On The Existence of Uniformly Unbiased Estimators

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE M.Sc DEGREE

Supervised by: Dr. Joseph Tabrikian

By Roni Winik

December 2010
BEN GURION UNIVERSITY OF THE NEGEV
FACULTY OF ENGINEERING SCIENCES
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

On The Existence of Uniformly Unbiased Estimators

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE M.Sc. DEGREE

By Roni Winik

Supervised by: Dr. Joseph Tabrikian

December 2010

Author: ___________________________ Date: ___________

Supervisor: _________________________ Date: ___________

Chairman of Graduate Studies Committee: _________ Date: ___________
Abstract

In this thesis, the problem of minimum-variance unbiased (MVU) parameter estimation in low signal-to-noise ratios (SNRs) or few number of samples was studied. Performance lower bounds are commonly used in statistical signal processing for system design and analysis. In most practical problems there exists a threshold SNR below which the performance of the estimators, such as the maximum likelihood estimator (MLE) rapidly degrades. This phenomenon cannot be predicted by the Cramér-Rao bound, therefore, different approximations of the Barankin bound are usually used to find the threshold region. In this work, we show some popular problems in which the uniformly unbiased assumption is too demanding, and in fact no uniformly unbiased estimator exists. In particular, we analytically show that in periodic problems the Barankin bound can go to infinity by taking test-points at specific locations. This proof was extended also to the case of Gaussian models with bounded expectation which carries the information about the unknown parameters of interest. The conclusions of this work are as follows. First, calculation of the Barankin bound is heavily dependent on the chosen test-points, and in some problems it can go to infinity. Accordingly, the Barankin bound in these problems is meaningless. Second, the uniformly unbiasedness constraint is too demanding and may not be satisfied in many practical cases. Therefore, new performance lower bounds with modified unbiasedness constraints should be investigated.
Acknowledgment

The author wishes to express his gratitude to his supervisor, Dr. Joseph Tabrikian who was abundantly helpful and offered invaluable assistance, support and guidance. Deepest gratitude are also due to Koby Todros.
Contents

1 Introduction 1

2 Background 5

2.1 Preliminaries .............................................. 5
2.2 The Barankin Theorem: Review ............................. 6
2.3 Approximation of the Barankin bound ..................... 8
   2.3.1 The Cramér-Rao lower bound ......................... 9
   2.3.2 Other lower bounds on the mean-square error of unbiased estimators 9
2.4 Estimation via the scoring method - theorem and disadvantages 10

3 On the Meaningless of Barankin Bound Approximations in Common Estimation Problems 12

3.1 The nonexistence of UU estimators ....................... 12
   3.1.1 Periodic likelihood functions case .................. 13
   3.1.2 Parameter estimation of a bounded deterministic signal in Gaussian noise .......................... 15
3.2 Examples ................................................................. 16
  3.2.1 Single tone estimation ............................................ 17
  3.2.2 Exponential parameter estimation in Gaussian noise ............ 21
3.3 Calculation of the approximated MH bound in single-tone parameter estima-
  tion ................................................................. 24
3.4 Appendix .......................................................... 29

4 Approximated minimum-variance unbiased estimator ...................... 31
  4.1 A new approach to minimum-variance unbiased estimation .......... 31
  4.2 MVU estimation via an integral equation .......................... 32
  4.3 Approximated MVU estimator ....................................... 33
    4.3.1 Derivation of scoring method from the approximated MVU estimator 36
    4.3.2 Approximated MVU estimator based on the Barankin kernel ........ 37
  4.4 Bias evaluation of the approximated MVU estimator based on the Barankin
    kernel ............................................................ 38
  4.5 MSE evaluation of the approximated MVU estimator based on the Barankin
    kernel ............................................................ 40
  4.6 Examples and simulation ........................................... 40
    4.6.1 Example 1 .................................................... 40
    4.6.2 Example 2 .................................................... 43

5 Summary ........................................................................... 46
  5.1 Discussion and Conclusions ........................................... 46
Abbreviations

AWGN - Additive white Gaussian noise
WGN - White Gaussian noise
MLE - Maximum likelihood estimator
MSE - Mean-square error
pdf - Probability density function
SNR - Signal-to-noise ratio
CRLB - Cramér-Rao lower bound
BB - Barankin bound
HCR - Hammersley-Chapman-Robbins
MH - McAulay-Hofstetter
UUB - Uniformly unbiased
r.h.s - right hand side
MVU - Minimum variance unbiased
AMVU - Approximate minimum variance unbiased
RBLs - Rao-Blackwell-Lehmann-Scheffe
LS - Least squares
EM - Expectation-maximization
LR - Likelihood ratio
STD - Standard deviation
Notations

\((\cdot)^{-1}\) - Matrix inverse

\((\cdot)^T\) - Matrix transpose

\((\cdot)^*\) - Conjugate

\((\cdot)^H\) - Conjugate transpose (Hermitian)

diag(\(\cdot\)) - Diagonal matrix with the same diagonal as its matrix argument

\(|\cdot|\) - Absolute value of a scalar

Re - Real part of a complex number

Im - Imaginary part of a complex number

I - Identity matrix

\([A]_{ij}\) - \((i,j)\)th element of the matrix \(A\)

\([a]_i\) - \((i)\)th element of the vector \(a\)

\(\theta_0\) - Unknown parameter to be estimated

\(g(\theta_0)\) - Function to be estimated

\(\Theta\) - Parameter space

\(\mathbb{R}\) - Real numbers space

\(\mathbb{C}\) - Complex numbers space

\(L_2(\Omega, F, P_{\theta_0})\) - Probability space with finite second moment

\(\Omega\) - Obseration space

\(F\) - \(\sigma\)-algebra on \(\Omega\)
\( p_{\theta_0} \) - Probability measure w.r.t. \( \theta_0 \)

\( \mathbf{x} \) - Observation vector

\( f(\mathbf{x}; \theta) \) - pdf of the random vector \( \mathbf{x} \), with deterministic parameter \( \theta \).

\( \hat{g}(\mathbf{x}) \) - Estimator of \( g(\cdot) \)

\( MSE[\cdot; \theta_0] \) - MSE of an estimator at \( \theta_0 \)

\( E[\cdot; \theta] \) - Expectation using pdf parameterized by \( \theta \)

\( L(\mathbf{x}; \theta, \theta_0) \) - Likelihood ratio

\( \phi \) - Class of UUB estimators

\( N \) - Number of samples

\( W \) - Period of likelihood function

\( M \) - Number of test-points

\( \mathbf{a}(\cdot) \) - Main signal or steering vector

\( \mathbf{n} \) - WGN

\( s \) - Amplitude of signal

\( J^{-1} \) - Fisher information matrix

\( \mathbf{t} \) - test-points vector

\( \mathbf{B} \) - Barankin matrix

\( \mathbf{d} \) - Barankin-Cramér-Rao Projection vector

\( \delta \) - Proximity of the second test-point to the full cycle of the likelihood function

\( \mathbf{C}_{MH} \) - McAulay-Hofstetter bound

\( \mathbf{C}_{HCR} \) - Hammersley-Chapman-Robbins bound
List of Figures

3.1 Comparison of MLE, CRLB and the MH bounds for various values of δ versus SNR. ................................. 17
3.2 Threshold SNR of the MLE and MH bound versus δ. ....................... 18
3.3 Comparison of the MH bounds and CRLB for various SNRs versus δ. ... 18
3.4 Comparison of MLE, CRLB and the MH bounds for various θ_M, versus the number of samples, N. ......................... 22
3.5 Threshold region of the MH bounds relative to number of samples for various values of θ_M, in comparison with the threshold region of the MLE. ... 23
3.6 Comparison of the MH bounds and CRLB for different number of samples, versus θ_M. ................................. 24
4.1 Bias versus the number of samples. .............................................. 42
4.2 RMSE versus the number of samples. ........................................... 42
4.3 Bias of the approximated MVU estimator for a specific θ_k, versus θ_0. ... 44
4.4 MSE of the approximated MVU estimator for each θ_k, versus θ_0. ........ 44
List of Tables

4.1  Iterations of the Barankin kernel based estimator. .................. 41
Chapter 1

Introduction

Lower bounds on the mean-square-error (MSE) of unbiased estimators enable performance prediction and constitute a benchmark for performance evaluation. Under the assumption of uniform unbiasedness over the entire parameter space, the Barankin bound (BB) is the tightest lower bound on the MSE of any uniformly unbiased (UU) estimator.

It is well known that under some mild assumptions the maximum likelihood estimator (MLE) achieves the Cramér-Rao lower bound (CRLB) [1, 2], asymptotically, where the number of samples or the signal-to-noise (SNR) goes to infinity [3]. However, in most practical problems, there exists a threshold region below which the performance of the MLE rapidly degrades. This phenomenon cannot be predicted by the CRLB, therefore, different versions of the BB are usually used to find the threshold region [6]-[16].

Some existing MSE bounds may be interpreted as the minimum MSE obtained over all estimators whose bias satisfies certain constraints [35]. The CRLB is derived from a weak formulation of unbiasedness at the vicinity of the true value of the unknown pa-
rameter, such that, the derivative of the bias at the true parameter is constrained to be zero. The popularity of the CRLB is largely due to its simplicity, and the fact that under some mild conditions it can be achieved by the MLE [3]. This initial characterization of locally unbiased estimators has been improved by Bhattacharyya [4], which used local statistical information, similar to the CRLB, such that, higher-order derivatives of the bias at the true parameter are constrained to be zero. Generalization of the bound was made by Barankin [5], who established the general form of the greatest lower bound of any uniformly unbiased estimator with finite MSE. The exact BB assumes strict uniform unbiasedness, while the MLE is not necessarily a uniform unbiased estimator in the threshold region. Accordingly, the BB may be inadequate for prediction of the MLE threshold region.

The BB is generally incomputable and numerous works [17]-[23] have been devoted to deriving computable approximations of the BB. These approximations involve the unbiasedness constraint, only on a chosen set of points in the parameter space, named, test-points. Recently, several new approximations of the BB have been proposed [9]-[16]. The BB approximations are usually used to predict the threshold region of the MLE. The quality of the BB approximations has been mainly measured by their ability to predict this threshold region.

In this work, we show that in some common problems this task is meaningless. By using Barankin theorem [5], we analytically show that in problems with periodic likelihood function, the constraint of unbiasedness at specific locations can result in a bound which goes to infinity. Moreover, we show that BB approximations can be heavily dependent on
the chosen test-points and consequently, the prediction of the threshold region is greatly dependent on the chosen test-points. This threshold can be below or above the MLE threshold. Accordingly, the BB approximations in these problems can not be used for finding the threshold region of the MLE. Note that the MLE performance can potentially be better than the performance predicted by the bound because the MLE is biased.

The proof of the nonexistence of UU estimator is extended also to the case of Gaussian models with bounded mean, which carries the information about the unknown parameters of interest. We show that the uniformly unbiasedness constraint is too demanding and may not be satisfied. In particular, the nonexistence of UU estimator is exemplified in an exponential parameter estimation problem with Gaussian observation model. Furthermore, we show that calculation of the BB is heavily dependent on the chosen test-points, and can go to infinity with specific selection of the test-points. We conclude that the BB in these problems is meaningless, and its approximations are not useful in finding the threshold phenomenon of the MLE.

Additionally, a new approach to minimum-variance unbiased estimates is presented. In [27] it was shown that if a MVU estimator exists, it can be obtained by applying a linear transformation on the likelihood ratio (LR) function. This linear transformation can be obtained via the solution of an integral equation, which in many cases can not be solved in a direct manner. Therefore, in practice, calculation of the linear transformation of the LR function is intractable. In order to overcome this limitation, an approximation of the integral equation is performed. The integral transform generalizes the derivative operator applied in the scoring method [3] for obtaining the MLE. Therefore, by modifying
the kernel of the integral transform, a variety of new approximate MVU estimators can be derived. We note that a similar approach was adopted in [30] to obtain a tight and computationally simple lower bound for parameter estimation.
Chapter 2

Background

2.1 Preliminaries

We begin by stating the following definitions and assumptions.

- **Parameter space and function to be estimated**: The estimation of $g(\theta_0)$ is considered, where $g : \Theta \rightarrow \mathbb{R}$ is deterministic known, $\theta_0 \in \Theta$ is deterministic unknown, and $\Theta$ is the parameter space.

- **Probability space**: The space $(\Omega, \mathcal{F}, P_{\theta_0})$ denotes a probability space with a finite second statistical moment w.r.t. $P_{\theta_0}$, where $\Omega$ is an observation space of points $x \in \Omega$, $\mathcal{F}$ is a $\sigma$-algebra on $\Omega$, and $\{P_\theta, \theta \in \Theta\}$ is a family of probability measures.

- **Likelihood function**: The family $P_\theta$ is assumed to have density functions $f(x; \theta)$, such that the probability of observing $A \in \mathcal{F}$ is,

$$P_\theta (A) \triangleq \int_A f(x; \theta) dx.$$
Given the random observations vector $x$, the family $\{f(x; \theta), \theta \in \Theta\}$ is termed as the likelihood function.

- **Locally unbiased estimator**: $\hat{g}(x)$ is a locally unbiased estimator of $g(\theta_0)$, if
  \[
  E[\hat{g}(x); \theta_0] \triangleq \int_{\Omega} \hat{g}(x)f(x; \theta_0)dx = g(\theta_0),
  \]
  and
  \[
  \frac{d}{d\theta}E[\hat{g}(x); \theta] \bigg|_{\theta_0} = 1.
  \]

- **Uniformly unbiased (UU) estimator**: $\hat{g}(x)$ is a UU estimator of $g(\theta)$, if
  \[
  E[\hat{g}(x); \theta] \triangleq \int_{\Omega} \hat{g}(x)f(x; \theta)dx = g(\theta), \quad \forall \theta \in \Theta.
  \]

- **Mean-square-error (MSE)**: The MSE of an estimator $\hat{g}(x)$, evaluated at $\theta_0$ is given by
  \[
  MSE[\hat{g}(x); \theta_0] \triangleq \int_{\Omega} (\hat{g}(x) - g(\theta_0))^2f(x; \theta_0)dx. \tag{2.1}
  \]

### 2.2 The Barankin Theorem: Review

In this section, we review Barankin’s theorem [5] regarding the existence of UU estimators and the BB is presented. In [5], Barankin investigated the existence and the characterization of a UU estimator $\hat{g}(x)$, which is “best” for a prescribed parameter point $\theta_0 \in \Theta$, in the MSE sense. In Section 3.1, we show via Barankin theorem, that the assumption of uniform unbiasedness can not be satisfied in common estimation problems.

The Barankin theorem states a necessary and sufficient condition for the existence of a UU estimator.
**Theorem 1** (Barankin theorem [5]). Let \( \phi \) denote the class of all UU estimates of \( g(\theta) \), defined on \( \Theta \), with finite MSE. A necessary and sufficient condition for \( \phi \) to be non-empty is that there exists a constant \( C \) such that for every set of \( n \) functions \( L(x; \theta_1, \theta_0), \ldots, L(x; \theta_n, \theta_0) \), and every set of \( n \) real numbers \( a_1, \ldots, a_n \), we have,
\[
\left| \sum_{i=1}^{n} a_i [g(\theta_i) - g(\theta_0)] \right|^2 \leq C \int_{\Omega} \left| \sum_{i=1}^{n} a_i L(x; \theta_i, \theta_0) \right|^2 f(x; \theta_0)dx, \quad n = 1, 2, \ldots \quad (2.2)
\]

The family \( \theta_i \in \Theta, i = 1, \ldots, n \) is termed as test-points, and \( a_i \in \mathbb{R}, \ i = 1, \ldots, n \).

**Proof.** The proof of this theorem is given in detail in [5].

The following corollary states the relation of Theorem 1 to the BB.

**Corollary 2** ([5]). The infimum of the set of admissible constants \( C \) in (2.2) constitutes the BB. Thus, a necessary and sufficient condition such that a UU estimator exists with a finite MSE, is that the BB is finite.

The BB is defined as
\[
C_{BB} \triangleq \sup_{\{\theta_i, a_i\}_{i=1}^{n}} \frac{(\sum_{i=1}^{n} a_i[g(\theta_i) - g(\theta_0)])^2}{\int_{\Omega} (\sum_{i=1}^{n} a_i L(x; \theta_i, \theta_0))^2 f(x; \theta_0)dx}. \quad (2.3)
\]
where
\[
L(x; \theta_i, \theta_0) \triangleq \frac{f(x; \theta_i)}{f(x; \theta_0)}, \quad (2.4)
\]
is the likelihood ratio (LR) function. The supremum is to be taken over all \( n = 1, 2, \ldots \), and all families of \( \theta_i \in \Theta \) and \( a_i \in \mathbb{R}, \ i = 1, \ldots, n \). The \( MSE[\hat{g}(x); \theta_0] \) of the estimator \( \hat{g}(x) \) is lower bounded according to
\[
MSE[\hat{g}(x); \theta_0] \geq C_{BB}. \quad (2.5)
\]
Barankin showed [5] that this is the tightest bound in the sense that if a UU estimator of \( g(\theta) \) exists, then there exists a unique UU estimator that achieves this bound. However, the estimator that achieves the BB in (2.5) may depend upon the parameter \( \theta_0 \). When this occurs, there are only “locally best” UU estimates for \( g(\theta) \). However, in many practical problems, a UU estimator does not exist and as Corollary 2 states, the BB in those cases goes to infinity.

### 2.3 Approximation of the Barankin bound

“Small-error” and “large-error” bounds may be interpreted as the minimum MSE obtained over all estimators whose bias satisfies certain constraints [35]. The CRLB is derived from a weak formulation of unbiasedness at the vicinity of the true value of the unknown parameter. Its popularity is largely due to its simplicity, and the fact that in many cases it can be achieved by the MLE [3]. This initial characterization of locally unbiased estimators has been improved by Bhattacharyya [4], which used local statistical information, similar to the CRLB. Generalization of the bound was made by Barankin [5], who established the general form of the greatest lower bound of any uniformly unbiased estimator with finite MSE. However, the BB is generally incomputable and numerous works [17]-[23] have been devoted to deriving computable approximation of the BB. These approximations involve the unbiasedness constraint, only on a chosen set of points in the parameter space, named, test-points.
2.3.1 The Cramér-Rao lower bound

In this subsection, the CRLB theorem is presented.

**Theorem 3.** It is assumed that the probability density function (pdf), \( f(x; \theta) \), satisfies the regularity condition
\[
E \left[ \frac{\partial \log f(x; \theta)}{\partial \theta} ; \theta \right] = 0, \ \forall \theta \in \Theta,
\]
where the expectation is taken with respect to \( f(x; \theta) \). Then the CRLB states that the MSE of any locally unbiased estimator \( \hat{g}(x) \) must satisfy
\[
MSE[\hat{g}(x); \theta_0] \geq \frac{1}{-E \left[ \frac{\partial^2 \log f(x; \theta)}{\partial \theta^2} \bigg|_{\theta = \theta_0} ; \theta_0 \right]}, \quad (2.6)
\]
where the derivative is evaluated at the true value of \( \theta \) and the expectation is taken with respect to \( f(x; \theta) \). Furthermore, an unbiased estimator may be found that attains the bound for all \( \theta \) if and only if there exists a function \( \hat{g}(x) \) such that
\[
\frac{\partial \log f(x; \theta)}{\partial \theta} = J(\theta)(\hat{g}(x) - \theta) \ \forall \theta \in \Theta, \quad (2.7)
\]
where \( J(\theta) = -E \left[ \frac{\partial^2 \log f(x; \theta)}{\partial \theta^2} ; \theta \right] \). The function \( \hat{g}(x) \) is the MVU estimator of \( g(\theta) \) and its variance is \( 1/J(\theta) \).

**Proof.** The proof of this theorem is given in detail in [3].

2.3.2 Other lower bounds on the mean-square error of unbiased estimators

In this subsection, we represent the McAulay-Hofstetter (MH) bound [20], which is an approximation of the BB. The MH bound involves unbiasedness constraint at the true
parameter and at the “test-points”. The MH bound is given by [20]

\[ C_{MH} = J^{-1} + q^T R^{-1} q, \]  

(2.8)

where

\[ q = t - J^{-1} d, \]  

(2.9)

\[ R = B - d J^{-1} d^T, \]  

(2.10)

and the elements of the vector \( d \) and \( t \) are given by

\[ [d]_j \triangleq \mathbb{E} \left[ \frac{\partial \log f(x; \theta)}{\partial \theta} \bigg| L(x; \theta_j, \theta_0); \theta_0 \right], \quad j = 1, \ldots, M \]  

(2.11)

\[ [t]_j \triangleq \theta_j - \theta_0, \quad j = 1, \ldots, M. \]  

(2.12)

and \( \theta_1, \ldots, \theta_M \) denote the \( M \) chosen test-points,

\[ J(\theta_0) \triangleq \mathbb{E} \left[ \left( \frac{\partial \log f(x; \theta)}{\partial \theta} \bigg| \theta = \theta_0 \right)^2 ; \theta_0 \right], \]  

(2.13)

\( J(\cdot) \) denotes the Fisher information [3], and the matrix \( B \) denotes the Barankin matrix whose elements are given by

\[ [B]_{ij} \triangleq \mathbb{E} [L(x; \theta_i, \theta_0) L(x; \theta_j, \theta_0); \theta_0] \quad i, j = 1, \ldots, M. \]  

(2.14)

### 2.4 Estimation via the scoring method - theorem and disadvantages

In this section, the scoring method [3] is presented. The purpose of this method is to find the MLE estimate in an iterative manner. This method is useful especially when the ML
estimator can not be derived explicitly. The scoring method is based on a common optimization method named Newton-Raphson method, which is designed to find a maximum of a function. The recursive equation of the scoring method is given by,

\[
\theta_{k+1} = \theta_k - J^{-1}(\theta_k) \frac{\partial \log f(x; \theta)}{\partial \theta} \bigg|_{\theta=\theta_k}
\]  

(2.15)

where \( J(\theta_k) \) is evaluated at \( \theta_k \). This method has several disadvantages:

1. Convergence to the global maximum is not guaranteed and the algorithm can converge to a local maximum, depending on the initial point.

2. In cases where an efficient estimator does not exist, the MLE is not the MVU estimator. Thus, the MLE is not optimal in terms of MVU. Since the scoring method is designed to find the MLE estimate, it also would not be optimal.

3. In cases where the Fisher information is nearly singular, the scoring method may diverge.

4. There is no control on the convergence speed.
Chapter 3

On the Meaningless of Barankin Bound Approximations in Common Estimation Problems

3.1 The nonexistence of UU estimators

In this section, we show via Theorem 1 that in some common signal processing cases a UU estimator does not exist. First, we investigate cases, in which the likelihood function $f(x; \theta)$ is periodic in $\theta$, with period $W$. Second, we examine a complex additive white Gaussian noise (AWGN) observation model, in which the norm of the mean is bounded.
3.1.1 Periodic likelihood functions case

In this subsection, it is shown that in cases where the likelihood function, \( f(x; \theta) \), is periodic in \( \theta \), a UU estimator of \( g(\theta) = \theta \) does not exist.

**Theorem 4.** Assume that \( g(\theta) = \theta \) and \( f(x; \theta) \) is continuously differentiable w.r.t. \( \theta \in \mathbb{R}, \forall x \in \Omega \). If \( \forall x \in \Omega, f(x; \theta) \) is periodic in \( \theta \in \mathbb{R} \) with period \( W \), then the class of all UU estimates of \( g(\theta) = \theta \), denoted by \( \phi \), is empty.

**Proof.** Consider the inequality (2.2) and let \( \{a_1 = 1, a_2 = -1\} \). For any \( \theta_0 \in [-W/2, W/2) \), the inequality in (2.2) can be rewritten as

\[
|\theta_1 - \theta_2|^2 \leq C \int_{\Omega} [L(x; \theta_1, \theta_0) - L(x; \theta_2, \theta_0)]^2 f(x; \theta_0) dx. \tag{3.1}
\]

Choose \( \{\theta_1 = -W/2, \theta_2 = W/2 - \delta\} \) for \( \delta > 0 \), where \( \delta \) denotes the proximity of the second test-point to the full cycle of the likelihood function. Thus, by substituting (2.4) into (3.1) it is implied that

\[
|W - \delta|^2 \leq C \int_{\Omega} \left[ \frac{f(x; -W/2) - f(x; W/2 - \delta)}{f(x; \theta_0)} \right]^2 f(x; \theta_0) dx. \tag{3.2}
\]

Since \( f(x; \theta) \) is periodic in \( \Theta \) with period \( W \), (3.2) can be rewritten as

\[
|W - \delta|^2 \leq C \int_{\Omega} \left[ \frac{f(x; -W/2) - f(x; -W/2 - \delta)}{f(x; \theta_0)} \right]^2 f(x; \theta_0) dx. \tag{3.3}
\]

Hence, by setting \( \delta \to 0 \), the following inequality is obtained

\[
|W|^2 \leq C \lim_{\delta \to 0} \int_{\Omega} \frac{1}{f(x; \theta_0)} \cdot \frac{f(x; -W/2) - f(x; -W/2 - \delta)}{\delta} \right]^2 f(x; \theta_0) dx. \tag{3.4}
\]

Thus,

\[
|W|^2 \leq C \left( \lim_{\delta \to 0} \delta^2 \right) \cdot \lim_{\delta \to 0} \int_{\Omega} \left[ \frac{1}{f(x; \theta_0)} \cdot \frac{f(x; -W/2) - f(x; -W/2 - \delta)}{\delta} \right]^2 f(x; \theta_0) dx. \tag{3.5}
\]
Under the condition that $f(x; \theta)$ is differentiable w.r.t. $\theta \forall x \in \Omega$, with bounded differential, the dominated convergence theorem [32] states that

$$\lim_{\delta \to 0} \int_{\Omega} \left[ \frac{1}{f(x; \theta_0)} \cdot \frac{f(x; -W/2) - f(x; -W/2 - \delta)}{\delta} \right]^2 f(x; \theta_0) dx = \int_{\Omega} \left[ L'(x; -W/2, \theta_0) \right]^2 f(x; \theta_0) dx$$

(3.6)

where

$$L'(x; -W/2, \theta_0) = \frac{\partial L(x; \theta, \theta_0)}{\partial \theta} \bigg|_{\theta = -W/2}.$$

By substituting (3.6) into (3.5), we obtain

$$|W|^2 \leq C \cdot \left( \lim_{\delta \to 0} \delta^2 \right) \int_{\Omega} \left[ L'(x; -W/2, \theta_0) \right]^2 f(x; \theta_0) dx.$$  

(3.7)

Assuming that $\int_{\Omega} \left[ L'(x; -W/2, \theta_0) \right]^2 f(x; \theta_0) dx$ is finite, (3.7) can rewritten as

$$|W|^2 / \left( \lim_{\delta \to 0} \delta^2 \right) \leq C_1,$$

(3.8)

where $C_1 = C \cdot \int_{\Omega} \left[ L'(x; -W/2, \theta_0) \right]^2 f(x; \theta_0) dx$. Equation (3.8) implies that $C_1$ is not bounded and therefore the class of UU estimators $\phi$, defined on $[-W/2, W/2)$, is empty.

Theorem 4 states that if the likelihood function, $f(x; \theta)$, is periodic in $\Theta$, then a UU estimator of $g(\theta) = \theta$ does not exist. Hence, in this case, by Corollary 2, the BB is infinite, and it can not be used for prediction of the threshold region.
3.1.2 Parameter estimation of a bounded deterministic signal in Gaussian noise

In this subsection, the problem of bounded deterministic signal parameter estimation in Gaussian noise is investigated. It is shown that a UU estimator of \( g(\theta) = \theta \) does not exist in cases with bounded deterministic signal, which carry the information on the unknown parameter.

**Theorem 5.** Assume the following model,

\[
x = sa(\theta_0) + n
\]

(3.9)

where \( n \) denotes an \( N \times 1 \) proper complex Gaussian noise with known covariance matrix \( R_n \), \( a(\cdot) \) denotes a known \( N \times 1 \) vector function, \( x \) is the \( N \times 1 \) observation vector and \( s \in \mathbb{C} \) is a known signal amplitude. We will show that in cases where \( \|a(\theta)\| \in [a, b] \) for all \( \theta \in \mathbb{R} \), such that \( a, b < \infty \), a UU of \( g(\theta) = \theta \) does not exist and therefore, the BB goes to infinity even for large number of samples, \( N \).

*Proof.* The pdf of the observation vector \( x \) can be written as

\[
f(x; \theta) = \frac{1}{\det(\pi R_n)} \exp \left\{ -\| x - sa(\theta) \|^2_{R_n} \right\},
\]

(3.10)

where \( \|z\|^2_{R_n} \overset{\Delta}{=} z^H R_n^{-1} z, z \in \mathbb{C}^N \), and \( (\cdot)^H \) denotes the Hermitian operator. By setting \( a_1 = 1 \) and \( n = 1 \) one can notice that (1) can be rewritten as

\[
|\theta_1 - \theta_0|^2 \leq C \int_{\Omega} [L(x; \theta_1, \theta_0)]^2 f(x; \theta_0) dx.
\]

(3.11)
Using (3.10), the LR in (2.4) can be expressed as

\[ L(x; \theta_1, \theta_0) = \exp \left\{ - \left[ \| x - s a(\theta_1) \|^2_{\mathbb{R}^n} + \| x - s a(\theta_0) \|^2_{\mathbb{R}^n} \right] \right\}. \]  (3.12)

By substituting (3.12) into (3.11), followed by using the result in [7], and Eqs. (11)-(16), one obtains

\[ |\theta_1 - \theta_0|^2 \leq C \exp \left\{ 2|s|^2 \| a(\theta_1) - a(\theta_0) \|^2_{\mathbb{R}^n} \right\}. \]  (3.13)

Since \( \|a(\cdot)\| \in [a, b] \), the exponential term in (3.13) is finite. Therefore, if \( \theta_1 \to \pm \infty \), then \( C \to \infty \), and hence by Theorem 4, a UU estimator does not exist in this case.

We have shown that a UU estimator of \( g(\theta) = \theta \) does not exist in the problem of parameter estimation of a bounded deterministic signal in Gaussian noise. Hence, by Corollary 2, the BB is infinity and can not be used for prediction of the threshold region of the MLE.

### 3.2 Examples

Theorems 4 and 5 showed that in some common estimation problems, UU estimators do not exist. In practice, BB approximations are usually used to predict the threshold region of the MLE. In this section, we show that the BB approximations can be heavily affected by the locations in which the unbiasedness is constrained and therefore, the corresponding predicted threshold region is meaningless in the problems of single tone and exponential parameter estimation.
In the problems of single tone and exponential parameter estimation, the observation model is given by

$$x = sa(\theta_0) + n$$

(3.14)

where \(x\) denotes an \(N \times 1\) observation vector, \(s \in \mathbb{R}\) is a known signal amplitude. \(a(\cdot)\) is an \(N \times 1\) known vector function, \(\theta_0\) is the parameter to be estimated, and \(n\) denotes an \(N \times 1\) proper complex white Gaussian noise (WGN) vector, with known covariance matrix \(C_n = \sigma^2 I\).

### 3.2.1 Single tone estimation

![Figure 3.1: Comparison of MLE, CRLB and the MH bounds for various values of \(\delta\) versus SNR.](image)

In this example, the entries of \(a(\theta_0)\) are given by \([a(\theta_0)]_k = \exp(i2\pi(k - 1)\theta_0), \ k = \ldots\)
Figure 3.2: Threshold SNR of the MLE and MH bound versus $\delta$.

Figure 3.3: Comparison of the MH bounds and CRLB for various SNRs versus $\delta$.

1, \ldots, N$ where $\theta_0 \in [-W/2, W/2)$ is the true tone to be estimated. In this problem, the likelihood function $f(x; \theta)$ is periodic in $\theta$ with period $W$. Therefore, by Theorem 4,
the BB is infinite. The meaningless of the BB approximations is being shown via the derivation of the McAulay-Hofstetter (MH) bound [20], which is an approximation of the BB. The MH bound is given by [20]

\[ C_{MH} = J^{-1} + q^T R^{-1} q, \]  

(3.15)

where

\[ q = t - J^{-1} d, \]  

(3.16)

\[ R = B - d J^{-1} d^T, \]  

(3.17)

and the elements of the vector \( d \) and \( t \) are given by

\[ [d]_j \triangleq \int_{\Omega} \frac{\partial \log f(x; \theta)}{\partial \theta} \bigg|_{\theta = \theta_0} L(x; \theta_j, \theta_0)f(x; \theta_0)dx \quad j = 1, \ldots, M \]  

(3.18)

\[ [t]_j \triangleq \theta_j - \theta_0 \quad j = 1, \ldots, M \]  

(3.19)

and \( \theta_1, \ldots, \theta_M \) denote the \( M \) chosen test-points, \( J \) denotes the Fisher information [3]

\[ J \triangleq \int_{\Omega} \left( \frac{\partial \log f(x; \theta)}{\partial \theta} \bigg|_{\theta = \theta_0} \right)^2 f(x; \theta_0)dx, \]  

(3.20)

and \( B \) denotes the Barankin matrix, whose elements are given by

\[ [B]_{ij} \triangleq \int_{\Omega} L(x; \theta_i, \theta_0)L(x; \theta_j, \theta_0)f(x; \theta_0)dx \quad i, j = 1, \ldots, M \]  

(3.21)

Figs. 3.1-3.3 depict the MH bound, for \( M = 2, \theta_1 = -W/2, \theta_2 = W/2 - \delta \), where \( \delta \) denotes the proximity of the second test-point to the full cycle of the likelihood function and \( W = 1 \). The parameter to be estimated and the number of samples were
set to $\theta_0 = 0.2$ and $N = 5$, respectively. It can be shown that under these settings, the terms composing the MH bound in (3.15) are given by: $J = 2\rho \|\hat{a}(\theta_0)\|^2$, where $\rho = \frac{|s|^2}{\sigma^2}$ denotes the SNR. 

$$d_j = 2\rho \text{Re} \left\{ \hat{a}^H(\theta_0)[a(\theta_j) - a(\theta_0)] \right\}, j = 1, \ldots, M$$

and

$$[B]_{ij} = \exp \left( 2\rho \text{Re} \left\{ [a(\theta_i) - a(\theta_0)]^H [a(\theta_j) - a(\theta_0)] \right\} \right), i, j = 1, \ldots, M.$$ 

The MH bound is derived under the assumption of unbiasedness in the vicinity of the true parameter and at the selected test-points. Observing Fig. 3.1, one can notice that the MH bound is greatly dependent on the location of the chosen test-points, i.e, the bound greatly increases with minor changes in $\delta$. Furthermore, it can be observed that the bound approaches large values, as $\delta$ becomes small. This indicates, that as $\delta \to 0$, an estimator which is unbiased at the test-point does not exist. Furthermore, one can notice that the threshold region of the MH bound is greatly dependent on the choice of $\delta$, and with minor changes in $\delta$. The threshold region of the MH bound can be below or above the threshold region of the MLE. Observing Fig. 3.1, one can notice that above the threshold region, the CRLB successfully predicts the MSE performance of the MLE. The reason is that the MLE is asymptotically locally unbiased and the constraint of locally unbiasedness of the CRLB is valid at the vicinity of the true value. However, the MLE is not constrained to uniform unbiasedness in the strict sense, and moreover, unlike the unbiasedness constraint of the MH bound, the MLE is not unbiased at the chosen test-points. Hence, it should not be surprising that the MLE has better MSE performance and it is lower than the BB approximations. Fig. 3.2 compares the threshold region of the MLE and the MH bound versus $\delta$. Observing Fig. 3.2, one can notice that the threshold region of the MH bound can arbitrarily move, depending on $\delta$, and it can be below or
above the threshold region of the MLE.

Fig. 3.3 depicts the MH bound as a function of $\delta$, for various SNRs. It is observed, that for each SNR value, there exists a small enough $\delta$, such that the MH bound diverges from the CRLB. By comparing the MH bounds at each SNR, one can notice that the threshold region of the MH bound is heavily dependent on the location of the test-point. Furthermore, according to Theorem 4, it is shown in Fig. 3.3, that $C_{MH} \to \infty$ as $\delta \to 0$. It should be noticed that for some finite set of test-points and finite $\delta$, the MH is finite, thus, an estimator which satisfies the condition of unbiasedness in those test-points exist. However, with minor changes to the location of the test-points and as $\delta \to 0$, the MH bound greatly increases, and approaches infinity even for large number of samples. Hence, the BB approximation in this problem, is meaningless for the prediction of the MLE threshold region, and the strict unbiasedness constraint is not useful in finding the threshold phenomenon of the MLE.

3.2.2 Exponential parameter estimation in Gaussian noise

In this example, the entries of $a(\theta_0)$ are given by $[a(\theta_0)]_k = \exp(\theta_0)$, $k = 1, \ldots, N$, the amplitude is $s = 1$, and $\theta_0 \in \mathbb{R}$ is the unknown parameter to be estimated. In Theorem 5, a proof of the nonexistence of UU estimator of $\theta_0 \in \mathbb{R}$ was obtained under the condition that $\|a(\theta)\| \in [a, b]$. This is similar to the case presented here, although in this problem the norm of the vector function $a(\theta), \theta \in \mathbb{R}$, is only half bounded, $\|a(\theta)\| \in [0, \infty)$. However, in similar to the proof of Theorem 5, it is shown in the appendix that a UU estimator does not exist.
Figure 3.4: Comparison of MLE, CRLB and the MH bounds for various $\theta_M$, versus the number of samples, $N$.

Figs. 3.4-3.6 depict the MH bounds in exponential parameter estimation in WGN, where $\theta_0 = 0.3$. The test-points are equally spaced, such that $M = 4$, $\theta_1$ is fixed and $\theta_2, \ldots, \theta_M$ vary in the parameter space $\Theta \subseteq \mathbb{R}$, and $\theta_j = \theta_1 - \frac{\theta_0 - \theta_M}{M-1} (j - 1), j = 2, \ldots, M$.

According to Fig. 3.4, the MH bound is greatly dependent on the location of the chosen test-points, and with minor changes in $\theta_M$ the bound sharply increases. It is shown that the threshold region of the MH bound is heavily dependent on the location of the unbiasedness constraint, which is $\theta_M$. In Fig. 3.5, the threshold region of the MH is compared with the threshold region of the MLE. It is shown that as $\theta_M \to -\infty$ the threshold region goes to infinity. Moreover, observing Fig. 3.4, one can notice that the CRLB successfully predicts the MSE performance of the MLE. The reason is that the
MLE is asymptotically locally unbiased and the constraint of unbiasedness of the CRLB is valid at the vicinity of the true value. However, the MLE is not uniformly unbiased in the strict sense and unlike the constraint in the MH bound, the MLE does is not unbiased at the chosen test-points. Thus, the MLE can achieve asymptotic efficiency and its MSE is lower than the BB approximations.

Fig. 3.6 depicts the MH bound as a function of $\theta_M$, for various number of samples, $N$. It can be observed, that there exists a set of test-points, such that the MH bound deviates from the CRLB. Furthermore, Fig. 3.6 shows that $C_{MH} \rightarrow \infty$ as $\theta_M \rightarrow -\infty$, as implied by Corollary 2. Hence, the BB approximation is meaningless in this example for the prediction of the MLE threshold region, and it is not a useful tool in finding the
3.3 Calculation of the approximated MH bound in single-tone parameter estimation

In this section, the calculation of the approximated MH bound is presented. Denote by $C_{MH}$ the MH bound. Thus,

$$C_{MH} = J^{-1} + q^T R^{-1} q$$ (3.22)
where the matrices $q, R$ are defined as,

$$q = t - J^{-1}d$$  \hspace{1cm} (3.23) $$

$$R = B - dJ^{-1}dT.$$  \hspace{1cm} (3.24) $$

and,

$$J \triangleq \int_\Omega \left[ \frac{\partial \log f(x; \theta)}{\partial \theta} \right] _{\theta=\theta_0}^2 f(x; \theta_0)dx,$$  \hspace{1cm} (3.25) $$

$$[B]_{ij} \triangleq \int_\Omega L(x; \theta_i, \theta_0)L(x; \theta_j, \theta_0)f(x; \theta_0)dx, \hspace{1cm} i,j = 1, \ldots, M$$  \hspace{1cm} (3.26) $$

$$[d]_j \triangleq \int_\Omega \left[ \frac{\partial \log f(x; \theta)}{\partial \theta} \right] _{\theta=\theta_0} L(x; \theta_j, \theta_0)f(x; \theta_0)dx, \hspace{1cm} j = 1, \ldots, M$$  \hspace{1cm} (3.27) $$

$$[t]_j \triangleq \theta_j - \theta_0, \hspace{1cm} j = 1, \ldots, M$$  \hspace{1cm} (3.28) $$

where the parameters $\theta_j \in [-W/2, W/2], j = 1, \ldots, M$ denote the $M$ test-points and $\theta_0$ is the unknown parameter. Using the Woodberry identity [34], the MH bound can be evaluated by using a simpler form,

$$(B - dJ^{-1}dT)^{-1} = B^{-1} + B^{-1}d(J - d^TB^{-1}d)^{-1}dT^{-1}B.$$  \hspace{1cm} (3.29) $$

Thus, the MH bound is lower bounded by,

$$C_{MH} = J^{-1} + q^TB^{-1}q + q^TB^{-1}d(J - d^TB^{-1}d)^{-1}dT^{-1}B^{-1}q$$  \hspace{1cm} (3.30) $$

$$= J^{-1} + (t - J^{-1}d)^TB^{-1}(t - J^{-1}d) + q^TB^{-1}d(J - d^TB^{-1}d)^{-1}dT^{-1}B^{-1}q$$  \hspace{1cm} (3.31) $$

$$\geq J^{-1} + t^TB^{-1}t - t^TB^{-1}dJ^{-1} - J^{-1}dT^{-1}B^{-1}t$$  \hspace{1cm} (3.32) $$

where the last inequality is due to the fact that the scalar term $(J - d^TB^{-1}d)$ is non-negative and $B$ is a positive semi-definite (p.s.d) matrix. We will approximate the remaining expressions using first-order Taylor approximation. The pdf of the observation
vector $\mathbf{x}$ in can be written as

$$f(\mathbf{x}; \theta) = \frac{1}{(\pi \sigma^2)^N} \exp \left\{ -\|\mathbf{x} - s\mathbf{a}(\theta)\|^2 \right\}.$$  \hspace{1cm} (3.33)

Using (3.33), the LR in (2.4) can be expressed as

$$L(\mathbf{x}; \theta_j, \theta_0) = \exp \left\{ -\left[ \|\mathbf{x} - s\mathbf{a}(\theta_j)\|^2 + \|\mathbf{x} - s\mathbf{a}(\theta_0)\|^2 \right] \right\}, j = 1, \ldots, M.$$  \hspace{1cm} (3.34)

Denote $[\mathbf{B}]_{ij} = b_{ij}$. By substituting (3.34) into (3.26), followed by using the result in [7], Eqs. (11)-(16), one obtains

$$b_{11} = \exp \left[ 2\rho \|\mathbf{a}(\theta_1) - \mathbf{a}(\theta_0)\|^2 \right]$$

$$b_{12} = b_{21} = \exp \left[ 2\rho \Re \{ (\mathbf{a}(\theta_2) - \mathbf{a}(\theta_0))^H (\mathbf{a}(\theta_1) - \mathbf{a}(\theta_0)) \} \right]$$

$$b_{22} = \exp \left[ 2\rho \|\mathbf{a}(\theta_2) - \mathbf{a}(\theta_0)\|^2 \right].$$

where $\theta_1 = -W/2, \theta_2 = W/2 - \delta$. Using Taylor approximation of $\mathbf{a}(\theta_2)$ and the fact that the function $\mathbf{a}(\cdot)$ is periodic, with period $W$, results in,

$$\mathbf{a}(W/2 - \delta) = \mathbf{a}(-W/2 - \delta)$$

$$\approx \mathbf{a}(-W/2) - \delta \dot{\mathbf{a}}(-W/2).$$ \hspace{1cm} (3.35)

Thus, using Taylor approximation, the elements $b_{ij}$ can be approximated as

$$b_{21} \approx \exp \left[ 2\rho \Re \{ (\mathbf{a}(-W/2) - \delta \dot{\mathbf{a}}(-W/2) - \mathbf{a}(\theta_0))^H (\mathbf{a}(\theta_1) - \mathbf{a}(\theta_0)) \} \right]$$  \hspace{1cm} (3.37)

$$= b_{11} \exp \left[ -2\delta \rho \Re \{ \dot{\mathbf{a}}(-W/2)^H (\mathbf{a}(-W/2) - \mathbf{a}(\theta_0)) \} \right]$$  \hspace{1cm} (3.38)

$$b_{22} \approx \exp \left[ 2\rho \|\mathbf{a}(-W/2) - \delta \dot{\mathbf{a}}(-W/2) - \mathbf{a}(\theta_0)\|^2 \right]$$  \hspace{1cm} (3.39)

$$= \exp \left[ \{ 2\rho \{ \|\mathbf{a}(\theta_1) - \mathbf{a}(\theta_0)\|^2 \right.$$

$$\left. - 2\delta \Re \{ \dot{\mathbf{a}}(-W/2)^H (\mathbf{a}(-W/2) - \mathbf{a}(\theta_0)) \} + \delta^2 \|\dot{\mathbf{a}}(-W/2)\|^2 \} \right].$$
The determinant of matrix $\mathbf{B}$ is,

$$|\mathbf{B}| = b_{11}b_{22} - b_{12}^2$$

(3.41)

$$\approx b_{11}^2 \exp \left\{ 2\delta^2 \rho \| \dot{\mathbf{a}} (-W/2) \|^2 \right\}$$

(3.42)

$$- 4\delta \rho \mathrm{Re} \left\{ \dot{\mathbf{a}} (-W/2)^H (\mathbf{a} (-W/2) - \mathbf{a}(\theta_0)) \right\}$$

$$- b_{11}^2 \exp \left[ -2\delta \rho \mathrm{Re} \left\{ \dot{\mathbf{a}} (-W/2)^H (\mathbf{a} (-W/2) - \mathbf{a}(\theta_0)) \right\} \right]$$

(3.43)

$$= b_{11}^2 \exp \left[ -4\delta \rho \mathrm{Re} \left\{ \dot{\mathbf{a}} (-W/2)^H (\mathbf{a} (-W/2) - \mathbf{a}(\theta_0)) \right\} \right]$$

$$[\exp \left\{ 2\delta^2 \rho \| \dot{\mathbf{a}} (-W/2) \|^2 \right\} - 1]$$

(3.44)

$$= b_{11}^2 \left[ 2\delta^2 \rho \| \dot{\mathbf{a}} (-W/2) \|^2 \right]$$

$$\approx b_{11}^2 \left[ 2\delta^2 \rho \| \dot{\mathbf{a}} (-W/2) \|^2 \right],$$

(3.45)

where the last approximation is valid only if $\delta \cdot \rho \ll 1$. Denote $[\mathbf{d}]_j = d_j$, thus, the elements $d_j$ can be approximated as,

$$d_1 = 2\rho \mathrm{Re} \left\{ \dot{\mathbf{a}} (\theta_0)^H (\mathbf{a} (-W/2) - \mathbf{a}(\theta_0)) \right\}$$

(3.46)

$$d_2 = 2\rho \mathrm{Re} \left\{ \dot{\mathbf{a}} (\theta_0)^H (\mathbf{a} (-W/2) - \delta \dot{\mathbf{a}} (-W/2) - \mathbf{a}(\theta_0)) \right\}$$

(3.47)

$$\approx 2\rho \mathrm{Re} \left\{ \dot{\mathbf{a}} (\theta_0)^H (\mathbf{a} (-W/2) - \delta \dot{\mathbf{a}} (-W/2) - \mathbf{a}(\theta_0)) \right\}$$

(3.48)

$$= d_1 - 2\delta \rho \mathrm{Re} \left\{ \dot{\mathbf{a}} (\theta_0)^H \dot{\mathbf{a}} (-W/2) \right\}$$

(3.49)

$$\approx d_1 + O(\delta),$$

(3.50)
where $O(\cdot)$ is the big O notation function. Furthermore, the elements $b_{12}, b_{22}$ can be rewritten as

\[
b_{21} \approx \exp \left[ 2\rho \text{Re} \left\{ (a(-W/2) - \delta \dot{a}(-W/2) - a(\theta_0))^H (a(\theta_1) - a(\theta_0)) \right\} \right] \tag{3.51}
\]

\[
\approx b_{11} (1 + O(\delta)) \tag{3.52}
\]

\[
b_{22} \approx b_{11} \exp \left\{ 2\delta^2 \rho \|\dot{a}(-W/2)\|^2 - 4\delta \rho \text{Re} \left\{ \dot{a}(-W/2)^H (a(-W/2) - a(\theta_0)) \right\} \right\} \tag{3.53}
\]

\[
\approx b_{11} (1 + O(\delta)). \tag{3.54}
\]

Thus, the element $t^T B^{-1} t$ can be rewritten as,

\[
t^T B^{-1} t = \frac{1}{|B|} \left[ b_{22} (\theta_1 - \theta_0)^2 
\right.

- 2b_{12} (\theta_1 - \theta_0) (\theta_2 - \theta_0) + b_{11} (\theta_2 - \theta_0)^2 \right] \tag{3.55}
\]

\[
\approx \frac{1}{|B|} \left[ b_{11} (1 + O(\delta)) (\theta_1 - \theta_0)^2 
\right. 

- 2 (b_{11} (1 + O(\delta))) (\theta_1 - \theta_0) (\theta_2 - \theta_0) + b_{11} (\theta_2 - \theta_0)^2 \right] \tag{3.56}
\]

\[
\approx \frac{b_{11}}{|B|} \left[ (\theta_1 - \theta_0)^2 - 2 (\theta_1 - \theta_0) (\theta_2 - \theta_0) + (\theta_2 - \theta_0)^2 + O(\delta) \right] \tag{3.57}
\]

\[
= \frac{b_{11}}{|B|} \left[ (\theta_1 - \theta_2)^2 + O(\delta) \right] \tag{3.58}
\]

\[
= \frac{b_{11}}{|B|} \left[ (-W + \delta)^2 + O(\delta) \right] \tag{3.59}
\]

\[
= \frac{b_{11}}{|B|} \left[ W^2 + O(\delta) \right] \tag{3.60}
\]
and the element $t^T B^{-1} d J^{-1}$ can be written as,

$$t^T B^{-1} d J^{-1} = \frac{1}{|B|} J^{-1} [b_{22} (\theta_1 - \theta_0) d_1 - b_{12} (\theta_1 - \theta_0) d_2] \quad (3.61)$$

$$- [b_{21} (\theta_2 - \theta_0) d_1 + b_{11} (\theta_2 - \theta_0) d_2]$$

$$\approx \frac{b_{11}}{|B|} J^{-1} [(1 + O(\delta)) d_1 (\theta_1 - \theta_0)] \quad (3.62)$$

$$- (1 + O(\delta)) (\theta_1 - \theta_0) (d_1 + O(\delta))$$

$$- (1 + O(\delta)) (\theta_2 - \theta_0) d_1 + (\theta_2 - \theta_0) (d_1 + O(\delta))]$$

$$= \frac{b_{11}}{|B|} J^{-1} [O(\delta)] \quad (3.63)$$

Finally, (3.32) can be approximated as

$$t^T B^{-1} t - t^T B^{-1} d J^{-1} - J^{-1} d^T B^{-1} t = \frac{b_{11}}{|B|} [W^2 + O(\delta) - J^{-1} O(\delta)] \quad (3.64)$$

$$= \frac{b_{11}}{b_{11} [2\delta^2 \rho \|\hat{a}(-W/2)\|^2]} [W^2 + O(\delta)] \quad (3.65)$$

$$= \frac{1}{b_{11} [2\delta^2 \rho \|\hat{a}(-W/2)\|^2]} [W^2 + O(\delta)], \quad (3.66)$$

under the condition that $\delta \cdot \rho \ll 1$. Thus, the resulting approximation of the MH bound is,

$$C_{MH} \geq J^{-1} + \frac{W^2}{b_{11} [2\delta^2 \rho \|\hat{a}(-W/2)\|^2]} \quad (3.67)$$

$$= J^{-1} + \frac{W^2}{2\delta^2 \rho \|\hat{a}(-W/2)\|^2 \exp \left[2\rho \|a(-W/2) - a(\theta_0)\|^2\right]} \quad (3.68)$$

### 3.4 Appendix

In this appendix, the proof of the nonexistence of UU estimator in Theorem 5 is extended to the case of exponential parameter estimation in WGN. It should be noticed that (3.13)
is still valid under the condition of half boundedness, \( \|a(\theta)\| \in [0, \infty) \). Thus, (3.13) can be rewritten as

\[
|\theta_1 - \theta_0|^2 \leq C \cdot \exp \left\{ 2 \| \exp(\theta_1) - \exp(\theta_0) \|^2_{\mathbb{R}^n} \right\}.
\] (3.69)

Choosing \( \theta_1 \to -\infty \), the exponential term is finite, and for finite \( \theta_0 \) the constant \( C \to \infty \). Thus, the proof of nonexistence of UU estimator is extended to match this particular model. Furthermore, according to as Corollary 2, the nonexistence of UU estimator states that the BB is infinity.
Chapter 4

Approximated minimum-variance unbiased estimator

4.1 A new approach to minimum-variance unbiased estimation

Minimum variance unbiased (MVU) estimation is one of the main challenges in the theory of statistical signal processing. Currently, direct MVU estimation can be carried out via the following approaches. First, according to the Rao-Blackwell-Lehmann-Scheffe (RBLS) theorem [3], the MVU estimator can be derived under the existence of a complete sufficient statistic. Unfortunately, completeness of the sufficient statistic is not always satisfied and moreover, completeness proof of the sufficient statistic is not always easy. Hence, this approach for finding a MVU estimator is not practical in many cases. The MVU estimator
can also be obtained in cases where an efficient estimator exists. In these cases, the MVU is efficient and hence, it is given by the MLE. However, in many estimation problems an efficient estimator does not exist and unlike the MLE, no numerical technique has been proposed for approximate MVU estimation. In this chapter, a new iterative approach for approximate MVU estimation is presented.

4.2 MVU estimation via an integral equation

In this section, we represent a theorem presented in [25],[27] in which a direct approach for finding the locally best uniformly unbiased (LBU) estimator was derived. Furthermore, it was shown, that if a MVU estimator exists, then the LBU estimator is the MVU estimator.

Denote \( \varphi \) as the set of UU estimators in the parameter space, such that

\[
\varphi = \left\{ \hat{\theta} : E\left[ \hat{\theta} ; \theta \right] = \theta , \ \forall \theta \in \Theta \right\} ,
\]  

(4.1)

where \( \hat{\theta} = \hat{\theta}(x) \). The LBU estimator minimizes the MSE, evaluated at \( \theta_0 \), from the set \( \varphi \),

\[
\hat{\theta}_{LBU} = \arg \min_{\hat{\theta} \in \varphi} \text{MSE} \left[ \hat{\theta} ; \theta_0 \right] .
\]  

(4.2)

Under the assumption that the class of UU estimators is not empty, the LBU estimator can be obtained by the following theorem.

**Theorem 6 ([25],[27]).** Denote the LR function as,

\[
L(x; \theta, \theta_0) = \frac{f(x; \theta)}{f(x; \theta_0)} .
\]  

(4.3)
Then the LBU estimator, $\hat{\theta}_{LBU}$, is given by,

$$\hat{\theta}_{LBU} = \theta_0 + \int_{\Theta} \mu(\theta; \theta_0) L(x; \theta, \theta_0) d\theta$$ \hspace{1cm} (4.4)

where $\mu(\theta; \theta_0)$ is the solution of the integral equation,

$$\int_{\Theta} \mu(\theta'; \theta_0) B_{\theta_0}(\theta, \theta') d\theta' = \theta - \theta_0$$ \hspace{1cm} (4.5)

with

$$B_{\theta_0}(\theta, \theta') = E[L(x; \theta, \theta_0) L(x; \theta', \theta_0); \theta_0].$$ \hspace{1cm} (4.6)

**Proof.** The proof of this theorem is given in detail in [25],[27]. \hfill \Box

The LBU given in (4.4) depends on $\theta_0$. If, however, a MVU estimator exists, the uniformity property of the MVU estimator obligates that it should be independent of $\theta_0$. Thus, in practice, when a MVU estimator exists, it is possible to replace $\theta_0$ with any value. Moreover, if a MVU estimator exists, the LBU estimator is the MVU estimator, and it can be obtained by (4.4), as it was shown in [25],[27]. However, one can notice that (4.4) requires the solution of an integral equation, (4.5), which is analytically cumbersome, and in many cases not possible.

### 4.3 Approximated MVU estimator

In order to overcome the difficulty in solving the integral equation, (4.5), an approximation of the integral equation is performed. In [23], a new class of non-Bayesian lower bounds on the MSE for unbiased estimators was derived. Furthermore, generalization of known
lower bounds was obtained by applying an integral transform on the LR function. Thus, via the selection of the kernel of the integral transform, a tight and computationally simple approximation of the BB was obtained. We now show, that by using a similar approach, an approximation of the MVU estimator can be obtained, assuming that the MVU estimator does exist.

Let \( \eta(x; \tau, \theta_0) \) denote the integral transform of the LR function with kernel \( h(\tau, \theta) \),

\[
\eta(x; \tau, \theta_0) = \int_{\Theta} h(\tau, \theta)L(x; \theta, \theta_0)d\theta. \tag{4.7}
\]

By using the same integral transform, one can define a new function \( \lambda(\cdot; \theta_0) \), such that

\[
\mu(\theta; \theta_0) = \int_{\Lambda} \lambda(\tau; \theta_0)h(\tau, \theta)d\tau. \tag{4.8}
\]

Thus, the approximated MVU estimator, \( \hat{\theta}_{AMVU} \), can be obtained by substituting (4.8) into (4.4),

\[
\hat{\theta}_{AMVU} = \theta_0 + \int_{\Theta} \int_{\Lambda} h(\tau, \theta)\lambda(\tau; \theta_0)L(x; \theta, \theta_0)d\tau d\theta. \tag{4.9}
\]

By using the definition in (4.7), Eq. (4.9) can be rewritten as

\[
\hat{\theta}_{AMVU} = \theta_0 + \int_{\Lambda} \lambda(\tau; \theta_0)\eta(x; \tau, \theta_0)d\tau, \tag{4.10}
\]

where \( \lambda(\cdot; \theta_0) \) is the solution of the following integral equation. The new integral equation is derived as follows. The expectation operator with parameter \( \theta' \) is applied to (4.10),

\[
E\left[\hat{\theta}_{AMVU}; \theta'\right] = \theta_0 + \int_{\Lambda} \lambda(\tau; \theta_0)E[\eta(x; \tau, \theta_0); \theta']d\tau. \tag{4.11}
\]

Thus, by using the definition in (4.7), Eq. (4.11) can be rewritten as

\[
E\left[\hat{\theta}_{AMVU}; \theta'\right] = \theta_0 + \int_{\Lambda} \lambda(\tau; \theta_0)\int_{\Theta} h(\tau, \theta)B_{\theta_0}(\theta, \theta')d\theta d\tau. \tag{4.12}
\]
The constraint of uniform unbiasedness, $E\left[\hat{\theta}_{AMVU}; \theta'\right] = \theta'$, $\forall \theta' \in \Theta$, is applied to the approximated MVU estimator, thus (4.12) is

$$\int_{\Lambda} \lambda(\tau; \theta_0) \int_{\Theta} h(\tau, \theta) B_{\theta_0}(\theta, \theta') d\theta d\tau = \theta' - \theta_0. \quad (4.13)$$

Additionally, by applying the integral transform on (4.13), a more practical and computationally efficient integral equation can be obtained

$$\int_{\Theta} h(\tau', \theta') \int_{\Theta} \lambda(\tau; \theta_0) \int_{\Theta} h(\tau, \theta) B_{\theta_0}(\theta, \theta') d\theta d\tau d\theta' = \int_{\Theta} h(\tau', \theta')(\theta' - \theta_0) d\theta'. \quad (4.14)$$

We note that by applying the integral transform, the constraint of uniform unbiasedness is not valid. Finally, by integral order changes, (4.14) can be rewritten as

$$\int_{\Lambda} \lambda(\tau; \theta_0) \int_{\Theta} \int_{\Theta} h(\tau, \theta) B_{\theta_0}(\theta, \theta') h(\tau', \theta') d\theta d\theta' d\tau = \int_{\Theta} h(\tau', \theta')(\theta' - \theta_0) d\theta'. \quad (4.15)$$

One can notice that the approximated MVU estimator becomes the exact MVU estimator, if the integral transform is invertible. Additionally, by judicious selection of a non-invertible integral transform, the integral equation in (4.15) can be directly solved and an approximated MVU estimator can be derived explicitly by (4.10). In the following section, an iterative process is proposed, where at each step the denominator of the LR function and the kernel of the integral transform are updated by the estimate obtained in the previous step.
4.3.1 Derivation of scoring method from the approximated MVU estimator

Derivation of the scoring method [3] can be carried out by choosing $h(\tau, \theta)$ to be,

$$ h(\tau, \theta) = \frac{\partial \delta(\tau - \theta)}{\partial \tau} \delta(\tau - \theta_0). \quad (4.16) $$

where $\delta$ is the Dirac delta function. In [30] the same kernel was used in order to derive the CRLB. Thus, (4.10) can be rewritten as

$$ \hat{\theta}_{CR} = \theta_0 + \lambda(\theta_0) \left. \frac{\partial L(x; \tau, \theta_0)}{\partial \tau} \right|_{\tau = \theta_0}, \quad (4.17) $$

where $\lambda(\theta_0)$ can be obtained by substituting (4.16) into (4.15),

$$ \lambda(\theta_0) E \left[ \left. \frac{\partial L(x; \tau, \theta_0)}{\partial \tau} \right|_{\tau = \theta_0} \left. \frac{\partial L(x; \tau', \theta_0)}{\partial \tau'} \right|_{\tau' = \theta_0} ; \theta_0 \right] = 1. \quad (4.18) $$

One can notice that $E \left[ \left. \frac{\partial L(x; \tau, \theta_0)}{\partial \tau} \right|_{\tau = \theta_0} \left. \frac{\partial L(x; \tau', \theta_0)}{\partial \tau'} \right|_{\tau' = \theta_0} ; \theta_0 \right]$ is the Fisher information matrix evaluated in $\theta_0$. Thus, (4.18) can be rewritten as

$$ \lambda(\theta_0) = J^{-1}(\theta_0), \quad (4.19) $$

where $J(\theta_0)$ is the Fisher information evaluated in $\theta_0$. Thus, the resulting estimator is

$$ \hat{\theta}_{CR} = \theta_0 + J^{-1}(\theta_0) \left. \frac{\partial L(x; \tau, \theta_0)}{\partial \tau} \right|_{\tau = \theta_0}, \quad (4.20) $$

where $\theta_0$ is the unknown parameter. If a MVU exists, we can initialize $\theta_0$ by an arbitrary value, and at each iteration update $\theta_0$ by its estimate obtained in the previous step. Therefore, the following iterative method can be obtained, which is identical to the scoring method,

$$ \theta_{k+1} = \theta_k + J^{-1}(\theta_k) \left. \frac{\partial L(x; \tau, \theta_k)}{\partial \tau} \right|_{\tau = \theta_k}. \quad (4.21) $$
4.3.2 Approximated MVU estimator based on the Barankin kernel

In [30] the discrete form of the BB [5],[20],[21], was obtained by using the following kernel,

\[ h(\tau, \theta) = \sum_{i=1}^{M} \delta(\tau - \theta_i)\delta(\tau - \theta), \]  

(4.22)

where \{\theta_i \in \Theta, i = 1, \ldots, M\} denotes a set of test-points. Thus, an approximated MVU estimator based on the discrete BB can be derived, by substituting the kernel (4.22), into (4.10). This yields,

\[ \hat{\theta}_{\text{Barankin}} = \theta_0 + \lambda^T \nu_{\theta_0}(x), \]  

(4.23)

where \( \nu_{\theta_0} \) is the LR function evaluated at the test-points, such that, \( [\nu_{\theta_0}]_i = L(x; \theta_i, \theta_0), \ i = 1, \ldots, M \), and \( [\lambda]_i = \lambda(\theta_i; \theta_0), \ i = 1, \ldots, M \) is the function \( \lambda(\cdot; \theta_0) \) evaluated at the test-points. The vector \( \lambda \) can be obtained by substituting (4.22) into (4.15),

\[ \sum_{k=1}^{M} \lambda(\theta_k)B(\theta_i, \theta_k) = \theta_i - \theta_0, \ i = 1, \ldots, M. \]  

(4.24)

Equation (4.24) can be written in matrix form as

\[ \lambda^T B_{\theta_0} = t_{\theta_0}^T \]  

(4.25)

where \( [t_{\theta_0}]_i = \theta_i - \theta_0, \ i = 1, \ldots, M \), and \( B_{\theta_0} \) is the Barankin matrix evaluated at the test-points, such that, \( [B_{\theta_0}]_{ij} = B_{\theta_0}(\theta_i, \theta_j), \ i, j = 1, \ldots, M \). Thus,

\[ \lambda = B_{\theta_0}^{-1} t_{\theta_0}. \]  

(4.26)
where $\theta_0$ is the unknown parameter. Therefore, if a MVU exists, we can initialize $\theta_0$ by an arbitrary value, and at each iteration update $\theta_0$ by its estimate obtained in the previous step. Thus, the new iterative method can be obtained,

$$
\theta_{k+1} = \theta_k + t^{T}_{\theta_k} B^{-1}_{\theta_k} \nu_{\theta_k}(x).
$$

(4.27)

### 4.4 Bias evaluation of the approximated MVU estimator based on the Barankin kernel

In this section, we will analytically evaluate the bias of the approximated Barankin kernel MVU estimator, as a function of $\theta_0$. We define a family of approximated Barankin kernel MVU estimators \( \{\hat{\theta}_{AMVU}^k, k = 1, \ldots, K\} \), and a family of test-points \( \{\theta_k \in \Theta, k = 1, \ldots, K\} \), thus, according to (4.27) the approximated MVU estimator is

$$
\hat{\theta}_{AMVU}^k = \theta_k + t^{T}_{\theta_k} B^{-1}_{\theta_k} \nu_{\theta_k}(x).
$$

(4.28)

We will now evaluate the bias and MSE of each estimator at $\theta_0$. The expectation of the approximated MVU estimator at $\theta_0$ is,

$$
E \left[ \hat{\theta}_{AMVU}^k, \theta_0 \right] = \theta_k + t^{T}_{\theta_k} B^{-1}_{\theta_k} E \left[ \nu_{\theta_k}(x); \theta_0 \right]
$$

(4.29)

$$
= \theta_k + t^{T}_{\theta_k} B^{-1}_{\theta_k} \beta_{\theta_k}(\theta_0)
$$

(4.30)

where the elements of the vector $\beta_{\theta_k}(\theta_0)$ are

$$
[\beta_{\theta_k}(\theta_0)]_j = E[L(x; \theta_j, \theta_k)L(x; \theta_0, \theta_k); \theta_j], \ j = 1, \ldots, K.
$$

(4.31)
The bias of an estimator $\hat{\theta}$ at $\theta_0$, is defined as,

$$\text{bias} \left[ \hat{\theta}; \theta_0 \right] = E \left[ \hat{\theta} - \theta_0; \theta_0 \right].$$  \hspace{1cm} (4.32)

Accordingly, the bias of each approximated MVU estimator $\hat{\theta}_{AMVU}^k$, $k = 1, \ldots, K$, at $\theta_0$ is

$$\text{bias} \left[ \hat{\theta}_{AMVU}^k; \theta_0 \right] = \theta_k - \theta_0 + t_{\theta_k}^T B_{\theta_k}^{-1} \beta_k (\theta_0).$$  \hspace{1cm} (4.33)

We will now prove that under the conditions that $B_{\theta_k}^{-1}$ exists, and that $\theta_0$ is in the set of test-points, then each of the approximated MVU estimators $\hat{\theta}_{AMVU}^k$, $k = 1, \ldots, K$, are unbiased at all the chosen test-points.

**Theorem 7.** Let $B_{\theta_k}^{-1}$ be invertible and $\theta_0 \in \{ \theta_l \in \Theta, l = 1, \ldots, K \}$, thus, the estimator $\hat{\theta}_{AMVU}^k$ is unbiased at the test-points $\theta_l \in \Theta, l = 1, \ldots, K$ for all $k = 1, \ldots, K$ and it is unbiased at $\theta_0$.

**Proof.** Note that if $\theta_0 = \theta_i \in \{ \theta_1, \ldots, \theta_K \}, i = 1, \ldots, K$, then the vector $\beta_k (\theta_0)$ is the $i$th column of the matrix $B_{\theta_k}$. Since $B_{\theta_k}^{-1} B_{\theta_k} = I$, where $I$ is the identity matrix, then $B_{\theta_k}^{-1} \beta_k (\theta_0 = \theta_i) = e_i$, where $e_i$ is a standard basis vector. Accordingly, the bias of $\hat{\theta}_{AMVU}^k$ at $\theta_0 = \theta_i$ is,

$$\text{bias} \left[ \hat{\theta}_{AMVU}^k; \theta_i \right] = \theta_k - \theta_i + t_{\theta_k}^T e_i.$$  \hspace{1cm} (4.34)

Since

$$t_{\theta_k} e_i = \theta_0 - \theta_k$$

then,

$$\text{bias} \left[ \hat{\theta}_{AMVU}^k; \theta_i \right] = 0.$$  \hspace{1cm} (4.35)

Thus, $\hat{\theta}_{AMVU}^k$ is unbiased at $\theta_0$ if $\theta_0 \in \{ \theta_1, \ldots, \theta_K \}$. \hfill $\Box$
4.5 MSE evaluation of the approximated MVU estimator based on the Barankin kernel

In this section, we will analytically evaluate the MSE of the approximated MVU estimator. The family of approximated MVU estimators is defined in (4.28). The MSE of the approximated MVU estimator evaluated at $\theta_0$ is given by

$$\text{MSE}[\hat{\theta}_k^{AMVU}; \theta_0] \triangleq \mathbb{E}[(\hat{\theta}_k^{AMVU} - \theta_0)^2; \theta_0].$$  \hfill (4.36)

Substituting (4.28) into (4.36) yields

$$\text{MSE}[\hat{\theta}_k^{AMVU}; \theta_0] = \mathbb{E}[(\theta_k + t_{\theta_k} B_{\theta_k}^{-1}\nu_{\theta_k}(x) - \theta_0)^2; \theta_0]$$

$$= t_{\theta_k}^T B_{\theta_k}^{-1} S_{\theta_0} B_{\theta_k}^{-1} t_{\theta_k} + 2(\theta_k - \theta_0) t_{\theta_k}^T B_{\theta_k}^{-1} \mathbb{E}[\nu_{\theta_k}(x); \theta_0] + (\theta_k - \theta_0)^2,$$

where $S_{\theta_0} = \mathbb{E}[\nu_{\theta_k}(x)\nu_{\theta_k}^T(x); \theta_0]$.

4.6 Examples and simulation

4.6.1 Example 1

Consider the problem of estimating the upper limit of a uniformly distributed random variable. The observed data are modeled by,

$$x[n] = w[n], \ n = 0, 1, \ldots, N - 1$$  \hfill (4.39)
where \( \{w[n]\}_{n=0}^{N-1} \) is an independent and identically distributed (i.i.d) sequence with uniform pdf, \( U[0,\theta] \), for \( \theta > 0 \). It is known from [3] that the MVU estimator of \( \theta \) is 
\[
\hat{\theta}_{MVU} = \frac{N+1}{N} \max_{n=1,\ldots,N} \{x[n]\}.
\]
We wish to find the approximated MVU estimator of the upper limit \( \theta \), using the Barankin kernel based estimator. Calculation of the Barankin matrix at \( \theta_0 \) yields,
\[
[B_{\theta_0}]_{ij} = \theta_0^N \min \left\{ \theta_i, \theta_j \right\}^N, \quad 0 < \theta_i, \theta_j \leq \theta_0
\]
Hence, the iterative Barankin kernel based estimator is,
\[
\theta_{k+1} = \theta_k + t_{\theta_k}^T B_{\theta_k}^{-1} \nu_{\theta_k}(x)
\]
where
\[
[\nu_{\theta_k}(x)]_i = \begin{cases} 
\frac{\theta_k^N u(\theta_i - \max\{x[n]\}) u(\min\{x[n]\})}{\theta_i^N \theta_j^N} \min\{x[n]\}\theta_k^N, & 0 < \theta_i \leq \theta_k, \ i = 1, \ldots, M \\
0, & \text{else}
\end{cases}
\]
and \( t_{\theta_k} \), \( i = 1, \ldots, M \). In this example, \( M = 2 \) test-points are selected, at \( \{T(x) - \Delta, T(x) + \Delta\} \) where \( \Delta = 0.001 \) and \( T(x) = \max_{n=1,\ldots,N} \{x[n]\} \) denotes the sufficient statistic. The parameters chosen in this example are: \( \theta_0 = 10 \), \( N = 10 \).

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Barankin kernel based estimator</th>
<th>MVU estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial guess</td>
<td>12 17 30</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>10.5385 10.5385 10.5385</td>
<td>10.5385</td>
</tr>
</tbody>
</table>

Table 4.1: Iterations of the Barankin kernel based estimator.
Figure 4.1: Bias versus the number of samples.

Figure 4.2: RMSE versus the number of samples.
Table 4.1 presents the iterations of the Barankin kernel based estimator, with several initial points. It is shown that the Barankin kernel based estimator converges with one iteration to the MVU estimate. Fig. 4.1 depicts the bias of the MLE, the MVU estimator and the Barankin kernel based estimator. It is observed that the Barankin kernel based estimator has zero bias as the MVU estimator, as opposed to the MLE. Fig. 4.2 depicts the RMSE of the MLE and the Barankin kernel based estimator, compared to the standard deviation (STD) of the MVU, which is \( \text{STD} \left[ \hat{\theta}_{MVU} \right] = \frac{\theta_0}{\sqrt{N(N+2)}} \). Note that the STD of the MVU estimator is the tightest lower bound of all the unbiased estimators. It is observed that the Barankin kernel based estimator RMSE achieves the STD of the MVU estimator, as expected. Moreover, the Barankin kernel based estimator has better performance than the MLE.

### 4.6.2 Example 2

In this example, the observation model is given by

\[
x = s a(\theta_0) + n
\]

(4.43)

where \( x \) denotes an \( N \times 1 \) observation vector, \( s \in \mathbb{R} \) is a known complex amplitude. \( a(\cdot) \) is an \( N \times 1 \) vector function and the entries of \( a(\theta_0) \) are given by \( [a(\theta_0)]_k = \exp(i2\pi(k - 1)\theta_0), \ k = 1, \ldots, N. \theta_0 \) is the parameter to be estimated and \( n \) denotes an \( N \times 1 \) proper complex WGN vector, with known covariance \( C_n = \sigma^2 I \). The pdf of the observation vector \( x \) can be written as

\[
f(x; \theta) = \frac{1}{(\pi \sigma^2)^N} \exp \left\{ -\frac{1}{\sigma^2} \| x - s a(\theta) \|^2 \right\}. \quad (4.44)
\]
Figure 4.3: Bias of the approximated MVU estimator for a specific $\theta_k$, versus $\theta_0$.

Figure 4.4: MSE of the approximated MVU estimator for each $\theta_k$, versus $\theta_0$.

In Chapter 3, a proof of the nonexistence of UU estimator in periodic likelihood functions was derived. However, an estimator that satisfies the constraint of unbiasedness
at a finite set of test-points does exist. In the following figures we evaluate the bias and MSE of the approximated MVU estimator. We define a family of approximated MVU estimators, such that

$$\hat{\theta}_k^{AMVU} = \theta_k + t_{\theta_k} B_{\theta_k}^{-1} \nu_{\theta_k}(x),$$

(4.45)

where the index $k$ denotes the $k$th approximated MVU estimator, based on the point $\theta_k$. The $ij$th element of the Barankin matrix is

$$[B_{\theta_k}]_{ij} = \exp \left\{ 2 \rho (a(\theta_i) - a(\theta_k))^H (a(\theta_j) - a(\theta_k)) \right\},$$

(4.46)

where $\rho = \frac{|s|^2}{\sigma^2}$ and the LR function, $\nu_{\theta_k}$, is

$$[\nu_{\theta_k}]_i = \exp \left\{ -\frac{1}{\sigma^2} \left[ \|x - sa(\theta_i)\|^2 + \|x - sa(\theta_k)\|^2 \right] \right\},$$

(4.47)

and $\theta_1, \ldots, \theta_M$ are a set of $M$ test-points.

Figs. 4.3, 4.4 present the bias and MSE of the approximated MVU estimator evaluated by (4.33) and (4.39), respectively. The set of test-points are uniformly spaced such that $\theta_1 = -0.4291, \ldots, \theta_M = 0.4212$, and $M = 8$ is the number of test-points. Fig. 4.3 depicts the bias of the approximated MVU estimator, defined by $\theta_k = -0.0355$, for various $\theta_0$. It is observed that the approximated MVU estimator is unbiased at the chosen set of test-points. From the Barankin theorem, we conclude that in this problem, a locally best estimator exists, which is unbiased at the test-points. Fig. 4.4 depicts the MSE of each $k$th estimator, as a function of $\theta_0$. It is observed, that each $k$th estimator is best only if the true value $\theta_0$ is in the vicinity of $\theta_k$. If $\theta_0$ is not in the vicinity of $\theta_k$ the approximated MVU estimator results in very large MSE.
Chapter 5

Summary

5.1 Discussion and Conclusions

In this thesis, we presented some popular problems, in which the uniform unbiasedness assumption is not satisfied. We analytically showed that in cases with periodic likelihood function, or parameters of a bounded signal in Gaussian noise, the BB approaches infinity. This property of the BB was demonstrated in problems of single tone estimation, and exponential parameter estimation in AWGN. We showed, that the BB approximations are heavily dependent on the location of the test-points, and can go to infinity. Additionally, the BB approximations, such as MH bound, exhibit a threshold effect similar to the MLE. However, the threshold region of the bound can greatly vary and is heavily dependent on the location of the chosen test-points. Moreover, this threshold can be below or above the MLE threshold. Hence, the BB approximations are meaningless in these problems, in term of finding the threshold region of the MLE.
Additionally, the BB and its approximations, are lower bound on estimators that are conditioned to unbiasedness at the chosen test-points. However, unlike the BB or its approximations, the MLE is not constrained to be unbiased in the strict sense and does not satisfy the unbiasedness constraint at the chosen test-points. Hence, it should not be surprising that the MLE has better MSE performance and it is lower than the BB approximation. Thus, the MLE is not restricted to any strict unbiasedness constraint and therefore, efficiency of the MLE can be achieved.

5.2 Further Research

In this work, it was shown that the BB and its approximations are meaningless, in terms of threshold prediction, in cases with periodic likelihood function, or parameter estimation of a bounded signal in Gaussian noise. Moreover, it was shown that the uniformly unbiased constraint can not be satisfied in these cases. A possible solution to this problem, in cases with periodic likelihood function, can be a new constraint on estimators with periodic unbiasedness property. This constraint can result in a new class of lower bounds on estimators that are uniformly unbiased, with periodic consideration. Furthermore, a new class of estimators can be derived, which are uniformly unbiased in a periodic sense. In cases of parameter estimation of a bounded signal in Gaussian noise, it was shown that the uniformly unbiased constraint can not be satisfied. A possible solution of this problem can be a new constraint on estimators that are uniformly biased, with a known bias function. The bias function can be derived such that the estimators will not be
dependent on the unknown parameter. Additionally, this constraint can result in a new class of lower bounds, on estimators that are uniformly biased in the parameter space, with a known bias function.
Bibliography


