On the Cramer Rao Bound in determining scattering center parameters using high resolution radar

Anna Genell and David Lösaus

Supervisors:
Roland Jonsson
Ericsson Microwave Systems AB
and
Mats Viberg
Chalmers University of Technology

Examiner:
Patrik Albin
Department of Mathematical Statistics
University of Gothenburg and Chalmers University of Technology

April 21, 2004

Abstract

The radar transmitted electro-magnetic high frequency signal is scattered from objects - targets. In automatic target recognition the detected return signal can be modeled as scattering centers. Scattering centers represent dominant reflectors on the targets, such as corners or flat surfaces. Mathematically, the return signal can be modeled as a sum of complex sinusoids. In the more general, two-dimensional, case the model must take care of the amplitude aspect dependency and the range aspect dependency. For the range aspect dependency a second order polynomial has been used.

This master thesis in mathematical statistics presents results for some scattering center models, where polynomials have been used as models for the aspect dependent amplitudes. The theoretical amplitude approximation and estimation errors have been examined for different cases and have been used for determination of the amplitude approximation polynomial polynomial order. The thesis also include a comparison between the use of a simple polynomial type and a set of orthogonal polynomials. Besides the amplitude estimation error, we have also studied the range estimation error.

Conclusions drawn in this thesis are that the polynomial order should be chosen to be 2, given some restrictions on the aspect interval, and orthogonal polynomials are used. The orthogonal polynomial gives less numerical trouble but has also other advantages. This thesis can be seen as a part of a larger effort to develop better scattering center models.

Contents

1 Introduction

Traditionally radar, which is short for ra(dio) d(etecting) a(nd) r(anging), is a method of detecting distant objects and determining their position, velocity, or other characteristics by analysis of very high frequency radio waves reflected from their surfaces. It is the echo of a pulse of microwave radiation that is used for the detecting and locating of the objects. A new radar generation is the one including automatic target recognition (ATR) where the challenge is to make use of the return signal in such a way that the target can be identified. One way of dealing with ATR is using scattering center models.

When a radar signal hits an object the signal is reflected and scattered from several points and areas on the object. This high frequency electro-magnetic scattering can be modeled as coming from a number of scattering centers.

The idea is to build a reference library with known targets. One then uses measurements, or electro-magnetic simulations, of known targets. Such reference measurements are performed at a number of carrier frequencies and a number of aspects.

The returned radar signal is received and measured at a number of carrier frequencies. The measured value is modeled as a sum of complex sinusoids in the frequency domain, corresponding to different scattering centers. The use of complex signals here have have, among other, the following reasons. Generally if the Fourier transform of a real signal is nonzero over a frequency band, then the essential information is contained in its complex envelope. Using this complex envelope will allow us to sample the signal at a lower rate than for the real signal. For further understanding, see [**?**]

In the scattering center modeling each term in the sum corresponds to a scattering center. The scattering centers can be thought of as the significant points causing the scattering. For an airplane this could be for example nose, cockpit and wings, among other.
The mathematical model for the measured reflected signal is called a scattering center model. In the area of radar based ATR it is of great importance that good scattering center models are produced and used.

Each term that corresponds to a scattering center consists of an amplitude component and an phase component where the latter contains the frequency and the range between the radar and the scattering center. The central part of identifying a target is estimating the number of terms in the sum, which in this case means estimating the number of scattering centers, and also estimating the amplitude component and the range component of each term and then compare it to references in a reference library with known targets, see [**?**]. Another method, of more interest,that also uses a reference library of scattering centers and the measured signal directly, was presented in [**?**]. Regardless of the method used in the classifier, one needs scattering center models in the reference library. One then needs to estimate scattering centers from either measurements, or electro-magnetic simulations at different carrier frequencies and aspects.

Both the range and the amplitude are varying with changing aspect. The ranges aspect dependency is known to be well described by a second order polynomial for small aspect intervals. Today's models does not capture the aspect dependency of the amplitude for an arbitrary target. For well specified scattering

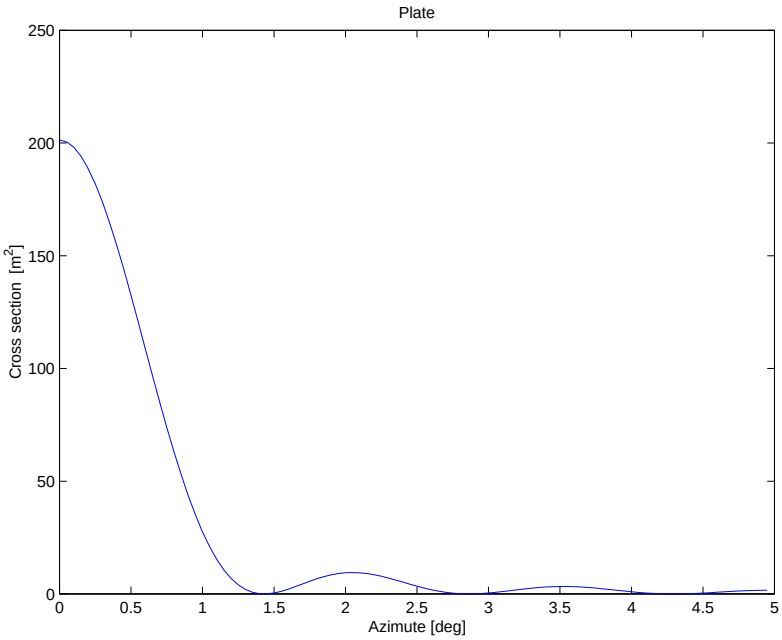

Figure 1: Amplitude for a signal reflected from a plate

centers such as a plate of a certain size and position this dependency is known as is seen in figure **??**. The amplitude of a signal reflected from plate is varying similar to the well known $\frac{\sin x}{x}$- function.
More background can be found in [**?**] and [**?**].

To be able to make a reliable estimation of the scattering from targets a suitable model for the amplitude should be developed.

2 Problem formulation

In this thesis we will:

1. Suggest a class of polynomial models for the aspect dependency of the amplitudes, over a small aspect interval, in order to extend the way of describing the aspect dependent amplitude in the scattering center model.

2. Apply some different tools from our mathematical statistical tool-box in order to find the most proper amplitude polynomial.

   (a) Compute the Fischer information matrix to obtain the Cramer-Rao lower bound for the variances of the estimations of components in the scattering center model.

(b) Make use of the Cramer-Rao lower bounds for the estimations to obtain error expressions.

i. There will occur an approximation error caused by approximating the unknown amplitude function with a polynomial. This approximation error must be small if the approximation is supposed to work.

ii. In the amplitude polynomial each coefficient is to be estimated and those estimations together generates an estimation error. The estimation error together with the approximation error will serve as a measurement to find the most appropriate order of the polynomial.

iii. The range polynomial will also generate an estimation error whose behavior will be studied as the order of the amplitude polynomial changes so that the error does not become to large.

3 The scattering center model

The scattering center model can be expressed as

$$s(f,\theta) = \sum_{k=1}^{K} a_k(\theta) e^{j2\pi f r_k(\theta)},$$

where s is the signal at frequency f and aspect θ. K is the number of scattering centers the signal comes from, a_k is amplitude for tone k (tone k is the reflection from scattering center number k) and r_k is the range between radar and scattering center number k. We use the N frequency samples $2\pi f = \omega = \omega_0 + \delta n$, for $n = 0, \ldots, (N-1)$ and the M aspect samples $\theta = dm$, for $m = 0, \ldots, (M-1)$

The observed data consists of a matrix $\boldsymbol{X}(n,m)$ of samples for different frequencies n and aspects m. The reader should note that

$$\boldsymbol{X}(n,m) = [\boldsymbol{X}]_{(n+1)(m+1)}$$

in the sample matrix. Each sample will, except from the signal, contain complex Gaussian noise $w(n,m)$ with noise variance σ^2. With this noise added the so far used model for the two-dimensional signal will have the matrix form

$$\boldsymbol{x} = \boldsymbol{F}\boldsymbol{a} + \boldsymbol{w},$$

where the columns in $\boldsymbol{X}$ have been "stacked" with the vec-operator so that

$$\boldsymbol{x} = \mathrm{vec}(\boldsymbol{X}(n,m))$$

and the vector $\boldsymbol{a}$ contains the amplitudes according to

$$\boldsymbol{a} = [a_1(0), ..., a_1(M-1), a_2(0), ..., a_2(M-1), ..., a_K(0), ..., a_K(M-1)]^T.$$

If $a(m)$ is constant, closely spaced centers can be resolved but the true aspect dependency is not accounted for. See [**?**].

If $a(m)$ is completely unknown only one center at each range can be estimated. See [**?**]. We will in the subsection **??** describe a compromise constituted by a polynomial model.

Further we have $\boldsymbol{F} = [\boldsymbol{F}_1 ... \boldsymbol{F}_K]$, where each $\boldsymbol{F}_k$ is a column vector of diagonal matrices and can be constructed as

$$\boldsymbol{F}_k = \text{Diag}(\text{vec}(\boldsymbol{P}_k^T))(\boldsymbol{1}_{N\times 1} \otimes \boldsymbol{I}_M)$$

where, for matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, $\boldsymbol{A} \otimes \boldsymbol{B}$ is the Kronecker product of $\boldsymbol{A}$ and $\boldsymbol{B}$. Definition of and reading about the Kronecker product can be found in [**?**]. The result of the product will, if $\boldsymbol{A}$ is a $p \times q$ matrix, be

$$\begin{bmatrix} A_{11}\boldsymbol{B} & \dots & A_{1q}\boldsymbol{B} \\ \vdots & \vdots & \\ A_{p1}\boldsymbol{B} & \dots & A_{pq}\boldsymbol{B} \end{bmatrix}$$

Each element in $\boldsymbol{P}_k$ is given by

$$[\boldsymbol{P}_k]_{nm} = e^{i[\psi_k]_{nm}}$$

The Diag-operator takes a vector and puts it along the diagonal of an identity matrix of size $M \times M$. We have

$$\boldsymbol{F}_k = \begin{bmatrix} e^{i[\psi_k]_{11}} & 0 & \dots & 0 & e^{i[\psi_k]_{n1}} & 0 & \dots & 0 \\ 0 & e^{i[\psi_k]_{12}} & \dots & 0 & 0 & e^{i[\psi_k]_{n2}} & \dots & 0 \\ \vdots & & & & \dots & \vdots & & \\ 0 & 0 & \dots & e^{i[\psi_k]_{1m}} & 0 & 0 & \dots & e^{i[\psi_k]_{nm}} \end{bmatrix}$$

The ψ components in the exponential terms, for scattering center k, comes from

$$\boldsymbol{\psi}_k = \boldsymbol{h}\boldsymbol{r}_k^T$$

where r_k is the range to scattering center k.

$$\boldsymbol{h} = \omega_0 \boldsymbol{1}_{N\times 1} + \delta \boldsymbol{u}_{N\times 1},$$

where ω_0 is the initial frequency multiplied by 2π, δ is the frequency step length and

$$\boldsymbol{u}_{N\times 1} = [0, ..., N-1]^T.$$

3.1 Polynomial model for aspect dependent range

The range is aspect dependent but since a relatively small aspect interval is used it is possible to approximate the range with a polynomial of low order. Here the range term r for a specific aspect will be approximated with a polynomial of order two so that for each scattering center k there is an α_k, a β_k and a γ_k to describe the range $r_k(m)$ to that scattering center.

$$r_k(m) = \alpha_k + \beta_k dm + \gamma_k \frac{d^2}{2} m^2$$

If the scattering center behaves like a fixed point on the target, the range r will be described by

$$r = \xi \cos(dm) + \eta \sin(dm)$$

and that gives us the relation $\alpha \approx \xi$, $\beta \approx \eta$ and $\gamma \approx -\xi$. The total matrix form for the range $\boldsymbol{r}$ will then be:

$$\boldsymbol{r}_k = \alpha_k \mathbf{1}_{M\times 1} + \beta_k d\boldsymbol{u}_{M\times 1} + \gamma_k \frac{d^2}{2}(\boldsymbol{u}_{M\times 1} \odot \boldsymbol{u}_{M\times 1})$$

where $\boldsymbol{u}_{M\times 1} = [0, ..., M-1]^T$. We then have a parameter vector

$$\boldsymbol{\phi} = [\phi_1 ... \phi_K]$$

with $\boldsymbol{\phi}_k = [\alpha_k \beta_k \gamma_k]^T$, that is, when $K = 2$ we have $\boldsymbol{\phi} = [\alpha_1 \beta_1 \gamma_1 \alpha_2 \beta_2 \gamma_2]^T$.

3.2 Polynomial model for aspect dependent amplitude

In order to find a proper model for the amplitudes, whos aspect dependency is not well known for arbitrary scatterers, we will try to approximate the magnitude of the amplitudes with polynomials and therefore we shall replace $\boldsymbol{a}_k$ with $\sum_{o=0}^{O} e^{i\nu_k} \boldsymbol{\lambda}_o b_k(o) = e^{i\nu_k} \boldsymbol{\Lambda b}$ where

$$\boldsymbol{Y} = \begin{bmatrix} e^{i\nu_1} & 0 & 0 & \dots & 0 \\ 0 & e^{i\nu_2} & 0 & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & e^{i\nu_K} \end{bmatrix}$$

$$\boldsymbol{b} = [b_1, \dots b_O]^T$$

The matrix $\boldsymbol{Y}$ contains the phases for the amplitudes. The phases are constant over each aspect. The vector $\boldsymbol{b}$ contains the polynomial coefficients and O is the order of the polynomials.

$$\boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{\lambda}_0 & \boldsymbol{\lambda}_1 & \dots & \boldsymbol{\lambda}_O \end{bmatrix}$$

Each column λ_o in the $\boldsymbol{\Lambda}$ matrix is some basis function of the polynomial so that for example $\boldsymbol{\Lambda}$ for ordinary polynomials is:

$$\boldsymbol{\Lambda} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & \dots & \\ 1 & 2 & 4 & \dots & \\ \vdots & & & & \\ 1 & (M-1) & (M-1)^2 & \dots & (M-1)^O \end{bmatrix}$$

This means $\boldsymbol{a}$ is being replaced by $(\boldsymbol{Y} \otimes \boldsymbol{\Lambda})\boldsymbol{b}$ in the matrix representation of the model.

We then have the polynomial coefficient vector $\boldsymbol{b}$ as a parameter and also the parameter vector

$$\boldsymbol{\nu} = [\nu_1, \nu_2, ..., \nu_K]$$

which contains the phases for the amplitudes.

The basis functions $\lambda_o(m)$ can be chosen arbitrarily but here they will be restricted to two cases. In the first case the $\mathbf{\Lambda}$ matrix will be identical with the suggested above. In the other case it will be based on orthogonal Legendre polynomials $P_o(\theta)$ generated by the recurrence formula

$$(o+1)P_{o+1}(\theta) = (2o+1)\theta P_o(\theta) - oP_{o-1}(\theta)$$

The matrix created from aspect sample values will not be orthogonal so therefore the matrix will be Single Value decompositioned. The orthogonal Legendre polynomials are further described in [**?**].

Now we have the scattering center model

$$X(n,m) = \sum_{k=1}^{K}\sum_{o=0}^{O} \boldsymbol{b}_k(o)\lambda_o(m)e^{i\nu_k}e^{i(\omega_0+\delta n)(\alpha_k+\beta_k dm+\gamma_k\frac{d^2}{2}m^2)} + w(n,m)$$

which corresponds to the matrix representation

$$\boldsymbol{x} = \boldsymbol{F}(\boldsymbol{Y}\otimes\Lambda)\boldsymbol{b} + \boldsymbol{w}$$

with $\boldsymbol{x}$ and $\boldsymbol{F}$ as earlier. The signal $\boldsymbol{x}$ is stochastic and to find out the probability density function we note that it consists of white complex Gaussian noise $\boldsymbol{w}$ where we have added a deterministic signal

$$\boldsymbol{F}(\boldsymbol{Y}\otimes\Lambda)\boldsymbol{b}$$

That is, $\boldsymbol{x}$ will have $\boldsymbol{F}(\boldsymbol{Y}\otimes\Lambda)\boldsymbol{b}$ as its mean. We also note that the observations are stochastically independent. The probability density function $f(\boldsymbol{x}|[\boldsymbol{b},\boldsymbol{\nu},\boldsymbol{\phi}])$ for $\boldsymbol{x}$ given the parameter vector $[\boldsymbol{b},\boldsymbol{\nu},\boldsymbol{\phi}]$ is then given by

$$f(\boldsymbol{x}|[\boldsymbol{b},\boldsymbol{\nu},\boldsymbol{\phi}]) = (\pi\sigma^2)^{-NM}e^{-\frac{1}{\sigma^2}\|\boldsymbol{x}-\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b}\|^2}$$

4 Cramer-Rao lower bound

Since we want to obtain the total estimated parameter vector $[\boldsymbol{b},\boldsymbol{\nu},\boldsymbol{\phi}]$ and find out the quality of the estimation, the Cramer-Rao-Lower-Bound (CRLB) is of interest.

The CRLB for an estimation of a parameter vector like the one above is

$$\boldsymbol{J}^{-1}([\boldsymbol{b},\boldsymbol{\nu},\boldsymbol{\phi}]),$$

where $\boldsymbol{J}([\boldsymbol{b},\boldsymbol{\nu},\boldsymbol{\phi}])$ is the so called Fisher information matrix. In the diagonal of the CRLB matrix we then have the CRLB:s for the individual parameters. To determine the CRLB we therefore need to compute the Fisher information matrix, which contains the derivatives, with respect to the components in the parameter vector, of the "log-likelihood", l, for the probability density function. The log-likelihood is

$$l = -NM\ln(\pi\sigma^2) - \frac{1}{\sigma^2}\|\boldsymbol{x} - \boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b}\|^2$$

and the Fisher information matrix is

$$\boldsymbol{J}([\boldsymbol{b},\boldsymbol{\nu},\boldsymbol{\phi}]) = \begin{bmatrix} E[\frac{\partial l}{\partial \boldsymbol{b}}(\frac{\partial l}{\partial \boldsymbol{b}})^T] & E[\frac{\partial l}{\partial \boldsymbol{b}}(\frac{\partial l}{\partial \boldsymbol{\nu}})^T] & E[\frac{\partial l}{\partial \boldsymbol{b}}(\frac{\partial l}{\partial \boldsymbol{\phi}})^T] \\ E[\frac{\partial l}{\partial \boldsymbol{\nu}}(\frac{\partial l}{\partial \boldsymbol{b}})^T] & E[\frac{\partial l}{\partial \boldsymbol{\nu}}(\frac{\partial l}{\partial \boldsymbol{\nu}})^T] & E[\frac{\partial l}{\partial \boldsymbol{\nu}}(\frac{\partial l}{\partial \boldsymbol{\phi}})^T] \\ E[\frac{\partial l}{\partial \boldsymbol{\phi}}(\frac{\partial l}{\partial \boldsymbol{b}})^T] & E[\frac{\partial l}{\partial \boldsymbol{\phi}}(\frac{\partial l}{\partial \boldsymbol{\nu}})^T] & E[\frac{\partial l}{\partial \boldsymbol{\phi}}(\frac{\partial l}{\partial \boldsymbol{\phi}})^T] \end{bmatrix}$$

The reason why we want to compute the Cramer-Rao lower bound is that the amplitude will be approximated with a polynomial and a good approximation should include several terms. A problem is that as we increase the number of terms in our polynomial the Cramer-Rao lower bound for the total sum of estimated parameters also increases. So the purpose is to make use of the Cramer-Rao lower bound in some way to determine the number of terms that should be included in our polynomial to give the best approximation possible, without having excessively high variance for the total sum of estimated parameters.

In order to compute the Fischer information matrix some tedious calculations are needed. They can be found in appendix **??**. The Fischer information matrix can be expressed using the matrices $\boldsymbol{G}$, $\boldsymbol{H}$ and $\boldsymbol{Q}$ defined in appendix **??**.

The elements for the first row of blocks in the Fischer information matrix are

$$E[\frac{\partial l}{\partial \boldsymbol{b}}(\frac{\partial l}{\partial \boldsymbol{b}})^T] =$$

$$= \frac{2}{\sigma^2}\Re[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})] = \frac{2}{\sigma^2}\Re[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{G}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})]$$

$$E[\frac{\partial l}{\partial \boldsymbol{b}}(\frac{\partial l}{\partial \boldsymbol{\nu}})^T] =$$

$$= \frac{2}{\sigma^2}\Im[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{G}^T(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\mathrm{Diag}(\boldsymbol{b})(\boldsymbol{I}_k\otimes\boldsymbol{1}_{1\times O})^H$$

$$E[\frac{\partial l}{\partial \boldsymbol{b}}(\frac{\partial l}{\partial \boldsymbol{\phi}})^T] =$$

$$= \frac{-2}{\sigma^2}\Im[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{H}^H\mathrm{Diag}((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})(\boldsymbol{I}_K\otimes\boldsymbol{D})^H]$$

and for the second row of blocks they are

$$E[\frac{\partial l}{\partial \boldsymbol{\nu}}(\frac{\partial l}{\partial \boldsymbol{b}})^T] =$$

$$= \frac{2}{\sigma^2}\Im[(\boldsymbol{I}_K\otimes\boldsymbol{1}_{1\times O})\mathrm{Diag}(\boldsymbol{b})^H(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{G}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})]$$

$$E[\frac{\partial l}{\partial \boldsymbol{\nu}}(\frac{\partial l}{\partial \boldsymbol{\nu}})^T] =$$

$$= \frac{2}{\sigma^2}\Re[(\boldsymbol{I}_k \otimes \boldsymbol{1}_{1\times O})\mathrm{Diag}(\boldsymbol{b})^H(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{G}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\mathrm{Diag}(b)(\boldsymbol{I}_k \otimes \boldsymbol{1}_{1\times O})^H]$$

$$E[\frac{\partial l}{\partial \boldsymbol{\nu}}(\frac{\partial l}{\partial \boldsymbol{\phi}})^T] =$$

$$= \frac{2}{\sigma^2}\Re[(\boldsymbol{I}_k \otimes \boldsymbol{1}_{1\times O})\mathrm{Diag}(\boldsymbol{b})^H(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{H}^H\mathrm{Diag}((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})(\boldsymbol{I}_K \otimes \boldsymbol{D})^H]$$

and for the last row of blocks the second order derivatives are

$$E[\frac{\partial l}{\partial \boldsymbol{\phi}}(\frac{\partial l}{\partial \boldsymbol{b}})^T] =$$

$$= \frac{2}{\sigma^2}\Im[(\boldsymbol{I}_K \otimes \boldsymbol{D})\mathrm{Diag}((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H\boldsymbol{H}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})]$$

$$E[\frac{\partial l}{\partial \boldsymbol{\phi}}(\frac{\partial l}{\partial \boldsymbol{\nu}})^T] =$$

$$= \frac{2}{\sigma^2}\Re[(\boldsymbol{I}_K \otimes \boldsymbol{D})\mathrm{Diag}((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H\boldsymbol{H}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\mathrm{Diag}(\boldsymbol{b})(\boldsymbol{I}_k \otimes \boldsymbol{1}_{1\times O})^H]$$

$$E[\frac{\partial l}{\partial \boldsymbol{\phi}}(\frac{\partial l}{\partial \boldsymbol{\phi}})^T] =$$

$$= \frac{2}{\sigma^2}\Re[(\boldsymbol{I}_K \otimes \boldsymbol{D})\mathrm{Diag}((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H\boldsymbol{Q}\mathrm{Diag}((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})(\boldsymbol{I}_K \otimes \boldsymbol{D})^H]$$

When the scattering centers are well separated the block matrices $\boldsymbol{G}$, $\boldsymbol{H}$ and $\boldsymbol{Q}$ contains real diagonals and the non-diagonal elements are close to zero. All off-diagonal blocks of the Fischer information matrix are imaginary parts of expressions containing $\boldsymbol{G}$, $\boldsymbol{H}$ or $\boldsymbol{Q}$, so these blocks are zero when the centers are well separated. This makes the Fischer matrix block diagonal and the CRLB matrix is then simply the inverse of the corresponding blocks in the Fischer matrix. See also appendix **??**

An interesting detail is that when the centers are well separated the CRLB for $\boldsymbol{b}$ only depends on the form of the basis functions. For example, look at the first block $\frac{2}{\sigma^2}\Re[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{G}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})]$. Studying the well separated centers is like studying one single center at a time. If $\boldsymbol{G}$ is an identity matrix and the$\boldsymbol{Y}$ matrix is just a constant, all that is left after simplifying the expression is $\frac{2}{\sigma^2}\Re[\boldsymbol{\Lambda}^H\boldsymbol{\Lambda}]$. This can be compared to a discussion in [**?**].

5 Errors in the scattering center model

When we use the scattering center model for the signal we approximate aspect dependencies with polynomials and we estimate several parameters. This will cause errors of various kind. The different errors are described here below and also the use of them.

5.1 Amplitude approximation and estimation error

When we approximate the amplitude aspect dependency with a polynomial we will systematically obtain an error which is the difference between the real amplitude function and the polynomial we have used to approximate it.

We will also get another error which is caused by estimating the coefficients for the polynomial. This error is stochastic but it will tend to increase as we use more terms in the polynomial, because the error is a sum of the estimation error generated in each term.

The systematical approximation error however will decrease as we increase the number of polynomial terms since it gives us a better opportunity to more accurately approximate the real amplitude function. It is therefore of interest to look at the behavior of the two errors together and hopefully find some minimal total error that tells us what order the polynomial should have. The systematic errors are also of interest on its own, since we can't expect any estimator to be able to find weak scattering centers that are of strength comparable to the errors left by the strong ones.
Given our earlier notations we define an approximation error, $\tilde{\boldsymbol{a}_k}$, for the amplitude of scattering center k:

$$\tilde{\boldsymbol{a}_k} = \mathring{\boldsymbol{a}_k} - \boldsymbol{\Lambda}\mathring{\boldsymbol{b}_k}$$

where $\mathring{\boldsymbol{a}_k}$ is the true amplitude aspect dependency and $\mathring{\boldsymbol{b}_k}$ is the ideal coefficients (coefficients without estimation error) for this dependency. We choose coefficients that minimize $\|\tilde{\boldsymbol{a}_k}\|$ and get

$$\mathring{\boldsymbol{b}_k} = \arg\min_{\boldsymbol{b}} \|\mathring{\boldsymbol{a}_k} - \boldsymbol{\Lambda}\boldsymbol{b}\|^2 = (\boldsymbol{\Lambda}^T\boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}^T\mathring{\boldsymbol{a}_k}$$

In "real life" we use estimated parameters $\hat{\boldsymbol{b}_k}$ that has an estimation error

$$\tilde{\boldsymbol{b}_k} = \mathring{\boldsymbol{b}_k} - \hat{\boldsymbol{b}_k}$$

A total amplitude error will then appear which is

$$\mathring{\boldsymbol{a}_k} - \boldsymbol{\Lambda}\hat{\boldsymbol{b}_k} = \mathring{\boldsymbol{a}_k} - \boldsymbol{\Lambda}(\mathring{\boldsymbol{b}_k} + \tilde{\boldsymbol{b}_k}) = \mathring{\boldsymbol{a}_k} - \boldsymbol{\Lambda}\mathring{\boldsymbol{b}_k} - \boldsymbol{\Lambda}\tilde{\boldsymbol{b}_k} = \tilde{\boldsymbol{a}_k} - \boldsymbol{\Lambda}\tilde{\boldsymbol{b}_k}$$

We compute $\epsilon_{\boldsymbol{a}_k}$ - the expected value of the squared total amplitude error norm. The error is normalized with the number of aspect samples.

$$\epsilon_{\boldsymbol{a}_k} = \frac{1}{M}E[\|\tilde{\boldsymbol{a}_k} - \boldsymbol{\Lambda}\tilde{\boldsymbol{b}_k}\|^2] = \frac{1}{M}E[\|\tilde{\boldsymbol{a}_k}\|^2] - \frac{1}{M}E[2\tilde{\boldsymbol{a}_k^T}\boldsymbol{\Lambda}\tilde{\boldsymbol{b}_k}] + \frac{1}{M}E[\tilde{\boldsymbol{b}_k}^T\boldsymbol{\Lambda}^T\boldsymbol{\Lambda}\tilde{\boldsymbol{b}_k}] =$$

$$= \frac{1}{M}\|\tilde{\boldsymbol{a}_k}\|^2 + \frac{1}{M}E[\tilde{\boldsymbol{b}_k}^T\boldsymbol{\Lambda}^T\boldsymbol{\Lambda}\tilde{\boldsymbol{b}_k}] = \frac{1}{M}\|\tilde{\boldsymbol{a}_k}\|^2 + \frac{1}{M}E[\mathrm{Tr}(\boldsymbol{\Lambda}^T\boldsymbol{\Lambda}\tilde{\boldsymbol{b}_k}\tilde{\boldsymbol{b}_k}^T)] =$$

$$= \frac{1}{M}\|\tilde{\boldsymbol{a}_k}\|^2 + \frac{1}{M}\mathrm{Tr}(\boldsymbol{\Lambda}^T\boldsymbol{\Lambda}E[\tilde{\boldsymbol{b}_k}\tilde{\boldsymbol{b}_k}^T]) = \frac{1}{M}\|\tilde{\boldsymbol{a}_k}\|^2 + \frac{1}{M}\mathrm{Tr}(\boldsymbol{\Lambda}^T\boldsymbol{\Lambda}\mathrm{Cov}(\tilde{\boldsymbol{b}_k}\tilde{\boldsymbol{b}_k}^T))$$

where the $\mathrm{Cov}(\tilde{\boldsymbol{b}_k}\tilde{\boldsymbol{b}_k}^T)$ part is approximated by the CRLB matrix and is the corresponding block with CRL bounds. This is due to the fact that the estimation error $\tilde{\boldsymbol{b}_k}$ is not known.

5.2 Range estimation error

In the process of estimating all parameters we will also get errors for the range. The approximation error tough is approximately zero for an aspect interval as small as in this case.

The estimation error for the range is however of interest. Since the main problem is to determine what order should be used for the amplitude polynomial, it is essential to study how the range estimation error behaves when the amplitude polynomial order is changing. The range estimation error is defined as

$$\tilde{\boldsymbol{r}_k} = \boldsymbol{D}^T\tilde{\boldsymbol{\phi}}$$

where $\tilde{\boldsymbol{\phi}}$ is the estimation error of $\boldsymbol{\phi}$.

We compute $\epsilon_{\boldsymbol{r}_k}$ the expected value of the squared range estimation error norm.

$$\epsilon_{\boldsymbol{r}_k} = \frac{1}{M}E[\|\tilde{\boldsymbol{r}_k}\|^2] = \frac{1}{M}E[\tilde{\boldsymbol{r}_k}^T\tilde{\boldsymbol{r}_k}] = \frac{1}{M}E[\tilde{\boldsymbol{\phi}_k}^T\boldsymbol{D}\boldsymbol{D}^T\tilde{\boldsymbol{\phi}_k}] = \frac{1}{M}E[\mathrm{Tr}(\boldsymbol{D}\boldsymbol{D}^T\tilde{\boldsymbol{\phi}_k}\tilde{\boldsymbol{\phi}_k}^T)] =$$

$$= \frac{1}{M}\mathrm{Tr}(\boldsymbol{D}\boldsymbol{D}^TE[\tilde{\boldsymbol{\phi}_k}\tilde{\boldsymbol{\phi}_k}^T]) = \frac{1}{M}\mathrm{Tr}(\boldsymbol{D}\boldsymbol{D}^T\mathrm{Cov}(\tilde{\boldsymbol{\phi}_k}\tilde{\boldsymbol{\phi}_k}^T))$$

where the $\mathrm{Cov}(\tilde{\boldsymbol{\phi}_k}\tilde{\boldsymbol{\phi}_k}^T)$ part, for the same reason as mentioned above, has been taken from the CRLB matrix.

6 Determining the amplitude polynomial

As mentioned earlier, to determine the order of the polynomial we look at the different types of errors and we try to minimize them. We plot the errors against the order of the amplitude polynomial and compare the behavior of the errors for different cases.

6.1 The plotted errors

The scattering centers for all cases are one plate, sometimes called SC1, and one point scatterer, sometimes called SC2, with amplitude strength $1m^2$. See figure **??** to get an approximate picture of the situation.

The standard settings for all runs are

1. Plate size 0.1×0.3.
2. Point positioned on the edge of the plate. That is down-range difference $\mathrm{d}\alpha = 0$ and cross-range difference $\mathrm{d}\beta = 0.15$.

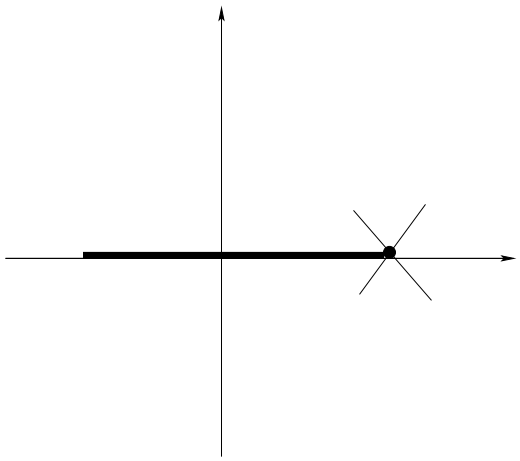

Figure 2: A plate and a point scatterer viewed from above, positioned in a coordinate system.

3. Total number of aspect samples M= 50.
4. Aspect step length d= 0.05 degrees.
5. Plate unrotated.
6. The orthogonal polynomials have been used, since they give less numerical problem in the computer work with the Fischer information matrix.
7. The noise level is $\sigma^2 = 0.01$

The errors plotted are

1. The amplitude total error, including both the amplitude approximation error and the amplitude estimation error in m^2. The amplitude squared is a natural measure of radar cross section. The plot pertains to the plate scatterer.
2. The range estimation error and the β coefficient estimation error in m. The β error is the cross-range error for the range. The estimation of β is expected to give a larger error than the α estimation. This is the reason for plotting the β error separately. The plot pertains to the plate scatterer.
3. The amplitude estimation error in m^2. The plot pertains to the point scatterer.
4. The range estimation error and the β coefficient estimation error in m. The plot pertains to the point scatterer.

The amplitude function, together with "the most realistic candidate or candidates" among amplitude approximation polynomial orders, is also plotted. Note that both the amplitude function and the approximation have been squared for the same reason as for the amplitude errors mentioned above. An approximation of order 2 can therefore be mistaken for a 4 order polynomial if it "makes

an extra turn".

Typically, the approximation part of the amplitude total error will decrease for increasing order of the amplitude polynomial while the estimation error part will increase. The idea is to find some order where the approximation error has vanished and the estimation error is not yet to large. Taking into consideration the range errors for both the plate and the point, and also the amplitude estimation error for the point scatterer, we then draw some conclusion.

6.1.1 Varying aspect step length

The problem of finding an appropriate order for the amplitude polynomial is more complicated than just plotting the different errors for different orders. The behavior of the estimation errors will also be affected by several factors in the measuring and estimation process.

In the first collection of plots the errors have been plotted against the order of the amplitude polynomial for three different aspect step lengths ($d = 0.04$, $d = 0.05$ and $d = 0.1$). Figure **??** shows the amplitude total error for the plate. It can be seen there that for smaller step length the (systematical part of the)

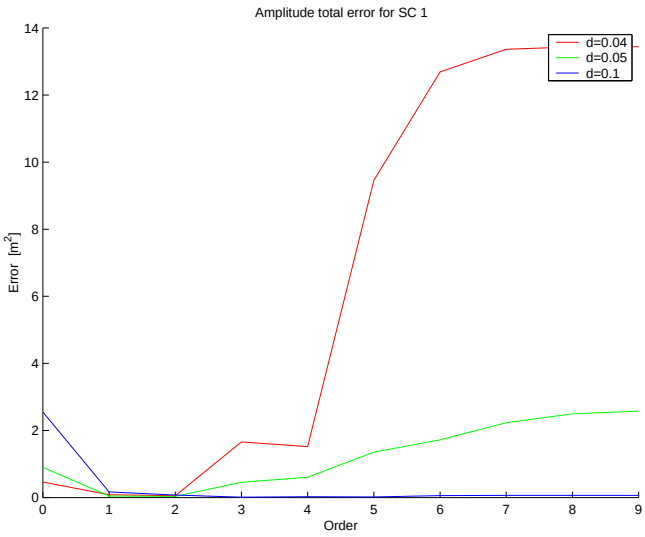

Figure 3: Amplitude total error for the plate, plotted for three different aspect step lengths

error decreases faster. However, the error increases more slowly for larger step lengths and for that larger step the approximation part of the error does not reach small values until the order is 2. The reason is that the use of the larger step length, while keeping the total number of aspect samples constant, results in using a wider aspect interval. Hence the slower decrease in approximation error. Figure **??** shows the range estimation error for the plate, with the β parameter estimation error plotted specifically. In this figure there is almost no noticeable range estimation error increase for any of the cases and the values of

the range error is small. However, we can notice an increase for the separately plotted β parameter estimation error. In figure **??** the amplitude estimation

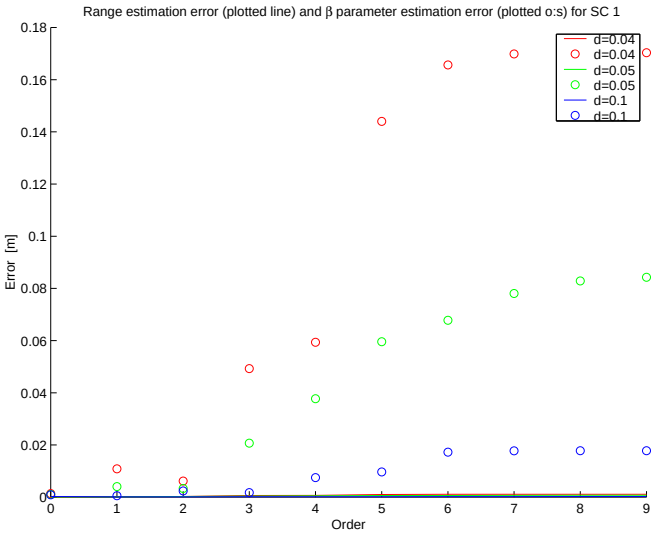

Figure 4: Range estimation error error for the plate, plotted for three different aspect step lengths

error for the point scatterer is almost constant with increasing order but only until order 2 where it starts to increase for all step length cases. Also for the

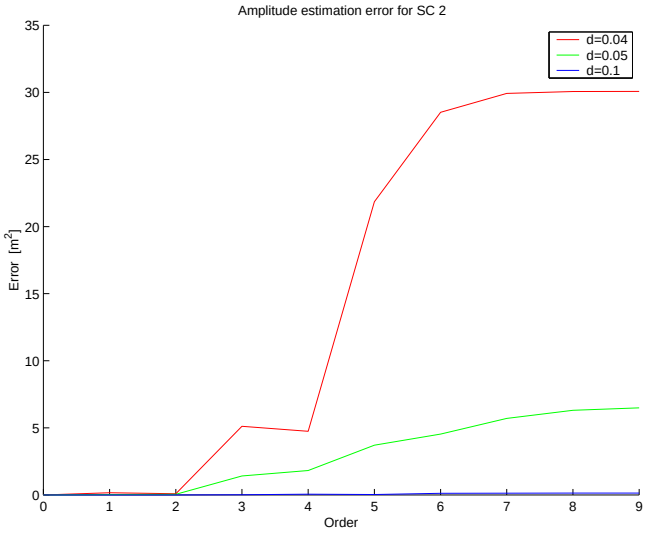

Figure 5: Amplitude estimation error for the point scatterer, plotted for three different aspect step lengths

range estimation error in figure **??** the increase is almost non-existing until order 2 where it starts to increase but only slowly. The more increasing β parameter

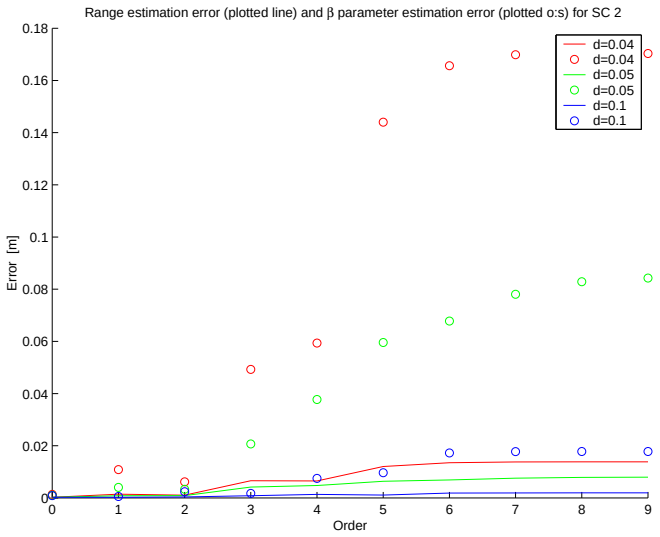

Figure 6: Range estimation error for the point scatterer, plotted for three different aspect step lengths

estimation error starts to increase fast at order 2. In figure **??** the amplitude function for the largest step length $d = 0.1$ is plotted together with amplitude approximation polynomial using order 2 and also an approximation with order 3. The approximation with order 2 is not perfect and leaves an approximation error for this large step length case. The approximation with order 3 is much better and does not leave any noticeable approximation error but it will probably make the estimation of the point impossible.

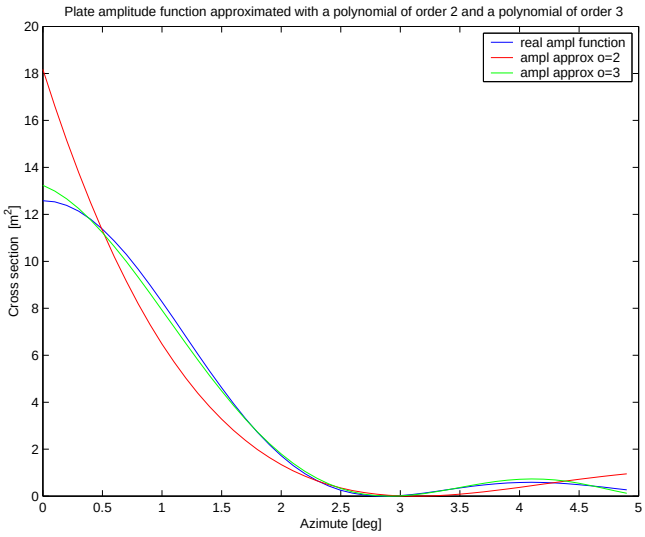

Figure 7: The amplitude function for the largest step length plotted with two different approximations. The polynomial orders are 2 and 3.

6.1.2 Varying total aspect sample size

The second collection of plots contains the errors plotted for three different cases where the total number of aspect samples have been varied between $M = 50$, $M = 100$ and $M = 150$. In figure **??** we see the amplitude total error for

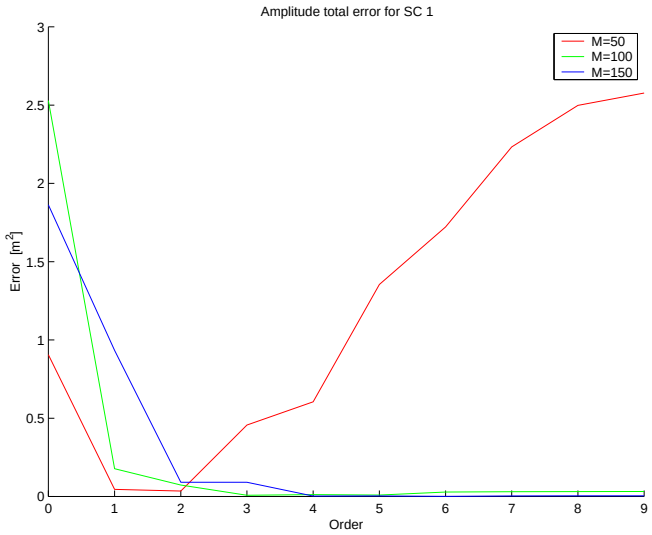

Figure 8: Amplitude total error for the plate, plotted for three different aspect sample sizes

the plate. Clearly, it looks like the approximation needs higher polynomial order for larger aspect sample sizes. This is due to the fact that we keep the aspect step length constant. The aspect interval is the product of aspect sample size and aspect step length and thus it increases as we increase the sample size. A wider aspect interval includes more of the amplitude function and needs therefore a higher order to be approximated. The highest order needed, among the three cases, to guarantee that the approximation part of the error has become approximately zero is 4 but it is small already at order 2. This is also the last polynomial order before the estimation part of the error starts to increase significantly with increasing order, for the smallest sample size case. The range estimation error for the plate, which can be seen in figure **??** does not increase significantly with increasing order for any of the cases. The estimation error of the linear coefficient β shows a relatively large increase, beginning at order 2, for the smallest sample size case. The amplitude estimation error, in figure **??**, for the point scatterer does not show almost any increase except for the smallest sample size case for which the error starts its fast increase at order 2. The behavior of the range estimation error, in figure **??**, for the point scatterer is similar to the corresponding error for the plate in figure **??**. At order 2 the range error for all three cases begins to increase slowly and the β error begins to increase faster. In figure **??** the amplitude function for sample size $M = 150$ is plotted together with the amplitude approximation polynomial using polynomial order 2 and 4. The approximation with order 4 is very good but using this polynomial order means accepting an amplitude estimation error that exceeds the strength of the point scatterer tone. Order 2 means a worse

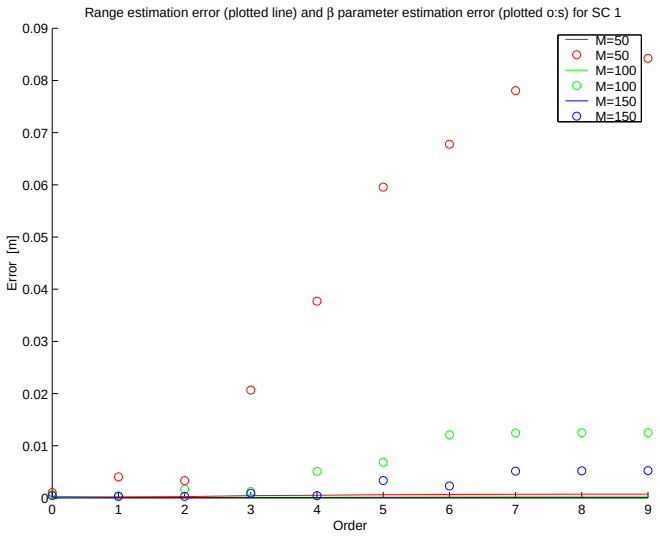

Figure 9: Range estimation error error for the plate, plotted for three different aspect sample sizes

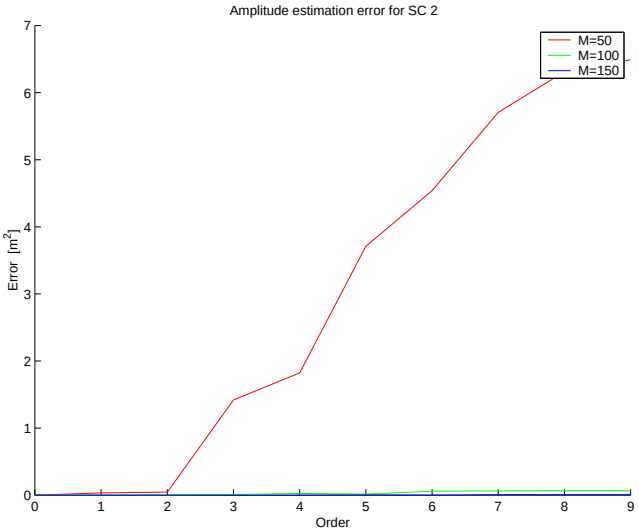

Figure 10: Amplitude estimation error for the point scatterer, plotted for three different aspect sample sizes

approximation but a possibility to estimate the tone from the point scatterer.

6.1.3 Varying scattering center positions

Other factors affecting the estimation is the combination of scattering centers and the position of them. For example, two plates of similar size, well separated in position will be much more easy to estimate than a large plate and a small point reflector placed on the edge of the plate. The signal from the plate

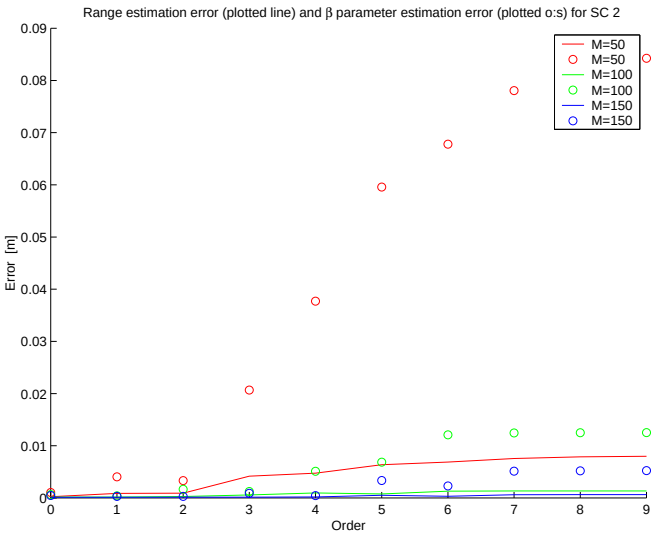

Figure 11: Range estimation error for the point scatterer, plotted for three different aspect sample sizes

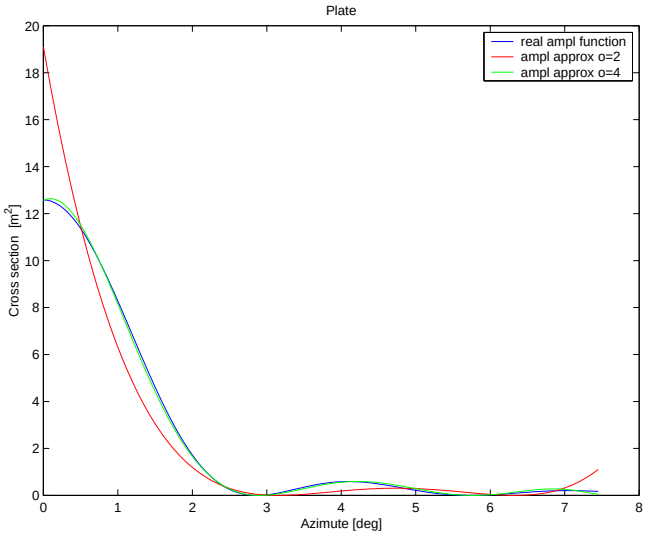

Figure 12: The amplitude function for the largest sample size plotted with two different approximations.

is much stronger, and thus have much higher amplitude, than the signal from the small point reflector. Both cases must be taken care of in the creation of a good estimation model. In figures **??** to **??** the errors have been plotted for three different positions between a plate and a point scatterer. The first position between them is $\mathrm{d}\alpha = 0$, $\mathrm{d}\beta = 0.15$ (unseparated), the second is $\mathrm{d}\alpha = 0.05$, $\mathrm{d}\beta = 0.15$ and the third is $\mathrm{d}\alpha = 0.5$, $\mathrm{d}\beta = 0.5$. In figure **??** the approximation part of the error has become approximately zero at amplitude polynomial order

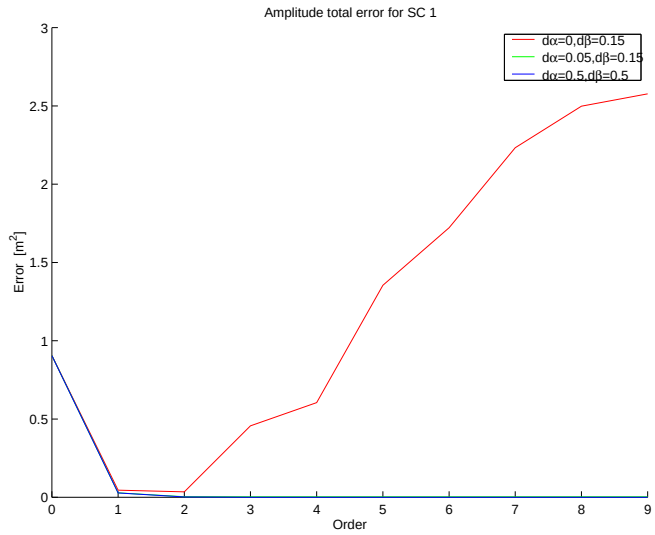

Figure 13: Amplitude total error for the plate, plotted for three different positions of the plate and the point scatterer

1, for all three different positions of the plate and the point scatterer. The estimation part of the error has a clear increasing behavior from order 2 for the case where $d\alpha = 0$ and $d\beta = 0.15$, that is for the unseparated positioning of the centers. The range estimation error for any of the three cases does not increase

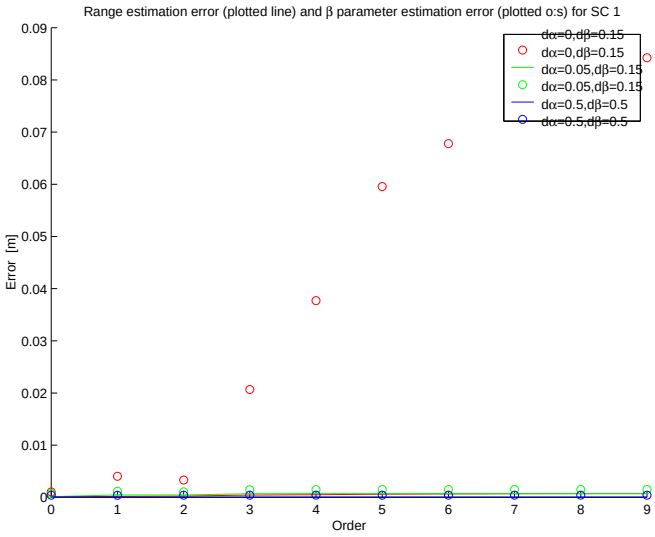

Figure 14: Range estimation error error for the plate, plotted for three different positions of the plate and the point scatterer

noticeably in figure **??** except from the β parameter estimation error for the case with unseparated centers. That β parameter error increases from order 2. The amplitude estimation error for the point scatterer can be seen in figure **??**.

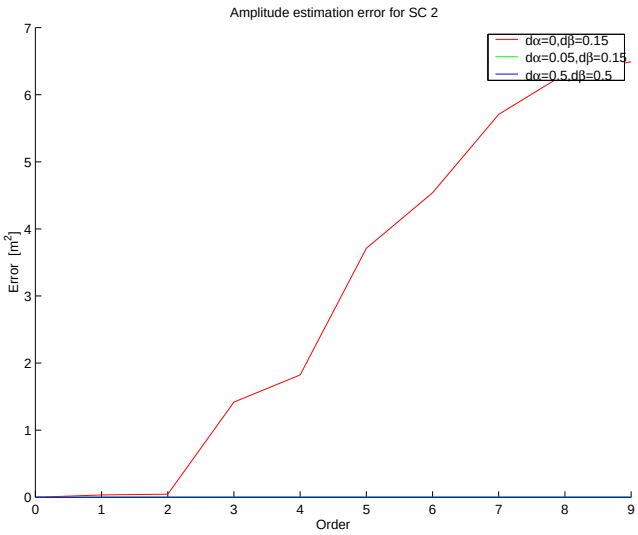

Figure 15: Amplitude estimation error for the point scatterer, plotted for three different positions of the plate and the point scatterer.

The error increases from order 2 for the unseparated case. The two other cases has no noticeable increase.

In figure **??** we notice only slow increase in range error, for any of the position cases, except from the β parameter estimation error for the unseparated case which increases from order 2. In figure **??** the amplitude function is plotted

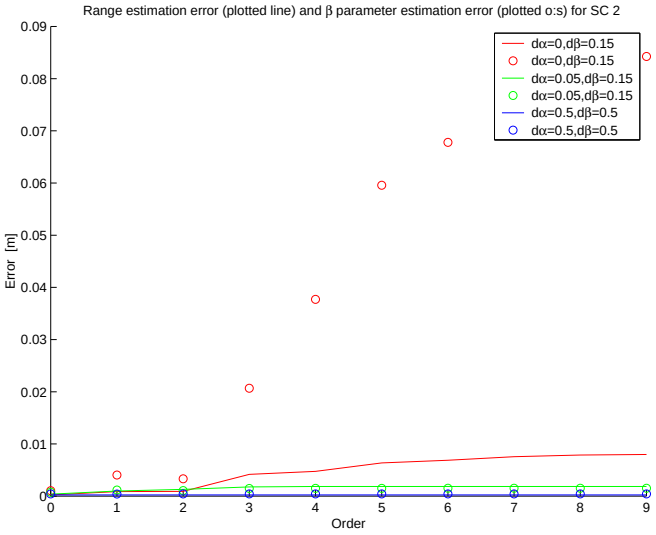

Figure 16: Range estimation error for the point scatterer, plotted for three different positions of the plate and the point scatterer.

together with the amplitude approximation polynomial using order 2.

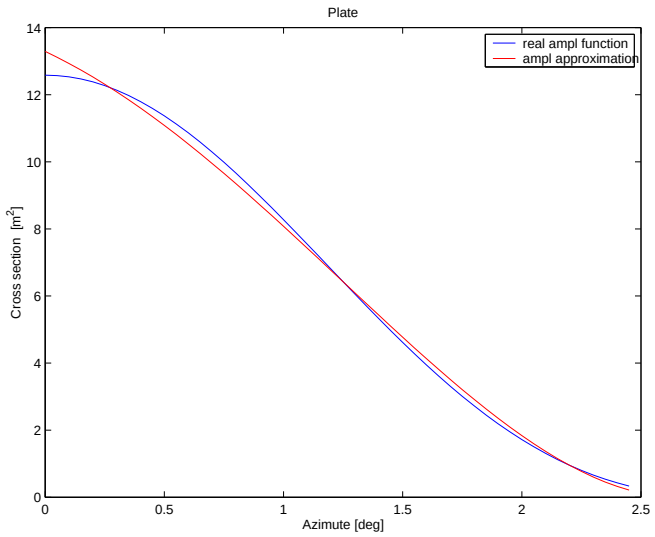

Figure 17: The amplitude function plotted with an approximation with polynomial order 2.

6.1.4 Varying plate rotation

If the plate is rotated the amplitude function is dislocated from its symmetric location about zero. In this collection of figures **??** to **??** the plate rotation has been varied between -2 deg, 0 deg and 4 deg. We see in figure **??** that there

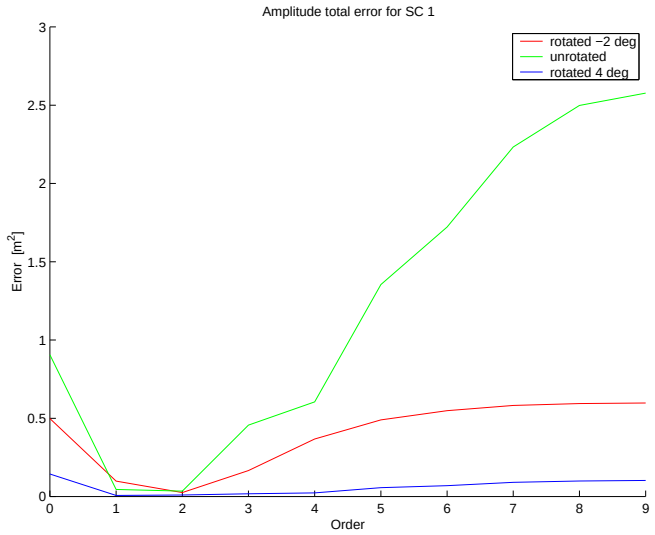

Figure 18: Amplitude total error for the plate, plotted for three different rotations of the plate

is a joint minimum at polynomial order 2 where the approximation part of the amplitude error has become almost zero for all three cases and the estimation

error has not started to increase. The range error in figure **??** does not increase

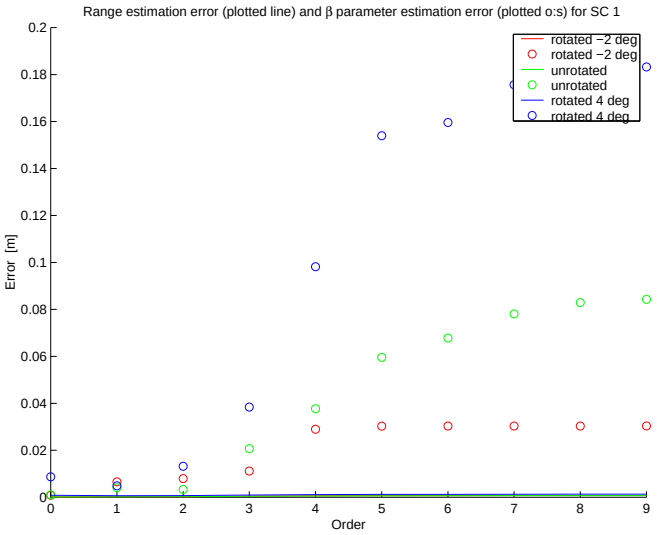

Figure 19: Range estimation error for the plate, plotted for three different rotations of the plate

noticeably for any of the three cases. The β error increases significantly from order 2 for all three cases but the actual values of it are not high. In figure

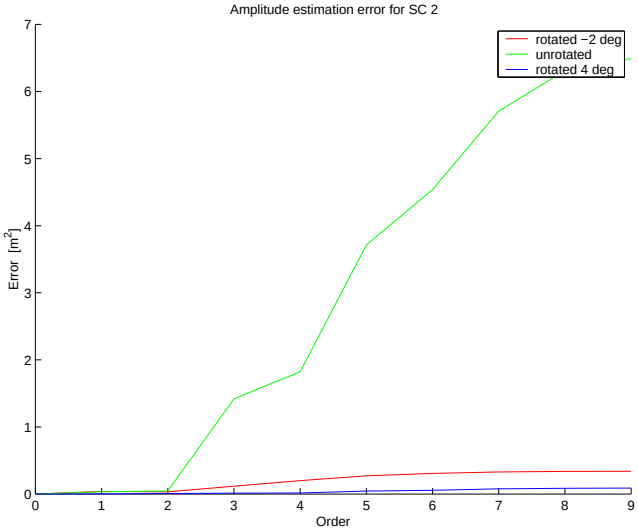

Figure 20: Amplitude estimation error for the point scatterer, plotted for three different rotations of the plate

?? we see that for all three cases the amplitude estimation error increases from order 2. Already at order 3 the error for the unrotated plate case has become larger than the amplitude of the tone from the point.

In figure **??** the range estimation error for the point only increases very slow for all plotted cases and the β error, which increases much more, starts to increase faster from order 2. In figure **??** the amplitude function for the last case of

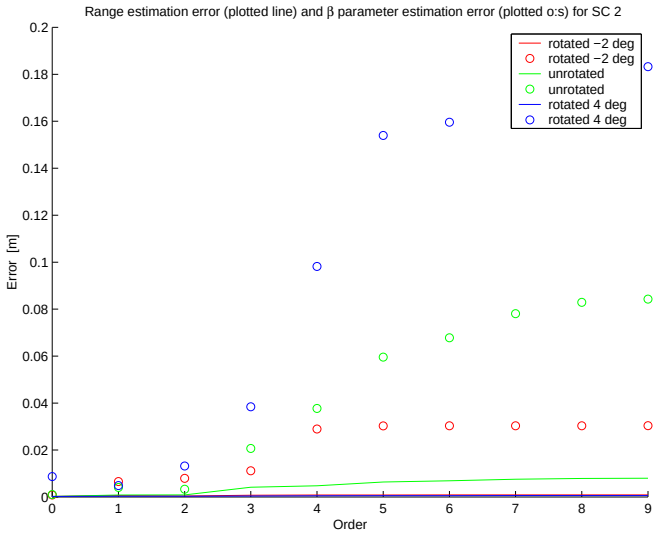

Figure 21: Range estimation error for the point scatterer, plotted for three different rotations of the plate

rotation is plotted together with the amplitude approximation polynomial using order 2.

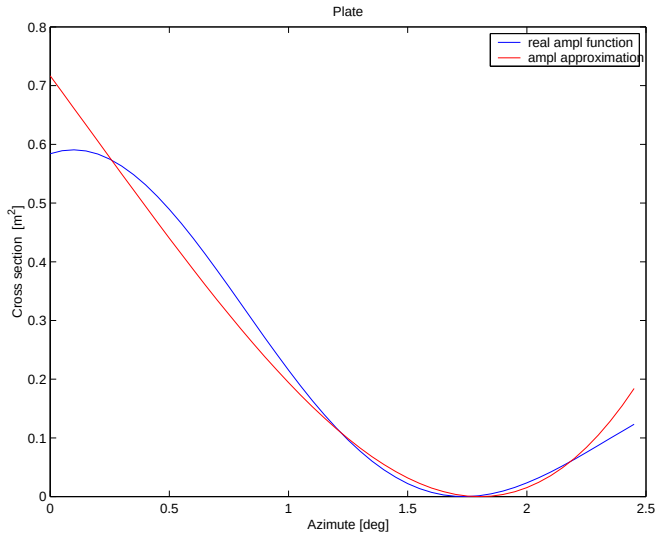

Figure 22: The amplitude function for rotation 4 degrees plotted with an approximation with polynomial order 2

6.1.5 Varying plate size

The most obvious factor affecting the amplitude approximation error is the order of the polynomial. The more terms included in the polynomial the better the approximation and hence less approximation error. Now, if the aspect interval is increased, when looking at the scattering from a plate, it will allow for the aspect dependent amplitude to vary more and thus it will be more to approximate. Something similar happens if the plate size is increased. See figures **??** and **??**. The amplitude function will then vary faster, but in the same aspect

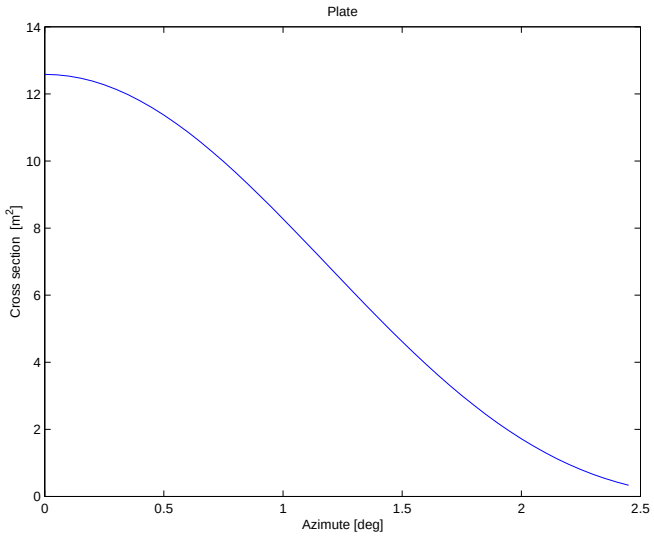

Figure 23: Amplitude varying with the aspect interval 0 to 2.5. Plate size 0.1×0.3

interval. This calls for higher order of the amplitude approximation polynomial and therefore the amplitude approximation error will not decrease with increasing polynomial order as fast as for a smaller plate size. In the last collection of plots the errors have been plotted for three different cases of target plate size. The sizes are 0.1×0.1, 0.1×0.3 and 0.1×0.6 m. As mentioned above the approximation error decreases slower for larger plate sizes because the amplitude aspect dependency becomes more and more difficult to approximate for larger plates. The approximation then needs higher degree of the polynomial to give a proper result. This is obvious in figure **??**. For the largest plate the systematical part of the amplitude error does not become approximately zero until at order 3 while it is almost zero for order 0 and 1 for plate sizes 0.1×0.1 and 0.1×0.3. The range error in figure **??** increases slowly except for the *beta* error for the first plate size case. It starts to increase fast at order 2. The amplitude estimation error for the point in figure **??** starts to increase dramatically from order 2 for the smallest plate case. Up to order 2 the error for all three cases is approximately zero or small. In figure **??** the range error for the point acts very similar to the range error in figure **??**. In figure **??** the amplitude function for the largest plate is plotted together with an amplitude approximation polynomial using order 2 and an approximation with order 3. The approximation with order 3 i much better than the approximation with order 2 but a higher

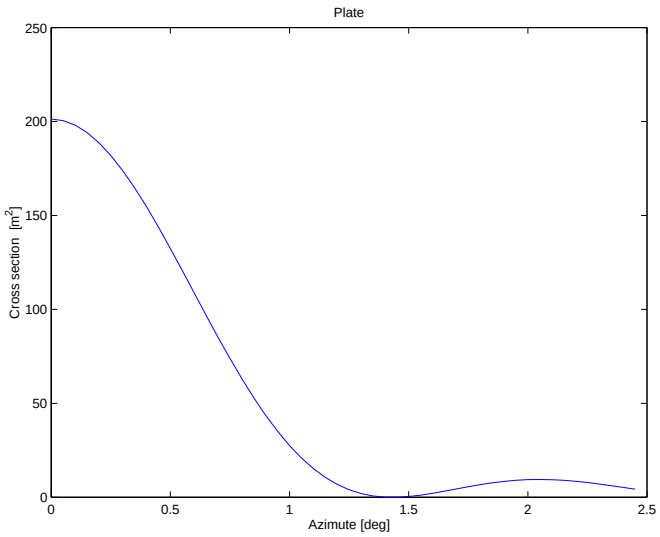

Figure 24: Amplitude varying with the aspect interval 0 to 2.5. Plate size 0.1×0.6

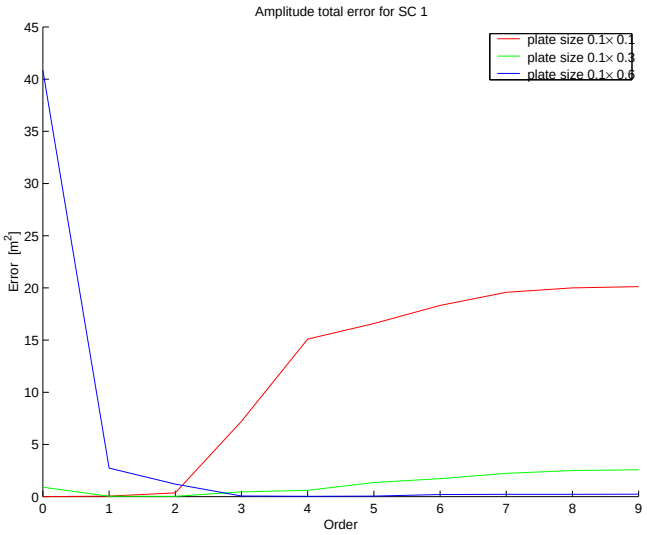

Figure 25: Amplitude total error for the plate, plotted for three different sizes of the plate

order than 2 makes the estimation of the point very unlikely.

6.2 Polynomial order discussion

It will turn out here below that it is not impossible to get an acceptable signal estimation model when letting the aspect dependent amplitude be approximated by a polynomial. The model choice is then between some different orders of that polynomial.

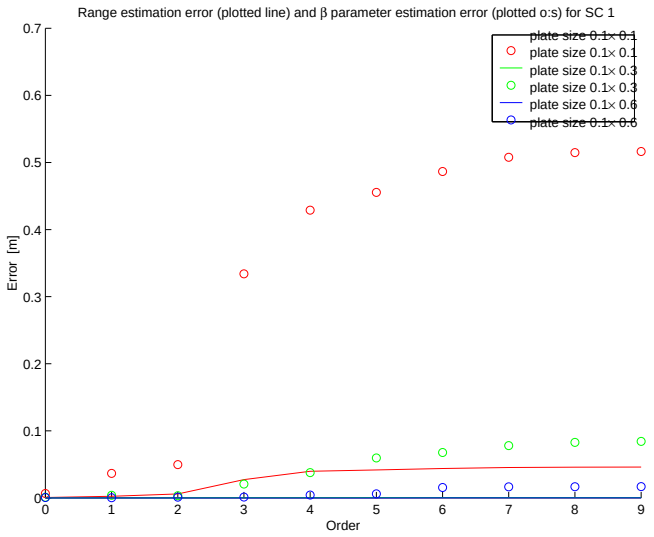

Figure 26: Range estimation error for the plate, plotted for three different sizes of the plate

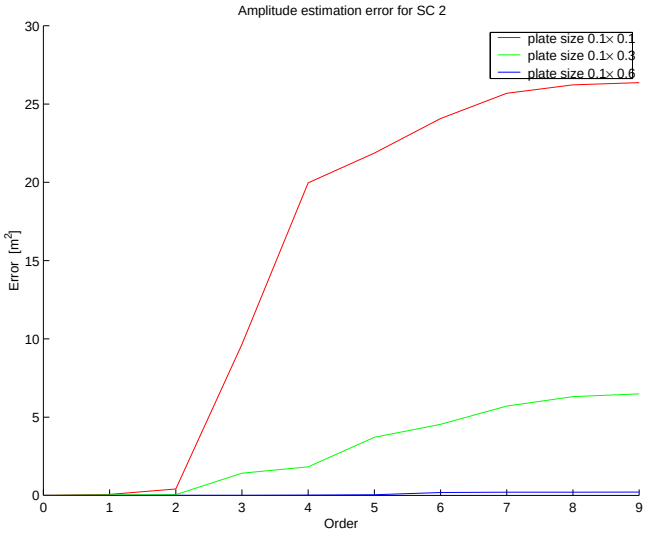

Figure 27: Amplitude estimation error for the point, plotted for three different sizes of the plate

For the total amplitude error for the plate, including the estimation error, there is almost always be possible to find a minimum for some relatively low order. In most cases there also is one or more error value greater than the minimum that could be acceptable. These values occur somewhere around order 1, 2 or 3. The approximation part of the error does sometimes not get approximately zero until at order 4. At order 4 the estimation error has become large for a

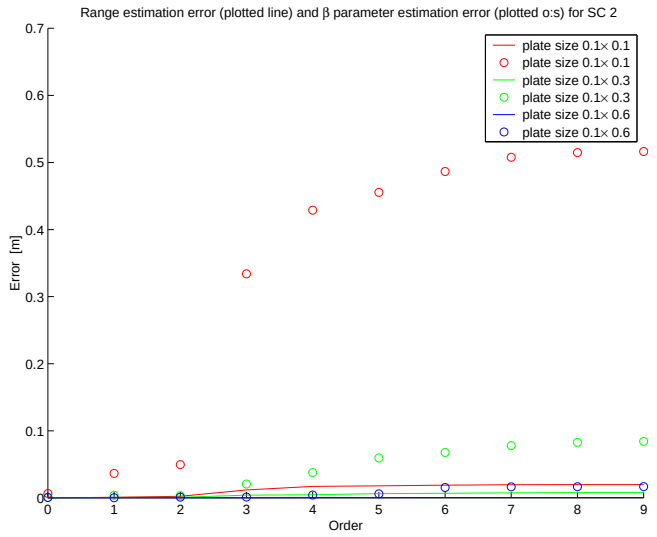

Figure 28: Range estimation error for the point, plotted for three different sizes of the plate

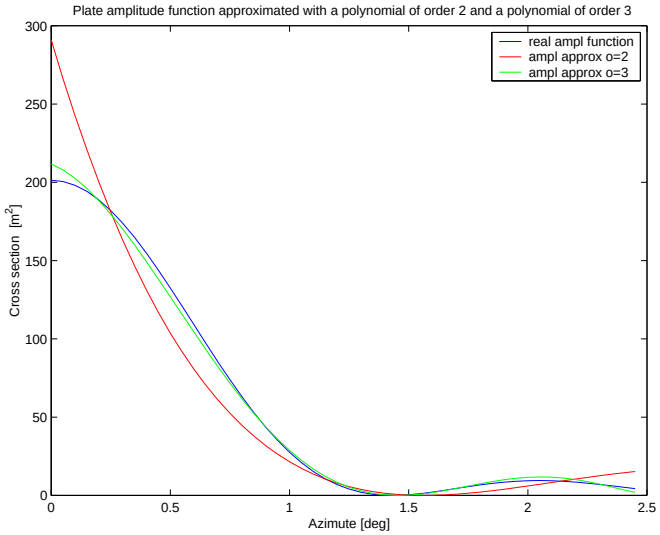

Figure 29: The amplitude function for plate size 0.1×0.6m plotted with two different approximations. The polynomial orders are 2 and 3.

small plate so it might not be advisable to choose as high order as 4 if small plates are supposed to be estimated.

The amplitude estimation error for the plate consequently gets to large to make the estimation of the point possible when amplitude polynomial order gets higher than 2 but is approximately zero or very small for order 0, 1 and 2.

The range estimation never appears to be a an actual problem. The estimation error does vary with amplitude polynomial order but is consequently relatively small. The error increase is dominated by the cross range parameter (β) error while the range estimation error in total is small.

When choosing a permanent order for the amplitude polynomial, it should be chosen so that the systematical error for the amplitude is approximately zero for all relevant estimation cases. Perhaps some increase in estimation error should be accepted in order to prepare for a more complicated amplitude aspect dependency (e.g. for a wider aspect interval or a larger plate) where a higher order will be nessecary.It could be wise to do this rather than risking the approximation and maybe therefore risking the estimation. A signal "rest" could cause an estimator to "think" that there exists some extra scattering center. Choosing an higher order could mean that we have to sacrifice the estimation of a weak point scatterer and for order 4 maybe even a small plate. If estimating large plates is highest priority and we are willing to do the sacrifice we should choose order 4. If we can sacrifice only the estimation of a weak point scatterer, we can choose order 3.

However, if we restrict the relevant cases to mean that a plate has maximum size 0.1×0.3m and that the aspect interval is at most 5 deg, then it is enough to use order 2 for the approximation.

6.3 Polynomial basis functions discussion

In order to examine whether the ordinary polynomials in the $\mathbf{\Lambda}$ matrix should be kept or replaced by the orthogonal Legendre polynomials, different comparison between them can be made. First we will study the errors to see if they show any difference between the two kinds of polynomials. Figures **??** and **??**

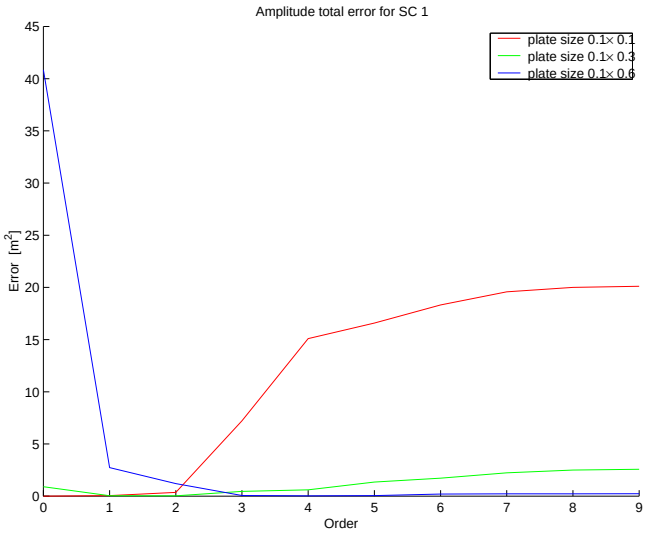

Figure 30: Amplitude total error for the plate, plotted for three different plate sizes, using the ordinary polynomials.

show the same error plots as figures **??** to **??** except that here we use ordinary polynomials. As seen in the figures there is not much difference between the

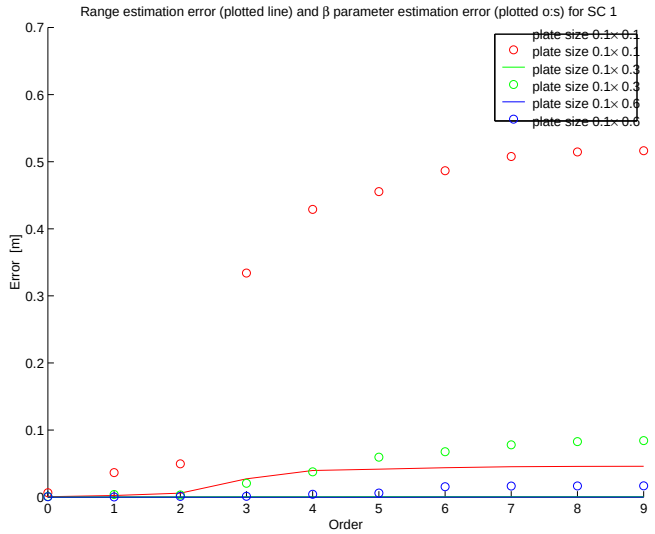

Figure 31: Range estimation error for the plate, plotted for three different plate sizes, using the ordinary polynomials.

plots with ordinary polynomials and the ones with orthogonal polynomials. In

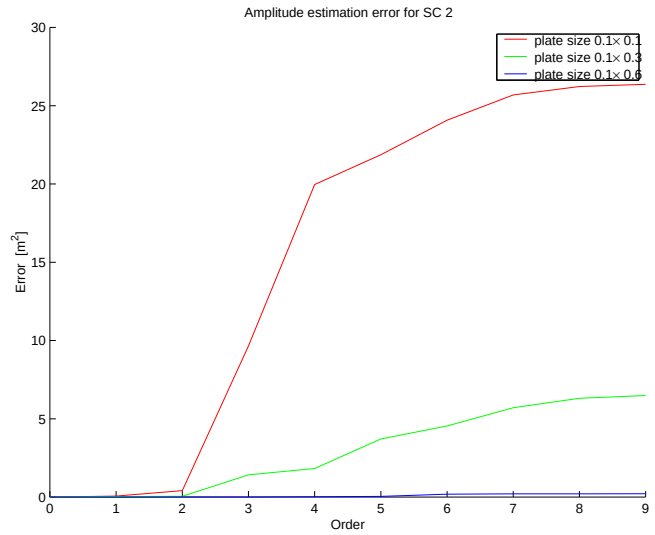

Figure 32: Amplitude estimation error for the point, plotted for three different plate sizes, using the ordinary polynomials.

fact they look identical. The orthogonality does not seem to affect the approximation or estimation itself. However, when using the ordinary polynomial basis functions there was numerical problems in the computing of the CRLB matrix.

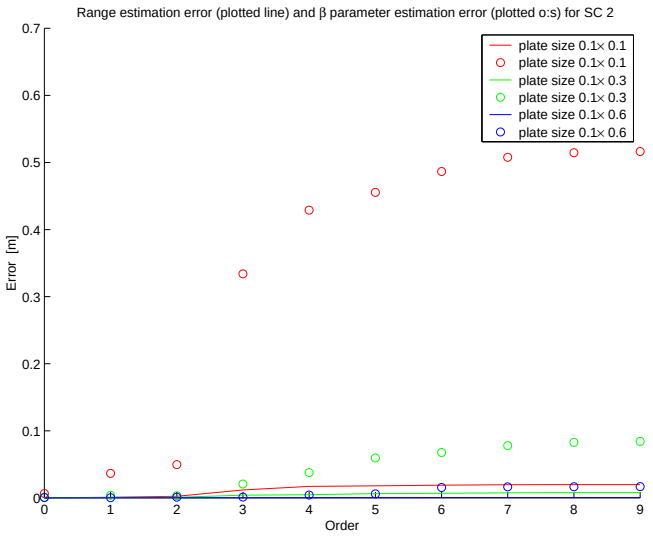

Figure 33: Range estimation error for the point, plotted for three different plate sizes, using the ordinary polynomials.

The inverting of the Fischer information matrix causes complains about bad conditioning. This problem does not occur when replacing the ordinary $\mathbf{\Lambda}$ as an orthogonal matrix using orthogonal Legendre polynomials.

7 Possible extension

Since this thesis is a part of a larger effort aiming to develope good and useful scattering center models, a natural extension of the thesis would be to implement an estimator needed for estimating signals using our new model.

The estimation of all signal parameters in the model, using a detected signal, is a difficult task. One possible approach is to estimate one parameter at a time, holding the others constant. This can be done by the so called RELAX method which is described in [**?**]. The original RELAX method by Lee and Stoica, see [**?**], is a way of estimating the parameters in models like

$$x_n = \sum_{k=1}^{K} a_k e^{i2\pi f_k n} + w_n$$

This can be generalized to more general multicomponent signal. The RELAX algorithm approaches least square problems by estimating some of the unknown parameters while keeping the others constant. Consider the estimation of the parameters for the complex sinusoids in complex Gaussian noise. This will be carried out in a maximum likelihood manner and it will result in minimizing the following "cost" function:

$$C(f_1, a_1, .., f_k, a_k) = ||\boldsymbol{x} - \sum_{k=1}^{K} \omega(f_k) a_k||$$

Now we take one component

$$y_k = \boldsymbol{x} - \sum_{i=1, i\neq k}^{K} \hat{a}_i \omega(\hat{f}_i)$$

We minimize

$$\|y_k - a_k \omega(f_k)\|^2$$

for the component k, holding the others constant, and this iterates trough the K components until the residual error

$$\epsilon_K = \|\boldsymbol{y} - \sum_{k=1}^{K} \hat{a}_k \omega(\hat{f}_k)\|^2$$

is no longer changing during the last iteration.

To apply RELAX to our problem we will have to find the minimum of least square expressions like

$$\underset{\boldsymbol{b},\boldsymbol{\phi},\boldsymbol{\nu}}{Min} ||\boldsymbol{y}_k - \boldsymbol{F}_k e^{j\nu_k} \boldsymbol{\Lambda} \boldsymbol{b}_k||^2$$

In that work the smart choice of orthogonal Legendre polynomials will be a great help. Making the $\boldsymbol{\Lambda}$ matrix orthogonal will cause $\boldsymbol{\Lambda}^T \boldsymbol{\Lambda}$ to become $\boldsymbol{I}$ and hence large expressions will collapse. The computationally heavy RELAX algorithm

can then be a bit less heavy.

The Following calculations are an attempt to find a good expression for the estimator. Here all expressions are for one component k.

$$\boldsymbol{J_0} = ||\boldsymbol{x} - \boldsymbol{F\Lambda b}e^{i\boldsymbol{v}}||^2$$

This is to be minimized, by estimating one parameter at a time, holding the others constant:

$$\boldsymbol{\Lambda}^T\boldsymbol{\Lambda} = \boldsymbol{I}_0$$

$$\boldsymbol{F}^H\boldsymbol{F} = N\boldsymbol{I}_M$$

$$\frac{\partial \boldsymbol{J}}{\partial \boldsymbol{b}} = -2\Re[e^{-i\boldsymbol{\nu}}\boldsymbol{\Lambda}^T\boldsymbol{F}^H(\boldsymbol{x} - \boldsymbol{F\Lambda b}e^{i\boldsymbol{\nu}})] =$$

$$= -2\Re[e^{-i\boldsymbol{\nu}}\boldsymbol{\Lambda}^T\boldsymbol{F}^H\boldsymbol{x}] + 2N\boldsymbol{b}$$

which implies that

$$\hat{\boldsymbol{b}} = \frac{1}{N}\Re[e^{i\boldsymbol{\nu}}\boldsymbol{\Lambda}^T\boldsymbol{F}^H\boldsymbol{x}] = \frac{1}{N}\Re[e^{i\boldsymbol{\nu}}\boldsymbol{U}]$$

where $\boldsymbol{U}$ is a notation:

$$\boldsymbol{U} = \boldsymbol{\Lambda}^T\boldsymbol{F}^H\boldsymbol{x}$$

After computing $\hat{b}$ (or estimating b) the result is inserted and from the new expression we derive an estimation for $\hat{\nu}$.

$$\boldsymbol{J_1} = ||\boldsymbol{x} - \boldsymbol{F\Lambda}\hat{\boldsymbol{b}}e^{i\boldsymbol{\nu}}||^2 = ||\boldsymbol{x}||^2 + N||\hat{\boldsymbol{b}}||^2 - 2\Re[e^{-i\boldsymbol{\nu}}\hat{\boldsymbol{b}}^T\boldsymbol{\Lambda}^T\boldsymbol{F}^H\boldsymbol{x}] =$$

$$= ||\boldsymbol{x}||^2 + N(||\hat{\boldsymbol{b}}||^2 - 2||\tilde{\boldsymbol{b}}||^2) = ||\boldsymbol{x}||^2 - N||\hat{\boldsymbol{b}}||^2 =$$

$$= ||\boldsymbol{x}||^2 - N\Re[e^{i\boldsymbol{\nu}}\boldsymbol{U}]^T\Re[e^{i\boldsymbol{\nu}}\boldsymbol{U}] =$$

$$||\boldsymbol{x}||^2 - \frac{N}{4}(e^{i2\boldsymbol{\nu}}\boldsymbol{U}^T\boldsymbol{U} + \boldsymbol{U}^H\boldsymbol{U} + \boldsymbol{U}^T\boldsymbol{U}^* + e^{-i2\boldsymbol{\nu}}\boldsymbol{U}^H\boldsymbol{U}^*) =$$

$$||\boldsymbol{x}||^2 - \frac{N}{2}\boldsymbol{U}^H\boldsymbol{U} - \frac{N}{2}Re[e^{i2\boldsymbol{\nu}}\boldsymbol{U}^T\boldsymbol{U}],$$

$$\hat{\boldsymbol{\nu}} = \underset{\boldsymbol{\nu}}{ArgMax}\ J_1 = -Arg[\boldsymbol{U}^T\boldsymbol{U}]$$

This results in:

$$\boldsymbol{J_2} = ||\boldsymbol{x} - \boldsymbol{F\Lambda}\hat{\boldsymbol{b}}e^{i\tilde{\boldsymbol{\nu}}}||^2 = ||\boldsymbol{x}||^2 - \frac{N}{2}\boldsymbol{U}^H\boldsymbol{U} - \frac{N}{2}|\boldsymbol{U}^T\boldsymbol{U}|$$

The above expression should be minimized with respect to $\Phi = [\alpha, \beta, \gamma]$. The procedure will be carried out in a similar way as for b and ν.

8 Conclusions

We have derived the CRLB matrix for a scattering center model with amplitudes modeled as polynomials. We have shown that it is possible tot use the CRLB together with the approximation error for determining a suitable polynomial model for the amplitude aspect dependency.

The judgment about amplitude polynomial model order must be that for the plate and point configuration of scattering centers order 2 is the most wise one to use here, if we choose some suitable sample size step length so that the aspect interval is small enough and if a plate has maximum size 0.1×0.3m. For example the step length could be chosen to be 0.05 and the sample size 50.

If prioritize finding plates and are willing to sacrifice amplitude estimation of the point, the order should be 3.

We can also draw the conclusion that the use of an orthogonal Legendre polynomial $\mathbf{\Lambda}$ gives us the benefit of less numerical problems in the computer work with the CRLB matrix and, as motivated, is an intelligent choice for further work with scattering center models.

A Notation

An abbreviation of the symbols used:

$\boldsymbol{X}(n, m)$ The matrix of sampled data

$\boldsymbol{x}$ The vector of sampled data. Build from the matrix of sampled data by stacking the columns.

$\boldsymbol{F}$ Matrix. Part of the model. Contains phases, frequencies and range terms.

$\boldsymbol{a}$ Vector of amplitudes.

$\boldsymbol{w}$ Vector of noise complex Gaussian noise terms. able

$\boldsymbol{P}_k$ Matrix

$\boldsymbol{\psi}_k$ Matrix of range terms.

$\boldsymbol{h}$ Vector of frequencies.

ω_0 Initial frequency.

δ Frequency step length.

$\boldsymbol{u}_{N\times 1}$ Column vector of length N. containing the integers from 1 to N.

$\mathbf{1}_{N\times 1}$ Column vector of length N. containing 1's.

$\boldsymbol{I}_M$ Identity matrix of size $M \times M$.

$\boldsymbol{r}_k$ The range to scattering center k.

α_k Coefficient to the constant term in the range polynomial.

β_k Coefficient to the linear term in the range polynomial.

γ_k Coefficient to the quadratic term in the range polynomial.

d Aspect step length.

N The total number of frequency samples.

M The total number of aspect samples.

O The (total) order of the amplitude polynomial.

ϕ The vector of coefficients to the range polynomial.

$\boldsymbol{b}$ The vector of coefficients to the amplitude polynomial.

$\boldsymbol{\Lambda}$ The matrix of amplitude polynomials.

$\boldsymbol{Y}$ The matrix containing the amplitude phase terms.

$\boldsymbol{\nu}$ The vector of amplitude phases.

l The log-likelihood function for the data vector $\boldsymbol{x}$.

$\boldsymbol{D}$ The matrix of the range derivative with respect to the coefficient vector ϕ.

$\boldsymbol{G}$ Matrix, see appendix **??**.

$\boldsymbol{H}$ Matrix, see appendix **??**.

$\boldsymbol{Q}$ Matrix, see appendix **??**.

B Computing the Fischer information matrix

In order to be to determine the second order derivatives we first compute the first order derivatives, then compute and give several expressions for different derivatives.

B.1 First order derivatives

We begin by again giving definitions $\boldsymbol{F} = [\boldsymbol{F}_1 ... \boldsymbol{F}_K]$, where each $\boldsymbol{F}_k$ is a column vector of diagonal matrices and can be constructed as

$$\boldsymbol{F}_k = \text{Diag}(\text{vec}(\boldsymbol{P}_k^T))(\boldsymbol{1}_{N\times 1} \otimes \boldsymbol{I}_M)$$

$$[\boldsymbol{P}_k]_{nm} = e^{i[\psi_k]_{nm}}$$

$$\boldsymbol{\psi}_k = \boldsymbol{h}\boldsymbol{r}_k^T$$

We have

$$\boldsymbol{Y} = \begin{bmatrix} e^{i\nu_1} & 0 & 0 & \dots & 0 \\ 0 & e^{i\nu_2} & 0 & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & e^{i\nu_K} \end{bmatrix}$$

which contains the phases $\boldsymbol{\nu} = [\nu_1, \dots, \nu_K]$ for the amplitudes and

$$\boldsymbol{b} = [b_1, \dots b_O]^T$$

which contains the coefficients to the amplitude polynomial. Further we have

$$\boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{\lambda}_0 & \boldsymbol{\lambda}_1 & \dots & \boldsymbol{\lambda}_O \end{bmatrix}$$

Each column λ_o in the $\boldsymbol{\Lambda}$ matrix is some basis function of the polynomial. We then have a parameter vector

$$\boldsymbol{\phi} = [\phi_1 ... \phi_K]$$

with $\boldsymbol{\phi}_k = [\alpha_k \beta_k \gamma_k]^T$ which are the coefficients for the range polynomial

To compute the derivatives of the "log-likelihood" we need the following derivation rules, where we allow for the vectors and the matrix to be complex. For a vector $\boldsymbol{c}$, a matrix $\boldsymbol{M}$ and for a function $\boldsymbol{z}$ of a vector $\boldsymbol{v}$ we have the identities

$$\frac{\partial}{\partial \boldsymbol{v}} \|\boldsymbol{c} - \boldsymbol{M}\boldsymbol{v}\|^2 = -2\Re[\boldsymbol{M}^H(\boldsymbol{c} - \boldsymbol{M}\boldsymbol{v})]$$

and

$$\frac{\partial}{\partial \boldsymbol{v}} \|\boldsymbol{z}(\boldsymbol{v})\|^2 = -2\Re[(\frac{\partial}{\partial \boldsymbol{v}} \boldsymbol{z}^H(\boldsymbol{v}))\boldsymbol{z}(\boldsymbol{v})]$$

The first order derivatives are then

$$\frac{\partial l}{\partial \boldsymbol{b}} = \frac{\partial}{\partial \boldsymbol{b}}(-\frac{1}{\sigma^2}\|\boldsymbol{x} - \boldsymbol{F}(\boldsymbol{Y} \otimes \boldsymbol{\Lambda})\boldsymbol{b}\|^2) = -\frac{2}{\sigma^2}\Re[(\boldsymbol{Y} \otimes \boldsymbol{\Lambda})^H \boldsymbol{F}^H \boldsymbol{w}]$$

$$\frac{\partial l}{\partial \boldsymbol{\nu}} = \frac{\partial}{\partial \boldsymbol{\nu}}(-\frac{1}{\sigma^2}\|\boldsymbol{x} - \boldsymbol{F}(\boldsymbol{Y} \otimes \boldsymbol{\Lambda})\boldsymbol{b}\|^2) = -\frac{2}{\sigma^2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y} \otimes \boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial \boldsymbol{\nu}}\boldsymbol{w}]$$

$$\frac{\partial l}{\partial \boldsymbol{\phi}} = \frac{\partial}{\partial \boldsymbol{\phi}}(-\frac{1}{\sigma^2}\|\boldsymbol{x} - \boldsymbol{F}(\boldsymbol{Y} \otimes \boldsymbol{\Lambda})\boldsymbol{b}\|^2) = -\frac{2}{\sigma^2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y} \otimes \boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial \boldsymbol{\phi}}\boldsymbol{w}]$$

where $\boldsymbol{x} - \boldsymbol{F}(\boldsymbol{Y} \otimes \Lambda)\boldsymbol{b} = \boldsymbol{w}$ which we recognize as the noise.

B.2 Useful expressions for some derivatives

In order to determine the derivative $\frac{\partial l}{\partial \boldsymbol{\phi}}$ above we first note that $\text{vec}(\boldsymbol{\psi}_k^T) = \boldsymbol{h} \otimes \boldsymbol{r}_k$, and then look at the derivative

$$\frac{\partial}{\partial \boldsymbol{\phi}_k}(\text{vec}(\boldsymbol{P}_k^T))^T = i(\frac{\partial}{\partial \boldsymbol{\phi}_k}(\text{vec}(\psi^T))^T)(\text{vec}(\boldsymbol{P}_k^T))^T$$

with

$$\frac{\partial}{\partial \boldsymbol{\phi}_k}(\text{vec}(\boldsymbol{\psi}_k^T))^T = \frac{\partial}{\partial \boldsymbol{\phi}_k}(\boldsymbol{h}^T \otimes \boldsymbol{r}_k^T) = \boldsymbol{h}^T \otimes \boldsymbol{D}$$

where we have

$$\boldsymbol{D} = \frac{\partial \boldsymbol{r}_k^T}{\partial \boldsymbol{\phi}_k} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & \frac{d^2}{2} \end{bmatrix} \begin{bmatrix} \boldsymbol{1}_{1\times M} \\ \boldsymbol{u}_{1\times M} \\ \boldsymbol{u}_{1\times M} \odot \boldsymbol{u}_{1\times M} \end{bmatrix}$$

and $\boldsymbol{u}_{1\times M} = [0, \ldots, M-1]$.

For later use we introduce three notations for three combinations of the matrices $\boldsymbol{F}$ and $\boldsymbol{Z}$, where $\boldsymbol{F}$ is defined as earlier and $\boldsymbol{Z} = [Z_1, \ldots, Z_K]$ with $\boldsymbol{Z}_k = \text{Diag}(\text{vec}(\boldsymbol{P}_k^T))(\boldsymbol{h} \otimes \boldsymbol{I}_M)$. The first notation is

$$\boldsymbol{F}^H\boldsymbol{F} = \boldsymbol{G} = \begin{bmatrix} \boldsymbol{G}_{11} & \ldots & \boldsymbol{G}_{K1} \\ \vdots & \ddots & \vdots \\ \boldsymbol{G}_{1K} & \ldots & \boldsymbol{G}_{KK} \end{bmatrix}$$

where each entry is a matrix according to

$$\boldsymbol{G}_{lk} = \text{Diag}((\boldsymbol{P}_l^T \odot \boldsymbol{P}_k^H)\boldsymbol{1}_{N\times 1})$$

and similar to that we let

$$\boldsymbol{Z}^H\boldsymbol{F} = \boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}_{11} & \ldots & \boldsymbol{H}_{K1} \\ \vdots & \ddots & \vdots \\ \boldsymbol{H}_{1K} & \ldots & \boldsymbol{H}_{KK} \end{bmatrix}$$

where each entry is

$$\boldsymbol{H}_{lk} = \text{Diag}((\boldsymbol{P}_l^T \odot \boldsymbol{P}_k^H)\boldsymbol{h})$$

and the third one is

$$\boldsymbol{Z}^H\boldsymbol{Z} = \boldsymbol{Q} = \begin{bmatrix} \boldsymbol{Q}_{11} & \cdots & \boldsymbol{Q}_{K1} \\ \vdots & \ddots & \vdots \\ \boldsymbol{Q}_{1K} & \cdots & \boldsymbol{Q}_{KK} \end{bmatrix}$$

where each entry is

$$\boldsymbol{Q}_{lk} = \text{Diag}((\boldsymbol{P}_l^T \odot \boldsymbol{P}_k^H)(\boldsymbol{h} \odot \boldsymbol{h}))$$

When the scattering centers are well separated the block matrices $\boldsymbol{G}$, $\boldsymbol{H}$ and

$\boldsymbol{Q}$ will be approximately like identity matrices. They will contain real diagonal blocks and the non-diagonal blocks will be close to zero. We then get that the imaginary part of them is zero or approximately zero. Since some blocks of the Fischer information matrix are imaginary parts of expressions containing $\boldsymbol{G}$, $\boldsymbol{H}$ or $\boldsymbol{Q}$, these blocks will be zero when the centers are well separated. This is a truly nice property because this will make the Fischer matrix block diagonal. The CRLB matrix is then simply build by the inverse of the corresponding blocks in the Fischer matrix.

When the scattering centers are equally positioned $(\boldsymbol{P}_l^T \odot \boldsymbol{P}_k^H)\boldsymbol{1}_{N\times 1}$ will become $\boldsymbol{1}_M$ because it contains phase differences in the exponents. In the M long diagonal of for example $\boldsymbol{G}_{lk}$ element number m will be

$$\sum_{n=0}^{N-1} e^{i(\omega_0+\delta n)([r_k]_m-[r_l]_m)}$$

and now we see that $\boldsymbol{G}_{kk} = N\boldsymbol{I}_M$.The non-diagonal elements will grow as the centers separates.

The matrices $\boldsymbol{G}$, $\boldsymbol{H}$ and $\boldsymbol{Q}$ will be useful in the computing of second order derivatives for the Fischer information matrix. There is also a computational reason for building the three matrices. When it comes to programming we don't want to compute $\boldsymbol{F}$ or $\boldsymbol{Z}$ which are of size $NMK \times NMK$. It could take excessively long time. Building $\boldsymbol{G}$, $\boldsymbol{H}$ and $\boldsymbol{Q}$ is much more reasonable. They have the size $MK \times MK$. The difference between NMK and MK can be in the same range as the difference between 100 and 10000.
Now we build expressions for the $\boldsymbol{\nu}$- and $\boldsymbol{\phi}$- derivatives we saw in the first order derivatives. We have

$$\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}} = \begin{bmatrix} \frac{\partial(\boldsymbol{F}_1(e^{i\nu_1}\boldsymbol{\Lambda})\boldsymbol{b}_1)^H}{\partial\boldsymbol{\nu}_1} \\ \vdots \\ \frac{\partial(\boldsymbol{F}_K(e^{i\nu_K}\boldsymbol{\Lambda})\boldsymbol{b}_K)^H}{\partial\boldsymbol{\nu}_K} \end{bmatrix}$$

where each component is

$$\frac{\partial(\boldsymbol{F}_k(e^{i\nu_k}\boldsymbol{\Lambda})\boldsymbol{b}_k)^H}{\partial\boldsymbol{\nu}_k} = -i(\boldsymbol{F}_k(e^{i\nu_k}\boldsymbol{\Lambda})\boldsymbol{b}_k)^H$$

so that

$$\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}} = -i(\boldsymbol{I}_K\otimes\boldsymbol{1}_{1\times O})\text{Diag}(\boldsymbol{b})^H(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H$$

Further we have

$$\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}}=\begin{bmatrix}\frac{\partial(\boldsymbol{F}_1(e^{i\nu_1}\boldsymbol{\Lambda})\boldsymbol{b}_1)^H}{\partial\boldsymbol{\phi}_1}\\ \vdots\\ \frac{\partial(\boldsymbol{F}_K(e^{i\nu_K}\boldsymbol{\Lambda})\boldsymbol{b}_K)^H}{\partial\boldsymbol{\phi}_K}\end{bmatrix}$$

where each component is

$$\frac{\partial(\boldsymbol{F}_k(e^{i\nu_k}\boldsymbol{\Lambda})\boldsymbol{b}_k)^H}{\partial\boldsymbol{\phi}_k}=-i\boldsymbol{D}\text{Diag}((e^{i\nu_k}\boldsymbol{\Lambda})\boldsymbol{b}_k)^*(\boldsymbol{h}^T\otimes\boldsymbol{I}_M)\text{Diag}(\text{vec}(\boldsymbol{P}_k^T)^*)$$

because it follows from

$$\frac{\partial(\boldsymbol{F}_k(e^{i\nu_k}\boldsymbol{\Lambda})\boldsymbol{b}_k)^T}{\partial\boldsymbol{\phi}_k}=\frac{\partial}{\partial\boldsymbol{\phi}_k}(\text{Diag}(\text{vec}(\boldsymbol{P}_k^T)))(\boldsymbol{1}_{N\times1}\otimes(e^{i\nu_k}\boldsymbol{\Lambda})\boldsymbol{b}_k)^T=$$

$$=\frac{\partial}{\partial\boldsymbol{\phi}_k}(\text{vec}(\boldsymbol{P}_k^T))^T\text{Diag}(\boldsymbol{1}_{N\times1}\otimes(e^{i\nu_k}\boldsymbol{\Lambda})\boldsymbol{b}_k)=$$

$$=i(\frac{\partial}{\partial\boldsymbol{\phi}_k}(\text{vec}(\psi^T))^T)\text{Diag}(\text{vec}(\boldsymbol{P}_k^T))(\boldsymbol{I}_N\otimes\text{Diag}((e^{i\nu_k}\boldsymbol{\Lambda})\boldsymbol{b}_k))=$$

$$=i(\boldsymbol{h}^T\otimes\boldsymbol{D})(\boldsymbol{I}_N\otimes\text{Diag}((e^{i\nu_k}\boldsymbol{\Lambda})\boldsymbol{b}_k))\text{Diag}(\text{vec}(\boldsymbol{P}_k^T))=i(\boldsymbol{h}^T\otimes(\boldsymbol{D}\text{Diag}((e^{i\nu_k}\boldsymbol{\Lambda})\boldsymbol{b}_k))\text{Diag}(\text{vec}(\boldsymbol{P}_k^T))=$$

$$=i\boldsymbol{D}\text{Diag}((e^{i\nu_k}\boldsymbol{\Lambda})\boldsymbol{b}_k)(\boldsymbol{h}^T\otimes\boldsymbol{I}_M)\text{Diag}(\text{vec}(\boldsymbol{P}_k^T))$$

In the above calculations we have used the rule that, for vectors $\boldsymbol{v}$ and $\boldsymbol{u}$,

$$(\text{Diag}(\boldsymbol{u}))\boldsymbol{v}=(\text{Diag}(\boldsymbol{v}))\boldsymbol{u}$$

Now we have

$$\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}}=-i(\boldsymbol{I}_K\otimes\boldsymbol{D})\text{Diag}((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H\boldsymbol{Z}^H$$

Now we build some expressions which are parts of, and therefore will be needed in the computing of, the second order derivatives. First we build

$$\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\Lambda)\boldsymbol{b})^H}{\partial\boldsymbol{\phi}}\boldsymbol{F}=$$

$$=-i(\boldsymbol{I}_k\otimes\boldsymbol{D})\text{Diag}((\boldsymbol{Y}\otimes\Lambda)\boldsymbol{b})^H\boldsymbol{Z}^H\boldsymbol{F}=$$

$$=-i(\boldsymbol{I}_k\otimes\boldsymbol{D})\text{Diag}((\boldsymbol{Y}\otimes\Lambda)\boldsymbol{b})^H\boldsymbol{H}$$

and then we build

$$\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\Lambda)\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\Lambda)\boldsymbol{b})^H}{\partial\boldsymbol{\nu}})^H =$$

$$= -i(\boldsymbol{I}_K\otimes\boldsymbol{1}_{1\times O})\text{Diag}(\boldsymbol{b})^H(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\text{Diag}(\boldsymbol{b})(\boldsymbol{I}_K\otimes\boldsymbol{1}_{1\times O})^H i =$$

$$= (\boldsymbol{I}_k\otimes\boldsymbol{1}_{1\times O})\text{Diag}(\boldsymbol{b})^H(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{G}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\text{Diag}(\boldsymbol{b})(\boldsymbol{I}_k\otimes\boldsymbol{1}_{1\times O})^H$$

We also build the expression

$$\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\Lambda)\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\Lambda)\boldsymbol{b})^H}{\partial\boldsymbol{\phi}})^H =$$

$$= -i(\boldsymbol{I}_k\otimes\boldsymbol{1}_{1\times O})\text{Diag}(\boldsymbol{b})^H(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{Z}\text{Diag}((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})(\boldsymbol{I}_K\otimes\boldsymbol{D})^H i =$$

$$= (\boldsymbol{I}_k\otimes\boldsymbol{1}_{1\times O})\text{Diag}(\boldsymbol{b})^H(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{H}^H\text{Diag}((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})(\boldsymbol{I}_K\otimes\boldsymbol{D})^H$$

and finally we build

$$\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\Lambda)\boldsymbol{b})^H}{\partial\boldsymbol{\phi}}(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\Lambda)\boldsymbol{b})^H}{\partial\boldsymbol{\phi}})^H =$$

$$= -i(\boldsymbol{I}_K\otimes\boldsymbol{D})\text{Diag}((\boldsymbol{Y}\otimes\Lambda)\boldsymbol{b})^H\boldsymbol{Z}^H\boldsymbol{Z}\text{Diag}((\boldsymbol{Y}\otimes\Lambda)\boldsymbol{b})(\boldsymbol{I}_K\otimes\boldsymbol{D})^H i$$

$$= (\boldsymbol{I}_K\otimes\boldsymbol{D})\text{Diag}((\boldsymbol{Y}\otimes\Lambda)\boldsymbol{b})^H\boldsymbol{Q}\text{Diag}((\boldsymbol{Y}\otimes\Lambda)\boldsymbol{b})(\boldsymbol{I}_K\otimes\boldsymbol{D})^H$$

B.3 Second order derivatives

First we introduce a rule, which is easily verified by straight forward calculation, for a product, $\Re[\boldsymbol{A}]\Re[\boldsymbol{B}]^T$, where $\boldsymbol{A}$ and $\boldsymbol{B}$ are matrices:

$$\Re[\boldsymbol{A}]\Re[\boldsymbol{B}]^T = \frac{1}{2}\Re[\boldsymbol{A}\boldsymbol{B}^H] + \frac{1}{2}\Re[\boldsymbol{A}\boldsymbol{B}^T]$$

This rule will consistently be used to compute the nine blocks of the Fischer information matrix. The idea is that the second of the two terms in the sum above will turn out to be equal to zero. This fact can be understood by noting that $E[\boldsymbol{w}\boldsymbol{w}^T] = 0$. Now after also noting that $E[\boldsymbol{w}\boldsymbol{w}^H] = \sigma^2$ let us compute the second order derivatives that will represent the blocks in the Fisher information matrix. For the first row of blocks in the matrix they are

$$E[\frac{\partial l}{\partial \boldsymbol{b}}(\frac{\partial l}{\partial \boldsymbol{b}})^T] =$$

$$= E[(-\frac{2}{\sigma^2}\Re[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{w}])(-\frac{2}{\sigma^2}\Re[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{w}])^T] =$$

$$= E[\frac{4}{\sigma^4}(\frac{1}{2}\Re[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{w}\boldsymbol{w}^H\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})]+$$

$$+\frac{1}{2}\Re[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{w}\boldsymbol{w}^T(\boldsymbol{F}^H)^T((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H)^T])] =$$

$$= \frac{2}{\sigma^2}\Re[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})] = \frac{2}{\sigma^2}\Re[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{G}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})]$$

$$E[\frac{\partial l}{\partial \boldsymbol{b}}(\frac{\partial l}{\partial \boldsymbol{\nu}})^T] =$$

$$= E[(-\frac{2}{\sigma^2}\Re[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{w}])(-\frac{2}{\sigma^2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}\boldsymbol{w}])^T]$$

$$= E[\frac{4}{\sigma^4}(\frac{1}{2}\Re[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{w}\boldsymbol{w}^H(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}})^H]+$$

$$+\frac{1}{2}\Re[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{w}\boldsymbol{w}^T(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}})^T])] =$$

$$= \frac{2}{\sigma^2}\Re[((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H)(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}})^H] =$$

$$= \frac{2}{\sigma^2}\Re[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\mathrm{Diag}(\boldsymbol{b})(\boldsymbol{I}_k\otimes\boldsymbol{1}_{1\times O})^H i] =$$

$$= \frac{2}{\sigma^2}\Im[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{G}^T(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\mathrm{Diag}(\boldsymbol{b})(\boldsymbol{I}_k\otimes\boldsymbol{1}_{1\times O})^H$$

$$E[\frac{\partial l}{\partial\boldsymbol{b}}(\frac{\partial l}{\partial\boldsymbol{\phi}})^T] =$$

$$= E[(-\frac{2}{\sigma^2}\Re[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{w}])(-\frac{2}{\sigma^2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}}\boldsymbol{w}])^T] =$$

$$E[\frac{4}{\sigma^4}(\frac{1}{2}\Re[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{w}\boldsymbol{w}^H(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}})^H]+$$

$$+\frac{1}{2}\Re[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{w}\boldsymbol{w}^T(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}})^T])] =$$

$$= \frac{2}{\sigma^2}\Re[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}})^H] =$$

$$= \frac{2}{\sigma^2}\Re[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{Z}\mathrm{Diag}((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})(\boldsymbol{I}_K\otimes\boldsymbol{D})^H i] =$$

$$= \frac{-2}{\sigma^2}\Im[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{H}^H\mathrm{Diag}((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})(\boldsymbol{I}_K\otimes\boldsymbol{D})^H]$$

and for the second row of blocks they are

$$E[\frac{\partial l}{\partial \boldsymbol{\nu}}(\frac{\partial l}{\partial \boldsymbol{b}})^T] =$$

$$= E[(-\frac{2}{\sigma^2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}\boldsymbol{w}])(-\frac{2}{\sigma^2}\Re[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{w}])^T] =$$

$$= E[\frac{4}{\sigma^4}(\frac{1}{2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}\boldsymbol{w}\boldsymbol{w}^H\boldsymbol{F}((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H)^H]+$$

$$+\frac{1}{2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}\boldsymbol{w}\boldsymbol{w}^T(\boldsymbol{F}^H)^T((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H)^T])] =$$

$$= \frac{2}{\sigma^2}\Re[(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}\boldsymbol{w})((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{w})^H] =$$

$$= \frac{2}{\sigma^2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})] =$$

$$= \frac{2}{\sigma^2}\Re[-i(\boldsymbol{I}_K\otimes\boldsymbol{1}_{1\times O})\mathrm{Diag}(\boldsymbol{b})^H(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})] =$$

$$= \frac{2}{\sigma^2}\Im[(\boldsymbol{I}_K\otimes\boldsymbol{1}_{1\times O})\mathrm{Diag}(\boldsymbol{b})^H(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{G}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})]$$

$$E[\frac{\partial l}{\partial \boldsymbol{\nu}}(\frac{\partial l}{\partial \boldsymbol{\nu}})^T] =$$

$$= E[(-\frac{2}{\sigma^2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}\boldsymbol{w}])(-\frac{2}{\sigma^2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}\boldsymbol{w}])^T] =$$

$$= E[\frac{4}{\sigma^4}(\frac{1}{2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}\boldsymbol{w}\boldsymbol{w}^H(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}})^H]+$$

$$+\frac{1}{2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}\boldsymbol{w}\boldsymbol{w}^T(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}})^T])] =$$

$$= \frac{2}{\sigma^2}\Re[(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}\boldsymbol{w})(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}\boldsymbol{w})^H] =$$

$$= \frac{2}{\sigma^2}\Re[-i(\boldsymbol{I}_K\otimes\boldsymbol{1}_{1\times O})\mathrm{Diag}(\boldsymbol{b})^H(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\mathrm{Diag}(b)(\boldsymbol{I}_K\otimes\boldsymbol{1}_{1\times O})^H i] =$$

$$= \frac{2}{\sigma^2}\Re[(\boldsymbol{I}_k\otimes\boldsymbol{1}_{1\times O})\mathrm{Diag}(\boldsymbol{b})^H(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{G}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\mathrm{Diag}(b)(\boldsymbol{I}_k\otimes\boldsymbol{1}_{1\times O})^H]$$

$$E[\frac{\partial l}{\partial\boldsymbol{\nu}}(\frac{\partial l}{\partial\boldsymbol{\phi}})^T] =$$

$$= E[(-\frac{2}{\sigma^2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}\boldsymbol{w}])(-\frac{2}{\sigma^2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}}\boldsymbol{w}])^T] =$$

$$= E[\frac{4}{\sigma^4}(\frac{1}{2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}\boldsymbol{w}\boldsymbol{w}^H(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}})^H]+$$

$$+\frac{1}{2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}\boldsymbol{w}\boldsymbol{w}^T(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}})^T])] =$$

$$= \frac{2}{\sigma^2}\Re[(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}\boldsymbol{w})(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}}\boldsymbol{w})^H] =$$

$$= \frac{2}{\sigma^2}\Re[(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}})^H]$$

$$= \frac{2}{\sigma^2}\Re[-i(\boldsymbol{I}_k\otimes\boldsymbol{1}_{1\times O})\mathrm{Diag}(\boldsymbol{b})^H(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{Z}\mathrm{Diag}((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})(\boldsymbol{I}_K\otimes\boldsymbol{D})^H i] =$$

$$= \frac{2}{\sigma^2}\Re[(\boldsymbol{I}_k \otimes \boldsymbol{1}_{1\times O})\mathrm{Diag}(\boldsymbol{b})^H(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{H}^H\mathrm{Diag}((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})(\boldsymbol{I}_K\otimes\boldsymbol{D})^H]$$

and for the last row of blocks the second order derivatives are

$$E[\frac{\partial l}{\partial \boldsymbol{\phi}}(\frac{\partial l}{\partial \boldsymbol{b}})^T] =$$

$$= E[(-\frac{2}{\sigma^2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}}\boldsymbol{w}])(-\frac{2}{\sigma^2}\Re[(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H\boldsymbol{F}^H\boldsymbol{w}])^T] =$$

$$= E[\frac{4}{\sigma^4}(\frac{1}{2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}}\boldsymbol{w}\boldsymbol{w}^H\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})]+$$

$$+\frac{1}{2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}}\boldsymbol{w}\boldsymbol{w}^T(\boldsymbol{F}^H)^T((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})^H)^T])] =$$

$$= \frac{2}{\sigma^2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}}\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})] =$$

$$= \frac{2}{\sigma^2}\Re[-i(\boldsymbol{I}_K\otimes\boldsymbol{D})\mathrm{Diag}((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H\boldsymbol{Z}^H\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})] =$$

$$= \frac{2}{\sigma^2}\Im[(\boldsymbol{I}_K\otimes\boldsymbol{D})\mathrm{Diag}((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H\boldsymbol{H}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})]$$

$$E[\frac{\partial l}{\partial \boldsymbol{\phi}}(\frac{\partial l}{\partial \boldsymbol{\nu}})^T] =$$

$$= E[(-\frac{2}{\sigma^2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}}\boldsymbol{w}])(-\frac{2}{\sigma^2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}\boldsymbol{w}])^T] =$$

$$= E[\frac{4}{\sigma^4}(\frac{1}{2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}}\boldsymbol{w}\boldsymbol{w}^H(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}})^H]+$$

$$+\frac{1}{2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}}\boldsymbol{w}\boldsymbol{w}^T(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}})^T])] =$$

$$= \frac{2}{\sigma^2}\Re[(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}}\boldsymbol{w})(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\nu}}\boldsymbol{w})^H] =$$

$$= \frac{2}{\sigma^2}\Re[-i(\boldsymbol{I}_K\otimes\boldsymbol{D})\mathrm{Diag}((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H\boldsymbol{Z}^H\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\mathrm{Diag}(\boldsymbol{b})(\boldsymbol{I}_k\otimes\boldsymbol{1}_{1\times O})^H i] =$$

$$= \frac{2}{\sigma^2}\Re[(\boldsymbol{I}_K\otimes\boldsymbol{D})\mathrm{Diag}((\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H\boldsymbol{H}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\mathrm{Diag}(\boldsymbol{b})(\boldsymbol{I}_k\otimes\boldsymbol{1}_{1\times O})^H]$$

$$E[\frac{\partial l}{\partial\boldsymbol{\phi}}(\frac{\partial l}{\partial\boldsymbol{\phi}})^T] =$$

$$= E[(-\frac{2}{\sigma^2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}}\boldsymbol{w}])(-\frac{2}{\sigma^2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}}\boldsymbol{w}])^T] =$$

$$= E[\frac{4}{\sigma^4}(\frac{1}{2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}}\boldsymbol{w}\boldsymbol{w}^H(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}})^H]+$$

$$+\frac{1}{2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}}\boldsymbol{w}\boldsymbol{w}^T(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}})^T])] =$$

$$= \frac{2}{\sigma^2}\Re[\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}}(\frac{\partial(\boldsymbol{F}(\boldsymbol{Y}\otimes\boldsymbol{\Lambda})\boldsymbol{b})^H}{\partial\boldsymbol{\phi}})^H] =$$

$$= \frac{2}{\sigma^2}\Re[-i(\boldsymbol{I}_K \otimes \boldsymbol{D})\mathrm{Diag}((\boldsymbol{Y} \otimes \boldsymbol{\Lambda})\boldsymbol{b})^H \boldsymbol{Z}^H \boldsymbol{Z}\mathrm{Diag}((\boldsymbol{Y} \otimes \boldsymbol{\Lambda})\boldsymbol{b})(\boldsymbol{I}_K \otimes \boldsymbol{D})^H i] =$$

$$= \frac{2}{\sigma^2}\Re[(\boldsymbol{I}_K \otimes \boldsymbol{D})\mathrm{Diag}((\boldsymbol{Y} \otimes \boldsymbol{\Lambda})\boldsymbol{b})^H \boldsymbol{Q}\mathrm{Diag}((\boldsymbol{Y} \otimes \boldsymbol{\Lambda})\boldsymbol{b})(\boldsymbol{I}_K \otimes \boldsymbol{D})^H]$$

So, using these blocks to build the Fischer information matrix and then inverting it results in the CRLB matrix for the parameter vector $[\boldsymbol{b}, \boldsymbol{\nu}, \boldsymbol{\phi}]$ which is to be estimated.

References

[1] Steven M. Kay, *Fundamentals of statistical signal processing estimation theory*, Prentice-Hall, 1993.

[2] Roland Jonsson, *Scattering Center Parameter Estimation from RCS-prediction Data*, Ericsson internal report, Jul. 2001.

[3] Roland Jonsson, Jan O. Hagberg and Fredrik Dicander, *Mastering Interfering Scatterers in HRR Data for ATR Systems*, Proceedings of SPIE, Vol. 4053, Apr. 2000.

[4] Benjamine Friedlander and Joseph M. Francos, *Estimation of Amplitude and Phase Parameters of Multicomponent Signals*, IEEE Transactions on signal processing, Vol. 43, Apr. 1995.

[5] Hung-Chih Chiang and Randolph L. Moses, *ATR Performance Prediction Using Attributed Scattering Features*, Proceedings of SPIE, Vol. 3721, Apr. 1999.

[6] Jan R. Magnus and Heinz Neudecker, *Matrix differential calculus with applications in statistics and econometrics*, John Wiley and Sons, 1999.

[7] Jian Li and Petre Stoica, *Efficient Mixed-Spectrum Estimation with Applications to Target Feature Extraction*, IEEE Transactions on signal processing, Vol. 44, Feb. 1996.

[8] Lennart Rade and Bertil Westergren, *Mathematics Handbook for Sience and engineering (BETA)*, Studentlitterattur, 1998.

[9] J. Li, R. Wu, Z. Bi, P. Stoica and E. G. Zelnio, *A robust semi-parametric method for feature extraction and SAR image formation of targets consisting of trihedrals and dihedrals*, Proceedings of SPIE, Vol. 3721, Apr. 1999.