Université de Rennes I
Master SISEA
20 January 2022

TP: Détection - Estimation
Report

AUTHOR:
Kevin MICHALEWICZ
<kevin.michalewicz@etudiant.univ-rennes1.fr>

TEACHER:
Prof. Di GE

Contents

1 Introduction

Given $x(n) = A\cos(2\pi f_0 n + \phi) + w(n)$ for $n = 0, ..., N-1$, it can be noticed that $w(n)$, a centered white gaussian noise with variance σ^2, is the only source of randomness present. A and ϕ are unknown but deterministic parameters and the goal is to estimate them. As for f_0, it is known and deterministic, but its possible values need to be specified.

The observation under consideration can be thought of as a sampled continuous time signal (frequency f) using a sampling frequency f_s. Due to Whittaker-Nyquist-Kotelnikov-Shannon sampling theorem, $f_s \geq 2f$, so with $f_0 \triangleq \frac{f}{f_s}$ it can be written that $0 < f_0 \leq \frac{1}{2}$ are the allowed values for the dimensionless quantity f_0.

2 Theory

2.1 Exercise 1

As the only stochastic source is a centered white gaussian noise $w(n)$, the $x(n)$ at different times are not correlated. In particular, a vector X_N containing them follows a gaussian distribution with mean $A\cos(2\pi f_0 n + \phi)$ and covariance matrix $\sigma^2 I$. The log-likelihood function can be written as

$$L(A, \phi) = log\, p(\{x(n)\}_{n=0,...,N-1} | A, \phi)$$

The observations are independent for the reasons mentioned in the previous paragraph, so

$$L(A, \phi) = log \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^N exp \left(-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} [x(n) - A\cos(2\pi f_0 n + \phi)]^2 \right)$$

And finally

$$L(A, \phi) = -\frac{N}{2} log \left(2\pi\sigma^2 \right) - \frac{1}{2\sigma^2} \sum_{n=0}^{N-1} [x(n) - A\cos(2\pi f_0 n + \phi)]^2$$

2.2 Exercise 2

The idea is to differentiate the log-likelihood function with respect to each of A and ϕ separately and set them equal to zero. For example, to obtain $\hat{A}_{ML}$:

$$\frac{\partial L(A, \phi)}{\partial A} = \frac{1}{2\sigma^2} \sum_{n=0}^{N-1} 2\, \cos(2\pi f_0 n + \phi) [x(n) - A\cos(2\pi f_0 n + \phi)] \Big|_{A=\hat{A}_{ML}} = 0$$

Just by rearranging the terms

$$\hat{A}_{ML} = \frac{\sum_{n=0}^{N-1} x(n)\cos(2\pi f_0 n + \phi)}{\sum_{n=0}^{N-1} \cos^2(2\pi f_0 n + \phi)}$$

But as N is big enough $\sum_{n=0}^{N-1} \cos^2(2\pi f_0 n + \phi) \approx \frac{N}{2}$ and the expected result is obtained

$$\hat{A}_{ML} \approx \frac{2}{N} \sum_{n=0}^{N-1} x(n)\cos(2\pi f_0 n + \phi)$$

To get $\hat{\phi}_{ML}$ the idea is similar but some additional properties need to be considered:

$$\frac{\partial L(A,\phi)}{\partial \phi} = -\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} 2\left[x(n) - A\cos(2\pi f_0 n + \phi)\right] A\sin(2\pi f_0 n + \phi)\bigg|_{\phi=\hat{\phi}_{ML}} = 0$$

Since $\sum_{n=0}^{N-1} \sin(2\pi f_0 n + \phi)\cos(2\pi f_0 n + \phi) \approx 0$, the previous expression simplifies to

$$\sum_{n=0}^{N-1} x(n)\sin(2\pi f_0 n + \hat{\phi}_{ML}) = 0$$

Using the trigonometric property of the sine of the sum of angles, namely $\sin(\alpha + \beta) = \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta)$

$$\sum_{n=0}^{N-1} x(n)\left[\sin(2\pi f_0 n)\cos(\hat{\phi}_{ML}) + \cos(2\pi f_0 n)\sin(\hat{\phi}_{ML})\right] = 0$$

Just by rearranging the terms

$$\sin(\hat{\phi}_{ML}) \sum_{n=0}^{N-1} x(n)\cos(2\pi f_0 n) = -\cos(\hat{\phi}_{ML}) \sum_{n=0}^{N-1} x(n)\sin(2\pi f_0 n)$$

Which gives the final result

$$\hat{\phi}_{ML} = \arctan\left\{-\frac{\sum_{n=0}^{N-1} x(n)\sin(2\pi f_0 n)}{\sum_{n=0}^{N-1} x(n)\cos(2\pi f_0 n)}\right\}$$

2.3 Exercise 3

The goal of this exercise is to find the conditions for which $\hat{\phi}_{ML} \xrightarrow{\text{a.s.}} \phi$, i.e. convergence in probability. This can also be written as $\lim_{N\to+\infty} \mathbb{P}[\hat{\phi}_{ML} = \phi] = 1$. As the arctan function is monotone and continuous, this is equivalent to considering $\tan(\hat{\phi}_{ML}) \xrightarrow{\text{a.s.}} \tan(\phi)$.

Let $X_N = \begin{bmatrix} X_0 & \dots & X_{N-1} \end{bmatrix}^T$, $S_N = \begin{bmatrix} \sin(2\pi f_0 0) & \dots & \sin(2\pi f_0(N-1)) \end{bmatrix}^T$ and $C_N = \begin{bmatrix} \cos(2\pi f_0 0) & \dots & \cos(2\pi f_0(N-1)) \end{bmatrix}^T$ be three column vectors used to simplify the notation. Then, the phase maximum likelihood estimator expression can be reduced to

$$\tan(\hat{\phi}_{ML}) = -\frac{X_N^T S_N}{X_N^T C_N}$$

Let's use the trigonometric property of the cosine of the sum of angles with $x(n)$

$$x(n) = A\left[\cos(2\pi f_0 n)\cos(\phi) - \sin(2\pi f_0 n)\sin(\phi)\right] + w(n)$$

Using the previous expression:

$$X_N^T C_N = A\cos(\phi)\sum_{n=0}^{N-1}\cos^2(2\pi f_0 n) - A\sin(\phi)\sum_{n=0}^{N-1}\sin(2\pi f_0 n)\cos(2\pi f_0 n) + \sum_{n=0}^{N-1} w(n)\cos(2\pi f_0 n)$$

Using once again that $\sum_{n=0}^{N-1}\cos^2(2\pi f_0 n) \approx \frac{N}{2}$ and that $\sum_{n=0}^{N-1}\sin(2\pi f_0 n)\cos(2\pi f_0 n) \approx 0$, the following simplification is made

$$X_N^T C_N = A\cos(\phi)\frac{N}{2} + \sum_{n=0}^{N-1} w(n)\cos(2\pi f_0 n)$$

By proceeding in an analogous way for the $X_N^T S_N$ term:

$$X_N^T S_N = -A\sin(\phi)\frac{N}{2} + \sum_{n=0}^{N-1} w(n)\sin(2\pi f_0 n)$$

Putting everything together

$$\tan(\hat{\phi}_{ML}) = \frac{A\sin(\phi)\frac{N}{2} - \sum_{n=0}^{N-1} w(n)\sin(2\pi f_0 n)}{A\cos(\phi)\frac{N}{2} + \sum_{n=0}^{N-1} w(n)\cos(2\pi f_0 n)}$$

It can be pointed out that the first term of both the numerator and denominator are deterministic, while the second one is a source of randomness in both cases. Furthermore, if the latter two terms were negligible, $\tan(\phi)$ would be obtained. This calls for an in-depth study.

By defining $V_1 \triangleq -\sum_{n=0}^{N-1} w(n)\sin(2\pi f_0 n)$ with $V_1 \sim \mathcal{N}(0, \frac{N\sigma^2}{2})$ and $V_2 \triangleq \sum_{n=0}^{N-1} w(n)\cos(2\pi f_0 n)$ with $V_2 \sim \mathcal{N}(0, \frac{N\sigma^2}{2})$, their non-correlation (and thus, their independence) can be proved as follows

$$\mathbb{E}[V_1 V_2] = \mathbb{E}\left[\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} w(n)w(m)sin(2\pi f_0 n)cos(2\pi f_0 m)\right]$$

Except for white Gaussian noise, all factors are deterministic, so:

$$\mathbb{E}[V_1V_2] = \sum_{n=0}^{N-1}\sum_{m=0}^{N-1} \mathbb{E}\left[w(n)w(m)\right] sin(2\pi f_0 n) cos(2\pi f_0 m)$$

If $n \neq m$, then noises are in dependant and $\mathbb{E}\left[w(n)w(m)\right] = \mathbb{E}\left[w(n)\right][w(m)] = 0$. For the case in which $n = m$, $\mathbb{E}\left[w(n)w(m)\right] = \mathbb{E}\left[w(n)^2\right] = \sigma^2$. Thus

$$\mathbb{E}[V_1V_2] = \sigma^2 \sum_{n=0}^{N-1} sin(2\pi f_0 n) cos(2\pi f_0 n) \approx 0$$

.

In order for these terms to be negligible, the 3σ criterion is used. In this case, $3\sigma = 3\sqrt{\frac{N\sigma^2}{2}}$, so the conditions become

$$\begin{cases} \left|A\cos(\phi)\frac{N}{2}\right| \gg 3\sqrt{\frac{N\sigma^2}{2}} \\ \left|A\sin(\phi)\frac{N}{2}\right| \gg 3\sqrt{\frac{N\sigma^2}{2}} \end{cases} .$$

This means that N is the key parameter to tune in order to guarantee $\hat{\phi}_{ML} \xrightarrow{\text{a.s.}} \phi$. The case in which $\cos(\phi) \approx 0$ should be avoided as this previous condition will never hold. However, there is not problem with $\sin(\phi) \approx 0$.

2.4 Exercise 4

In general, a given estimator $\hat{\theta}$ is unbiased if $b(\hat{\theta}, \theta) = \theta - \mathbb{E}[\hat{\theta}] = 0$. A reasonable idea is to compute the mean of A's estimator

$$\mathbb{E}[\hat{A}] = \mathbb{E}\left[\frac{2}{N}\sum_{n=0}^{N-1} x(n)\cos(2\pi f_0 n + \hat{\phi}_{ML})\right]$$

Writing $x(n)$ explicitly

$$\mathbb{E}[\hat{A}] = \frac{2}{N}\mathbb{E}\left[\sum_{n=0}^{N-1} A\cos^2(2\pi f_0 n + \hat{\phi}_{ML})\right] + \frac{2}{N}\mathbb{E}\left[\sum_{n=0}^{N-1} w(n)\cos(2\pi f_0 n + \hat{\phi}_{ML})\right]$$

Remembering that $\mathbb{E}[w(n)] = 0$

$$\mathbb{E}[\hat{A}] = \frac{2}{N}A\frac{N}{2} + \frac{2}{N}\sum_{n=0}^{N-1} \cancel{\mathbb{E}[w(n)]}\cos(2\pi f_0 n + \hat{\phi}_{ML}) = A$$

2.5 Exercise 5

By definition, the Fisher Information Matrix can be obtained as $J \triangleq -\mathbb{E}\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j}L(\theta)\right]$. Therefore, the log-likelihood second order partial derivatives will be computed in this exercise

$$\frac{\partial^2 L(A,\phi)}{\partial A^2} = -\frac{1}{\sigma^2}\sum_{n=0}^{N-1}\cos^2(2\pi f_0 n+\phi) \approx -\frac{N}{2\sigma^2}$$

$$\mathbb{E}\left[\frac{\partial^2 L(A,\phi)}{\partial \phi^2}\right] = \mathbb{E}\left[-\frac{A^2}{\sigma^2}\sum_{n=0}^{N-1}\sin^2(2\pi f_0 n+\phi)\right] + 0 \approx -\frac{A^2 N}{2\sigma^2}$$

$$\mathbb{E}\left[\frac{\partial^2 L(A,\phi)}{\partial A\partial \phi}\right] = \mathbb{E}\left[\frac{\partial^2 L(A,\phi)}{\partial \phi\partial A}\right] = \mathbb{E}\left[\frac{1}{\sigma^2}\sum_{n=0}^{N-1} 2\sin(2\pi f_0 n+\phi)\cos(2\pi f_0 n+\phi)\right] \approx 0$$

The result is the following 2x2 diagonal matrix:

$$J(A,\phi) = \begin{bmatrix} \frac{N}{2\sigma^2} & 0 \\ 0 & \frac{A^2 N}{2\sigma^2} \end{bmatrix}$$

As the unbiased case is considered, i.e. $\frac{\partial b(\theta)}{\partial \theta} = 0$, the Cramer-Rao lower bound can be computed in a straightforward way

$$\text{CRLB} = J^{-1}(A,\phi) = \begin{bmatrix} \frac{2\sigma^2}{N} & 0 \\ 0 & \frac{2\sigma^2}{A^2 N} \end{bmatrix}$$

As $J^{-1}(A,\phi)$ is diagonal each entry is lower bounded independantly.

3 MATLAB simulations

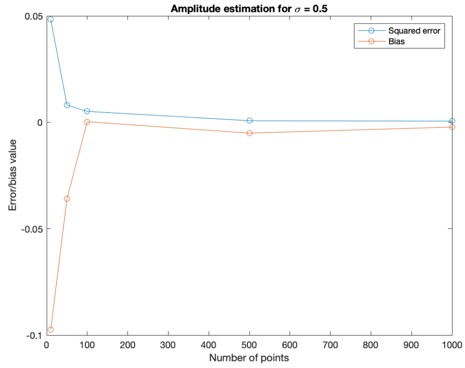
3.1 Exercise 6

An observation sequence $\{x(n)\}_{n=0,\ldots,N-1}$ was generated according to the expression given in the Introduction with $A = 1$, $f_0 = \frac{1}{5}$ and $\phi = 1$. Five different observation sizes N (10, 50, 100, 500, 1000) and three noise standard deviation σ values (0.25, 0.5, 1) were considered. Also, $K = 20$ realizations were obtained for each N and σ in order to compute the empirical means and empirical squared errors of the estimators.

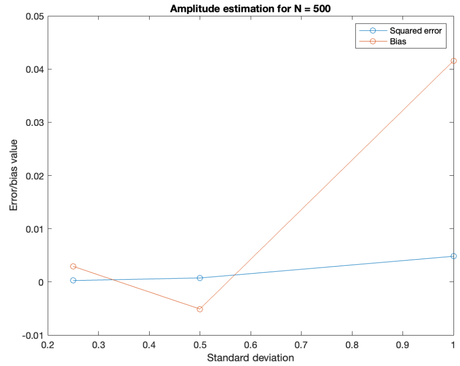
3.2 Exercise 7

$\hat{A}_{ML}$ and $\hat{\phi}_{ML}$ were computed according to the results from Exercise 2. However, $\hat{A}_{ML}$ was used to obtain $\hat{\phi}_{ML}$. Afterwards, the biases $b(\hat{A}_{ML}, A) = A - \mathbb{E}_{\text{EMP}}[\hat{A}_{ML}]$ and $b(\hat{\phi}_{ML}, \phi) = \phi - \mathbb{E}_{\text{EMP}}[\hat{\phi}_{ML}]$ were calculated for each of the above-mentioned values of N and σ, being $\mathbb{E}_{\text{EMP}}[.]$ the empirical mean, i.e. $\mathbb{E}_{\text{EMP}}[\hat{\theta}] = \frac{1}{K}\sum_{k=1}^{K}\hat{\theta}_k$. In the previous expression K is the number of realizations and $\hat{\theta}_k$ corresponds to each of the k^{th} estimator. Also, the squared errors were given by $C(\hat{A}_{ML}, A) = \mathbb{E}_{\text{EMP}}[(A - \hat{A}_{ML})^2] \approx \frac{1}{K}\sum_{k=1}^{K}(A - \hat{A}_{ML_k})^2$ and $C(\hat{\phi}_{ML}, \phi) = \mathbb{E}_{\text{EMP}}[(\phi - \hat{\phi}_{ML})^2] \approx \frac{1}{K}\sum_{k=1}^{K}(\phi - \hat{\phi}_{ML_k})^2$.

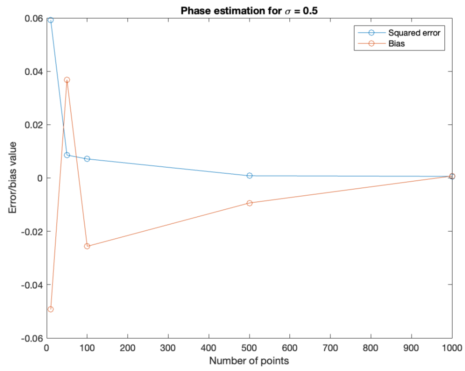
In Figure 3.1a, $b(\hat{A}_{ML}, A)$ and $C(\hat{A}_{ML}, A)$ are plotted for a fixed $\sigma = 0.5$ as a function of N. An overall decrease of the value of these metrics can be



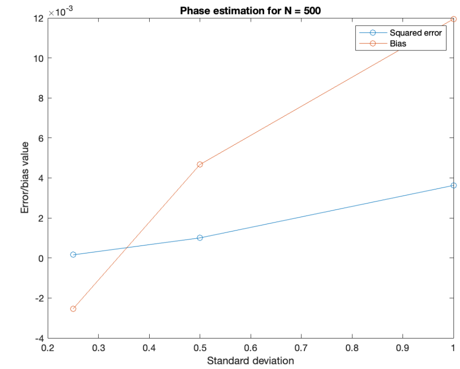
(a) Amplitude estimation - Bias and squared error as a function of N with a constant $\sigma = 0.5$.



(b) Amplitude estimation - Bias and squared error as a function of σ with a constant $N = 500$.



(c) Phase estimation - Bias and squared error as a function of N with a constant $\sigma = 0.5$.



(d) Phase estimation - Bias and squared error as a function of σ with a constant $N = 500$.

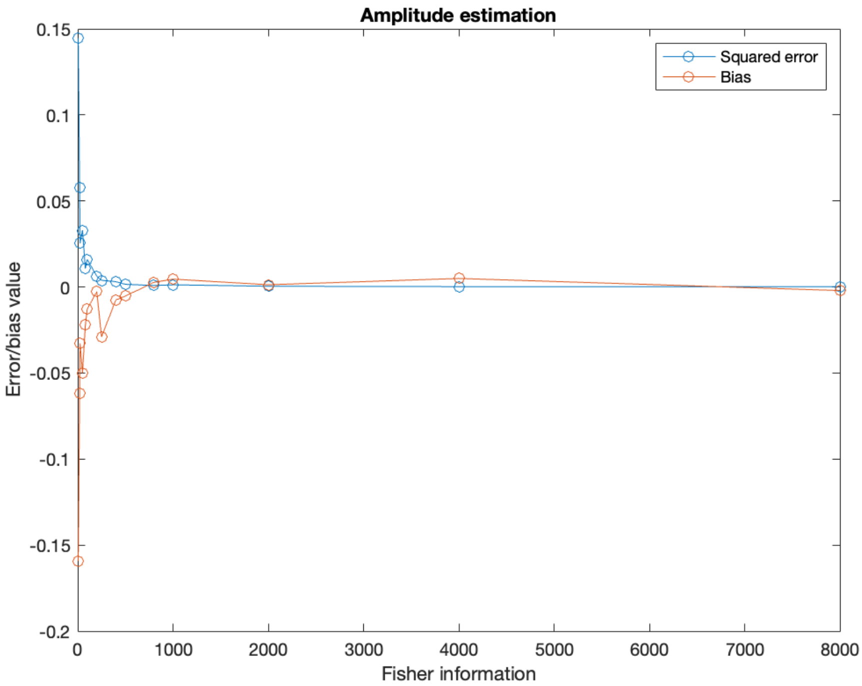
Figure 3.1: Exercise 7 figures.

ascertained as N increases. The same idea is exhibited in Figure 3.1c. For the bias, this corresponds to the result obtained for Exercise 4. On the other hand, in Figures 3.1b and 3.1d the opposite tendency is shown: both bias and squared error are prone to increase with a higher noise standard deviation σ. This last result makes sense as noise masks the useful signal and its parameters cannot be estimated as accurately as in the absence of it.

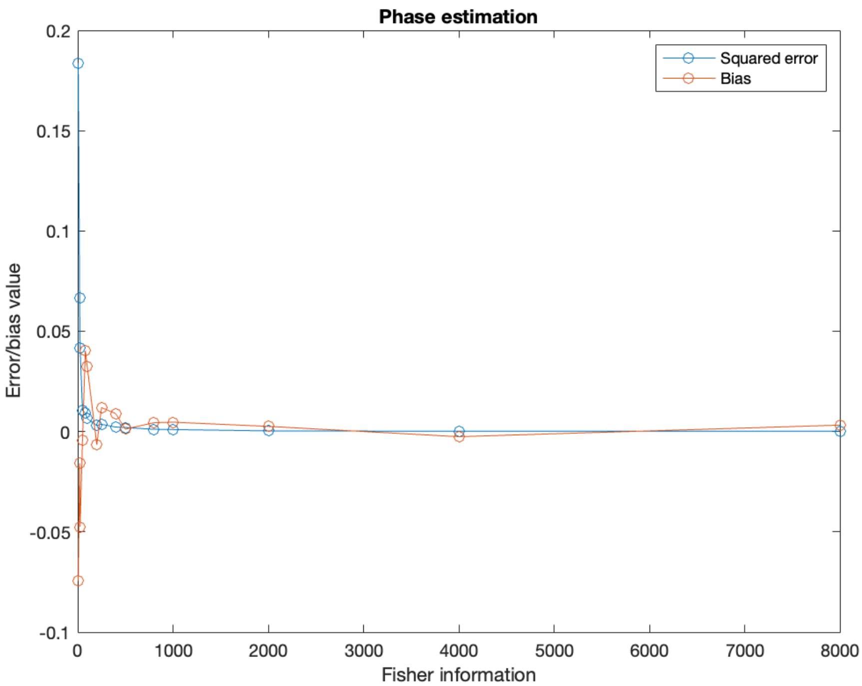
3.3 Exercise 8

In this exercise the biases and squared errors were plotted as a function of a more appropriate quantity: the Fisher Information. Due to the fact that five N and three σ cases are considered, there are fifteen Fisher Information values for A and fifteen for ϕ (all the possible combinations). In other words, for each pair (N, σ) bias and squared error values are associated to a particular $J_{11}(A, \phi)$ or $J_{22}(A, \phi)$, corresponding to the estimation of A and ϕ respectively. In both Figures 3.2a and 3.2b the bias and the squared errors decrease as the Fisher Information becomes larger. Seeing how $J(A, \phi)$ is written, this synthesizes the same dependencies analyzed in

Exercise 7.



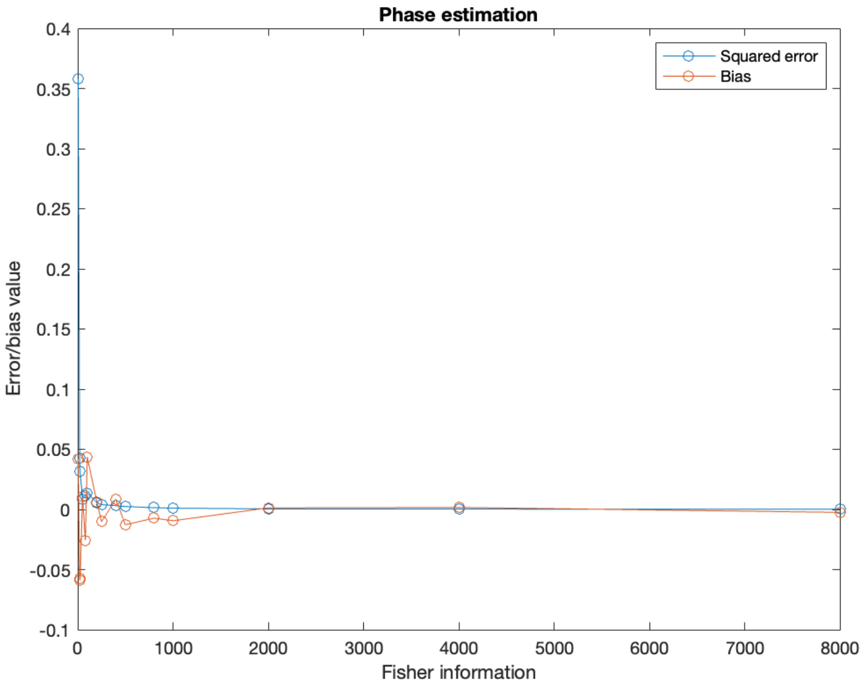
(a) Amplitude estimation - Bias and squared error as a function of the Fisher Information.



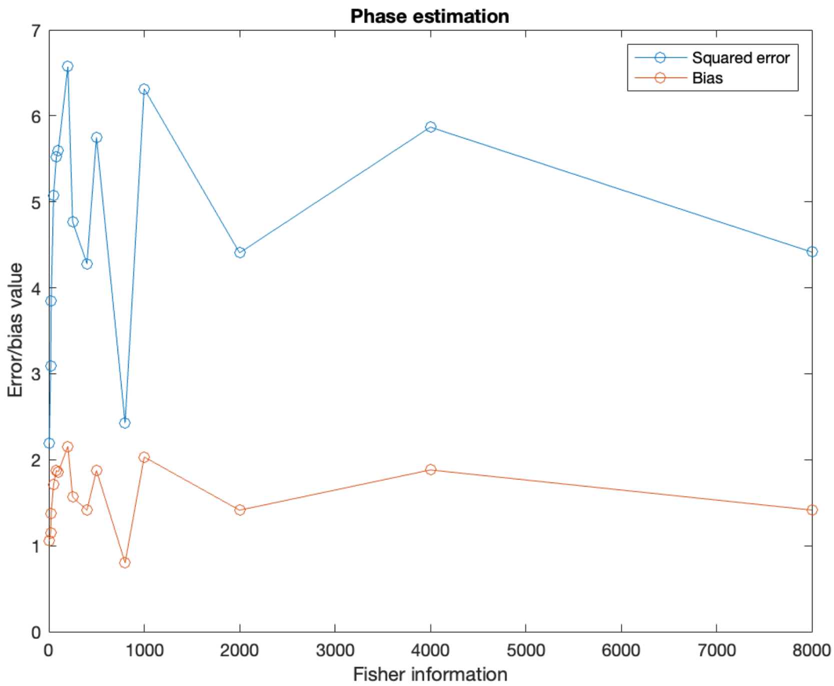
(b) Phase estimation - Bias and squared error as a function of the Fisher Information.

Figure 3.2: Exercise 8 figures.

In Exercise 3 it was stated that the condition $\sin(\phi) \approx 0$ was not a problem, while $\cos(\phi) \approx 0$ needed to be avoided. Figures 3.3a and 3.3b (respectively) show this result for the phase estimation.



(a) Phase estimation when $\sin(\phi) \approx 0$



(b) Phase estimation when $\cos(\phi) \approx 0$

Figure 3.3: Figures that make reference to Exercise 3.

4 Conclusions

In conclusion, this work first made it possible to make theoretical derivations concerning an estimation problem and then to analyze whether the results were consistent in practice. The outcome is satisfactory because the squared error and bias behave as expected as a function of N, σ and the Fisher Information. In particular, for a high number of points and a low noise level - and thus a high $J(A, \phi)$ -, they tend to zero experimentally.

5 Appendix - MATLAB Code

```
% Detection and Estimation - Simulations corresponding to
    exercises 6 to 8
% Master SISEA
% By Kevin Michalewicz
clear all
close all
clc

N_vec = [10 50 100 500 1000];
A = 1;
sigma_vec = [0.25 0.5 1];
freq0 = 0.2; % between 0 and 0.5
phi = 1;
K = 20; % how many estimators for each (sigma, N)
phi_ML = zeros(1,K);
A_ML = zeros(size(phi_ML));
A_squared_error = zeros(length(sigma_vec),length(N_vec));
phi_squared_error = zeros(size(A_squared_error));
A_bias = zeros(length(sigma_vec),length(N_vec));
phi_bias = zeros(size(A_bias));

N_idx = 4;
sigma_idx = 2;
A_fisher = reshape((1./(2*sigma_vec.^2))'*N_vec, [], 1);
[~,fisher_order] = sort(A_fisher); % getting sortIdx (
    increasing order)
phi_fisher = A_fisher*A^2;

for i=1:length(sigma_vec)
    sigma = sigma_vec(i);
    for j=1:length(N_vec)
        for k=1:K
            N = N_vec(j);
            n = 1:N;
            S_N = sin(2*pi*freq0*n);
```

```
                C_N = cos(2*pi*freq0*n);
                X_N = A*cos(2*pi*freq0*n+phi) + randn(1,N)*sigma;
                phi_ML(1,k) = atan(-X_N*S_N'/(X_N*C_N'));
                A_ML(1,k) = 2*mean(X_N.*cos(2*pi*freq0*n+phi_ML(1,k)));
            end
            A_squared_error(i,j) = mean((A - A_ML).^2);
            A_bias(i,j) = A - mean(A_ML);
            phi_squared_error(i,j) = mean((phi - phi_ML).^2);
            phi_bias(i,j) = phi - mean(phi_ML);
        end
end

% Plotting part of exercise 7

figure(1)
plot(N_vec, A_squared_error(sigma_idx,:), '-o')
hold on
plot(N_vec, A_bias(sigma_idx,:), '-o')
legend('Squared error', 'Bias');
title("Amplitude estimation for \sigma = " + sigma_vec(sigma_idx));
xlabel("Number of points");
ylabel("Error/bias value");

figure(2)
plot(sigma_vec, A_squared_error(:,N_idx), '-o')
hold on
plot(sigma_vec, A_bias(:,N_idx), '-o')
legend('Squared error', 'Bias');
title("Amplitude estimation for N = " + N_vec(N_idx));
xlabel("Standard deviation");
ylabel("Error/bias value");

figure(3)
plot(N_vec, phi_squared_error(sigma_idx,:), '-o')
hold on
plot(N_vec, phi_bias(sigma_idx,:), '-o')
legend('Squared error', 'Bias');
title("Phase estimation for \sigma = " + sigma_vec(sigma_idx));
xlabel("Number of points");
ylabel("Error/bias value");

figure(4)
```

```
plot(sigma_vec, phi_squared_error(:,N_idx), '-o')
hold on
plot(sigma_vec, phi_bias(:,N_idx), '-o')
legend('Squared error', 'Bias');
title("Phase estimation for N = " + N_vec(N_idx));
xlabel("Standard deviation");
ylabel("Error/bias value");

% Exercise 8

% Flattening the quadratic errors and bias
A_squared_error = reshape(A_squared_error,[],1);
A_bias = reshape(A_bias,[],1);
phi_squared_error = reshape(phi_squared_error,[],1);
phi_bias = reshape(phi_bias,[],1);

figure(5)
plot(A_fisher(fisher_order), A_squared_error(fisher_order),
    '-o')
hold on
plot(A_fisher(fisher_order), A_bias(fisher_order), '-o')
legend('Squared error', 'Bias');
title("Amplitude estimation");
xlabel("Fisher information");
ylabel("Error/bias value");

figure(6)
plot(phi_fisher(fisher_order), phi_squared_error(
    fisher_order), '-o')
hold on
plot(phi_fisher(fisher_order), phi_bias(fisher_order), '-o'
    )
legend('Squared error', 'Bias');
title("Phase estimation");
xlabel("Fisher information");
ylabel("Error/bias value");
```