 559, page 91, 2004.
- [33] David J Thomson. Polynomial phase demodulation in multitaper analysis. *IEEE Workshop on Statistical Signal Processing Proceedings*, pages 401–404, 2009.
- [34] Nergis Tomen and Jan C van Gemert. Spectral leakage and rethinking the kernel size in CNNs. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5138–5147, 2021.
- [35] Hadley Wickham. *reshape2: Flexibly Reshape Data: A Reboot of the Reshape Package*, 2020, R package version 1.4.4.